Data-Driven Control of Unknown Switched Linear Systems Using Scenario Optimization

Zheming Wang, *Member, IEEE*, Guillaume O. Berger, *Member, IEEE*, and Raphaël M. Jungers, *Senior Member, IEEE*

***Abstract*—We tackle *uniform* state feedback control of switched linear systems under arbitrary switching using scenario optimization. We propose a data-driven control framework, in which scenario programs are formulated to compute stabilizing state feedback control relying on a finite set of observations of trajectories with quadratic and sum of squares (SOS) Lyapunov functions. We do not require the exact dynamical model or the switching signal, and as a consequence, we aim at solving *uniform* stabilization problems, in which the feedback is stabilizing for all possible switching sequences. In order to generalize the solution obtained from trajectories to the actual system, probabilistic guarantees on the obtained quadratic or SOS Lyapunov function are derived in the spirit of scenario optimization. For the quadratic Lyapunov technique, the generalization relies on a geometric analysis argument, while, for the SOS Lyapunov technique, we follow a sensitivity analysis argument. In order to deal with high-dimensional systems, we also develop a parallelized scheme for the proposed approach. We show that, with some modifications, the data-driven quadratic Lyapunov technique can be extended to linear quadratic regulator (LQR) control design. Finally, the proposed data-driven control framework is demonstrated on several numerical examples.**

***Index Terms*—Data-driven control, scenario optimization, stabilization, switched linear systems.**

Manuscript received 6 December 2023; accepted 23 March 2024. Date of publication 28 March 2024; date of current version 25 October 2024. The work of Raphaël M. Jungers was supported in part by the Innoviris Foundation and in part by FNRS (Chist-Era Druid-net). The work of Zheming Wang was supported in part by the National Natural Science Foundation of China under Grant 62303416. The work of Guillaume O. Berger was supported by WBI and FNRS, in part by the funding of Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang under Grant 2023R01006, in part by the Key Research and Development Program of Zhejiang Province under Grant 2023C01144, and in part by the the European Research Council through the European Union's Horizon 2020 research and innovation programme under Grant 864017 - L2C. Recommended by Associate Editor A. Chaillet. *(Corresponding author: Zheming Wang.)*

Zheming Wang is with the Department of Automation, Zhejiang University of Technology, Hangzhou 310023, China (e-mail: wangzheming@zjut.edu.cn).

Guillaume O. Berger and Raphaël M. Jungers are with the ICTEAM Institute, UCLouvain, 1348 Louvain-la-Neuve, Belgium (e-mail: guillaume.berger@uclouvain.be; raphael.jungers@uclouvain.be).

Codes are written in Julia and available online at https://github.com/zhemingwang/DataDrivenSwitchControl.

Digital Object Identifier 10.1109/TAC.2024.3382610

I. Introduction

SWITCHED systems are typical hybrid dynamical systems, which consist of a number of dynamics modes and a switching rule selecting the current mode. The jump from one mode to another often causes complicated hybrid behaviors resulting in significant challenges in stability analysis and control design, see, e.g., [1], [2], [3]. This article focuses on stabilization and control of switched linear systems.

Many control techniques have been proposed for switched systems depending on the assumptions on the switching rule. In the case where the switching signal is the only control input, a standard technique for achieving stabilization is to impose constraints on the switching sequence, e.g., dwell time [4], [5], [6] and average dwell time [7], [8]. More advanced techniques use state-dependent switching rules to achieve stabilization, see, e.g., [5], [9], [10] and the references therein. In [9] and [11], necessary and sufficient conditions for stabilizability are also provided. For switched systems with affine control inputs, the stabilization problem becomes even more complicated due to extra freedom. In [12] and [13], the (time) varying nature of the dynamics is considered as uncertainty and uniform state feedback stabilization laws are proposed for all possible switching sequences. When both the affine control input and the switching signal are accessible, exponential stabilization can be achieved for instance by using piecewise quadratic control Lyapunov functions, which are essentially solutions to switched linear quadratic regulator (LQR) problems [14]. In the presence of state and input constraints, optimal control of switched linear systems is also addressed under the framework of model predictive control [15], [16]. However, these stabilization methods all require a model of the underlying switched system.

While there exist hybrid system identification techniques [17], identification of state-space models of switching systems is in general cumbersome and computationally demanding, in particular when the switching signal is not accessible. More specifically, identifying a switched linear system is NP-hard [18] and most of currently available techniques as mentioned in [17] rely on heuristics and lack of formal guarantees. Recently, data-driven control of complex systems has received a lot of attention, see, e.g., [19], [20], [21], [22]. In particular, for switched systems, the computational hardness in hybrid system identification has motivated several data-driven approaches that do not require a mathematical model of the system. For instance, data-driven stability analysis is considered in [23], [24], and [25] for arbitrarily switched linear systems, based on the observation of a finite set of trajectories. In [26], a direct data-driven control design approach is proposed for unknown switched linear systems under unknown switching signals. However, the stability analysis in [26] relies on the assumption that the dwell time of the switching signal is larger than a certain number, which is unknown and implicitly depends on the parameters of the system, that is in general stability is guaranteed only when the switching is slow enough. Such an assumption can be restrictive in real-world applications, especially when the model is unknown and there is no control over the switching signal.

In this article, we address feedback control design for arbitrarily switched linear systems without knowing the model of the system or the switching signal. As the switching is considered as a source of uncertainty, we need to design a uniform state feedback controller allowing to stabilize the system in the worst case, similar to [12], [13]. To do this, we compute a state feedback controller and a common Lyapunov function for all the switching modes of the closed-loop system using a finite set of trajectories. We use both quadratic and sum of squares (SOS) Lyapunov functions, which lead to constrained nonlinear optimization problems with a large number of Lyapunov inequalities. For numerical tractability, we then design algorithms to solve these problems by making use of the underlying structure. For example, with quadratic Lyapunov functions, the biconvex Lyapunov inequalities allow to use alternating minimization where each iteration solves convex problems.

One major issue of this data-driven feedback control design is that it is usually only valid for the regions where the data are sampled but may not stabilize the actual system in the whole space. In order to formally describe the properties of the controller, we derive probabilistic stability guarantees in the spirit of scenario optimization [27], [28], [29], [30], [31]. In this context, one trajectory can be considered as a scenario and the stabilization problem formulated based on a set of trajectories is a scenario program. As our problem is nonconvex, the convex chance-constrained theorems in [27] are not applicable. While chance-constrained theorems for nonlinear optimization problems also exist in [28] and [29], their probabilistic bounds rely on the knowledge of the essential set (which is basically the set of irremovable constraints). Identifying this set can be very expensive for general nonlinear problems, in particular for nonlinear semidefinite problems. Hence, the techniques in [28] and [29] are not suitable for our case, which involves a large number of polynomial constraints and linear matrix inequalities (LMI)/bilinear matrix inequalities (BMI). Instead, probabilistic stability guarantees in this article are derived relying on the notions of set covering and packing (see, e.g., [32, Ch. 27]) and geometric/sensitivity analysis of the underlying problem. Similar probabilistic guarantees are also developed in [23], [24], and [25] for autonomous systems. However, the probabilistic guarantees in [23], [24], and [25] all require the optimality of the obtained solution, while our techniques work with any feasible solution of the underlying optimization problem. This allows to parallelize our algorithms to substantially speed up the computations. Finally, we show that the proposed data-driven Lyapunov framework can be extended to switched LQR problems.

A preliminary version of this article appears as a conference paper in [33], which only considers quadratic stabilization, and does not contain complete proofs. In this article, we provide complete detailed proofs of all the results in [33]. In particular, this article contains the first published proof of our main result, Theorem 1. Furthermore, we present an extension to SOS stabilization, which calls for a new technique for deriving probabilistic stability guarantees. Besides the stabilization problem, we also consider a switched LQR problem. In addition, to circumvent computational issues, we also present a parallelized scheme for both quadratic and SOS stabilization.

The rest of this article is organized as follows. This section ends with the notation, followed by Section II on the review of preliminary results on stability of switched linear systems and the formulation of the state feedback stabilization problem. Section III presents the proposed data-driven quadratic Lyapunov technique, including an alternating minimization algorithm, probabilistic stability analysis and a parallelized scheme. In Section IV, the SOS Lyapunov technique is considered with a similar alternating minimization algorithm. In Section V, we extend the data-driven Lyapunov framework to switched LQR design. Numerical results are provided in Section VI. Finally, Section VII concludes this article.

Notation: The set of nonnegative integers is denoted by $\mathbb{Z}^+$. For a square matrix Q, $Q \succ (\succeq)\, 0$ means that Q is symmetric and positive definite (semidefinite). $\mathbb{S}$ and $\mathbb{B}$ are the unit sphere and the unit (closed) ball in $\mathbb{R}^n$, respectively. $\mu(\cdot)$ denotes the uniform spherical measure on $\mathbb{S}$ with $\mu(\mathbb{S}) = 1$. For any square matrix P, $\mathrm{tr}(P)$ denotes the trace of P. For any symmetric matrix P, we denote by $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$ the largest and smallest eigenvalues of P, respectively. For any matrix $P \succ 0$, let $\kappa(P) := \lambda_{\max}(P)/\lambda_{\min}(P)$ be the condition number. For any $p \geq 1$, the ℓ_p norm of a vector $x \in \mathbb{R}^n$ is $\|x\|_p$ ($\|x\|$ is the ℓ_2 norm by default) with $\|x\|_Q^2 = x^\top Q x$ for any $Q \succeq 0$. Finally, given $x \in \mathbb{S}$ and $\theta \in [0, \pi/2]$, let $\mathrm{Cap}(x, \theta) := \{v \in \mathbb{S} : |x^\top v| \geq \cos(\theta)\}$ be the *symmetric spherical cap* with direction x and angle θ.

II. Preliminaries and Problem Statement

We consider the following switched linear system:

$$x(t+1) = A_{\sigma(t)}x(t) + Bu(t), \quad t \in \mathbb{Z}^+ \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the input, and $\sigma : \mathbb{Z}^+ \to \mathcal{M} := \{1, 2, \ldots, M\}$ is a time-dependent switching signal that indicates the current active mode of the system among M possible modes in $\mathcal{A} := \{A_1, A_2, \ldots, A_M\}$. In this article, we consider the case in which the switching signal is changing arbitrarily and cannot be observed, i.e., the information on the switching signal is not available. Note that the input matrix B is constant. Our goal is to find a static linear state feedback stabilizing the system under arbitrary switching, that is, a feedback matrix $K \in \mathbb{R}^{m\times n}$ such that the closed-loop system below is stable for all switching signals

$$x(t+1) = (A_{\sigma(t)} + BK)x(t), \quad t \in \mathbb{Z}^+. \tag{2}$$

For notational convenience, let $\mathcal{A}_K := \{A_1 + BK, A_2 + BK, \ldots, A_M + BK\}$ for a given $K \in \mathbb{R}^{m\times n}$. The stability of system (2) under arbitrary switching can be described by the joint spectral radius (JSR) [34] of the matrix set $\mathcal{A}_K$ defined by

$$\rho(\mathcal{A}_K) := \lim_{k\to\infty} \max_{\boldsymbol{\sigma}(k)\in\mathcal{M}^k} \|\bar{\boldsymbol{A}}_{\boldsymbol{\sigma}(k)}(K)\|^{1/k} \tag{3}$$

where $\boldsymbol{\sigma}(k) := \{\sigma(0), \sigma(1), \ldots, \sigma(k-1)\}$, $\mathcal{M}^k$ is the k-ary Cartesian power of $\mathcal{M}$ and $\bar{\boldsymbol{A}}_{\boldsymbol{\sigma}(k)}(K) = (A_{\sigma(k-1)} + BK)\cdots(A_{\sigma(1)} + BK)(A_{\sigma(0)} + BK)$. System (2) is asymptotically stable when $\rho(\mathcal{A}_K) < 1$. Hence, the state feedback stabilization problem for System (1) amounts to finding a $K \in \mathbb{R}^{m\times n}$ such that $\rho(\mathcal{A}_K) < 1$. However, the computation of the JSR of a set of matrices is known to be a very challenging problem in general, let alone its optimization in the context of control design. For this reason, we use tractable upper bounds on the JSR, providing sufficient conditions for stability or stabilization, see [34]. The following proposition provides a sufficient condition based on a common quadratic Lyapunov function which can be computed via semidefinite programming (SDP) [35].

Proposition 1 ([34, Prop. 2.8]): Consider the closed-loop matrices $\mathcal{A}_K$ for some state feedback $K \in \mathbb{R}^{m\times n}$. If there exist $\gamma \ge 0$ and $P \succ 0$ such that $A^\top PA \preceq \gamma^2 P \ \forall A \in \mathcal{A}_K$, then $\rho(\mathcal{A}_K) \le \gamma$.

