Input Design for Asymptotic Robust $\mathcal{H}_2$-Filtering

Cristian R. Rojas and Håkan Hjalmarsson

***Abstract*— In this paper we study the problem of robust discrete-time $\mathcal{H}_2$ filtering using a Linear Matrix Inequality approach. By assuming that the number of samples available for the identification of the system is large enough, we describe the filter design problem as a semidefinite program. Afterwards, the problem of designing an input signal for the identification of the system, to improve the performance of the conceived filter, is examined, and we show how to solve this problem using convex optimization.**

I. INTRODUCTION

The problem of linear filtering, or in general, linear estimation, is ubiquitous in engineering, and it has been studied since the birth of signal processing, with pioneering contributions by N. Wiener [28], A. N. Kolmogorov [17] and R. E. Kalman [16].

A main drawback of the standard filtering approaches is that they rely on perfect knowledge of the stochastic properties of the signals involved. This assumption does not hold in general, since these properties have typically to be estimated from data. Several kinds of approaches have been proposed in the literature, including the use of $\mathcal{H}_\infty$, minimax $\mathcal{H}_2$ and averaged $\mathcal{H}_2$ criteria; for a nice survey of the area, the reader is referred to [22].

Nowadays the literature on robust filtering is extensive. During the last decade several approaches have been proposed based on tools from the robust control community, such as Linear Fractional Transformations (LTFs), Integral Quadratic Constraints (IQCs), parameter-dependent Lyapunov functions (PDLFs), dynamic multipliers, and so on [31, 11, 24, 25]. As in the robust control problem, all these approaches require an estimation of the uncertainty region where the parameters of the signal properties lie. One idea to solve this issue, proposed by Bombois [3], consists in using as uncertainty regions the confidence ellipsoids delivered by the asymptotic (in sample size) Prediction Error Method (PEM) theory [18]. This allows to connect the results independently developed in the areas of system identification and robust filtering/control. As a consequence of this link, we see that it is possible to improve the performance of the robust filters by taking advantage of the degrees of freedom available during the identification step, such as the choice of the input signal [18].

This paper studies the problem of designing an input signal to optimize the performance of a robust $\mathcal{H}_2$ filter. By considering that the number of samples available for estimation is large enough, we derive expressions for the sensitivity of the performance of the filter with respect to first order perturbations of the parameters of the signal models. Then, we combine these expressions with the asymptotic covariance of a PEM estimator to develop a convex formulation of the input design problem using Linear Matrix Inequalities (LMIs) [5], which can be efficiently solved using standard interior point solvers [6].

It should be mentioned that this problem has already been studied in [4], where LFTs and dynamic multipliers are used to attack the problem. Unlike the approach followed in that paper, here we consider a Taylor expansion of the cost function, thus arriving at a simpler convex formulation, which can be globally solved with standard techniques.

To simplify the developments, we assume that during the input design step we have perfect knowledge of the plant. This assumption is clearly unrealistic, but it can be overcome by overimposing an extra robustness layer with respect to the lack of prior knowledge [27, 19]. Another approach is to do a sequential or adaptive design, where one redesigns the input signal as more data becomes available, see [7, 9, 10].

The structure of the paper is as follows: In Section II some preliminaries are given, and the problem is stated in three stages: the determination of the worst case performance of a filter, studied in Section III; the design of a robust filter for a given ellipsoidal uncertainty region, considered in Section IV; and the design of an optimal input for the construction of a robust filter, carried out in Section V. In Section VI we present an illustrative example of the techniques developed here, and Section VII concludes the paper.

The notation used in the paper is standard. $\mathbb{Z}$ denotes the set of integers. $\mathbb{R}^n$ ($\mathbb{C}^n$) is the space of real (complex) vectors of dimension n. If $x \in \mathbb{C}^n$, then x^T, x^* and x^H denote the transpose, complex conjugate and complex conjugate transpose of x, respectively. If P is a square matrix, $P > 0$ ($P \geq 0$) means that P is positive (semi-)definite.

II. PRELIMINARIES

Let us consider the system given by

$$y_t = G_o(q)u_t + H_o(q)e_t, \quad t \in \mathbb{Z},$$

where G_o and H_o are stable transfer functions, q is the forward shift operator, $\{e_t\}$ is white noise of zero mean and unit variance, and $\{u_t\}$ is a wide-sense stationary process, independent of $\{e_t\}$, of zero mean and spectrum Φ_u.

