Identification for control: Optimal input design with respect to a worst-case ν-gap cost function

R. Hildebrand †
Center for Operations Research and Econometrics (CORE)
Université Catholique de Louvain
1348 Louvain-la-Neuve, Belgium
hildebrand@core.ucl.ac.be

M. Gevers ‡
Centre for Systems Engineering and Applied Mechanics (CESAME)
Université Catholique de Louvain
1348 Louvain-la-Neuve, Belgium
gevers@auto.ucl.ac.be

***Abstract*— The aim of this contribution is to demonstrate efficient applicability of modern convex optimization techniques in control theory. We solve the problem of designing an input for a parameter identification experiment such that the worst-case ν-gap over all plants in the resulting uncertainty region between the identified plant and plants in this region is as small as possible. The motivation for choosing this cost criterion is robust controller design, where the controller has to stabilize all plants in the identified uncertainty region.**

I. Introduction

In this contribution we deal with a problem that connects prediction error identification methods with robust control theory. A series of investigations in this direction has been undertaken recently [2]. In this work we focus on computational aspects, specifically we show that the existing apparatus of convex analysis is capable of tackling this kind of problems efficiently.

Subject to investigation are discrete time SISO real-rational stable LTI plants, which are to be identified in open loop within an ARX model structure. We assume the true plant to lie in the model set. Hence the model error is determined only by the covariance of the estimated parameter vector.

Since the aim of the identification experiment is control design, it is desirable to obtain an uncertainty region with good stability robustness properties. The set of controllers that stabilize all models in the uncertainty set should be large. A suitable measure of robust stability that allows one to connect the "size" of an uncertainty set with a set of robustly stabilizing controllers is the worst-case ν-gap $\delta_{WC}(\hat{G},\mathscr{D})$ introduced in [2]. It is the supremum of the Vinnicombe ν-gap [10] between the identified model $\hat{G}$ and all plants in the uncertainty set $\mathscr{D}$ which emerges from the experiment. The problem we deal with is to minimize the worst-case ν-gap of the uncertainty region $\mathscr{D}$ by choosing a suitable input $u(t)$ for the identification experiment.

† The European Commission is herewith acknowledged for its financial support in part to the research reported on in this contribution. The support is provided via the Program Training and Mobility of Researchers (TMR) and Project System Identification (ERB FMRX CT98 0206) to the European Research Network System Identification (ERNSI).

‡ This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors.

The problem setting of experiment design first arose in statistics and was extensively studied throughout the last century. We adopt the most common viewpoint and optimize the input power spectrum with respect to a cost function that depends on the average per data sample information matrix $\bar{M}$ of the experiment. This matrix is defined as the limit of the ratio between the information matrix and the number of data as the number of data tends to infinity (see e.g. [12]). Thus we effectively optimize the average information matrix with respect to the considered cost function, and then construct an input power spectrum and an input that produces this information matrix.

For different classes of cost functions iterative procedures were designed to find the optimal input power spectrum up to a prespecified precision. Most common cost functions are $\ln(\det\bar{M}^{-1})$ (D-optimality), $\mathrm{tr}\bar{M}^{-1}$ (A-optimality), $\mathrm{tr}W\bar{M}^{-1}$, where $W \geq 0$ (L-optimality), $\lambda_{\max}(\bar{M}^{-1})$ (E-optimality). All mentioned cost functions depend analytically on the entries of $\bar{M}$ and Kiefer-Wolfowitz theory can effectively be applied to them (see [5]). These criteria are convex and monotonic with respect to $\bar{M}$ (see [12, p.39]).

In this contribution, we optimize the input power spectrum with respect to the worst-case ν-gap of the uncertainty region $\mathscr{D}$. We shall also introduce another cost function, which approximates the worst-case ν-gap, but is somewhat simpler. Both cost functions are compound criteria (see [5, section 4G]) and application of Kiefer-Wolfowitz theory does not make them more tractable. However, the proposed criteria satisfy the natural condition of monotonicity with respect to $\bar{M}$, as well as the condition of quasiconvexity, which is slightly weaker than convexity.

