<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *18th European Control Conference, ECC 2019; Naples; Italy; 25 June 2019 through 28 June 2019*.

Citation for the original published paper:

Müller, M I., Rojas, C R. (2019)

Risk-Coherent H_∞ -optimal Filter Design Under Model Uncertainty with Applications to MISO Control

In: *2019 18th European Control Conference, ECC 2019*, 8795947 (pp. 1461-1466).

Institute of Electrical and Electronics Engineers (IEEE)

<https://doi.org/10.23919/ECC.2019.8795947>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-256077>

Risk-Coherent $\mathcal{H}_\infty$ -optimal Filter Design Under Model Uncertainty with Applications to MISO Control

Matias I. Müller¹ and Cristian R. Rojas¹

Abstract—This work presents a framework to address the problem of designing discrete-time LTI (linear and time-invariant) multiple-input and multiple-output (MIMO) filters, aiming to optimize the performance of a system when model uncertainty is considered. Additionally, we present an interesting application to control design for disturbance rejection under model uncertainty. To account for this uncertainty we employ coherent measures of risk, which are a family of measures in theory of risk. We particularly discuss which measures are suitable by comparing the conditional value-at-risk (CVaR) to other three common designs. Using a scenario approach, we derive a convex optimization problem based on linear matrix inequalities (LMIs), whose solution minimizes the risk of falling into poor $\mathcal{H}_\infty$ performance. Finally, we present an application to multiple-input and single-output (MISO) control design under model uncertainty in the auto-covariance function of the output noise, comparing approaches minimizing different notions of risk.

I. INTRODUCTION

The different sources of uncertainty in model-based control design have different consequences in the final control performance. One of them regards the lack of knowledge we have about the plant which ends up in a mismatch between the true dynamics of the plant and the model used for designing a controller, commonly known as *modeling error*.

The problem of quantifying the impact of the modeling error in the closed-loop $\mathcal{H}_2$ performance [6] has been widely studied in the literature (see, e.g., [2], [3], [4]), where it has been shown that the performance of the control loop depends on the largest gain, in frequency, of the modeling error, also known as the $\mathcal{H}_\infty$ -norm [5] of the modeling error. Here, the classic approach is to synthesize a controller that attains a given performance for any modeling error whose $\mathcal{H}_\infty$ -norm is bounded by some given value. This approach is known as robust (worst-case) control design, since the controller attains (at least) that performance for any modeling error.

The modeling error is a naturally unknown function whose largest gain might be hard to upper bound. Instead, one usually knows a parametric family to which the true process belongs, having different confidence levels on different sets of parameters. From a Bayesian perspective, this knowledge can be summarized as a probability density function (pdf) over the set of parameters parametrizing the family of plants, which allows us to quantify the uncertainty of the model.

Since the true process belongs to a parametrized family, so does the modeling error. At this point, it is worth to mention that the deterministic problem of controlling a parametric family of plants has been previously addressed when their parameters belong to a polytopic convex set, also known as polytopic control. Although robust stability analysis for polytopic systems has been properly studied in the literature [3], the extensions to robust control synthesis of these systems are rather limited. For example, authors in [11] introduce a clever relaxation of the exact problem, in order to avoid non-affine functions of decision variables, that leads to a set of LMIs for $\mathcal{H}_2/\mathcal{H}_\infty$ synthesis.

In this work, we focus on minimizing the $\mathcal{H}_\infty$ performance of the closed loop by considering the probabilistic information provided by the distributions of the parameters in the parametric model, leading to -hopefully- better results than what the robust synthesis provides. The proposed framework aims to design a controller that minimizes the uncertainty of falling into poor $\mathcal{H}_\infty$ performance. A method for appropriately measuring the uncertainty has been studied in the theory of risk in finance [7], giving rise to the notion of *coherent measures of risk* [8]. Among the many different coherent measures of risk, we mention *expectation*, *worst-case scenario*, and the so called *conditional value-at-risk* (CVaR) [9]. By means of an affine parametrization, we end up with a filter design problem casted as an LMI (linear matrix inequality).

The notion of risk in control theory has been already employed by several authors. In [12], CVaR has been used to impose convex constraints in LQR (linear-quadratic regulator) under uncertainty problems. Recently, the authors in [13] have derived a framework for risk-coherent control design when the goal is to minimize the impact of disturbances in the ℓ_2 performance of the closed loop, with an interesting application to cyberphysical security [14]. Other notions of non-coherent measures of risk have been also invoked in control theory, such as the risk sensitive criterion [15].

The contributions of this paper are twofold:

- 1) We cast the problem of minimizing the CVaR of the $\mathcal{H}_\infty$ -norm in the filter design problem as an LMI by means of the Kalman-Yakubovic-Popov lemma [16], where the decision variables are the coefficients of a finite impulse response (FIR) filter.
- 2) We extend the notion of the filter design problem to the control problem of $\mathcal{H}_\infty$ disturbance rejection under model uncertainty, comparing how different coherent-risk measures (nominal design, expectation, worst-case and CVaR) perform on an illustrative example.

