On the Sample Complexity of Uncertain Linear and Bilinear Matrix Inequalities

Mohammadreza Chamanbaz, Fabrizio Dabbene, Roberto Tempo,
Venkatakrishnan Venkataramanan, Qing-Guo Wang

Abstract— In this paper, we consider uncertain linear and bilinear matrix inequalities which depend in a possibly non-linear way on a vector of uncertain parameters. Motivated by recent results in statistical learning, we show that probabilistic guaranteed solutions can be obtained by means of randomized algorithms. In particular, we show that the Vapnik-Chevonenkis dimension (VC-dimension) of the two problems is finite, and we compute upper bounds on it. In turn, these bounds allow us to derive explicitly the sample complexity of the problems. Using these bounds, in the second part of the paper, we derive a sequential scheme, based on a sequence of optimization and validation steps. The algorithm is on the same lines of recent schemes proposed for similar problems, but improves both in terms of complexity and generality.

I. INTRODUCTION

Statistical learning theory is a very effective tool in dealing with various applications, which include neural networks and control systems, see for instance [21]. The main objective of this theory is to extend convergence properties of the empirical mean, which can be computed with a Monte Carlo simulation, from finite families to infinite families of functions. For finite families, the convergence properties can be easily established by means of a repeated application of the so-called Hoeffding inequality, and are related to the well-known law of large numbers, see [16], [21]. On the other hand, for infinite families much deeper technical tools are needed and have been developed in the seminal work of Vapnik and Chervonenkis [19]. In this case, the main issue is to establish *uniform* convergence of empirical means. In particular, this requires to establish whether or not a combinatorial parameter called the Vapnik-Chevonenkis dimension (VC-dimension) is finite, see [18].

Continuing the line of research on statistical learning theory, subsequent contributions by Vidyasagar [20] followed two main research directions: First, to demonstrate that this theory is indeed an effective tool for control of systems affected by uncertainty. Second, to "invert" the bounds provided by Vapnik and Chervonenkis, introducing the concept of *sample complexity*. Roughly speaking, when dealing with control of uncertain systems, the sample complexity provides the number of random samples of the uncertainty that should be drawn to derive a stabilizing controller (or a controller

which attains a given bound on the closed-loop sensitivity function), with sufficiently high probabilistic accuracy and confidence. Since the sample complexity is a function of accuracy, confidence and the VC-dimension, specific bounds on this combinatorial parameter should be derived. In turn, this involves a problem reformulation in terms of Boolean functions, and the evaluation of the number of required polynomial inequalities, their order and the number of design variables.

For various stabilization problems, which include stability of interval matrices and simultaneous stabilization with static output feedback, bounds on the VC-dimension have been derived in [22]. In this paper, we continue this specific line of research, and we compute the VC-dimension for control problems formulated in terms of *uncertain* Linear Matrix Inequalities (LMIs) and Bilinear Matrix Inequalities (BMIs). It is well-known that many robust and optimal control problems can be indeed formulated in these forms, see for instance [4], [12], [13], [17]. However, due to the presence of uncertainty it is often unclear how uncertain LMIs and BMIs can be effectively solved, for example when the uncertainty enters nonlinearly into the control system. In these cases, relaxation techniques are usually introduced, leading to conservative results.

In this paper, we provide new bounds for the VC-dimension for uncertain LMIs and BMIs. These bounds are then combined with results in [2] to establish the sample complexity of uncertain LMIs and BMIs. We remark that the sample complexity is independent from the number of uncertain parameters entering into the LMIs and BMIs, and on their functional relationship. For this reason, the related randomized algorithms run in polynomial-time. However, for relatively small values of the probabilistic accuracy and confidence, the sample complexity turns out to be very large, as usual in the context of Statistical Learning Theory. For this reason, randomized algorithms based on a direct application of these bounds may be of limited use in practice. To alleviate this difficulty, in the second part of the paper we propose a new sequential algorithm. This algorithm has some similarities with other sequential algorithms previously developed for other problems in the area of randomized algorithms for control of uncertain systems, see [7], and in particular [1], [2], [9], [14].

M. Chamanbaz and V. Venkatakrishnan are with Data Storage Institute, Singapore (emails: {Mohammad_C, Venka_V}@dsi.a-star.edu.sg).

R. Tempo and F. Dabbene are with IEIIT-CNR Torino, Italy (emails: {roberto.tempo, fabrizio.dabbene}@polito.it).

M. Chamanbaz and Q-G. Wang are with Department of Electrical and Computer Engineering, National University of Singapore (emails: {Chamanbaz,Elewqg}@nus.edu.sg).

II. PROBLEM FORMULATION

Most robust and optimal control problems can be formulated as linear or bilinear matrix inequality (LMI or BMI).

