

Chance-constrained quasi-convex optimization with application to data-driven switched systems control

Guillaume O. Berger
Raphaël M. Jungers
Zheming Wang

GUILLAUME.BERGER@UCLouvain.BE
RAPHAEL.JUNGERS@UCLouvain.BE
ZHEMING.WANG@UCLouvain.BE

Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Department of Mathematical Engineering (ICTEAM/INMA) at UCLouvain, 1348 Louvain-la-Neuve, Belgium

Abstract

We study quasi-convex optimization problems, where only a subset of the constraints can be sampled, and yet one would like a probabilistic guarantee on the obtained solution with respect to the initial (unknown) optimization problem. Even though our results are partly applicable to general quasi-convex problems, in this work we introduce and study a particular subclass, which we call “quasi-linear problems”. We provide optimality conditions for these problems. Thriving on this, we extend the approach of chance-constrained convex optimization to quasi-linear optimization problems. Finally, we show that this approach is useful for the stability analysis of black-box switched linear systems, from a finite set of sampled trajectories. It allows us to compute probabilistic upper bounds on the JSR of a large class of switched linear systems.

Keywords: Data-driven control, chance-constrained optimization, quasi-convex programming, switched systems.

1. Introduction

Data-driven control has gained a lot of interest from the control community in recent years; see, e.g., [Duggirala et al. \(2013\)](#); [Huang and Mitra \(2014\)](#); [Blanchini et al. \(2016\)](#); [Kozarev et al. \(2016\)](#); [Balkan et al. \(2017\)](#); [Boczar et al. \(2018\)](#). In many modern applications of control systems, one cannot rely on having a model of the system, but rather has to design a controller in a *blackbox*, *data-driven* fashion. This is the case for instance for *proprietary* systems; more usually, this happens because the system is too complex to be modeled, or because the obtained model is too complicated to be analyzed with classical control techniques. In these situations, the control engineer can only rely on data — which sometimes come in huge amounts —, but make the problem of very different nature than the classical, model-based control problems. Examples of such situations include self-driving cars, where the input to the controller is made of huge heterogeneous data (harvested from cameras, lidars, etc.); or smart grid applications, where the heterogeneous parts of the system (prosumers, smart buildings, etc.) are best described with data harvested from observing these parts than with a rigid, closed-form model ([Aswani et al., 2012](#); [Zhou et al., 2017](#)).

Data collected from a control system can be seen as *samples* extracted from a large set of possible behaviors. Controller design can then be approached by synthesizing controllers based on the sampled set of behaviors; the challenge is then to provide guarantees on the correctness of the controller for the whole behavior of the system. In optimization, this approach is known as *chance-constrained optimization*, which consists in sampling a *subset* of the constraints of an optimization

problem and solving the problem with these constraints only. The solution obtained in this way will in general not satisfy *all of the constraints* of the problem; however, probabilistic guarantees can be obtained on the measure of the set of constraints that are compatible with this solution (Calafiore, 2010; Margellos et al., 2014; Campi et al., 2018).

The approach of chance-constrained optimization has already proved useful in several areas of control, like robust control design (Calafiore and Campi, 2006) or quantized control (Campi et al., 2018). Recently, it has been successfully applied to data-driven control problems, as a technique to bridge the gap between data and model-based control; see, e.g., applications in data-enabled predictive control (Van Parys et al., 2015; Coulson et al., 2020) and stability analysis of black-box dynamical systems (Kenanian et al., 2019; Wang and Jungers, 2020).

A large class of control problems can be formulated as Linear Matrix Inequalities or Bilinear Matrix Inequalities (VanAntwerp and Braatz, 2000; Safonov et al., 1994). For some of these problems, replacing the matrix inequalities with their sampled counter-part leads to Linear Programs or Quadratic Programs. The approach of chance-constrained optimization then allows to bridge the gap between the original and the sampled formulations. To enable the systematic analysis of such problems in a data-driven fashion, we introduce the class of *quasi-linear optimization problems*, as a subclass of quasi-convex optimization problems (see, e.g., Eppstein, 2005), and we study it in the context of chance-constrained optimization. As we will see, this special class of optimization problems arises for instance from data-driven stability analysis of some hybrid systems.

In particular, we extend the results from Calafiore (2010) for chance-constrained optimization of convex problems to quasi-linear optimization problems. This is achieved by showing that for any such optimization problem there is a subset of constraints, called an *essential set*, with bounded cardinality, which provides the same optimal solution as the original problem. This result draws on an akin result for quasi-convex problems in Eppstein (2005), and improves it in two ways: we get a better upper bound on the cardinality of essential sets, while removing the assumption that the constraints are “continuously shrinking” (Eppstein, 2005).

As a proof of concept, we apply the setting of chance-constrained quasi-linear optimization for the stability analysis of *black-box switched linear systems*. Switched Linear Systems are systems described by a finite set of linear modes among which the system can switch over time. They constitute a paradigmatic class of hybrid and cyber-physical systems, and appear naturally in many engineering applications, or as abstractions of more complicated systems (Alur et al., 2009; Jadbabaie et al., 2003). These systems turn out to be extremely challenging in terms of control and analysis, even for basic questions like stability or stabilizability. In particular, the computation of the *Joint Spectral Radius* (JSR), a measure of stability of switched linear systems, has been used as a benchmark for testing new approaches in complex systems (Blondel and Nesterov, 2005; Parrilo and Jadbabaie, 2008; Jungers et al., 2017).

Recently, the problem of JSR approximation was introduced for *black-box* switched linear systems. It is well known that bounds on the JSR of switched linear systems can be obtained from the resolution of adequate quasi-convex optimization problems built from the system (Jungers et al., 2017). In Kenanian et al. (2019), the authors extend this approach when the system is not known but only a few trajectories are observed, and apply chance-constrained optimization techniques to obtain probabilistic upper bounds and lower bounds on the JSR of the system. In this work, we show that this approach fits in fact into the framework of chance-constrained quasi-linear optimization. From this, probabilistic upper and lower bounds on the JSR of the system can be obtained

straightforwardly, by applying the results introduced in this paper; the bounds obtained in that way are also better than the ones proposed in [Kenanian et al. \(2019\)](#).

