

# OPTIMAL EXPERIMENT DESIGN FOR OPEN AND CLOSED-LOOP SYSTEM IDENTIFICATION\*

MICHEL GEVERS<sup>†</sup>, XAVIER BOMBOIS<sup>‡</sup>, ROLAND HILDEBRAND<sup>§</sup>, AND GABRIEL SOLARI<sup>¶</sup>

**This paper is dedicated to Brian Anderson, who has been such a tremendous source of inspiration for us during the four decades that are covered in this survey paper.**

**Abstract.** This article reviews the development of experiment design in the field of identification of dynamical systems, from the early work of the seventies on input design for open loop identification to the developments of the last decade that were spurred by the research on identification for control. While the early work focused entirely on criteria based on the asymptotic parameter covariance, the results of the last decade aim at minimizing a wide range of possible criteria, including measures of the estimated transfer function, or of functions of this estimated transfer function. Two important recent developments are the solution of the experiment design problem for closed loop identification, and the formulation and solution of the dual optimal design problem in which the cost of identification is minimized subject to a quality constraint on the estimated model. We shall conclude this survey with new results on the optimal closed loop experiment design problem, where the optimization is performed jointly with respect to the controller and the spectrum of the external excitation.

**Key words.** Prediction error identification, experiment design, identification for control

**1. Introduction.** Optimal experiment design has its roots in the statistical literature, where it was applied to the choice of regression variables in classical regression problems of the type  $y = f^T(x)\theta + v$ , where  $\theta$  is to be estimated from the observed variables  $x$  and  $y$ ,  $f(\cdot)$  is a known vector function,  $v$  is a random error, and the design consists of choosing the independent variables  $x$  from a compact set  $X$ . The most relevant contributions date back to the 1960's and are due to Kiefer [44], Kiefer and Wolfowitz [45], Karlin and Studden [43] and Fedorov [15]. The first contributions of optimal experiment design to dynamical system identification date back to around 1970, and the main contributors of this early period are Aoki and Staley [3], Mehra [58], Goodwin, Payne, Zarrop, Ng and Söderström [25, 27, 26, 60, 74]. An excellent survey of the early results in system identification can be found in [57].

The setup adopted in most of the literature on optimal experiment design for dynamical systems is that of a linear time-invariant discrete-time system identified using a parametric model structure that is able to represent the true system, and a Prediction Error (PE) identification criterion. For simplicity of exposition, we shall in this paper consider the single-input single-output (SISO) case only. Thus the “true system” can be represented by:

$$\mathcal{S} : y(t) = \overbrace{G(z, \theta_0)}^{G_0(z)} u(t) + \overbrace{H(z, \theta_0)e(t)}^{v(t)} \quad (1.1)$$

---

\*This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

<sup>†</sup>CESAME, Université Catholique de Louvain, Bâtiment EULER, 4 av. Georges Lemaitre, 1348 Louvain-la-Neuve, Belgium (Michel.Gevers@uclouvain.be).

<sup>‡</sup>Delft Center for Systems and Control, Delft University of Technology, The Netherlands. (x.j.a.bombois@tudelft.nl).

<sup>§</sup>LJK, Université Grenoble 1/CNRS, BP 53, 38041 Grenoble Cedex 9, France (roland.hildebrand@imag.fr).

<sup>¶</sup>Dalmine SpA, Piazza Caduti 6 Luglio 1944, 1, 24044 Dalmine, Italy (gsolari@dalmine.it).

for some unknown parameter vector  $\theta_0 \in \mathbf{R}^d$ , where  $e(t)$  is white noise of variance  $\lambda_0$ , while  $G(z, \theta_0)$  and  $H(z, \theta_0)$  are stable discrete-time transfer functions, with  $H(z, \theta_0)$  minimum-phase and monic, i.e.  $H(\infty, \theta_0) = 1$ . The system is identified with a model structure  $M(\theta) = [G(z, \theta), H(z, \theta)]$ ,  $\theta \in \mathbf{R}^d$ , which represents the true system exactly when  $\theta = \theta_0$ . The model set  $\mathcal{M}$  is defined as  $\mathcal{M} = \{M(\theta), \theta \in D_M \subset \mathbf{R}^d\}$ , where  $D_M$  is a specified subset.

When Prediction Error identification is used with a full order model structure, then under mild assumptions the estimated parameter vector  $\hat{\theta}_N$  is known to be asymptotically normally distributed and to converge as the number of collected data tends to infinity, such that:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{N \rightarrow \infty} N(0, P_\theta), \quad (1.2)$$

where the asymptotic parameter covariance matrix  $P_\theta$  can be estimated from the data. When the true system cannot be represented within the chosen model structure, the parameter estimate converges under reasonable assumptions to an asymptotic value  $\theta^*$  defined as:

$$\theta^* = \arg \min_{\theta \in D_M} \bar{V}(\theta), \quad \text{where } \bar{V}(\theta) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E[y(t) - \hat{y}(t|t-1, \theta)]^2, \quad (1.3)$$

where  $\hat{y}(t|t-1, \theta)$  is the one-step ahead predictor of  $y(t)$  computed from the model  $M(\theta)$ : see [52] for details. The error in the transfer function estimate at a given frequency  $\omega$  can then be decomposed as:

$$G(e^{j\omega}, \hat{\theta}_N) - G(e^{j\omega}, \theta_0) = \underbrace{G(e^{j\omega}, \hat{\theta}_N) - G(e^{j\omega}, \theta^*)}_{\text{variance error}} + \underbrace{G(e^{j\omega}, \theta^*) - G(e^{j\omega}, \theta_0)}_{\text{bias error}}. \quad (1.4)$$

The optimal design research of the seventies focused for the most part on the identification of systems operating in open loop, and the design variables were either the choice of input signal or of the sampling time. In this article we shall not discuss optimal sampling time design. The criteria for optimal input design were scalar functions of the parameter covariance matrix, or of the asymptotic parameter covariance matrix, and the optimization was performed either in the time domain or in the frequency domain. In both cases a statistical framework is assumed.

In the time domain approach, a Gaussian distribution is assumed for the noise and the likelihood function  $L(y, u, \theta)$ , evaluated at input and output data vectors  $u \in \mathbb{R}^N$  and  $y \in \mathbb{R}^N$ , is used to estimate the unknown parameter vector  $\theta$ . Assuming an unbiased estimator  $\hat{\theta}$  for  $\theta$ , the Cramér-Rao inequality states that the covariance matrix of  $\hat{\theta}$  is bounded from below by the inverse of the Fisher Information Matrix (FIM)  $M_\theta(u)$  constructed from  $L(y, u, \theta)$ . The input design problem is then to select a sequence  $u \in \mathbb{R}^N$  in some admissible compact set  $\mathcal{U}$  so as to maximize some scalar function of  $M_\theta(u)$ . The set  $\mathcal{U}$  is a subset of  $\mathbb{R}^N$  such as, e.g.  $\mathcal{U} = \{u \mid u^T u \leq \alpha\}$ . The problem thus reduces to a nonlinear optimal control problem; however, the computational burden increases dramatically as  $N$  gets large.

As a result, this time-domain approach has for the most part given way to a frequency domain approach that is based on a quasi-stationarity assumption of the input-output process and the replacement of the information matrix by the asymptotic per sample information matrix  $\bar{M}_\theta$ , which is the inverse of the asymptotic per sample covariance matrix  $P_\theta$ . Using Parseval theory, this asymptotic information matrix

is then expressed as a frequency weighted integral of the input spectrum, and the optimal design on the input signal  $u$  then reduces to an optimization with respect to the input spectrum  $\Phi_u(\omega)$  subject to a constraint on  $\Phi_u(\omega)$ , yielding an optimal spectrum  $\Phi_u^*(\omega)$ . Any input signal that realizes this optimal spectrum  $\Phi_u^*(\omega)$  is then the solution of the optimal input design problem. In Section 2 we shall briefly review the main results of the early work of the pioneers of experiment design in system identification. The main outcome of this period is that optimal inputs were obtained for a range of classical criteria, leading to convex optimization problems. The most commonly used design criteria were  $A$ -optimal design which minimizes  $tr(P_\theta)$ ,  $D$ -optimal design which minimizes  $det(P_\theta)$ ,  $E$ -optimal design which minimizes  $\lambda_{max}(P_\theta)$ , and  $L$ -optimal design which minimizes  $tr(WP_\theta)$ , where  $W$  is a nonnegative weighting matrix. It was shown that, under a constraint on the input power (i.e.  $\int \Phi_u(\omega) \leq c$ ), the optimal solutions can always be generated as multisines even though the set of optimal solutions is infinite [26]. The number of frequencies required is a function of the number of unknown parameters. One major difficulty of optimal experiment design problems is that the optimal solution depends on the unknown system. To circumvent this problem, it was suggested to use a Bayesian viewpoint, regarding the unknown parameter vector  $\theta$  as a random vector with a prior distribution.

Research on input design essentially came to a halt after the mid-seventies. But in the mid-eighties, Ljung and collaborators produced variance formulas [51, 53] directly for the transfer function estimates, rather than for the parameter estimates which only serve as auxiliary variables in the representation of these transfer functions. The asymptotic variance formulas were derived under the assumption that the model order  $n$  tends to infinity in some appropriate way when the data length  $N$  tends to infinity. Thus, for the variance of the transfer function estimates  $G(z, \hat{\theta}_N)$  and  $H(z, \hat{\theta}_N)$ , the following approximation was obtained in [51] under an assumption of high model order:

$$Cov \begin{pmatrix} G(e^{j\omega}, \hat{\theta}_N) \\ H(e^{j\omega}, \hat{\theta}_N) \end{pmatrix} \cong \frac{n}{N} \Phi_v(\omega) \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{eu}(\omega) & \lambda_0 \end{bmatrix}^{-1} \triangleq \Phi_{\chi_0}(\omega) \quad (1.5)$$

where  $n$  is the model order,  $N$  is the number of data,  $\Phi_u(\omega)$  is the input spectrum,  $\Phi_{ue}(\omega)$  is the cross-spectrum between  $u$  and  $e$ ,  $\Phi_v(\omega)$  is the output disturbance spectrum, and  $\lambda_0$  is the variance of the white noise signal  $e(t)$  in (1.1). Strictly speaking, the result assumes that the model order  $n$  tends to infinity, but experience has shown that this formula most often gives a reasonable approximation also for model orders that do not exceed that of the true system. The formula explicitly contains the effect of the experimental conditions (number of data, input spectrum, noise spectrum) on the error measure. For open loop identification this formula reduces to

$$Var(G(e^{j\omega}, \hat{\theta}_N)) \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \quad (1.6)$$

where  $n$  is the model order,  $N$  is the number of data,  $\Phi_u(\omega)$  is the input spectrum, and  $\Phi_v(\omega)$  is the output disturbance spectrum.