In the case where problem data involves some uncertain parameters the LMI and BMI problems are in the form of semi-infinite optimization programs, due to the infinite number of constraints involved. We now formally state the uncertain LMI and BMI problems.

Problem 1 (Uncertain LMI Optimization Problem): Find the optimal value of x, if it exists, which solves the optimization problem

$$\underset{x}{\text{minimize}} \quad c_x^T x$$

$$\text{subject to} \quad F_{\text{LMI}}(x, q) \doteq F_0(q) + \sum_{i=1}^{m} x_i F_i(q) \succ 0, \ \forall q \in \mathbb{Q} \tag{1}$$

where $x \in \mathbb{R}^m$ is the vector of optimization variables, $q \in \mathbb{Q} \subset \mathbb{R}^\ell$ is the vector of uncertain parameters bounded in the set $\mathbb{Q}$ and $F_i = F_i^T \in \mathbb{R}^{n \times n}$, $i = 0, \ldots, m$. The inequality $F_{\text{LMI}}(x, q) \succ 0$ means that $F_{\text{LMI}}(x, q)$ is positive definite.

Problem 2 (Uncertain BMI Optimization Problem): Find the optimal values of x and y, if they exist, which solve the optimization problem

$$\underset{x,y}{\text{minimize}} \quad c_x^T x + c_y^T y$$

$$\text{subject to} \quad F_{\text{BMI}}(x, y, q) \doteq F_0(q) + \sum_{i=1}^{m} x_i F_i(q)$$

$$+ \sum_{j=1}^{p} y_j G_j(q) + \sum_{i=1}^{m} \sum_{j=1}^{p} x_i y_j H_{ij}(q) \succ 0, \ \forall q \in \mathbb{Q} \tag{2}$$

where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ are the vectors of optimization variables, $q \in \mathbb{Q} \subset \mathbb{R}^\ell$ is the vector of uncertain parameters and $G_i = G_i^T$, $H_{ij} = H_{ij}^T \in \mathbb{R}^{n \times n}$, $i = 1, \ldots, m$.

In the present paper, we study a probabilistic framework for solving Problems 1 and 2 in which the uncertain parameters are assumed to be random variables. Furthermore, the constraints in (1) and (2) are allowed to be violated for some $q \in \mathbb{Q}$, provided that this violation is sufficiently small. This concept is formally expressed using the notion of "probability of violation".

Definition 1 (Probability of Violation): The probability of violation of θ for the function $g : \Theta \times \mathbb{Q} \to \{0, 1\}$ is defined as:

$$V_g(\theta) \doteq \Pr \{ q \in \mathbb{Q} \ : \ g(\theta, q) = 1 \} \tag{3}$$

where $\theta = [x] \in \Theta \subseteq \mathbb{R}^m$ and

$$g(\theta, q) \doteq \begin{cases} 0 & \text{if } F_{\text{LMI}}(\theta, q) \succ 0 \\ 1 & \text{otherwise} \end{cases} \tag{4}$$

for Problem 1; and similarly, $\theta = [x^T, y^T]^T \in \Theta \subseteq \mathbb{R}^{m+p}$ and

$$g(\theta, q) \doteq \begin{cases} 0 & \text{if } F_{\text{BMI}}(\theta, q) \succ 0 \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

for Problem 2.

The probability (3) is in general very difficult to evaluate due to the complexity of the multiple integrals associated with its computation. Nevertheless, we can "estimate" this

probability using randomization. To this end, we assume that a probability measure is given over the set $\mathbb{Q}$, and extract N independent and identically distributed (i.i.d) samples from the set $\mathbb{Q}$

$$\mathbf{q} = \{q^{(1)}, \ldots, q^{(N)}\} \in \mathbb{Q}^N,$$

based on the given density function, where $\mathbb{Q}^N \doteq \mathbb{Q} \times \mathbb{Q} \times \cdots \times \mathbb{Q}$ (N times). Next, a Monte Carlo approach is employed to obtain the so called "empirical violation".

Definition 2 (Empirical Violation): For given $\theta \in \Theta$ the empirical violation of $g(\theta, q)$ with respect to the multisample $\mathbf{q} = \{q^{(1)}, \ldots, q^{(N)}\}$ is defined as

$$\widehat{V}_g(\theta, \mathbf{q}) \doteq \frac{1}{N} \sum_{i=1}^{N} g(\theta, q^{(i)}). \tag{6}$$

A. Randomized Strategy to Optimization Problems

There are a number of randomized methodologies in the literature which are based on randomization in the uncertainty space, design parameter space or both. For example, in [20] randomization in both uncertainty and design parameter spaces is employed for minimizing the empirical mean. Similarly, a bootstrap learning method and a min-max approach are presented in [14] and [11], respectively. We remark that these papers deal with finite families. In [2] the authors proposed a randomized algorithm for infinite families which is applicable to convex and non-convex problems. Finally, a non-sequential randomized methodology for uncertain convex problems is introduced in [5], [6], [8]. Now we present a randomized strategy for solving Problems 1 and 2. Consider the following non-sequential randomized strategy.