The paper is organized as follows. In Section 2, we introduce the class of quasi-linear optimization problems and discuss their properties. In Section 3, we state and prove the main theorem of this paper, which extends the results of chance-constrained *convex* optimization to quasi-linear optimization problems. Then, in Section 4, we apply the framework of chance-constrained quasi-linear optimization to the problem of stability analysis of black-box switched linear systems, and we show how this framework can be used to obtain probabilistic bounds on the JSR of the system. Finally, in Section 5, we demonstrate the applicability of our results with several numerical examples.

Notation. \mathbb{N} denotes the set of nonnegative integers, and \mathbb{N}_* the set of positive integers. For a set of vectors $\mathcal{V} \subseteq \mathbb{R}^d$, $\text{conv}(U)$ denotes the *convex hull* of U , and $\text{cone}(U)$ its *conic hull*. For a convex function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and $x \in \mathbb{R}^d$, we let $\text{Sub}_x(f)$ be the *subdifferential* of f at x , i.e., the set of vectors $g \in \mathbb{R}^d$ such that $f(y) - f(x) \geq g^\top(y - x)$ for all $y \in \mathbb{R}^d$; for a convex set $\mathcal{C} \subseteq \mathbb{R}^d$, we let $\text{Nor}_x(\mathcal{C})$ be the *normal cone* of \mathcal{C} at x , i.e., the set of vectors $g \in \mathbb{R}^d$ such that $g^\top(y - x) \leq 0$ for all $y \in \mathcal{C}$. If Δ is a set, $\bar{\omega} := (\delta_1, \dots, \delta_N) \in \Delta^N$ and $\delta_{N+1} \in \Delta$, we use $\bar{\omega} \parallel \delta_{N+1}$ to denote their *concatenation*: $\bar{\omega} \parallel \delta_{N+1} := (\delta_1, \dots, \delta_{N+1})$; in Section 3, for the sake of simplicity, we will slightly abuse the notation and write ω to denote the *set* obtained from the elements of $\bar{\omega} := (\delta_1, \dots, \delta_N)$, i.e., $\omega = \{\delta_1, \dots, \delta_N\}$.

2. Quasi-linear optimization problems

In this section, we introduce a novel class of optimization problems, which are a particular case of quasi-convex problems. We particularize and improve some classical results of quasi-convex programming to this class.

Let \mathcal{X} be a compact convex subset of \mathbb{R}^d , with nonempty interior and with $0 \notin \mathcal{X}$. Let Δ be a set, and $\{a_\delta\}_{\delta \in \Delta}$ and $\{b_\delta\}_{\delta \in \Delta}$ be two collections — indexed by $\delta \in \Delta$ — of vectors in \mathbb{R}^d and such that $b_\delta^\top x > 0$ for all $x \in \mathcal{X}$ and $\delta \in \Delta$. Consider the following optimization problem:

$$\min_{x \in \mathbb{R}^d, \lambda \geq 0} (\lambda, c(x)) \quad \text{s.t.} \quad x \in \mathcal{X}, \quad \text{and} \quad a_\delta^\top x \leq \lambda b_\delta^\top x, \quad \forall \delta \in \Delta, \quad (1)$$

where $c : \mathcal{X} \rightarrow \mathbb{R}$ is a strongly convex function. The objective of (1) is to minimize $(\lambda, c(x))$ in the *lexicographical order*¹, while respecting the constraints defined by Δ and $x \in \mathcal{X}$. See Figure 1 for an illustration.

Example 1 Consider the problem of finding the smallest $\lambda \geq 0$ for which there is $x \in \mathcal{X} \subseteq \mathbb{R}^d$, with \mathcal{X} compact, such that $\mathcal{A}(x) \preceq \lambda \mathcal{B}(x)$, where \mathcal{A} and \mathcal{B} are linear functions from \mathbb{R}^d to the set of $n \times n$ positive semi-definite matrices. This problem belongs to the class of *Generalized Eigenvalue problems* (see, e.g., [Boyd and El Ghaoui, 1993](#)). The sampled counter-part of the matrix inequality consists in sampling points $\{\xi_\delta\}_{\delta \in \Delta} \subseteq \mathbb{R}^n$, and imposing that $\xi_\delta^\top \mathcal{A}(x) \xi_\delta \leq \lambda \xi_\delta^\top \mathcal{B}(x) \xi_\delta$ for all $\delta \in \Delta$. The latter constraints are quasi-linear in λ and x . The resulting quasi-linear optimization problem will appear for instance in the application to the data-driven stability analysis of switched linear systems, in Section 4.

1. “First component first”: $(\lambda_1, c_1) < (\lambda_2, c_2)$ if $\lambda_1 < \lambda_2$, or else $\lambda_1 = \lambda_2$ and $c_1 < c_2$.

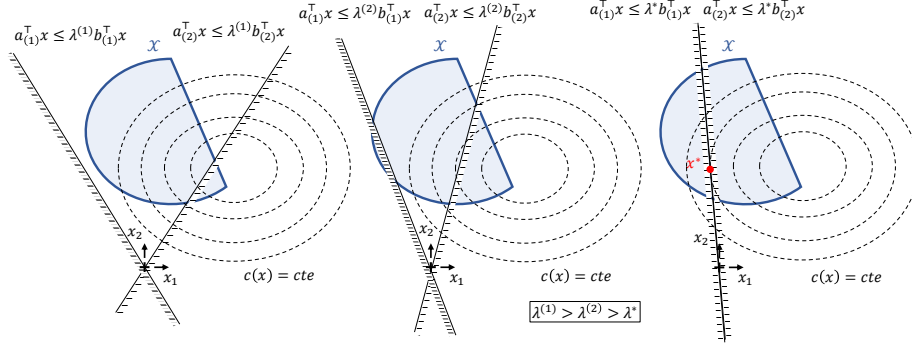


Figure 1: Set of feasible points $x \in \mathbb{R}^d$ of a quasi-linear optimization problem (1), for three different values of λ . The blue set \mathcal{X} represents the fixed constraints, while the two quasi-linear constraints are represented in black. The dotted curves are level-curves of the secondary cost function $c(x)$. The smallest λ for which there is a feasible point x is the optimal λ , denoted λ^* . For this value of λ , the feasible point x that minimizes c is the optimal point x , denoted x^* .

