On learning guaranted bounds for black-box/dark data switched systems ;)

Ayça Balkan, Raphael Jungers, Joris Kenanian, Paulo Tabuada

Abstract— **To do - Ayça?**

I. INTRODUCTION

As the computational power of the computers has increased, so is the complexity of dynamical systems we can model with them. Today, the industrial dynamical systems does not only consist of simple differential or difference equations; they are multimodal, hybrid, and often contain variety of subcomponents such as lookup tables, delay differential equations, actuation delays and thermodynamic models. The current modeling paradigm is also highly distributed, in the sense that, some of the model subcomponents are developed by different parties and therefore their internal structure is partially, if not completely unknown to the end user. Hence, it is often virtually impossible to obtain simple analytical formulas for today's industrial scale dynamical system models. However, thanks to the advances in our computational power again, in spite of this increased complexity, for a notable subset of these models we can still efficiently perform simulations. Therefore, it is a natural question to ask whether we can provide formal analyses about certain properties of these systems based solely on the collected simulation data. In this paper, we focus on the analysis of one of the most important properties of dynamical systems in the context of control theory: stability.

More formally, given a dynamical system as in:

$$x_{k+1} = f(k, x_k), \tag{1}$$

where, $x_k \in \mathbb{R}^n$, k is index of time. Let $y_k := x_{k+1}$ We ask the following question: *given N input-output pairs, (x_1, y_1), (x_2, y_2), ..., (x_N, y_N) such that $y_k = f(k, x_k)$, what can we say about the stability of the system* (**??**)*?* The answer is immediate when (**??**) is a linear time-invariant system, since we can simply identify the system by *n* linearly independent output traces. In this paper, we seek the answer to this question for switched linear systems for which the problem immediately becomes nontrivial. A switched linear systems is in the form:

$$x_{k+1} = A_{\sigma(k)} x_k, \tag{2}$$

where, $\sigma : \mathbb{N} \to \{1, 2, \ldots, m\}$ is the switching sequence and $A_{\sigma(k)} \in \mathcal{M}$, for all σ and k. Aside from their theoretical value, switched systems model the behavior of dynamical systems in the presence of known or unknown varying parameters. These parameters can model internal properties of the dynamical system such as uncertainties, look-up tables, values in a discrete register as well as exogenous inputs provided by a controller in a closed-loop control system.

Assessing the stability of nonlinear systems by leveraging simulations has been an active area of research in the recent years. Simulation data has been used in both construction and verification of Lyapunov functions. Topcu et.al. [**?**] and Kapinski et.al. [**?**] construct Lyapunov function candidates using the simulation traces, however to be able to formally verify the constructed Lyapunov function, they require the knowledge of the full dynamics. In [**?**] and [**?**] Bobiti and Lazar address this and provide sampling based probabilistic and deterministic guarantees of a given Lyapunov function candidate. The presented method requires the knowledge of how fast the output of the system can change as the initial condition changes, and moreover the number of required samples increases exponentially in the dimension of the state, *n*.

The stability of switched systems closely relates closely to the *joint spectral radius* (JSR) of the matrices appearing in (2). Under certain conditions deciding stability amounts to deciding whether JSR less than one or not. There has been a lot of work on developing algorithms to approximate this quantity, when the matrices appearing in (2) are known. Therefore, our work is also connected to the identification of switched systems, since once the system (2) is identified one can then apply these well-established results. However, there are two main reasons behind our quest to directly work on input-output pairs and bypassing the identification phase: (1) Even when $\mathcal{M}$ is known, approximating the JSR is NP-hard [**?**]. (2) Identifying the set $\mathcal{M}$ is also NP-hard. Therefore, the existing identification techniques can identify $\mathcal{M}$ up to an approximation error. As a result, how to relate this identification error to an error on the stability of (2) is still nontrivial.

In this paper, we present an algorithm to approximate the JSR of a switched linear system from *N* input-output pairs. This algorithm provides an upper bound on the JSR with a user-defined confidence level. As the number of samples increases, this bound gets tight. Moreover, we characterize with a closed form expression what the exact trade-off between the tightness of this bound and the number of samples is. In order to understand the quality of our technique, the algorithm also provides a deterministic lower-bound.

The organization of the paper is follows: Section introduces our notation and definitions from the switched linear system literature to present our results, Section formalizes the problem definition, Section provides an algorithm to compute a deterministic lower bound and

II. PRELIMINARIES AND PROBLEM DEFINITION

We consider the usual Hilbert finite normed vector space $(\mathbb{R}^n, \ell_2)$, $n \in \mathbb{N}_{>0}$, ℓ_2 the classical euclidean norm. We denote a unit ball in $\mathbb{R}^n$ with B and unit sphere in $\mathbb{R}^n$ of radius r as S. We only denote the radius r explicitly as in B_r and S_r, when r is different than 1. We denote the set of real symmetric matrices of size n by $\mathbb{S}^n$, and the set of linear functions in $\mathbb{R}^n$ by $\mathscr{L}(\mathbb{R}^n)$. We denote the ellipsoid described by the matrix $P \in \mathbb{S}^n$ as E_P. We denote the homothety of ratio λ by $\mathscr{H}_\lambda$.

We denote by M the set $\{1, 2, \ldots, m\}$, $(m \in \mathbb{N}_{>0})$ and consider a switched linear system of modes $\mathscr{M} = \{A_1, A_2, \ldots, A_m\} \subset \mathbb{R}^{n \times n}$ indexed by M. We observe the system:

$$x_{k+1} = A_{\sigma(k)} x_k, \tag{3}$$

where, $x_k \in \mathbb{R}^n$, k is index of time and $\sigma : \mathbb{N} \to M$ is the switching sequence. Let $y_k := x_{k+1}$. We assume that we only know the number of modes m, the input x_k and the output y_k at each time k. We ignore what are the matrices and the index of the matrix which is applied at every time. We consider the following problem.

Given N input-output-matrix pairs, (x_1, y_1), (x_2, y_2), ..., (x_N, y_N) such that

$$y_k = A_{\sigma(k)} x_k, \tag{4}$$

for some $\sigma(k)$, (but not necessarily $x_{k+1} = y_k$) how can we decide if the switching system is stable?

To make our reasoning clearer, we introduce the *Lyapunov exponent* of the system, which is a numerical quantity describing its stability.