*This work was partially supported by the Swedish Research Council under contract 2016-06079.

¹ The authors are with the Department of Automatic Control, KTH Royal Institute of Technology, SE 100 44 Stockholm, Sweden. Corresponding author e-mail: mimr2@kth.se.

In the light of the results, we expose how CVaR can be considered as a useful tool for control design under model uncertainty that can be used efficiently by means of the proposed design method, which only involves LMI optimization.

The remainder of the paper is organized as follows. Section II describes the problem of $\mathcal{H}_\infty$ -optimal filter design. Section III states the fundamental tool to impose frequency domain restrictions (such as $\mathcal{H}_\infty$ restrictions) in terms of an LMI. Section IV establishes a framework for optimal risk-theoretical $\mathcal{H}_\infty$ design when the coherent measure of risk is CVaR, while Section V shows how the CVaR filter design problem can be extended to control design under model uncertainty, comparing its performance against three other coherent measures of risk. Finally, Section VI presents the conclusions of this work.

A. Notation

The positive part function is $[X]_+ = \max\{0, X\}$, and Dirac's delta function $\delta(t - t_0)$ integrates to 1 and it is equal to zero everywhere else but at $t = t_0$. The $n \times n$ identity matrix is I_n and $\mathbf{1}_n := [1 \ 1 \ \dots \ 1]^\top \in \mathbb{R}^n$. The transpose and Hermitian operators are $(\cdot)^\top$ and $(\cdot)^H$, respectively. $\mathbb{S}_n$ is the set of $n \times n$ Hermitian matrices, while for $A, B \in \mathbb{S}_n$, $A \succ B$ ($A \succeq B$) means that $A - B$ is positive definite (semi-definite). The Z-transform variable is z , and the induced ℓ_2 norm of a function F is $\|F\|_2^2 := \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|^2 d\omega = \lim_{t \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |f(t)|^2$, where $\{f(t)\}_{t=1}^\infty$ is the inverse Z-transform of $F(z)$. The set $\mathcal{H}_2$ is composed by all analytic functions F such that $\|F\|_2 < \infty$. The 1-sample delay operator is q^{-1} and it operates on time series. For a random variable X , $\text{supp } X$ is the support of X , while $\mathbb{E}\{X\}$ denotes expectation of X . $X \sim \text{Unif}(\mathcal{A})$ means that X is uniformly distributed over set $\mathcal{A}$. The Hilbert space composed by all random variables with bounded mean and covariance is $\mathbb{L}$.

II. PROBLEM STATEMENT

We consider the general filter design problem depicted in Fig. 1, where the goal is to minimize some spectral property of the output signal $y := \{y(t)\}_{t=1}^\infty$, where $y(t) = P_0(q)u(t)$ for a given input sequence $u := \{u(t)\}_{t=1}^\infty$, and where P_0 is a causal, stable, linear and time-invariant (LTI) multivariable filter. At the same time, P_0 is composed by deterministic multivariable filters F_0 , E_0 and a design variable Q (being a multivariable system itself). Under this assumption, the traditional $\mathcal{H}_\infty$ filtering problem aims to design Q such that the maximum gain (in the frequency domain) of signal $y(t) = P_0(q)u(t) = [F_0(q) - E_0(q)Q(q)]u(t)$ is minimized. In other words, it attempts to solve the deterministic minimization (DM) problem

$$(DM) \quad \min_{Q \in \mathcal{H}_2} \|F_0 - QE_0\|_\infty.$$

In a more general setup, we consider parametric models $F(q; \theta)$ and $E(q; \theta)$, where $\theta \in \mathbb{L}^{n_\theta}$ is a random vector in $\mathbb{R}^{n_\theta}$ of compact support $\text{supp } \theta = \Theta$, with a known probability density function (pdf) $p(\theta)$, i.e., $\theta \sim p(\theta)$. The deterministic problem is retrieved by considering $p(\theta) =$

Fig. 1. The optimization problem can be casted as minimizing the $\mathcal{H}_\infty$ norm of the augmented system P_0 .

$\delta(\theta - \theta_0)$, that is, $P_0(q) = P(q; \theta_0)$, $F_0(q) = F(q; \theta_0)$ and $E_0(q) = E(q; \theta_0)$, which we refer from now on as the nominal models.

Remark 1: Within a Bayesian framework, $p(\theta)$ represents our belief of θ_0 being equal to each θ in Θ . The problem of acquiring this knowledge might be difficult and it is out of the scope of this work, however, we mention that it can be obtained from data-driven modeling, using techniques such as PEM [1]. ■

Assumption 1: Filters $F(q; \theta)$ and $E(q; \theta)$ are stable for every $\theta \in \Theta$. ■