These asymptotic (in the model order  $n$  and in the number of data  $N$ ) formulas paved the way for the formulation and solution of a range of goal-oriented experiment design problems, including control-oriented problems [23, 35, 17, 67]. We shall review some of the most important results based on Ljung's asymptotic variance formulas in Section 3.

The early nineties saw the emergence of identification for control as an important new research subject. The focus was initially on reducing the effect of the bias error of the identified model (see (1.4)) on the controller derived from the estimated model. In practical applications the complexity of the model is always lower than that of the true system, and such bias error is therefore inevitable, and may lead to the computation of a controller that destabilizes the true system if no precautions are taken. The main outcomes that emerged from this research period can be summarized as follows:

- the data should be collected in closed loop with a controller that is as close as possible to the optimal controller that would result if the true system were known;
- the identification criterion should match the control performance criterion as much as possible;
- to achieve these goals a succession of closed-loop identification steps followed by control design steps should be performed.

These ideas led to the widespread development of iterative identification and control design schemes, and the use of the worst-case  $\nu$ -gap as a measure for checking closed-loop stability [19, 69, 28, 20, 73, 71].

The last decade saw a significant resurgence of activity in optimal experiment design. This revival was stirred by new developments in identification for robust control with the focus shifting from the influence of the bias error on the computed controller to the influence of the variance error of the estimated model on the set of stabilizing controllers. The progress was made possible by the advent of powerful techniques for convex optimization under Linear Matrix Inequality (LMI) constraints. This activity saw a return to the use of the more accurate finite order covariance formulas based on the parameter covariance matrix  $P_\theta$ , after it was observed that the asymptotic (in the model order) transfer function variance formulas could sometimes lead to erroneous conclusions [61].

One commonly used approach to go from parameter covariance to transfer function covariance is to use the following first order Taylor series approximation:

$$\text{Var}(G(e^{j\omega}, \hat{\theta}_N)) \approx \frac{1}{N} \frac{\partial G^*(e^{j\omega}, \theta_0)}{\partial \theta} P_\theta \frac{\partial G(e^{j\omega}, \theta_0)}{\partial \theta} \quad (1.7)$$

Optimal design problems based on this formula have been addressed in [49, 48, 39, 38, 37], where it is shown that several useful  $H_\infty$  design criteria can be reformulated as weighted trace optimal input design problems subject to LMI constraints.

A rather different approach to optimal input design for robust control has been developed for the case where the system is in the model set, in which case the error reduces to the variance error. This approach, which directly uses the covariance matrix  $P_\theta$  without the need for an approximation, is based on the use of the ellipsoidal uncertainty set  $U_\theta$ :

$$U_\theta = \{\theta | (\theta - \hat{\theta}_N)^T N P_\theta^{-1} (\theta - \hat{\theta}_N) < \chi^2\}. \quad (1.8)$$

It follows from the property (1.2) that the true parameter vector  $\theta_0 \in \mathbf{R}^d$  belongs to  $U_\theta$  with probability  $\alpha(d, \chi^2) = \text{Pr}(\chi^2(d) \leq \chi^2)$ , where  $\chi^2(d)$  denotes the  $\chi^2$  distribution with  $d$  degrees of freedom. The results in [8, 22], which connect robust stability and robust performance measures directly to the ellipsoidal uncertainty region  $U_\theta$ , allow one to formulate experiment design problems for robust control in terms of the minimization of some appropriate function of  $U_\theta$  (or of  $P_\theta$ ) without the need for a mapping to an uncertainty set in the frequency domain, which typically requires

both a Taylor series approximation as in (1.7) and/or a conservative step of over-bounding of the uncertainty set. The first open-loop optimal input design problem for robust control based on the direct use of the uncertainty ellipsoid  $U_\theta$  was formulated in [31], which provided an optimal input design that minimizes the worst-case  $\nu$ -gap  $\delta_{WC}(G(z, \hat{\theta}_N), \mathcal{D})$  between the identified model  $G(z, \hat{\theta}_N)$  and all models in the Prediction Error uncertainty set  $\mathcal{D} \triangleq \{G(z, \theta) | \theta \in U_\theta\}$ , with  $U_\theta$  defined by (1.8). This worst-case  $\nu$ -gap, defined in [71], is directly related to the size of the set of its stabilizing controllers: the smaller the worst-case  $\nu$ -gap of the uncertainty set  $\mathcal{D}$ , the larger is the set of controllers that stabilize all models in  $\mathcal{D}$ .

The research on experiment design of the last decade has enabled to vastly expand the realm of optimal design criteria and constraints, and the solutions proposed to these problems are all based on the finite order expressions for the covariance matrix  $P_\theta$  rather than on the approximate expressions derived under the assumption of a model order tending to infinity. The major advances can be summarized as follows.

- The design criteria are no longer limited to the classical  $D$ -,  $A$ -,  $E$ - or  $L$ -optimal design criteria, or to the variance of the transfer function estimates. They have been extended to an array of more complicated but also more practically interesting functions of the parameter or transfer function estimates [37, 39, 38, 55, 9].
- The constraints are no longer limited to energy constraints such as  $\int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \leq \alpha$ , but can also handle frequency by frequency power constraints such as  $\Phi_u(\omega) \leq \alpha(\omega) \forall \omega$ , or quality constraints such as  $Var[\hat{G}_N(e^{j\omega})] \leq \alpha(\omega) \forall \omega$  [40].
- Whereas the classical experiment design formulation considered the minimization of a quality measure of the estimated quantity of interest under constraints on the input signal, a dual approach has been introduced in which the cost of the identification is minimized subject to a quality constraint on the estimated quantity of interest [12, 10, 11, 39]. From a practical point of view, the cost of the identification is a major issue [63]; recent work has dealt with the understanding and formalization of the cost of an identification experiment and its relation to the model complexity [65, 34]. The duality of these two approaches has been established in [64].
- Experiment design has been extended from open loop design to closed loop design with a fixed controller (in which the design variable becomes the spectrum of the external excitation signal) [37, 11] and to the joint closed loop design, in which the design variables are the choice of feedback controller and/or the spectrum of the external excitation [36, 7]. In Section 7 we will provide an optimal solution to the latter problem.

Underpinning these results of the last decade were some significant technical developments that can essentially be summarized as follows. The first contribution was to show that many optimal design problems can be formulated or reformulated as convex optimization problems under LMI constraints. The second technical work consisted in showing that these optimization problems can be transformed into problems involving a finite dimensional parametrization of the design variables.

Because of the dependence of  $P_\theta$  on the true system (i.e.  $P_\theta = P_{\theta_0}$ ), another important feature of optimal experiment design is that the optimal experimental conditions are dependent on the to-be-identified system (1.1). This could appear to be a critical problem since an optimal design of the identification experiment requires knowledge of the system one attempts to identify. This problem was already acknowl-

edged in the early age of optimal experiment design (see e.g. [57]) and was generally solved by replacing the true system (1.1) by an initial estimate, i.e.  $P_\theta = P_{\theta_{init}}$ . However, in the last few years, more attention has been devoted to this issue. Two different approaches have been considered: robust experiment design [68, 11] and adaptive experiment design [35, 18]. In a nutshell, in robust experiment design, one considers a set of possible initial estimates instead of just one, and the optimal experimental conditions are the ones which give a satisfactory objective function for all possible values in that set. In adaptive experiment design, the initial estimate of  $\theta_0$  as well as the spectrum of the excitation signal is adapted throughout the identification experiment based on the information gathered since the beginning of this experiment.

In the remainder of this paper we shall first briefly review the major results derived from the seventies to the turn of the century. The bulk of the paper will be to present the techniques and results developed in the last decade in the context of experiment design for robust control. We shall conclude with some new results for the optimal design of closed loop experiments.

**2. The early work of the seventies.** We shall only refer in this section to the frequency domain approach. Whereas in the time-domain approach, a finite set of input data  $u(1), \dots, u(N)$  are computed as the solution of a nonlinear optimal control problem that maximizes some scalar measure of the information matrix, here the input signal is assumed quasi-stationary and the open loop optimal design problem then reduces to computing the optimal spectrum  $\Phi_u^*(\omega)$  that minimizes a scalar measure of the asymptotic per sample covariance matrix  $P_\theta$  or, equivalently, that maximizes a scalar measure of the average per sample information matrix  $\bar{M}_\theta$ . When the system is in the model set, the dependence of  $P_\theta$  on the input spectrum is given by the following expression, using Parseval's theorem:

$$\begin{aligned} \bar{M}_\theta = P_\theta^{-1} = & \left( \frac{1}{2\pi\lambda_0} \int_{-\pi}^{\pi} F_u(e^{j\omega}, \theta_0) F_u(e^{j\omega}, \theta_0)^* \Phi_u(\omega) d\omega \right) \\ & + \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} F_e(e^{j\omega}, \theta_0) F_e(e^{j\omega}, \theta_0)^* d\omega \right) \end{aligned} \quad (2.1)$$

Here,  $F_u(z, \theta_0) = \frac{\Lambda_G(z, \theta_0)}{H(z, \theta_0)}$ ,  $F_e(z, \theta_0) = \frac{\Lambda_H(z, \theta_0)}{H(z, \theta_0)}$ ,  $\Lambda_G(z, \theta) = \frac{\partial G(z, \theta)}{\partial \theta}$  and  $\Lambda_H(z, \theta) = \frac{\partial H(z, \theta)}{\partial \theta}$ . The formula shows that the input spectrum  $\Phi_u(\omega)$  appears linearly in the expression of the information matrix  $P_\theta^{-1}$ , and that, for a given data length  $N$ , the input spectrum is the only design quantity that can shape the parameter covariance matrix.

Most of the results of this early period were derived under the following simplifying assumptions:

- The length  $N$  of the data set is large;
- The input signal  $u(t)$  is quasistationary and admits a spectral density  $\Phi_u(\omega)$ .
- The constraint is on the allowable input power:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \leq \alpha. \quad (2.2)$$

The main outcomes of the research in the 1970's can be summarized as follows: see e.g. [26, 74] for the relevant theorems and proofs.