Definition Given a dynamical system as in (3) its *Lyapunov exponent* is given by

$$\rho = \inf \{r : \forall x_0, \exists C \in \mathbb{R}^+ : \quad x(0) = x_0 \Rightarrow x(t) \leq Cr^t\}.$$

Under certain conditions [RJ: to be epxplicited], deciding stability amounts to decide whether $\rho < 1$. In order to understand the quality of our techniques, we will actually try to prove lower and upper bounds on ρ.
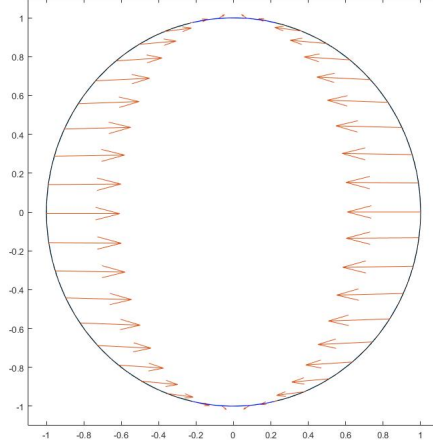
A. Examples

Many observations, coherent with the existence of a quadratic Lyapunov function, do not mean much, even for linear systems. We can see for the simple linear system

$$\dot{x}(t) = \begin{bmatrix} -2 & 0 \\ 0 & 0.1 \end{bmatrix} x(t)$$

that by sampling on the unit circle, as long as we do not pick points from the blue areas, we will observe a contraction of the points. Nevertheless, all the trajectories with initial condition $x \in \mathbb{R}^2 - \{(\lambda, 0), \lambda \in \mathbb{R}\}$ will eventually diverge.

B. Assumptions

It is clear that one has to pose assumptions on the dynamical system studied in order to be able to infer stability properties from the observations of samplings of its dynamics: one can easily build nonlinear systems such that the Euclidean norm of most points is dramatically decreased by one iteration of

the system, but however the system is unstable. Even more, as we have seen in Example **??**, there are linear systems that exhibit this property. One goal of this work is to understand what are the important properties that a system should enjoy, and what type of guarantees they could imply in terms of stability. We list here a few properties that are going to be studied in this work.

Property 2.1: The dynamical system is a polynomial of bounded degree d.

Property 2.2: The dynamical system is a homogeneous function of the statespace:

$$x_0' = \lambda x_0 \Rightarrow x'(t) = \lambda x(t).$$

Property 2.3: The dynamical system is a switching linear system

$$x_{k+1} = A_{\sigma(k)} x_k, \tag{5}$$

where, $x_k \in \mathbb{R}^n$, k is index of time and $\sigma : \mathbb{N} \to \{1, 2, \ldots, m\}$ is the switching sequence. This is the first example we'll study. In this case the Lyapunov exponent is known as the Joint Spectral Radius of the set of matrices, which can be alternatively defined as follows:

Definition [?] Given a set of matrices $\mathscr{M} \subset \mathbb{R}^{n \times n}$, its *joint spectral radius* (JSR) is given by

$$\rho(\mathscr{M}) = \lim_{t \to \infty} \max_{i_1, \ldots, i_t} \{||A_{i_1} \ldots A_{i_t}||^{1/t} : A_i \in \mathscr{M}\}.$$

Remark 2.1: Note that one can scale the problem because the JSR is homogeneous:

$$\rho(\mathscr{M}/\lambda) = \rho(\mathscr{M})/\lambda \, \forall \lambda > 0,$$

and $\mathscr{M}/\lambda$ can be studied by studying the scaled inputs

$$(x_t, y_t/\lambda, \sigma(t)).$$

Property 2.4: The function f is *convex-preserving*, meaning that for any set of points $X = \{x_i\}$, one has

$$f(\text{convhull } X) \subset \text{convhull } \{f(x_i)\}.$$

We do not know until now whether there are many functions that have the above property, or even whether there are non-linear functions that have it. It turns out that switching linear

systems have this property. There is a similar property, that seems to imply (something similar to) the previous one:

Property 2.5: [**?**]Let K be a (convex solid pointed) cone. The function f is *K-convex*, meaning that

- it is *K-monotone*: for any $x, y \in K$,

$$x \preceq_K y \Rightarrow f(x) \preceq_K f(y)$$

- for any $x, y \in \mathbb{R}^n$, and $\lambda \in \mathbb{R}$, $f(\lambda x + (1-\lambda)y) \preceq_K \lambda f(x) + (1-\lambda)f(y)$.

III. A LOWER BOUND

For the lower bound, we will leverage the following theorem from the switching system literature.

Theorem 3.1: [**?**, Theorem 2.11] For any bounded set of matrices such that $\rho(\mathcal{M}) < 1/\sqrt(n)$, there exists a Common Quadratic Lyapunov Function (CQLF) for $\mathcal{M}$, that is, a $P \succeq 0$ such that

$$\forall A \in \mathcal{M}, A^T P A \preceq P.$$

[RJ: We can probably generalize this result. What is needed is a Converse Convex Lyapunov Theorem: stability implies the existence of a convex (or, quasiconvex, it should suffice) Lyapunov function. Maybe we should use Mircea's result [?]. This would be nice per se, I think.]

The existence of a CQLF for our (potentially scaled) blackbox system is certainly something we can check: after collecting N observations, one can solve the following optimization problem efficiently.

$$
\begin{aligned}
\min \quad & \lambda \\
s.t. \quad & \forall 1 \le i \le N, \ (y_i^T P y_i)/\lambda^2 \le x_i^T P x_i \quad (6) \\
& P \succeq 0.
\end{aligned}
$$

Indeed, when λ is fixed, the problem is a set of LMIs, and λ can be optimized by bisection.

Theorem 3.2: Let λ^* be the minimum λ such that (6) above has a solution. If $\lambda^* < \infty$, one has

$$\rho(\mathcal{M}) \ge \lambda^*/\sqrt(n).$$

Proof: Just apply Theorem 3.1 to $\mathcal{M}/(\lambda^* - \varepsilon)$, for any $\varepsilon > 0$. Since this latter set has no CQLF, we obtain that $\rho(\mathcal{M})/\lambda^* \ge 1/\sqrt{n}$. ∎

This result could be improved in several ways: first, it is classical in the JSR literature to replace the CQLF with a SOS polynomial of degree $2d$, $d > 1$. this narrows the $1/\sqrt(n)$ accuracy factor, up to one when $d \to \infty$. [RJ: this is not 100 % clear but interesting: if we take a too large degree for a fixed number of points, we will always find a very small lower bound! Nevertheless I hope that it will be easy to understand how increasing the degree can improve the accuracy, and anyway there are interesting things there.]