Under the presented set-up, the cost function in (DP) is a random variable depending on the realization of parameter θ . Denote this cost function, parametrized by Q , as $J_Q(\theta) := \|F(\cdot; \theta) - E(\cdot; \theta)Q\|_\infty$. Observe that $J_Q(\theta) \in \mathbb{L}$, $\forall Q \in \mathcal{H}_2$, since $F(q; \theta) - Q(q)E(q; \theta)$ is stable for every $\theta \in \Theta$. To account for the random nature of J_Q , we consider a functional $\mathcal{R}: \mathbb{L} \rightarrow (-\infty, \infty]$ that operates on random variables. Functional $\mathcal{R}$ is what we call a risk measure, and leads us to the following risk minimization (RM) problem:

$$(RM) \quad \begin{cases} Q^* &= \arg \min_{Q \in \mathcal{H}_2} \mathcal{R}\{J_Q(\theta)\}, \\ J_Q(\theta) &= \|F(\cdot; \theta) - QE(\cdot; \theta)\|_\infty. \end{cases}$$

The filter variable Q manipulates the pdf of J_Q , in the sense that different Q 's lead to different pdf's of J . One might choose Q^* such that pdf of J_{Q^*} attains the least risk of falling into large values. Notice that an important issue in (RM) is, precisely, choosing the risk measure.

A. The risk framework

As previously discussed, there are several ways to choose $\mathcal{R}$, but we will restrict our attention to the family of coherent measures of risk due to their attractive properties such as convexity [8]. Within this family, we will derive a solution to (RM) when the risk measure is chosen as CVaR defined, for a random variable $X \sim p_x(X)$ (that can be linked to the random cost function J_Q) and for a user-defined parameter $\alpha \in (0, 1)$, as [17]

$$\text{CVaR}_\alpha(X) := \frac{1}{1 - \alpha} \int_{x \geq \text{VaR}_\alpha(X)} x p_x(x) dx, \quad (1)$$

where $\text{VaR}_\alpha(X)$ is the α -value at risk of X [18]. It is known [19] that the definition in (1) is equivalent to the following optimization problem:

$$\text{CVaR}_\alpha(X) = \min_{\mu \in \mathbb{R}} \mu + \frac{1}{1 - \alpha} \mathbb{E}\{[X - \mu]_+\}. \quad (2)$$

Under this fact, we can cast the solution of (RM) as the solution of the CVaR minimization (CM)

$$(CM) \quad \begin{cases} Q^* &= \arg \min_{Q \in \mathcal{H}_2, \mu \in \mathbb{R}} \mu + \frac{1}{1-\alpha} \mathbb{E} \{ [J_Q - \mu]_+ \}, \\ J_Q(\theta) &= \|F(\cdot; \theta) - QE(\cdot; \theta)\|_\infty^2. \end{cases}$$

This is the problem we aim to solve in the rest of the paper. The squared performance is to preserve convexity in the final optimization problem, as it will be shown in the following sections. Technically, we are solving a relaxation of (DM).

Remark 2: As discussed in [13], coherent measures of risk such as expectation and worst-case scenario are retrieved by $\lim_{\alpha \rightarrow 0^+} \text{CVaR}_\alpha$ and $\lim_{\alpha \rightarrow 1^-} \text{CVaR}_\alpha$, respectively. This means that α can be seen as a tuning parameter that balances between average and worst-case performance. ■

B. State-space realization

It is important, for future reference, to remark that since $(F(\cdot; \theta), E(\cdot; \theta), Q)$ are LTI filters, they admit state-space representations $(\mathcal{S}_F(\theta), \mathcal{S}_E(\theta), \mathcal{S}_Q)$, where $\mathcal{S}_\Sigma$ defines a state-space realization $(A_\Sigma(\theta), B_\Sigma(\theta), C_\Sigma(\theta), D_\Sigma(\theta))$, $\Sigma \in \{F(\cdot, \theta), E(\cdot, \theta), Q\}$, i.e., each filter satisfies the following state-space model

$$\begin{cases} x_\Sigma(t+1) &= A_\Sigma(\theta)x_\Sigma(t) + B_\Sigma(\theta)u_\Sigma(t) \\ y_\Sigma(t) &= C_\Sigma(\theta)x_\Sigma(t) + D_\Sigma(\theta)u_\Sigma(t) \end{cases} \quad (3)$$

By defining $P(q; \theta) := F(q; \theta) - Q(q)E(q; \theta)$, the augmented system depicted in Fig. 1, considering the augmented state vector as $x := [x_F \ x_Q \ x_E]^\top$, can be expressed as

$$x(t+1) = \underbrace{\begin{bmatrix} A_F(\theta) & 0 & 0 \\ 0 & A_Q & 0 \\ 0 & B_E(\theta)C_Q & A_E(\theta) \end{bmatrix}}_{:=A(\theta)} x(t) + \underbrace{\begin{bmatrix} B_F(\theta) \\ B_Q \\ B_E(\theta)D_Q \end{bmatrix}}_{:=B(\theta)} u(t), \quad (4)$$

$$y(t) = \underbrace{\begin{bmatrix} C_F(\theta) & -C_Q & -D_Q C_E(\theta) \end{bmatrix}}_{:=C(\theta)} x(t) + \underbrace{(D_F(\theta) - D_Q D_E(\theta))}_{:=D(\theta)} u(t), \quad (5)$$

which means that $P(z; \theta) = C(\theta)(zI - A(\theta))^{-1}B(\theta) + D(\theta)$.