Algorithm 1 A RANDOMIZED STRATEGY FOR UNCERTAIN LMIS/BMIS

- Given the underlying probability density function (pdf) over the uncertainty set $\mathbb{Q}$ and the level parameter $\rho \in [0, 1)$, extract N independent identically distributed samples from $\mathbb{Q}$ based on the underlying pdf

$$\mathbf{q} = \{q^{(1)}, \ldots, q^{(N)}\}.$$

- Find the optimal value, if it exists, of the following optimization problem

$$\underset{\theta \in \Theta}{\text{minimize}} \quad c^T \theta$$

$$\text{subject to} \quad \widehat{V}_g(\theta, \mathbf{q}) \le \rho \tag{7}$$

where $c = [c_x]$ for Problem 1 and $c = [c_x^T, c_y^T]$ for Problem 2.

We remark that introducing the level parameter $\rho > 0$ enables us to handle probabilistic (soft) constraints. The main objective of the present paper is to derive the explicit sample bound N which guarantees the obtained solution having the desired probabilistic behavior. We use statistical learning theory to derive a bound on N.

III. Vapnik-Chervonenkis Theory

In this section, first we give a very brief overview of Vapnik-Chervonenkis theory. The material presented is classical, but a summary is instrumental to our next developments. In particular, we review some bounding inequalities which are later used in the subsequent sections to derive the explicit sample bounds for solving Problems 1 and 2.

Definition 3 (Probability of Two-sided Failure): Given $N, \varepsilon \in (0,1)$ and $g : \Theta \times \mathbb{Q} \to \{0,1\}$, the probability of two-sided failure denoted by $q_g(N,\varepsilon)$ is defined as

$$q_g(N,\varepsilon) \doteq \Pr \left\{ \mathbf{q} \in \mathbb{Q}^N : \sup_{\theta \in \Theta} |V_g(\theta) - \widehat{V}_g(\theta, \mathbf{q})| > \varepsilon \right\}. \tag{8}$$

The probability of two-sided failure determines how close the empirical violation is to the true probability of violation. In other words, if we extract a multisample $\mathbf{q}$ with cardinality N from the uncertainty set $\mathbb{Q}$, we guarantee that the empirical violation (6) is within ε of the true probability of violation (3) for all $q \in \mathbb{Q}$ except for a subset having probability measure at most $q_g(N,\varepsilon)$. The parameter $1 - \varepsilon$ is called "accuracy level".

Let $\mathcal{G}$ denotes the family of functions $\{g(\theta, q) : \theta \in \Theta\}$ where $g : \Theta \times \mathbb{Q} \to \{0,1\}$ is defined in (4) or in (5). The family $\mathcal{G}$ is said to have the property of *uniform convergence of empirical mean* (UCEM) if $q_g(N,\varepsilon) \to 0$ as $N \to \infty$ for any $\varepsilon \in (0,1)$. We remark that if $\mathcal{G}$ includes finite family of functions, it indeed has the UCEM property. However, infinite families do not necessarily enjoy the UCEM property, see [21] for several examples of this type. The formulated problems in the present paper (Problems 1 and 2) belong to the class of infinite family of functions. We define the family $\mathcal{S}_g$ containing all possible sets $S_g \doteq \{q \in \mathbb{Q} : g(\theta, q) = 1\}$, for g varying in $\mathcal{G}$. Now consider a multisample $\mathbf{q} = \{q^{(1)}, \ldots, q^{(N)}\}$ of cardinality N. For the family of functions $\mathcal{G}$, let

$$\mathbb{N}_{\mathcal{G}}(\mathbf{q}) \doteq \mathrm{Card}\left(\mathbf{q} \cap S_g, S_g \in \mathcal{S}_g\right).$$

We say that $\mathcal{S}_g$ shatters $\mathbf{q}$ when $\mathbb{N}_{\mathcal{G}}$ is equal to 2^N. The notion of "shatter coefficient" also known as "growth function" is now defined formally.

Definition 4 (Shatter Coefficient): The shatter coefficient of the family $\mathcal{G}$ denoted by $\mathbb{S}_{\mathcal{G}}(N)$ is defined as

$$\mathbb{S}_{\mathcal{G}}(N) \doteq \max_{\mathbf{q} \in \mathbb{Q}^N} \mathbb{N}_{\mathcal{G}}(\mathbf{q}).$$

A bound on the shatter coefficient can be obtained by Sauer lemma [15], which in turn requires the computation of the VC-dimension, defined next.