To tackle the considered problem we will use the theory of Tchebycheff systems and their moment spaces. The set of possible average information matrices $\bar{M}$ can be represented as the feasible set of a linear matrix inequality (LMI) [4, chapter VI, Theorem 4.1]. This allows to apply convex analysis and the theory of LMI's to this optimization problem. For recent results in convex optimization see e.g. [8].

It follows from a well-known fact of Tchebycheff system theory that any admissible average information matrix $\bar{M}$ can be obtained by applying an input with discrete power spectrum, and that there exist admissible $\bar{M}$ which can be realized only by discrete power spectra. A restatement of this assertion is provided in Theorem 1 in this contribution. In view of this, we propose an algorithm that yields optimal input power spectra which are discrete. Given the result just quoted, this is in no way a restriction. There are different ways to choose an input sequence with a desired power spectrum. We can choose the input e.g. as a multisine function. However, in many cases one could use also binary signals (see e.g. [12, p.29]) or other functions. For a comprehensive treatment of Tchebycheff systems see textbook [4] by Karlin and Studden.

In the last years several authors successfully treated input design problems arising in Identification for Control with convex optimization methods. In [6], the input spectrum for an open loop identification experiment was designed to minimize the closed-loop system performance. By a Taylor series truncation, the cost function reduced to the weighted-trace criterion (L-optimality). However, the input spectra were restricted to those which can be realized by white noise filtered through an FIR filter. An LMI description of the corresponding set of information matrices can be derived from the positive-real lemma [1],[11].

We stress that the assumption of an ARX model structure and an input energy constraint are in no way restrictive. The ideas and methods proposed here easily carry over to other model structures and to input power or output power/energy constraints.

The remainder is structured as follows. In the next section the considered identification problem as well as the cost functions are formally defined. In section 3 we show that the set over which the optimization takes place is amenable to an LMI formulation. In section 4 we prove that the optimization problem is quasiconvex. In section 5 we construct cutting planes to the different cost functions. Sections 3 to 5 are the key part. The results obtained therein allow the problem to be treated with standard convex analysis methods. Since the optimization takes place in an abstract parameter space, it is necessary to convert values in this space into power spectra and input sequences. This task is accomplished in section 6. In Section 7 we demonstrate the benefits of the computed input in a numerical example. Finally, in section 8 we draw some conclusions.

II. Problem Setting

Let us consider an ARX model structure

$$\begin{aligned} y(t)+a_1y(t-1)+\cdots+a_{n_a}y(t-n_a) &= \\ = \quad b_1u(t-n_k)+\cdots+b_{n_b}u(t-n_k-n_b+1)+e(t), \end{aligned}$$

where $u(t)$ is the input signal, $y(t)$ is the output signal, both onedimensional, $\theta=(a_1,\ldots,a_{n_a},b_1,\ldots,b_{n_b})^T$ is the parameter vector, and $e(t)$ is normally distributed white noise with covariance λ_0. Let us assume that the true system dynamics can be described within this structure and corresponds to a parameter value $\theta=\theta_0$. Assume further that the true system is stable. Denote by z^{-1} the delay operator. Then we can write

$$y=z^{-n_k+1}\frac{B(\theta)}{A(\theta)}u+\frac{1}{A(\theta)}e=G(\theta)u+\frac{1}{A(\theta)}e,$$

where A,B are obviously defined polynomials in the delay operator. Note that by our stability assumption A has no zeros on the unit circle and hence $|A|^2$ is strictly positive there.