In the following section we restate a generalized version of the KYP lemma, which will allow us to recast the problem as LMIs.

III. GENERALIZED

KALMAN-YAKUBOVIC-POPOV LEMMA

The Kalman-Yakubovic-Popov (KYP) lemma [20], [21], [22] is one of the most important tools in the field of dynamical systems analysis, feedback control and signal processing. The result establishes a necessary and sufficient condition for frequency-domain constraints in terms of LMIs

that can be exploited for systems synthesis. Let $T(z) = C(zI - A)^{-1}B + D \in \mathbb{C}^{n \times m}$ for some matrices (A, B, C, D) . The KYP lemma translates a frequency domain inequality (FDI), defined by matrices $(\Pi, \Phi, \Psi) \in \mathbb{S}_{m+n} \times \mathbb{S}_2 \times \mathbb{S}_2$, of the form

$$(FDI) \quad \begin{cases} \begin{bmatrix} T(z) \\ I_m \end{bmatrix}^H \Pi \begin{bmatrix} T(z) \\ I_m \end{bmatrix} < 0, \quad \forall z \in \Lambda(\Phi, \Psi), \\ \Lambda(\Phi, \Psi) = \left\{ z \in \mathbb{C} : \begin{bmatrix} z \\ 1 \end{bmatrix}^H \Phi \begin{bmatrix} z \\ 1 \end{bmatrix} = 0, \right. \\ \left. \begin{bmatrix} z \\ 1 \end{bmatrix}^H \Psi \begin{bmatrix} z \\ 1 \end{bmatrix} \geq 0 \right\}, \end{cases}$$

into an LMI that depends on (Π, Φ, Ψ) . Here Π defines a quadratic cost, while (Φ, Ψ) define the curve in the complex plane for which the frequency inequality should hold. For completeness, we now proceed to state a generalized version of the KYP lemma, provided in [23], [24], that will be useful to solve (CM).

Lemma 1 (Generalized KYP lemma): Define

$$\Gamma(\lambda) := \begin{cases} \begin{bmatrix} I_n & -\lambda I_n \\ 0 & -I_n \end{bmatrix}, & \lambda \in \mathbb{C}, \\ \begin{bmatrix} 0 & -I_n \end{bmatrix}, & \lambda = \infty. \end{cases} \quad (6)$$

Let $\Pi \in \mathbb{S}_{n+m}$, $M \in \mathbb{C}^{2n \times (n+m)}$, and $\Phi, \Psi \in \mathbb{S}_2$ be given. Denote by $N(\lambda)$ the null space of $\Gamma(\lambda)M$, as a function of λ . Then, the following statements are equivalent:

- 1) (FDI) $N^H(\lambda)\Pi N(\lambda) \prec 0$, $\forall \lambda \in \Lambda(\Phi, \Psi)$.
- 2) (LMI) There exist $\Gamma, S \in \mathbb{S}_n$, such that

$$M^H(\Phi \otimes \Gamma + \Psi \otimes S)M + \Pi \prec 0, \\ S \succ 0.$$

Proof: See [23, Theorem 2]. ■

Remark 3: Variable λ can be either used as the Laplace transform variable $s \in \mathbb{C}$ or as the Z-transform variable $z \in \mathbb{C}$. ■

A. Application to $\mathcal{H}_\infty$ design

Let (A_0, B_0, C_0, D_0) be the state space realization of the $n \times m$ transfer-function matrix $P(z; \theta_0)$, described by (4) and (5) with $p(\theta) = \delta(\theta - \theta_0)$. The traditional $\mathcal{H}_\infty$ design in (DM) can be casted as the following FDI:

$$\begin{aligned} \min_{A_Q, B_Q, C_Q, D_Q} \|P_0\|_\infty &= \min_{A_Q, B_Q, C_Q, D_Q, \gamma} \gamma \\ \text{s.t.} \quad &\begin{bmatrix} P_0(z) \\ I_m \end{bmatrix}^H \begin{bmatrix} I_m & 0 \\ 0 & -\gamma^2 I_m \end{bmatrix} \begin{bmatrix} P_0(z) \\ I_m \end{bmatrix} < 0, \\ &\forall z \in \Lambda \left(0, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right), \end{aligned}$$

since $\|G_0\|_\infty \leq \gamma \iff P_0^H(e^{j\omega})P_0(e^{j\omega}) \preceq \gamma^2 I_m$. By using Lemma 1 with $M = \begin{bmatrix} A_0 & B_0 \\ I_m & 0 \end{bmatrix}$, $\Pi = \begin{bmatrix} C_0 & D_0 \\ 0 & I_m \end{bmatrix}^H \begin{bmatrix} I_m & 0 \\ 0 & \gamma^2 I_m \end{bmatrix} \begin{bmatrix} C_0 & D_0 \\ 0 & I_m \end{bmatrix}$, $N(z) = [A_0 - zI \ B_0]^\top$,

the above FDI is equivalent to the following LMI:

$$\begin{aligned} \min_{A_Q, B_Q, C_Q, D_Q, \gamma, S} \quad & \\ \text{s.t.} \quad & \begin{bmatrix} A_0^H S A_0 - S & A_0^H S B_0 & C_0^H \\ B_0^H S A_0 & B_0^H S B_0 - \gamma^2 I_m & D_0^H \\ C_0 & D_0 & -I_m \end{bmatrix} \prec 0, \\ & S \succ 0. \end{aligned}$$

As it is well known [10], the constraints in the latter problem involve products of the design parameters $(A_Q, B_Q, C_Q, D_Q, S, \gamma)$ when (A_0, B_0, C_0, D_0) are defined as in (4) and (5). In the following section we describe how to overcome this issue for solving (CM) by restricting Q to be an FIR filter.