From this proposition, we formulate the following nonlinear semidefinite optimization problem for stabilization of switched linear systems

$$\gamma^* := \min_{\gamma\ge 0, P, K} \gamma \tag{4a}$$

$$\text{s.t.} \quad (A + BK)^\top P(A + BK) \preceq \gamma^2 P \quad \forall A \in \mathcal{A} \tag{4b}$$

$$P \succ 0. \tag{4c}$$

Using the Schur complement formula [36, Th. 1.12] with $S = P^{-1}$ and $Y = KS$, the nonlinear constraints in (4) can be converted into LMI when γ is fixed

$$\min_{\gamma\ge 0, S, Y} \gamma \tag{5a}$$

$$\text{s.t.} \quad \begin{pmatrix} \gamma^2 S & SA^\top + Y^\top B^\top \\ AS + BY & S \end{pmatrix} \succeq 0 \quad \forall A \in \mathcal{A} \tag{5b}$$

$$S \succ 0. \tag{5c}$$

Such a transformation is widely used in stability analysis and control design, see, e.g., [37]. When the matrices $\mathcal{A}$ are known, problem (5) can be efficiently solved via SDP and bisection on γ.

In this article, we aim to solve the stabilization problem of switched linear systems when the matrices $\mathcal{A}$ are unknown and only a finite set of trajectories of the system are observed. Such systems are called black-box switched linear systems as in [23]. To this end, we reformulate problem (4) as a problem with an infinite number of constraints as follows:

$$\min_{\gamma\ge 0, P, K} \gamma \tag{6a}$$

$$\text{s.t.} \quad (Ax + BKx)^\top P(Ax + BKx) \le \gamma^2 x^\top Px$$
$$\forall A \in \mathcal{A} \quad \forall x \in \mathbb{R}^n \tag{6b}$$

$$P \succ 0. \tag{6c}$$

By homogeneity, one can restrict the constraints in (6b) to the set of points in the unit sphere $\mathbb{S}$ instead of the whole space $\mathbb{R}^n$. As we will show later, the formulation in (6) allows us to develop a data-driven control design. For that, we make the following assumption about the observations available.

Assumption 1: The state $x(t)$ can be fully observed for all $t \in \mathbb{Z}^+$, the input matrix B is time-invariant and known, and the number of modes (or an upper bound) is available.

The assumption that B is time-invariant is not restrictive in many applications, for instance, when the switching only occurs in some parameters of the dynamics. Such an assumption is often made in the literature, see, e.g., [12] and [38]. In the literature of learning-based control, the assumption on the knowledge of B is also considered, see, e.g., [39] and [40]. From a practical point of view, the control input dynamics sometimes can be known a priori or identified with proper initialization.

III. State Feedback Stabilization

This section presents our data-driven quadratic Lyapunov technique for stabilizing unknown switched linear systems under arbitrary switching. We first formulate a data-based stabilization problem, which consists of a set of biconvex constraints. Then, to solve this problem, we present an alternating minimization algorithm that generates feasible iterates. With the concepts of covering/packing numbers, probabilistic guarantees on the obtained solution are then provided using geometric analysis. Finally, we also show that the algorithm can be parallelized to speed up the computation.

A. Data-Based Stabilization Problem

For the data-driven design, we sample a finite number of pairs of the initial state and switching mode. More precisely, we randomly and uniformly generate N initial states on $\mathbb{S}$ and N modes in $\mathcal{M}$, which are denoted by $\omega_N := \{(x_i, \sigma_i) \in \mathbb{S} \times \mathcal{M} : i = 1, 2, \ldots, N\}$. The sampled points in ω_N are called scenarios. From this random sampling, we observe the trajectories of the system in (1) with $u = 0$ and obtain the observed dataset $\{(x_i, x_i^+) : i = 1, 2, \ldots, N\}$, where $x_i^+ := A_{\sigma_i} x_i$ is the successor of the initial state x_i with respect to mode σ_i. Note that although $A_{\sigma_i} x_i$ depends on σ_i, σ_i is not directly observed. It is well known that identifying a switched linear system without the information on the switching signal is NP-hard [18]. While heuristics techniques are available in [17] for switched systems identification, they lack of formal guarantees.

For the given sample set ω_N, we define the following *sampled* problem or scenario program:

$$\min_{\gamma\ge 0, P, K} \gamma \tag{7a}$$

$$\text{s.t.} \quad \left(x_i^+ + BKx_i\right)^\top P\left(x_i^+ + BKx_i\right) \le \gamma^2 x_i^\top P x_i$$
$$i = 1, 2, \ldots, N \tag{7b}$$

$$P \succ 0 \tag{7c}$$

which can be equivalently written in a compact form as

$$\min_{\gamma\ge 0, P, K} \gamma \tag{8a}$$

$$\text{s.t.} \quad (A_\sigma x + BKx)^\top P(A_\sigma x + BKx) \le \gamma^2 x^\top Px \quad \forall\,(x,\sigma) \in \omega_N \tag{8b}$$

$$P \succ 0. \tag{8c}$$

From the homogeneity property of (8b), the inequality $P \succ 0$ can be replaced by $P \succeq I$ due to the fact that the feasibility of P implies the feasibility of $P/\lambda_{\min}(P)$. When the size of ω_N is small, the sampled problem (8) can be solved using polynomial optimization toolboxes [41], [42], [43]. As P is invertible, the Schur complement formula [36, Th. 1.12] can be applied to (8b) using the reformulation as follows:

$$(A_\sigma x + BKx)^\top PP^{-1}P(A_\sigma x + BKx) \le \gamma^2 x^\top Px \quad \forall\,(x,\sigma) \in \omega_N. \tag{9}$$

We can then convert the constraints in (8b) into a set of BMI and reformulate Problem (8) as the following BMI problem:

$$\min_{\gamma\ge 0, P\succeq I, K} \gamma \tag{10a}$$

$$\text{s.t.} \quad \begin{pmatrix} \gamma^2 x^\top Px & (A_\sigma x + BKx)^\top P \\ P(A_\sigma x + BKx) & P \end{pmatrix} \succeq 0 \quad \forall\,(x,\sigma) \in \omega_N. \tag{10b}$$

This reformulation allows us to solve Problem (8) using BMI solvers [37], [44], [45].

B. Alternating Algorithm

As the size of ω_N increases, it becomes numerically intractable to find a (local) optimum of (8) or (10) using the aforementioned polynomial or BMI solvers. Since we do not seek to have optimality, we can use a less costly approach described below. We propose an alternating minimization algorithm between P and K for its numerical tractability and simple implementation. As we will show later, this alternating algorithm also enables us to parallelize the computation. Given a fixed K, we also define

$$\bar{\mathcal{P}}(\omega_N; K) := \min_{\gamma\ge 0, P\succeq I} \gamma \tag{11a}$$

$$\text{s.t.} \quad (A_\sigma x + BKx)^\top P(A_\sigma x + BKx) \le \gamma^2 x^\top Px \quad \forall\,(x,\sigma) \in \omega_N. \tag{11b}$$

For fixed values of γ, the constraints (11b) reduce to LMIs, so that Problem (11) can be solved efficiently using SDP solvers [35] and bisection on γ, with the solution of (12) being the initial feasible guess. Given a fixed P, we define

$$\hat{\mathcal{P}}(\omega_N; P) := \min_{\gamma\ge 0, K} \gamma \tag{12a}$$

$$\text{s.t.} \quad (A_\sigma x + BKx)^\top P(A_\sigma x + BKx) \le \gamma^2 x^\top Px \quad \forall\,(x,\sigma) \in \omega_N. \tag{12b}$$

Problem (12) is a second-order cone program that can be solved by well-documented convex solvers, such as interior point methods [35]. The overall procedure is summarized in Algorithm 1. Note that this alternating algorithm always terminates since the sequence $\{\gamma_k\}$ is nonincreasing, though it does not necessarily converge to a (local) optimum of problem (10).

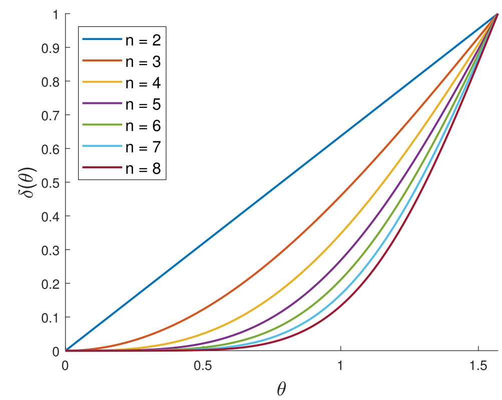


Fig. 1. Measure μ of the symmetric spherical cap $\mathrm{Cap}(x,\theta)$ in $\mathbb{R}^n$ for different values of n.

Algorithm 1: Alternating Minimization for Quadratic Stabilization.

Input: $\{(x_i, x_i^+)\}_{i=1}^N$, B and some tolerance $\epsilon_{tol} > 0$
Output: $\gamma(\omega_N)$, $P(\omega_N)$, and $K(\omega_N)$
1: *Initialization:* Let $k \leftarrow 0$, $K_k \leftarrow 0$, $P_k \leftarrow I$, and $\gamma_k \leftarrow \max_{(x,\sigma)\in\omega_N} \frac{\|A_\sigma x\|}{\|x\|}$;
2: Obtain P_{k+1} from (11) with $K = K_k$ via bisection on γ starting from $\max_{(x,\sigma)\in\omega_N} \frac{\|(A_\sigma x + BK_k x)\|_{P_k}}{\|x\|_{P_k}}$;
3: Obtain K_{k+1} and γ_{k+1} from (12) with $P = P_{k+1}$;
4: **if** $\|\gamma_{k+1} - \gamma_k\| < \epsilon_{tol}$ **then**
5: $\quad\gamma(\omega_N) \leftarrow \gamma_{k+1}$, $P(\omega_N) \leftarrow P_{k+1}$, $K(\omega_N) \leftarrow K_{k+1}$;
6: $\quad$Terminate;
7: **else**
8: $\quad$Let $k \leftarrow k+1$ and go to Step 2.
9: **end if**

C. Probabilistic Stability Guarantees

We now derive formal stability guarantees on the solution obtained from Algorithm 1. We first show in Theorem 1 that any feasible solution (γ, P, K) to problem (8) leads to an upper bound on the closed-loop JSR $\rho(\mathcal{A}_K)$ when ω_N is an ϵ-covering of $\mathbb{S}\times\mathcal{M}$. We then provide the probability that ω_N is an ϵ-covering of $\mathbb{S}\times\mathcal{M}$ for independent and identically distributed (i.i.d.) sampling and apply Theorem 1 to Algorithm 1 in Corollary 1.

Let us introduce a few definitions. For any $\theta \in [0, \pi/2]$ and any $x \in \mathbb{S}$, we let $\delta(\theta)$ denote the relative area of the *symmetric spherical cap* $\mathrm{Cap}(x,\theta)$, given by $\{v \in \mathbb{S} : |x^\top v| \ge \cos(\theta)\}$. From [46], it holds that

$$\delta(\theta) = \mathcal{I}\left(\sin^2(\theta); \frac{n-1}{2}, \frac{1}{2}\right) \tag{13}$$

where $\mathcal{I}(x;a,b)$ is the regularized incomplete beta function defined as

$$\mathcal{I}(x;a,b) := \frac{\int_0^x t^{a-1}(1-t)^{b-1}dt}{\int_0^1 t^{a-1}(1-t)^{b-1}dt}. \tag{14}$$

The function δ is strictly increasing with θ, and thus we can define its inverse, denoted by δ^{-1}, see also Fig. 1 for an illustration. Let $\delta_v(\theta)$ denote the relative volume of the portion of the unit ball $\mathbb{B}$ enclosed by $\mathrm{Cap}(x,\theta)$ and the hyperplanes $|x^\top v| = \cos(\theta)$, expressed as $\{z \in \mathbb{B} : |x^\top z| \geq \cos(\theta)\}$. We recall again from [46] that $\delta_v(\theta)$ can be given by

$$\delta_v(\theta) = \mathcal{I}\left(\sin^2(\theta); \frac{n+1}{2}, \frac{1}{2}\right) \quad \forall \theta \in [0, \pi/2]. \tag{15}$$

Similarly, the function δ_v is also strictly increasing with θ and its inverse is denoted by δ_v^{-1}.

Let us also recall the notions of covering and packing numbers, see [32, Ch. 27] for details. We adapt the classical definitions to the unit sphere.

Definition 1: Given $\epsilon \in (0,1)$, a set $Z \subset \mathbb{S}$ is called an ϵ-*covering* of $\mathbb{S}$ if, for any $x \in \mathbb{S}$, there exists $z \in Z$ such that $|z^\top x| \geq \cos(\theta)$ where $\theta = \delta^{-1}(\epsilon)$. The *covering number* $\mathcal{N}_c(\epsilon)$ is the minimal cardinality of an ϵ-covering of $\mathbb{S}$.

Definition 2: Given $\epsilon \in (0,1)$, a set $Z \subset \mathbb{S}$ is called an ϵ-*packing* of $\mathbb{S}$ if, for any two $z, v \in Z$, $|z^\top v| < \cos(\theta)$ where $\theta = \delta^{-1}(\epsilon)$. The *packing number* $\mathcal{N}_p(\epsilon)$ is the maximal cardinality of an ϵ-packing of $\mathbb{S}$.

With these definitions, we also adapt fundamental results on set covering and packing (see [32, Ch. 27]) to the unit sphere, as stated below.

