Journal of Information and Optimization Sciences
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tios20

Input design for linear dynamic systems using maxmin criteria

Payman Sadegh [a], Lars H. Hansen [a b], Henrik Madsen [a] & Jan Holst [c]

[a] Institute of Mathematical Modeling, Technical University of Denmark, Lyngby, DK-2800, Denmark

[b] Institute of Mathematical Statistics, Lund Institute of Technology, Lund, S-22100, Sweden

[c] Technology R&D, Grundfos A/S, Bjerringbro, DK-8850, Denmark

Published online: 18 Jun 2013.

Input design for linear dynamic systems using maxmin criteria

Payman Sadegh*
Lars H. Hansen*0
Henrik Madsen*
Jan Holst+
*Institute of Mathematical Modeling
Technical University of Denmark
DK-2800 Lyngby
Denmark
+Institute of Mathematical Statistics
Lund Institute of Technology
Lund S-22100
Sweden
0Technology R&D
Grundfos A/S
DK-8850 Bjerringbro
Denmark

ABSTRACT

This paper considers the problem of input design for maximizing the smallest eigenvalue of the information matrix for linear dynamic systems. The optimization of the smallest eigenvalue is of interest in parameter estimation and parameter change detection problems. We describe a simple cutting plane algorithm to determine the optimal frequency power weights of the input, using successive solutions to linear programs. We present a case study related to estimation of thermal parameters of a building.

1. INTRODUCTION

Consider the problem of estimating a set of unknown parameters or detecting parametric changes in a dynamic system based on

Journal of Information & Optimization Sciences
Vol. 19 (1998), No. 2, pp. 223-240
 0252-2667/98 $2.00+.25

experimental input-output data. The accuracy of the estimates or the detectability of changes are often dependent upon the experimental conditions under which the data are collected. We regard the output data as a realization of some random process which is obviously affected by the controlled input to the system. We wish to determine the input sequence so as to maximize the amount of useful information in the output data. Similar to the usual approach of the statistical experiment design literature (see e.g. [22]), we use Fisher's information matrix as a measure of quantifying the amount of information in data. In [22], it is argued that it is not possible to design an experiment to maximize the information matrix in a strong (matrix inequality) sense. Therefore, we consider instead the optimization of some suitable *scalar* function of the information matrix, see e.g. [22] and [6] for a discussion of the widely used and statistically meaningful criteria. A particularly useful and important choice is the smallest eigenvalue of the information matrix or the so called E-optimality criterion (see e.g. [17]). We shall discuss the role of the E-optimality criterion (maxmin design criterion) later in connection with parameter estimation and change detection problems.

The problem of input design has been extensively studied in the literature using different approaches. The statistical approach of the present work is similar to the approach in [9]. Other selected references are [7], [25], [23], [8], [20], which treat different aspects of the input design problem. The present paper is distinguished by the fact that the considered optimality criterion is nonsmooth and special optimization techniques should be employed. The paper indeed shows that the maxmin input design can be addressed within the setting of another extensively studied problem that is maximizing (minimizing) the smallest (largest) eigenvalue of a linear combination of given symmetric matrices (see e.g. [2], [5], [11]). We discuss in some detail a cutting plane algorithm (see [12], [24]) for the optimization of the criterion. The algorithm is both efficient and relatively simple, requiring only successive solutions to linear programs.

The rest of the paper is organized as follows. In Section 2, we state the problem formulation. Section 3 presents the solution. In Section 4, we study the design of optimal inputs for estimating thermal parameters of a building. Finally, Section 5 offers concluding remarks.

2. PROBLEM FORMULATION

Consider a random variable y with the probability density function $h(y \mid \theta)$ where $\theta \in \mathbf{R}^p$ is a (p-dimensional) parameter. We introduce

Definition 1. Fisher's information matrix for the random variable y is defined by

$$M_F = E\left\{\left(\frac{\partial h(y\,|\,\theta)}{\partial\theta}\right)^T \left(\frac{\partial h(y\,|\,\theta}{\partial\theta}\right)\right\} \tag{2.1}$$

where $E\{\cdot\}$ denotes mean value. For a total number of samples N_T, the (average) information matrix per sample is defined by $M = \lim_{N_T\to\infty} M_F/N_T$.

We now consider the linear dynamic system given by

$$y_t = G_1(q^{-1})u_t + G_2(q^{-1})\epsilon_t\,,\quad t = \ldots, -1, 0, 1, \ldots \tag{2.2}$$

where $\{u_t\}$ and $\{y_t\}$ are the input and output sequences respectively, $\{\epsilon_t\}$ is a sequence of Gaussian i.i.d. random variables which without loss of generality can be assumed to have unit covariance, and G_1 and G_2 are transfer functions in the backward shift operator q^{-1}. The transfer functions G_1 and G_2 depend upon the parameter θ.

We shall be concerned with the problem of maximizing $\lambda_{\min}(M)$ with respect to the input sequence, where $\lambda_{\min}(\cdot)$ denotes smallest eigenvalue, and M is the information matrix per sample for the output data of the system (2.2). Note that all the eigenvalues are real since the information matrix is a real symmetric matrix (the information matrix is moreover nonnegative definite). The stated maxmin problem is of interest in a variety of areas such as parameter estimation and change detection.