Suppose an identification experiment with input $(u(1),\ldots,u(N))$ is performed, leading to an observed output $(y(1),\ldots,y(N))$ with N data samples, where $u(t)$ is quasistationary with power spectrum Φ_u. Suppose a parameter estimate $\hat{\theta}$ is obtained by least squares prediction error minimization. Then it is well-known [7] that the estimate $\hat{\theta}$ is asymptotically unbiased as $N\to\infty$ and its covariance for large N is given by $E(\theta_0-\hat{\theta})(\theta_0-\hat{\theta})^T\approx\frac{\lambda_0}{N}(\bar{E}\psi\psi^T)^{-1}$, where ψ^T is the gradient of the predictor with respect to θ at $\theta=\theta_0$. The asymptotic expression for the parameter covariance is then a function of the input power spectrum and the true values of the coefficients of A and B [7]. The inverse of the parameter covariance matrix is the Fisher information matrix. Let us denote the asymptotic expression for the information matrix by M and the average information matrix per data sample (see e.g. [12, p.24]) by $\bar{M}$, $\bar{M}=\frac{1}{N}M$.

Since the parameter estimate $\hat{\theta}$ is asymptotically normally distributed [7], we can assume, following [2], that the true parameter vector θ_0 lies with a prespecified probability $\alpha\in(0,1)$ in the uncertainty ellipsoid

$$U=\left\{\theta\,|\,\frac{N}{\chi^2_{n_a+n_b}(\alpha)}(\theta-\hat{\theta})^T\bar{M}(\theta-\hat{\theta})<1\right\}, \qquad (1)$$

where χ^2_l is the χ^2 probability distribution with l degrees of freedom.

The uncertainty ellipsoid U corresponds to an uncertainty set $\mathscr{D}=\left\{G(z,\theta)=z^{-n_k+1}\frac{B(\theta)}{A(\theta)}|\theta\in U\right\}$in the space of transfer functions.

The worst-case ν-gap between the identified model $G(\hat{\theta})$ and the uncertainty region $\mathscr{D}$ is defined by

$$\delta_{WC}(G(\hat{\theta}),\mathscr{D})=\sup_{\theta\in U}\delta_\nu(G(\hat{\theta}),G(\theta)), \qquad (2)$$

where δ_ν denotes the Vinnicombe ν-gap between two plants [10]. Since $G(\hat{\theta})$ belongs to $\mathscr{D}$, the worst-case ν-gap can be expressed in the following way [2, Lemma 5.1].

$$\delta_{WC}(G(\hat{\theta}),\mathscr{D})=\sup_{\omega\in[0,\pi]}\kappa_{WC}(G(e^{j\omega},\hat{\theta}),\mathscr{D}), \qquad (3)$$

where $\kappa_{WC}(G(e^{j\omega},\hat{\theta}),\mathscr{D})$ is called the worst-case chordal distance between $G(\hat{\theta})$ and $\mathscr{D}$ at frequency ω and is defined by

$$\sup_{\theta\in U}\frac{|G(e^{j\omega},\hat{\theta})-G(e^{j\omega},\theta)|}{\sqrt{(1+|G(e^{j\omega},\hat{\theta})|^2)(1+|G(e^{j\omega},\theta)|^2)}}. \qquad (4)$$

Our goal shall be to minimize the quantity $\delta_{WC}(G(\hat{\theta}),\mathscr{D}) = \max_{\omega\in[0,\pi]} \kappa_{WC}(G(e^{j\omega},\hat{\theta}),\mathscr{D})$ by choosing an input with an appropriate power spectrum.

To restrict the class of admissible power spectra we impose an input energy constraint

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_u(\omega)d\omega \le c, \tag{5}$$

where $c > 0$ is a prespecified positive constant.

Problem 1 Find Φ_u satisfying (5) such that $\bar{M}(\Phi_u)$ minimizes the cost function $\mathscr{J}_1 = \delta_{WC}(G(\hat{\theta}),\mathscr{D})$ defined by equations (3),(4).