Lemma 1: For any $\epsilon \in (0,1)$

$$\mathcal{N}_c(\epsilon) \leq \mathcal{N}_p(\epsilon) \leq \frac{1}{\delta(\frac{1}{2}\delta^{-1}(\epsilon))}. \tag{16}$$

Proof: The first inequality follows from the fact that any ϵ-packing with maximal cardinality is also an ϵ-covering. To prove the second inequality, let Z be the ϵ-packing with the maximal cardinality. Let $\theta = \delta^{-1}(\epsilon)$. From the definition of an ϵ-packing, the spherical caps $\{\mathrm{Cap}(z,\theta/2)\}_{z\in Z}$ are disjoint. Hence, $\sum_{z\in Z} \mu(\mathrm{Cap}(z,\theta/2)) \leq 1$, which leads to the second inequality. ■

Remark 1: The definitions above are similar to those in [47], except that we consider *symmetric spherical caps* in the form of $\{v \in \mathbb{S} : |x^\top v| \geq \cos(\theta)\}$ given any $\theta \in [0, \pi/2]$ and any $x \in \mathbb{S}$, to take into account the symmetry of the problem.

We then extend the definition of ϵ-covering to the joint set $\mathbb{S} \times \mathcal{M}$ as follows.

Definition 3: Given $\epsilon \in (0,1)$, a set $\omega \subset \mathbb{S} \times \mathcal{M}$ is called an ϵ-*covering* of $\mathbb{S} \times \mathcal{M}$ if, for any $(x,\sigma) \in \mathbb{S} \times \mathcal{M}$, there exists $(z,\sigma) \in \omega$ such that $|z^\top x| \geq \cos(\theta)$ where $\theta = \delta^{-1}(\epsilon)$.

The following lemma shows probabilistic properties of the sample set ω_N, which will be needed below in order to achieve formal guarantees on the controller.

Lemma 2: Given $N \in \mathbb{Z}^+$, let $\omega_N = \{(x_i,\sigma_i)\}_{i=1}^N$ be i.i.d. with respect to the uniform distribution $\mathbb{P}$ over $\mathbb{S} \times \mathcal{M}$. Then, given any $\epsilon \in (0,1)$, with probability no smaller than $1 - \mathcal{B}(\epsilon; N)$, ω_N is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$, where

$$\mathcal{B}(\epsilon; N) := \frac{M\left(1 - \dfrac{\delta\left(\frac{1}{2}\delta^{-1}(\epsilon)\right)}{M}\right)^N}{\delta\left(\frac{1}{4}\delta^{-1}(\epsilon)\right)}. \tag{17}$$

Proof: Consider a maximal ϵ'-packing Z of $\mathbb{S}$ with $\epsilon' = \delta(\frac{1}{2}\delta^{-1}(\epsilon))$ and let $\theta = \delta^{-1}(\epsilon') = \frac{1}{2}\delta^{-1}(\epsilon)$. From the proof of Lemma 1, $\{\mathrm{Cap}(z,\theta)\}_{z\in Z}$ covers $\mathbb{S}$. Suppose ω_N is sampled randomly according to the uniform distribution, then the probability that each set in $\{\mathrm{Cap}(z,\theta)\}_{z\in Z}$ contains M points with M different modes is no smaller than $1 - \mathcal{N}_p(\epsilon')M(1 - \frac{\epsilon'}{M})^N \geq 1 - \mathcal{B}(\epsilon; N)$. When this happens, for any $(x,\sigma) \in \mathbb{S} \times \mathcal{M}$, there exists a pair $(z,\sigma) \in \omega_N$ such that $|x^\top z| \geq \cos(2\theta)$. This completes the proof. ■

Remark 2: A similar result to Lemma 2 can be obtained if we assume that $\{\sigma_i\}_{i=1}^N$ are i.i.d. and independent of $\{x_i\}_{i=1}^N$, and that every $z \in Z$ satisfies that $\mathrm{Cap}(z,\theta)$ contains at least K points from $\{x_i\}_{i=1}^N$, where Z and θ are as in the proof of Lemma 2 and $K > 0$. Namely, in that case $\mathcal{B}$ in (17) becomes $\mathcal{B}(\epsilon; K) = \frac{M(1-\frac{1}{M})^K}{\delta(\frac{1}{4}\delta^{-1}(\epsilon))}$. For the sake of brevity, the details of the derivation are omitted. This result may be of practical interest, because the set $\{x_i\}_{i=1}^N$ is observed so that we can check whether it satisfies the constraints above. Furthermore, in some cases, the control input can be used to induce that $\{x_i\}_{i=1}^N$ is a good covering of $\mathbb{S}$ (i.e., satisfies the constraints with small ϵ and large K).

We are now able to present a key result of this section, which provides a stability certificate from the solution of the sampled problem (8) assuming sufficient covering by the sample set.

Theorem 1: Given a sample set $\omega_N \subset \mathbb{S} \times \mathcal{M}$, consider problem (8). Let (γ, P, K) be a feasible solution to problem (8). Suppose that ω_N is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$ for some $\epsilon \in (0,1)$. Then

$$\rho(\mathcal{A}_K) \leq \frac{\gamma}{\max\{\varphi_P(\epsilon), \psi_P(\epsilon)\}} \tag{18}$$

where $\rho(\mathcal{A}_K)$ is defined in (3) and

$$\varphi_P(\epsilon) := 1 - \kappa(P)\left(1 - \cos\left(\delta^{-1}(\epsilon)\right)\right) \tag{19}$$

$$\psi_P(\epsilon) := \cos\left(\delta_v^{-1}\left(1 - \sqrt{\frac{\det(P)}{\lambda_{\max}(P)^n}}\cos\left(\delta^{-1}(\epsilon)\right)^n\right)\right) \tag{20}$$

with $\delta(\cdot)$ and $\delta_v(\cdot)$ being given in (13) and (15), respectively.

Proof: We drop the subscript N in ω_N in the proof for convenience. Let $P = L^\top L$ be the Cholesky decomposition of P, and let

$$\tilde{\omega} := \left\{\left(\frac{Lz}{\|Lz\|}, \sigma\right) : (z,\sigma) \in \omega\right\} \subset \mathbb{S} \times \mathcal{M}. \tag{21}$$

We first show that $\rho(\mathcal{A}_K) \leq \frac{\gamma}{\varphi_P(\epsilon)}$. The proof is divided into two steps.

Step 1: We show that if ω is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$, then $\tilde{\omega}$ is an $\tilde{\epsilon}$-covering of $\mathbb{S} \times \mathcal{M}$ for some $\tilde{\epsilon} > 0$ defined below, that is, for any $(\tilde{x},\sigma) \in \mathbb{S} \times \mathcal{M}$, we want to show that there exists $(\tilde{z},\sigma) \in \tilde{\omega}$ such that $|\tilde{z}^\top \tilde{x}| \geq \cos(\tilde{\theta})$ where $\tilde{\theta} = \delta^{-1}(\tilde{\epsilon})$. Note that any $\tilde{x} \in \mathbb{S}$ can be uniquely expressed as $\tilde{x} = Lx/\|Lx\|$ for some $x \in \mathbb{S}$. Let $\tilde{x} = Lx/\|Lx\| \in \mathbb{S}$. Since ω is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$, from the definition, there exists $(z,\sigma) \in \omega$ such that $|x^\top z| \geq \cos(\theta)$ where $\theta = \delta^{-1}(\epsilon)$, which implies that $\|x - z\| \leq \sqrt{2 - 2\cos(\theta)}$ or $\|x + z\| \leq \sqrt{2 - 2\cos(\theta)}$. Now, let us look at the value $|\tilde{x}^\top \tilde{z}|$ where $\tilde{z} = Lz/\|Lz\| \in \tilde{\omega}$. Without loss of generality, we consider the case that $\|x - z\| \leq \sqrt{2 - 2\cos(\theta)}$. Hence

$$
\begin{aligned}
\frac{|(Lx)^\top Lz|}{\|Lx\|\|Lz\|} &= \frac{|x^\top Pz|}{\|Lx\|\|Lz\|} \\
&= \frac{|x^\top Px + z^\top Pz - (x-z)^\top P(x-z)|}{2\|Lx\|\|Lz\|} \\
&\geq \frac{x^\top Px + z^\top Pz - |(x-z)^\top P(x-z)|}{2\sqrt{x^\top Px}\sqrt{z^\top Pz}} \\
&\geq 1 - \frac{|(x-z)^\top P(x-z)|}{2\sqrt{x^\top Px}\sqrt{z^\top Pz}} \qquad (22) \\
&\geq 1 - \frac{\|x-z\|^2 \lambda_{\max}(P)}{2\lambda_{\min}(P)} \qquad (23) \\
&\geq 1 - \kappa(P)(1 - \cos(\theta)) = \varphi_P(\epsilon) \qquad (24)
\end{aligned}
$$

where (22) follows from the fact that $x^\top Px + z^\top Pz \geq 2\sqrt{x^\top Px}\sqrt{z^\top Pz}$ and (23) is a direct consequence of the inequality $\lambda_{\min}(P)I \preceq P \preceq \lambda_{\max}(P)I$. Hence, $\tilde{\omega}_N$ is a $\tilde{\epsilon}$-covering of $\mathbb{S} \times \mathcal{M}$ with $\tilde{\epsilon} = \delta(\tilde{\theta})$ and $\cos(\tilde{\theta}) = 1 - \kappa(P)(1 - \cos(\theta))$.

Step 2: Now, we are in a position to show $\rho(\mathcal{A}_K) \leq \frac{\gamma}{\varphi_P(\epsilon)}$. Let us define

$$\tilde{\omega}^\sigma := \{x : (x, \sigma) \in \tilde{\omega}\} \quad \forall \sigma \in \mathcal{M}. \qquad (25)$$

From Definition 3, we know that $\tilde{\omega}^\sigma$ is a $\tilde{\epsilon}$-covering of $\mathbb{S}$ for all $\sigma \in \mathcal{M}$. This implies that $\cos(\tilde{\theta})\mathbb{B} \subseteq \mathrm{conv}(\pm\tilde{\omega}^\sigma)$ for all $\sigma \in \mathcal{M}$, where $\pm\tilde{\omega}^\sigma$ denotes the union of $\tilde{\omega}^\sigma$ and $-\tilde{\omega}^\sigma$, i.e., $\tilde{\omega}^\sigma \cup -\tilde{\omega}^\sigma$, and $\mathrm{conv}(\cdot)$ denotes the convex hull of a set. Hence

$$\cos\left(\tilde{\theta}\right)\mathbb{B} \subseteq \bigcap_{\sigma \in \mathcal{M}} \mathrm{conv}(\pm\tilde{\omega}^\sigma). \qquad (26)$$

Let $\tilde{A}_\sigma := LA_\sigma L^{-1} + LBKL^{-1}$ for all $\sigma \in \mathcal{M}$. It holds that for all $\sigma \in \mathcal{M}$ and $z \in \pm\tilde{\omega}^\sigma$, $\|\tilde{A}_\sigma z\| \leq \gamma\|z\|$. This, together with (26), implies that $\forall\, \sigma \in \mathcal{M}$

$$\frac{\cos\left(\tilde{\theta}\right)}{\gamma}\tilde{A}_\sigma\mathbb{B} \subseteq \frac{1}{\gamma}\tilde{A}_\sigma \mathrm{conv}(\pm\tilde{\omega}^\sigma) \subseteq \frac{1}{\gamma}\mathrm{conv}\left(\pm\tilde{A}_\sigma\tilde{\omega}^\sigma\right) \subseteq \mathbb{B}.$$

As a consequence, we obtain that $\forall \sigma \in \mathcal{M}$

$$(A_\sigma + BK)^\top P(A_\sigma + BK) \preceq \left(\frac{\gamma}{\cos\left(\tilde{\theta}\right)}\right)^2 P. \qquad (27)$$

Finally, by combining (27) with Proposition 1, we get that $\frac{\gamma}{\cos(\tilde{\theta})} = \frac{\gamma}{\varphi_P(\epsilon)}$ is an upper bound on $\rho(\mathcal{A}_K)$.