IV. MINIMAL- $\mathcal{H}_\infty$ -RISK FILTER DESIGN

In this section we present the main result of this work, where we solve (CM) by a sequence of approximations based on N samples from $p(\theta)$ and restricting Q to be an FIR (finite impulse response) filter of length L .

Finding the optimal $n_y \times n_u$ filter Q^* in (CM) is equivalent to finding an optimal state-space realization $\mathcal{S}_Q = (A_Q, B_Q, C_Q, D_Q)$. Let $\mathcal{Q}_L$ be the set of FIR filters of length L , that is

$$\mathcal{Q}_L := \left\{ Q : Q(z) = \sum_{i=0}^L w_i z^{-i} \right\}, \quad (7)$$

where $w_0, w_1, \dots, w_L$ are $n_y \times n_u$ matrices. Then, every $Q \in \mathcal{Q}_L$ has a state space representation of the form

$$\begin{bmatrix} A_Q & B_Q \\ C_Q & D_Q \end{bmatrix} = \left[\begin{array}{c|c} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ I_{n_u(L-1)} & \mathbf{0} \end{bmatrix} & \begin{bmatrix} I_{n_u} \\ \mathbf{0} \end{bmatrix} \\ \hline \begin{bmatrix} w_1 & w_2 & \dots & w_L \end{bmatrix} & w_0 \end{array} \right], \quad (8)$$

where $\mathbf{0}$ are zero matrices of adequate dimensions.

It is well known that $\mathcal{Q}_L$ converges to $\mathcal{H}_2$ as $L \rightarrow \infty$, since the functions $g_i(z) \rightarrow z^{-i}$, $i = 0, 1, \dots$ form a complete orthonormal set in $\mathcal{H}_2$ [25]. This means that (CM) can be well approximated by an equivalent problem where $Q \in \mathcal{Q}_L$, with L sufficiently large. On the other hand, even when the distribution of θ is known to the designer, the explicit

pdf of $J_Q(\theta)$ for every $Q \in \mathcal{Q}_L$ might be hard to obtain. Therefore, we employ a scenario approach by sampling $p(\theta)$. In other words, (CM) can be well approximated by solving the following approximated minimization (AM) problem:

$$(AM) \begin{cases} Q_L^* = \arg \min_{Q \in \mathcal{Q}_L, \mu \in \mathbb{R}} \mu + \frac{\sum_{i=1}^N [J_Q(\theta_i) - \mu]_+}{N(1 - \alpha)}, \\ J_Q(\theta) = \|F(\cdot; \theta_i) - QE(\cdot; \theta_i)\|_2^2. \end{cases}$$

Theorem 1: Let $\{\theta_i\}_{i=1}^N$ denote a set of N independent samples from $\theta \sim p(\theta)$, and let $\mathbf{t} = [t_1 \ t_2 \ \dots \ t_N]^\top \in \mathbb{R}^N$. Let also $\mathbf{w} := [w_0 \ w_1 \ \dots \ w_L] \in \mathbb{R}^{n_y \times (L+1)n_u}$. Then, under Assumption 1, the minimum in (AM) converges to the one of (CM) as $N, L \rightarrow \infty$, whose minimizer, for fixed N and L , is given by $Q_L^*(q) = \sum_{i=0}^L w_i^* q^{-i}$, where $\mathbf{w}^*$ is the solution of Problem 1 (P1). ■

Proof: Since Θ is compact and $\theta \in \mathbb{L}$, it follows, by Assumption 1, that $J_\theta < \infty$ for every $\theta \in \Theta$. In particular, it means that $\sup_{Q \in \mathcal{Q}_L} |\mathbb{E}\{J_Q(\theta)\}| < \infty$ and that $\mathbb{E}\{|[J_Q(\theta) - \mu]_+|\} < \infty$. Thus, by the law of large numbers [26],

$$\frac{\sum_{i=1}^N J_Q(\theta_i) - \mu}{N} \xrightarrow{\text{a.s.}} \int_{\theta \in \Theta} [J_Q(\theta) - \mu]_+ p(\theta) d\theta, \quad (9)$$

which guarantees convergence of the cost functions of (CM) and (AM) for sufficiently large L .