Definition 5 (VC-dimension): The VC-dimension of the family of functions $\mathcal{G}$ is defined as the largest integer d for which $\mathbb{S}_{\mathcal{G}}(N) = 2^d$.

The following result, establishes a bound on the probability of two-sided failure in terms of VC-dimension.

Theorem 1: Let d denote the VC-dimension of the family of functions $\mathcal{G}$. Then

$$q_g(N,\varepsilon) \leq 4e^{2\varepsilon} \left(\frac{2eN}{d}\right)^d e^{-N\varepsilon^2} \tag{9}$$

where e is the Euler number.

Proof: See Theorem 4.4 in [18] and Corollary 1 in [2]. ∎

IV. Main Results

In view of Theorem 1, we conclude that families with finite VC-dimension $d < \infty$ enjoy the UCEM property. Hence, it is important i) to show that the collection $\mathcal{G}$ of functions has finite VC-dimension and furthermore, ii) to derive upper bounds on its VC-dimension.

A. Computation of Vapnik-Chervonenkis Dimension

In the following theorem, which is one of the main contributions of this paper, we derive an upper bound on the VC-dimension of the uncertain LMI and BMI in Problems 1 and 2.

Theorem 2: The VC-dimensions of uncertain LMI problem (1) and uncertain BMI problem (2) are upper bounded by $2m \lg(4en^2)$ and $2(m + p) \lg(4en^2)$, respectively, where $\lg(.)$ denotes the logarithm to the base 2.

Proof: See Theorem 2 in [10]. ∎

It is interesting to observe that the VC-dimension of both LMIs and BMIs is linear in the number of design variables. We remark that it is not possible to compute the VC-dimension for the general case of nonlinear matrix inequality (NMI). The extension of the results from LMI to BMI is due to the fact that there is no optimization variable of degree larger than one in the BMI, and this is clearly not the case for NMI. In the next subsection, we derive explicit sample bounds to be used in Algorithm 1 for solving Problems 1 and 2.

B. Sample Complexity Bounds

In this section, we study a number of sample bounds guaranteeing probability of failures to be bounded by a confidence parameter $\delta \in (0,1)$. We remark that there are several results in the literature to derive the explicit sample size N which guarantees the right hand side of (9) to be bounded by the confidence parameter δ. To the best of our knowledge, the least conservative is stated in Corollary 3 in [2]: For given $\varepsilon, \delta \in (0,1)$, the probability of two-sided failure (9) is bounded by δ provided that at least

$$N \geq \frac{1.2}{\varepsilon^2} \left(\ln \frac{4e^{2\varepsilon}}{\delta} + d \ln \frac{12}{\varepsilon^2} \right) \tag{10}$$

samples are drawn, where $d < \infty$ denotes the VC-dimension of the family of functions $\mathcal{G}$. This result is exploited in the next theorem, that provides the explicit sample complexity bounds for the probability of two-sided failure.

Theorem 3: Suppose that $\varepsilon \in (0,1)$ and $\delta \in (0,1)$ are given. Then, the probability of two-sided failure is bounded by δ if at least

$$N_{\mathrm{LMI}} \geq \frac{1.2}{\varepsilon^2} \left(\ln \frac{4e^{2\varepsilon}}{\delta} + 2m \lg(4en^2) \ln \frac{12}{\varepsilon^2} \right) \tag{11}$$

and

$$N_{\mathrm{BMI}} \geq \frac{1.2}{\varepsilon^2} \left(\ln \frac{4e^{2\varepsilon}}{\delta} + 2(m + p) \lg(4en^2) \ln \frac{12}{\varepsilon^2} \right) \tag{12}$$

samples are drawn for the Problems 1 and 2 respectively.

Proof: The statement of Theorem 3 follows immediately by combining (10) and the results of Theorem 2. ∎

A weaker notion than the probability of two-sided failure is the "probability of one-sided constrained failure" introduced in the following definition.

Definition 6 (Probability of One-sided Constrained Failure): Given N, $\varepsilon \in (0,1)$, $\rho \in [0,1)$ and $g : \Theta \times \mathbb{Q} \to \{0,1\}$, the probability of one-sided constrained failure, denoted by $p_g(N, \varepsilon, \rho)$ is defined as

$$p_g(N, \varepsilon, \rho) \doteq \Pr \left\{ \mathbf{q} \in \mathbb{Q}^N : \text{ there exist } \theta \in \Theta \right.$$

$$\left. \text{such that } \widehat{V}_g(\theta, \mathbf{q}) \leq \rho \text{ and } V_g(\theta) - \widehat{V}_g(\theta, \mathbf{q}) > \varepsilon \right\}.$$
(13)

On the same lines of Theorem 3 , sample complexity bounds for the probability of one-sided constrained failure are derived.