Along with the worst-case ν-gap of the uncertainty region $\mathscr{D}$, we will consider another cost function, which is easier to compute and is an approximation of δ_{WC}. For a fixed positive definite matrix $\bar{M}_0$ the size of the parameter ellipsoid U defined by any multiple $\bar{M} = \beta\bar{M}_0$ of $\bar{M}_0$, where $\beta > 0$, is proportional to $\beta^{-1/2}$. Since for small ellipsoids the worst-case ν-gap is asymptotically proportional to the size of the former, it follows that for large β the value of $\mathscr{J}_1(\bar{M})$ diminishes asymptotically proportionately to $\beta^{-1/2}$. Thus we can approximate $\mathscr{J}_1$ by the leading Taylor series term

$$\mathscr{J}_2 = \lim_{\varepsilon\to 0}\frac{\mathscr{J}_1(\varepsilon^{-2}\bar{M})}{\varepsilon}. \tag{6}$$

Problem 2 Find Φ_u satisfying (5) such that $\bar{M}(\Phi_u)$ minimizes cost function $\mathscr{J}_2$ defined by equation (6).

The goal of the present contribution is the development of numerical algorithms for solving both Problems 1 and 2. There is a two-fold reason for introducing cost function $\mathscr{J}_2$. Beside its much lower computational complexity, it turns out that identification with an input power spectrum minimizing $\mathscr{J}_2$ in many cases gives better results than one with an input power spectrum minimizing $\mathscr{J}_1$. This apparently counter-intuitive observation has the following reason. Both cost functions depend on the identified parameter value $\hat{\theta}$, the true parameter value θ_0 and the noise covariance λ_0. These quantities are unknown and must be replaced by estimates obtained e.g. from a preliminary identification experiment. This approximation introduces an error to the argument of the minimum of the cost functions $\mathscr{J}_1$ and $\mathscr{J}_2$, i.e. to the solutions of Problems 1 and 2. Now simulations show that the impact of this effect on $\arg\min\mathscr{J}_2$ is lower than that on $\arg\min\mathscr{J}_1$ and that this difference as a rule overweighs the error introduced by approximating cost function $\mathscr{J}_1$ by $\mathscr{J}_2$. We will address this issue again in the simulation section.

III. LMI DESCRIPTION OF THE SEARCH SPACE

In this section we shall describe the set of possible average information matrices $\bar{M}$, over which the optimization takes place, as the feasible set of an LMI.

Proposition 1: [9] The average information matrix $\bar{M}$ is contained in a (n_a+n_b)-dimensional affine subspace of the space of symmetric $(n_a+n_b)\times(n_a+n_b)$-matrices.

This subspace can be parameterized by the *trigonometric moments* of the measure $\frac{\Phi_u}{\pi\lambda_0|A|^2}$, i.e. the numbers $x_k = \frac{1}{\pi}\int_0^{\pi}\frac{\Phi_u}{\lambda_0|A|^2}\cos(k\omega)\,d\omega$, $k = 0,\dots,n$, where $n = n_a+n_b-1$. Let us compose a vector $\tilde{x}\in\mathbf{R}^{n+1}$ of the real numbers x_k, $k = 0,\dots,n$. It lies in the moment space $\mathscr{M}^{(n+1)}$ of the Tchebycheff system $\{1,\cos\omega,\dots,\cos n\omega\}$ on $[0,\pi]$ (see e.g. [12]). Thus the set of feasible information matrices $\bar{M}$ is the affine image of the trigonometric moment cone $\mathscr{M}^{(n+1)}$. It therefore can be characterized as the feasible set of an LMI (see e.g. [4, Chapter VI, Theorem 4.1]). Let us denote the interior of the feasible set by $\mathscr{M}$.

Definition 1: (see e.g. [4]) Let Φ_u be a discrete power spectrum with support $supp\,\Phi_u \subset [0,\pi]$. The number $\#[supp\,\Phi_u\cap(0,\pi)] + \frac{1}{2}\#[supp\,\Phi_u\cap\{0,\pi\}]$, where # denotes the cardinality, is called the *index* of Φ_u.

The notion of the index allows us to characterize the interior of the moment space $\mathscr{M}^{(n+1)}$. The following theorem is a standard result on moment spaces.