We then prove that $\rho(\mathcal{A}_K) \leq \frac{\gamma}{\psi_P(\epsilon)}$. Similarly, let $\omega^\sigma := \{x : (x, \sigma) \in \omega\}\, \forall \sigma \in \mathcal{M}$. We consider an arbitrary $\sigma \in \mathcal{M}$. Since ω^σ is an ϵ-covering of $\mathbb{S}$, $\cos(\delta^{-1}(\epsilon))\mathbb{B} \subseteq \mathrm{conv}(\pm\omega^\sigma)$. Hence

$$\mu_L(\mathrm{conv}(\pm\omega^\sigma)) \geq \cos\left(\delta^{-1}(\epsilon)\right)^n \lambda(\mathbb{B}) \qquad (28)$$

where $\mu_L(\cdot)$ denotes the Lebesgue measure. Note that $\tilde{\omega}^\sigma$ as defined in (25) can be expressed as $\tilde{\omega}^\sigma = \{\frac{Lz}{\|Lz\|} : z \in \omega^\sigma\}$, which leads to the following relation:

$$\mathrm{conv}(\pm\tilde{\omega}^\sigma) \supseteq \frac{L}{\sqrt{\lambda_{\max}(P)}}\mathrm{conv}(\pm\omega^\sigma). \qquad (29)$$

Combining (28) and (29) yields

$$\frac{\mu_L(\mathrm{conv}(\pm\tilde{\omega}^\sigma))}{\mu_L(\mathbb{B})} \geq \sqrt{\frac{\det(P)}{\lambda_{\max}(P)^n}}\cos(\delta^{-1}(\epsilon))^n \qquad (30)$$

which implies that

$$\frac{\mu_L(\mathbb{B} \setminus \mathrm{conv}(\pm\tilde{\omega}^\sigma))}{\mu_L(\mathbb{B})} \leq 1 - \sqrt{\frac{\det(P)}{\lambda_{\max}(P)^n}}\cos(\delta^{-1}(\epsilon))^n. \qquad (31)$$

We claim that the distance from $\partial(\mathrm{conv}(\pm\tilde{\omega}^\sigma))$ to the origin is bounded from below by $\psi_P(\epsilon) = \cos(\delta_v^{-1}(1 - \sqrt{\frac{\det(P)}{\lambda_{\max}(P)^n}}\cos(\delta^{-1}(\epsilon))^n))$. We go by contradiction. Suppose there exists a point $x \in \partial(\mathrm{conv}(\pm\tilde{\omega}^\sigma))$ such that $\|x\| < \psi_P(\epsilon)$. Then, there exists a hyperplane $h^\top x = 1$ such that $\frac{1}{\|h\|} < \psi_P(\epsilon)$ and $h^\top x \leq 1$ for any $x \in \mathrm{conv}(\pm\tilde{\omega}^\sigma)$. By symmetry, it also holds that $-h^\top x \leq 1$ for any $x \in \mathrm{conv}(\pm\tilde{\omega}^\sigma)$. Let $\tilde{B}$ denote the set $\{x \in \mathbb{B} : h^\top x \geq 1\} \cup \{x \in \mathbb{B} : -h^\top x \geq 1\}$. By construction, $\tilde{B} \subseteq \mathbb{B} \setminus \mathrm{conv}(\pm\tilde{\omega}^\sigma)$, which means that $\mu_L(\tilde{B}) \leq (1 - \sqrt{\frac{\det(P)}{\lambda_{\max}(P)^n}}\cos(\delta^{-1}(\epsilon))^n)\mu_L(\mathbb{B})$ from (31). We recall from [46] that the volume of $\tilde{B}$ is $\mathcal{I}(1 - \frac{1}{\|h\|^2}; \frac{n+1}{2}, \frac{1}{2})\mu_L(\mathbb{B})$, which from the condition that $\frac{1}{\|h\|} < \psi_P(\epsilon)$, implies that $\mu_L(\tilde{B}) > \mathcal{I}(1 - \psi_P(\epsilon)^2; \frac{n+1}{2}, \frac{1}{2})\mu_L(\mathbb{B}) = (1 - \sqrt{\frac{\det(P)}{\lambda_{\max}(P)^n}}\cos(\delta^{-1}(\epsilon))^n)\mu_L(\mathbb{B})$, where the equality is by the definition of $\psi_P(\epsilon)$ in (20). This leads to a contradiction. Therefore, we conclude that $\psi_P(\epsilon)\mathbb{B} \subseteq \mathrm{conv}(\pm\tilde{\omega}^\sigma)$. As σ is chosen arbitrarily, this implies that

$$\frac{\psi_P(\epsilon)}{\gamma}\tilde{A}_\sigma\mathbb{B} \subseteq \frac{1}{\gamma}\tilde{A}_\sigma\mathrm{conv}(\pm\tilde{\omega}^\sigma) \subseteq \frac{1}{\gamma}\mathrm{conv}\left(\pm\tilde{A}_\sigma\tilde{\omega}^\sigma\right) \subseteq \mathbb{B}.$$

Thus, $\forall \sigma \in \mathcal{M}$

$$(A_\sigma + BK)^\top P(A_\sigma + BK) \preceq \left(\frac{\gamma}{\psi_P(\epsilon)}\right)^2 P. \qquad (32)$$

Putting (27) and (32) together, we arrive at (18). ■

The result in Theorem 1 allows to establish a probabilistic stability certificate from the solution of the sampled problem (8).

Corollary 1: Given $N \in \mathbb{Z}^+$, let $\omega_N = \{(x_i, \sigma_i)\}_{i=1}^N$ be i.i.d. with respect to the uniform distribution $\mathbb{P}$ over $\mathbb{S} \times \mathcal{M}$. Let $\gamma(\omega_N), P(\omega_N)$, and $K(\omega_N)$ be obtained from Algorithm 1. Then, for any $\epsilon \in (0,1)$, with probability no smaller than $1 - \mathcal{B}(\epsilon; N)$

$$\rho(\mathcal{A}_{K(\omega_N)}) \leq \frac{\gamma(\omega_N)}{\max\left\{\varphi_{P(\omega_N)}(\epsilon), \psi_{P(\omega_N)}(\epsilon)\right\}} \qquad (33)$$

where $\rho(\mathcal{A}_{K(\omega_N)})$ is defined in (3) with $K = K(\omega_N)$ and $\mathcal{B}(\epsilon; N)$, $\varphi_{P(\omega_N)}(\epsilon)$ and $\psi_{P(\omega_N)}(\epsilon)$ are given in (17), (19), and (20), respectively.

Proof: From Lemma 2, with probability no smaller than $1 - \mathcal{B}(\epsilon; N)$, ω_N is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$. Combining this with Theorem 1, we obtain the result above, since Algorithm 1 always generates feasible iterations. ∎

Remark 3: The results above bear some similarities with the probabilistic stability guarantees in [23], [24], and [25], which are concerned with autonomous systems, the major difference is that the bound in this article is applicable to any feasible solution while [23], [24], [25] rely on the optimality of the solution. Let us also highlight that the bound in Theorem 1 can be considered as an improvement to the one in [33] in the sense that the additional term $\psi_P(\epsilon)$ prevents the bound from going unbounded when $\kappa(P)(1-\cos(\delta^{-1}(\epsilon))) \geq 1$, which happens when ϵ is not sufficiently small. From a practical point of view, ϵ has to be small for the bound in (18) to be meaningful. In such cases, $\varphi_P(\epsilon)$ is often larger than $\psi_P(\epsilon)$. Hence, the bound that is really used in practice is $\frac{\gamma}{\varphi_P(\epsilon)}$ in (18).

Remark 4: In some practice situations, what we receive is a set of input-state data, i.e., $\{(x_i, u_i, x_i^+) : i = 1, 2, \ldots, N\}$ where $x_i^+ = A_{\sigma_i} x_i + B u_i$ and u_i is the ith input. As B is known, we can convert this data set into $\{(x_i, x_i^+ - Bu_i) : i = 1, 2. \cdots, N\}$. We can then apply our approach on this converted dataset. Furthermore, the states may not lie on the unit sphere. While the solution of the sampled problem in (8) does not change from a theoretical point of view, we can use the scaled data $\{(x_i/\|x_i\|, (x_i^+ - Bu_i)/\|x_i\|) : i = 1, 2. \cdots, N\}$ to improve numerical stability. If the samples follow an isotropic Gaussian distribution centered at zero (with the covariance matrix being a scalar variance multiplied by the identity matrix) and are generated independently, the scaled points are uniformly distributed on the unit sphere and hence our probabilistic guarantees in Corollary 1 are still valid. We are also able to deal with noisy data following the idea in [48] provided that the noise is bounded.

D. Parallel Scheme for Quadratic Stabilization

To achieve high confidence in Theorem 3, a large number of samples are typically needed. As a result, the problems (11) and (12) at each iteration of Algorithm 1 quickly become demanding due to an increasing number of constraints. To circumvent this issue, inspired by the stochastic gradient descent and its variants [49], we propose a parallelized scheme for Algorithm 1 by dividing these constraints into small batches. More precisely, given $L \in \mathbb{Z}^+$, we build L disjoint subsets of ω_N, denoted by $\{\omega_N^i\}_{i=1}^L$, with $\cup_{i=1}^L \omega_N^i = \omega_N$. The choice of L and $\{\omega_N^i\}_{i=1}^L$ depends on the number of computing resources available and their computation power. With this partition, we then solve the alternating minimization problems as defined in (11) and (12) individually for each batch, as shown in Algorithm 2. The solution of each subproblem provides a candidate descent direction. We then choose the solution that provides the lowest convergence rate via a line search heuristic (34) and (35), which guarantees feasibility of the iterate for all the constraints. The line search step also guarantees that $\{\gamma_k\}$ does not increase along iterations. Note that (34) and (35) are both scalar optimization problems that can be easily solved. The probabilistic guarantee in Corollary 1 is still applicable as it only requires a feasible solution, though Algorithm 2 may produce a more conservative solution $\gamma(\omega_N)$ compared with Algorithm 1.

Algorithm 2: Parallel Alternating Minimization for Quadratic Stabilization.

Input: $\{(x_i, A_{\sigma_i} x_i)\}_{i=1}^N$, B, and some tolerance $\epsilon_{tol} > 0$
Output: $\gamma(\omega_N)$, $P(\omega_N)$, and $K(\omega_N)$

1: *Initialization:* Create a partition $\{\omega_N^\ell\}_{\ell=1}^L$ on ω_N; Let $k \leftarrow 0$, $K_k \leftarrow 0$, $P_k \leftarrow I$, and $\gamma_k \leftarrow \max_{(x,\sigma)\in\omega_N} \frac{\|A_\sigma x\|}{\|x\|}$;
Minimization on P ....
2: **for** $\ell = 1, 2, \ldots, L$ **do**
3: Solve $\bar{\mathcal{P}}(\omega_N^\ell; K_k)$ and let the solution be denoted by P_k^ℓ;
4: Compute

$$\bar{\gamma}_k^\ell \leftarrow \min_{\lambda\in[0,1]} \max_{(x,\sigma)\in\omega_N} \frac{\|A_\sigma x + BK_k x\|_{\bar{P}_k^\ell(\lambda)}}{\|x\|_{\bar{P}_k^\ell(\lambda)}} \tag{34}$$

where $\bar{P}_k^\ell(\lambda) := (1-\lambda)P_k + \lambda P_k^\ell$ and let $\bar{\lambda}_k^\ell$ be the solution of (34).
5: **end for**
6: Find the minimum among $\{\bar{\gamma}_k^\ell\}$ and let $\bar{\ell}_k := \arg\min_\ell \bar{\gamma}_k^\ell$;
7: Let $P_{k+1} \leftarrow \bar{P}_k^{\bar{\ell}_k}(\bar{\lambda}_k^{\bar{\ell}_k})$;
Minimization on K ....
8: **for** $\ell = 1, 2, \ldots, L$ **do**
9: Solve $\hat{\mathcal{P}}(\omega_N^\ell; P_{k+1})$ and let the solution be denoted by K_k^ℓ;
10: Compute

$$\hat{\gamma}_k^\ell \leftarrow \min_{\lambda\in[0,1]} \max_{(x,\sigma)\in\omega_N} \frac{\|A_\sigma x + B\hat{K}_k^\ell(\lambda)x\|_{P_{k+1}}}{\|x\|_{P_{k+1}}} \tag{35}$$

where $\hat{K}_k^\ell(\lambda) := (1-\lambda)K_k + \lambda K_k^\ell$, and let $\hat{\lambda}_k^\ell$ denote the solution of (35);
11: **end for**
12: Find the minimal among $\{\hat{\gamma}_k^\ell\}$ and let $\hat{\ell}_k := \arg\min_\ell \hat{\gamma}_k^\ell$;
13: Let $\gamma_{k+1} \leftarrow \hat{\gamma}_k^{\hat{\ell}_k}$ and $K_{k+1} \leftarrow \mathcal{K}_k^{\hat{\ell}_k}(\hat{\lambda}_k^{\hat{\ell}_k})$;
Stopping criterion ....
14: **if** $\|\gamma_{k+1} - \gamma_k\| < \epsilon_{tol}$ **then**
15: $\gamma(\omega_N) \leftarrow \gamma_{k+1}$, $P(\omega_N) \leftarrow P_{k+1}$, $K(\omega_N) \leftarrow K_{k+1}$, and terminate;
16: **else**
17: Let $k \leftarrow k+1$ and go to Step 2.
18: **end if**

IV. SOS Lyapunov Framework

Quadratic stabilization of switched systems can be very restrictive in some cases. To reduce conservatism, we can use SOS techniques, which have already been used in [50] and [51] to improve the bound on the JSR. In the framework of data-driven stability analysis, the application of SOS optimization has already proven useful for the case of autonomous systems in [24]. Here, we want to show that SOS techniques are also applicable for the stabilization problem.

Let us first recall some definitions in SOS optimization. We refer to [50] for the details. Given $x \in \mathbb{R}^n$ and $d \in \mathbb{Z}^+$, let $x^{[d]} \in$

$\mathbb{R}^{\binom{n+d-1}{d}}$ denote the d-lift of x which consists of all possible monomials of degree d, indexed by all the possible exponents α of degree d

$$x_\alpha^{[d]} = \sqrt{\alpha!}x^\alpha \tag{36}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\sum_{i=1}^n \alpha_i = d$ and $\alpha!$ denotes the multinomial coefficient

$$\alpha! := \frac{d!}{\alpha_1! \cdots \alpha_n!}. \tag{37}$$

The d-lift of a matrix $A \in \mathbb{R}^{n\times n}$ is defined as: $A^{[d]} : x^{[d]} \to (Ax)^{[d]}$. The following proposition provides an upper bound for the JSR based on SOS Lyapunov functions.