IV. AN UPPER BOUND

[RJ: it is not clear how this upper bound can be useful for switching systems, because we need to know the modes here, and then the system can be easily identified, but maybe for general f it is? Or, maybe it's possible to reuse the same type of ideas without knowing the nodes?] It is classical to try to upper bound the JSR by constructing a polytope norm [**?**], defined by a set of points $X = \{x_i : i = 1, \ldots, N\}$ satisfying the following property: if

$$(A_i/\lambda)\text{convhull } (X \cup -X) \subset \text{convhull } (X \cup -X), \quad \forall A_i \in \mathcal{M}$$

then,

$$\rho(\{A_1, \ldots, A_m\}) \le \lambda.$$

This can easily be checked by verifying that

$$\forall x \in X, \forall A \in \mathcal{M}, Ax/\lambda \in \text{convhull } (X \cup -X),$$

that is, by verifying the property only for the vertices of the polytope.

Now, in our situation, we would like to address situations where we do not command what matrix is applied, and/or which point in the state space is selected for the next triple $(x_t, y_t, \sigma(t))$. It turns out that this slight difference with the 'classical JSR computation' can be accounted for:

Theorem 4.1: Suppose that one has m sets of observations

$$S_j = \{(x_{j,1}, y_{j,1}), \ldots, (x_{j,t_j}, y_{j,t_j})\}$$

as in (4), and denote

$$X_j = \{x_{j,t} : t = 1, \ldots, t_j\},$$

then the solution of the following optimization problem

$$
\begin{aligned}
\min \quad & \lambda \quad \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (7)\\
s.t. \quad & \forall i, t, \quad y_{i,t}/\lambda \in \bigcap_{1 \le j \le m} \text{convhull } (X_j \cup -X_j)
\end{aligned}
$$

is a valid upper bound on $\rho(\mathcal{M})$.

Proof: Let us denote $S = \bigcap_{1 \le j \le m} \text{convhull } (X_j \cup -X_j)$. Fix any feasible λ for (7). Since for all $j: 1 \le j \le m$

$$S \subset \text{convhull } (X_j \cup -X_j)$$

and

$$(A_j/\lambda) \text{ convhull } (X_j \cup -X_j) \subset S,$$

we have that

$$(A_j/\lambda)S \subset S \quad \forall 1 \le j \le m,$$

and this proves that $\rho \le \lambda$. ∎

It is not difficult to see that the above bound could be improved further: indeed, it relies on sets of points X_i, for which we know the image by the matrix A_i. However, if we know that $A_i x = y$, then, we know of course that, for any real number λ, $A_i(x/\lambda) = y/\lambda$. In other words, we can scale the couples $(x_{j,i}, y_{j,i})$ above by any factor λ, and this would still give a valid upper bound. Thus, one could try to optimize

over this scaling factors in order to get the best possible upper bound. This gives the following optimization problem:

Theorem 4.2: Suppose that one has m sets of observations

$$S_j = \{(x_{j,1}, y_{j,1}), \ldots, (x_{j,t_j}, y_{j,t_j})\}$$

as in (4), and denote

$$X_j = \{x_{j,t} : t = 1, \ldots, t_j\},$$

then the solution of the following optimization problem

$$
\begin{aligned}
\min_{\lambda, \lambda_{j,t}} \quad & \lambda \\
\text{s.t.} \quad & y_{j,t}/(\lambda \lambda_{j,t}) \in \bigcap_{1 \leq j \leq m} \text{convhull}\left(\{x_{j,t}/\lambda_{j,t}, -x_{j,t}/\lambda_{j,t}\}\right).
\end{aligned}
$$

is a valid upper bound on $\rho(\mathcal{M})$.

A. A heuristic to improve the upper bound

The above program (8) aims at finding scaling factors of the couples $(x_{j,i}, y_{j,i})$ so as to minimize the upper bound. Intuitively, there is an optimal scaling factor which can be defined as follows: let us denote $||\cdot||_*$ the *extremal norm* of the set of matrices ; that is, the norm satisfying

$$\forall x \in \mathbb{R}^n \quad \max_{A_i \in \mathcal{M}} ||A_i x||_* = \rho ||A_i||_*. \quad (9)$$

Such a norm always exists for any bounded set of matrices $\mathcal{M}$, provided that the matrices do not share a common invariant subspace [**?**, Section 2.1.2] (if they do, this does not change our discourse, since there exists in any case a norm satisfying (9), with ρ replaced by $\rho + \varepsilon$, for any $\varepsilon > 0$).

Suppose that every set (X_j) is the complete unit ball $B_* = \{x : ||x||_* = 1\}$. Then, by the definition (9) of $||\cdot||_*$, we have that

$$y_{j,t}/\rho \in \bigcap_{1 \leq j \leq m} \text{convhull}\left(X_j \cup -(X_j)\right),$$

because convhull $(X_j) \cup -(X_j) = B_*$, and by (7) one has a perfect upper bound on ρ. This suggests to scale the $x_{j,t}$ with an approximation of the extremal norm. Now, remark that Equation (6) is precisely aiming at approximating the Definition (9) of an extremal norm. Of course, it is only an approximation, because it only takes into account the finite number of points available, and restricts the possible norms to ellipsoidal ones. However, if no other way is available for optimizing (7), it is hopeful that this approximation could work well. Summarizing, we propose to fix

$$\lambda_{j,t} = \sqrt{x_{j,t}^T P x_{j,t}}, \quad (10)$$

where P is the solution of (6) as a heuristic solution to (8). [RJ: again we might have to trash all the above because knowing the modes allows us to identify the matrices, but it seems that there are good ideas to use for more general models...]

V. AN UPPER BOUND WITH A LEVEL OF CONFIDENCE.

In this section we suppose that we do not observe the modes, but only the pairs (x_k, y_k). Our goal here is to improve the upper bound by paying a little bit of uncertainty. Also, it allows us to provide a bound even if we do not know the modes. The idea is that if we have a polytope that is contracted for all our observed pairs (x, y), it is not enough to imply stability, because we observed only some possible behaviours of the dynamics. However, suppose that we observed that the λ−contraction of the polytope $\mathscr{P}$ is satisfied for many values of x, then, it is tempting to extrapolate our observation, and claim that (noting $y = f(x)$ as always),

$$\forall x \in \mathbb{R}^n \text{ s.t. } Cx \leq b \Rightarrow \quad Cy/\lambda \leq b, \quad (11)$$

where C, b is the canonical representation of the polytope $\mathscr{P}$.