1. The set of average information matrices corresponding to a power constrained input is a convex set.

2. Let  $\Phi_u^{(1)}(\omega)$  be an input spectrum that satisfies the power constraint (2.2) and yields an average information matrix  $\bar{M}_\theta(\Phi_u^{(1)})$ . Then there exists an input spectrum  $\Phi_u^{(2)}(\omega)$  generated by  $\frac{1}{2}d(d+1) + 1$  sinusoids, where  $d$  is the dimension of  $\theta$ , that yields the same average information matrix:  $\bar{M}_\theta(\Phi_u^{(2)}) = \bar{M}_\theta(\Phi_u^{(1)})$ . This means in particular that an optimal input signal can always be generated by no more than  $\frac{1}{2}d(d+1) + 1$  sinusoids. Note that for some optimal design criteria an optimal solution can often be generated with a smaller number of sinusoids.
3. For a Box-Jenkins model structure

$$y(t) = \frac{b_1 z^{-1} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} u(t) + \frac{1 + c_1 z^{-1} + \dots + c_n z^{-n}}{1 + d_1 z^{-1} + \dots + d_n z^{-n}} e(t), \quad (2.3)$$

an optimal input signal minimizing  $\det \bar{M}_\theta$  under the input power constraint (2.2) can be constructed that comprises no more than  $2n$  sinusoids [26]. Observe that in such model structure the input signal yields information only about the parameters  $b_i$  and  $a_i$ , i.e.  $2n$  parameters in the structure (2.3). We notice in passing that the number  $2n$  of sinusoids required for an optimal solution is twice the minimum number of sinusoids required to make the information matrix nonsingular [21].

Some preliminary results were also obtained that compared open loop and closed loop experimental conditions [26, 60]. It was shown that when the design criterion is to minimize  $\det \bar{M}_\theta$ , open loop identification is optimal if the constraint (2.2) is imposed on the input power. For the simple ARX model  $y(t) = a_1 y(t-1) + \dots + a_n y(t-n) + bu(t-1) + e(t)$  with the same design criterion and a constraint on the output power, it was shown that closed loop identification with a minimum variance controller and a white noise external excitation signal is optimal.

**3. Optimal design based on Ljung's asymptotic variance formulas.** In the mid-eighties, Ljung [51] produced the asymptotic variance formulas (1.5) for the transfer function estimates under the assumptions that not only the number of data  $N$  but also the model order  $n$  tends to infinity. When identification is performed in open loop,  $\Phi_{ue}(\omega) = 0$  and the variance expression for  $G(e^{j\omega}, \hat{\theta}_N)$  reduces to the simplified expression (1.6). These formulas led to the formulation of a range of new optimal design criteria and the solution of the corresponding input design problems [23, 35, 17, 67]. The optimal design criteria used in these design problems are the variance of the error between an application dependent signal obtained from the estimated model and the corresponding signal that would be obtained if the true system were known. Thus, let  $s(t)$  be a signal derived from an application of the model; for example,  $s(t)$  could be the output of the closed loop system obtained using a model-based controller. If the true system (1.1) were known, this signal would be the optimal signal  $s_0(t) = f(G_0(z), H_0(z))w(t)$ , where  $w(t)$  is a driving signal in the application. Since the signal is based on an estimated model  $G(z, \hat{\theta}_N)$  based on  $N$  data, the achieved signal is a degraded version of  $s_0(t)$ , namely  $\hat{s}_N(t) = f(G(z, \hat{\theta}_N), H(z, \hat{\theta}_N))w(t)$ , yielding an error  $\Delta s_N(t) \triangleq \hat{s}_N(t) - s_0(t)$ . By expanding  $f(G(z, \hat{\theta}_N), H(z, \hat{\theta}_N))$  around the true system and linearizing, the variance of this degradation can be approximated by  $E[\Delta s_N(t)]^2 \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr}[Q(\omega)\Phi_{\chi_0}(\omega)]d\omega$ , where  $Q(\omega)$  is a weighting matrix that reflects the application and  $\Phi_{\chi_0}(\omega)$  is the spectrum defined in (1.5). An optimal experiment design criterion, in the framework proposed

in the mid-eighties and based on the formula (1.5), can be formulated as

$$\min_{\Phi_{\chi_0}} \int_{-\pi}^{\pi} \text{tr}[Q(\omega)\Phi_{\chi_0}(\omega)]d\omega \quad (3.1)$$

subject to  $\int_{-\pi}^{\pi} \Phi_u(\omega)d\omega \leq \alpha$ , or  $\int_{-\pi}^{\pi} \Phi_y(\omega)d\omega \leq \alpha$ .

A much studied problem, for example, is when the application dependent signal is the output of a closed loop system under minimum variance control: the error signal is then the difference between the output of the closed loop system under exact minimum variance control and the output of the system controlled by a minimum variance controller derived from the estimated model; the driving signal  $w(t)$  is then the external reference signal applied to the closed loop system. It was shown in [23] that, for an ARMAX model structure, the experimental conditions that minimize the variance of this error signal  $\Delta s_N(t)$  consist of identifying the system in closed loop, controlled by the ideal minimum variance controller. A remarkable side result showed that the accuracy of the minimum variance controller estimated from the identified model is independent of the power of the externally applied excitation signal, even though the model accuracy does improve by increasing the power of that signal. Experiment design for the minimum variance control application was further developed in [33, 56] using the novel techniques developed in the last decade, with results based on finite order covariance formulas.

**4. The nineties: the advent of closed loop identification.** The result of [23] on optimal experimental conditions for a minimum variance control application was the first instance in which closed loop identification was shown to be beneficial compared to open loop identification. The street consensus up to that point was that closed loop experimental conditions should be avoided by all means.

Around 1990, the new topic of *identification for control* emerged. The question addressed in this research can be summarized as follows. Given that the objective of the identification is to design a model-based controller, what are the experimental conditions and the identification criterion that should be used so that the degradation in control performance due to bias and variance error in the model are as small as possible? The early work focused on the effect of the identification error and criterion on the bias error and led to the finding that, when the identification is performed for the purpose of designing a controller, it should be performed in closed loop with a controller that should ideally match the to-be-identified controller. This quickly led to the idea of iterative identification and control design schemes, which flourished during the early nineties [19, 69, 6, 28, 47, 73].

In the second part of the nineties, attention turned to the design of identification experiments that minimize some measure of the controller variance or the expected value of the noise-induced control performance degradation, reaching the similar conclusion that closed loop experimental conditions are to be preferred if the identified model is to be used for control design [35, 17].

The system identification community realized during the nineties that there was a significant gap between the tools, methods and uncertainty sets developed by the robust control community and those delivered by the identification community, and major attempts were made to match these tools and uncertainty sets [19, 14, 54, 29, 24, 8].

The recognition that identification for control should be performed in closed loop led to a renewed interest in closed loop identification, and a number of new closed loop identification methods were developed whose purpose was to address and eliminate



the bias error induced on the estimated input-output model when the noise model is incorrect [30, 70, 16]. As for the iterative identification and control schemes, they led to the question as to whether the newly computed controller can safely replace the existing controller, i.e. can we guarantee closed loop stability, and possibly also closed loop performance? This led to the introduction of caution in iterative identification and control design [2].

**5. Optimal experiment design in the last decade.** The last decade has seen an intense new activity in experiment design for system identification, with a shift from the asymptotic variance formulas described in the previous section to the more accurate formulas based on finite order models and parametrizations. As explained in the introduction, the recent work has seen an expansion of the set of admissible criteria and constraints (and in particular the handling of frequency by frequency criteria and constraints), the solution of both open loop and closed loop experiment design problems (including the simultaneous optimization of controller and external excitation), the introduction of the concept of ‘identification cost’ and the solution to the dual problem of least costly identification, and some preliminary results on robust experiment design. Optimal experiment design can now handle criteria that are a function of the estimated parameters or of the estimated transfer functions, such as, e.g. the optimal design of inputs for the estimation of the zeros of a transfer function [55, 66]. These new results have been made possible by the important developments in convex optimization under LMI constraints over the last two decades [59].

**5.1. The main ingredients of current experiment design.** We shall assume that the identification data are collected on the ‘true system’ (1.1). During the data collection, this system may be operating under feedback control with an operating controller  $C_{id}(z)$ :

$$u(t) = r(t) - C_{id}(z)y(t) \quad (5.1)$$

where  $r(t)$  is an external excitation signal. The closed-loop system can be written as:

$$y(t) = G_0 S_{id} r(t) + S_{id} v(t) \quad (5.2)$$

$$u(t) = S_{id} r(t) - C_{id} S_{id} v(t), \quad \text{with } S_{id} \triangleq 1/(1 + C_{id} G_0). \quad (5.3)$$

Assuming quasistationary signals and a frequency domain approach, one can distinguish between three possible formulations of optimal experiment design problems, of increasing complexity.

- Open loop design:  $\min_{\Phi_u(\omega)} J$  subject to constraints
- Closed-loop design with fixed controller  $C_{id}$  :  $\min_{\Phi_r(\omega)} J$  subject to constraints
- Joint design of  $C_{id}(z)$  and  $\Phi_r(\omega)$ :  $\min_{C_{id}(z), \Phi_r(\omega)} J$  subject to constraints

The criterion  $J$  can take a range of possible expressions; examples are a quality criterion such as  $J = \text{Var}(C(\hat{G}_N))$  where  $C(\hat{G}_N)$  is a model-based controller, or an energy criterion such as  $J = \int_{-\pi}^{\pi} \Phi_r(\omega) d\omega$ . The constraints can also cover a wide range of possibilities such as  $\int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \leq \alpha$ , or  $\int_{-\pi}^{\pi} \Phi_y(\omega) d\omega \leq \alpha$ , or  $|T(e^{j\omega}) \frac{G_0(e^{j\omega}) - G(e^{j\omega}, \theta)}{G(e^{j\omega}, \theta)}| \leq \alpha \quad \forall \omega$  where  $T(e^{j\omega})$  is the complementary sensitivity function, i.e. a typical robust stability criterion.

The design variables  $C_{id}$  and  $\Phi_r$  in a closed loop design are related to the spectra  $\Phi_u$  and  $\Phi_{ue}$  defined earlier through the following one-to-one relationships:

$$\Phi_u(\omega) = \lambda_0 |(1 + C_{id}G_0)^{-1}C_{id}H_0|^2 + |1 + C_{id}G_0|^{-2}\Phi_r(\omega), \quad (5.4)$$

$$\Phi_{ue}(\omega) = -\lambda_0(1 + C_{id}G_0)^{-1}C_{id}H_0. \quad (5.5)$$

As we shall see, it will turn out to be much simpler to perform the optimization with respect to the design variables  $\Phi_u$  and  $\Phi_{ue}$ , because the optimization criterion can often be made affine in these variables. The optimal controller  $C_{id}^*$  and the optimal external excitation spectrum  $\Phi_r^*$  can then be reconstructed from the optimal  $\Phi_u^*$  and  $\Phi_{ue}^*$  by inverting the relationships above:

$$C_{id}^* = -\Phi_{ue}^*(\lambda_0 H_0 + G_0 \Phi_{ue}^*)^{-1} \quad (5.6)$$

$$\Phi_r^* = |1 + C_{id}^* G_0|^2 (\Phi_u^* - \lambda_0^{-1} |\Phi_{ue}^*|^2). \quad (5.7)$$

The stability of the closed loop system formed from  $C_{id}^* G_0$  puts some constraint on the optimal cross-spectrum  $\Phi_{ue}^*$  as we shall discuss in Section 7.2.