Let $\hat{G}_\theta, \hat{H}_\theta$ be a model of the system, where θ lies in the ellipsoid $U := \{\theta \in \mathbb{R}^p : (\theta - \theta_o)^T P^{-1}(\theta - \theta_o) < 1\}$, with $P = P^T > 0$, where θ_o is such that $\hat{G}_{\theta_o} = G_o$ and $\hat{H}_{\theta_o} = H_o$ (i.e. we assume that there is no undermodelling). This ellipsoid is usually provided by an identification experiment with large data samples, so it will be assumed that P is "small" in some sense. This allows us to consider the following Taylor approximation of $\hat{G}_\theta, \hat{H}_\theta$:

$$\begin{aligned}\hat{G}_\theta &\approx G_o + G'^T \Delta\theta \\ \hat{H}_\theta &\approx H_o + H'^T \Delta\theta, \quad \Delta\theta := \theta - \theta_o,\end{aligned} \tag{1}$$

where G' and H' denote the gradients of $\hat{G}_\theta$ and $\hat{H}_\theta$ with respect to θ at $\theta = \theta_o$, respectively.[1]

The purpose of filtering is to design a transfer function F (based on the knowledge of P and an estimate of θ, say $\hat{\theta}$) such that $\hat{u}_t := F(q)y_t$ is as close as possible to u_t, where "closeness" is defined in terms of the following cost function:

$$\begin{aligned}
&J_\theta(F) \\
&:= E\{(u_t - \hat{u}_t)^2\} \\
&= E\{[u_t - F(q)\hat{G}_\theta(q)u_t - F(q)\hat{H}_\theta(q)e_t]^2\} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\left|1 - F(e^{j\omega})\hat{G}_\theta(e^{j\omega})\right|^2 \Phi_u(e^{j\omega}) \right. \\
&\qquad\qquad \left. + \left|F(e^{j\omega})\hat{H}_\theta(e^{j\omega})\right|^2\right] d\omega \qquad (2) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\left(1 - F(e^{j\omega})\hat{G}_\theta(e^{j\omega}) - F^*(e^{j\omega})\hat{G}^*_\theta(e^{j\omega}) \right.\right. \\
&\qquad\qquad \left. + \left|F(e^{j\omega})\right|^2 \left|\hat{G}_\theta(e^{j\omega})\right|^2\right) \Phi_u(e^{j\omega}) \\
&\qquad\qquad \left. + \left|F(e^{j\omega})\right|^2 \left|\hat{H}_\theta(e^{j\omega})\right|^2\right] d\omega
\end{aligned}$$

where the third line comes from Parseval's relation. Here $E\{\cdot\}$ denotes expectation with respect to $\{u_t\}$ and $\{e_t\}$. To simplify the expressions, in the sequel we will omit the arguments of the functions and the limits of integration.

Since the input design problem involves quantities which cannot be determined at the time when the input is to be designed, such as the actual robust filter F, the complete problem of determining the best $\hat{u}$ is a stochastic multi-stage decision problem, which must be solved backwards (i.e. in a dynamic programming fashion). This implies solving the following 3 sub-problems, which are interesting in their own right:

1) **Filter Performance Verification:**
   Given a filter F and a positive scalar γ, we want to find if $J_\theta(F) < \gamma$ for all $\theta \in U$. In many cases, we actually need to know the minimum value of γ for which this holds, i.e., the worst case performance of the filter F.
2) **Robust $\mathcal{H}_2$-Filter Design:**
   Given an ellipsoidal uncertainty region U, we want to find a filter F such that $J_\theta(F) < \gamma$ for all $\theta \in U$ for the minimum possible value of γ.
3) **Input Design for Robust $\mathcal{H}_2$ Filtering**
   We want to design an input signal, described by its spectrum Φ_u^{id} and subject to a power constraint, to be used in an experiment to determine a model $\hat{G}_\theta, \hat{H}_\theta$ of the system with an uncertainty region U such that there is a filter F for which $J_\theta(F) < \gamma$ for all $\theta \in U$ for the minimum possible value of γ.

The first two problems are standard, and several solutions have been proposed in the literature (c.f. the references given in the introduction), even though very few consider ellipsoidal parametric uncertainty regions [3]. However, the last one is relatively new, and to the best of our knowledge, it has only been studied in [4], hence it is the actual focus of our paper. Notice, though, that these are nested problems, since each reduces to the previous one by fixing some of its decision variables. Therefore, they are sorted in increasing order of difficulty, and hence will be treated consecutively in the following three sections, by formulating a convex optimization program for each problem based on the solution of the previous one.

III. FILTER PERFORMANCE VERIFICATION PROBLEM

The first problem we will consider consists in verifying whether a given filter F satisfies some performance requirement for all models of the system with parameter θ in a given uncertainty ellipsoid U. The performance requirement has the form $J_\theta(F) \le \gamma$ (recall (2)), for a pre-specified constant $\gamma > 0$.

