Robust Experiment Design for System Identification via Semi-Infinite Programming Techniques

Dimitrios Katselis * **Cristian R. Rojas** * **James S. Welsh** **
Håkan Hjalmarsson *

* *Automatic Control Lab and ACCESS Linnaeus Center, Electrical Engineering, KTH Royal Institute of Technology, S-100 44 Stockholm, Sweden (Tel: +46 (0)8 790 8464; e-mail: dimitrik@ kth.se, cristian.rojas@ee.kth.se, hjalmars@kth.se).*
** *School of Electrical Engineering and Computer Science,* The *University* of *Newcastle, NSW 2308, Australia (Tel: +61 (02) 4921 6087; e-mail: James.Welsh@newcastle.edu.au).*

Abstract: Robust optimal experiment design for dynamic system identification is cast as a min-max optimization problem, which is infinite-dimensional. If the input spectrum is discretized (either by considering a Riemmann approximation, or by restricting it to the span of a finite dimensional linear space), this problem falls within the class of semi-infinite convex programs. One approach to this optimization problem of infinite constraints is the so called "scenario approach", which is based on a probabilistic description of the uncertainty to deliver a finite program that attempts to approximate the optimal solution with a prescribed probability. In this paper, we propose as an alternative an exchange algorithm based on some recent advances in the field of semi-infinite programming to tackle the same problem. This method is compared with the scenario approach both from the aspects of accuracy and computational efficiency. Furthermore, the comparison includes the MATLAB semi-infinite solver *fseminf* to provide a general palette of methods approximating the robust optimal design problem.

1. INTRODUCTION

The goal of experiment design is to tune the experimental conditions in such a way that all the available information about the unknown system is teased from the experimental data. The reason for this is that a good design can have significant effects in the accuracy of the estimated system model. This has motivated substantial research on experiment design during the last century. Early research in the statistics literature includes [Cox, 1958, Fedorov, 1972, Wald, 1943, Whittle, 1973, Wynn, 1972], and, in the engineering literature, [Gagliardi, 1967, Goodwin et al., 1973, Goodwin and Payne, 1977, Hildebrand and Gevers, 2003b, Levadi, 1966, Mehra, 1974, Zarrop, 1979]. More recent surveys are contained in [Gevers, 2005, Hjalmarsson, 2005, Pronzato, 2008] where many additional references can be found.

In dynamical systems, the model is, typically, nonlinearly parameterized. This implies that the Fisher information matrix (FIM) [Goodwin and Payne, 1977], typically used as the basis for experiment design, depends, inter alia, on the true system parameters, i.e., the very thing that the experiment is aimed at finding.

Robust experiment design includes substantial work on iterative design [Gevers, 2005, Hjalmarsson, 2005], and a sub-optimal min-max solution for a one parameter problem in [Walter and Pronzato, 1997]. Also, recent publications refer to the idea of min-max optimal experiment design [Gevers and Bombois, 2006, Mårtensson and Hjalmarsson, 2006, Rojas et al., 2007, Welsh and Rojas, 2009].

⋆ This work was supported by the Swedish Research Council under contract 621-2009-4017.

In the robust experiment design literature, uncertainty is usually described as a continuous set. This yields an infinite number of constraints, i.e., a semi-infinite programming (SIP) problem, known to be difficult to solve and possibly NP-hard [Ben-Tal and Nemirovski, 1998]. Several review papers cover the topic of semi-infinite programming [Hettich and Kortanek, 1993, López and Still, 2006], where they present alternative methods to tackle such problems in a numerical fashion. In [Calafiore and Campi, 2005, 2006] the "scenario approach" is presented. The advantage of this method is that solvability can be obtained through random sampling of constraints provided that a probabilistic relaxation of the worst case robust paradigm is accepted. The probabilistic relaxation amounts to being content with robustness against the large majority of situations rather than against all situations. In the scenario approach the number of situations is under the control of the designer and can be made arbitrarily close to the set of 'all' situations modulo, of course, computational limitations. Clearly, the scenario approach corresponds to a probabilistic way of converting a semi-infinite program into a finite program and then applying well-established finite nonlinear programming tools to find an approximate solution of the original problem.

In this paper, together with the scenario approach, we consider the class of exchange methods presented in [Hettich and Kortanek, 1993]. We choose not to consider discretization methods, which have the advantage to work with finite subsets of the uncertainty set, because they are computationally costly and the cost per iteration increases dramatically as the size of these finite subsets grows [Hettich and Kortanek, 1993]. Our main proposition is an exchange algorithm based on results presented in [Zhang

et al., 2010]. The approaches considered in this paper are compared with the MATLAB Optimization Toolbox's semi-infinite solver *fseminf*, both from the aspects of accuracy and computational efficiency. In this sense, this paper offers a wide palette of methods to tackle the min-max robust experiment design problem.