Theorem 1: (see e.g. [4]) Let $\tilde{x}$ be a point in $\mathscr{M}^{(n+1)}$. Then the following conditions hold.

i) $\tilde{x}\in Bd(\mathscr{M}^{(n+1)})$ if and only if there exists a discrete nonnegative measure on $[0,\pi]$ with index less than $\frac{n+1}{2}$ that induces $\tilde{x}$. This measure is unique.

ii) $\tilde{x}\in Int(\mathscr{M}^{(n+1)})$ if and only if there exists a discrete nonnegative measure on $[0,\pi]$ with index $\frac{n+1}{2}$ that induces $\tilde{x}$. There are exactly two such measures. Exactly one of them contains the frequency π.

iii) Let $\tilde{x}\in Int(\mathscr{M}^{(n+1)})$ and $\omega\in[0,\pi]$. Then there exists a unique discrete nonnegative measure on $[0,\pi]$ which induces $\tilde{x}$, has index not exceeding $\frac{n+2}{2}$, and contains the frequency ω. □

Proposition 2: Let Φ_u be a power spectrum and $\bar{M}$ the corresponding average information matrix. Then $\bar{M}$ is singular if and only if Φ_u is discrete and its index is less than $\frac{n_b}{2}$.

The proposition follows from the above theorem by considering the special structure of $\bar{M}$.

Corollary 1: Any $\bar{M}\in\mathscr{M}$ is strictly positive definite.

This corollary ensures the existence of the inverse $\bar{M}^{-1}$ in the interior of the search space.

The minimum of the considered cost functions under constraint (5) is attained when equality holds, i.e. we can replace (5) by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_u(\omega)d\omega = c. \tag{7}$$

This determines an affine hyperplane in the space of feasible average information matrices [12]. Moreover, (7) defines sections of the moment cone. Expressing the variable x_0 affinely through $x_1,\dots,x_n$, we obtain a compact feasible set described by an LMI on the variables $x_1,\dots,x_n$. Denote by $\mathscr{X}_c$ the interior of this set and by $\mathscr{M}_c$ the corresponding set of information matrices $\bar{M} = \bar{M}_0 + \sum_{i=1}^{n} x_i\bar{M}_i$. Here $\bar{M}_0$, $\bar{M}_i$ are known constant matrices.

Thus we reduced the infinite-dimensional problem of searching the minimum of the cost functions over the set of all admissible input power spectra to the finite-dimensional problem of searching the minimum over a convex compact section of the trigonometric moment cone, which can be described by an LMI.

IV. Quasiconvexity

In this section we prove quasiconvexity of cost functions $\mathscr{J}_1, \mathscr{J}_2$ and thus of Problems 1 and 2.

Proposition 3: On $\mathscr{M}$ cost function $\mathscr{J}_1$ is quasiconvex with respect to $\bar{M}$.

The proposition follows from a general assertion on quasiconvexity of cost functions depending on a quasiconvex constraint. Let us consider the following constrained optimization problem.

$$F = \max_{x \in X, g(x,y) \geq 0} f(x), \tag{8}$$

where X is an arbitrary set, $f(x)$ is an arbitrary function, and $g(x,y)$ is a constraint function picked out from a family of constraint functions parameterized by the variable y. The only assumption we make is that $g(x,y)$ is quasiconvex in y. The following lemma is easily proven by set-theoretic arguments.

Lemma 1: The value of problem (8), considered as a function of y, $F = F(y)$, is quasiconvex in y.

Note that cost function $\mathscr{J}_1$ is the maximum of a function of θ over the set U given by (1). But U is defined by an inequality which is linear in $\bar{M}$. Thus the above lemma applies.

Proposition 4: On $\mathscr{M}$ cost function $\mathscr{J}_2$ is quasiconvex with respect to $\bar{M}$.