Sometimes, it is not possible to solve (1) with all the constraints defined by Δ , either because only a subset of these constraints are known (as it is the case for instance in data-driven control problems), or because the set Δ is so large (or even infinite) that it is algorithmically impracticable to enforce all of these constraints. In these cases, for a finite set $\omega \subseteq \Delta$, we consider the following *sampled* optimization problem:

$$\mathcal{P}(\omega) : \min_{x \in \mathbb{R}^d, \lambda \geq 0} (\lambda, c(x)) \quad \text{s.t.} \quad x \in \mathcal{X}, \quad \text{and} \quad a_\delta^\top x \leq \lambda b_\delta^\top x, \quad \forall \delta \in \omega. \quad (2)$$

We let $\text{Opt}(\omega)$ be the optimal solution² of $\mathcal{P}(\omega)$ and we let $\text{Cost}(\omega)$ be its optimal cost. The constraints of $\mathcal{P}(\omega)$ defined by ω will be called the *sampled constraints*, while the constraint $x \in \mathcal{X}$ is the *common constraint*.

For a fixed value of λ , the sampled constraints of $\mathcal{P}(\omega)$ are linear in x . Therefore, we will say that $\mathcal{P}(\omega)$ is a *quasi-linear optimization problem*. Note that $\mathcal{P}(\omega)$ is a particular instance of quasi-convex optimization problems, as defined in Eppstein (2005). It is shown there that, under some technical assumption on the continuity of the constraints, the cardinality of any *essential set* (see Definition 1 below) of a quasi-convex problem is upper bounded by $d + 1$, where d is the dimension of x . In this paper, we provide for *quasi-linear problems* a better upper bound on the cardinality of their essential sets, and without the technical assumption of “continuously shrinking” constraints, present in Eppstein (2005).

Definition 1 (Calafiore, 2010, Definition 2.9) *An essential set for $\mathcal{P}(\omega)$ is a set $\beta \subseteq \omega$, with minimal cardinality, satisfying $\text{Cost}(\beta) = \text{Cost}(\omega)$.*

Theorem 2 *Let $\mathcal{P}(\omega)$ be as in (2). The cardinality of any essential set β of $\mathcal{P}(\omega)$ satisfies $|\beta| \leq d$.*

To prove this theorem, we will need the following lemma.

2. By the strong convexity of c , $\text{Opt}(\omega)$ is unique.

Lemma 3 (Rockafellar, 1970, Theorem 27.4) *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a convex function and $\mathcal{C} \subseteq \mathbb{R}^d$ a nonempty convex set. Then, $x \in \mathcal{C}$ is a minimizer³ of f over \mathcal{C} if and only if $0 \in \text{Sub}_x(f) + \text{Nor}_x(\mathcal{C})$.*

Proof of Theorem 2 Let β be an essential set for $\mathcal{P}(\omega)$ and let $(\lambda^*, x^*) = \text{Opt}(\omega)$. For each $\delta \in \omega$, let $h_\delta = a_\delta - \lambda^* b_\delta$. Let $\gamma \subseteq \omega$ be the set of all $\delta \in \omega$ such that $h_\delta^\top x^* = 0$. We divide the proof in two cases.

Case 1: First, we consider the case when $\lambda^* = 0$. Assume that $x \in \mathcal{X}$ is a *support constraint*, meaning that the optimal cost of $\mathcal{P}(\omega)$ without this constraint is strictly smaller than $\text{Cost}(\omega)$. Then, by the classical argument⁴, there is a set of at most d constraints among those of $\mathcal{P}(\omega)$ (i.e., among the constraints defined by ω , and the constraint $x \in \mathcal{X}$) such that the optimal solution of the problem with these constraints only is equal to $\text{Cost}(\omega)$. Because $x \in \mathcal{X}$ is a support constraint, it must belong to this set of constraints. Hence, there is a set $\beta' \subseteq \omega$, with $|\beta'| \leq d - 1$, such that $\text{Cost}(\beta') = \text{Cost}(\omega)$. This shows that $|\beta| \leq d - 1$ when $x \in \mathcal{X}$ is a support constraint.

Now, assume that $x \in \mathcal{X}$ is not a support constraint, i.e., the optimal cost of $\mathcal{P}(\omega)$ without this constraint is the same as $\text{Cost}(\omega)$. By Lemma 3, it holds that $0 \in \text{Sub}_{x^*}(c) + \text{cone}(\{h_\delta\}_{\delta \in \gamma})$. Note that, by definition of γ , the vectors $\{h_\delta\}_{\delta \in \gamma}$ are all orthogonal to x^* , so that they belong to a $(d - 1)$ -dimensional subspace. Hence, by Caratheodory theorem⁵, there is a set $\gamma' \subseteq \gamma$, with $|\gamma'| \leq d - 1$, such that $0 \in \text{Sub}_{x^*}(c) + \text{cone}(\{h_\delta\}_{\delta \in \gamma'})$. By Lemma 3, it thus follows that $\text{Cost}(\gamma') = \text{Cost}(\omega)$. This shows that $|\beta| \leq d - 1$ when $x \in \mathcal{X}$ is not a support constraint; concluding the proof for the first case.

Case 2: Now, we consider the case when $\lambda^* > 0$. By Lemma 3 applied on $f(x) = \sup_{\delta \in \gamma} h_\delta^\top x$ and $\mathcal{C} = \mathcal{X}$, it follows that $0 \in \text{conv}(\{h_\delta\}_{\delta \in \gamma}) + \text{Nor}_{x^*}(\mathcal{X})$. Let $\gamma' \subseteq \gamma$ be a nonempty subset with minimal cardinality such that $0 \in \text{conv}(\{h_\delta\}_{\delta \in \gamma'}) + \text{Nor}_{x^*}(\mathcal{X})$. By Caratheodory theorem, it holds that $|\gamma'| \leq d - f + 1$ where f is the dimension of the linear subspace orthogonal to $\{h_\delta\}_{\delta \in \gamma'}$. We conclude the proof by using the same argument as in case 1: since the problem is now restricted to an f -dimensional problem (because x is in the subspace orthogonal to $\{h_\delta\}_{\delta \in \gamma'}$), we may find a set $\beta' \subseteq \omega$, with $|\beta'| \leq f - 1$, such that $\text{Opt}(\gamma' \cup \beta') = (\lambda^*, x^*)$. This shows that $|\beta| \leq |\gamma'| + |\beta'| \leq d$; concluding the proof for the second case. ■