Parameter Estimation : Consider the system (2.2). Let in the following $\hat{\theta}_{N_T}$ denote the maximum likelihood estimate of θ based on N_T observations, and $||\cdot||$ denote the usual Euclidean norm. It is well-known that under mild stationarity and regularity conditions (see e.g. [9])

$$\sqrt{N_T}(\hat{\theta}_{N_T} - \theta) \xrightarrow{\text{dist}} \Delta\theta \sim N(0, M^{-1}). \tag{2.3}$$

We have that $\Delta\theta^T M\Delta\theta \sim \chi^2_p$, and hence $Pr(\Delta\theta^T M\Delta\theta \le \chi^2_{1-\delta;\,p}) = 1 - \delta$, $0 < \delta < 1$, where $\chi^2_{1-\delta;\,p}$ is the $1-\delta$ fractile of a χ^2-distribution with p degrees of freedom. Obviously

$$\Delta\theta^T M\Delta\theta \ge \lambda_{\min}(M)\ ||\Delta\theta||^2. \tag{2.4}$$

The equality in (2.4) is reached for $\Delta\theta$ being along any eigenvector corresponding to $\lambda_{\min}(M)$. It then follows that for any $0 \leq \delta \leq 1$, we have $||\Delta\theta||^2 \leq \chi^2_{1-\delta;\,p}/\lambda_{\min}(M)$ with probability $1-\delta$. Upon maximizing $\lambda_{\min}(M)$ with respect to the experiment, we minimize the largest (probabilistic) uncertainty bound on the estimate.

Change Detection : Assume that a change in the system given by (2.2) is characterized by a change in the parameter from θ to $\theta + \Delta\theta_c$. Under general regularity conditions, the quantity $\Delta\theta_c^T M_F \Delta\theta_c$ for $\Delta\theta_c \to 0$ tends to the divergence between the model under no change hypothesis and the model under the change hypothesis ([13]) where M_F is the information matrix for the model given by (2.2). It is then obvious that the average value of the divergence per sample tends to $\Delta\theta_c^T M \Delta\theta_c$. Divergence is a suitable measure of the detectability of a parametric change, see e.g. [1]. We have $\Delta\theta_c^T M \Delta\theta_c \geq \lambda_{\min}(M)\ ||\Delta\theta_c||^2$ where the equality is reached for $\Delta\theta_c$ being along an eigenvector corresponding to $\lambda_{\min}(M)$. For any fixed change magnitude $||\Delta\theta_c||$, maximizing the smallest eigenvalue with respect to the input sequence is related to maximizing the smallest (with respect to the direction of $\Delta\theta_c$) divergence between the two models.

We introduce the following assumptions:

A1 : the input and the noise sequences are uncorrelated (i.e., the experiments are performed in open loop),

A2 : the input is generated by a finite register with length N, i.e. the input sequence repeats periodically with cycle N,

A3 : the total number of samples N_T is large,

A4 : the input power is constrained.

We further assume that the general regularity and stationarity conditions that ensure the convergence result (2.3) hold. For simplicity, we restrict attention to single input systems.

Consider the system given by (2.2) and denote the one step ahead prediction error at time t by e_t. Using (2.1), it follows that the information matrix, M_F, for the system is given by ([9])

$$M_F = \sum_{t=0}^{N_T-1} \mu_t \mu_t^T + M_c\,, \tag{2.5}$$

where

$$\mu_t = G_2^{-1}(q^{-1})\left(\frac{\partial G_1(q^{-1})}{\partial\theta}\right)u_t\,, \tag{2.6}$$

and

$$M_c = E\left\{\sum_{t=0}^{N_T-1}\left[-G_2^{-1}(q^{-1})\left(\frac{\partial G_2(q^{-1})}{\partial\theta}\right)e_t\right]\cdot\left[-G_2^{-1}(q^{-1})\left(\frac{\partial G_2(q^{-1})}{\partial\theta}\right)e_t\right]^T \,|\,\theta\right\}.$$

This result is obtained from the definition of the information matrix.

Considerable simplicity is obtained if we represent the input sequence in the frequency domain. The assumption A2 implies that we can represent the input as

$$u_t = c_0 + \sum_{k=2}^{N}(\sqrt{2}c_{k-1})\sin(2\pi(k-1)t/N + \psi_k).$$

Without loss of generality, the input power restriction (see A4) can be expressed as $\sum_{k=1}^{N} c_{k-1}^2 = 1$. Now note that

$$\lim_{N_T\to\infty}\frac{1}{N_T}\sum_{N_T}\sin(2\pi k_1 t/N + \psi)\sin(2\pi k_2(t-T)/N + \psi') = 0$$

for all integer T, all ψ, ψ', and all $k_1, k_2 \in \{0, 1, \ldots, N-1\}$, $k_1 \neq k_2$. Denoting the information matrix per sample corresponding to the input u_t by $M(u_t)$, it follows immediately that $M(\sin(\omega t + \psi)) = M(\sin(\omega t))$ for all ψ and ω. This result together with (2.5) yield

$$M = c_0^2 M(1) + \sum_{k=2}^{N} c_{k-1}^2\, M(\sqrt{2}\sin(2\pi(k-1)t/N)). \tag{2.7}$$