Similarly, if we observed the λ−contraction for the quadratic Lyapunov function $x^\top P x$, it is tempting to generalize to the whole state space:

$$\forall x \in \mathbb{R}^n, (y^T P y)/\lambda^2 \leq x^T P x. \quad (12)$$

Both equations imply that $\rho \leq \lambda$.

Unfortunately, the true observations only imply that these inequalities are valid for a few points x_i, not for the whole statespace. However we will use classical probability results to bridge the gap between our observations, and a stability guarantee for the complete statespace. Observe that a priori this gap is not negligible, and for instance in Example **??** one can see that there exists a quadratic function which satisfies the Lyapunov equation for almost all the points, but yet, almost all the trajectories are unstable. We recall a classical result on optimization with a sampled subset of constraints:

Theorem 5.1: [**?**], [**?**] For any convex program, with a measure on the constraints, in any fixed dimension n, and for any $0 < \varepsilon, \beta < 1$, there exists a number $N(\varepsilon, \beta)$ such that, if one samples N constraints, then, with probability larger than β, the solution of the optimization program with these N constraints satisfies a subset of the full constraints set of measure at least $1 - \varepsilon$.

We first consider a simpler version of our theorem, which corresponds to the case $n = 0$ above, and where the polytope is fixed a priori. This is clearly conservative but is there for the sake of clarity. [RJ: We are trying to prove something as clean as possible... probably this simpler version will be removed later.]

Theorem 5.2: Consider a polytope[1]

$$\mathscr{P}_{C,b} = \{x : Cx \leq b\}$$

which is fixed a priori. Consider a black-box switching system and N couples (randomly, uniformly sampled) of its dynamics as in (4). Suppose that the polytope is invariant for all the observations (one can always scale the set of matrices and

---

[1]all our polytopes are supposed to be symmetric wrt the origin, so that they can be considered as the unit ball of a norm.

apply the theorem to the scaled version). That is, if the observed couples (x,y) are scaled so that x belongs to the boundary of the polytope, one has

$$Cy \leq b. \qquad (13)$$

Then, for any factor $1 < \delta$, one can compute the level of confidence β such that $\rho < \delta$.

Proof:

Let us fix $\delta > 1$. Define ε as the smallest measure of a set

$$X \subset \partial \mathscr{P} \text{ s.t. } \delta^{-1} = \sup\{d : d\mathscr{P} \subset \text{convhull } \partial \mathscr{P} \setminus X\}. \qquad (14)$$

In the equation above, ∂ represents the boundary of a set, and by measure, we mean the uniform measure on the Euclidean ball B (more precisely, $\mu(X) = \mu'(\{x \in B : \exists x' \in X, \lambda \in \mathbb{R}^+; x = \lambda x'\})$; and μ' is the uniform measure on the Euclidean ball).

Since Equation (13) does not contain any variable, it is a trivial application of Theorem 7.2, (or an application of first principles in probability) that, if the (scaled) vectors x are sampled randomly (say, on the unit sphere), and if we have N of these observations all of them satisfying (13), then we can claim that the measure of points x that violate (13) (more precisely, the points x such that, denoting $\hat{x}$ the scaled vector such that $\hat{x}$ belongs to the boundary of the polytope, we have that the image $y = f(\hat{x})$ does not satisfy $Cy \leq b$) is smaller than ε with confidence $1 - (1 - \varepsilon)^N$. Indeed, suppose by contradiction that the measure is $\varepsilon' \geq \varepsilon$. Then, the probability that our N samples satisfy (13) is equal to $(1 - \varepsilon')^N \leq (1 - \varepsilon)^N$. Thus, the level of confidence of our Hypothesis that only an ε measure of points x violate (13) is $1 - (1 - \varepsilon)^N$.

Summarizing, the equation above means that with confidence $\beta := 1 - (1 - \varepsilon)^N$, one has that (13) is satisfied for all $x \in \mathbb{R}^n$, except for a set of measure ε. Let us denote S_ε this set of violated constraints. That is,

$$f(\partial \mathscr{P} \setminus S_\varepsilon) \subset \mathscr{P}.$$

By our assumption 2.4,

$$f(\text{convhull } (\partial \mathscr{P} \setminus S_\varepsilon) \subset \text{convhull } (f(\partial \mathscr{P} \setminus S_\varepsilon)).$$

Combining the above two equations with our choice of ε in Equation (14), we have that

$$f(\text{convhull } (\partial \mathscr{P} \setminus S_\varepsilon) \subset \text{convhull } (\partial \mathscr{P} \setminus S_\varepsilon)/\delta,$$

and so $\rho \leq \delta$ with confidence $\beta := 1 - (1 - \varepsilon)^N$.

∎

A. the stronger version of the theorem

[RJ: This stronger version looks for an ellipsoid. We believe that everything is correct below too, but there's a little catch in the 'one can compute': for the moment we do not really know how to algorithmically compute ε like in the proof above, essentially because we have ellipsoids (but note that even

for the polytope above, everything is not crystal clear yet).]

Theorem 5.3: Consider a black-box switching system and N samples of its dynamics as in (4). Consider the optimal solution (λ^*, P) which minimizes λ in (6). For any factor $1 < \delta$, one can compute the level of confidence β such that $\rho < \delta \cdot \lambda^*$.

Proof:

Let us fix $\delta > 1$ and denote E_P the ellipsoid described by P (i.e., $\{x : x^T P x = 1\}$), and denote ε such that for any subset S_ε of measure ε,

$$E_{\delta^2 P} \subset \text{convhull } (E_P \setminus S_\varepsilon).$$

([RJ: this is where we still have to find a solution] First prove that the worst-case S_ε should be one single connected set, then, it should be centered around the longest axis of the ellipsoid.) [RJ: Alternatively, can't we simply rescale the state space so that the ellipsoid becomes a sphere? Then, we know how to compute the ε. We hope that this would come at the only price of increasing the measure of the points we miss, because we know how we distorted the measure. To be checked, fingers crossed.]

Now, denoting N the number of observations available, compute $0 < \beta < 1$ such that

$$N = N(\varepsilon, \beta)$$

in Theorem 7.2 above.