The paper is structured as follows. Section 2 presents an overview of the problem of robust experiment design. In Section 3, the scenario approach and an exchange algorithm for semi-infinite programming are introduced, in the context of our problem. Section 4 summarizes some theoretical properties of the exchange algorithm. Section 5 illustrates and compares these two methods (together with the MATLAB *fseminf* solver) via some simulation examples. Finally, Section 6 concludes the paper.

2. ROBUST EXPERIMENT DESIGN

2.1 The FIM

In the context of parametric estimation, we usually employ measures related to the estimator's accuracy so as to mathematically assess the objective of the experiment design. However, the aforementioned accuracy is not only dependent on the experimental conditions, but also on the form of the estimator. Therefore, it is desirable to employ an "estimator-independent" measure. To this end, we assume that the employed estimator is efficient in the sense that its error covariance achieves the Cramér-Rao lower bound [Goodwin and Payne, 1977], i.e, cov $\hat{\theta} = M^{-1}$, where M is the FIM [Casella and Berger, 2002, Silvey, 1970] and $\hat{\theta}$ the parameter estimator.

Consider a single-input single-output (SISO) linear continuous time system, with input $u(t)$ and output $y(t)$, of the form

$$y(t) = G(p)u(t) + H(p)w(t)$$

where G and H are stable rational transfer functions, p is the time derivative operator, H is minimum phase and normalized as $H(\infty) = 1$, and $w(t)$ is zero mean Gaussian white noise of intensity σ^2. We assume open loop operation of the system, i.e., that $u(t)$ and $w(t)$ are independent. We let $\theta := [\rho^T \ \eta^T \ \sigma^2]^T$ where ρ denotes the parameters in G and η denotes the parameters in H. Therefore, we assume that G, H and σ^2 are independently parameterized.

Assume that the sampling period is h, and that we sample the output $y(t)$ with the same sampling period. Using N samples $\{u(kh), y(kh)\}_{k=1}^{N}$ for estimation purposes, the FIM M is given by [Goodwin and Payne, 1977]

$$M = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}$$

where M_1 is the part of the information matrix related to ρ, and M_2 to the rest of the parameters. Clearly, M_2 is independent of the input. Assuming that N is large, we define [Goodwin and Payne, 1977, Walter and Pronzato, 1997],

$$\begin{aligned} \overline{M}(\theta, \Phi_u) &:= \lim_{N\to\infty} \frac{1}{Nh} M_1 \sigma^2 \\ &= \int_0^\infty \widetilde{M}(\theta, \omega)\Phi_u(\omega)d\omega. \end{aligned} \tag{1}$$

where

$$\widetilde{M}(\theta, \omega) := \mathrm{Re}\left\{ \frac{\partial G(j\omega)}{\partial \rho} |H(j\omega)|^{-2} \left[\frac{\partial G(j\omega)}{\partial \rho}\right]^H \right\},$$

and Φ_u is the continuous time input spectrum. We choose to work with $\overline{M}(\theta, \Phi_u)$, rather than M, for convenience purposes.

2.2 Criteria for Nominal Experiment Design

To proceed with the task of experiment design, we need to employ some sort of scalar measure of $\overline{M}$. In the nominal case, i.e., when a predefined estimator of θ is used, several measures of the "size" of $\overline{M}$ have been proposed. Some examples include

(i) D - optimality [Goodwin and Payne, 1977]
$$J_d(\theta, \Phi_u) := [\det \overline{M}(\theta, \Phi_u)]^{-1}.$$

(ii) Experiment design for robust control [Hildebrand and Gevers, 2003a,b, Hjalmarsson, 2005].
$$J_{rc}(\theta, \Phi_u) := \sup_\omega g(\theta, \omega)^H \overline{M}^{-1} g(\theta, \omega)$$
where g is a frequency dependent vector related to the ν-gap [Hildebrand and Gevers, 2003a,b].

(iii) and finally A-optimality ($\mathrm{tr}\, \overline{M}(\theta, \Phi_u)^{-1}$), L-optimality ($\mathrm{tr}\, W\overline{M}(\theta, \Phi_u)^{-1}$, for some $W \geq 0$) and E-optimality $\left(\lambda_{\max}(\overline{M}(\theta, \Phi_u)^{-1})\right)$ [Kiefer, 1974].