The connection between the design parameters and the average information matrix is obtained, using Parseval's theorem, by the following relationship:

$$\bar{M}_\theta \triangleq P_\theta^{-1} = \frac{1}{2\pi\lambda_0} \int_{-\pi}^{\pi} F(e^{j\omega}, \theta) \Phi_{\chi_0}(\omega) F^*(e^{j\omega}, \theta) d\omega \quad (5.8)$$

where

$$F(z, \theta) \triangleq \begin{bmatrix} H^{-1}(z, \theta) \frac{\partial G(z, \theta)}{\partial \theta} & H^{-1}(z, \theta) \frac{\partial H(z, \theta)}{\partial \theta} \end{bmatrix}, \quad (5.9)$$

$$\Phi_{\chi_0}(\omega) \triangleq \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{ue}^*(\omega) & \lambda_0 \end{bmatrix} \quad (5.10)$$

Observe that  $\bar{M}_\theta$  is affine in the design variables  $\Phi_u(\omega)$  and  $\Phi_{ue}(\omega)$ . The main strategy that leads to the new optimal design results of the last decade consist of the following ingredients:

- Express the design criterion as a linear function of the design variables.
- Express the constraints as an LMI on  $\bar{M}_\theta$ .
- Find a finite dimensional parametrization of the design variables.

The key technical tool for the realization of the first two ingredients is the Schur complement of a block-partitioned matrix. As for the obtention of a finite dimensional parametrization, there are essentially two approaches: one is based on approximating the spectrum  $\Phi_{\chi_0}(\omega)$  by a finite dimensional parametrization (e.g. by restricting it to be generated by a Finite Impulse Response (FIR) filter driven by white noise; the other is by using the so-called partial correlation approach, based on Tchebycheff moment theory introduced into experiment design theory by Zarrop [74]. In the next few subsections, we shall briefly explain or illustrate these key technical devices.

**5.2. Converting to an affine criterion.** An important contribution of the recent work on experiment design has been to show that many practically useful design criteria can be rewritten, after some manipulations involving the Schur complement of a partitioned matrix, as a convex criterion subject to LMI constraints [13, 49, 37]. We illustrate this procedure with a simple example.

Consider an open loop optimal experiment design problem in which the criterion is the largest eigenvalue of the covariance matrix  $P_\theta$ :

$$\min_{\Phi_u(\omega)} \lambda_{\max}(P_\theta)$$

This can be rewritten as

$$\min_{\Phi_u(\omega)} \gamma \quad \text{subject to } \lambda_{\max}(P_\theta) \leq \gamma$$

Using the Schur complement, this can be rewritten as:

$$\begin{aligned} & \min_{\Phi_u(\omega)} \gamma \quad \text{subject to} \\ & \begin{bmatrix} \gamma I & I \\ I & P_\theta^{-1} \end{bmatrix} \succeq 0 \end{aligned}$$

The important outcome of this technical trick is that the initial criterion, which was non convex as a function of the design parameters, has been transformed into a convex criterion, since  $P_\theta^{-1}$  is affine in  $\Phi_u$ , whereas  $P_\theta$  is not.

**5.3. Convex representation of the quality constraints.** Similar efforts were developed to show that a wide range of useful quality constraints can also be transformed into LMI's in which the design variables appear in an affine way. The following example, taken from [41], illustrates the tools that make this possible. In identification for robust control, an important quantity of interest is

$$\Delta(e^{j\omega}, \hat{\theta}_N) \triangleq T(e^{j\omega}, \hat{\theta}_N) \frac{G_0(e^{j\omega}) - G(\hat{\theta}_N)}{G(\hat{\theta}_N)} \quad (5.11)$$

where  $T(e^{j\omega}, \hat{\theta}_N)$  is the complementary sensitivity obtained from the estimated model  $G(\hat{\theta}_N)$ , since  $\|\Delta(\hat{\theta}_N)\|_\infty < 1$  is a robust stability condition. One way to approximate this constraint is to impose an upper bound  $b(\omega)$  on the variance of  $\Delta(\hat{\theta}_N)$  which, using a Taylor series expansion can be approximated by

$$\text{Var}(\Delta(e^{j\omega}, \hat{\theta}_N)) \approx \left| \frac{T}{G_0} \right|^2 \frac{1}{N} \frac{\partial G(e^{-j\omega}, \theta_0)}{\partial \theta} P_\theta \frac{\partial G(e^{j\omega}, \theta_0)}{\partial \theta} \quad (5.12)$$

Imposing a bound  $b(\omega)$  on  $\text{Var}(\Delta(e^{j\omega}, \hat{\theta}_N))$  is equivalent with imposing the following weighted trace constraint on  $P_\theta$ :

$$\text{tr}\{W(\omega)P_\theta\} \leq 1 \quad \forall \omega \quad (5.13)$$

where

$$W(\omega) \triangleq \frac{1}{Nb(\omega)} \left| \frac{T(e^{j\omega})}{G_0(e^{j\omega})} \right|^2 \frac{\partial G(e^{j\omega}, \theta_0)}{\partial \theta} \frac{\partial G(e^{-j\omega}, \theta_0)}{\partial \theta}$$

Since  $W(\omega)$  is Hermitian, it admits a factorization  $W(\omega) = V(e^{j\omega})V(e^{-j\omega})$  where  $V(e^{j\omega}) \in \mathbf{C}^{p \times d}$  with  $p \triangleq \sup_\omega \text{rank}(W(\omega))$ . The constraint  $\text{tr}\{W(\omega)P_\theta\} \leq 1 \quad \forall \omega$  is then equivalent with

$$\text{tr}\{Z\} \leq 1 \quad \text{and} \quad Z - V(e^{-j\omega})P_\theta V(e^{j\omega}) \succeq 0$$

Using the Schur complement again shows that the latter is equivalent with

$$\text{tr}\{Z\} \leq 1 \quad \text{and} \quad \Gamma(\omega) \triangleq \begin{bmatrix} Z & V(e^{-j\omega}) \\ V(e^{j\omega}) & P_\theta^{-1} \end{bmatrix} \succeq 0 \quad \forall \omega. \quad (5.14)$$

Thus, the frequency by frequency constraint  $|\Delta(e^{j\omega}, \hat{\theta}_N)| < 1 \quad \forall \omega$  has been converted into a constraint that is convex in  $P_\theta^{-1}$ .

One problem is that (5.14) represents an infinite set of constraints, since it needs to hold for every frequency. However, when the weighting matrix  $W(\omega)$  is a finite dimensional spectrum, then  $V(e^{j\omega})$  can be taken as a minimum phase stable rational factor of  $W(\omega)$  which can be represented by a controllable state-space realization  $\{A_V, B_V, C_V, D_V\}$ , i.e.  $V(z) = C_V(zI - A_V)^{-1}B_V + D_V$ . Using the Kalman Yakubovich Popov (KYP) Lemma [72, 1], the infinite set of constraints can then be replaced by a finite set. It can indeed be shown [40] that the constraint on  $\Gamma(\omega)$  in (5.14) holds for all  $\omega$  if and only if there exists a real symmetric matrix  $Q$  such that

$$\begin{bmatrix} Q - A^T Q A & -A^T Q B \\ -B^T Q A & -B^T Q B \end{bmatrix} + \begin{bmatrix} 0 & C^T \\ C & D + D^T \end{bmatrix} \succeq 0 \quad (5.15)$$

where  $\{A, B, C, D\}$  are constructed from  $\{A_V, B_V, C_V, D_V\}$ ,  $Z$  and  $P_\theta$  as follows:

$$A = A_V, \quad B = \begin{bmatrix} 0 & B_V \end{bmatrix}, \quad C = \begin{bmatrix} C_V \\ 0 \end{bmatrix}, \quad D + D^T = \begin{bmatrix} Z & D_V \\ D_V^T & P_\theta^{-1} \end{bmatrix}$$

We have illustrated how many experiment design criteria and constraints can be reformulated as convex criteria and constraints in which the design parameters appear linearly:  $\Phi_u(\omega)$  in the case of open loop design, or  $\Phi_{\chi_0}(\omega)$  in the case of closed loop design. We have also shown how the set of admissible design variables are equivalent to the satisfaction of a LMI, possibly involving auxiliary variables such as  $\gamma$  in the example of Section 5.2 or  $Z$  in the example of Section 5.3. However, these design variables still have infinite dimension. The next step is to show how these can be replaced by a finite set of design variables.

**6. Finite dimensional parametrization of criterion and constraints.** Two basic approaches to the choice of the design variables can be distinguished.

One is the *finite dimensional spectrum parametrization* (see e.g. [37, 50, 41]). Here the spectrum is developed into an infinite series of basis functions and the design variables are given by the coefficients of a truncated version of this infinite series. If the basis functions are rational, by the KYP lemma the positivity of the spectrum is expressed by an LMI in the truncated coefficient vector. Every coefficient vector satisfying this LMI corresponds to a valid spectrum, but not every spectrum corresponds to a coefficient vector, because all coefficients in the series after the cut-off are supposed to be zero. Therefore the feasible set of the LMI approximates the feasible set of the original problem from the inside, i.e., the former is contained in the latter. Speaking in the terminology of optimization, the finite dimensional spectrum parametrization leads to an inner semi-definite relaxation of the original input design problem. This has the advantage that the optimal solution of the relaxation yields a realizable input design. However, this approach considers only a finite-dimensional subspace cut out of the infinite-dimensional variety of possibly useful spectra. This results in a performance loss, as the optimisation procedure returns only suboptimal solutions. For simple systems this loss can be compensated by considering a long

truncated coefficient sequence, but in practice the gap is significant and optimisation over large enough numbers of coefficients is computationally prohibitive.

The other approach is the *partial correlation approach* (see e.g. [31, 37]), which is in some sense dual to the finite dimensional spectrum parametrization. Namely, an infinite sequence of *linear functionals* on the space of spectra is considered, and the design variables are the values of a finite number of these functionals on the spectrum in question. These values are called (generalized) *moments* [42, 46] of the spectrum. The linear functionals are chosen in such a way that both the constraints and the cost function depend only on a finite number of these moments. Geometrically, the optimisation is performed over a finite-dimensional *projection* of the infinite-dimensional cone of possible spectra, as opposed to a finite-dimensional *section* in the finite dimensional spectrum parametrization approach. Each point in the finite-dimensional truncated moment space thus still corresponds to an infinite set of spectra rather than a single spectrum, and the points of the finite-dimensional moment cone exhaust all possible spectra. Thus the partial correlation approach does not suffer from the performance loss characteristic of the finite dimensional spectrum parametrization approach.