Now, define $\alpha_k = c_{k-1}^2$, $k = 1, \ldots, N$. Then the input power restriction can be written as $\sum_{k=1}^{N}\alpha_k = 1$ and

$$M = \sum_{k=1}^{N} \alpha_k M_k \tag{2.8}$$

where $M_1 = M(1)$, and $M_k = M(\sqrt{2}\sin(2\pi(k-1)t/N)$ for $k \geq 2$. From the input power restriction $\sum_k \alpha_k = 1$ and the equation (2.8), it is evident that the symmetric nonnegative definite information matrix per sample, M, lies in the convex hull of the symmetric nonnegative definite matrices M_k ([9]).

Remark 1. It follows from (2.5) that the actual values of M_k (and M) are dependent upon the true parameter value θ. However, the true parameter is in general unknown at the experiment design stage, especially when the experiment concerns estimation of the parameters. In this paper, we assume that the M_k matrices are evaluated at an *a priori* value for the parameter, say its prior mean. The sensitivity and the robustness of the design to other parameter values should usually be checked, see [20] for the design of robust experiments using a Bayesian formulation.

Remark 2. Using a slightly different assumption than A2, we can obtain a result analogous to (2.8). Assume that the input can be represented as the linear combination $u_t = \sum_{k=0}^{N-1} c_k \phi_t^{(k)}$ where the $\phi_t^{(k)}$ are *given* functions satisfying $\lim_{N_T\to\infty} \sum_{N_T} \phi_t^{(k_1)}\phi_{t-T}^{(k_2)}/N_T = 0$ for all integer T and all $k_1, k_2 \in \{0, 1, ..., N-1\}$, $k_1 \neq k_2$. Again, defining $\alpha_k = c_{k-1}^2$, $k = 1, ..., N$, it is straightforward to show that $M = \sum_{k=1}^{N} \alpha_k M_k$ where M_k is the information matrix per sample under application of the input $\phi_t^{(k-1)}$ to the system. The M_k can be easily obtained using e.g. simulations where the simulations involve application of the input $\phi_t^{(k-1)}$ to the system and calculation of the relevant quantities in (2.5). The numerical case study of the paper (Section 4) illustrates such procedures. Assuming that $\lim_{N_T\to\infty} \sum_{t=1}^{N_T} [\phi_t^{(k_1)}]^2/N_T = \lim_{N_T\to\infty} \sum_{t=1}^{N_T} [\phi_t^{(k_2)}]^2/N_T$ for all $k_1, k_2 \in \{0, 1, ..., N-1\}$, the input power constraint can without loss of generality be stated by $\sum_k \alpha_k = 1$, and the input design problem concerns optimal allocation of the input power among the $\phi_t^{(k)}$. Note than A2 simply implies that we can select $\phi_t^{(0)} = 1$, and $\phi_t^{(k)} = \sqrt{2}\,\sin(2\pi kt/N)$, $k \geq 1$.

Now, denoting $\alpha = (\alpha_1, \ldots, \alpha_N)$, the maxmin problem can be stated as

$$\max_{\alpha \in \mathbf{A}} \left\{ \lambda_{\min}\left(\sum_{k=1}^{N} \alpha_k M_k \right) \right\} \tag{2.9}$$

where $\mathbf{A} = \left\{ \alpha \mid \sum_{k=1}^{N} \alpha_k = 1,\ \alpha_k \geq 0 \right\}$.

The optimization problem (2.9) can be equivalently formulated as a problem with linear objective function as follows. For convenience, we define $f(\alpha) = \lambda_{\min}\left(\sum_{k=1}^{N} \alpha_k M_k \right)$. It is also more convenient to consider the equivalent optimization problem

$$\max_{\alpha \in \mathbf{A}} \{ f(\alpha) \} \tag{2.10}$$

$$f(\alpha) \leq \lambda_{\min}\left(\sum_{k=1}^{N} \alpha_k \, M_k \right).$$

Without loss of generality, we can assume that all the M_k are positive definite implying that $f(\alpha) > 0$ for all $\alpha \in \mathbf{A}$. To ensure the positive definiteness of the M_k, we possibly need to add a constant matrix $\epsilon_0 I$ to each nonnegative definite M_k where I is the unity matrix of proper order and ϵ_0 is some positive number. This modifies the objective function of (2.9) to $\lambda_{\min}\left(\sum_{k=1}^{N} \alpha_k M_k + \epsilon_0 I \right)$. Recalling $\lambda_{\min}(M) = \min_{\|w\|=1} w^T M w$, it follows that the addition of $\epsilon_0 I$ to the M_k merely adds the constant ϵ_0 to the objective function. Assuming the positive definiteness of the M_k, it is then allowable to define the variable $\beta = (\beta_1, \ldots, \beta_N)$, $\beta = \alpha/f(\alpha)$ since $f(\alpha) \geq \epsilon_0$. Since $\sum_{k=1}^{N} \alpha_k = 1$ for all $\alpha \in \mathbf{A}$, we have that $\sum_{k=1}^{N} \beta_k = 1/f(\alpha)$. Furthermore, it holds that $\lambda_{\min}\left(\sum_{k=1}^{N} \alpha_k M_k\right)/f(\alpha) = \lambda_{\min}\left(\sum_{k=1}^{N} \beta_k M_k\right)$, and consequently we obtain the equivalent optimization problem