The experimental design goal is to choose Φ_u to minimize one of the above criteria. Since most criteria are convex in Φ_u, we will consider that the chosen criterion has this desirable property in the following.

2.3 Min-Max Robust Design

We pursue a min-max robust design on the basis that the parameters can take any value in a compact set Θ. A set of admissible input signals is defined and a constraint on the input energy is imposed. Here, we define the constraint as [1]

$$\mathcal{S}(\mathbb{R}_0^+) := \left\{ \Phi_u : \mathbb{R} \to \mathbb{R}_0^+ : \ \Phi_u \text{ is even and } \int_{-\infty}^{\infty} \Phi_u(\omega)d\omega = 1 \right\}.$$

The min-max robust optimal input spectral density, Φ_u^{opt}, is then chosen as

$$\Phi_u^{opt} = \arg \min_{\Phi_u \in \mathcal{S}(\mathbb{R}_0^+)} \sup_{\theta\in\Theta} J(\theta, \overline{M}(\theta, \Phi_u)) \tag{2}$$

where J is an appropriate scalar differentiable measure of $\overline{M}$. We assume that Φ_u^{opt} exists and is unique; see [Rojas et al., 2007]. Notice also that we allow J to depend explicitly on θ.

2.4 Discrete Approximation to the Optimal Input

The min-max optimization problem (2) is clearly infinite dimensional. For its solution, we use discretization of the design space. First we restrict the positive support of Φ_u to a compact interval, say $K := [\underline{\omega}, \overline{\omega}] \subset \mathbb{R}_0^+$, hence $\Phi_u \in \mathcal{S}(K)$. Next, we approximate the integral in equation (1) by a Riemann sum. Specifically, we choose a grid of $d+1$

[1] In general, given a set $X \subseteq \mathbb{R}_0$, we will denote by $\mathcal{S}(X)$ the set of all even generalized functions Φ_u on $\mathbb{R}$ [Rudin, 1973] such that Φ_u is the derivative of some probability distribution function on $\mathbb{R}$, and $\mathrm{supp}\, \Phi_u \subseteq X \cup (-X)$, where $\mathrm{supp}\, \Phi_u$ is the support of Φ_u (i.e. roughly speaking, $\mathcal{S}(X)$ is the set of all even (generalized) probability density functions on $X \cup (-X)$).

points $\omega_m \in [\underline{\omega}, \overline{\omega}]$ for $m = 0, \ldots, d$ such that $\omega_0 = \underline{\omega}$, $\omega_d = \overline{\omega}$. Then

$$\begin{aligned}\overline{M}(\theta, \Phi_u) &:= \int_{\underline{\omega}}^{\overline{\omega}} \widetilde{M}(\theta, \omega)\Phi_u(\omega)d\omega \\ &\approx \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n)\Phi_u(\omega_n)(\omega_{n+1} - \omega_n) \\ &= \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n)E_n\end{aligned}$$

where $E_n := \Phi_u(\omega_n)(\omega_{n+1} - \omega_n)$. We can now state the following discrete semi-infinite convex programming approximation to (2):

$$\begin{aligned}\min_{t\in\mathbb{R},\ E\in\mathbb{R}^d}\ & t \\ \text{s.t.}\quad & J\left(\theta, \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n)E_n\right) \le t, \qquad \forall\theta \in \Theta \\ & \sum_{n=0}^{d-1} E_n = 1 \\ & E_n \ge 0, \qquad n = 0, \ldots, d-1.\end{aligned} \tag{3}$$

where 's.t.' stands for 'subject to'. In the sequel, we call this problem robust SIP ($RSIP(\Theta)$).

3. COMPUTATIONAL APPROACHES TO ROBUST EXPERIMENT DESIGN

3.1 The Scenario Approach

The scenario approach presumes a probabilistic description of uncertainty, that is, the uncertainty is characterized through a set Δ describing the set of admissible situations, and a probability distribution P_r over Δ.

As shown in Section 2.4, the min-max optimization problem, when converted to a robust convex optimization program yields an unwieldy number of constraints, c.f., (3). The scenario approach involves selecting a small number of these constraints to include in the optimization problem. Therefore by extracting, at random, N instances or "scenarios" of the uncertainty parameter δ according to some probability P_r we consider only the corresponding constraints in the scenario optimization problem.

Consider the following general Robust Convex Program:

$$RCP: \begin{array}{ll}\min\limits_{\gamma\in R^d} & c^T\gamma \\ \text{s.t.} & f_\delta(\gamma) \le 0, \quad \delta \in \Delta.\end{array}$$

where $f_\delta : \mathbb{R}^d \to \mathbb{R}$ is convex for every $\delta \in \Delta$. The scenario-based approximation is described as Algorithm 1.