Proof. Direct calculation shows that $\mathscr{J}_2$ can be expressed as follows.

$$\mathscr{J}_2 = const \cdot \sup_{\omega \in [0,\pi]} \frac{\sqrt{\lambda_{\max}(T(\omega)\bar{M}^{-1}T(\omega)^T)}}{1 + |G(e^{j\omega}, \hat{\theta})|^2},$$

where $T(\omega)$ is a $2 \times (n_a + n_b)$-matrix given by the gradient $\frac{\partial G(e^{j\omega},\theta)}{\partial \theta}$. The inverse P^{-1} of a symmetric positive definite matrix P and the maximal eigenvalue $\lambda_{\max}(Q)$ of a symmetric positive semidefinite matrix Q are convex functions with respect to P or Q respectively. Hence $\lambda_{\max}(T\bar{M}^{-1}T^T)$ is convex with respect to $\bar{M}$ for fixed ω. Since the operation of taking the maximum over a family of functions preserves convexity, we have that $\mathscr{J}_2^2$ is a convex function with respect to $\bar{M}$. This yields quasiconvexity of $\mathscr{J}_2$. □

V. Cutting planes

In this section we provide the necessary tools that allow the user to apply standard convex algorithms to solve Problems 1 and 2 numerically.

Most black-box methods in convex analysis are based on the notion of a cutting plane [1]. If $S \subset \mathbf{R}^m$ is a convex set and $f : S \to \mathbf{R}$ is a quasiconvex function defined on S, then a cutting plane to f at a point $x^{(0)} \in S$ is defined by a nonzero vector $g \in \mathbf{R}^m$ such that $f(x^{(0)}) \leq f(x)$ for any $x \in S$ satisfying the inequality $g^T(x - x^{(0)}) \geq 0$. We compute cutting planes for cost functions $\mathscr{J}_1, \mathscr{J}_2$ at an arbitrary point $x^{(0)} \in \mathscr{X}_c$. Along with the LMI description of the feasible set this allows the user to employ standard convex black-box methods for solving Problems 1 and 2. For a description of different methods see e.g. [1],[8].

Cutting planes for cost function $\mathscr{J}_1$ can be computed using the special structure of this function. Namely, the worst-case chordal distance can be expressed as a solution to a generalized eigenvalue problem (GEVP) [2, Theorem 5.1]. The parameters of this GEVP enter in the components of the normal $g(x)$ to a cutting plane at x. Details can be found in [3] and are omitted here.

Let us now compute a cutting plane for cost function $\mathscr{J}_2$. Denote by $\omega^{(0)}$ the frequency at which the function $\frac{\lambda_{\max}(T(\omega)\bar{M}^{-1}T(\omega)^T)}{(1+|G(e^{j\omega},\hat{\theta})|^2)^2}$ attains its maximum. Let $v \in \mathbf{R}^2$ be a unit length eigenvector to the maximal eigenvalue of the matrix $T(\omega^{(0)})\bar{M}^{-1}T(\omega^{(0)})^T$.

Proposition 5: Let $g \in \mathbf{R}^n$ be defined componentwise by $g_i = -v^T T(\omega^{(0)})\bar{M}^{-1}\bar{M}_i\bar{M}^{-1}T(\omega^{(0)})^T v$. If $g \neq 0$, then g defines a cutting plane for the cost function $\mathscr{J}_2$ at $x^{(0)}$. If $g = 0$, then $\mathscr{J}_2$ attains a minimum at $x^{(0)}$.

The proof is by computing the gradient of the function $f(x) = \mathrm{tr}(T(\omega^{(0)})^T vv^T T(\omega^{(0)})(\bar{M}(x))^{-1})$.

VI. Design of input signals

Now we show how to design an input signal from an obtained solution $x^{(0)} \in \mathscr{X}_c$. By Theorem 1, there exist moment points which can be realized only by discrete spectra. On the other hand, any moment point can be realized by a discrete spectrum. Therefore we propose the following two-step procedure. First a discrete input power spectrum generating the moment point $x^{(0)}$ is computed, and then a multisine input with the desired spectrum is generated. The latter is a standard task.