To solve this problem, we start by approximating $J_\theta(F)$ using (1), which gives

$$\begin{aligned}
&J_\theta(F) \approx \\
&\frac{1}{2\pi}\int \left[\left(1 - 2\,\mathrm{Re}\{FG_o\} + \|F\|^2\|G_o\|^2\right)\Phi_u + \|F\|^2\|H_o\|^2\right] \\
&- \frac{1}{\pi}\int \left[F\Phi_u G'^T - \|F\|^2\Phi_u G_o^* G'^T - \|F\|^2 H_o^* H'^T\right]\Delta\theta \\
&+ \Delta\theta^T\left(\frac{1}{2\pi}\int \left[G'\|F\|^2\Phi_u G'^T + H'\|F\|^2 H'^T\right]\right)\Delta\theta
\end{aligned}$$

Notice that this corresponds to a quadratic expression in $\Delta\theta$, of the form

$$J_\theta(F) \approx A + B\Delta\theta + \Delta\theta^T B^T + \Delta\theta^T C \Delta\theta \qquad (3)$$

where

$$\begin{aligned}
&A := \\
&\frac{1}{2\pi}\int \left[\left(1 - 2\,\mathrm{Re}\{FG_o\} + \|F\|^2\|G_o\|^2\right)\Phi_u + \|F\|^2\|H_o\|^2\right] \\
&B := \frac{1}{2\pi}\int \left[-F\Phi_u G'^T + \|F\|^2\Phi_u G_o^* G'^T + \|F\|^2 H_o^* H'^T\right] \\
&C := \frac{1}{2\pi}\int \left[G'\|F\|^2\Phi_u G'^T + H'\|F\|^2 H'^T\right]
\end{aligned}$$

Therefore, by homogenizing (3) we obtain

$$\begin{aligned}
&J_\theta(F) \le \gamma \\
&\Leftrightarrow \quad \gamma - A - B\Delta\theta - \Delta\theta^T B^T - \Delta\theta^T C\Delta\theta \ge 0 \\
&\Leftrightarrow \quad (\gamma - A)\tau^2 - \tau B\widetilde{\Delta\theta} - \tau\widetilde{\Delta\theta}^T \widetilde{B}^T - \widetilde{\Delta\theta}^T C\widetilde{\Delta\theta} \ge 0 \\
&\Leftrightarrow \quad \begin{bmatrix}\tau \\ \widetilde{\Delta\theta}\end{bmatrix}^T \begin{bmatrix}\gamma - A & -B \\ -B^T & -C\end{bmatrix} \begin{bmatrix}\tau \\ \widetilde{\Delta\theta}\end{bmatrix} \ge 0
\end{aligned}$$

where $\tau \in \mathbb{R}$ is an arbitrary value, and $\widetilde{\Delta\theta} := \tau\Delta\theta$. In terms of $\widetilde{\Delta\theta}$, the ellipsoid U can be written for a fixed τ as

$$\begin{aligned}
U &= \{\widetilde{\Delta\theta} \in \mathbb{R}^p : \widetilde{\Delta\theta}^T P^{-1}\widetilde{\Delta\theta}^T \le \tau^2\} \\
&= \left\{\widetilde{\Delta\theta} \in \mathbb{R}^p : \begin{bmatrix}\tau \\ \widetilde{\Delta\theta}\end{bmatrix}^T \begin{bmatrix}1 & 0 \\ 0 & -P^{-1}\end{bmatrix} \begin{bmatrix}\tau \\ \widetilde{\Delta\theta}\end{bmatrix} \ge 0\right\}
\end{aligned}$$

[1]Notice that we do not assume that $\hat{G}_\theta$ and $\hat{H}_\theta$ are linearly parameterized, but only that they are differentiable functions of θ, which is a very mild and standard assumption [18, Definition 4.3].

To proceed, we need to use the so-called S-procedure (see e.g. [5, page 23]), according to which

$$
\begin{aligned}
&J_\theta(F) \leq \gamma, \quad \forall \theta \in U \\
&\Leftrightarrow \quad \exists \lambda \geq 0, \quad \begin{bmatrix} \gamma - A - \lambda & -B \\ -B^T & -C + \lambda P^{-1} \end{bmatrix} \geq 0
\end{aligned}
$$

This is an LMI feasibility problem, which can be easily solved using standard LMI solvers.

Remark 3.1: The evaluation of the integrals in A, B and C can be done by standard numerical techniques [8]. However, for rational model structures it is also possible to use residue calculus or to consider the integrals as covariance expressions and use special methods for their computation [1, 20].

IV. ROBUST $\mathcal{H}_2$- FILTER DESIGN PROBLEM

We now proceed to the second problem, consisting in designing a robust $\mathcal{H}_2$-filter. Let

$$
F(q) = \sum_{i=0}^{n} f_i q^{-i} = \Lambda_n^H(q) f \tag{4}
$$

be an FIR parametrization of a filter, where $\Lambda_n(q) := [1 \ \cdots \ q^n]^T$ and $f := [f_0 \ \cdots \ f_n]^T$. To proceed, we need to expand the cost function (2) in terms of $\Delta\theta$ and f, as in the previous section, but now we have to first make explicit the dependence on f. Thus, (2) becomes