On the other hand, (AM) can be written as

$$\begin{aligned} \min_{Q \in \mathcal{Q}_L, \mu, \mathbf{t}} \quad & \mu + \frac{\mathbf{1}_N^\top \mathbf{t}}{N(1 - \alpha)} \\ \text{s.t.} \quad & t_i \geq J_Q(\theta_i) - \mu \\ & t_i \geq 0, \quad i = 1, \dots, N, \end{aligned}$$

since the minimizer of $\sum_{i=1}^N [f_i(x)]_+$ satisfies either $[f_i(x)]_+ = f_i(x)$ or $[f_i(x)]_+ = 0$, for every $i = 1, 2, \dots, N$, i.e., one of them must be an active constraint. Recall that $J_Q(\theta_i)$ denotes the $\mathcal{H}_\infty$ -norm of P for a fixed Q . Then, $J_Q(\theta_i)$ satisfies $|P_i^H(e^{j\omega}) P_i(e^{j\omega})| \leq J_Q^2(\theta_i) I_m$, $\forall \omega \in [0, \pi]$, where $P_i(z) := F(z; \theta_i) - QE(z; \theta_i)$ has a state-space realization $(A_i, B_i, C_i, D_i) := (A(\theta_i), B(\theta_i), C(\theta_i), D(\theta_i))$, as defined in (4) and (5).

$$(P1) \left\{ \begin{aligned} & \mathbf{w}^* = \arg \min_{\substack{\mathbf{w}, \mu, \mathbf{t} \\ \{\gamma_i\}_{i=1}^N, \{S_i\}_{i=1}^N}} \mu + \frac{1}{N(1 - \alpha)} \mathbf{1}^\top \mathbf{t} \\ & \text{s.t.} \quad \begin{bmatrix} A_i^H S_i A_i - S_i & A_i^H S_i B_i \\ B_i^H S_i A_i & B_i^H S_i B_i - \gamma_i^2 I_m \\ \left[C_H^{(i)} \quad -\mathbf{w} \begin{bmatrix} I_{L n_u} \\ \mathbf{0} \end{bmatrix} \right] C_E^{(i)} & D_H^{(i)} - [0 \quad I_{n_u}] \mathbf{w} D_E^{(i)} \end{bmatrix} \end{bmatrix} \prec 0 \\ & S_i \succ 0 \\ & t_i \geq \gamma_i^2 - \mu, \\ & t_i \geq 0, \quad \forall i = 1, \dots, N, \end{aligned} \right. \quad \left[\begin{array}{c} (C_H^{(i)})^H \\ - \begin{bmatrix} I_{L n_u} & \mathbf{0} \end{bmatrix} \mathbf{w}^H \\ - (C_E^{(i)})^H \begin{bmatrix} \mathbf{0} & I_{n_u} \end{bmatrix} \mathbf{w}^H \\ (D_H^{(i)} - [0 \quad 1] \mathbf{w} D_E^{(i)})^H \end{array} \right] \prec 0 \end{aligned}$$

Fig. 2. Control loop in the application to risk-theoretical control design.

Now, by means of Lemma 1, we have that

$$J_Q^2(\theta_i) = \min_{\mu_i, S_i} \gamma_i^2$$

$$\text{s.t.} \begin{bmatrix} A_i^H S_i A_i - S_i & A_i^H S_i B_i & C_i^H \\ B_i^H S_i A_i & B_i^H S_i B_i - \gamma_i^2 I_m & D_i^H \\ C_i & D_i & -I_m \end{bmatrix} \prec 0, \quad S_i \succ 0.$$

Therefore, the result follows from replacing $J_Q^2(\theta_i)$ by a decision variable γ_i^2 .

Remark 4: When $p(\theta) = \delta(\theta - \theta_0)$ only one sample is enough, i.e., $N = 1$. Then, for $\alpha = 1$, the solution of (P1) is also the solution to the problem in Section III.A. ■

V. APPLICATION TO MISO CONTROL DESIGN

In this section we provide an interesting application to risk-theoretical control design, together with an insightful example showing the effectiveness of the framework proposed along the previous sections.

Consider a strictly proper MISO (multiple-input multiple-output) $1 \times n_G$ plant G and a noise shaping SISO (single-input single-output) block H , being controlled by an $n_G \times 1$ controller $C \in \mathcal{H}_2$, as depicted in Fig. 2. The disturbance v is zero-mean unit-variance white noise, and the transfer function to the output y is denoted by $T_{v \rightarrow y}$. Additionally, r is some reference signal.

Assumption 2: $G(z) = G_0(z)$ is known (at least with sufficient certainty for control purposes) and strictly proper, and $H = H(z, \theta) \in \mathbb{C}$ is parametrized by a random vector $\theta \in \Theta$, Θ compact, such that $H(z, \theta)$ is minimum-phase, biproper and stable for every $\theta \in \Theta$. ■

The contribution of the disturbance v to output y is given by $T_{v \rightarrow y} = (1 - GC)^{-1}H$. The specific problem is to design C so that the risk of $\|T_{v \rightarrow y}\|_\infty$ being large is minimized.