3. Chance-constrained quasi-linear optimization

Let \mathbb{P} be a probability measure on Δ . Suppose that the constraints $\delta_1, \dots, \delta_N$ are sampled from Δ according to \mathbb{P} , and that we solve the problem $\mathcal{P}(\omega_N)$ where $\omega_N = \{\delta_1, \dots, \delta_N\}$. This approach of solving the optimization problem for a few randomly sampled constraints is called *chance-constrained optimization*. Under certain assumptions, probabilistic guarantees can be obtained on the measure of the set of constraints $\delta \in \Delta$ that are compatible with the optimal solution of $\mathcal{P}(\omega_N)$. This is the case, for instance, for a large class of convex optimization problems (see, e.g., Calafiore, 2010) and non-convex optimization problems (though with weaker probabilistic guarantees; see, e.g., Campi et al., 2018). In this section, we extend the results from chance-constrained convex optimization (Calafiore, 2010) to chance-constrained quasi-linear problems.

3. I.e., $f(x^*) = \inf_{x \in \mathcal{C}} f(x)$.

4. Indeed, $\mathcal{P}(\omega)$ with λ fixed to zero is a convex optimization problem, and the cardinality of essential sets of feasible convex optimization problems is bounded by d (see, e.g., Calafiore and Campi, 2006, Theorem 3).

5. See, e.g., Rockafellar (1970, Corollary 17.1.2).

Therefore, we make the following standing assumption (Assumption 5) on the set Δ and on its probability measure \mathbb{P} . First, let us introduce the notion of *non-degenerate* vector of constraints.

Definition 4 (Calafiore, 2010, Definition 2.11) *Consider Problem (1). Let $N \in \mathbb{N}_*$. We say that $\bar{\omega}_N := (\delta_1, \dots, \delta_N) \in \Delta^N$ is non-degenerate if there is a unique set $I \subseteq \{1, \dots, N\}$ such that $\{\delta_i\}_{i \in I}$ is an essential set for $\mathcal{P}(\omega_N)$.*

Assumption 5 (Calafiore, 2010, Assumption 2) *For every $N \in \mathbb{N}_*$, the vector $\bar{\omega}_N \in \Delta^N$ is non-degenerate with probability one.*

For any vector of constraints $\bar{\omega}_N \in \Delta^N$, we define the *violating probability* associated to $\bar{\omega}_N$:

$$V(\bar{\omega}_N) = \mathbb{P}(\{\delta \in \Delta : \text{Cost}(\omega_N \cup \{\delta\}) > \text{Cost}(\omega_N)\}).$$

We are now able to present the extension of Calafiore (2010, Theorem 3.3) to chance-constrained quasi-linear programs. Therefore, let $\zeta \in \mathbb{N}_*$ be an upper bound on the cardinality of any essential set of $\mathcal{P}(\omega)$, with finite $\omega \subseteq \Delta$. From Theorem 2, it holds that $\zeta \leq d$.

Theorem 6 *Consider the sampled quasi-linear optimization problem (2), and let $V(\bar{\omega}_N)$ and ζ be as above. Let Assumption 5 hold. Let $N \in \mathbb{N}$, $N \geq \zeta$, and let $\varepsilon \in (0, 1)$. Then,*

$$\mathbb{P}^N(\{\bar{\omega}_N \in \Delta^N : V(\bar{\omega}_N) > \varepsilon\}) \leq \Phi(\varepsilon, \zeta - 1, N),$$

where $\Phi(\cdot, \zeta - 1, N)$ is the regularized incomplete beta function⁶.

Using Theorem 2, different proofs can be used to get the above result; see, e.g., Campi and Garatti (2018, Theorem 3). Here, we present a proof based on Calafiore (2010, Theorem 3.3).

Proof (Adapted from Calafiore, 2010) Fix $N \in \mathbb{N}$, $N \geq \zeta$. By Assumption 5, we may assume without loss of generality that $\bar{\omega}_N$ is non-degenerate for all $\bar{\omega}_N \in \Delta^N$. Hence, for each $\bar{\omega}_N := (\delta_1, \dots, \delta_N) \in \Delta^N$, we let $J(\bar{\omega}_N)$ be the unique set $I \subseteq \{1, \dots, N\}$ such that $\{\delta_i\}_{i \in I}$ is an essential set for $\mathcal{P}(\omega_N)$. Label the elements of Δ with labels belonging to a totally order set.⁷ Let $J^*(\bar{\omega}_N)$ be a completion of $J(\bar{\omega}_N)$ with the $\zeta - |J(\bar{\omega}_N)|$ elements of $\{1, \dots, N\} \setminus J(\bar{\omega}_N)$ such that $\{\delta_i\}_{i \in J^*(\bar{\omega}_N) \setminus J(\bar{\omega}_N)}$ have the largest labels among the elements of ω_N . From Assumption 5, it follows that $J^*(\bar{\omega}_N)$ is well defined with probability one; hence, in the following, we will assume without loss of generality that $J^*(\bar{\omega}_N)$ is well defined for all $\bar{\omega}_N \in \Delta^N$.

Let $\{I_1, \dots, I_M\}$ be the set of all subsets of $\{1, \dots, N\}$ with ζ elements; in particular, $M = N! / (\zeta!(N - \zeta)!) \triangleq C(N, \zeta)$. For each $i = 1, \dots, M$, let $S_i = \{\bar{\omega}_N \in \Delta^N : J^*(\bar{\omega}_N) = I_i\}$. The sets $\{S_i\}_{1 \leq i \leq M}$ are disjoint and their union is equal to Δ^N . Moreover, by the symmetry of their definition, they have the same probability; hence $\mathbb{P}(S_i) = 1/C(N, \zeta)$.