$$
\begin{aligned}
J_\theta(F) &:= \frac{1}{2\pi} \int \Big[\Phi_u - \hat{G}_\theta \Phi_u \Lambda_n^H f - f^T \Lambda_n \hat{G}_\theta^* \Phi_u \\
&\qquad + f^T \Lambda_n \left(\left| \hat{G}_\theta \right|^2 \Phi_u + \left| \hat{H}_\theta \right|^2 \right) \Lambda_n^H f \Big] \\
&= \frac{1}{2\pi} \int \Phi_u - 2 \operatorname{Re} \left[\frac{1}{2\pi} \int \hat{G}_\theta \Phi_u \Lambda_n^H \right] f \\
&\quad + f^T \left[\frac{1}{2\pi} \int \Lambda_n \left(\left| \hat{G}_\theta \right|^2 \Phi_u + \left| \hat{H}_\theta \right|^2 \right) \Lambda_n^H \right] f \\
&= \frac{1}{2\pi} \int \Phi_u - 2 \operatorname{Re} \left[\frac{1}{2\pi} \int G_o \Phi_u \Lambda_n^H \right] f \\
&\quad - 2\Delta\theta^T \operatorname{Re} \left[\frac{1}{2\pi} \int G' \Phi_u \Lambda_n^H \right] f \\
&\quad + f^T \left[\frac{1}{2\pi} \int \Lambda_n \left(\left| \hat{G}_\theta \right|^2 \Phi_u + \left| \hat{H}_\theta \right|^2 \right) \Lambda_n^H \right] f \\
&= J_0 - 2\Delta\theta^T J_1 + f^T J_2 f
\end{aligned}
$$

where

$$
\begin{aligned}
J_0 &:= \frac{1}{2\pi} \int \Phi_u - 2 \left[\frac{1}{2\pi} \int G_o \Phi_u \Lambda_n^H \right] f \\
J_1 &:= \frac{1}{2\pi} \operatorname{Re} \left[\int G' \Phi_u \Lambda_n^H \right] f \\
J_2 &:= \frac{1}{2\pi} \int \Lambda_n \left(\left| \hat{G}_\theta \right|^2 \Phi_u + \left| \hat{H}_\theta \right|^2 \right) \Lambda_n^H
\end{aligned}
$$

Notice that both J_0 and J_1 are affine in the parameter vector of the filter, f. If we factorize Φ_u as $|U|^2$, where U is stable and minimum phase, then J_2 can be approximated as

$$
\begin{aligned}
J_2 &= \frac{1}{2\pi} \int \Lambda_n \left(\left| \hat{G}_\theta \right|^2 \Phi_u + \left| \hat{H}_\theta \right|^2 \right) \Lambda_n^H \\
&= \frac{1}{2\pi} \int \begin{bmatrix} \Lambda_n \hat{G}_\theta^* U^* & \Lambda_n \hat{H}_\theta^* \end{bmatrix} \begin{bmatrix} \hat{G}_\theta U \Lambda_n^H \\ \hat{H}_\theta \Lambda_n^H \end{bmatrix} \\
&\approx \frac{1}{M} \sum_{k=0}^{M-1} \begin{bmatrix} \hat{G}_\theta(e^{j\omega_k}) U(e^{j\omega_k}) \Lambda_n^H(e^{j\omega_k}) \\ \hat{H}_\theta(e^{j\omega_k}) \Lambda_n^H(e^{j\omega_k}) \end{bmatrix}^H \\
&\qquad \cdot \begin{bmatrix} \hat{G}_\theta(e^{j\omega_k}) U(e^{j\omega_k}) \Lambda_n^H(e^{j\omega_k}) \\ \hat{H}_\theta(e^{j\omega_k}) \Lambda_n^H(e^{j\omega_k}) \end{bmatrix} \\
&= \left[D_0 + \sum_{k=1}^{p} D_k \Delta\theta_k \right]^H \left[D_0 + \sum_{k=1}^{p} D_k \Delta\theta_k \right]
\end{aligned} \tag{5}
$$

where $M > 0$, $\omega_k := (2\pi/M)k$ for $k = 0, \ldots, M-1$, and

$$
D_k := \begin{bmatrix} \frac{1}{\sqrt{M}} G_k(e^{j\omega_0}) U(e^{j\omega_0}) \Lambda_n^H(e^{j\omega_0}) \\ \vdots \\ \frac{1}{\sqrt{M}} G_k(e^{j\omega_{M-1}}) U(e^{j\omega_{M-1}}) \Lambda_n^H(e^{j\omega_{M-1}}) \\ \frac{1}{\sqrt{M}} H_k(e^{j\omega_0}) \Lambda_n^H(e^{j\omega_0}) \\ \vdots \\ \frac{1}{\sqrt{M}} H_k(e^{j\omega_{M-1}}) \Lambda_n^H(e^{j\omega_{M-1}}) \end{bmatrix}; \qquad k = 0, \ldots, p
$$

where

$$
G_k := \left. \frac{\partial \hat{G}_\theta}{\partial \theta_k} \right|_{\theta=\theta_o}; \quad H_k := \left. \frac{\partial \hat{H}_\theta}{\partial \theta_k} \right|_{\theta=\theta_o}; \qquad k = 1, \ldots, p
$$