$$\min_{\beta} \sum_{k=1}^{N} \beta_k$$

$$f(\beta) = \lambda_{\min}\left(\sum_{k=1}^{N} \beta_k M_k\right) \geq 1 \tag{2.11}$$

$$\beta_k \geq 0\,. \qquad k = 1, \ldots, N.$$

Provided that a solution to (2.11) is available, the solution to (2.9) is readily obtained using the simple transformation $\alpha = \beta/\sum_{k=1}^{N} \beta_k$. The equivalence of (2.9) and (2.11) is quite analogous to the equivalence of the matrix games and linear programs in the game theory ([4]).

3. OPTIMIZATION PROCEDURE

The constraint function $f(\beta)$ in (2.11) is nondifferentiable at those values of β where the multiplicity of $\sum_{k=1}^{N_c} \beta_k M_k$ is larger than one (similarly, the objective function of (2.9) is in general nondifferentiable). However, it can be readily verified that $f(\beta)$ is concave, i.e. for $0 \leq \gamma \leq 1$ and any β', β'':

$$f(\gamma\beta' + (1-\gamma)\beta'') \geq \gamma f(\beta') + (1-\gamma) f(\beta'').$$

Subdifferentials of a nonsmooth concave function play the same important role as the gradients of a differentiable function. We therefore introduce the following definition.

Definition 2. The subdifferential of a concave function $F(x)$, $x \in \mathbf{R}^n$, is the set of all vectors $z \in \mathbf{R}^n$, such that $F(v) \leq F(x) +$

$z^T(v - x)$ for all $v \in \mathbf{R}^n$. The subdifferential of F at x is denoted by $\partial F(x)$.

We can compute the subdifferential for the concave function $f(\beta)$ using basic rules of subdifferential calculus, see [19]. Analogous to proposition 2.8.8 of the reference [3], we obtain that at a point β where the multiplicity of the smallest eigenvalue of $\sum_{k=1}^{N} \beta_k M_k$ is equal to r, the subdifferential of f is given by:

$$\partial f(\beta) = co\{\, (w^T Q(\beta)^T M_1 Q(\beta) w, \ldots, w^T\, Q(\beta)^T M_N Q(\beta) w)^T : w \in S_r\} \quad (3.1)$$

where each column of the $p \times r$ matrix $Q(\beta)$ is equal to one of the r orthonormal eigenvectors of $\sum_{k=1}^{N} \beta_k M_k$ corresponding to the smallest eigenvalue (recall that $\sum_{k=1}^{N} \beta_k M_k$ is symmetric), S_r is the r-dimensional unit sphere, and $co\{\cdot\}$ denotes convex hull.

The optimization problem (2.9) can be addressed within the setting of maximizing the smallest eigenvalue of a linear combination of symmetric matrices. In the reference [2], some standard techniques for solving similar nondifferentiable problems are reviewed. Methods based on a smooth approach to nondifferentiable optimization have been recently reported, see e.g. [21], [11]. A particularly simple and efficient method which is suitable for the maxmin optimization of (2.11) is the cutting plane method (Kelley's cutting plane method, see [12]). In the following, we describe the method in some detail.

3.1. *Cutting Plane Algorithm*

In [12] a cutting plane algorithm for optimization problems of the form:

$$\begin{gathered} \min_{\beta} q^T \beta \\ g(\beta) \geq 0, \end{gathered} \quad (3.2)$$

is considered with the assumptions that the scalar valued function $g(\beta)$ is real, continuous, concave, and the set $B = \{\beta : g(\beta) \geq 0\}$ is compact. Moreover, the elements of $\partial g(\beta)$ are assumed to be uniformly bounded on some compact polytope containing B. The general form of the cutting plane algorithm is as follows ([14]):

PROCEDURE 1. *Cutting Plane Algorithm*: Select a polytope P_i containing B.

Step 1 : Minimize $q^T\beta$ over P_i to obtain $\beta^{(i)}$. If $\beta^{(i)} \in B$, then $\beta^{(i)}$ is optimal. Otherwise,

Step 2 : Add the hyperplane $a^{(i)T}(\beta - \beta^{(i)}) + g(\beta^{(i)}) \geq 0$ where $a^{(i)}$ is any element of $\partial g(\beta(i))$ to obtain a new polytope (update P_i) and go to Step 1.