Summarizing, the equation above means that with high probability, one has that (6) is satisfied for all $x \in \mathbb{R}^n$, except for a set of measure ε. Let us denote S_ε this set of violated constraints. Thus,

$$(\mathscr{M}/\lambda^*)\text{convhull } (E_P \setminus S_\varepsilon) \subset \text{convhull } (E_P \setminus S_\varepsilon).$$

Now, by definition of ε, one has

$$E_{\delta^2 P} \subset \text{convhull } (E_P \setminus S_\varepsilon),$$

and so

$$(\mathscr{M}/\delta\lambda^*)\text{convhull } (E_P \setminus S_\varepsilon) \subset \text{convhull } (E_P \setminus S_\varepsilon).$$

Then, $\delta\lambda$ is un upper bound on ρ, with a confidence β.

∎

VI. CAMPI'S PRELIMINARIES

Let us consider $X = \text{S} \times M$ the Cartesian product of the unit sphere S with M. Every element of X can be written as $x = (s_x, k_x)$ with $s_x \in \text{S}$ and $k_x \in M$. For notational simplicity, we drop the subscript x whenever it is clear from the context.

We define the projections:

$$\pi_{\text{S}} : \text{S} \times M \to \text{S}, (s,k) \mapsto s$$

$$\pi_M : \text{S} \times M \to M, (s,k) \mapsto k.$$

We consider the classical unsigned and finite uniform spherical measure on S, commonly denoted by σ^{n-1}. It is

associated to the spherical Borelian σ-algebra and is derived from the Lebesgue measure (a measure of subset of the sphere will be related to the measure of the sector of the unit ball it induces).

More formally, we will have $\mathscr{B}_S$ with $\mathscr{A} \in \mathscr{B}_S$ if and only if the sector $t\mathscr{A}$, $t \in [0,1]$ is in $\mathscr{B}_{\mathbb{R}^n}$. The spherical measure σ^{n-1} is defined by

$$\forall\ \mathscr{A} \in \mathscr{B}_S, \sigma(A) = \frac{\lambda(tA)}{\lambda(B)}.$$

We can notice that $\sigma^{n-1}(S) = 1$.

We assume from now that S is provided with $(\mathscr{B}_S, \sigma^{n-1})$ and that M is provided with the classical σ-algebra associated to finite sets: $\Sigma_M = \wp(M)$, where $\wp(M)$ is the power set of M.

We consider an unsigned finite measure[2] μ_M on (M, Σ_M) with $\mathrm{supp}(\mu_M) = M$. In other words, $\forall k \in M$, $\mu_M(\{k\}) > 0$.

We denote the product σ-algebra $\mathscr{B}_S \bigotimes \Sigma_M$ engendered by $\mathscr{B}_S$ and Σ_M: $\Sigma = \sigma(\pi_S^{-1}(\mathscr{B}_S), \pi_M^{-1}(\Sigma_M))$. On this set, we define the product measure $\mu = \mu_S \otimes \mu_M$ which is an unsigned finite measure on X.

VII. CAMPI

We are interested in solving the following optimization problem for a given $\gamma \in (0,1)$:

find $\qquad P$
subject to $\quad (A_i s)^T P(A_i s) \leq \gamma^2 s^T P s, \quad \forall i = \{1, 2, \ldots, m\}, \forall s \in S,$
$\qquad\qquad P \succ 0.$

(15)

Note that if P is a solution to (15), then so is αP for any $\alpha \in \mathbb{R}_{>0}$. Therefore, we can rewrite (15) as the following optimization problem:

find $\qquad P$
subject to $\quad (A_k s)^T P(A_k s) \leq \gamma^2 s^T P s, \quad \forall i = \{1, 2, \ldots, m\}, \forall s \in S,$
$\qquad\qquad P \succeq I.$

(16)

We define the linear isomorphism Φ as the natural mapping $\Phi : \mathbb{R}^{\frac{n(n+1)}{2}} \to \mathbb{S}^n$. Using this mapping, for a fixed $\gamma \in (0,1]$ we can rewrite (16) as:

find $\qquad p$
subject to $\quad f(p,x) \leq 0, \forall x \in X.$

(17)

where $f(p,x) = \max(f_1(p,x), f_2(p))$, and

$$\begin{aligned} f_1(p,x) &:= (A_k s)^T \Phi(p)(A_k s) - \gamma^2 s^T \Phi(p) s \\ f_2(p) &:= \lambda_{\max}(\Phi(-p)) + 1. \end{aligned}$$

Proposition 7.1: The optimization problem (17) is convex.

Proof: The function $f_1(p,x)$ is clearly convex in p for a fixed $x \in X$. The function $\lambda_{max} : \mathbb{S}^n \to \mathbb{R}$ maps a symmetric positive matrix to its maximum eigenvalue. It is well-known that the function $\lambda_{\max}$ is a convex function of P. [?]. This

[2]Recall that the support of a measure μ defined on a measurable space (X, Σ) is $\mathrm{supp}(\mu) = \{A \in \Sigma | \mu(A) > 0\}$

means that, $p \mapsto \Phi(\lambda_{\max}(p))$ is convex in p. Moreover, maximum of convex functions is also convex, which shows that $f(p,x)$ is convex in p. ∎

Note that the optimization problem (17) has infinitely many constraints. We next consider the following optimization problem where we sample N constraints of (17) independently and identically with the probability measure $\mathbb{P}(A) = \frac{\mu(A)}{\mu(X)}, \forall A \in \Sigma$, where $N \geq d+1$, and $d := \frac{n(n+1)}{2}$. We denote this sampling by $\omega := \{x_1, x_2, \ldots, x_N\} \subset X$, and obtain the following convex optimization problem $\mathrm{Opt}(\omega)$:

find $\qquad p$
subject to $\quad f(p,x) \leq 0, \forall x \in \omega.$

(18)

Let $p^*(\omega)$ be the solution of $\mathrm{Opt}(\omega)$. We are interested in the probability of $p^*(\omega)$ violating at least one constraint in the original problem (17). Therefore, we define constraint violation property next.

Constraint violation probability [?] The constraint violation probability is defined as:

$$\mathscr{V}^*(\omega) = \begin{cases} \mathbb{P}\{x \in X : f(p^*(\omega), x) > 0\} & \text{if } \omega \in X^{N*}, \\ 1, & \text{otherwise} \end{cases}$$

where $X^{N*} := \{\omega \in X^N : \text{the solution of } \mathrm{Opt}(\omega) \text{ exists}\}$. Note that, since we have $\mathbb{P}(A) = \frac{\mu(A)}{\mu(X)}$, we can rewrite this as:

$$\mathscr{V}^*(\omega) = \begin{cases} \frac{\mu\{x \in X : f(p^*(\omega), x) > 0\}}{\mu(X)} & \text{if } \omega \in X^{N*}, \\ 1, & \text{otherwise} \end{cases}$$

We make the following assumptions on the problem $\mathrm{Opt}(\omega)$:

1) Uniqueness of solution: Note that this can be enforced by adding a tie-break rule of at most $\frac{n(n-1)}{2}$ convex conditions discriminating our solutions.
2) Nondegenaracy: with probability 1, there is no redundancy in the constraint obtained from the sampling.