Therefore, for a fixed $\gamma > 0$ we have by homogenization that

$$
\begin{aligned}
&J_\theta(F) \leq \gamma \\
&\Leftrightarrow \ J_0 - 2\Delta\theta^T J_1 + f^T J_2 f \leq \gamma \\
&\Leftrightarrow \ \gamma - J_0 + 2\Delta\theta^T J_1 \\
&\qquad - f^T \left[D_0 + \sum_{k=1}^{p} D_k \Delta\theta_k \right]^H \left[D_0 + \sum_{k=1}^{p} D_k \Delta\theta_k \right] f \geq 0 \\
&\Leftrightarrow \ (\gamma - J_0)\tau^2 + 2\widetilde{\Delta\theta}^T J_1 \tau \\
&\quad - \left[D_0 f \tau + \sum_{k=1}^{p} D_k f \widetilde{\Delta\theta}_k \right]^H \left[D_0 f \tau + \sum_{k=1}^{p} D_k f \widetilde{\Delta\theta}_k \right] \geq 0 \\
&\Leftrightarrow \ \begin{bmatrix} \tau \\ \widetilde{\Delta\theta} \end{bmatrix}^T \left(\begin{bmatrix} \gamma - J_0 & J_1^T \\ J_1 & 0 \end{bmatrix} - \right. \\
&\qquad \left. [D_0 f \ \cdots \ D_p f]^H [D_0 f \ \cdots \ D_p f] \right) \begin{bmatrix} \tau \\ \widetilde{\Delta\theta} \end{bmatrix} \geq 0
\end{aligned}
$$

According to the S-procedure and an application of Schur complements [30], we obtain

$$
\begin{aligned}
& J_\theta(F) \leq \gamma, \quad \forall \theta \in U \\
\Leftrightarrow\ & \exists \lambda \geq 0,\ \begin{bmatrix} \gamma - J_0 - \lambda & J_1^T \\ J_1 & \lambda P^{-1} \end{bmatrix} \\
& \qquad - [D_0 f\ \cdots\ D_p f]^H [D_0 f\ \cdots\ D_p f] \geq 0 \\
\Leftrightarrow\ & \exists \lambda \geq 0,\ \left[\begin{array}{cc|c} \gamma - J_0 - \lambda & J_1^T & \multirow{2}{*}{$[D_0 f\ \cdots\ D_p f]^H$} \\ J_1 & \lambda P^{-1} & \\ \hline \multicolumn{2}{c|}{[D_0 f\ \cdots\ D_p f]} & I \end{array}\right] \geq 0
\end{aligned} \tag{6}
$$

Since J_0 and J_1 are affine in f, while the other matrices are known, (6) is an LMI in f. Therefore, the problem of designing a robust filter which solves $\min_F \max_{\theta \in U} J_\theta(F)$ is equivalent to the following semidefinite program

$$
\begin{aligned}
\min_{\gamma, \lambda, f} \quad & \gamma \\
\text{s.t.} \quad & \left[\begin{array}{cc|c} \gamma - J_0 - \lambda & J_1^T & [D_0 f\ \cdots\ D_p f]^H \\ J_1 & \lambda P^{-1} & \\ \hline \multicolumn{2}{c|}{[D_0 f\ \cdots\ D_p f]} & I \end{array}\right] \geq 0, \\
& \lambda \geq 0
\end{aligned} \tag{7}
$$

Remark 4.1: The procedure presented in this section follows similar lines to that described in [2, Section 3.2].

Remark 4.2: The FIR parametrization of F, (4), can be easily replaced by any other linear parametrization, e.g. one based on Laguerre basis functions [26].

Remark 4.3: The Riemann sum approximation in (5) seems unavoidable, because in order to obtain a robust convex optimization problem without such an approximation we need to factorize the integral in the second line of (5) into terms affine in the $\Delta\theta_k$'s. As shown in [12], the existence of such a factorization in general is possible only if $n \leq 2$ and $p \leq 2$, which is quite restrictive.

V. INPUT DESIGN FOR ROBUST $\mathcal{H}_2$ FILTERING

Now consider the problem of designing an input signal $\{u_t^{id}\}$, considered as a wide-sense stationary process of zero mean and spectrum Φ_u^{id}, with which we estimate a model of the system, $(\hat{G}_\theta, \hat{H}_\theta)$, with an associated uncertainty ellipsoid U described by a covariance matrix P (appropriately scaled so that U corresponds to a confidence region of a prescribed confidence level). If the number of samples, N, is sufficiently large, and an identification method such as PEM is used, then P is given by an expression of the form

$$
P^{-1} = N \int_{-\pi}^{\pi} K(e^{j\omega}) \Phi_u^{id}(e^{j\omega}) d\omega + N P_0^{-1}
$$

where $K : \mathbb{T} \to S_+^p$ ($\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ and S_+^p is the space of positive semi-definite matrices of dimension $p \times p$) and $P_0^{-1} = (P_0^{-1})^T \geq 0$ is constant. Notice that P^{-1} is an affine function of Φ_u^{id}, so if we parameterize this spectrum in a linear fashion, P^{-1} will be affine in those parameters. This is explained in detail in [13, 14]. Therefore, P^{-1} can be written as

$$
P^{-1} = N P_0 + \sum_{i=1}^{q} P_i r_i \tag{8}
$$

where $P_0, \cdots, P_q$ are fixed matrices which depend on K and the parametrization of Φ_u^{id}, and the r_i's are generalized moments of Φ_u^{id} (i.e., weighted integrals of Φ_u^{id} [29]), which now correspond to the design variables of the input design problem.