Proposition 2 (See [34], Th. 2.13): Consider the closed-loop matrices $\mathcal{A}_K$ for some state feedback $K \in \mathbb{R}^{m\times n}$. For any $d \in \mathbb{Z}^+ (d \geq 1)$, if there exist $\gamma \geq 0$ and $P \succ 0$ such that $\forall\, A \in \mathcal{A}_\mathcal{K}$, $x \in \mathbb{S}$

$$\left((Ax)^{[d]}\right)^\top P(Ax)^{[d]} \leq \gamma^{2\,d} \left(x^{[d]}\right)^\top Px^{[d]} \tag{38}$$

where $P \in \mathbb{R}^{D\times D}$ with $D = \binom{n+d-1}{d}$, then $\rho(\mathcal{A}_K) \leq \gamma$.

A. Data-Driven SOS Stabilization

With SOS Lyapunov functions, we formulate the following sampled problem using the given dataset ω_N:

$$\mathcal{P}_d(\omega_N) : \min_{\gamma\geq 0, P\succeq I, K} \gamma \tag{39a}$$

$$\text{s.t.} \quad \left((A_\sigma x + BKx)^{[d]}\right)^\top P\left(A_\sigma x + BKx\right)^{[d]}$$
$$\leq \gamma^{2\,d}\left(x^{[d]}\right)^\top Px^{[d]} \quad \forall\,(x,\sigma) \in \omega_N. \tag{39b}$$

From the definition of d-lift of vectors above, the variable P above is a $D \times D$ matrix, where $D = \binom{n+d-1}{d}$. Thus, problem (39) is a polynomial optimization problem with $mn + D(D+1)/2 + 1$ variables and N polynomial constraints, which is much more computationally demanding than Problem (8) depending on the degree d. For a small dataset, this problem can be solved by polynomial toolboxes [41], [42], [43] or general nonlinear solvers, such as IPOPT [52] and NLopt [53]. For a large dataset, to reduce the complexity, we again make use of the structure in (39b) to develop an alternating minimization algorithm between P and K as Algorithm 1 for the quadratic case. For a given K, we define

$$\tilde{\mathcal{P}}_d(\omega_N; K) : \min_{\gamma, P\succeq I} \gamma \tag{40a}$$

$$\text{s.t.} \quad \left((A_\sigma x + BKx)^{[d]}\right)^\top P\left(A_\sigma x + BKx\right)^{[d]}$$
$$\leq \gamma^{2\,d}\left(x^{[d]}\right)^\top Px^{[d]} \quad \forall\,(x,\sigma) \in \omega_N. \tag{40b}$$

For a given P, we define

$$\hat{\mathcal{P}}_d(\omega_N; P) : \min_{\gamma\geq 0, K} \gamma \tag{41a}$$

$$\text{s.t.} \quad \left((A_\sigma x + BKx)^{[d]}\right)^\top P\left(A_\sigma x + BKx\right)^{[d]}$$
$$\leq \gamma^{2\,d}\left(x^{[d]}\right)^\top Px^{[d]} \quad \forall\,(x,\sigma) \in \omega_N. \tag{41b}$$

Algorithm 3: Alternating Minimization for SOS Stabilization.

Input: $\{(x_i, x_i^+)\}_{i=1}^N$, B, d and some tolerance $\epsilon_{tol} > 0$
Output: $\gamma_{sos}(\omega_N)$, $P_{sos}(\omega_N)$, and $K_{sos}(\omega_N)$
1: *Initialization:* $k \leftarrow 0$, $K_k \leftarrow 0$, $P_k \leftarrow I$, and $\gamma_k \leftarrow \max_{(x,\sigma)\in\omega_N} \frac{\|(A_\sigma x)^{[d]}\|}{\|x^{[d]}\|}$;
2: Obtain P_{k+1} by solving $\tilde{\mathcal{P}}_d(\omega_N; K_k)$ via bisection on γ starting from $\max_{(x,\sigma)\in\omega_N} \frac{\|(A_\sigma x + BK_k x)^{[d]}\|_{P_k}}{\|x^{[d]}\|_{P_k}}$;
3: Obtain γ_{k+1} and K_{k+1} by solving $\tilde{\mathcal{P}}_d(\omega_N; P_{k+1})$ initialized at $K = K_k$;
4: **if** $\|\gamma_{k+1} - \gamma_k\| < \epsilon_{tol}$ **then**
5: $\quad\gamma_{sos}(\omega_N) \leftarrow \gamma_{k+1}$, $P_{sos}(\omega_N) \leftarrow P_{k+1}$, $K_{sos}(\omega_N) \leftarrow K_{k+1}$;
6: $\quad$Terminate;
7: **else**
8: $\quad$Let $k \leftarrow k+1$ and go to Step 2.
9: **end if**

Indeed, one may observe that problem (40) can be solved efficiently by using SDP solvers (like Mosek [54]) and bisection on γ. While Problem (41) is still a polynomial optimization problem, there are only mn decision variables, which makes it easier to handle than problem (39). To solve the problems in (40) and (41), we typically need an initial solution, in particular for the polynomial problem (41). In this alternating minimization algorithm, we set the initial solution to be the solution from the previous iteration, which leads to a convergent sequence of γ. The details of this procedure is given in Algorithm 3. Similar to the case of quadratic stabilization, Algorithm 3 can be also parallelized following Algorithm 2.

B. Stability Guarantees Via Sensitivity Analysis

With the aforementioned alternating minimization algorithm for SOS stabilization, we get a feasible solution, denoted by $(\gamma_{\text{sos}}(\omega_N), P_{\text{sos}}(\omega_N), K_{\text{sos}}(\omega_N))$. Similar to Section III-C, we now derive stability guarantees for this solution. However, due to the polynomial lifting in (39), we lose convexity in the Lyapunov function, which prevents us from applying the geometric results in Section III-C to the SOS framework. In particular, the reasoning about (27) and (32) in the proof of Theorem 1 does not hold. Hence, we use an alternative way to derive probabilistic guarantees on the solution obtained from Algorithm 3.

Given any $\epsilon \in (0,1)$ and $d \in \mathbb{Z}^+$ $(d \geq 1)$, let us define

$$\phi(\epsilon, d) := \sum_{k=1}^d \binom{d}{k} \left(2 - 2\cos\left(\delta^{-1}(\epsilon)\right)\right)^{\frac{k}{2}}. \tag{42}$$

Similar to Theorem 1, we derive a stability certificate for SOS stabilization as stated in the following theorem.

Theorem 2: Given a sample set $\omega_N \subset \mathbb{S} \times \mathcal{M}$ and an integer $d \geq 1$, consider Problem (39). Let (γ, P, K) be a feasible solution to Problem (39). Suppose ω_N is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$ for some $\epsilon > 0$. Then

$$\rho(\mathcal{A}_K) \leq \sqrt[d]{\gamma^d + (\gamma^d + \bar{\rho}(\mathcal{A}_K)^d)\sqrt{\kappa(P)}\phi(\epsilon, d)} \tag{43}$$

where $\rho(\mathcal{A}_K)$ is defined in (3), $\phi(\epsilon, d)$ is given in (42), and

$$\bar{\rho}(\mathcal{A}_K) := \max_{A\in\mathcal{A}_K} \|A\|. \tag{44}$$

Proof: We drop the subscript N in ω_N in the proof for convenience. Since ω is an ϵ-covering of $\mathbb{S}\times\mathcal{M}$, from Definition 3, for any $(x,\sigma)\in\mathbb{S}\times\mathcal{M}$, there exists $(z,\sigma)\in\omega$ such that $|z^\top x| \geq \cos(\delta^{-1}(\epsilon))$, which implies that $\|x-z\| \leq \sqrt{2-2\cos(\delta^{-1}(\epsilon))}$ or $\|x+z\| \leq \sqrt{2-2\cos(\delta^{-1}(\epsilon))}$. Due to the fact that $(-z,\sigma)$ also satisfies (39b), we consider the case that $\|x-z\| \leq \sqrt{2-2\cos(\delta^{-1}(\epsilon))}$ without loss of generality. From [50], it can be shown that

$$\|x^{[d]} - z^{[d]}\| = \|x^{\otimes d} - z^{\otimes d}\| \leq \sum_{k=1}^{d} \binom{d}{k} \|x-z\|^k \|z\|^{d-k}$$
$$\leq \sum_{k=1}^{d} \binom{d}{k} \left(2 - 2\cos(\delta^{-1}(\epsilon))\right)^{\frac{k}{2}}$$

where $x^{\otimes d}$ denotes the d-fold Kronecker product. With this and some manipulations, we get that

$$\begin{aligned}
&\|(A_\sigma x + BKx)^{[d]}\|_P \\
&\leq \|(A_\sigma z + BKz)^{[d]} + (A_\sigma + BK)^{[d]}\left(x^{[d]} - z^{[d]}\right)\|_P \\
&\leq \|(A_\sigma z + BKz)^{[d]}\|_P \\
&\quad + \sqrt{\lambda_{\max}(P)}\|A_\sigma + BK\|^d \|x^{[d]} - z^{[d]}\| \\
&\leq \gamma^d \|z^{[d]}\|_P + \sqrt{\lambda_{\max}(P)}\|A_\sigma + BK\|^d \phi(\epsilon, d) \\
&\leq \gamma^d \|z^{[d]}\|_P + \sqrt{\lambda_{\max}(P)}\bar{\rho}(\mathcal{A}_K)^d \phi(\epsilon, d) \\
&\leq \gamma^d \left(\|x^{[d]}\|_P + \sqrt{\lambda_{\max}(P)}\phi(\epsilon, d)\right) \\
&\quad + \sqrt{\lambda_{\max}(P)}\bar{\rho}(\mathcal{A}_K)^d \phi(\epsilon, d) \\
&\leq \left(\gamma^d + \left(\gamma^d + \bar{\rho}(\mathcal{A}_K)^d\right)\sqrt{\kappa(P)}\phi(\epsilon, d)\right)\|x^{[d]}\|_P.
\end{aligned}$$

From Proposition 2, we then obtain (43). ■

Remark 5: When $d=1$, it becomes exactly the quadratic case, which means that Theorem 2 provides an alternative bound for the quadratic case. However, the results for the quadratic case in Section III are not applicable to the SOS case. We also observe by numerical simulation that, for reasonable values of P and ϵ ($\kappa(P)$ and ϵ are not too large), the bound in (18) is better than the one in (43) with $d=1$. In addition, Theorem 2 requires the information of $\bar{\rho}(\mathcal{A}_K)$, which is yet to be estimated.

Following the same arguments as in Section III-C, we can then establish probabilistic guarantees for this data-driven SOS framework. However, the bound in (43) relies on $\bar{\rho}(\mathcal{A}_K)$ which is not available. To handle this issue, we use the data set to estimate $\bar{\rho}(\mathcal{A}_K)$. Given a sample set $\omega_N \subset \mathbb{S}\times\mathcal{M}$ and any $K\in\mathbb{R}^{m\times n}$, we define the following problem:

$$\eta^*(\omega_N, K) := \min_{\eta\geq 0} \eta \tag{45a}$$
$$\text{s.t.} \quad \|A_\sigma x + BKx\| \leq \eta \quad \forall (x,\sigma)\in\omega_N. \tag{45b}$$

An upper bound on $\bar{\rho}(\mathcal{A}_K)$ is then given in the following proposition.

Proposition 3: Given a sample set $\omega_N \subset \mathbb{S}\times\mathcal{M}$ and any $K\in\mathbb{R}^{m\times n}$, let $\eta^*(\omega, K)$ be defined as in (45). Suppose ω_N is an ϵ-covering of $\mathbb{S}\times\mathcal{M}$ for some $\epsilon > 0$. Then

$$\bar{\rho}(\mathcal{A}_K) \leq \frac{\eta^*(\omega, K)}{\cos(\delta^{-1}(\epsilon))} \tag{46}$$

where $\bar{\rho}(\mathcal{A}_K)$ is given in (44).

Proof: This result is a special case of Theorem 1 where $P=I$, which implies that $\varphi_P(\epsilon) = \cos(\delta^{-1}(\epsilon))$. ■

Based on the results above, we now present the main result of this section, which is a probabilistic stability certificate from the solution of the sampled problem (39).

Theorem 3: Given $N\in\mathbb{Z}^+$, let ω_N be i.i.d with respect to the uniform distribution $\mathbb{P}$ over $\mathbb{S}\times\mathcal{M}$. Let $(\gamma_{sos}(\omega_N), P_{sos}(\omega_N), K_{sos}(\omega_N))$ be obtained from Algorithm 3 and $\eta^*(\omega_N, K_{sos}(\omega_N))$ be defined as in (45). For any $\epsilon\in(0,1)$, with probability no smaller than $1-\mathcal{B}(\epsilon; N)$

$$\begin{aligned}
&\rho(\mathcal{A}_{K_{\text{sos}}(\omega_N)}) \\
&\leq \gamma_{\text{sos}}(\omega_N)\sqrt[d]{1 + \left(1 + \frac{\bar{\eta}(\omega_N)^d}{\gamma^d_{\text{sos}}(\omega_N)}\right)\sqrt{\kappa(P_{\text{sos}}(\omega_N))}\phi(\epsilon, d)}
\end{aligned} \tag{47}$$

where $\phi(\epsilon, d)$ is given in (42), $\mathcal{B}(\epsilon; N)$ is defined as in (17), and

$$\bar{\eta}(\omega_N) := \frac{\eta^*(\omega_N, K_{sos}(\omega_N))}{\cos(\delta^{-1}(\epsilon))} \tag{48}$$

where $\delta(\cdot)$ denotes the measure of a *symmetric spherical cap* as shown in (13).