The following theorem from [?] explicitly gives a relationship between $V^*(\omega)$ and N, n.

Theorem 7.2 (from [?]): Consider the optimization problem $\mathrm{Opt}(\omega)$ given in (18). Let Assumption 1 and Assumption 2 hold. Then, for all $\varepsilon \in (0,1)$ the following holds:

$$\mathbb{P}^N\{\{\mathscr{V}^*(\omega) \leq \varepsilon\} \cap X^{N*}\} \geq 1 - \sum_{j=0}^{d} \binom{N}{j} \varepsilon^j (1-\varepsilon)^{N-j}.$$

Note that $\varepsilon = 1 - I_\beta^{-1}(N - d, d + 1)$ and can be interpreted as the ratio of the measure of points in X that might violate at least one of the constraints in (16) to the measure of all points in X.

We now state our main theorem, which is based on Theorem 7.2 and devote the next section to proving it step by step. We denote by γ^*, the optimum value of the following optimization problem:

$$\min_{P, \gamma} \qquad \gamma$$
subject to $\quad (A_i s)^T P(A_i s) \leq \gamma^2 s^T P s, \quad \forall i = \{1, 2, \ldots, m\}, \forall s \in S,$
$\qquad\qquad P \succ 0.$

(19)

Theorem 7.3 (Main Theorem): For any $\eta > 0$, given $N \geq n+1$ and $\beta \in [0,1)$, we can compute $\delta < \infty$ such that with probability at least β, $\rho \leq \delta(1+\eta)\gamma^*$. Moreover, as $N \to \infty$, $\delta \to 1$.

VIII. PROOFS MEASURE

For a given sampling $\omega \in X^{N*}$, let $V := \{x \in X : f(p^*(\omega),x) > 0\}$, i.e., the set of points for which at least one constraint is violated, and V_S, V_M be its projections on S and M, respectively.

Lemma 8.1: $\mu_S(V_S) \leq \frac{\mu(V)}{m_1}$, where $m_1 = \min\{\mu_M(\{k\}), k \in M\}$.

Proof:

Let $A \subset X$, $A_S = \pi_S(A)$ and $A_M = \pi_M(A)$. We notice that Σ_M is the disjoint union of its 2^m elements $\{B_i, i \in \{1,2,\ldots 2^m\}\}$. Then A is the disjoint union $A = \sqcup_{1 \leq i \leq 2^m}(A_i, B_i)$ where $A_i = \pi_M^{-1}(B_i) \in S$. We notice that $A_S = \sqcup_{1 \leq i \leq 2^m} A_i$, and

$$\sigma^{n-1}(A_S) = \sum_{1 \leq i \leq 2^m} \sigma^{n-1}(A_i).$$

We have

$$
\begin{aligned}
\mu(A) = \mu(\sqcup_{1 \leq i \leq 2^m}(A_i, B_i)) &= \sum_{1 \leq i \leq 2^m} \mu((A_i, B_i)) \\
&= \sum_{1 \leq i \leq 2^m} \sigma^{n-1} \otimes \mu_M((A_i, B_i)) \\
&= \sum_{1 \leq i \leq 2^m} \sigma^{n-1}(A_i)\mu_M(B_i).
\end{aligned}
$$

Let m_1 be the minimum value of μ_M on its atoms: $m_1 = \min\{\mu_M(\{k\}), k \in M\}$ (recall that $m_1 > 0$). Then since $\forall i$, $\mu_M(B_i) \geq m_1$, we have

$$\sigma^{n-1}(A_S) \leq \frac{\mu(A)}{m_1}. \tag{20}$$

This proves our statement by taking $A = V_S$. ∎

Corollary 8.2: When the modes are sampled from the set M uniformly random,

$$\sigma^{n-1}(V_S) \leq m\mu(V).$$

We consider the linear transformation mapping S to E_P that denoted by $L \in \mathscr{L}(\mathbb{R}^n)$. Note that since $P \in \mathbb{S}^n$, it can be written in its Choleski form $P = UDU^{-1}$, where D diagonal matrix of its eigenvalues, and $U \in O_n(\mathbb{R})$. We define $D^{1/2}$ the positive square root of D as the matrix $\text{diag}(\sqrt{d_1},\ldots,\sqrt{d_n})$. Then, the positive square root of P is $VD^{1/2}V$. This means that, $L = P^{1/2}$. For the rest of the write-up, we denote

$$V' := \Pi_S(L^{-1}(V_S)),$$

and show how to upper bound $\sigma^{n-1}(V')$ in terms of $\mu(V)$.
****

Let us denote by $\mathscr{V}_S$ the sector of B defined by V_S. We recall the following result:

Lemma 8.3: Let ψ a smooth change of coordinates in $\mathbb{R}^n$ and $\mathscr{D} \subset \mathbb{R}^\setminus$, whose image under ψ is $\mathscr{D}' \subset \mathbb{R}^\setminus$. We provide $\mathbb{R}^n$ with the classical Borel σ-algebra and the Lebesgue measure λ. Then

$$\lambda(\mathscr{D}') = \int_{x \in \mathscr{D}} 1_{x \in \mathscr{D}} |\det J(\psi(x))| d\lambda(x), \tag{21}$$

which becomes when $\psi \in \mathscr{L}(\mathbb{R}^n)$ (and thus $\forall x \in \mathbb{R}^n, \det J(\psi(x)) = \det(\psi)$)

$$\lambda(\mathscr{D}') = |\det(\psi)|\lambda(\mathscr{D}). \tag{22}$$

Theorem 8.4:

$$\sigma n - 1(\Pi_S(L^{-1}(V_S))) \leq \det(L^{-1})\left(\frac{1}{\lambda_{\min}(L^{-1})}\right)^n \sigma^{n-1}(V_S). \tag{23}$$

Proof: We consider the set $\mathscr{V}'$ in $\mathbb{R}^n$ obtained after applying the mapping L^{-1} to $\mathscr{V}_S$. We relate now $\mathscr{V}'$ to a sector of B, which will give us access to the spherical measure of the set obtained after spherical projection of $L^{-1}(E_P)$.