Theorem 4: Suppose that $\varepsilon \in (0,1)$ and $\delta \in (0,1)$ are given. Then, the probability of one-sided constrained failure is bounded by δ if at least

$$N_{\text{LMI}} \geq \frac{5(\rho + \varepsilon)}{\varepsilon^2} \left(\ln \frac{4}{\delta} + 2m \lg(4en^2) \ln \frac{40(\rho + \varepsilon)}{\varepsilon^2} \right)$$
(14)

and

$$N_{\text{BMI}} \geq \frac{5(\rho + \varepsilon)}{\varepsilon^2} \left(\ln \frac{4}{\delta} + 2(m+p) \lg(4en^2) \ln \frac{40(\rho + \varepsilon)}{\varepsilon^2} \right)$$
(15)

samples are drawn for the Problems 1 and 2 respectively.

Proof: The results is an immediate consequence of Theorem 7 in [2], which states that, for given $\varepsilon, \delta \in (0,1)$ and $\rho \in [0,1)$, the probability of one-sided constrained failure is bounded by δ provided that at least

$$N \geq \frac{5(\rho + \varepsilon)}{\varepsilon^2} \left(\ln \frac{4}{\delta} + d \ln \frac{40(\rho + \varepsilon)}{\varepsilon^2} \right) \qquad (16)$$

samples are drawn, where $d < \infty$ denotes the VC-dimension of the family of functions $\mathcal{G}$. The statements in Theorem 4 is derived by substituting the results of Theorem 2 into (16). ∎

Note that the sample complexity of Theorem 4 is strictly smaller than that of Theorem 3. In particular, it is clear that the bounds (11) and (12) increase as $\mathcal{O}(\frac{1}{\varepsilon^2} \ln \frac{1}{\varepsilon^2})$, which implies that if the accuracy level ε is chosen to be very small, the sample bounds can be very large, while (14) and (15) grow as $\mathcal{O}(\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon})$.

On the other hand, we note that the bounds (11)-(12) and (14)-(15) can be quite large, as usual in the context of statistical learning, thus leading to computationally expensive optimization problems in (7). For this reason, in the next section we propose a sequential randomized algorithm for bounding the probability of one-sided constrained failure which can alleviate the computational burden of Algorithm 1.

V. SEQUENTIAL RANDOMIZED ALGORITHM

Sequential methods in probabilistic design usually follow an iterative scheme which includes optimization steps to update the design parameters, followed by randomization steps to check the feasibility of the candidate solution [16]. The first step is deterministic, while the second one is probabilistic. Examples of such scheme are probabilistic design methods based on gradient, ellipsoid and cutting plane update rules, see [16] for more details. Motivated by this, in this section we propose a sequential randomized algorithm which mitigates the conservatism of the bounds (14) and (15). This is accomplished by generating a sequence of "design" sample sets $\{q_d^{(1)}, \ldots, q_d^{(N_k)}\}$ with increasing cardinality N_k which are used in (7) for solving the optimization problem. In parallel, "validation" sample sets $\{q_v^{(1)}, \ldots, q_v^{(M_k)}\}$ of cardinality M_k are also generated by the algorithm in order to check if the given candidate solution, obtained from solving (7), satisfies the desired violation probability or not. The proposed scheme is reported in Algorithm 2. For simplicity of notation, we denote the sample bounds derived in Theorem 4 by N_{MI}. The theoretical properties of Algorithm 2 are summarized in the next theorem.

Theorem 5: If at iteration k Algorithm 2 exits with a probabilistic solution $\widehat{\theta}_{N_k}$, then it holds that $V(\widehat{\theta}_{N_k}) \leq \rho + \varepsilon$ with probability no smaller than $1 - \delta$ that is

$$\Pr \left\{ V_g(\widehat{\theta}_{N_k}) \leq \rho + \varepsilon \right\} \geq 1 - \delta.$$

Proof: See Appendix I. ∎

We now make a few comments regarding Algorithm 2. The optimal values of the constants a and α depend on other parameters of the algorithm, such as the termination parameter k_t, the accuracy level ε, and the level parameter ρ. Since the dependence of the sample bound (18) on the parameters a and α is highly nonlinear, this makes it difficult to compute their optimal value analytically. Nevertheless, a suboptimal solution can be found through numerical simulation: suboptimal values of a and α for which the sample bound (18) is minimized for a wide range of probabilistic levels are $a = 3.05$ and $\alpha = 0.9$.