**6.1. Finite dimensional spectrum parametrization.** The simplicity of this method is that it is the spectrum itself that is expressed as a finite set of real parameters. Since we have shown how a large set of criteria and constraints can be written as affine functions of the spectrum, the remainder of the procedure reduces to the solution of a convex optimization problem with LMI constraints in which this finite set of parameters, as well as other auxiliary variables, appear linearly. The drawback is that only a suboptimal solution is obtained since the finite truncation of the spectrum limits the search space of admissible solutions.

Consider that a closed loop optimization is performed with respect to the spectral matrix  $\Phi_{\chi_0}(\omega)$  defined above. This matrix can always be expressed as

$$\Phi_{\chi_0}(\omega) = \sum_{k=-\infty}^{\infty} \tilde{C}_k B_k(e^{j\omega}) \quad (6.1)$$

where  $B_k$  are scalar proper stable rational basis functions with  $B_{-k}(e^{j\omega}) = B_k(e^{-j\omega})$ ,  $\tilde{C}_k = \tilde{C}_{-k}^T$  are real matrix coefficients that must satisfy  $\Phi_{\chi_0}(\omega) \succeq 0 \forall \omega$ . A standard choice for the basis functions is  $B_k(e^{j\omega}) = e^{-jk\omega}$ . A finite dimensional parametrization is then obtained by replacing the spectrum  $\Phi_{\chi_0}(\omega)$  in (6.1) by the following approximation:

$$\Phi_{\chi_0}(\omega) = \Psi(e^{j\omega}) + \Psi^*(e^{j\omega}) \quad (6.2)$$

where  $\Psi(e^{j\omega})$  is a finite dimensional approximation of the positive real part of  $\Phi_{\chi_0}(\omega)$ , and  $\Psi^*(e^{j\omega}) \triangleq \Psi^T(e^{-j\omega})$ :

$$\Psi(e^{j\omega}) = \sum_{k=0}^{M-1} C_k B_k(e^{j\omega}) \quad (6.3)$$

where  $C_k = C_{-k}^T$  and  $M$  is chosen large enough. A constraint on the matrix coefficients  $C_k$  must then be imposed to ensure the positivity of  $\Phi_{\chi_0}(\omega)$  defined in (6.2). This can be achieved using the Positive Real Lemma, based on the KYP Lemma [72]. The procedure is to let  $\Psi(e^{j\omega})$  be parametrized via a controllable state space realization  $\{A, B, C, D\}$ , where the coefficients  $C_k$  appear linearly in  $C$  and  $D$  while  $A$  and  $B$

contain only zeros and ones: see e.g. [37] for details. The quantity  $\Phi_{\chi_0}(\omega)$  defined by (6.2) is then positive semidefinite for all  $\omega$  if and only if there exists a symmetric matrix  $Q$  such that condition (5.15) holds.

A special case of this finite dimensional parametrization method is when the standard basis functions  $B_k(e^{j\omega}) = e^{j\omega}$  are chosen, in which case the coefficients  $C_k$  are the autocorrelation coefficients of the output of a white noise driven Finite Impulse Response (FIR) filter. In this case, the constraint (5.15) is expressed directly as an LMI on  $Q$  and on these correlation coefficients [40].

The main drawback of the finite dimensional parametrization method are that it delivers a suboptimal solution. The accuracy grows with the number of coefficients  $M$  taken in the finite dimensional approximation, but a large  $M$  may lead to computational complexity problems. The main advantage of this method is that it can handle frequency by frequency constraints, which cannot be handled by the *partial correlation parametrization method* to be studied next.

**6.2. Partial correlation parametrization.** The key feature of the partial correlation parametrization approach is that the criterion and the constraints of the optimal experiment design problem can be expressed as a linear function and as an LMI, respectively, of a finite set of so-called *generalized moments* or *Tchebycheff moments*  $\{m_0, m_1, \dots, m_n\}$  of the design spectrum  $\Phi_{\chi_0}(\omega)$  and, possibly, an additional finite set of auxiliary variables  $\{x_1, x_2, \dots, x_N\}$ . These generalized moments are defined as the following  $2 \times 2$  matrices:

$$m_k = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{|d(e^{j\omega})|^2} \Phi_{\chi_0}(\omega) e^{jk\omega} d\omega, \quad k \in \mathbb{N}, \quad (6.4)$$

where  $d(z) = \sum_{l=0}^m d_l z^l$  is a polynomial of degree  $m$  with  $m \geq 0$  that is selected in such a way as to give to the criterion and the constraints the properties just described. The coefficients  $d_l$  of  $d(z)$  are real, they obey  $d_0 \neq 0$ ,  $d_m \neq 0$ , and  $d(z)$  has all roots outside of the closed unit disk. The matrices  $m_k$  defined by (6.4) are called the *generalized moments* of the spectrum  $\Phi_{\chi_0}$ . Note that the moments  $m_k$  are real, obey the relation  $m_k = m_{-k}^T$  and they are linear functions of  $\Phi_{\chi_0}(\omega)$ .

It is thus assumed that, by proper choice of  $d(z)$ , the experiment design criterion can be expressed as a linear function

$$J(m_0, \dots, m_n, x_1, \dots, x_N) = \sum_{k=0}^n \langle C_k, m_k \rangle + \sum_{l=1}^N c_l x_l, \quad (6.5)$$

where  $x_1, x_2, \dots, x_N$  are the auxiliary variables mentioned above,  $C_k$  are known matrices that depend on the particular problem,  $c_l$  are known reals that are also dependent on the particular problem, and  $\langle A, B \rangle = \text{trace}(AB^T)$ , and that the constraints can be expressed as an LMI

$$\mathcal{A}(m_0, m_1, \dots, m_n, x_1, \dots, x_N) \succeq 0 \quad (6.6)$$

in the same variables. The condition (6.5) is representative of a wide variety of problem formulations in open and closed-loop optimal experiment design; see e.g. [49, 37, 36]. In particular, all classical designs ( $D$ -optimal,  $A$ -optimal,  $L$ -optimal etc.) fall within this framework. As for condition (6.6) on the constraints, it is representative of all  $L_2$ -type constraints such as constraints on the possibly frequency weighted total input

or output power. However, as far as we know it cannot handle frequency by frequency constraints.

To illustrate the choice of the polynomial  $d(z)$ , consider an open loop design with the criterion  $\min_{\Phi_u} \det(\bar{M}_\theta)$  with an ARX model structure  $A(z)y(t) = B(z)u(t) + \epsilon(t)$ . The average information matrix  $\bar{M}_\theta$ , defined in (2.1), takes the form

$$\begin{aligned} \bar{M}_\theta &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Phi_u}{\lambda_0 |A|^2} \begin{pmatrix} -z^{-1}B(z) \\ \vdots \\ -z^{-n_a}B(z) \\ z^{-1}A(z) \\ \vdots \\ z^{-n_b}A(z) \end{pmatrix} \begin{pmatrix} -z^{-1}B(z)^* \\ \vdots \\ -z^{-n_a}B(z)^* \\ z^{-1}A(z)^* \\ \vdots \\ z^{-n_b}A(z)^* \end{pmatrix} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|A|^2} \begin{pmatrix} -z^{-1} \\ \vdots \\ -z^{-n_a} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} -z^{-1} \\ \vdots \\ -z^{-n_a} \\ 0 \\ \vdots \\ 0 \end{pmatrix}^* d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Phi_u}{\lambda_0 |A|^2} d\omega \tilde{M}_0 + \sum_{k=1}^n \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Phi_u}{\lambda_0 |A|^2} e^{jk\omega} d\omega \frac{\tilde{M}_k + \tilde{M}_k^T}{2} \right) + \tilde{M} \end{aligned}$$

It is easy to see that  $\bar{M}_\theta$  can be written as

$$\bar{M}_\theta = m_0 L_0 + \sum_{k=1}^n m_k L_k + L$$

with the moments defined by (6.4), provided the polynomial  $d(z)$  is chosen as  $d(z) = A(z^{-1})$ ; here  $n = n_a + n_b - 1$  and the  $L_k$  depend on the coefficients of  $A(z)$  and  $B(z)$ .

Now a typical experiment design problem, known as the *weighted trace design problem*, is

$$\min_{\Phi} \text{tr}\{P_\theta W\} \quad (6.7)$$

where  $\Phi$  is either  $\Phi_u(\omega)$  in open loop or  $\Phi_{\chi_0}(\omega)$  in closed loop,  $P_\theta$  is the asymptotic per sample covariance matrix, and  $W$  is a positive semi-definite (PSD) weighting matrix, typically the Hessian of the identification criterion. Such criterion has been studied e.g. in [49, 37]. Using the Schur complement, the optimal design problem (6.7) can be transformed into  $\min_{\Phi} \text{tr} X$ , subject to

$$\begin{bmatrix} X & W^{1/2} \\ W^{1/2} & \bar{M}_\theta \end{bmatrix} \succeq 0. \quad (6.8)$$

Thus, the criterion (6.7) can be re-expressed in the form of Assumption 6.5, where  $f_0$  takes the simple form  $f_0 = \sum_{k=1}^l x_{kk}$  where  $l$  is the size of  $X$ .

Since the criterion is now expressed as a function of the finite set of moments  $\{m_0, m_1, \dots, m_n\}$ , the optimization can be performed with respect to these moments, subject to the constraints imposed by the optimal design problem. To these constraints one must, however, add a realizability constraint given by the Carathéodory-Fejer theorem [43]. This theorem implies that the sequence  $\{m_0, m_1, \dots, m_n\}$  can be extended to an infinite sequence  $\{m_{n+1}, m_{n+2}, \dots\}$  that defines a positive power

spectrum if and only if the following LMI condition holds:

$$T_n = \begin{pmatrix} m_0 & m_1^T & \cdots & m_{n-1}^T & m_n^T \\ m_1 & m_0 & \cdots & m_{n-2}^T & m_{n-1}^T \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_n & m_{n-1} & \cdots & m_1 & m_0 \end{pmatrix} \succeq 0 \quad (6.9)$$

The partial correlation approach to optimal experiment design has been introduced by Zarrop [74] for the solution of open loop optimal design problems in which the optimization is with respect to the input spectrum  $\Phi_u(\omega)$  only. Many years later, in [31], this same approach allowed the solution of an open loop optimal design for a robust control problem, where the criterion is the worst-case  $\nu$ -gap  $\delta_{WC}(G(\hat{\theta}_N), \mathcal{D})$  between the estimated model and all models in an uncertainty set  $\mathcal{D}$ . The  $\nu$ -gap  $\delta_\nu(G_1, G_2)$  between two transfer functions  $G_1(z)$  and  $G_2(z)$  is a measure of distance between these two transfer functions, introduced by Vinnicombe [71], which allows one to evaluate the stability margin of a  $[G_2, C]$  loop as a function of the stability margin of the  $[G_1, C]$  loop and the distance  $\delta_\nu(G_1, G_2)$  between  $G_1$  and  $G_2$ . The worst-case  $\nu$ -gap  $\delta_{WC}(G(\hat{\theta}_N), \mathcal{D})$  is the supremum of the distance  $\delta_\nu(G(\hat{\theta}_N), G(\theta))$  between  $G(\hat{\theta}_N)$  and any member  $G(\theta) \in \mathcal{D}$ .