Now, for each $\bar{\omega}_\zeta \in \Delta^\zeta$, we let $V^*(\bar{\omega}_\zeta)$ be the *violating probability* of $\bar{\omega}_\zeta$ with respect to (2) and the labelling of the constraints: that is, $V^*(\bar{\omega}_\zeta) = \mathbb{P}(\{\delta \in \Delta : J^*(\bar{\omega}_\zeta \parallel \delta) \neq \{1, \dots, \zeta\}\})$. From the uniqueness of the optimal solution of the problems $\mathcal{P}(\omega)$, $\omega \subseteq \Delta$, it follows that for every $L \in \mathbb{N}_*$,

6. See, e.g., Kenanian et al. (2019, Definition 2).

7. This approach, from Calafiore (2010), requires the *axiom of choice* when Δ is a general set. However, it is not needed for instance if $\Delta \subseteq \mathbb{R}^n$, as it is the case in our application (see Section 4).

$\bar{\omega}_L \in \Delta^L$ and $\delta, \eta \in \Delta$, if $J^*(\bar{\omega}_L \|\delta) = J^*(\bar{\omega}_L \|\eta) = J^*(\bar{\omega}_L)$, then $J^*((\bar{\omega}_L \|\delta) \|\eta) = J^*(\bar{\omega}_L)$.⁸ It follows that, for any $v \in [0, 1]$,

$$\mathbb{P}[S_i \mid V^*(\bar{\omega}_{N,i}) = v] = (1 - v)^{N-\zeta}, \quad \forall N \geq \zeta, \quad i = 1, \dots, C(N, \zeta),$$

where $\bar{\omega}_{N,i}$ is the restriction of $\bar{\omega}_N$ to the indices in I_i : $\bar{\omega}_{N,i} = (\delta_i)_{i \in I_i}$. Hence, we get that

$$\mathbb{P}(S_i) = \int_0^1 (1 - v)^{N-\zeta} dF_i(v) = 1/C(N, \zeta), \quad \forall N \geq \zeta, \quad i = 1, \dots, C(N, \zeta), \quad (3)$$

where $F_i(v) = \mathbb{P}^N(\{\bar{\omega}_N \in \Delta^N : V^*(\bar{\omega}_{N,i}) \leq v\})$. Equation (3) describes a *Hausdorff moment problem*; it is shown in Calafiore (2010, p. 3436) that (3) implies that $F_i(v) = v^\zeta$.

Finally, for each $i = 1, \dots, C(N, \zeta)$, we let $B_i = \{\bar{\omega}_N \in \Delta^N : V^*(\bar{\omega}_{N,i}) > \varepsilon\}$. Using the expression of F_i , it can be shown⁹ that $\mathbb{P}^N(B_i) = \Phi(\varepsilon, \zeta - 1, N)/C(N, \zeta)$. By symmetry, we get that $\mathbb{P}^N(\bigcup_{1 \leq i \leq M} B_i) = \Phi(\varepsilon, \zeta - 1, N)$. Since $\{\bar{\omega}_N \in \Delta^N : V(\bar{\omega}_N) > \varepsilon\} \subseteq \bigcup_{1 \leq i \leq M} B_i$, we obtain the desired result. \blacksquare

4. Application to data-driven stability analysis of switched linear systems

Let $\mathcal{A} = \{A_1, \dots, A_m\}$ be a fixed set of matrices in $\mathbb{R}^{n \times n}$, and let \mathbb{S} be the unit sphere (boundary of the unit Euclidean ball) in \mathbb{R}^n . Let $\Delta = \mathcal{A} \times \mathbb{S}$, and let \mathbb{P} be the *uniform distribution* on Δ .¹⁰ For a finite set $\omega \subseteq \Delta$, we consider the following sampled quasi-linear optimization problem:

$$\mathcal{P}_{\text{jsr}}(\omega) : \min_{\substack{P=P^\top \in \mathbb{R}^{n \times n} \\ \gamma \geq 0}} (\gamma, \|P\|_F^2) \quad \text{s.t.} \quad P \in \mathcal{X} := \{P : P \succeq I \wedge \|P\|_F \leq C\}, \quad (4)$$

$$(Ax)^\top P(Ax) \leq \gamma^2 x^\top P x, \quad \forall (A, x) \in \omega,$$

for some fixed parameter $C \geq n$. Note that $\mathcal{P}_{\text{jsr}}(\omega)$ is a sampled, data-driven version of the classical quadratic Lyapunov framework for the approximation of the *Joint Spectral Radius* (JSR) of the *switched linear system* defined by \mathcal{A} ; see, e.g., Jungers (2009, Theorem 2.11). The JSR is a ubiquitous measure of stability of switched linear systems (Blondel and Nesterov, 2005; Parrilo and Jadbabaie, 2008; Jungers et al., 2017); it also appears in other areas of hybrid system control, like wireless networked control (Berger and Jungers, 2020).

In order to apply the results from Section 3 on $\mathcal{P}_{\text{jsr}}(\omega)$, we make the following assumption on the matrices in \mathcal{A} . First, let us introduce the notion of *Barabanov matrix* (referring to the fact that these matrices admit a quadratic Barabanov norm; see Jungers, 2009, Definition 2.3).

Definition 7 *A matrix $A \in \mathbb{R}^{n \times n}$ is said to be Barabanov if there exists a symmetric matrix $P \succ 0$ and $\gamma \geq 0$ such that $A^\top P A = \gamma^2 P$.*

Assumption 8 *There is no Barabanov matrix in \mathcal{A} .*

We claim that Assumption 8 is not restrictive in most of the practical situations. To motivate this claim, we provide an equivalent characterization of Barabanov matrices in the proposition below, whose proof can be found in Appendix A. For further work, we plan to investigate the possibility to relax or remove this technical assumption.

8. See, e.g., Calafiore (2010, §2.1) for details.

9. See, e.g., Calafiore (2010, Theorem 3.3).

10. I.e., $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$ where \mathbb{P}_1 and \mathbb{P}_2 are the uniform distributions on \mathcal{A} and \mathbb{S} respectively.

Proposition 9 *A matrix $A \in \mathbb{R}^{n \times n}$ is Barabanov if and only if it is diagonalizable and all its eigenvalues have the same modulus.*

We now show that Assumption 8 ensures that Assumption 5 holds for (4).

Proposition 10 *Consider the sampled problem (4). Let Assumption 8 hold. Then, for every $N \in \mathbb{N}_*$, the vector $\bar{\omega}_N \in \Delta^N$ is non-degenerate with probability one.*

We will need the following lemma.