Recalling that $\sum_{k=1}^{N} \beta_k \leq 1/\epsilon_0$ (ϵ_0 is a number such that $f(\alpha) \geq \epsilon_0$), the optimization problem (2.11) can be reformulated as (3.2) by letting $q = (1, \ldots, 1)^T$, and

$$g(\beta) = \min \left(1/\epsilon_0 - \sum_{k=1}^{N} \beta_k, f(\beta) - 1, \beta_1, \ldots, \beta_N \right).$$

It is easy to check that $g(\beta)$, as defined above, is continuous and concave, and the restriction defined by $g(\beta) \geq 0$ is compact. As the start polytope for solving (2.11) using the cutting plane algorithm, we select a polytope defined by the restrictions

$$1/\epsilon_0 - \sum_{k=1}^{N} \beta_k, \beta_1, \ldots, \beta_N \geq 0.$$

At any iteration i, it holds that either the algorithm stops or $f(\beta^{(i)}) < 1$, while $1/\epsilon_0 - \sum_{k=1}^{N} \beta_k^{(i)}, \beta_1^{(i)}, \ldots, \beta_N^{(i)} \geq 0$. This implies that at all the iterations where the algorithm does not stop $g(\beta^{(i)}) = f(\beta^{(i)}) - 1$. Since the convergence proof of Kelley's cutting plane algorithm is based on the asymptotic behavior of a limit sequence of $g(\beta^{(i)})$ (see [12]), the hyperplanes for optimization of (2.11) can be selected as $a^{(i)} \in \partial f(\beta^{(i)})$ (notice that $\partial(f(\beta) - 1) = \partial f(\beta)$). Since, the elements of $\partial f(\beta)$ are uniformly bounded on any compact set, the assumptions for applying Kelley's cutting plane algorithm to the optimization problem (2.11) hold. Furthermore, it can be readily verified that for any

$a^{(i)} \in \partial f(\beta^{(i)})$, we have $a^{(i)T}\beta^{(i)} = f(\beta^{(i)})$. Therefore, the hyperplanes at Step 2 of Procedure (1) are selected as $a^{(i)T}\beta \geq 1$.

Notice that each iteration of the algorithm requires solution to a linear program. The drawback of the method is that the number of constraints of the linear program at each iteration grows with the number of iterations. Simple devices may be used to circumvent this problem, e.g. by deleting the inactive constraints at the end of each iteration (see [14]). Different numerical experimentations indicate the efficiency of the algorithm for the maxmin problem of interest. For a detailed treatment of the cutting plane algorithm where convergence is established under more general conditions, see [24], Chapter 14.

4. CASE STUDY: DOMESTIC HEATING OF A HOUSE

An exemplification of the described theory is given in this case study, which is concerned with the domestic heating of a house. This case study is inspired by a low energy test house at the Department of Buildings and Energy, the Technical University of Denmark. A water based central heating system is used as the domestic heating system.

The low energy house and the central heating system are modelled and implemented in Matlab®. The goal is to find an optimal sequence of pump pressures in order to obtain accurate estimates of some thermal capacities in the house, using (indoor) room temperature measurements.

4.1. *The Model*

This subsection presents the model for the heat transfer dynamics. The house has a ground floor of approximately 120 m^2, and a wooden outer wall which is insulated with 300 mm mineral wool. The power needed to maintain 20°C at an ambient temperature of − 12°C is about 2.5 kW. For details, see e.g. [18], [16], and [15]. The house contains two separate rooms A and B each of 60 m^2.

The modeling objective here is to obtain accurate estimates of the parameters that are related to dominant time constants of the system. Therefore, lumped modeling of the heat transfer will be appropriate, provided that certain conditions hold (see [10]). Based on a second order lumped model for the heat transfer in each room, we obtain the thermal network model illustrated in Figure 1.

In Figure 1, R and C generically denote thermal resistance and thermal heat capacity, respectively. The indices A and B refer to the

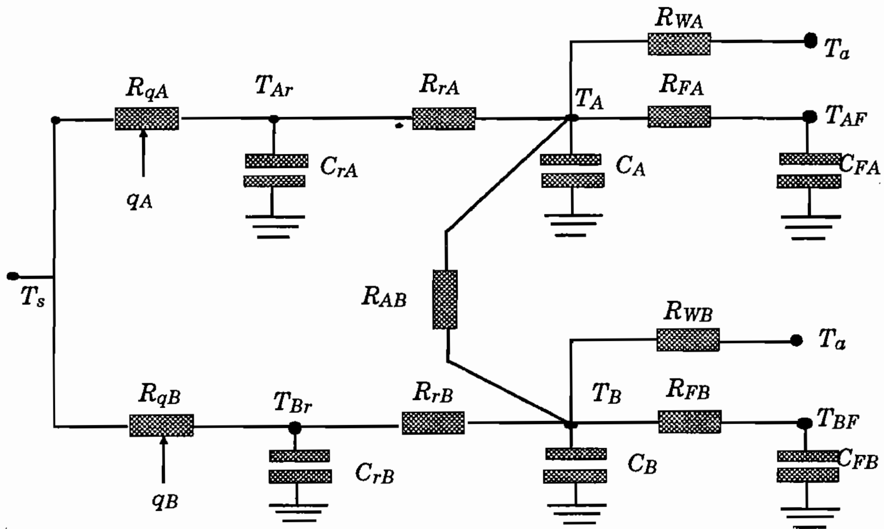

Figure 1. Thermal network equivalent model of the house

rooms A and B, and the indices F, W, r and q refer to the floor, the outer wall, the radiator, and the flow in the radiator, respectively, As indicated in Figure 1, R_{qA} and R_{qB} are dependent upon the actual flows. This makes the model nonlinear in q_A and q_B. Finally, T_s denotes the temperature of the supply water from the boiler and T_a denotes the ambient temperature. The outputs (measurements) are the two room temperatures T_A and T_B. The measurements are taken in the presence of mutually uncorrelated i.i.d. Gaussian noise with unit covariance.