In both cases, constraint (6.9) imposed by the Carathéodory-Fejer theorem on the optimal set of moments  $\{m_0, m_1, \dots, m_n\}$  is all that is needed to guarantee the realizability of the solution. However, in the case of a closed loop design, the  $2 \times 2$  moments  $m_k$  are related to the joint spectrum  $\Phi_{\chi_0}(\omega)$  through (6.4) and this spectrum need not just be positive semi-definite. Indeed, there are constraints imposed on all but the (1, 1) elements of  $\Phi_{\chi_0}(\omega)$  (e.g. stability constraints on the (1, 2) and (2, 1) elements) and, therefore, the same holds for the moments  $m_k$ . This difficulty has prevented the application of the partial correlation method for closed loop experiment design, until a solution was obtained by Hildebrand, Gevers and Solari [32], which we shall briefly explain in Section 7.2. A suboptimal solution based on the finite dimensional spectrum parametrization approach was presented by Hjalmarsson and Jansson [36]: stability of the closed loop is enforced by the use of a Youla parameter which is approximated by a finite parametrization. In the next section we present these two solutions to the closed loop experiment design problem, where the optimization is with respect to both the controller and the external reference.

**7. Optimal joint closed-loop experiment design.** In this section we present some recent solutions to the closed loop optimal design problem where the optimization is performed with respect to both the controller  $C_{id}(z)$  that is present during the data collection, and the spectrum  $\Phi_r(\omega)$  of the external signal excitation. For reasons explained in Section 5.1, these design variables are replaced by the equivalent design variables  $\Phi_u(\omega)$  and  $\Phi_{ue}(\omega)$ . Thus the design problem treated in this section is

$$\min_{\Phi_u(\omega), \Phi_{ue}(\omega)} J \quad (7.1)$$

**7.1. Solution using the finite dimensional spectrum parametrization.**

The difficulty in the closed loop design, as explained in Section 6.2, is that the spectrum  $\Phi_{\chi_0}(\omega)$  defined in (5.10) is constrained, as shown by the fact that the (2, 2) element is a constant and that  $\Phi_{ue}(\omega)$  must be realizable as in (5.5), which imposes a



stability constraint. A solution to this problem using the finite dimensional spectrum parametrization approach has been proposed in [36] where a Youla parameter is used to parametrize the set of stabilizing controllers  $C_{id}(z)$ , with the property that the closed loop quantity  $\Phi_{ue}$  is affine in this Youla parameter. A finite parametrization is then chosen to approximate both  $\Phi_u(\omega)$  and this Youla parameter, and hence the whole spectrum  $\Phi_{\chi_0}(\omega)$ . The procedure can be briefly summarized as follows.

Let us denote  $G(z, \theta_0)$  and  $H(z, \theta_0)$  in (1.1) by  $G_0$  and  $H_0$  for simplicity, and let us assume that any unstable poles of  $H_0$  are also poles of  $G_0$ . Let us then factorize  $G_0$  as  $G_0 = \frac{N}{M}$ , where  $N$  and  $M$  are proper stable real rational transfer functions, i.e. they are in  $\mathcal{RH}_\infty$ . Then there exists a uniquely defined pair of transfer functions  $U, V \in \mathcal{RH}_\infty$  such that the Bezout identity  $UN + VM = 1$  holds. The set of all controllers  $C_{id}$  that stabilize  $G_0$  is then defined by  $C_{id} = \frac{U+QM}{V-QN}$ , where  $Q \in \mathcal{RH}_\infty$  is a free parameter called the Youla parameter. The set of all spectra  $\Phi_{\chi_0}(\omega)$  that can be generated from the system (1.1) with an internally stabilizing controller  $C_{id}(z)$  is given by

$$\Phi_{\chi_0} = \begin{pmatrix} \Phi_u & -\lambda_0 H_0 M (U + MQ) \\ -\lambda_0 H_0^* M^* (U + MQ)^* & \lambda_0 \end{pmatrix}$$

where  $\Phi_u(\omega)$  and  $Q$  are free variables subject to the constraints  $Q \in \mathcal{RH}_\infty$  and  $\Phi_{\chi_0}(\omega) \succeq 0$ .

The asymptotic average information matrix  $\bar{M}$ , related to  $\Phi_{\chi_0}(\omega)$  via (5.8), is now affine in the design variables  $\Phi_u(\omega)$  and  $Q$ ; however, these are infinite dimensional quantities. It then remains to select a finite dimensional parametrization for the spectrum  $\Phi_u(\omega)$  and the rational transfer function  $Q$ , using the procedures described in Section 6.1 and to perform the optimization with respect to this finite set of parameters, yielding an approximate solution to the closed loop experiment design problem.

**7.2. Solution using the partial correlation parametrization.** In Section 6.2 we have explained that the difficulty in the application of the partial correlation parametrization approach to the closed loop experiment design problem is that the Carathéodory-Fejer theorem, which provides necessary and sufficient conditions for the extension of a finite moment sequence  $\{m_0, \dots, m_n\}$  to an infinite sequence, has been proved only under the assumption that there are no constraints on these extended moments. In the case of closed loop identification, the definition of the  $\{m_k\}$  through (6.4) together with the expression (5.10) for  $\Phi_{\chi_0}(\omega)$  and (5.5) for  $\Phi_{ue}(\omega)$  implies that the (1, 2), (2, 1) and (2, 2) elements of all moments  $m_k$  must be constrained for them to be feasible, i.e. for them to be the moments of a spectrum  $\Phi_{\chi_0}(\omega)$  that results from the closed loop system defined by (1.1) and (5.1). We now detail these constraints.

Let  $\mathbb{T} \subset \mathbb{C}$  be the unit circle and let us define the function  $f_{ue} : \mathbb{T} \subset \mathbb{C}$  as  $f_{ue}(e^{j\omega}) \triangleq \Phi_{ue}(\omega)$ . The stability of the transfer function  $(1 + C_{id}G_0)^{-1}C_{id}H_0$  relating  $e$  to  $u$  in (5.5) requires that  $f_{ue}$  be holomorphic outside the unit disk. The defining relationship

$$m_{k,21} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{\Phi_{ue}(-\omega)}{d(e^{j\omega})d(e^{-j\omega})} e^{jk\omega} d\omega.$$

then implies

$$\sum_{l=0}^m d_l m_{k-l,21} = \frac{1}{2\pi j} \int_{\mathbb{T}} \frac{f_{ue}(z^{-1})}{d(z)} z^{k-1} dz \quad \forall k > 0.$$

Since all zeros of  $d(z)$  are outside the closed unit disc, the ratio  $f_{ue}(z^{-1})/d(z)$  is also holomorphic inside the unit disc, which implies the following constraint on the  $(2,1)$  elements of the  $\{m_k\}$ :

$$\sum_{l=0}^m d_l m_{k-l,21} = 0 \quad \forall k > 0. \quad (7.2)$$

In a similar manner it follows that

$$m_{k,22} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{\lambda_0}{|d(e^{j\omega})|^2} e^{jk\omega} d\omega, \quad \text{and hence} \quad \sum_{l=0}^m d_l m_{k-l,22} = 0 \quad \forall k > 0. \quad (7.3)$$

It thus follows that if  $\Phi_{\chi_0}(\omega)$  is a positive semidefinite spectrum defined by (5.4), (5.5) and (5.10), and if the  $\{m_k\}$  are the generalized moments of  $\Phi_{\chi_0}(\omega)$  defined by (6.4), then the constraints (7.2) and (7.3) must also hold. In addition the block-Toeplitz matrix  $T_n$  defined by (6.9) must be positive semidefinite since  $\Phi_{\chi_0}(\omega)$  is a spectrum.

We have shown in [32] that the converse holds as well. More precisely, if the  $2 \times 2$  matrices  $m_0, m_1, \dots, m_n$  satisfy  $m_0 = m_0^T$  and if they obey the constraints (7.2), (7.3) and (6.9) together with  $m_{-k} = m_k^T$  for all  $k = 1, \dots, n$ , then this finite sequence is extendable to an infinite sequence of moments  $\{m_{n'}\}$  for  $n' = n, n+1, \dots$ , which satisfy the the same constraints for all  $n'$ . The relation (6.4) then defines a spectrum  $\Phi_{\chi_0}(\omega)$ , via the M. Riesz extension theorem [62], that satisfies the constraints imposed by (5.4), (5.5) and (5.10). The proof of this result is too lengthy for this survey paper; it can be found in [32].

This result allows one to reformulate the optimal experiment design problem (6.5)-(6.6) into the following semi-definite program.

$$\min \left( \sum_{k=0}^n \langle C_k, m_k \rangle + \sum_{l=1}^N c_l x_l \right) \quad (7.4)$$

with respect to the set of constraints

$$\begin{aligned} \mathcal{A}(m_0, m_1, \dots, m_n, x_1, x_2, \dots, x_N) &\succeq 0, \\ m_{k,22} &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{\lambda_0}{|d(e^{j\omega})|^2} e^{jk\omega} d\omega, \quad k = -n, \dots, n, \\ \sum_{l=0}^m d_l m_{k-l,21} &= 0, \quad k = 1, \dots, n, \\ \begin{pmatrix} m_0 & m_1^T & \cdots & m_{n-1}^T & m_n^T \\ m_1 & m_0 & \cdots & m_{n-2}^T & m_{n-1}^T \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_n & m_{n-1} & \cdots & m_1 & m_0 \end{pmatrix} &\succeq 0, \end{aligned}$$

where  $m_{-k} = m_k^T$ . By solving this semi-definite program, the user obtains the optimal truncated moment sequence  $(m_0, \dots, m_n)$  and the optimal value of the cost function. In order to perform the identification experiment, the spectrum  $\Phi_r(\omega)$  and the controller  $C_{id}(z)$  must now be computed from the optimal moment matrices  $m_k$ . A solution to this recovery step, which is by no means a trivial problem, can also be found in [32].