Proof: From Lemma 2, ω_N is an ϵ-covering of $\mathbb{S}\times\mathcal{M}$ with probability no smaller than $1-\mathcal{B}(\epsilon; N)$. Combining Theorem 2 and Proposition 3, we obtain the result above. ■

V. Data-Driven Switched LQR

In this section, we show that the proposed data-driven framework can be extended to infinite-horizon LQR problems of arbitrary switched linear systems. We consider the following infinite-horizon quadratic cost

$$J_\infty(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\sigma}) = \sum_{\ell=0}^{\infty} \mathcal{L}(x(\ell), u(\ell)) \tag{49}$$

where $\boldsymbol{x}$, $\boldsymbol{u}$, and $\boldsymbol{\sigma}$ denote the state, control, and switching sequences, respectively, and $\mathcal{L}(x,u) = x^\top Qx + u^\top Ru$ is the stage cost with $Q\succ 0$ and $R\succ 0$. With these definitions, the infinite-horizon LQR problem of System (1) can be cast as

$$J^*(x) := \inf_{\boldsymbol{u}} \sup_{\boldsymbol{\sigma}\in\mathcal{M}^\infty} J_\infty(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\sigma}) \tag{50}$$

with $x(0)=x$. Without any information on the switching signal, we only consider static linear feedback in the infinite-horizon LQR problem. Our goal is to find a quadratic upper bound $x^\top Px$ of $J^*(x)$ with a static feedback $u=Kx$, i.e., we want to find a pair (K,P) such that $J_\infty(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\sigma}) \leq \|x(0)\|_P^2$ for all $\boldsymbol{\sigma}\in\mathcal{M}^\infty$

with $u(\ell) = Kx(\ell)$ for all $\ell \in \mathbb{Z}^+$. When (K, P) satisfies

$$(A + BK)^\top P(A + BK) \preceq P - Q - K^\top RK \tag{51}$$

for all $A \in \mathcal{A}$, following standard manipulations (see, e.g., [55] and [56]), it can be shown that, for any switching sequence and any $k \in \mathbb{Z}^+$

$$\|x(0)\|_P^2 - \sum_{\ell=0}^{k} \mathcal{L}(x(\ell), u(\ell)) \geq \|x(k+1)\|_P^2 \tag{52}$$

with $u(\ell) = Kx(\ell)$, which implies that $J^*(x) \leq \|x\|_P^2$ for all $x \in \mathbb{R}^n$. For linear systems (when $\mathcal{A}$ is a singleton), (K, P) can be obtained by solving the algebraic Riccati equation, see [57, Ch. 2] for details. For multiple dynamics matrices, using the Schur complement formula with $S = P^{-1}$ and $Y = KS$, a model-based solution of (51) can be obtained by solving the following LMI problem, see [58, Th. 1]

$$\min_{S,Y} -\log\det(S) \tag{53a}$$

$$\text{s.t.} \begin{pmatrix} S & SA^\top + Y^\top B^\top & S & Y^\top \\ AS + BY & S & \mathbf{0} & \mathbf{0} \\ S & \mathbf{0} & Q^{-1} & \mathbf{0} \\ Y & \mathbf{0} & \mathbf{0} & R^{-1} \end{pmatrix} \succeq 0$$

$$\forall A \in \mathcal{A}. \tag{53b}$$

A. Sampled LQR Problem

In the case where the dynamics matrices $\mathcal{A}$ are unknown, given a sample set $\omega_N \subset \mathbb{S} \times \mathcal{M}$, a sample-based relaxation of (51) is given as follows:

$$\begin{aligned}&(A_\sigma x + BKx)^\top P(A_\sigma x + BKx)\\ &\leq x^\top \left(P - Q - K^\top RK\right) x \quad \forall (x, \sigma) \in \omega_N.\end{aligned} \tag{54}$$

However, a solution to (54) may not be valid for (51) due to the randomness in the sampling. To deal with this issue, we introduce a scaling parameter $\xi \in (0, 1)$ and formulate the following problem:

$$\min_{P,K} \operatorname{tr}(P) \tag{55a}$$

$$\text{s.t.} \quad (A_\sigma x + BKx)^\top P(A_\sigma x + BKx) \tag{55b}$$

$$\leq \xi^2 x^\top \left(P - Q - K^\top RK\right) x \quad \forall (x, \sigma) \in \omega_N$$

$$P \succeq Q + K^\top RK \tag{55c}$$

where the constraint (55c) is imposed as (55b) does not guarantee that $P - Q - K^\top RK$ is positive semidefinite. When the dataset is sufficiently rich, the problem (55) leads to the following robustness property.

Theorem 4: Given a sample set $\omega_N \subset \mathbb{S} \times \mathcal{M}$, and $\xi \in (0, 1)$, let (P, K) be a feasible solution to problem (55) and $Z = P - Q - K^\top RK$. Suppose ω_N is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$ and Z is invertible. Then

$$\begin{aligned}&(A_\sigma + BK)^\top P(A_\sigma + BK)\\ &\preceq \frac{\xi^2}{\max\{\varphi_Z(\epsilon), \psi_Z(\epsilon)\}^2} Z \quad \forall \sigma \in \mathcal{M}\end{aligned} \tag{56}$$

where $\varphi_Z(\epsilon)$ and $\psi_Z(\epsilon)$ are given in (19) and (20), respectively.

Proof: We drop the subscript N in ω_N in the proof for convenience. The proof follows similar arguments as in Theorem 1. Consider the Cholesky decomposition of $Z = P - Q - K^\top RK = \bar{L}^\top \bar{L}$, define

$$\varpi := \left\{ \left(\frac{\bar{L}z}{\|\bar{L}z\|}, \sigma \right) : (z, \sigma) \in \omega \right\} \subset \mathbb{S} \times \mathcal{M}. \tag{57}$$

From the same reasoning as in the proof of Theorem 1, it can be shown that ϖ is an $\bar{\epsilon}$-covering of $\mathbb{S} \times \mathcal{M}$ with $\bar{\epsilon} = \delta(\bar{\theta})$ and $\cos(\bar{\theta}) = \varphi_Z(\epsilon) = 1 - \kappa(Z)(1 - \cos(\theta))$ where $\theta = \delta^{-1}(\epsilon)$. Again, as in (25), we define

$$\varpi^\sigma := \{x : (x, \sigma) \in \varpi\} \quad \forall \sigma \in \mathcal{M}. \tag{58}$$

Following (26), we have $\varphi_Z(\epsilon)\mathbb{B} \subseteq \bigcap_{\sigma \in \mathcal{M}} \operatorname{conv}(\pm\varpi^\sigma)$. With Cholesky decomposition of $P = L^\top L$, it also holds that $\|\bar{A}_\sigma z\| \leq \xi\|z\|$ for all $\sigma \in \mathcal{M}$ and $z \in \pm\varpi^\sigma$, where $\bar{A}_\sigma := LA_\sigma\bar{L}^{-1} + LBK\bar{L}^{-1}$. Then, with (27), we have, $\forall \sigma \in \mathcal{M}$

$$\frac{\varphi_Z(\epsilon)}{\xi} \bar{A}_\sigma \mathbb{B} \subseteq \frac{1}{\xi} \bar{A}_\sigma \operatorname{conv}(\pm\varpi^\sigma) \subseteq \frac{1}{\xi} \operatorname{conv}(\pm\bar{A}_\sigma\varpi^\sigma) \subseteq \mathbb{B}.$$

Hence, $\|\frac{\varphi_Z(\epsilon)}{\xi}\bar{A}_\sigma x\| \leq \|x\|$ for any $x \in \mathbb{R}^n$ and $\sigma \in \mathcal{M}$, which implies $\bar{A}_\sigma^\top \bar{A}_\sigma \preceq (\frac{\xi}{\varphi_Z(\epsilon)})^2 I$ for any $\sigma \in \mathcal{M}$. Finally, we get

$$(A_\sigma + BK)^\top P(A_\sigma + BK) \preceq \frac{\xi^2}{\varphi_Z(\epsilon)^2} Z \quad \forall \sigma \in \mathcal{M} \tag{59}$$

which implies (56). Following the reasoning above and the arguments in the proof of Theorem 1, we also have

$$(A_\sigma + BK)^\top P(A_\sigma + BK) \preceq \frac{\xi^2}{\psi_Z(\epsilon)^2} Z \quad \forall \sigma \in \mathcal{M}. \tag{60}$$

■

From Theorem 4, we also impose the constraint $\max\{\varphi_{P-Q-K^\top RK}(\epsilon), \psi_{P-Q-K^\top RK}(\epsilon)\} \geq \xi$ in addition to (55b) and (55c) to ensure that the pair (K, P) is a feasible solution to (51). As mentioned in Remark 3, in practice, we only need to consider $\varphi_{P-Q-K^\top RK}(\epsilon)$ as it is often larger for reasonable values of ϵ. Motivated by this fact, we co-design (P, K) and the parameter ξ and modify problem (55) by imposing an additional constraint as follows:

$$\min_{P,K,\xi} \operatorname{tr}(P) \tag{61a}$$

$$\text{s.t.} \quad \text{(55b), (55c)} \tag{61b}$$

$$\kappa\left(P - Q - K^\top RK\right) \leq \frac{1 - \xi}{1 - \cos\left(\delta^{-1}(\epsilon)\right)} \tag{61c}$$

$$0 \leq \xi \leq \cos\left(\delta^{-1}(\epsilon)\right) \tag{61d}$$

where (61c) is a reformulation of $\varphi_{P-Q-K^\top RK}(\epsilon) \geq \xi$, (61d) is due to the fact that $\kappa(P - Q - K^\top RK) \geq 1$, and ϵ is a user-defined parameter. Note that the constraint (61c) is nonconvex.

B. Alternating LQR Design

To solve the nonconvex problem (61), we also develop an alternating minimization algorithm. While the general implementation is similar to the algorithms in the previous sections, the technical details are quite different due to the additional complexity arising from the constraint (61c). As K may not be a stabilizing feedback in the initial steps, the value of ξ can be larger than 1, which means that (61c) is not valid. In view of this, we propose a heuristic in which we relax this constraint with an a priori upper bound $\bar{\kappa}$ on $\kappa(P - Q - K^\top RK)$ and continuously minimize ξ at each iteration. We then check the constraint (61c) when ξ is less than 1. Given any $\epsilon \in (0,1)$ and $\alpha \geq 0$, let us define

$$\xi^*(\epsilon, \alpha) := 1 - \alpha(1 - \cos(\delta^{-1}(\epsilon))). \tag{62}$$

With this definition, it can be verified that, when $\kappa(P - Q - K^\top RK) \leq \bar{\kappa}$ and $\xi \leq \xi^*(\epsilon, \bar{\kappa})$, the constraint (61c) is satisfied.

We now present the overall procedure. At the initialization, we find the value of ξ that is closest to 1 such that (55b) is feasible with $K = \mathbf{0}$, denoted by ξ_0. Note that, given the a priori upper bound $\bar{\kappa}$, the constraint $\kappa(P - Q - K^\top RK) \leq \bar{\kappa}$ can be rewritten as $\lambda_{\max}(P - Q - K^\top RK) \leq \bar{\kappa}\nu$ and $\lambda_{\min}(P - Q - K^\top RK) \geq \nu$ for some $\nu > 0$, which are equivalent to $\nu I \preceq P - Q - K^\top RK \preceq \bar{\kappa}\nu I$. We then proceed by optimizing (K, ξ) and P in an alternating way. In the optimization of (K, ξ), the goal is to decrease ξ to a certain value below 1 but not to minimize ξ as much as possible. For this reason, we penalize the distance of the current ξ to the previous value at each iteration in order to generate a smooth sequence $\{\xi_k\}$. Even when P is fixed, the constraints (55b) and (55c) are still nonlinear. Again, we use the Schur complement formula to convert these constraints into LMIs as follows:

$$\begin{pmatrix} x^\top Px - x^\top Qx & (A_\sigma x + BKx)^\top P & x^\top K^\top \\ P(A_\sigma x + BKx) & \xi^2 P & \mathbf{0} \\ Kx & \mathbf{0} & R^{-1} \end{pmatrix} \succeq 0 \quad \forall (x, \sigma) \in \omega_N \tag{63}$$

$$\begin{pmatrix} P - Q & K^\top \\ K & R^{-1} \end{pmatrix} \succeq 0. \tag{64}$$

The optimization problem involving (K, ξ) is formulated in (66). Replacing ξ^2 with a new variable ζ yields a convex problem. To ensure the constraint $\kappa(P - Q - K^\top RK) \leq \bar{\kappa}$ at each iteration, the optimization of (K, ξ) is followed by a backtracking step. After K and ξ are updated, we optimize over P subject to the constraints (55b), (55c), and (65c). The details are given in Algorithm 4. This algorithm can also be parallelized following the same idea as in Algorithm 2.