Based on Figure 1, the following coupled first order differential equations for the room A can be derived

$$
\begin{aligned}
C_{rA}\frac{dT_{Ar}}{dt} &= \frac{1}{R_{qA}}(T_s - T_{Ar}) + \frac{1}{R_{rA}}(T_A - T_{Ar}) \\
C_{FA}\frac{dT_{AF}}{dt} &= \frac{1}{R_{FA}}(T_A - T_{AF}) \qquad (4.1) \\
C_A\frac{dT_A}{dt} &= \frac{1}{R_{rA}}(T_{Ar} - T_A) + \frac{1}{R_{FA}}(T_{AF} - T_A) + \frac{1}{R_{WA}}(T_a - T_A) \\
&\quad + \frac{1}{R_{AB}}(T_B - T_A).
\end{aligned}
$$

The relationship between the resistance R_{qA} and the flow q_A is given by $R_{qA} = 1/c_p\rho q_A$ where c_p and ρ denote the specific heat capacity and density of water, respectively. Identical equations hold for the room B. The total hydraulic flow to the radiator, q, is obviously the sum of q_A and q_B. The relationship between q_A and q (q_B and q) is in general nonlinear. However, assuming small flow perturbations for q_A and q_B around some nominal values allows the linearizations $\Delta q_A = k_A \Delta q$ and $\Delta q_B = k_B \Delta q$, where Δq_A, Δq_B, and Δ_q denote perturbations around the nominal values of q_A, q_B, and q respectively, and $k_A + k_B = 1$. The small perturbation assumption also allows linearization of the system of equations (4.1) and the similar equations corresponding to the room B. We therefore obtain a total linear model from the pump pressure perturbations to the indoor temperatures. The smallest time constant for the total linearized model is approximately 4 minutes (see Appendix for numerical values) which allows a sampling time of 1 minute.

4.2. *Optimal Design of Inputs*

The pump pressure around the nominal value (used for linearization) is the designed input to the system. We denote the designed input by Δp_t. The input power restriction is a realistic constraint in this case study, implying restricted pump power.

A complete design of inputs should be based on including all the unknown physical constants in the parameter vector θ. However, the physical knowledge of the system confirmed by numerical experimentation shows that the worst estimable parameters are related to the slow dynamics of the system (due to the large floor capacities). Therefore, we select $\theta = (C_{FA}, C_{FB})^T$ as the parameter vector.

In order to design optimal inputs, it is required to specify the functions $\phi_t^{(k)}$ and compute the corresponding matrices M_k (see Remark 2). We select $\phi_t^{(k-1)}$ as $25c_k \sin(\omega_k t)$, i.e. the input is represented as

$$\Delta p_t = 25 \sum_k c_k \sin(\omega_k t) \ \text{[mBar]} \tag{4.2}$$

where $\sum c_k^2 = 1$, and the frequencies ω_k, $k \in \{1, ..., 20\}$, are selected as $\omega_k = \dfrac{2\pi}{60 \times 50 \times k}$. The reason for this selection is that each sinusoid $\sin(\omega_k t)$ has a period of $50 \times k$ hours, and the largest time constant of the system ($\simeq$ 180 hours) is included within the time range [50,1000] hours. The above values for the frequencies ω_k are used throughout the case study. The M_k matrices are computed by numerical differentiation (with respect to θ) of the simulated noise free output, under the application of the input 25 $\sin(\omega_k t)$, see (2.6) and note that in this example $G_2(q^{-1}) = I$.