**8. Least costly experiment design.** Traditionally, experiment design criteria have always been formulated as the minimization of a function of the covariance of the estimated quantities subject to constraints on the excitation signals, whether it be frequency by frequency constraints or total energy constraints. One of the new contributions of the last decade is to consider a dual to that problem: it consists of minimizing some measure of the *identification cost* subject to constraints on the achieved quality of the estimated quantity. The identification cost can be measured in many different ways, including the energy of the excitation signal, but it can also be the time required for the identification (i.e. the data length) or a measure of the performance degradation of the system induced by the application of the excitation signal required for the identification. In this section we present some of the key ideas and results of this dual approach to experiment design.

This dual approach to experiment design was first proposed in [12] in the context of identification for robust control; the proposed solution to the problem was based on Ljung's high model order variance formulas for estimated transfer functions presented in Section 3. A solution based on the finite order covariance formulas followed quickly thereafter [10, 11]. The experiment design problem solved in these papers is in the context of identification for robust control where the control objective is one of disturbance rejection. The data are collected in closed loop with a fixed controller  $C_{id}(z)$ . The optimization is thus performed with respect to the spectrum  $\Phi_r(\omega)$  of the external excitation only, and the design objective is to minimize the cost of the identification subject to the requirement that some model uncertainty measure is within the bounds set by some robust control specifications. The cost of the identification can be the data length  $N$ , or the performance degradation induced by the added excitation signal  $r$ , or a combination of both. We briefly explain the main ideas and concepts of this dual approach.

Consider that the 'true' system (1.1) is controlled by an existing controller  $C_{id}$  as in (5.1) which is not considered satisfactory. The objective is to identify a model  $G(e^{j\omega}, \hat{\theta}_N), H(e^{j\omega}, \hat{\theta}_N)$  of the plant in order to design a better (i.e. satisfactory) controller  $\hat{C}(z)$  that will achieve some robust control specification. In [11] a prediction error uncertainty set was defined for the set of transfer functions:

$$\mathcal{D}(\hat{\theta}_N, P_\theta) = \left\{ G(z, \theta) \mid (\theta - \hat{\theta}_N)^T P_\theta^{-1} (\theta - \hat{\theta}_N) < \chi \right\} \quad (8.1)$$

The constant  $\chi$  is chosen such that the true plant  $G(z, \theta_0)$  belongs to  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  with any desired probability level  $\beta$  (see [8]). The robust control design criterion adopted in [11] is of the form

$$J(G(z, \theta), C(G(z, \hat{\theta}_N)), W(z)) \leq \gamma < 1 \quad \forall G(z, \theta) \in \mathcal{D}(\hat{\theta}_N, P_\theta), \quad (8.2)$$

where  $\gamma$  is a fixed scalar,  $C(G(z, \hat{\theta}_N))$  is the controller defined from  $G(z, \hat{\theta}_N)$  through some model-based design method,  $W(z)$  is a frequency weighting filter, and  $J$  is an  $H_\infty$  measure, e.g.  $J(G, C, W) = \left\| \frac{W}{1+CG} \right\|_\infty$ . Many other criteria can of course be

adopted and have actually been adopted in the context of this dual experiment design method since then.

The experiment design variables are the data length  $N$  and the spectrum  $\Phi_r(\omega)$  of the external excitation, and the constraint on these design variables is that the quality criterion (8.2) must hold for all models in the uncertainty set  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  resulting from the identification. Observe that this is a real robust control constraint that directly relates the identification experiment with the quality required of the set of ‘estimated’ controllers resulting from the identification.

We now turn to the *identification cost*. In a disturbance rejection control problem, the input and output signals in normal operation are given by  $u(t) = -C_{id}S_{id}v(t)$  and  $y(t) = S_{id}v(t)$ , respectively, where  $S_{id} = \frac{1}{1+C_{id}G_0}$ : see (1.1) and (5.1). When an external signal  $r(t)$  is added for the purpose of closed loop identification, these signals are perturbed as follows:

$$y(t) = S_{id}v(t) + \overbrace{G_0 S_{id} r(t)}^{y_r(t)} \quad (8.3)$$

$$u(t) = -C_{id}S_{id}v(t) + \underbrace{S_{id}r(t)}_{u_r(t)} \quad (8.4)$$

This entails a performance degradation during the data collection. The power of this degradation could, e.g., be measured by

$$\begin{aligned} \mathcal{J}_r &= \alpha_y \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{y_r} d\omega \right) + \alpha_u \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_r} d\omega \right) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\alpha_y |G_0(e^{j\omega})S_{id}(e^{j\omega})|^2 + \alpha_u |S_{id}(e^{j\omega})|^2) \Phi_r(\omega) d\omega \end{aligned} \quad (8.5)$$

It is well known that in closed loop identification the covariance of the estimated parameters can be made arbitrarily small, even without external excitation, provided the controller is sufficiently complex, even though this may require very long data sets. This requires that the average per sample information matrix  $\bar{M}_\theta$  be nonsingular. Necessary and sufficient conditions for this to happen have been given in [21]. As a result, the performance degradation due to the external excitation can possibly be avoided if the cost due to the experiment time is negligible. Alternatively, if the cost of the experiment time is high, one can fix the experiment time and compute the experiment that minimizes the power  $\mathcal{J}_r$  of the degradation subject to satisfaction of a required level for the control performance cost. In [11] optimal experiment designs have been computed for these two scenarios as well as for a combination of them.

The concept of this dual approach to experiment design has quickly been adopted, and optimal design solutions have been proposed for a large number of constraints and criteria in this dual framework [37, 41, 40, 5, 4] and others. Finally, the direct and dual approach to optimal experiment design are of course closely related. The equivalence between these two approaches have been precisely established in [64].

**9. Conclusions.** We have reviewed the development of experiment design for system identification over the last four decades. The foundational tools of experiment design originate from the statistics community. Their introduction into the framework of the identification of dynamical systems was the work of a very limited number of pioneers in the seventies. For some reason, and with the exception of some results based on Ljung’s asymptotic formulas for the variance of estimated transfer functions,

the progress in experiment design essentially came to a halt for almost two decades. It was revived in the last decade, and has now become a flourishing activity, for essentially two reasons: the availability of a range of new tools in convex optimization under LMI constraints, and the development of identification for robust control as a major area of research and applications. Identification is now viewed as a goal-oriented design problem in which the experimental conditions under which the data are collected has become an essential part of the design. The recent work on experiment design is opening new research questions that will undoubtedly influence the research agenda in the future, such as the cost of the identification and its connection to the model complexity, as well as the connection between the application and the required model complexity.

#### REFERENCES

- [1] B. ANDERSON AND S. VONGPANITLERD, *Network Analysis and Synthesis?A Modern Systems Approach*, Prentice-Hall, Englewood Cliffs, New Jersey, 1973, 2006.
- [2] B. D. O. ANDERSON, X. BOMBOIS, M. GEVERS, AND C. KULCSÁR, *Caution in iterative modeling and control design*, in IFAC Workshop on Adaptive Control and Signal Processing, Glasgow, 1998, pp. 13–19.
- [3] M. AOKI AND R. STALEY, *On input signal synthesis in parameter identification*, Automatica, 6 (1970), pp. 431–440.
- [4] M. BARENTHIN, X. BOMBOIS, H. HJALMARSSON, AND G. SCORLETTI, *Identification for control of multivariable systems: controller validation and experiment design via LMIs*, Automatica, 44 (2008), pp. 3070–3078.
- [5] M. BARENTHIN, H. JANSSON, AND H. HJALMARSSON, *Applications of mixed  $H_2$  and  $H_\infty$  input design in identification*, in 16th IFAC World Congress on Automatic Control, paper 03882, Prague, July 2005. Paper Tu-A13-TO/1.
- [6] D. S. BAYARD, Y. YAM, AND E. METTLER, *A criterion for joint optimization of identification and robust control*, IEEE Trans. Automatic Control, 37 (1992), pp. 986–991.
- [7] A. BAZANELLA, M. GEVERS, AND L. MIŠKOVIĆ, *Closed-loop identification of MIMO systems: a new look at identifiability and experiment design*, European Journal of Control, 16 (2010), pp. 228–239.
- [8] X. BOMBOIS, M. GEVERS, G. SCORLETTI, AND B. D. O. ANDERSON, *Robustness analysis tools for an uncertainty set obtained by prediction error identification*, Automatica, 37 (2001), pp. 1629–1636.
- [9] X. BOMBOIS, H. HJALMARSSON, AND G. SCORLETTI, *Identification for robust  $H_2$  deconvolution filtering*, Automatica, 46 (2010), pp. 577–584.
- [10] X. BOMBOIS, G. SCORLETTI, M. GEVERS, R. HILDEBRAND, AND P. VAN DEN HOF, *Cheapest open-loop identification for control*, in CD-ROM Proc. 33rd IEEE Conf on Decision and Control, The Bahamas, December 2004, pp. 382–387.
- [11] X. BOMBOIS, G. SCORLETTI, M. GEVERS, P. VAN DEN HOF, AND R. HILDEBRAND, *Least costly identification experiment for control*, Automatica, 42 (2006), pp. 1651–1662.
- [12] X. BOMBOIS, G. SCORLETTI, P. VAN DEN HOF, AND M. GEVERS, *Least costly identification experiment for control: a solution based on a high-order model approximation*, in CD-ROM Proc. American Control Conference, Boston, MA, USA, July 2004, pp. 2818–2823.
- [13] B. COOLEY, J. LEE, AND S. BOYD, *Control-relevant experiment design: a plant-friendly LMI-based approach*, in Proc. American Control Conference, Philadelphia, Pennsylvania, USA, 1998.
- [14] D. K. DE VRIES AND P. M. J. VAN DEN HOF, *Quantification of uncertainty in transfer function estimation: a mixed probabilistic - worst-case approach*, Automatica, 31 (1995), pp. 543–558.
- [15] V. FEDOROV, *Theory of Optimal Experiments (Russian translation)*, Academic Press, New York, 1972.
- [16] U. FORSSELL AND L. LJUNG, *Closed-loop identification revisited*, Automatica, 35 (1999), pp. 1215–1241.
- [17] ———, *Some results on optimal experiment design*, Automatica, 36 (2000), pp. 749–756.
- [18] L. GERENCSÉR, H. HJALMARSSON, AND J. MARTENSSON, *Identification of ARX systems with non-stationary inputs ? Asymptotic analysis with application to adaptive input design*, Automatica, 45 (2009), pp. 623–633.