Algorithm 1 Scenario Approach [Calafiore and Campi, 2006]

1: Generate N independent identically distributed samples $\delta^{(1)}, \ldots, \delta^{(N)} \in \Delta$, according to P_r
2: Solve the scenario convex program

$$SCP_N: \begin{array}{ll}\min\limits_{\gamma\in\mathbb{R}^d} & c^T\gamma \\ \text{s.t.} & f_{\delta^{(i)}}(\gamma) \le 0, \quad i = 1, \ldots, N.\end{array} \tag{4}$$

It can be seen from (4) (below) that it is a standard finite dimensional convex optimization problem with a finite number of constraints.

3.2 An Exchange Algorithm

To present this approach, we make first some introductory comments on SIP problems. A convex SIP problem is an optimization problem with a finite number of decision variables, but with a feasibility set described by an infinite number of constraints. Its general form is given by

$$SIP(\Theta): \begin{array}{ll}\min\limits_{\gamma} & f(\gamma) \\ \text{s.t.} & g(\gamma, \theta) \le 0, \quad \forall\theta \in \Theta.\end{array} \tag{5}$$

where $f, g(\cdot, \theta) : \mathbb{R}^n \to \mathbb{R}$ are continuous convex functions (for a fixed θ) and Θ is a given nonempty compact set.

There are several algorithms in the literature for solving the SIP (5); see e.g. the survey [Hettich and Kortanek, 1993]. An important class is that of the exchange algorithms. These methods replace (5) by a sequence of finite dimensional convex programs, whose constraints are of the form $g(\gamma, \theta) \le 0$, where θ ranges over a finite subset of Θ. Each program of this sequence is built solving the previous convex program and then adding some "violated" constraints of (5) (i.e., those for which $g(x^*, \theta) > 0$, where x^* is the optimal solution of the previous convex program of the sequence).

Most of the existing exchange algorithms devote most of their computational effort in solving the so-called "auxiliary problem", i.e., in determining the most violated constraint at each iteration. The proposed exchange algorithm, due to Zhang et al. [2010], maintains only the active constraints, i.e., those corresponding to positive Lagrange multipliers, at each step, and no global search for the most violated constraint is needed.

In order to explain the proposed algorithm, denote by Ψ a given finite subset of Θ, say, $\Psi = \{\theta_1, \theta_2, \ldots, \theta_m\}$. We associate with this set the following finitely constrained convex problem:

$$\begin{aligned}\min_{t\in\mathbb{R},\ E\in\mathbb{R}^d}\ & t \\ \text{s.t.}\quad & J\left(\theta_i, \sum_{n=0}^{d-1} \widetilde{M}(\theta_i, \omega_n)E_n\right) \le t, \quad i = 1, \ldots, m \\ & \sum_{n=0}^{d-1} E_n = 1 \\ & E_n \ge 0, \quad n = 0, \ldots, d-1.\end{aligned}$$

This problem is denoted as $RSIP(\Psi)$. Let θ_0 be an arbitrary element of Θ (e.g., its centroid, if Θ is a convex polytope). For this problem, a feasible solution $(t^*, E^*) \in \mathbb{R}^{d+1}$ is optimal if and only if there exist Lagrange multipliers $\mu^* \in \mathbb{R}^m_+$, $\lambda^* \in \mathbb{R}^d_+$, $\nu^* \in \mathbb{R}$ such that $(t^*, E^*, \mu^*, \lambda^*, \nu^*)$ satisfy the Karush-Kuhn-Tucker (KKT) conditions [Boyd and Vandenberghe, 2003] (if Slater's condition is satisfied; see Section 4 below). We are now ready to give the steps of the proposed exchange algorithm.

Algorithm 2 Exchange Algorithm [Zhang et al., 2010]

Require: θ_0, a finite set $\Psi_0 = \{\theta_1^0, \dots, \theta_{m_0}^0\} \subset \Theta$ such that $\theta_0 \in \Psi_0$, and a small number $\varepsilon > 0$