Let us now parametrize the controller as [27]

$$C = (1 - QG)^{-1}\tilde{Q} = Q(I_{n_G} - G\tilde{Q})^{-1}, \quad (10)$$

where $\tilde{Q}$ is an $n_G \times 1$ transfer-function matrix, known as the Youla Parameter. This affine parametrization guarantees that the control loop in Fig. 2 is internally stable if and only if $\tilde{Q} \in \mathcal{Q} := \{\tilde{Q} : \tilde{Q} \text{ is real-rational and stable}\}$ when the controller is given as in (10). This affine parametrization allows us to write $T_{v \rightarrow y} = (1 - G\tilde{Q})H = H - HG\tilde{Q}$. Note that the invertibility of this parametrization is guaranteed only under Assumption 2. Now, due to the invariance of the $\mathcal{H}_\infty$ -norm, it follows that $\|T_{v \rightarrow y}\|_\infty = \|T_{v \rightarrow y}^\top\|_\infty = \|H - H\tilde{Q}^\top G^\top\|_\infty$.

Then, the risk-theoretical control problem we aim to solve is given by

$$\tilde{Q}^* := \min_{\tilde{Q} \in \mathcal{Q}} \text{CVaR}_\alpha(\underbrace{\|H(\cdot; \theta) - \tilde{Q}^\top H(\cdot; \theta)G^\top\|_\infty}_{=: F(\cdot; \theta)}), \quad (11)$$

$$\underbrace{\tilde{Q}^\top}_{=: Q} \underbrace{H(\cdot; \theta)G^\top}_{=: E(\cdot; \theta)}$$

which can be solved efficiently as the semidefinite program provided by Theorem 1, for fixed α , by finding Q^* , and then $\tilde{Q}^* = (Q^*)^\top$.

Remark 5: The extension to the full MIMO case can be addressed by a more complicated argument that involves deriving a state-space representation of $G\tilde{Q}$ in state-space matrices $(A_{G\tilde{Q}}, B_{G\tilde{Q}}, C_{G\tilde{Q}}, D_{G\tilde{Q}})$ such that $(A_{G\tilde{Q}}, B_{G\tilde{Q}})$ does not depend on the parameters defining $\tilde{Q}$. For the sake of brevity, we only present here the MISO case. ■

In what follows, we present a simulated study case that illustrates the properties of CVaR with respect to other measures of risk.

A. An illustrative example

Consider the control loop depicted in Fig. 2, for a plant

$$G(z) = \begin{bmatrix} \frac{z}{(z-2z0.5 \cos(\pi/6)+0.5)^2} & \frac{1}{(z-0.3)} & \frac{-3}{(z-0.8)} \end{bmatrix},$$

and an uncertain noise filter H parametrized by $\theta := [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4]^\top$ as

$$H(z; \theta) = \frac{z^2 + \theta_1 z + \theta_2}{z^2 + \theta_3 z + \theta_4}, \quad (12)$$

where $(\theta_1, \theta_2, \theta_3, \theta_4) \sim \text{Unif}(-0.02148 \pm 0.6, 0.210 \pm 0.6, 0.164 \pm 0.6, -0.0896 \pm 0.6)$. The goal is to design a controller $C_L^* = (1 - (\tilde{Q}_L^\top G)^{-1}\tilde{Q}_L)$ where $\tilde{Q}_L^* = (Q_L^*)^\top$ the L -FIR solution to the following optimization problem:

$$Q_L^* := \arg \min_{Q \in \mathcal{Q}_L} \mathcal{R}\{\|H(\cdot; \theta) - QH(\cdot; \theta)G^\top\|_\infty^2\},$$

where $\mathcal{R}$ is picked among 4 different designs, namely *Nominal*, *Mean* (average/expected value), *Robust* (worst-case) and *CVaR* $_\alpha$ design, with $\alpha = 0.95$. The solution for the CVaR design is approximated (in N) by solving (AM) using Theorem 1. The Nominal design assumes nominal models with $\theta_0 = \mathbb{E}\{\theta\}$, the Mean design minimizes the cost function $\mathbb{E}\{\|H(\cdot; \theta) - QH(\cdot; \theta)G^\top\|_\infty^2\}$, and the Robust design minimizes $\max_{\xi \in \Theta} \|H(\cdot; \xi) - QH(\cdot; \xi)G^\top\|_\infty^2$. All controllers consider $L = 20$ and are obtained by using $N = 1000$ samples from $p(\theta)$, described above, and evaluated using $N_{\text{val}} = 5000$ validation points.

The results for different designs are presented in Fig. 3 where, for each design, the top and the bottom ends of the blue boxes represent the 25th and 75th percentiles of the samples, respectively, the red line inside the box stands for the sample median, and observations beyond 1.5 times the interquartile length are marked as outliers, displayed with red crosses. From Fig. 3 we can firstly notice how bad the nominal design performs in terms of median and outliers compared to the other three designs, showing that it is wiser to consider the information provided in $p(\theta)$. Secondly, we notice that the behavior of the Mean design is slightly better than the CVaR $_{0.95}$ and the Robust design but, as expected,

Fig. 3. Results of the numerical example in terms of performance comparison among 4 different designs. The performance is measured as the $\|(1 - G\tilde{Q})H(\cdot; \theta_i)\|_\infty^2$, where $\{\theta_i\}_{i=1}^{N_{\text{val}}}$ is the set of independent validation samples.

there is no control on the tail distribution, leading to larger outliers. On the other hand, the $\text{CVaR}_{0.95}$ and the Robust design have similar behavior in terms of outliers, presumably due to the fact that α is close to 1. This shows how well $\text{CVaR}_{0.95}$ compares to the other two designs.