As mentioned before, P^{-1} should be scaled according to the α-level of the confidence ellipsoids being considered. This means in particular that P^{-1} should be divided by $\chi^2_{1-\alpha}(p)$ (the $(1-\alpha)$-percentile of a χ^2 distribution with p degrees of freedom) [18, Section II.2].

Additionally, the parameterization of the input spectrum Φ_u^{id} should typically satisfy some power, nonnegativity and/or covariance extension constraints, which can be written as LMI constraints on the r_i's, i.e.,

$$
F_0 + \sum_{i=1}^{q} F_i^{(1)} r_i + \sum_{i=1}^{s} F_i^{(2)} h_i \geq 0 \tag{9}
$$

where $F_0^{(1)}, \ldots, F_q^{(1)}, F_1^{(1)}, \ldots, F_s^{(2)}$ are constant hermitian matrices, and $h_1, \ldots, h_s$ are auxiliary variables (which might appear e.g. from the KYP Lemma [13, 14]).

We wish to design the input signal Φ_u^{id} in order to minimize the cost of the robust optimal filter. Plugging (8) and (9) into (7) gives

$$
\begin{aligned}
\min_{\gamma, \lambda, f, r_1, \ldots, r_q} \quad & \gamma \\
\text{s.t.} \quad & \left[\begin{array}{cc|c} \gamma - J_0 - \lambda & J_1^T & [D_0 f\ \cdots\ D_p f]^H \\ J_1 & \lambda N P_0 + \sum_{i=1}^{q} P_i \lambda r_i & \\ \hline \multicolumn{2}{c|}{[D_0 f\ \cdots\ D_p f]} & I \end{array}\right] \\
& \qquad \geq 0, \\
& F_0 + \sum_{i=1}^{q} F_i^{(1)} r_i + \sum_{i=1}^{s} F_i^{(2)} h_i \geq 0, \\
& \lambda \geq 0
\end{aligned} \tag{10}
$$

This problem is non convex in the decision variables, because of the products λr_i present in (10). This can be easily solved by reformulating (10) as

$$
\begin{aligned}
\min_{\gamma, \lambda, f, \tilde{r}_1, \ldots, \tilde{r}_q} \quad & \gamma \\
\text{s.t.} \quad & \left[\begin{array}{cc|c} \gamma - J_0 - \lambda & J_1^T & [D_0 f\ \cdots\ D_p f]^H \\ J_1 & \lambda N P_0 + \sum_{i=1}^{q} P_i \tilde{r}_i & \\ \hline \multicolumn{2}{c|}{[D_0 f\ \cdots\ D_p f]} & I \end{array}\right] \\
& \qquad \geq 0, \\
& \lambda F_0 + \sum_{i=1}^{q} F_i^{(1)} \tilde{r}_i + \sum_{i=1}^{s} F_i^{(2)} \tilde{h}_i \geq 0, \\
& \lambda \geq 0
\end{aligned} \tag{11}
$$

where $\tilde{r}_i := \lambda r_i$ $(i = 1, \ldots, q)$ and $\tilde{h}_i := \lambda h_i$ $(i = 1, \ldots, s)$. This is a standard semidefinite program whose solution provides the optimal input for the robust filtering problem.

Remark 5.1: Notice that $\lambda = 0$ cannot yield a feasible solution of (10), since otherwise the first constraint in (10) would become sign indefinite. The same happens in (11), under some input constraints (e.g. for an input power constraint, with P^{-1} parameterized using a Tchebycheff system approach [29]), for which $\lambda = 0$ implies that the $\tilde{r}_i$ and $\tilde{h}_i$ terms are also zero.