Based on Theorem 4, a probabilistic guarantee is presented below for the proposed data-driven LQR.

Corollary 2: Consider Problem (51) with $Q, R \succ 0$. Given $N \in \mathbb{Z}^+$, let $\omega_N = \{(x_i, \sigma_i)\}_{i=1}^N$ be i.i.d. with respect to the uniform distribution $\mathbb{P}$ over $\mathbb{S} \times \mathcal{M}$. For any $\epsilon \in (0,1)$, we define $\xi^*(\epsilon)$ as in (62). Let $(\xi(\omega_N), P(\omega_N), K(\omega_N))$ be a solution obtained from Algorithm 4 and $Z(\omega_N) := P(\omega_N) - Q - K(\omega_N)^\top RK(\omega_N)$. Then, with probability no smaller than $1 - \mathcal{B}(\epsilon; N)$, the following statement holds: if $\xi(\omega_N) \leq \xi^*(\epsilon, \kappa(Z(\omega_N)))$, $(P(\omega_N), K(\omega_N))$ is a feasible solution to (51).

Proof: From Lemma 2, with probability no smaller than $1 - \mathcal{B}(\epsilon; N)$, ω_N is an ϵ-covering of $\mathbb{S} \times \mathcal{M}$. Then, according to Theorem 4, when $\xi(\omega_N) \leq \xi^*(\epsilon, \kappa(Z(\omega_N)))$, $(P(\omega_N), K(\omega_N))$ is a feasible solution. ■

Algorithm 4: Alternating Minimization for LQR Design.

Input: $\{(x_i, x_i^+)\}_{i=1}^N$, B, Q, R, $\bar{\kappa}$, c, ϵ and ϵ_{tol}
Output: $\xi(\omega_N)$, $P(\omega_N)$, and $K(\omega_N)$
1: *Initialization:* Let $k \leftarrow 0$ and $K_k \leftarrow \mathbf{0}$; Obtain P_k and ξ_k by solving

$$\min_{P,\xi} \xi \tag{65a}$$
$$\text{s.t.} \quad (55b), (55c), \tag{65b}$$
$$\kappa(P - Q - K^\top RK) \leq \bar{\kappa} \tag{65c}$$
$$\xi \geq 1 \tag{65d}$$

with $K = K_k$.
2: Solve

$$(\bar{K}_k, \bar{\xi}_k) \leftarrow \arg\min_{K,\xi} \xi^2 + c(\xi^2 - \xi_k^2)^2 \tag{66a}$$
$$\text{s.t.} \quad (63), (64), \tag{66b}$$
$$\xi \geq \xi^*(\epsilon, \bar{\kappa}), \tag{66c}$$

with $P = P_k$;
3: Obtain the stepsize via backtracking

$$\lambda_k \leftarrow \arg\min_{\lambda \in [0,1]} \lambda : \tag{67a}$$
$$\text{s.t.} \quad \kappa(P_k - Q - \mathcal{K}_k(\lambda)^\top R\mathcal{K}_k(\lambda)) \leq \bar{\kappa} \tag{67b}$$

where $\mathcal{K}_k(\lambda) := \lambda K_k + (1 - \lambda)\bar{K}_k$.
4: Let $K_{k+1} = \lambda_k K_k + (1 - \lambda_k)\bar{K}_k$ and $\xi_{k+1} = \sqrt{\lambda_k \xi_k^2 + (1 - \lambda_k)\bar{\xi}_k^2}$;
5: Obtain P_{k+1} by solving

$$\min_P \operatorname{trace}(P) \tag{68a}$$
$$\text{s.t.} \quad (55b), (55c), (65c) \tag{68b}$$

with $K = K_{k+1}$ and $\xi = \xi_{k+1}$.
6: **if** $\|\xi_{k+1} - \xi_k\| < \epsilon_{tol}$ or $\kappa(P_{k+1} - Q - K_{k+1}^\top RK_{k+1}) \leq \frac{1 - \xi_{k+1}}{1 - \cos(\delta^{-1}(\epsilon))}$ **then**
7: $\xi(\omega_N) \leftarrow \xi_{k+1}$, $P(\omega_N) \leftarrow P_{k+1}$, $K(\omega_N) \leftarrow K_{k+1}$;
8: Terminate;
9: **else**
10: Let $k \leftarrow k + 1$ and go to Step 1.
11: **end if**

Remark 6: From (66), the sequence $\{\xi_k\}$ obtained from Algorithm 4 is monotonically nonincreasing. With the constraint in (65d), the sequence $\{\xi_k\}$ is bounded from below and thus is convergent.

VI. Numerical Experiments

In this section, we demonstrate the proposed data-driven control framework on several numerical examples.

A. Quadratic Stabilization for Multiple Examples

We first consider an example that can be quadratically stabilized. Consider the following switched linear system with $n = 2, m = 1$ and $M = 3$

$$A_1 = \begin{pmatrix} 0.7 & 0.16 \\ 1.1 & -1.1 \end{pmatrix}, A_2 = \begin{pmatrix} 0.4 & -0.84 \\ 0.83 & 0.35 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 0.37 & 0.96 \\ 0.34 & -1.2 \end{pmatrix}, B = \begin{pmatrix} -0.9 \\ -1.2 \end{pmatrix}.$$

To show that this is not a trivial example, we compute $\rho(\mathcal{A}) = 1.544$, the JSR of the open-loop system, using the JSR toolbox [59]. We also compute $\gamma^* = 0.8756$ as defined in (4) by solving (5) with bisection on γ to check feasibility of quadratic stabilization for this example.

First, let $N = 2000$ and set the confidence level to $\mathcal{B}(\epsilon; N) = 0.01$. The corresponding value of ϵ can be computed via bisection using (17). With this setting, the upper bound in (18) is valid with probability larger than 99%. For convenience, let

$$\overline{\gamma}(\omega_N) := \frac{\gamma(\omega_N)}{1 - \kappa(P(\omega_N))(1 - \cos(\delta^{-1}(\epsilon)))}.$$

We then sample ω_N according to the uniform distribution on $\mathbb{S} \times \mathcal{M}$ and apply Algorithm 1 with the tolerance being $\epsilon_{tol} = 10^{-3}$. The obtained solution is

$$\gamma(\omega_N) = 0.8836,\ K(\omega_N) = \begin{pmatrix} 0.6626 & -0.4359 \end{pmatrix}$$

$$P(\omega_N) = \begin{pmatrix} 1.1302 & 0.5480 \\ 0.5480 & 3.3064 \end{pmatrix}.$$

The bound obtained from (18) is $\overline{\gamma}(\omega_N) = 0.8873$. To empirically verify the solution, we compute $\rho(\mathcal{A}_{K(\omega_N)}) = 0.8767$, the JSR of the closed-loop system with $K(\omega_N)$, using the JSR toolbox [59].

We then apply the proposed approach to higher dimensional examples. Again, the confidence level is set to 0.01, i.e., $\mathcal{B}(\epsilon; N) = 0.01$. We choose different values of N and compute the corresponding ϵ that satisfies $\mathcal{B}(\epsilon; N) = 0.01$. The dynamics matrices $\mathcal{A}$ and the input matrix B are generated randomly in a way that each entry is chosen from the uniform distribution over $[-1, 1]$. Note that these random examples may not be stabilizable by a static linear feedback. Hence, in the simulation, we only compute the upper bound $\overline{\gamma}(\omega_N)$ and compare it to the true solution γ^* defined in problem (6). Finally, we divide the dataset into a number of subsets of size 1000 and use the parallelized scheme in Algorithm 2. The results are shown in Fig. 2. As expected, the sample size needed to reach the true white-box solution increases as the system dimension and the number of modes increase.

B. Quadratic Stabilization Versus SOS Stabilization

For stability analysis of switched linear systems, SOS Lyapunov functions often provide tighter bounds than quadratic Lyapunov functions, see [50] and [51] for a few examples. Recently, we have also found out that it is beneficial to use SOS

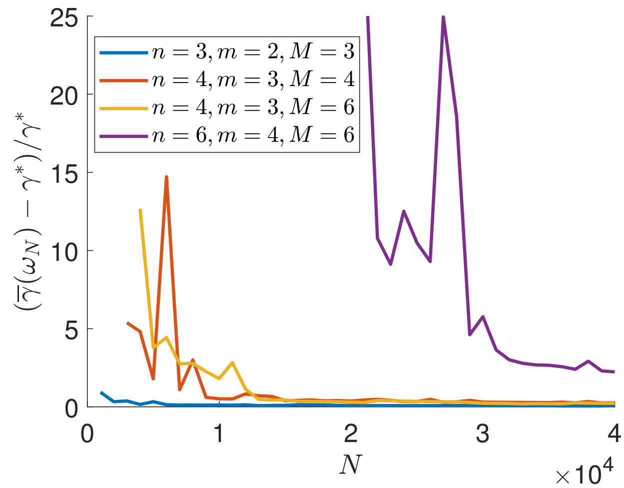


Fig. 2. Convergence of the sample-based solution to the true solution for systems of different dimensions and modes.

Lyapunov functions in data-driven stability analysis in [24] with numerical examples. Here, we show that this is also the case in the stabilization problem. More importantly, we show how SOS stabilization performs compared with quadratic stabilization as the sample size increases. Consider the following switched linear system with two modes, which is generated randomly with a procedure similar as above

$$A_1 = \begin{pmatrix} -1.6856 & -0.1665 \\ 0.7785 & -1.6321 \end{pmatrix}, B = \begin{pmatrix} 0.1975 \\ 0.8640 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} -0.2915 & -3.2824 \\ 3.9761 & -0.02274 \end{pmatrix}.$$

For different values of N, let ϵ be chosen such that $\mathcal{B}(\epsilon; N) = 0.01$. We then compute the upper bounds on the JSR according to the results in Sections III-C and IV-B. For SOS stabilization, we consider the case of $d = 2$. The results are given in Fig. 3, which shows that the JSR upper bound from SOS stabilization becomes tighter as N increases. When $N = 25\,000$, the solutions for the two algorithms are: $K(\omega_N) = [-2.0257\ 0.5407]$ and $K_{\text{sos}}(\omega_N) = [-2.2517\ 0.9063]$. In addition to the probabilistic bounds in Fig. 3, we also compute the actual JSR of the closed-loop system using the JSR toolbox [59]: $\rho(\mathcal{A}_{K(\omega_N)}) = 2.6002$ and $\rho(\mathcal{A}_{K_{\text{sos}}(\omega_N)}) = 2.1188$. As a reference, we also solve the white-box quadratic stabilization problem as given in (5) and obtain the white-box solution $K^* = [-0.2831\ \ -0.2965]$. With this, we again compute the actual JSR [59]: $\rho(\mathcal{A}_{K^*}) = 2.5582$, as indicated by the dashed line in Fig. 3. From these computations, we can see that the data-driven solution from SOS stabilization even outperforms the solution from the white-box quadratic stabilization. We can also conclude from Fig. 3 that SOS stabilization requires more samples to converge as it has more design variables. In this example, we slightly abuse the concept of stabilization, because the closed-loop JSR is larger than 1. When the example is generated randomly, it often happens that the switched linear system cannot be stabilized by a common feedback gain. However, we can still use such

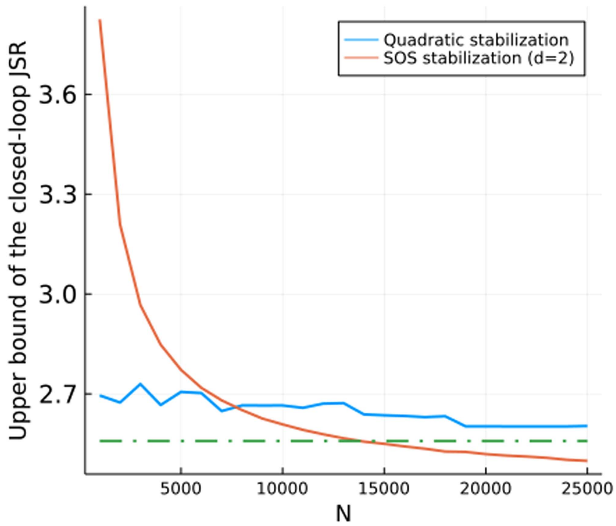


Fig. 3. Comparison of quadratic stabilization and SOS stabilization: the dashed line is the JSR of the white-box quadratic stabilization solution.

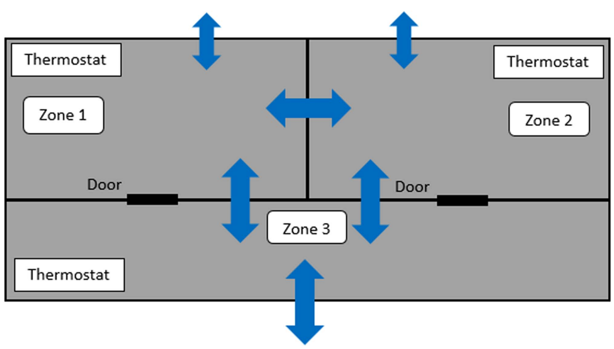


Fig. 4. System schematic of the building.

an example to have some idea on the performance difference between quadratic and SOS stabilization in a data-driven setting.