The point $x^{(0)}$ corresponds to a point $\tilde{x} = (x_0, x_1, \ldots, x_n)$ in moment space $\mathscr{M}^{(n+1)}$. Denote by $\tilde{x}^s(\omega)$ the moment point induced by the design measure that satisfies constraint (7) and concentrates all power at the single frequency ω. The point $\tilde{x}$ is a convex combination $\sum_k \lambda_k \tilde{x}^s(\omega_k)$ of points on the curve $\{\tilde{x}^s(\omega) \,|\, \omega \in [0,\pi]\}$. The weights λ_k and frequencies ω_k determine a power spectrum which induces the moment point $\tilde{x}$.

In order to find λ_k, ω_k we exploit an idea that is used to prove Theorem 1 [4]. Namely, the expression of a point on the boundary of the feasible set as convex combination of points $\tilde{x}^s(\omega)$ is unique and the corresponding frequencies ω_k are the roots of a trigonometric polynomial whose coefficients can be computed from a supporting plane at that point. The weights are obtained by solving a standard linearly constrained least squares problem. But we can easily represent any feasible point $\tilde{x}$ as convex combination of two points on the boundary.

VII. Simulation results

Consider the true system $y = G_0 u + H_0 e = \frac{B(z)}{A(z)} u + \frac{1}{A(z)} e$ with $G_0 = \frac{B(z)}{A(z)} = \frac{0.1047z^{-1}+0.0872z^{-2}}{1-1.5578z^{-1}+0.5769z^{-2}}$. The input u is subject to the energy constraint $Eu^2(t) = 1$, and the noise e has unit variance.

The system is identified within an ARX model structure of order two. The number of data points to be collected is $N = 1000$. We minimize the worst-case ν-gap of the uncertainty region around the identified model corresponding to a confidence level of $\alpha = 0.95$.

In a Monte-Carlo simulation, 500 runs were performed. Each run consisted of five identification experiments: one preliminary and four mutually independent second experiments based on this preliminary experiment, corresponding to the four different cost functions $\mathscr{J}_1$, $\mathscr{J}_2$, D-optimality and E-optimality.

In the preliminary experiment, the input was chosen to be white Gaussian noise with variance 1. The identified parameter vector and noise variance were used as a priori estimates of the true parameter vector and the true noise variance for designing the input power spectrum for the series of second experiments. After each identification experiment the worst-case ν-gap of the identified uncertainty region was recorded.

The noise realizations for the five experiments within one run and for different runs were different, as well as the input realizations for the preliminary experiments of the different runs.

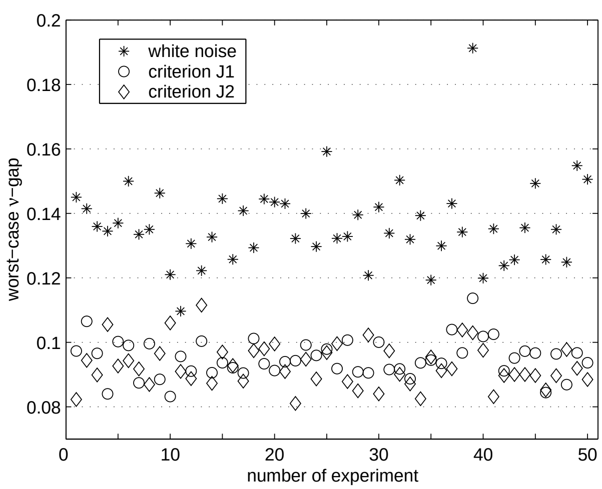

Fig. 1. Identification with white and subsequently estimated optimal input

In figure 1 the worst-case ν-gap obtained from the preliminary experiment with white noise input, as well as from the experiments with inputs optimized with respect to $\mathscr{J}_1$ and $\mathscr{J}_2$ respectively, are shown for the first 50 simulation runs. The mean over 500 runs of the worst-case ν-gap resulting from the preliminary experiments equals 0.1345. The means of the worst-case ν-gap resulting from the experiments with multisine input optimized with respect to criteria $\mathscr{J}_1, \mathscr{J}_2$ are 0.0937 and 0.0927, respectively. The difference between them is statistically significant (2×1.64 standard deviations). The means of the worst-case ν-gap resulting from the experiments with D- and E-optimal multisine input are equal to 0.1434 and 0.1055.