1: $k \leftarrow 0$
2: Solve $RSIP(\Psi_0)$ to obtain an optimum (t^0, E^0)
3: **loop**
4: Find a $\theta_{new}^k \in \Theta$ such that
$$J\left(\theta_{new}^k, \sum_{n=0}^{d-1} \widetilde{M}(\theta_{new}^k, \omega_n) E_n^k\right) - t^k > \varepsilon.$$
5: **if** θ_{new}^k does not exist **then**
6: STOP
7: $\overline{\Psi}_{k+1} \leftarrow \Psi_k \cup \{\theta_{new}^k\}$
8: Solve $RSIP(\overline{\Psi}_{k+1})$ to obtain an optimum (t^{k+1}, E^{k+1}) and Lagrange multipliers $(\mu^{k+1}, \lambda^{k+1}, \nu^{k+1})$
9: $\Psi_{k+1} \leftarrow \{\theta_i \in \overline{\Psi}_{k+1} | \theta_i = \theta_0 \text{ or } \mu_i^{k+1} > 0\}$
10: $k \leftarrow k + 1$

The set Ψ_0, in Step 1, can be chosen, for example, as a uniform grid on Θ. In order to carry out Step 4, following [Zhang et al., 2010], we can choose a set of refined grids of Θ, $T_1 \subset T_2 \subset \cdots \subset T_l$, where $l > 0$ is a predetermined value, and perform the following sub-steps:

Algorithm 3 Steps 4-7 of Algorithm 2

1: $r \leftarrow 1$
2: **loop**
3: **if** $r > l$ **then**
4: STOP
5: **if** there is a $\bar{\theta} \in T_r$ such that $J\left(\bar{\theta}, \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n) E_n^k\right) - t^k > \varepsilon$ **then**
6: $\theta_{new}^k \leftarrow \bar{\theta}$
7: $\overline{\Psi}_{k+1} \leftarrow \Psi_k \cup \{\theta_{new}^k\}$
8: Go to Step 8 of Algorithm 2
9: **if** $\max_{\theta \in T_r} J\left(\theta, \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n) E_n^k\right) - t^k < -\varepsilon$ **then**
10: $r \leftarrow r + 1$
11: **else**
12: Let $\tilde{\theta} \in T_r$ be such that $\left|J\left(\tilde{\theta}, \sum_{n=0}^{d-1} \widetilde{M}(\tilde{\theta}, \omega_n) E_n^k\right) - t^k\right| < \varepsilon$.
13: Apply Newton's method to $\max_{\theta \in \Theta} J\left(\theta, \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n) E_n^k\right)$, starting from $\theta = \tilde{\theta}$
14: **if** Newton's method delivers a $\hat{\theta} \in \Theta$ such that $J\left(\hat{\theta}, \sum_{n=0}^{d-1} \widetilde{M}(\hat{\theta}, \omega_n) E_n^k\right) - t^k > \varepsilon$ **then**
15: $\theta_{new}^k \leftarrow \hat{\theta}$
16: $\overline{\Psi}_{k+1} \leftarrow \Psi_k \cup \{\theta_{new}^k\}$
17: Go to Step 8 of Algorithm 2
18: $r \leftarrow r + 1$

Remark 1. The Lagrange multipliers μ^*, λ^*, ν^* are usually provided by standard convex optimization packages, because they typically rely on primal-dual interior point algorithms. This means that it is not necessary to solve the KKT conditions explicitly in order to obtain such quantities for Step 8 of Algorithm 2. □

4. THEORETICAL PROPERTIES OF THE EXCHANGE METHOD

According to Zhang et al. [2010], under certain conditions, Algorithm 2 terminates in a finite number of iterations, if $\Psi_1 \setminus \{\theta_0\} \neq \emptyset$. Furthermore, if the algorithm terminates, and we denote its result as E_ε^*, then every accumulation point of E_ε^* as $\varepsilon \to 0$ is an optimal solution of $RSIP(\Theta)$. These conclusions suggest that the choice of a ε should give a good approximation to the $RSIP(\Theta)$, which is the problem we are actually interested in.

A set of conditions under which these results hold are:

- $J : \Theta \times \mathbb{R}^{n \times n} \to \mathbb{R}_0^+ \cup \{+\infty\}$ is continuous, strictly convex, and such that $J(\theta, \overline{M}) < \infty$ for all non-singular $\overline{M}$, and that $J(\theta, \cdot)$ is continuously differentiable for all $\theta \in \Theta$.
- $\overline{M}$ is non-singular whenever $E_n > 0$ for all $n = 0, \dots, d-1$.

The first condition is quite mild, and is typically satisfied by most experiment design criteria. The second condition, on the other hand, is satisfied for most rational model structures, since it states that a persistently exciting signal should imply model identifiability [Ljung, 1999].