Lemma 11 *Let $P(x_1, \dots, x_n)$ be a nonzero polynomial on \mathbb{R}^n . The zero set of P , i.e., the set of points $x \in \mathbb{R}^n$ such that $P(x) = 0$, has zero Lebesgue measure.*

We skip the proof of this well-known fact (see, e.g., [Teschl](#), Problem 2.15).

Proof of Proposition 10 Let $\bar{\omega}_N = \{\delta_1, \dots, \delta_N\}$. Let $1 \leq i \leq N-1$. Let us look at the probability that $\beta := \{\delta_1, \dots, \delta_i\}$ is an essential set for $\mathcal{P}_{\text{jsr}}(\omega)$ and that δ_N is in another essential set. This probability is smaller than or equal to the probability that β is an essential set for $\mathcal{P}_{\text{jsr}}(\omega)$ and that $(Ax)^\top P(Ax) = \gamma^2 x^\top Px$, where $(\gamma, P) = \text{Opt}_{\text{jsr}}(\beta)$ and $\delta_N = (A, x)$.

Assume that the above probability is nonzero. Then, since \mathcal{A} is finite, that there is $A \in \mathcal{A}$ such that $(Ax)^\top P(Ax) = \gamma^2 x^\top Px$ for all x in a set $S \subseteq \mathbb{S}$ with nonzero measure. Thus, by Lemma 11, it holds that $(Ax)^\top P(Ax) = \gamma^2 x^\top Px$ for all $x \in \mathbb{S}$. This contradicts the assumption that there is no Barabanov matrix in \mathcal{A} . Hence, the probability that β is a basis for $\mathcal{P}_{\text{jsr}}(\omega)$ and that δ_N is in another basis is zero. Since β and δ_N were arbitrary, this concludes the proof. \blacksquare

Theorem 6 can thus be applied to $\mathcal{P}_{\text{jsr}}(\omega)$.

Corollary 12 *Consider the sampled problem (4). Let Assumption 8 hold. Let $N \in \mathbb{N}$, $N \geq d := \frac{n(n+1)}{2}$, and let $\varepsilon \in (0, 1)$. Then,*

$$\mathbb{P}^N(\{\bar{\omega}_N \in \Delta^N : V_{\text{jsr}}(\bar{\omega}_N) > \varepsilon\}) \leq \Phi(\varepsilon, d-1, N), \quad (5)$$

where $V_{\text{jsr}}(\bar{\omega}_N) = \mathbb{P}(\{\delta \in \Delta : \text{Cost}_{\text{jsr}}(\omega_N \cup \{\delta\}) > \text{Cost}_{\text{jsr}}(\omega_N)\})$.

Remark 13 *We note the improvement of the right-hand side term of (5), compared to [Kenanian et al. \(2019, Theorem 10\)](#); this term becomes $\Phi(\varepsilon, d-1, N)$ instead of $\Phi(\varepsilon, d, N)$ in [Kenanian et al. \(2019\)](#). This is due to the improvement of the bound on the cardinality of essential sets of quasi-linear problems; see [Theorem 2](#).*

From Corollary 12, we deduce the following probabilistic guarantee on the upper bound on the JSR of the switched linear system given by \mathcal{A} , that we can get from the solution of the sampled problem $\mathcal{P}_{\text{jsr}}(\omega)$. The derivation of this result follows the same lines as in [Kenanian et al. \(2019, Theorems 14 and 15\)](#), so that the details are omitted here.

Corollary 14 *Consider the sampled problem (4). Let Assumption 8 hold. Let $N \in \mathbb{N}$, $N \geq d := \frac{n(n+1)}{2}$, and let $\varepsilon \in (0, 1)$. Then, for all $\bar{\omega}_N \in \Delta^N$, except possibly those $\bar{\omega}_N$ in a subset $\Omega \subseteq \Delta^N$ with measure $\mathbb{P}^N(\Omega) \leq \Phi(\varepsilon, d-1, N)$, it holds that*

$$\rho(\mathcal{A}) \leq \gamma^* / \sqrt{1 - I^{-1}\left(\frac{\varepsilon \kappa(P^*)}{m}; \frac{d-1}{2}; \frac{1}{2}\right)},$$

where $(\gamma^*, P^*) = \text{Opt}_{\text{JSR}}(\omega_N)$, $\kappa(P) = \sqrt{\frac{\det(P)}{\lambda_{\min}(P)^n}}$, I^{-1} is the inversed regularized incomplete beta function¹¹ and $\rho(\mathcal{A})$ is the JSR of the switched linear system defined by \mathcal{A} .

5. Numerical experiments: consensus of hidden network

We consider the problem of consensus in a switching and hidden network. The interaction between the nodes in the network over time can be modeled as a switched linear dynamical system:

$$x(t+1) = A_{\sigma(t)}x(t), \quad x(t) \in \mathbb{R}^n, \quad A_{\sigma(t)} \in \mathcal{A} := \{A_1, \dots, A_m\} \subseteq \mathbb{R}^{n \times n},$$

where $x(t)$ is the state vector (n is the number of nodes) at time t and $A_{\sigma(t)}$ is the interaction matrix at time t , with A_i being *unknown* row-stochastic matrices, i.e., $A_i \mathbf{1} = \mathbf{1}$, $i = 1, \dots, m$, where $\mathbf{1}$ is the all-one vector in \mathbb{R}^n . The goal is to verify that $x(t) = A_{\sigma(t-1)} \cdots A_{\sigma(1)} A_{\sigma(0)} x(0)$ converges to $c\mathbf{1}$ for some c as $t \rightarrow \infty$. As shown by [Jadbabaie et al. \(2003\)](#), this question boils down to the computation of the JSR of $\mathcal{A}' := \{A'_1, \dots, A'_m\} \subseteq \mathbb{R}^{(n-1) \times (n-1)}$ where $A'_i = BA_iB^\top$, for $i = 1, \dots, m$, and $B \in \mathbb{R}^{(n-1) \times n}$ is a fixed orthogonal matrix ($BB^\top = I_{n-1}$) with kernel spanned by $\mathbf{1}$. In our experiment, we consider a network of 8 nodes, switching among 3 modes, as shown in Figure 2. The possible networks are not known, and only the state of the different agents is available. Hence, we use the data-driven framework in Section 4 to estimate the JSR of \mathcal{A}' .