It is evident that using inputs optimized with respect to criteria $\mathscr{J}_1, \mathscr{J}_2$ gives better results than using white noise input or input optimized with respect to the classical D- and E-optimality criteria. Note also that the inputs optimized with respect to the cost function $\mathscr{J}_2$ give better results than $\mathscr{J}_1$, despite the fact that the plotted quantity is in fact $\mathscr{J}_1$. This tendency was observed also in simulations with other systems. As mentioned already in section 2, the reason is that the optimum of the input power spectrum with respect to $\mathscr{J}_2$ is less dependent on the preliminary estimate of the true parameter vector. Given the lower complexity of $\mathscr{J}_2$ and hence the lower computational effort in comparison with $\mathscr{J}_1$, it is recommendable to use primarily the former.

VIII. Conclusions

Let us summarize the results obtained in the present paper. We have to design an input sequence for an identification experiment that makes the worst-case ν-gap between the identified model and the uncertainty region around it as small as possible. The design takes place via power spectrum optimization. Two nonstandard cost criteria $\mathscr{J}_1$ and $\mathscr{J}_2$ are defined, which reflect the optimization task with different accuracy. $\mathscr{J}_1$ is the exact worst-case ν-gap one would want to minimize, while $\mathscr{J}_2$ is an approximation of $\mathscr{J}_1$. Both fulfil the natural conditions of monotonicity and quasiconvexity with respect to the power spectrum.

It was shown that optimization of the input power spectrum with respect to these cost criteria can be reduced to a convex optimization problem involving LMI constraints.

Simulations show clearly the superiority of the proposed cost functions over classical design criteria. They also suggest to use cost function $\mathscr{J}_2$ rather than $\mathscr{J}_1$, due to both lower computational effort and higher performance.

IX. References

[1] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15 of *SIAM Stud. Appl. Math.* SIAM, 1994.

[2] Michel Gevers, Xavier Bombois, Benoît Codrons, Gérard Scorletti, and Brian Anderson. Model validation for control and controller validation: a prediction error identification approach. In *Proceedings of the 12th IFAC Symposium on System identification (SYSID 2000)*, pages 319–324, Santa Barbara, June 2000. Pergamon, Elsevier Science.

[3] Roland Hildebrand and Michel 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):1586–1605, 2003.

[4] Samuel Karlin and William Studden. *Tchebycheff systems: with applications in analysis and statistics*, volume XV of *Pure Appl. Math.* Interscience Publishers, 1966.

[5] Jack Kiefer. General equivalence theory for optimum designs (approximate theory). *Ann. Statist.*, 2(5):849–879, 1974.

[6] K. Lindqvist and H.Hjalmarsson. Identification for control: Adaptive input design using convex optimization. In *Proceedings of the 40th Conference on Decision and Control*, Orlando, Florida, USA, 12 2001.

[7] Lennart Ljung. *System identification: theory for the user*. Prentice-Hall Information and System Sciences Series. Prentice Hall, second edition, 1999.

[8] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Stud. Appl. Math.* SIAM, Philadelphia, 1994.

[9] Robert Payne and Graham Goodwin. Simplification of frequency domain experiment design for SISO systems. Publication 74/3, Dept. of Computing and Control, Imperial College, London, 1974.

[10] Glenn Vinnicombe. Frequency domain uncertainty and the graph topology. *IEEE Trans. Automat. Control*, AC-38(9):1371–1383, 1993.

[11] S.P. Wu, Stephen P. Boyd, and L. Vandenberghe. FIR filter design via semidefinite programming and spectral factorization. In *Proceedings of the 35th CDC*, Kobe, Japan, 1996.

[12] Martin Zarrop. *Optimal experiment design for dynamic system identification*, volume 21 of *Lecture Notes in Control and Inform. Sci.* Springer, Berlin, New York, 1979.