Remark 5.2: In the formulation of the input design problem of this section, the following issue arises: the robust filter to be designed should depend on $\hat{\theta}$, an estimate of the true plant to be obtained after the experiment is performed, not on θ_o, which is unknown. One way to overcome this problem is to notice that $\|x\|_{P^{-1}} := \sqrt{x^T P^{-1} x}$ is a norm in $\mathbb{R}^p$, so by the triangular inequality we have that

$$\|\theta - \theta_o\|_{P^{-1}} \leq \|\theta - \hat{\theta}\|_{P^{-1}} + \|\hat{\theta} - \theta_o\|_{P^{-1}}$$

where θ is the vector which is "believed" to be the true parameter (at the time when the robust filter is to be designed). Since the robust filter is designed to give a reasonable performance for all plants parameterized with θ such that $\|\theta - \hat{\theta}\|_{P^{-1}} < 1$, and we have that $\|\hat{\theta} - \theta_o\|_{P^{-1}} < 1$ from the identification experiment (with high probability), it follows that the input design should consider the performance cost for all plants with θ such that $\|\theta - \theta_o\|_{P^{-1}} < 2$. Therefore, a simple (but perhaps conservative) solution is to divide P^{-1} by 4 in the previous computations in order to obtain an input signal which guarantees a worst case performance γ when the robust filter is designed based on $\hat{\theta}$.

VI. ILLUSTRATIVE EXAMPLE

Consider the system

$$y_t = \frac{2}{1 - 0.8q^{-1}} u_t + (1 - 0.6q^{-1}) e_t \tag{12}$$

for which the following model (with $p = 4$ parameters) is to be estimated:

$$y_t = \frac{b_0}{1 + a_1 q^{-1}} u_t + (c_0 + c_1 q^{-1}) e_t$$

The input signal to be reconstructed is white noise of variance 10000, thus $\Phi_u = 10000$. To keep the example simple, we will assume that during the identification stage, the power of the input signal should be at most equal to 1, and we take only $N = 1000$ samples. We will consider confidence ellipsoids with an α-level of 95% (which gives $\chi^2_{0.05}(4) \approx 9.49$ [21]).

If we had perfect knowledge of the true plant, the optimal filter (by the Wiener approach [15]) is given by

$$\hat{u}_t = \frac{0.5 - 0.4q^{-1}}{1 - 5.1 \cdot 10^{-5} q^{-1} + 1.2 \cdot 10^{-5} q^{-2}} y_t \tag{13}$$

The frequency response of this nominal filter together with the plant and the disturbance transfer functions is shown in Figure 1.

To design a design a robust filter, we will consider an FIR parametrization of the filter with[2] $n = 5$, and a discretization of the Riemann sum in (5) with $M = 100$. For the input design problem, we consider a parametrization of P^{-1} based on a Tchebycheff system approach [29], which in this case depends on 4 parameters. As a benchmark, we also design a robust filter based on a model identified using white noise of unit variance as input.

The optimal cost of the filters designed using:

1) Perfect knowledge of the true system (nominal Wiener filter)
2) A model identified using white noise of unit variance
3) A model identified using an optimal input of unit variance

are presented in Table I. For the robust filters we have scaled the information matrix according to Remark 5.2. The optimal input is a sinusoid[3] of frequency 0.031 (rad/s) and amplitude $\sqrt{2}$.

From Table I we can see the great improvement on the worst case performance obtained by identifying the model using a carefully designed input.

VII. CONCLUSIONS

In this paper an approach to design optimal input signals for robust $\mathcal{H}_2$ filtering has been developed. The idea is based on the assumption that the number of samples for identification is large, so that asymptotic closed-form expressions can be derived for the performance of the filters. The optimization problem for the design of the optimal input turns out to be convex and can be efficiently solved using standard interior point SDP solvers. Finally, with an illustrative example we have shown the great improvement

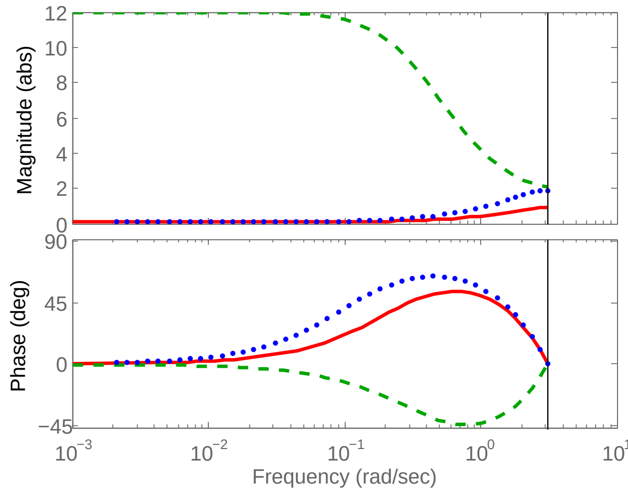

Fig. 1. Bode diagram of the Wiener filter (13) (red solid), the plant G_o (green dashed) and the disturbance H_o (blue dotted) from equation (12).

TABLE I
OPTIMAL COSTS FOR THE FILTERS OF SECTION VI

Nominal Filter	Robust Filter with White Noise	Robust Filter with Optimal Input
0.79752	192.94635	13.30903

[2]The rationale behind this choice of n comes from the observation that only the first 2 pulse response coefficients of the Wiener filter are significant for this example, so we expect that its robust counterpart could be well described by a small order FIR system.