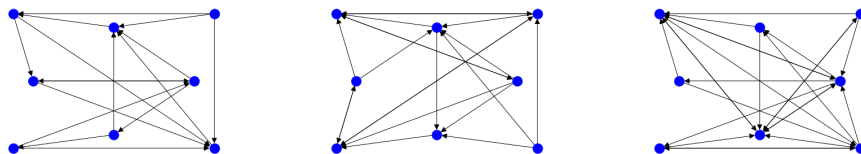


Figure 2: Example of switching network with 3 modes.

First, we sample a data set of N pairs: (x_i, y_i) , with $i = 1, \dots, N$, where x_i is sampled uniformly at random on \mathbb{S} and $y_i = A_{\sigma_i}x_i$, with σ_i sampled uniformly at random in $\{1, \dots, m\}$. This data set is projected onto \mathbb{R}^{n-1} as follows: $(x_i, y_i) \mapsto (x'_i, y'_i)$ where $x'_i = Bx_i$ and $y'_i = By_i$ and B is as above.¹² We then solve the problem in Section 4 with the projected data set. We fix the confidence level at $\beta = 0.05$. The probabilistic upper bound on the JSR obtained from Corollary 14 is shown in Figure 3 for different sizes of the sample set. For a comparison, the bound of [Kenanian et al. \(2019\)](#) is also given. While both bounds converge when the number of samples increases, the bound in this paper requires fewer samples to deduce convergence of the system to consensus, with the same confidence level.

6. Conclusions

In this work, we generalized the theory of chance-constrained optimization to quasi-convex problems, and pushed further the effort initiated in [Kenanian et al. \(2019\)](#), demonstrating its use for data-driven stability analysis of complex systems. More precisely, we introduced the class of quasi-linear optimization problems, which is a subclass of quasi-convex problems. We particularized

11. See, e.g., [Kenanian et al. \(2019, Definition 2\)](#).

12. The orthogonality of B is important to ensure that $x'_i/\|x'_i\|$ is distributed uniformly on \mathbb{S} .

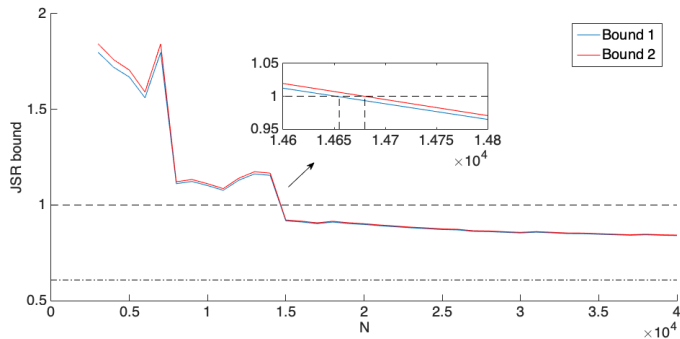


Figure 3: Data-driven upper bounds on the JSR for different sizes of the sample set. Bound 1 refers to the bound of this paper, Bound 2 refers to the bound of [Kenanian et al. \(2019\)](#), and the dash-dotted line is the bound computed from the white-box model, using the JSR toolbox ([Jungers, 2009](#)).

and improved some classical results of quasi-convex programming to this class. This allowed us to extend the results of chance-constrained convex optimization to quasi-linear optimization problems. Thriving on this, we provided a proof of concept that quasi-linear problems are useful for data-driven control applications. In particular, we applied our framework to the problem of JSR approximation of black-box switched linear systems, introduced in [Kenanian et al. \(2019\)](#).

For future work, we plan to investigate other applications of chance-constrained quasi-linear optimization for data-driven control. For instance, we believe that by replacing the conic constraints with their sampled counterpart, one could transform many optimization problems in control theory into quasi-linear programs, and then use chance-constrained optimization to bridge the gap between the original and the sampled formulations. We also plan to investigate the possibility of relaxing or removing the assumption that there are no Barabanov matrices involved in the switched linear system. Finally, we plan to provide other approaches for the data-driven stability analysis of switched linear systems based on chance-constrained quasi-linear optimization (e.g., thriving on sum-of-square optimization or path-complete Lyapunov frameworks).

Appendix A. Proof of Proposition 9

First, we prove the *if* direction: Assume that A is diagonalizable and all its eigenvalues have the same modulus. Then, there is $T \in \mathbb{R}^{n \times n}$ invertible such that $A = T^{-1}DT$, where $D \in \mathbb{R}^{n \times n}$ is block-diagonal with diagonal blocks of size 1 or 2, corresponding to eigenvalues with the same modulus. Denote this common modulus by γ . Now, let $P = T^T T$, which is positive definite. We verify that $A^T P A = T^T D^T D T = \gamma^2 T^T T = \gamma^2 P$. Hence, A is Barabanov.

Now, we show the *only if* direction: Assume that $A^T P A = \gamma^2 P$ for some $P \succ 0$ and $\gamma \geq 0$. Let $P = L^T L$ be a Cholesky factorization of P . It follows that $B^T B = \gamma^2 I$, where $B = L A L^{-1}$. If $\gamma = 0$, this implies that $B = 0$ and thus $A = 0$, proving the *only if* direction when $\gamma = 0$. If $\gamma > 0$, this implies that B/γ is a unitary matrix. It follows that B is diagonalizable and all its eigenvalues have modulus γ . Now, since A is similar to B , the same holds for A , proving the *only if* direction when $\gamma > 0$. ■