We have that $\pi_S(\mathscr{V}') \subset \mathscr{H}_{1/\lambda_{\min}}(\mathscr{V}')$.

Hence $\lambda(\pi_S(\mathscr{V}')) \leq \frac{1}{\lambda_{\min}}^n(\mathscr{V}')$.

By the former lemma, we have that $\sigma n - 1(\Pi_S(L^{-1}(V_S))) = \lambda(\pi_S(\mathscr{V}')) \leq \det(L^{-1})\left(\frac{1}{\lambda_{\min}(L^{-1})}\right)^n \sigma^{n-1}(V_S)$. ∎

Theorem 8.5: $\sigma^{n-1}(V') \leq m\varepsilon \sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}$, where $\mu(V) = \varepsilon$.

IX. PROOFS DELTA

We denote $\varepsilon' := \frac{\varepsilon}{2}\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}$, where the additional factor $\frac{1}{2}$ follows from the homogeneity of the dynamics. In this section, we show how to relate ε' to δ in the statement of the Theorem 7.3. We start by a few definitions that will help us along the way. Let d be a distance on $\mathbb{R}^n$. We define the distance between a set $X \subset \mathbb{R}^n$ and a point $p \in \mathbb{R}^n$ is $d(X,p) := \inf_{x \in X} d(x,p)$.

Spherical Cap We define the *spherical cap* on S for a given hyperplane $c^T x = k$ as:

$$\mathscr{C}_{c,k} := \{x \in S : c^T x > k\}.$$

Proposition 9.1 (see e.g. [?]): The distance between the point $x = 0$ and the hyperplane $c^T x = k$ is $\frac{|k|}{\|c\|}$.

We define the function $\Delta : 2^S \to [0,1]$ as:

$$\Delta(X) := \sup\{r : B_r \subseteq \text{convhull }(S \setminus X)\}. \tag{24}$$

Note that, $\Delta(X)$ can be rewritten as in:

$$\Delta(X) = d(\partial\text{convhull }(S \setminus X),0). \tag{25}$$

Lemma 9.2: $\Delta(\mathscr{C}_{c,k}) = \min\left(1, \frac{|k|}{\|c\|}\right)$.

Proof: Note that convhull $(S \setminus X) = \{x \in B : c^T x \leq k\}$.

$$
\begin{aligned}
\Delta(X) &= d(\partial\text{convhull }(S \setminus X),0) \\
&= \min(d(\partial B,0), d(\partial\{x : c^T x \leq k\},0)) \\
&= \min(d(S,0), d(\{x : c^T x = k\},0)) \\
&= \min\left(1, \frac{|k|}{\|c\|}\right).
\end{aligned}
$$

∎

Corollary 9.3: $\Delta(\mathscr{C}_{c,k_1}) < \Delta(\mathscr{C}_{c,k_2})$ when $k_1 < k_2$.

Lemma 9.4: $\sigma^{n-1}(\mathscr{C}_{c,k_1}) < \sigma^{n-1}(\mathscr{C}_{c,k_2})$, for $k_1 > k_2$.

Proof: convhull $(S \setminus \{x \in S : c^T x > k_1\}) \subseteq$ convhull $(S \setminus \{x \in S : c^T x > k_2\})$, for $k_1 > k_2$. ∎

Now we are ready to present the following lemma which is the key to proving our main result.

Lemma 9.5: For any set $X \subseteq S$, there exists c and k such that $\mathscr{C}_{c,k}$ satisfies:

$$\mathscr{C}_{c,k} \subseteq X,$$

and

$$\Delta(\mathscr{C}_{c,k}) = \Delta(X). \tag{26}$$

Proof: Let $X_S := \text{convhull}\,(S \setminus X)$. Since the distance function d is continuous and the set ∂X_S is compact there exists a point $x^* \in \partial X_S$, such that:

$$\Delta(X) = d(\partial X_S, 0) = \inf_{x \in \partial X_S} d(x, 0) = \min_{x \in \partial X_S} d(x, 0) = d(x^*, 0). \tag{27}$$

Next, consider the supporting hyperplane of X_S at x^*, which we denote by $\{x : c^T x = k\}$. Note that, this supporting hyperplane is unique because it is also a supporting hyperplane of the ball $B_{\Delta(X)}$ at x^* as well, which is unique. This can be seen from the fact that:

$$\partial B_{\Delta(X)} \subseteq \partial X_S \subseteq \{x : c^T x = k\}.$$

Then we have:

$$\Delta(X) = d(x^*, 0) = d(\{x : c^T x = k\}, 0) = \frac{|k|}{\|c\|}.$$

Now, consider the spherical cap $\mathscr{C}_{c,k}$. Then, by Lemma we have $\Delta(\mathscr{C}_{c,k}) = \frac{|k|}{\|c\|}$. Therefore, $\Delta(X) = \Delta(\mathscr{C}_{c,k})$.

We next show $\mathscr{C}_{c,k} \subseteq X$. We prove this by contradiction. Assume $x \in \mathscr{C}_{c,k}$ and $x \notin X$. Note that, if $x \notin X$, then $x \in S \setminus X \subseteq \text{convhull}\,(S \setminus X)$. Since $x \in \mathscr{C}_{c,k}$ we have $c^T x > k$. But due to the fact that $x \in \text{convhull}\,(S \setminus X)$, we also have $c^T x \leq k$, which leads to a contradiction. Therefore, $\mathscr{C}_{c,k} \subseteq X$. ∎

We now prove our main result.

Theorem 9.6: Let $X_{\varepsilon'} = \{X \subset S : \sigma^{n-1}(X) = \varepsilon'\}$. Then, for any $\varepsilon' \in (0,1)$, the function $\Delta(X)$ attains its minimum over $X_{\varepsilon'}$ for some X which is a spherical cap.

Proof: We prove this via contradiction. Assume that there exists no spherical cap in $X_{\varepsilon'}$ such that $\Delta(X)$ attains its minimum. This means there exists an $X^* \in X_{\varepsilon'}$, where X^* is not a spherical cap and $\arg\min_{X \in X_{\varepsilon'}}(\Delta(X)) = X^*$. By Lemma 9.5 we can construct a spherical cap $\mathscr{C}_{c,k}$ such that $\mathscr{C}_{c,k} \subseteq X^*$ and $\mathscr{C}_{c,k} = \Delta(X^*)$. Note that, we further have $\mathscr{C}_{c,k} \subset X^*$, since X^* is assumed not to be a spherical cap. This means that, there exists a spherical cap $\sigma^{n-1}(\mathscr{C}_{c,k})$ such that $\sigma^{n-1}(\mathscr{C}_{c,k}) < \varepsilon'$.