To verify that the stated conditions imply those required in [Zhang et al., 2010], first notice that $RSIP(\Theta)$ can be written as (5) by replacing E_0 with $1 - \sum_{n=1}^{d-1} E_n$, thus reducing $E \in \mathbb{R}^d$ to $\tilde{E} = (E_1, \dots, E_{d-1}) \in \mathbb{R}^{d-1}$, and where $f(t, E) = t$ and g combines the two groups of inequality constraints in $RSIP(\Theta)$: $J(\theta, \sum_{n=0}^{d-1} \widetilde{M}(\theta, \omega_n) E_n) - t \leq 0$ for all $\theta \in \Theta$, and $-E_n \leq 0$ for $n = 0, \dots, d-1$. This means that Θ has to be replaced by a larger set $\overline{\Theta} \supset \Theta$ containing d additional points, in order to fit the nonnegativity constraints into the structure of (5).

The algorithm in [Zhang et al., 2010] requires a set Ω_0 such that f is level bounded in $SIP(\Omega_0)$. In our case, $\Omega_0 = \{\theta_0\} \cup (\overline{\Theta} \setminus \Theta)$ (i.e., θ_0 and the nonnegativity assumptions on E_n should be included in Ω_0) plays such role, as it can be easily verified for our assumptions.

Most of the conditions in Assumption A of [Zhang et al., 2010] hold. In particular, Assumptions A(i) (convexity and continuous differentiability of f) and A(iv) (level boundedness of f in the feasibility set of $SIP(\Omega)$ can be easily verified. Assumption A(iii) (Slater's condition) also holds, since Slater's condition involves the existence of a feasible solution (t, E) which satisfies all inequality constraints with a strict sign [Boyd and Vandenberghe, 2003], and for such pair we may take $E_n = 1/d$ $(n = 0, \dots, d-1)$ and $t = 1 + \max_{i=1,\dots,m} J(\theta_i, \sum_{n=0}^{d-1} \widetilde{M}(\theta_i, \omega_n) E_n)$.

Assumption A(ii) in [Zhang et al., 2010] does not directly hold, but it can be relaxed to the requirement that $g(\cdot, \theta)$ is convex for all $\theta \in \Theta$ and that $\nabla_x g(x, \theta)$ exists and is continuous on $\mathbb{R}^{d-1} \times \Theta_k$, $k = 1, \dots, K$, where $\{\Theta_1, \dots, \Theta_K\}$ is a finite partition of Θ. The proofs in [Zhang et al., 2010] continue to hold without modifications under this relaxed assumption A(ii), which applies to our case.

In order to apply Theorem 3.1 in [Zhang et al., 2010], condition (ii) needs to be relaxed to requiring only that

$J(\theta,\cdot)$ is strictly convex [2] for every $\theta \in \Theta$. Unfortunately, for these relaxed assumptions the proofs of Lemma 3.1 and Theorem 3.1 of [Zhang et al., 2010] need further modifications, and we cannot develop these changes here, for reasons of space.

On the other hand, Theorem 3.2(a) of [Zhang et al., 2010] can be directly applied to our setup. This establishes the conclusions at the beginning of this section.

5. NUMERICAL EXAMPLES

In this section, an example is presented to demonstrate the numerical aspects of the presented methods. This example utilizes a one parameter first order system, which can be solved in principle using linear programming (if the robust experiment design problem is approximated by a finite dimensional program).

Consider a model given by $H(s) = 1$ and

$$G(s) = \frac{1}{s/\theta + 1},$$

where it is assumed that $\theta \in [0.1, 10]$. For this model structure, the 'single frequency' normalized information matrix is given by

$$\widetilde{M}(\theta,\omega) = \frac{\omega^2/\theta^4}{(\omega^2/\theta^2+1)^2}.$$

Consider a criterion of the form

$$J(\theta, \overline{M}(\theta,\Phi_u)) = \frac{1}{\theta^2 \overline{M}(\theta,\Phi_u)}.$$

The reason for multiplying $\overline{M}$ by θ^2 is that $\overline{M}^{-1}$ is a variance measure and thus $[\theta^2\overline{M}]^{-1}$ gives relative (mean square) errors.

This robust experiment design problem can be solved by discretizing the interval for θ, and rewriting the problem as a linear program [Rojas et al., 2007]. This approach is similar to the one described in Subsection 3.1, except for the fact that in [Rojas et al., 2007] a deterministic (in fact, uniform) sampling of the constraints has been used.