Remark 1 (Related Results): Algorithm 2 follows the general scheme of other sequential algorithms previously developed in the area of randomized algorithms for control of uncertain systems, see [7], and in particular [1], [2], [9], [14]. However, the sample bound M_k in Algorithm 2 is strictly less conservative than the bound derived in [1] because the infinite sum (Riemann Zeta function) is replaced with a finite sum which leads to a considerable improvement in the sample complexity. Another important difference is on how the cardinality of the design sample set N_k appears in the sequential algorithm. In [9, Algorithm 1], the constraints are required to be satisfied for all the samples extracted from the set $\mathbb{Q}$ while, in Algorithm 2, we allow a limited number of samples to violate the constraints in (1) and (2), or their semidefinite versions, in both "design" and "validation" steps. Finally, we note that the sequential randomized algorithm in [9, Algorithm 2] is purely based on

Algorithm 2 A SEQUENTIAL RANDOMIZED ALGORITHM

1) INITIALIZATION

Set the iteration counter to zero ($k = 0$). Choose the desired accuracy $\varepsilon \in (0, 1)$, confidence $\delta \in (0, 1)$ and level $\rho \in [0, 1)$ parameters and the desired number of iterations $k_t > 1$.

2) UPDATE

Set $k = k + 1$ and $N_k = N_{\text{MI}} \frac{k}{k_t}$.

3) DESIGN

- Draw N_k i.i.d samples $\mathbf{q}_d = \{q_d^{(1)} \ldots q_d^{(N_k)}\}$ from the uncertainty set $\mathbb{Q}$ based on the underlying distribution.

- Solve the following optimization problem

$$\underset{\theta \in \Theta}{\text{minimize}} \quad c^T \theta$$
$$\text{subject to} \quad \widehat{V}_g(\theta, \mathbf{q}_d) \leq \rho. \qquad (17)$$

- If the optimization problem (17) is not feasible, the original problem (1) or (2) is not feasible as well. Else, continue to the next step.

4) VALIDATION

- Draw

$$M_k > \frac{\alpha \ln k + \ln\left(\mathcal{S}_{k_t}(\alpha)\right) + \ln \frac{1}{\delta}}{\ln\left(\frac{1}{(\rho + \varepsilon)a^{\rho-1} + a^\rho(1 - (\rho + \varepsilon))}\right)} \qquad (18)$$

i.i.d samples $\mathbf{q}_v = \{q_v^{(1)} \ldots q_v^{(M_k)}\}$ from the uncertainty set $\mathbb{Q}$ based on the underlying distribution. In (18), the parameters $a \geq 1$ and $\alpha > 0$ are real and $\mathcal{S}_{k_t}(\alpha)$ is a finite hyperharmonic series also known as p-series that is

$$\mathcal{S}_{k_t}(\alpha) = \sum_{k=1}^{k_t} \frac{1}{k^\alpha}.$$

- If

$$\widehat{V}_g(\widehat{\theta}_{N_k}, \mathbf{q}_v) \leq \rho$$

then, $\widehat{\theta}_{N_k}$ is a probabilistic solution and Exit. Else, goto step (2).

---

additive and multiplicative Chernoff inequalities and hence may provide larger sample complexity than (18).

Remark 2 (Zero Level Case): Note that the sample bound (18) in the case when the level parameter ρ is zero reduces to the following bound

$$M_k > \frac{\alpha \ln k + \ln\left(\mathcal{S}_{k_t}(\alpha)\right) + \ln \frac{1}{\delta}}{\ln\left(\frac{1}{1-\varepsilon}\right)}.$$

In this case, it can be shown that the optimal value of α is $\alpha = 0.1$.

The termination parameter k_t defines the maximum number of iterations of the algorithm which can be chosen by the user. For problems in which the bound[1] N_{MI} in Algorithm 2 is large, larger values of k_t may be used. In this way, the sequence of sample bounds N_k would start from a reasonably small number and would not increase dramatically with the iteration counter k.

VI. CONCLUSIONS

In this paper, we computed explicit bounds on the Vapnik-Chervonenkis dimension (VC-dimension) of two problems frequently arising in robust control, namely the solution of uncertain LMIs and BMIs. In both cases, we have shown that the VC-dimension is linear in the number of design variables. These bounds are then used in a sequential randomized algorithm that can be efficiently applied to obtain probabilistic optimal solutions to uncertain LMI/BMI. Since the sample complexity is independent of the number of uncertain parameters, the proposed algorithm runs in polynomial time.

APPENDIX I
PROOF OF THEOREM 5

The following lemma is instrumental to prove the result of Theorem 5.

Lemma 1: At the iteration k of Algorithm 2, if a candidate solution $\widehat{\theta}_{N_k}$ is declared as feasible by the "validation" step then, with probability no larger than $\left((\rho + \varepsilon)a^{\rho-1} + a^\rho\left(1 - (\rho + \varepsilon)\right)\right)^{M_k}$ it holds that $V_g(\widehat{\theta}_{N_k}) > \rho + \varepsilon$.