Then, the spherical cap $\mathscr{C}_{c,\tilde{k}}$ with $\sigma^{n-1}(\mathscr{C}_{c,\tilde{k}}) = \varepsilon'$, satisfies $\tilde{k} < k$, due to Lemma 9.4. This implies $\Delta(\mathscr{C}_{c,\tilde{k}}) < \Delta(\mathscr{C}_{c,k}) = \Delta(X^*)$ due to Lemma 9.3. Therefore, $\Delta(\mathscr{C}_{c,\tilde{k}}) < \Delta(X^*)$. This is a contradiction since we initially assumed that $\Delta(X)$ attains its minimum over $X_{\varepsilon'}$ at X^*. ∎

Theorem 9.7: Given a spherical cap $\mathscr{C}_{c,k} \subseteq S$ such that $\sigma^{n-1}(\mathscr{C}_{c,k}) = \varepsilon'$,

$$\Delta(\mathscr{C}_{c,k}) = \sqrt{(1-\alpha)},$$

where $\alpha := I^{-1}\left(\frac{\varepsilon' \Gamma(\frac{d}{2})}{\pi^{d/2}}, \frac{d-1}{2}, \frac{1}{2}\right)$ and $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. Here I^{-1} is the inverse incomplete beta function, i.e., $I^{-1}(y, a, b) = x$ where $I_x(a, b) = y$.

Proof: Let $h := 1 - \Delta(\mathscr{C}_{c,k})$. It is well known [?] that the area of the spherical cap $\mathscr{C}_{c,k} \subseteq S$ is given by the equation:

$$\varepsilon' = \sigma^{n-1}(\mathscr{C}_{c,k}) = \frac{\pi^{d/2}}{\Gamma[\frac{d}{2}]} I_{2h-h^2}\left(\frac{d-1}{2}, \frac{1}{2}\right), \tag{28}$$

where I is the incomplete beta function. From this, we get the following set of equations:

$$
\begin{aligned}
\frac{\varepsilon' \Gamma[\frac{d}{2}]}{\pi^{d/2}} &= I_{2h-h^2}\left(\frac{d-1}{2}, \frac{1}{2}\right) \\
2h - h^2 &= I^{-1}\left(\frac{\varepsilon' \Gamma(\frac{d}{2})}{\pi^{d/2}}, \frac{d-1}{2}, \frac{1}{2}\right) \\
2h - h^2 &= \alpha \\
h^2 - 2h + \alpha &= 0. \tag{29}
\end{aligned}
$$

From (29), we get $h = 1 \pm \sqrt{(1-\alpha)}$. Since $h \leq 1$, we conclude that $\Delta(\mathscr{C}_{c,k}) = \sqrt{(1-\alpha)}$. Note that, $\Delta(\mathscr{C}_{c,k})$ only depends on ε for fixed n. ∎

Corollary 9.8: For a fixed $\beta \in (0,1)$, $\lim_{N \to \infty} \delta_\beta(N) = 1$.

Proof: We first prove that $\lim_{N \to \infty} \varepsilon_\beta(N) = 0$. Note that, we can upper bound $1 - \beta$ as follows:

$$\left(1 - \beta = \sum_{j=0}^d N \atop j \varepsilon^j (1-\varepsilon)^{N-j} \leq (d+1) N^d (1-\varepsilon)^{N-d}.\right) \tag{30}$$

We prove $\lim_{N \to \infty} \varepsilon_\beta(N) = 0$ by contradiction and assume that $\lim_{N \to \infty} \varepsilon_\beta(N) \neq 0$. This means that, there exists some $\delta > 0$ such that $\varepsilon_\beta(N) > \delta$ infinitely often. Then, consider the subsequence N_k such that $\varepsilon_\beta(N_k) > \delta, \forall k$. By (30) we have:

$$1 - \beta \leq (d+1) N_k^d (1-\varepsilon)^{N_k-d} \leq (d+1) N_k^d (1-\delta)^{N_k-d} \forall k \in \mathbb{N}.$$

Note that $\lim_{k \to \infty} (d+1) N_k^d (1-\delta)^{N_k-d} = 0$. Therefore, there exists a k' such that, we have $(d+1) N_{k'}^d (1-\delta)^{N_{k'}-d} < 1 - \beta$, which is a contradiction. Therefore, we must have $\lim_{N \to \infty} \varepsilon_\beta(N) = 0$.

Showing I^{-1} in its first parameter, $\delta = \sqrt{1-\alpha}$ tends to 1 as $\varepsilon \to 0$. ∎

X. FURTHER IDEAS

- develop the 'brute force campi application' (without our reasoning) on an example, and show with counterexamples and so that more hypotheses have to be used (as we do in Section V) in order to make it useful. Maybe this can be done on the first example (which still needs to be written down)?
- how to present/argue of usefulness against the fact that we can identify the matrices if we have a lot of points? Partial answer: identify the matrices is not easy (e.g. Paulo found a recent paper where they prove that it's NP-hard.[?].)
- We should certainly try to improve our estimate by playing with sos of higher degree rather than cqlf
- Noise   [RJ: Following Paulo's idea]
  If one has Property 2.1, then the Lyapunov inequality can be replaced by an equality. This

could allow for a better decision algorithm. [RJ: it is not completely clear how we can leverage that, and what nontrivial results we can claim (especially in view of the fact that we could identify the system). We thought that the more sensible was to add noise; and then to study how does a direct Lyapunov construction bypassing the identification step would compare with a 2 step (identification+stability analysis) method. We think that nontrivial (new?) questions might already appear for linear systems.]

- Find how to exploit the switching systems approach to other systems. How about a linear system which is intermittently controlled by a linear controller? Homogeneity, convex-preservation, $K-$convexity, seem to be key properties, but none of this is clear. (Also, Convex Converse Lyapunov Theorem seems to be important, at least for the lower bound.)
- How about continuous time? There, it seems that identification of the system is not possible. Thus, it is maybe a better framework to develop our ideas?

XI. EXAMPLES

XII. CONCLUSION

To do

A. *Future Works*

To do

XIII. ACKNOWLEDGMENTS

?

REFERENCES

[1] To do