[3]Due to the sinusoidal nature of the optimal input, the signal cannot be realized as an AR process (and the Yule-Walker equations are ill-conditioned, so it is difficult to obtain an AR approximation). To find the frequency of the sinusoid from the moments of the Tchebycheff system, the MUSIC method can be employed [23].

that this technique can yield on the performance of a robust filter.

References

[1] Åström, K. J. [1970]. *Introduction to Stochastic Control Theory*, Academic Press, New York.

[2] Ben-Tal, A. and Nemirovski, A. [1998]. Robust convex optimization, *Mathematics of Operations Research* **23(4)**: 769–805.

[3] Bombois, X. [2000]. *Connecting Prediction Error Identification and Robust Control Analysis: a new framework*, PhD thesis, Université Catholique de Louvain.

[4] Bombois, X. and Hjalmarsson, H. [2009]. Optimal input design for robust H_2 deconvolution filtering, *15th IFAC Symposium on System Identification*, Saint-Malo, France.

[5] Boyd, S., El Ghaoui, L., Feron, E. and Balakrishnan, V. [1994]. *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, Philadelphia, USA.

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

[7] Chernoff, H. [1972]. *Sequential Analysis and Optimal Design*, SIAM Press, Philadelphia.

[8] Davis, P. J. and Rabinowitz, P. [1984]. *Methods of Numerical Integration, 2nd Edition*, Academic Press, Orlando.

[9] Gerencsér, L. and Hjalmarsson, H. [2005]. Adaptive input design in system identification, *44th IEEE Conference on Decision and Control*, Seville, Spain.

[10] Gerencsér, L. and Hjalmarsson, H. [2009]. Identification of ARX systems with non-stationary inputs – asymptotic analysis with application to adaptive input design, *Automatica* **45(3)**: 623–633.

[11] Geromel, J. C., De Oliveira, M. C. and Bernussou, J. [2002]. Robust filtering of discrete-time linear systems with parameter dependent Lyapunov functions, *SIAM Journal on Control and Optimization* **41(3)**: 700–711.

[12] Hildebrand, R. [2007]. Exactness of sums of squares relaxations involving 3×3 matrices and Lorentz cones, *Linear Algebra and Its Applications* **426**: 815–840.

[13] Jansson, H. [2004]. *Experiment design with applications in identification for control*, PhD thesis, Royal Institute of Technology (KTH).

[14] Jansson, H. and Hjalmarsson, H. [2005]. Input design via LMIs admitting frequency-wise model specifications in confidence regions, *IEEE Transactions on Automatic Control* **50(10)**: 1534–1549.

[15] Kailath, T., Sayed, A. H. and Hassibi, B. [2000]. *Linear Estimation*, Prentice Hall, Upper Saddle River, NJ.

[16] Kalman, R. E. [1960]. A new approach to linear filtering and prediction problems, *Transactions of the ASME* **82**: 34–45.

[17] Kolmogorov, A. N. [1941]. Interpolation and extrapolation of stationary random processes, *Bull. Acad. Sci. USSR* **5**.

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

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

[20] Söderström, T., Šebek, M., Ježek, J., Kučera, V. and Henrion, D. [2002]. An evaluation of algorithms for computing the covariance function of a multivariable ARMA process, *European Journal of Control* **8(4)**: 315–325.

[21] Söderström, T. and Stoica, P. [1989]. *System Identification*, Prentice Hall, Hertfordshire, United Kingdom.

[22] Sternad, M. and Ahlén, A. [1996]. $\mathcal{H}_2$ design of nominal and robust discrete time filters, *in* M. Grimble and V. Kučera (eds), *A Polynomial Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Robust Control Design*, Springer-Verlag, London, chapter 5.

[23] Stoica, P. and Moses, R. L. [2005]. *Spectral Analysis of Signals*, Prentice Hall.

[24] Sun, K. and Packard, A. [2005]. Robust H_2 and H_∞ filters for uncertain LFT systems, *IEEE Transactions on Automatic Control* **50(5)**: 715–720.

[25] Veenman, J., Köroğlu, H. and Scherer, C. W. [2008]. An IQC approach to robust estimation against perturbations of smoothly time-varying parameters, *47th IEEE Conference on Decision and Control*, Cancun, Mexico, pp. 2533–2538.

[26] Wahlberg, B. [1991]. System identification using Laguerre models, *IEEE Transactions on Automatic Control* **36**(5): 551–562.

[27] Walter, E. and Pronzato, L. [1997]. *Identif. of Parametric Models from Experimental Data*, Springer-Verlag, Berlin.

[28] Wiener, N. [1949]. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series, With Engineering Applications*, John Wiley & Sons, New York.

[29] Zarrop, M. [1979]. *Optimal Experiment Design for Dynamic System Identification*, Springer, Berlin.

[30] Zhang, F. [2005]. *The Schur Complement and Its Applications*, Springer, New York.

[31] Zhou, K., Doyle, J. C. and Glover, K. [1996]. *Robust and Optimal Control*, Prentice Hall, Englewood Cliffs, New Jersey.