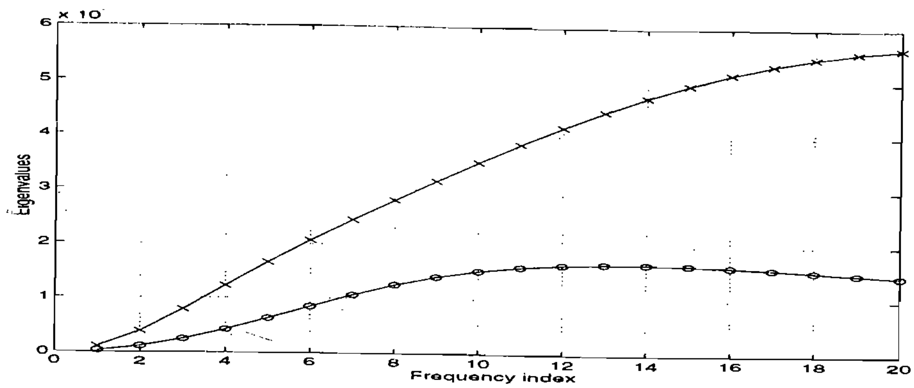

Figure 2. Eigenvalues of M_k versus k

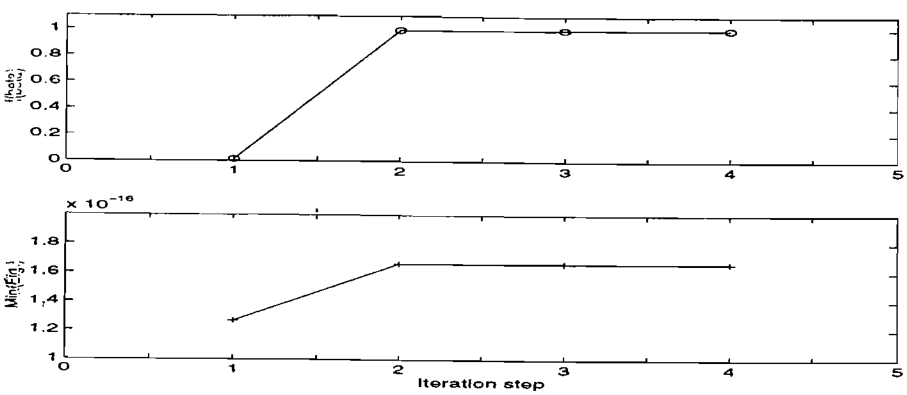

Figure 3. Convergence of the cutting plane algorithm

4.3. *Simulation results*

In Figure 2, the two eigenvalues of M_k versus the (frequency) index k are shown and at $k \simeq 13$ the smallest eigenvalue has a maximum with multiplicity one.

The behaviour of the cutting plane algorithm is illustrated in Figure 3. The top figure shows $f(\beta^{(i)})$ as a function of iteration number i, while the bottom figure shows $\lambda_{\min}\left(\sum_k a_k^{(i)} M_k\right)$ as a function of i.

Notice that the value of $f(\beta)$ can be used as a stop criterion for the algorithm (convergence follows if $f(\beta) \geq 1$). The optimal input is computed to be

$$\Delta p_t = 25\left(0.74 \sin\left(\frac{2\pi t}{60 \times 50 \times 13}\right) + 0.67 \sin\left(\frac{2\pi t}{60 \times 50 \times 14}\right)\right).$$

The result is in agreement with the optimal input suggested by Figure 2.

Let us now examine the case where R_{AB} tends to infinity which means that the two rooms become thermally independent. The eigenvalues of M_k are plotted as a function of k in Figure 4. In this case the maximum of the smallest eigenvalue has multiplicity 2 at $k \simeq 16$.

Figure 5 illustrates $f(\beta^{(i)})$ (top figure) and $\lambda_{\min}\left(\sum_k \alpha_k^{(i)} M_k\right)$ (bottom figure) as a function of i. The optimal solution given by the cutting plane algorithm is

$$\Delta p_t = 25\left(0.89 \sin\left(\frac{2\pi t}{60 \times 50 \times 16}\right) + 0.45 \sin\left(\frac{2\pi t}{60 \times 50 \times 17}\right)\right).$$

which is in agreement with Figure 4.

5. CONCLUDING REMARKS

We have studied the problem of input design for maximizing the smallest eigenvalue of the information matrix. It is established that the design problem can be addressed within the setting of maximizing the smallest eigenvalue of a linear (indeed convex) combination of given symmetric (nonnegative definite) matrices. We have presented a cutting plane algorithm for the optimization, requiring only successive solutions to linear programs. Numerical experience indicates the efficiency of the algorithm for input design problem. The method is illustrated by a case study related to domestic heating of a house.

APPENDIX

Here we list the thermal data of the test house which are used throughout the case study.