We consider an interval $[0.1, 10]$ for the support of Φ_u (which, according to [Rojas et al., 2007], actually contains the optimal spectrum), $d = 30$ and $N = 7864$. The choice of this value for the number of scenarios is motivated in [Welsh and Rojas, 2009]. For this value of N, we also use a distribution P_r, which is uniform on $\ln\theta$. To set up a fair comparison with the MATLAB semi-infinite solver *fseminf*, we use the same number of constraints. We achieve this by a uniform discretization of $[0.1, 10]$ with a spacing equal to $(10 - 0.1)/N$ and by neglecting one of the generated discrete θ's, e.g., the last one. We maintain the "options" structure of this solver to its default values. Finally, for the exchange algorithm we choose Ψ_0 as a uniform grid on $[0.1, 10]$ with spacing equal to $(10 - 0.1)/50$. Solving $RSIP(\Psi_0)$, we obtain the initial solution (t^0, E^0). We set $l = 4$. The set of refined grids $T_1 \subset T_2 \subset T_3 \subset T_4$ is generated via gridding on $[0.1, 10]$, with constant spacing equal to $1/10^l, l = 1, 2, 3, 4$. Instead of solving the auxiliary problem using Newton's method, we employ a more direct technique similar to ideas used in discretization methods for semi-infinite programming problems [Hettich and Kortanek, 1993]: Upon obtaining a $\tilde{\theta} \in T_r$ such that $\left|J\left(\tilde{\theta}, \sum_{n=0}^{d-1} \widetilde{M}(\tilde{\theta},\omega_n)E_n^k\right) - t^k\right| < \varepsilon$, we search for $\arg\max_{\theta\in\Theta} J\left(\theta, \sum_{n=0}^{d-1} \widetilde{M}(\theta,\omega_n)E_n^k\right)$ in a refined grid around $\tilde{\theta}$. The end points of this grid are selected to be $T_r^{previous}(\tilde{\theta})$ and $T_r^{next}(\tilde{\theta})$, i.e., the points of T_r immediately before and after $\tilde{\theta}$, while the spacing is assumed to be $\left(T_r^{next}(\tilde{\theta}) - T_r^{previous}(\tilde{\theta})\right)/1000$.

[2] In [Zhang et al., 2010, Sections 4.2-4.3], Theorem 3.1 is applied, assuming erroneously that the constraint functions are strictly convex.

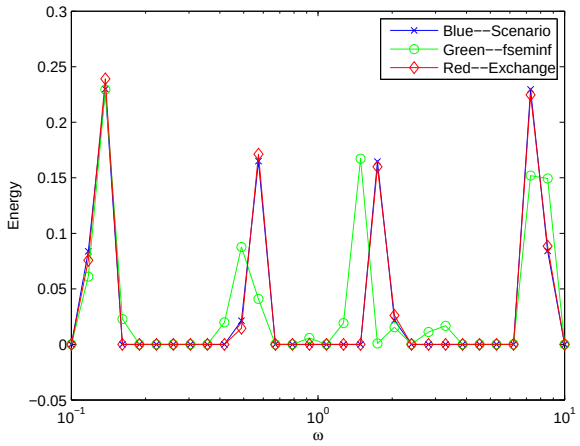

Fig. 1. Values of E as a function of θ.

The three methods give the spectrums presented in Figure 1. We observe that the Scenario approach and the Exchange Algorithm give almost identical spectra. This can also be verified in Figure 2, where the cost $J(\theta, \overline{M}(\theta, \Phi_u^{opt}))$ as a function of θ is demonstrated for all methods, while in every case, we also give the optimal cost produced by SeDuMi.

The methods have been solved using CVX with the SeDuMi as solver in a PC with Intel(R) Core(TM) 2 CPU at 1.83 GHz and 1 Gb of RAM. The Scenario Approach terminated in 121.453 seconds, *fseminf* in 74.093 seconds and the Exchange Algorithm in 48.5 seconds. It can be noted that if Newton's method is employed in the exchange algorithm, then we should get an ever better time to termination [Zhang et al., 2010]. We choose to implement a grid search for the solution of the auxiliary problem due to its direct implementation.

6. CONCLUSIONS

In this paper, the problem of robust experiment design based on semi-infinite programming techniques was investigated. The so-called "scenario" approach was presented, while a new exchange algorithm was proposed for the solution of the same problem. These two methods were compared numerically with the MATLAB semi-infinite solver *fseminf*. The comparison for a simple one parameter example showed that the scenario approach is more accurate than *fseminf*, but less computationally efficient. The exchange algorithm seems to provide an accuracy similar to the scenario approach, but with a much better computationally efficiency. This comparison verifies that the exchange techniques should be further investigated in the context of the robust experiment design problem.