References

- Rajeev Alur, Alessandro D’Innocenzo, Karl H Johansson, George J Pappas, and Gera Weiss. Modeling and analysis of multi-hop control networks. In *2009 15th IEEE Real-Time and Embedded Technology and Applications Symposium*, pages 223–232. IEEE, 2009. doi: 10.1109/RTAS.2009.40.
- Anil Aswani, Neal Master, Jay Taneja, Virginia Smith, Andrew Krioukov, David Culler, and Claire J Tomlin. Identifying models of HVAC systems using semiparametric regression. In *2012 American Control Conference (ACC)*, pages 3675–3680. IEEE, 2012. doi: 10.1109/ACC.2012.6315566.
- Ayca Balkan, Paulo Tabuada, Jyotirmoy V Deshmukh, Xiaoqing Jin, and James Kapinski. Underminer: a framework for automatically identifying nonconverging behaviors in black-box system models. *ACM Transactions on Embedded Computing Systems (TECS)*, 17(1):1–28, 2017. doi: 10.1145/3122787.
- Guillaume O Berger and Raphaël M Jungers. Worst-case topological entropy and minimal data rate for state observation of switched linear systems. In *Proceedings of the 23rd International Conference on Hybrid Systems: Computation and Control*, pages 1–11. ACM, 2020. doi: 10.1145/3365365.3382195.
- Franco Blanchini, Gianfranco Fenu, Giulia Giordano, and Felice Andrea Pellegrino. Model-free plant tuning. *IEEE Transactions on Automatic Control*, 6(62):2623–2634, 2016. doi: 10.1109/TAC.2016.2616025.
- Vincent D Blondel and Yurii Nesterov. Computationally efficient approximations of the joint spectral radius. *SIAM Journal on Matrix Analysis and Applications*, 27(1):256–272, 2005. doi: 10.1137/040607009.
- Ross Boczar, Nikolai Matni, and Benjamin Recht. Finite-data performance guarantees for the output-feedback control of an unknown system. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 2994–2999. IEEE, 2018. doi: 10.1109/CDC.2018.8618658.
- Stephen Boyd and Laurent El Ghaoui. Method of centers for minimizing generalized eigenvalues. *Linear algebra and its applications*, 188:63–111, 1993. doi: 10.1016/0024-3795(93)90465-Z.
- Giuseppe C Calafiore. Random convex programs. *SIAM Journal on Optimization*, 20(6):3427–3464, 2010. doi: 10.1137/090773490.
- Giuseppe C Calafiore and Marco C Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006. doi: 10.1109/TAC.2006.875041.
- Marco C Campi and Simone Garatti. Wait-and-judge scenario optimization. *Mathematical Programming*, 167(1):155–189, 2018. doi: 10.1007/s10107-016-1056-9.
- Marco C Campi, Simone Garatti, and Federico A Ramponi. A general scenario theory for non-convex optimization and decision making. *IEEE Transactions on Automatic Control*, 63(12):4067–4078, 2018. doi: 10.1109/TAC.2018.2808446.

- Jeremy Coulson, John Lygeros, and Florian Dörfler. Distributionally robust chance constrained data-enabled predictive control. *arXiv:2006.01702*, 2020.
- Parasara Sridhar Duggirala, Sayan Mitra, and Mahesh Viswanathan. Verification of annotated models from executions. In *2013 Proceedings of the International Conference on Embedded Software (EMSOFT)*, pages 1–10. IEEE, 2013. doi: 10.1109/EMSOFT.2013.6658604.
- David Eppstein. Quasiconvex programming. *Combinatorial and Computational Geometry*, 52: 287–331, 2005.
- Zhenqi Huang and Sayan Mitra. Proofs from simulations and modular annotations. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, pages 183–192. ACM, 2014. doi: 10.1145/2562059.2562126.
- Ali Jadbabaie, Jie Lin, and A Stephen Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003. doi: 10.1109/TAC.2003.812781.
- Raphaël M Jungers. *The joint spectral radius: theory and applications*, volume 385 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag Berlin Heidelberg, 2009. doi: 10.1007/978-3-540-95980-9.
- Raphaël M Jungers, Amir Ali Ahmadi, Pablo A Parrilo, and Mardavij Roozbehani. A characterization of Lyapunov inequalities for stability of switched systems. *IEEE Transactions on Automatic Control*, 62(6):3062–3067, 2017. doi: 10.1109/TAC.2017.2671345.
- Joris Kenanian, Ayca Balkan, Raphaël M Jungers, and Paulo Tabuada. Data driven stability analysis of black-box switched linear systems. *Automatica*, 109:108533, 2019. doi: 10.1016/j.automatica.2019.108533.
- Alexandar Kozarev, John Quindlen, Jonathan How, and Ufuk Topcu. Case studies in data-driven verification of dynamical systems. In *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, pages 81–86. ACM, 2016. doi: 10.1145/2883817.2883846.
- Kostas Margellos, Paul J Goulart, and John Lygeros. On the road between robust optimization and the scenario approach for chance constrained optimization problems. *IEEE Transactions on Automatic Control*, 59(8):2258–2263, 2014. doi: 10.1109/TAC.2014.2303232.
- Pablo A Parrilo and Ali Jadbabaie. Approximation of the joint spectral radius using sum of squares. *Linear Algebra and Its Applications*, 428(10):2385–2402, 2008. doi: 10.1016/j.laa.2007.12.027.
- R Tyrrell Rockafellar. *Convex analysis*. Princeton University Press, 1970.
- Michael G Safonov, Keat-Choon Goh, and JH Ly. Control system synthesis via bilinear matrix inequalities. In *Proceedings of 1994 American Control Conference – ACC '94*, pages 45–49. IEEE, 1994. doi: 10.1109/ACC.1994.751690.
- Gerald Teschl. *Topics in real analysis*. Graduate studies in mathematics. American Mathematical Society. <https://www.mat.univie.ac.at/~gerald/ftp/book-ra/ra.pdf>.

Bart PG Van Parys, Daniel Kuhn, Paul J Goulart, and Manfred Morari. Distributionally robust control of constrained stochastic systems. *IEEE Transactions on Automatic Control*, 61(2):430–442, 2015. doi: 10.1109/TAC.2015.2444134.

Jeremy G VanAntwerp and Richard D Braatz. A tutorial on linear and bilinear matrix inequalities. *Journal of process control*, 10(4):363–385, 2000. doi: 10.1016/S0959-1524(99)00056-6.

Zheming Wang and Raphaël M Jungers. A data-driven method for computing polyhedral invariant sets of black-box switched linear systems. *arXiv:2009.10984*, 2020.

Datong P Zhou, Qie Hu, and Claire J Tomlin. Quantitative comparison of data-driven and physics-based models for commercial building HVAC systems. In *2017 American Control Conference (ACC)*, pages 2900–2906. IEEE, 2017. doi: 10.23919/ACC.2017.7963391.