- [19] M. GEVERS, *Connecting identification and robust control: A new challenge*, in Proc. IFAC/IFORS Symposium on Identification and System Parameter Estimation, Budapest, Hungary, 1991, pp. 1–10.
- [20] ———, *Towards a joint design of identification and control?*, in Essays on Control: Perspectives in the Theory and its Applications, Birkhauser, New York, H. Trentelman and J. Willems, eds., New York, 1993, Birkhauser, pp. 111–151.
- [21] M. GEVERS, A. BAZANELLA, X. BOMBOIS, AND L. MIŠKOVIĆ, *Identification and the information matrix: how to get just sufficiently rich?*, IEEE Transactions on Automatic Control, 54 (2009), pp. 2828–2840.
- [22] M. GEVERS, X. BOMBOIS, B. CODRONS, G. SCORLETTI, AND B. D. O. ANDERSON, *Model validation for control and controller validation in a prediction error identification framework - Part I: theory*, Automatica, 39 (2003), pp. 403–415.
- [23] M. GEVERS AND L. LJUNG, *Optimal experiment designs with respect to the intended model application*, Automatica, 22 (1986), pp. 543–554.
- [24] L. GIARRÉ AND M. MILANESE, *Model quality evaluation in  $H_2$  identification*, IEEE Trans. Automatic Control, 42 (1997), pp. 691–698.
- [25] G. C. GOODWIN, *Optimal input signals for nonlinear system identification*, Proc. Inst. Elec. Eng., 118 (1971), pp. 922–926.
- [26] G. C. GOODWIN AND R. L. PAYNE, *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York, 1977.
- [27] G. C. GOODWIN, M. ZARROP, AND R. PAYNE, *Coupled design of test signals, sampling intervals and filters for system identification*, IEEE Trans. Automatic Control, 19 (1974), pp. 748–752.
- [28] R. G. HAKVOORT, R. J. P. SCHRAMA, AND P. M. J. VAN DEN HOF, *Approximate identification with closed-loop performance criterion and application to LQG feedback design*, Automatica, 30 (1994), pp. 679–690.
- [29] R. G. HAKVOORT AND P. M. J. VAN DEN HOF, *Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors*, IEEE Trans. Automatic Control, 42 (1997), pp. 1516–1528.
- [30] F. HANSEN, G. FRANKLIN, AND R. L. KOSUT, *Closed-loop identification via the fractional representation: Experiment design*, Proc. American Control Conference, (1989), pp. 1422–1427.
- [31] R. HILDEBRAND AND M. GEVERS, *Identification for control: optimal input design with respect to a worst-case  $\nu$ -gap cost function*, SIAM Journal on Control and Optimization, 41 (2003), pp. 1586–1608.
- [32] R. HILDEBRAND, M. GEVERS, AND G. SOLARI, *Closed-loop optimal experiment design: the partial correlation approach*, in 49th IEEE Conf. on Decision and Control, accepted, Atlanta, Georgia, USA, 2010.
- [33] R. HILDEBRAND AND G. SOLARI, *Identification for control: optimal input intended to identify a minimum variance controller*, Automatica, 43 (2007), pp. 758–767.
- [34] H. HJALMARSSON, *System identification of complex and structured systems*, European Journal of Control, 15 (2009), pp. 275–310.
- [35] H. HJALMARSSON, M. GEVERS, AND F. DE BRUYNE, *For model-based control design, closed-loop identification gives better performance*, Automatica, 32 (1996), pp. 1659–1673.
- [36] H. HJALMARSSON AND H. JANSSON, *Closed loop experiment design for linear time invariant dynamical systems via LMIs*, Automatica, 44 (2008), pp. 623–636.
- [37] H. JANSSON, *Experiment design with applications in identification for control*, PhD thesis, Royal Institute of Technology (KTH), December 2004. TRITA-S3-REG-0404.
- [38] H. JANSSON AND H. HJALMARSSON, *Convex computation of worst-case criteria with applications in identification and control*, in IEEE Conference on Decision and Control, The Bahamas, December 2004, pp. 3132–3137.
- [39] ———, *A general framework for mixed  $H_\infty$  and  $H_2$  input design*, submitted for publication to IEEE Trans. Auto. Control, (2004).
- [40] H. JANSSON AND H. HJALMARSSON, *Input design via LMIs admitting frequency-wise model specifications in confidence regions*, IEEE Transactions on Automatic Control, 50 (2005), pp. 1534–1549.
- [41] H. JANSSON AND H. HJALMARSSON, *Optimal experiment design in closed loop*, in 16th IFAC World Congress on Automatic Control, paper 04528, July 2005.
- [42] S. KARLIN AND L. SHAPLEY, *Geometry of moment spaces*, vol. 12 of Memoirs of the AMS, American Mathematical Society, 1953.
- [43] S. KARLIN AND W. STUDDEN, *Tchebycheff systems with applications in analysis and statistics*, Pure Appl. Math. 15, Wiley-Interscience, New York, 1966.

- [44] J. KIEFER, *Optimum experimental designs*, Journal of the Royal Statistical Society, ser. B, 21 (1959), pp. 273–319.
- [45] J. KIEFER AND J. WOLFOWITZ, *Optimum designs in regression problems*, Ann. Math. Statistics., 30 (1959), pp. 271–294.
- [46] M. G. KREIN, *The ideas of P. L. Chebyshev and A. A. Markov in the theory of limiting values of integrals and their further development*, Uspechi Mat. Nauk, 6 (1951), pp. 3–120.
- [47] W. S. LEE, B. D. O. ANDERSON, I. M. Y. MAREELS, AND R. L. KOSUT, *On some key issues in the windsurfer approach to adaptive robust control*, Automatica, 31 (1995), pp. 1619–1636.
- [48] K. LINDQVIST, *On experiment design in identification of smooth linear systems*, PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2001.
- [49] K. LINDQVIST AND H. HJALMARSSON, *Optimal input design using Linear Matrix Inequalities*, in CD-ROM Proc. of IFAC Symp. on System Identification, Santa Barbara, CA, USA, 2000.
- [50] ———, *Identification for control: adaptive input design using convex optimization*, in CD-ROM Proc. of 40th IEEE Conf. on Decision and Control, Orlando, Florida, 2001, pp. 4326–4331.
- [51] L. LJUNG, *Asymptotic variance expressions for identified black-box transfer function models*, IEEE Trans. Automatic Control, AC-30 (1985), pp. 834–844.
- [52] ———, *System Identification: Theory for the User, 2nd Edition*, Prentice-Hall, Englewood Cliffs, NJ, 1999.
- [53] L. LJUNG AND Z. YUAN, *Asymptotic properties of black-box identification of transfer functions*, IEEE Trans. Automatic Control, AC-30 (1985), pp. 514–530.
- [54] P. MÄKILÄ, J. PARTINGTON, AND T. GUSTAFSSON, *Worst-case control-relevant identification*, Automatica, 31 (1995), pp. 1799–1819.
- [55] J. MÅRTENSSON, H. JANSSON, AND H. HJALMARSSON, *Input design for identification of zeros*, in 16th IFAC World Congress on Automatic Control, July 2005.
- [56] J. MÅRTENSSON, C. R. ROJAS, AND H. HJALMARSSON, *Finite model order optimal input design for minimum variance control*, in Proc. European Control Conference, Budapest, August 2009.
- [57] R. K. MEHRA, *Optimal input signals for parameter estimation in dynamic systems - survey and new results*, IEEE Trans. on Automatic Control, AC-19 (1974), pp. 753–768.
- [58] ———, *Optimal inputs for linear system identification*, IEEE Trans. on Automatic Control, AC-19 (1974), pp. 192–200.
- [59] Y. NESTEROV AND A. NEMIROVSKII, *Interior Point polynomial algorithms in convex programming*, vol. 13, SIAM Studies in Applied Mathematics, Philadelphia, 1994.
- [60] T. S. NG, G. C. GOODWIN, AND T. SÖDERSTRÖM, *Optimal experiment design for linear systems with input-output constraints*, Automatica, 13 (1977), pp. 571–577.
- [61] B. NINNESS AND H. HJALMARSSON, *Variance error quantifications that are exact for finite model order*, IEEE Transactions on Automatic Control, 49 (2004), pp. 1275–1291.
- [62] M. RIESZ, *Sur le problème des moments: troisième note*, Ark. Math. Fys., 16 (1923), pp. 1–52.
- [63] D. E. RIVERA, H. LEE, M. BRAUN, AND H. MITTELMANN, *“plant-friendly” system identification: a challenge for the process industries*, in Proc. 13th IFAC Symp. on System Identification, Rotterdam, The Netherlands, 2003, pp. 917–922.
- [64] C. ROJAS, J. AGÜERO, J. WELSH, AND G. GOODWIN, *On the equivalence of least costly and traditional experiment design for control*, Automatica, 44 (2008), pp. 2706–2715.
- [65] C. ROJAS, M. BARENTIN, J. WELSH, AND H. HJALMARSSON, *The cost of complexity in identification of FIR systems*, in CD-ROM Proc. 17th IFAC World Congress, Seoul, Korea, 2008, pp. 11451–11456.
- [66] C. ROJAS, H. HJALMARSSON, L. GERENCSEK, AND J. M. RTENSSON, *An adaptive method for consistent estimation of real-valued non-minimum phase zeros in stable lti systems*, Automatica, to appear, 47 (2011).
- [67] C. ROJAS, H. HJALMARSSON, AND R. HILDEBRAND, *MIMO experiment design based on asymptotic model order theory*, in CD-ROM Proc. 48th IEEE Conf. on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R. China, 2009, pp. 488–493.
- [68] C. ROJAS, J. WELSH, G. GOODWIN, AND A. FEUER, *Robust optimal experiment design for system identification*, Automatica, 43 (2007), pp. 993–1008.
- [69] R. J. P. SCHRAMA, *Accurate identification for control: The necessity of an iterative scheme*, IEEE Trans. on Automatic Control, 37 (1992), pp. 991–994.
- [70] P. M. J. VAN DEN HOF AND R. J. P. SCHRAMA, *An indirect method for transfer function estimation from closed loop data*, Automatica, 29 (1993), pp. 1523–1527.
- [71] G. VINNICOMBE, *Frequency domain uncertainty and the graph topology*, IEEE Trans Automatic Control, AC-38 (1993), pp. 1371–1383.
- [72] V. A. YAKUBOVICH, *Solution of certain matrix inequalities occurring in the theory of automatic control*, Docl. Acad. Nauk. SSSR, (1962), pp. 1304–1307.

- [73] Z. ZANG, R. R. BITMEAD, AND M. GEVERS, *Iterative weighted least-squares identification and weighted LQG control design*, *Automatica*, 31 (1995), pp. 1577–1594.
- [74] M. ZARROP, *Optimal experiment design for dynamic system identification*, *Lecture Notes in Control and Information Sciences*, Vol. 21, Springer Verlag, Berlin, New York, 1979.