VI. CONCLUSIONS

We have presented a framework for $\mathcal{H}_\infty$ -risk minimization of multi-input multi-output LTI systems as an LMI based on the celebrated KYP lemma. The only information the framework requires is to have parametric models for $F(q, \theta)$ and $E(q, \theta)$, where the prior distribution of θ , of compact support, is known. This information can be obtained in the process of modeling the true systems, which commonly provides confidence bounds on the estimations.

We have also presented an interesting application of this framework to *robust* control design when the goal is to minimize the risk of falling into poor tracking error performance. An illustrative example shows how CVaR_α can be employed to balance the behavior of such traditional risk measures as average and as worst-case performance.

REFERENCES

- [1] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Prentice Hall, 1999.
- [2] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall, 1996.
- [3] S. P. Bhattacharyya, H. Chapellat, and L. H. Keel, *Robust Control: The Parametric Approach*. Prentice Hall, 1995.
- [4] B. A. Francis, *A Course in $\mathcal{H}_\infty$ Control Theory*. Springer, 1987.
- [5] F. W. Fairman, *Linear Control Theory: The State Space Approach*. John Wiley & Sons, 1998.
- [6] A. Saberi, A. A. Stoorvogel, and P. Sannuti, *Control of linear systems with regulation and input constraints*. Springer, 2000.
- [7] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. John Wiley and Sons, 1959.
- [8] P. Artzner, F. Delbaen, J. Eber, and D. Heath, "Thinking coherently," *Risk*, pp. 68–71, 1999.

- [9] R. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, pp. 21–42, 2000.
- [10] C. Scherer, P. Gahinet, and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *IEEE Transactions on Automatic Control*, vol. 42, no. 7, pp. 896–911, 1997.
- [11] M. C. de Oliveira, J. C. Geromel, and J. Bernussou, "Extended H_2 and H_∞ norm characterizations and controller parametrizations for discrete-time systems," *International Journal of Control*, vol. 75, no. 9, pp. 666–679, 2002.
- [12] B. P. G. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, "Distributionally robust control of constrained stochastic systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 430–442, Feb 2016.
- [13] M. I. Müller, P. E. Valenzuela, and C. R. Rojas, "Risk-coherent $\mathcal{H}_2$ -optimal disturbance rejection under model uncertainty," *Proceedings of the 20th IFAC World Congress*, vol. 50, no. 1, pp. 15 530 – 15 535, 2017.
- [14] M. I. Müller, J. Milošević, H. Sandberg, and C. R. Rojas, "A risk-theoretical approach to $\mathcal{H}_2$ -optimal control under covert attacks," in *Proceedings of the 57th IEEE Conference on Decision and Control*, To appear, 2018.
- [15] P. Whittle, *Risk-sensitive Optimal Control*. John Wiley & Sons, 1990.
- [16] A. Rantzer, "On the Kalman-Yakubovich-Popov lemma," *Systems & Control Letters*, vol. 28, no. 1, pp. 7–10, 1996.
- [17] C. Acerbi, "A coherent representation of subjective risk aversion," *Journal of Banking and Finance*, vol. 26, pp. 1505–1518, 2002.
- [18] G. A. Holton, *Value-at-Risk: Theory and Practice*. Academic Press, 2014.
- [19] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.
- [20] V. M. Popov, "On absolute stability of non-linear automatic control systems," *Avtomat. i Telemekh.*, vol. 22, no. 8, pp. 961–979, 1961.
- [21] V. A. Yakubovich, "The solution of some matrix inequalities encountered in automatic control theory," *Doklady Akademii Nauk*, vol. 143, no. 6, pp. 1304–1307, 1962.
- [22] R. E. Kalman, "Lyapunov functions for the problem of lur'e in automatic control," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 49, no. 2, pp. 201–205, 1963.
- [23] T. Iwasaki and S. Hara, "Generalized KYP lemma: Unified frequency domain inequalities with design applications," *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 41–59, 2005.
- [24] S. Hara, T. Iwasaki, and D. Shiokata, "Robust pid control using generalized KYP synthesis: direct open-loop shaping in multiple frequency ranges," *IEEE Control Systems Magazine*, vol. 26, no. 1, pp. 80–91, 2006.
- [25] N. Young, *An introduction to Hilbert space*. Cambridge University Press, 1988.
- [26] A. Papoulis, *Probability, Random Processes, and Stochastic Processes*. New York: McGraw-Hill, 1991.
- [27] G. C. Goodwin, S. F. Graebe, and M. E. Salgado, *Control System Design*. Prentice Hall, 2000.