Proof: The objective is to bound the probability of obtaining $\widehat{\theta}_{N_k}$ which satisfies $\widehat{V}_g(\widehat{\theta}_{N_k}, \mathbf{q}_v) \leq \rho$ and $V_g(\widehat{\theta}_{N_k}) > \rho + \varepsilon$. In the following chain of inequalities, we bound this probability using a binomial distribution

$$\Pr\left\{\widehat{V}_g(\widehat{\theta}_{N_k}, \mathbf{q}_v) \leq \rho \text{ and } V_g(\widehat{\theta}_{N_k}) > \rho + \varepsilon\right\} =$$

$$\Pr\left\{\sum_{i=1}^{M_k} g(\widehat{\theta}_{N_k}, q) \leq \rho M_k \text{ and } V_g(\widehat{\theta}_{N_k}) > \rho + \varepsilon\right\} \leq$$

$$\Pr\left\{\sum_{i=1}^{M_k} g(\widehat{\theta}_{N_k}, q) \leq \rho M_k \text{ and } V_g(\widehat{\theta}_{N_k}) = \rho + \varepsilon\right\} =$$

$$\sum_{i=1}^{\rho M_k} \binom{M_k}{i} (\rho + \varepsilon)^i \left(1 - (\rho + \varepsilon)\right)^{M_k - i} \leq$$

$$a^{\rho M_k} \left(\frac{\rho + \varepsilon}{a} + 1 - (\rho + \varepsilon)\right)^{M_k}.$$

We note that the last inequality results from Lemma 1 in [3]. ∎

---

[1] Obtained respectively by (14) and (15) for LMIs and BMIs.

We begin the proof of Theorem 5 by introducing the following events

$\text{Iter}_k \doteq \{\text{the } k\text{th outer iteration is reached}\}$,

$\text{Feas}_k \doteq \{\widehat{\theta}_{N_k} \text{ is declared as feasible in the "validation" step}\}$,

$\text{Bad}_k \doteq \{V_g(\widehat{\theta}_{N_k}) > \rho + \varepsilon\}$,

$\text{ExitBad}_k \doteq \{\text{Algorithm 2 exits at iteration } k \cap \text{Bad}_k\}$,

$\text{ExitBad} \doteq \{\text{Algorithm 2 exits at some unspecified iteration } k \cap \text{Bad}_k\}$.

The goal is to bound the probability of the event "ExitBad". Since $\text{ExitBad}_i \cap \text{ExitBad}_j = \emptyset$ for $i \neq j$, the probability of the event "ExitBad" can be reformulated in terms of the event 'ExitBad$_k$" as

$$\begin{aligned}
\Pr\{\text{ExitBad}\} &= \Pr\{\text{ExitBad}_1 \cup \text{ExitBad}_2 \cup \ldots \cup \text{ExitBad}_{k_t}\} \\
&= \Pr\{\text{ExitBad}_1\} + \Pr\{\text{ExitBad}_2\} + \ldots + \\
&\quad \Pr\{\text{ExitBad}_{k_t}\}.
\end{aligned} \quad (19)$$

From the definition of the event "ExitBad$_k$" and by considering the point that to exit at iteration k, Algorithm 2 needs to reach kth iteration and declares $\widehat{\theta}_{N_k}$ as feasible, we arrive at

$$\begin{aligned}
\Pr\{\text{ExitBad}_k\} &= \Pr\{\text{Feas}_k \cap \text{Bad}_k \cap \text{Iter}_k\} \\
&= \Pr\{\text{Feas}_k \cap \text{Bad}_k \mid \text{Iter}_k\} \Pr\{\text{Iter}_k\} \\
&\leq \Pr\{\text{Feas}_k \cap \text{Bad}_k \mid \text{Iter}_k\} \\
&= \Pr\{\text{Feas}_k \mid \text{Bad}_k \cap \text{Iter}_k\} \Pr\{\text{Bad}_k \mid \text{Iter}_k\} \\
&\leq \Pr\{\text{Feas}_k \mid \text{Bad}_k \cap \text{Iter}_k\}. \quad (20)
\end{aligned}$$

Using the result of Lemma 1, we bound the right hand side of (20)

$$\Pr\{\text{Feas}_k \mid \text{Bad}_k \cap \text{Iter}_k\} <$$
$$\left((\rho + \varepsilon)a^{\rho-1} + a^{\rho}\left(1 - (\rho+\varepsilon)\right)\right)^{M_k}. \quad (21)$$

Combining (19) and (21) results in

$$\Pr\{\text{ExitBad}\} < \sum_{k=1}^{k_t} \left((\rho + \varepsilon)a^{\rho-1} + a^{\rho}\left(1 - (\rho+\varepsilon)\right)\right)^{M_k}. \quad (22)$$