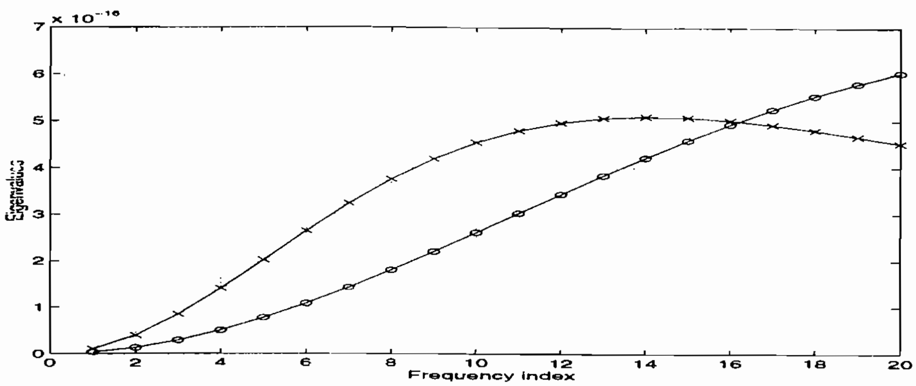

Figure 4. Eigenvalues of M_k versus k

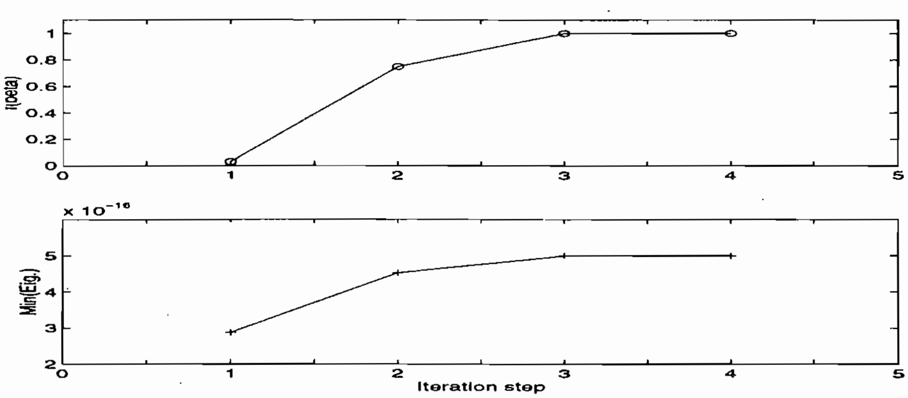

Figure 5. Convergence of the cutting plane algorithm

From measurements on the central heating system it is known that a pump pressure of $\Delta p = 0.1$ Bar yields $q_0 = 28.7l/h$. These values are used as nominal values. At $q_0 = 28.7l/h$, the flow fractions are given by $k_A = 0.25$ and $k_B = 0.75$.

Parameter	value	unit
C_{rA} , C_{rB}	6.9	kJ/K
C_A, C_B	158	kJ/K
C_{FA}	11.6	MJ/K
C_{FB}	5.80	MJ/K
R_{rA} , R_{rB}	0.333	K/W
R_{AB}	0.15	K/W
R_{WA} , R_{WB}	0.186	K/W
R_{FA} , R_{FB}	5.56	mK/W
c_p	4.2	kJ/(kg K)
ρ	992	kg/m^3

REFERENCES

1. M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1993.
2. S. Boyd and Q. Yang, Structured and simultaneous Lyapunov functions for stability problems, *International Journal of Control,* Vol. 49, pp. 2215-2240, 1989.
3. F. H. Clarke, *Optimization and Nonsmooth Analysis,* Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley and Sons, New York, 1983.
4. G. B. Dantzig, *Linear Programming and Extensions,* Princeton University Press, 1963.
5. M. K. H. Fan and B. Nekooie, On minimizing the largest eigenvalue of a symmetric matrix, *Linear Algebra and its Applications,* Vol. 214, p. 225, January 1992.
6. V. V. Fedorov, *Theory of Optimal Experiments,* Academic, New York, 1972.
7. M. Gevers and L. Ljung, Optimal experiment design with respect to the intended model application, *Automatica,* Vol. 22(5), pp. 543-554, 1986.
8. K. R. Godfrey, *Perturbation Signals for System Identification,* Prentice Hall International Series in Acoustics, Speech, and Signal Processing, Prentice Hall, 1993.
9. Graham C. Goodwin and Robert L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis,* Academic, New York, 1977.
10. L. H. Hansen, Preliminary analysis of a low energy test house, IMM and Grundfos A/S, July 1996.
11. F. Jarre, An interior point method for minimizing the maximum eigenvalue of a linear combination of matrices, *SIAM Journal on Control and Optimization,* Vol. 31 (5), pp. 1360-1377, 1993.

12. J. E. Kelley, The cutting plane method for solving convex problems, *J. Soc. Indus, Appl. Math.*, Vol. VIII (4), pp. 703-712, 1960.

13. S. Kullback, *Information Theory and Statistics*, John Wiley & Sons Ltd., 1959.

14. D. Luenberger, *Linear and Nonlinear Programming*, 2nd Edition, Addison Wesley Publishing Company, Inc., 1984.

15. H. Madsen, A. A. Nielsen and B. Saxhof, Identification of models for the heat dynamics of building, to appear.

16. H. Madsen, H. Melgaard and J. Holst, Identification of building performance parameters, in J. J. Bloem (ed.), *Workshop on Advanced Identification Tools in Solar Energy Research*, Non Nuclear Energy, pp. 37-60, Commission of the European Communities, DG XII, 1990.

17. A. Pazman, *Foundations of Optimum Experimental Design*, D. Reidel Publishing Company, 1986.

18. N. H. Rasmussen and B. Saxhof, Simultaneous testing of heating systems, Technical Report 128, Thermal Insulation Laboratory, The Technical University of Denmark, 1982.

19. R. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

20. P. Sadegh, J. Holst, H. Madsen and H. Melgaard, Experiment design for greybox identification, *International Journal of Adaptive Control and Signal Processing*, Vol. 9 (6), pp. 491-507, 1995.

21. A. Shapiro and M. K. H. Fan, On eigenvalue optimization, *SIAM Journal on Optimization*, Vol. 5(3), pp. 552-569, 1995.

22. S. D. Silvey, *Optimal Design*, Chapman & Hall, 1980.

23. Herbet J. A. F. Tulleken, Generalized binary noise test-signal concept for improved identification-experiment design, *Automatica*, Vol. 26 (1), pp. 37-49, 1990.

24. W. I. Zangwill, *Nonlinear Programming: A Unified Approach*, Prentice Hall, Inc., Englewood Cliffs, N. J., 1969.

25. X. J. Zhang, *Auxiliary Signal Design in Fault Detection and Diagnosis*, Vol. 134 of Lecture Notes in Control and Information Sciences, Springer-Verlag, 1989.

Received September, 1996