REFERENCES

A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–

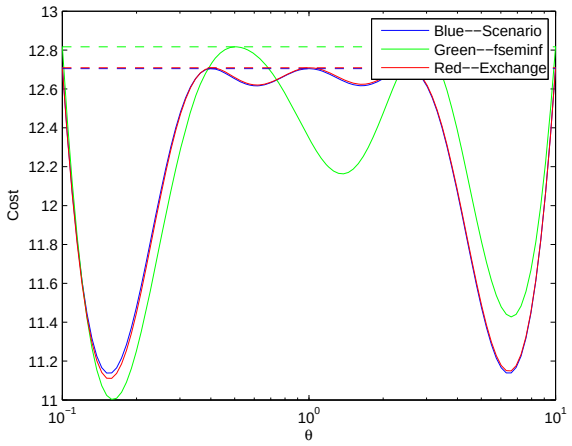

Fig. 2. Cost $J(\theta, \overline{M}(\theta, \Phi_u^{opt}))$ as a function of θ (solid lines), and the corresponding optimal costs (dashed lines).

805, 1998.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2003.

G. Calafiore and M. C. Campi. Uncertain convex programs: Randomized solutions and confidence levels. *Mathematical Programming, Series A*, 102:25–46, 2005.

G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

G. Casella and R. L. Berger. *Statistical Inference, 2nd Edition*. Duxbury, California, 2002.

D. R. Cox. *Planning of Experiments*. Wiley, New York, 1958.

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

R. M. Gagliardi. Input selection for parameter idenfication in discrete systems. *IEEE Transactions on Automatic Control*, AC-12(5):133–147, 1967.

M. Gevers. Identification for control: From the early achievements to the revival of experiment design. *European Journal of Control*, 11:1–18, 2005.

M. Gevers and X. Bombois. Input design: From open-loop to control-oriented design. In *14th IFAC Symposium on System Identification (SYSID'06)*, pages 1329–1334, Newcastle (Australia), 2006.

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

G. C. Goodwin, J. C. Murdoch, and R. L. Payne. Optimal test signal design for linear single input – single output system identification. *International Journal of Control*, 17(1):45–55, 1973.

R. Hettich and K. O. Kortanek. Semi-infinite programming: Theory, methods, and applications. *SIAM Review*, 35(3):380–429, September 1993.

R. Hildebrand and M. Gevers. Minimizing the worst-case ν-gap by optimal input design. In *Proceedings of 13th IFAC Symposium on System Identification*, pages 665–670, 2003a.

R. Hildebrand and M. Gevers. Identification for control: Optimal input design with respect to a worst-case ν-gap cost function. *SIAM Journal on Control and Optimization*, 41(5):1586–1608, 2003b.

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

J. Kiefer. General equivalence theory for optimum designs (aproximate theory). *Annals of Statistics*, 2(5):849–879, 1974.

V. S. Levadi. Design of input signals for parameter estimation. *IEEE Transactions on Automatic Control*, 11(2):205–211, 1966.

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

M. López and G. Still. Semi-infinite programming. *European Journal of Operattional Research*, 180:491–518, November 2006.

J. Mårtensson and H. Hjalmarsson. Robust input design using sum of squares constraints. In *14th IFAC Symposium on System Identification (SYSID'06)*, pages 1352–1357, Newcastle, Australia, 2006.

R. K. Mehra. Optimal input signals for parameter estimation in dynamic systems – survey and new results. *IEEE Transactions on Automatic Control*, 19(6):753–768, December 1974.

L. Pronzato. Optimal experimental design and some related control problems. *Automatica*, 44(2):303–325, February 2008.

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

W. Rudin. *Functional Analysis*. McGraw-Hill, New York, 1973.

S. D. Silvey. *Statistical Inference*. Chapman and Hall, London, 1970.

A. Wald. On the efficient design of statistical investigations. *Annals of Mathematical Statistics*, 14:134–140, 1943.

E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*. Springer-Verlag, Berlin, Heidelberg, New York, 1997.

J. S. Welsh and C. R. Rojas. A scenario based approach to robust experiment design. In *Proceedings of the 15th IFAC Symposium on System Identification*, Saint-Mailo, France, July 2009.

P. Whittle. Some general points in the theory of optimal experimental design. *Journal of the Royal Statistical Society*, 35(1):123–130, 1973.

H. P. Wynn. Results in the theory and construction of D-optimum experimental designs. *Journal of the Royal Statistical Society*, 34(2):133–147, 1972.

M. Zarrop. *Optimal Experiment Design for Dynamic System Identification*, volume 21 of *Lecture Notes in Control and Information*. Springer, Berlin, New York, 1979.

L. Zhang, S.-Y. Wu, and M. A. López. A new exchange method for convex semi-infinite programming. *Siam J. Optim.*, 20(6):2959–2977, October 2010.