The summation in (22) can be made arbitrary small by an appropriate choice of M_k. By choosing

$$\left((\rho + \varepsilon)a^{\rho-1} + a^{\rho}\left(1 - (\rho+\varepsilon)\right)\right)^{M_k} = \frac{1}{k^{\alpha}}\frac{1}{\mathcal{S}_{k_t}(\alpha)}\delta \quad (23)$$

where $\delta \in (0,1)$ is a (small) desired probability level, we have

$$\begin{aligned}
\Pr\{\text{ExitBad}\} &< \sum_{k=1}^{k_t} \frac{1}{k^{\alpha}}\frac{1}{\mathcal{S}_{k_t}(\alpha)}\delta \\
&= \frac{1}{\mathcal{S}_{k_t}(\alpha)}\delta \sum_{k=1}^{k_t} \frac{1}{k^{\alpha}} \\
&= \frac{1}{\mathcal{S}_{k_t}(\alpha)}\delta \mathcal{S}_{k_t}(\alpha) = \delta.
\end{aligned}$$

Therefore, the appropriate choice of M_k which guarantees $\Pr\{\text{ExitBad}\} < \delta$ can be computed by solving (23) for M_k which results in the bound (18).

REFERENCES

[1] T. Alamo, A. Luque, D.R. Ramirez, and R. Tempo. Randomized control design through probabilistic validation. In *Proc. American Control Conference (ACC)*, pages 839–844, 2012.

[2] T. Alamo, R. Tempo, and E.F. Camacho. Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems. *IEEE Transactions on Automatic Control*, 54:2545–2559, 2009.

[3] T. Alamo, R. Tempo, and A. Luque. On the sample complexity of probabilistic analysis and design methods. In J. Willems, S. Hara, Y. Ohta, and H. Fujioka, editors, *Perspectives in Mathematical System Theory, Control, and Signal Processing*, volume 398 of *Lecture Notes in Control and Information Sciences*, pages 39–50. Springer Berlin / Heidelberg, 2010.

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

[5] G.C. Calafiore and M.C. Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102:25–46, 2004.

[6] G.C. Calafiore and M.C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51:742–753, 2006.

[7] G.C Calafiore, F. Dabbene, and R. Tempo. Research on probabilistic methods for control system design. *Automatica*, 47:1279–1293, 2011.

[8] M.C. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM J. on Optimization*, 19:1211–1230, 2008.

[9] M. Chamanbaz, F. Dabbene, R. Tempo, V. Venkataramanan, and Q-G. Wang. Sequential randomized algorithms for convex optimization in the presence of uncertainty. *IEEE Transactions on Automatic Control*, 2013, submitted. See also arXiv:1304.2222 [cs.SY].

[10] M. Chamanbaz, F. Dabbene, R. Tempo, V. Venkataramanan, and Q-G. Wang. A statistical learning theory approach for uncertain linear and bilinear matrix inequalities. *Automatica*, 2013, submitted. See also http://arxiv.org/abs/1305.4952.

[11] Y. Fujisaki and Y. Kozawa. Probabilistic robust controller design: Probable near minimax value and randomized algorithms. In G.C. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design under Uncertainty*, pages 317–329. Springer-Verlag, London, 2006.

[12] K.C. Goh, M.G. Safonov, and G.P. Papavassilopoulos. Global optimization for the biaffine matrix inequality problem. *Journal of Global Optimization*, 7:365–380, 1995.

[13] S. Kanev, C. Scherer, M. Verhaegen, and B. De Schutter. Robust output-feedback controller design via local BMI optimization. *Automatica*, 40:1115–1127, 2004.

[14] V. Koltchinskii, C.T. Abdallah, M. Ariola, P. Dorato, and D. Panchenko. Improved sample complexity estimates for statistical learning control of uncertain systems. *IEEE Transactions on Automatic Control*, 45:2383–2388, 2000.

[15] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13:145–147, 1972.

[16] R. Tempo, G.C. Calafiore, and F. Dabbene. *Randomized Algorithms for Analysis and Control of Uncertain Systems: With Applications*. Springer, 2nd edition, 2013.

[17] J. G. VanAntwerp and R. D. Braatz. A tutorial on linear and bilinear matrix inequalities. *Journal of Process Control*, 10:363–385, 2000.

[18] V.N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1 edition, 1998.

[19] V.N. Vapnik and A.Y. Chervonenkis. On the uniform convergence of relative frequencies to their probabilities. *Theory of Probability and Its Applications*, 16:264–280, 1971.

[20] M. Vidyasagar. Randomized algorithms for robust controller synthesis using statistical learning theory. *Automatica*, 37:1515–1528, 2001.

[21] M. Vidyasagar. *Learning and Generalization: With Applications to Neural Networks*. Springer, 2nd edition, 2002.

[22] M. Vidyasagar and V. Blondel. Probabilistic solutions to some NP-hard matrix problems. *Automatica*, 37:1397–1405, 2001.