C. Building Temperature Regulation With LQR

Buildings can be viewed as complex control systems with both continuous and discrete dynamics. For instance, events, such as opening/closing doors and windows, instantly affect the dynamic evolution of the zone temperature. To capture these hybrid behaviors, hybrid *RC* networks are often used in the modeling of building systems [60], [61]. Consider a building with three zones as shown in Fig. 4. Its thermal *RC* model is described as follows, $i = 1, 2, 3$:

$$c_i \dot{T}_i = \sum_{j \neq i} \frac{T_j - T_i}{R_{ji}} + \frac{T_o - T_i}{R_i^o} + m_i c_p (T_s - T_i) + q_i$$

where i is the index of the zone, T_i is the temperature of zone i, T_o is the temperature of outside air, c_i is the thermal capacitance of the air in zone i, R_{ij} denotes the thermal resistances between zone i and zone j, R_i^o denotes the thermal resistance between zone i and the outside environment, c_p is the specific heat capacity of air, T_s is the temperature of the supply air delivered

TABLE I
SYSTEM PARAMETERS

Symbol	Value	Units
$c_i, \forall i$	1.375×10^3	$\mathrm{kJ/K}$
c_p	1.012	$\mathrm{kJ/(kg \cdot K)}$
$R_{12} = R_{21}$	1.5	$\mathrm{K/kW}$
$R_1^o = R_2^o$	3	$\mathrm{K/kW}$
R_3^o	2.7	$\mathrm{K/kW}$
T_s	16	$^\circ\mathrm{C}$
T_o	32	$^\circ\mathrm{C}$

TABLE II
THERMAL RESISTANCES OF DIFFERENT MODES

	"open"	"closed"
$R_{13} = R_{31}$ (K/kW)	0.8	1.2
$R_{23} = R_{32}$ (K/kW)	0.8	1.2

to zone i, m_i is the flow rate into zone i, and q_i is the thermal disturbance from internal loads, such as occupants and lighting. The temperatures of the supply air and the outside environment are known, i.e., T_s and T_o are available. The thermal disturbance is estimated as: $q_1 = 0.1$ kJ/s, $q_2 = 0.1$ kJ/s, $q_3 = 0.12$ kJ/s. Other system parameters are given in Table I. The control objective is to steer the temperature of each zone to $T_{\text{target}} = 24\,^\circ\mathrm{C}$.

By letting $Q_i^{AC} = m_i c_p (T_s - T_i)$, an equivalent linearized model is obtained

$$c_i \dot{T}_i = \sum_{j \neq i} \frac{T_j - T_i}{R_{ji}} + \frac{T_o - T_i}{R_i^o} + Q_i^{AC} + q_i, i = 1, 2, 3.$$

The steady input for the given targeted temperature is $\bar{Q}_i^{AC} = -\frac{T_o - T_{\text{target}}}{R_i^o} - q_i$ for all i. Let $x_i = T_i - T_{\text{target}}$ and $u_i = Q_i^{\text{AC}} - \bar{Q}_i^{\text{AC}}$ for all i. We get

$$\dot{x}_i = \sum_{j \neq i} \frac{x_j - x_i}{c_i R_{ji}} - \frac{x_i}{c_i R_i^o} + \frac{1}{c_i} u_i, i = 1, 2, 3.$$

As the doors are frequently and unpredictably open and closed, this is a typical switching system, in which the thermal resistances R_{13} (or R_{31}) and R_{23} (or R_{32}) are changing arbitrarily. For each door, we consider two modes: "open" and "closed." Hence, for the overall system, there are four modes in total, see Table II. The values of these thermal resistances are assumed to be unknown. In the simulation, we only use them to generate synthetic data.

We discretize the continuous-time system with the sampling time $\tau = 3$ min. Here, we use the Euler forward method for its simple implementation. See [62] for an extensive study on different discretization methods for buildings. The discretized system is given as follows:

$$\frac{x_i(t+1) - x_i(t)}{\tau} = \sum_{j \neq i} \frac{x_j(t) - x_i(t)}{c_i R_{ji}} - \frac{x_i(t)}{c_i R_i^o} + \frac{u_i(t)}{c_i}, i = 1, 2, 3.$$

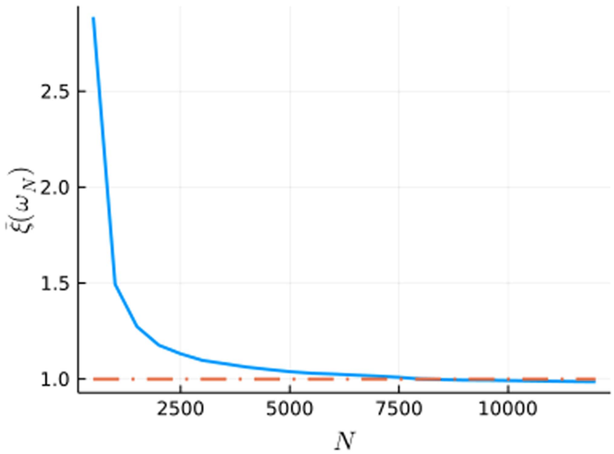


Fig. 5. Feasibility measure of the LQR solution.

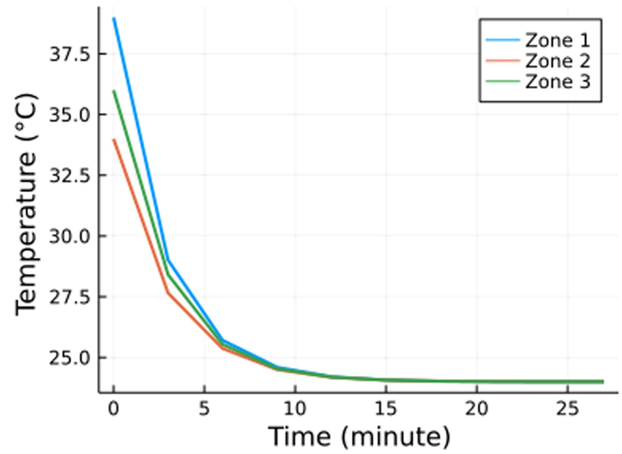


Fig. 6. Temperatures of the three zones with the LQR controller.

We now design the LQR for the switching system above using scenario optimization. Let the LQR parameters be $Q = I$ and $R = 0.02I$. For different values of N, we generate the dataset ω_N. The confidence level is set to be $\mathcal{B}(\epsilon; N) = 0.01$ in all the cases and the corresponding ϵ is computed via bisection. We then use Algorithm 4 with $\bar{\kappa} = 100$ to obtain a feasible solution to the sampled LQR problem in (55). Let

$$\bar{\xi}(\omega_N) := \frac{\xi(\omega_N)}{\xi^*\left(\epsilon, \kappa\left(P(\omega_N) - Q - K(\omega_N)^\top RK(\omega_N)\right)\right)}.$$

From the discussions in Section V, the value $\bar{\xi}(\omega_N)$ can be considered as an indicator for feasibility (provided that ω_N is an ϵ-covering of of $\mathbb{S} \times \mathcal{M}$): $(P(\omega_N), K(\omega_N))$ is a feasible solution when $\bar{\xi}(\omega_N) \leq 1$. We show the values of $\bar{\xi}(\omega_N)$ as the size of the dataset ω_N increases in Fig. 5. From this curve, we can see that $\bar{\xi}(\omega_N)$ becomes less than 1 when $N \geq 8000$. We also show the LQR solution when $N = 12\,000$ as follows:

$$K(\omega_N) = \begin{pmatrix} -3.3773 & -0.5579 & -0.6681 \\ -0.5580 & -3.3763 & -0.6660 \\ -0.6683 & -0.6686 & -3.2397 \end{pmatrix}$$

$$P(\omega_N) = \begin{pmatrix} 1.4041 & 0.1138 & 0.1333 \\ 0.1138 & 1.4041 & 0.1333 \\ 0.1333 & 0.1333 & 1.3787 \end{pmatrix}.$$

As a way to validate the sample-based solution, we compute the white-box LQR solution by solving the LMI inequalities in (53) and the solution is given as follows:

$$K^* = \begin{pmatrix} -3.1736 & -0.4840 & -0.5938 \\ -0.4840 & -3.1736 & -0.5938 \\ -0.5882 & -0.5882 & -3.0320 \end{pmatrix}$$

$$P^* = \begin{pmatrix} 1.3844 & 0.1085 & 0.1270 \\ 0.1085 & 1.3844 & 0.1270 \\ 0.1270 & 0.1270 & 1.3602 \end{pmatrix}.$$

The relative differences between the two solutions are given by $\|K(\omega_N) - K^*\|/\|K^*\| \times 100\% = 8.44\%$ and $\|P(\omega_N) - P^*\|/\|P^*\| \times 100\% = 1.93\%$, which suggests that the sample-based solution is a quite good approximation. Finally, we show the evolution of the temperatures of the three zones with the controller $u = K(\omega_N)x$ in Fig. 6.

VII. Conclusion

We have presented a data-driven control framework for stabilization of black-box switched linear systems, in which the dynamics matrices and the switching signal are unknown using quadratic and SOS Lyapunov functions. With quadratic Lyapunov functions, the stabilization problem is formulated as a biconvex problem using a finite number of trajectories. With SOS Lyapunov functions, we end up with a nonlinear optimization problem with a set of polynomial constraints. We then propose alternating minimization algorithms to solve these problems by making use of the underlying structure. In both cases, we develop parallelized schemes that allow to handle high-dimensional systems. Using the notions of covering/packing numbers, we also provide probabilistic stability guarantees via geometric analysis for the quadratic Lyapunov technique and sensitivity analysis for the SOS Lyapunov technique. Finally, we show that the proposed data-driven framework can be extended to LQR design of switched linear systems.

For future work, we want to consider the impact of measurement noise in our framework. We also want to investigate the use of the control input (aka. excitation) to 1) induce that the sampled states provide a good covering of the states around the equilibrium point (cf., Remark 2), and 2) to control the impact of the noise on the outcome of the algorithm. We also plan to investigate the case where the matrix B is unknown.

Acknowledgment

Raphaël M. Jungers is a FNRS honorary Research Associate.

References

[1] D. Liberzon and A. S. Morse, "Basic problems in stability and design of switched systems," *IEEE Control Syst. Mag.*, vol. 19, no. 5, pp. 59–70, Oct. 1999.

[2] D. Liberzon, *Switching in Systems and Control*. Berlin, Germany: Springer, 2003.
[3] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: A survey of recent results," *IEEE Trans. Autom. Control*, vol. 54, no. 2, pp. 308–322, Feb. 2009.
[4] A. S. Morse, "Supervisory control of families of linear set-point controllers—Part I. exact matching," *IEEE Trans. Autom. Control*, vol. 41, no. 10, pp. 1413–1431, Oct. 1996.
[5] J. C. Geromel and P. Colaneri, "Stability and stabilization of continuous-time switched linear systems," *SIAM J. Control Optim.*, vol. 45, no. 5, pp. 1915–1930, 2006.
[6] G. Chesi, P. Colaneri, J. C. Geromel, R. Middleton, and R. Shorten, "Computing upper-bounds of the minimum dwell time of linear switched systems via homogeneous polynomial Lyapunov functions," in *Proc. Amer. Control Conf.*, 2010, pp. 2487–2492.
[7] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *Proc. IEEE 38th Conf. Decis. Control*, 1999, vol. 3, pp. 2655–2660.
[8] G. Zhai, B. Hu, K. Yasuda, and A. N. Michel, "Qualitative analysis of discrete-time switched systems," in *Proc. Amer. Control Conf.*, 2002, vol. 3, pp. 1880–1885.
[9] M. Fiacchini, A. Girard, and M. Jungers, "On the stabilizability of discrete-time switched linear systems: Novel conditions and comparisons," *IEEE Trans. Autom. Control*, vol. 61, no. 5, pp. 1181–1193, May 2016.
[10] R. M. Jungers and P. Mason, "On feedback stabilization of linear switched systems via switching signal control," *SIAM J. Control Optim.*, vol. 55, no. 2, pp. 1179–1198, 2017.
[11] M. Fiacchini and M. Jungers, "Necessary and sufficient condition for stabilizability of discrete-time linear switched systems: A set-theory approach," *Automatica*, vol. 50, no. 1, pp. 75–83, 2014.
[12] F. Blanchini and S. Miani, "A new class of universal Lyapunov functions for the control of uncertain linear systems," *IEEE Trans. Autom. Control*, vol. 44, no. 3, pp. 641–647, Mar. 1999.
[13] J. Lee and G. E. Dullerud, "Uniform stabilization of discrete-time switched and Markovian jump linear systems," *Automatica*, vol. 42, no. 2, pp. 205–218, 2006.
[14] W. Zhang, A. Abate, J. Hu, and M. P. Vitus, "Exponential stabilization of discrete-time switched linear systems," *Automatica*, vol. 45, no. 11, pp. 2526–2536, 2009.
[15] C. J. Ong, Z. Wang, and M. Dehghan, "Model predictive control for switching systems with dwell-time restriction," *IEEE Trans. Autom. Control*, vol. 61, no. 12, pp. 4189–4195, Dec. 2016.
[16] B. A. H. Vicente and P. A. Trodden, "Switching tube-based MPC: Characterization of minimum dwell-time for feasible and robustly stable switching," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4345–4352, Oct. 2019.
[17] F. Lauer and G. Bloch, *Hybrid System Identification: Theory and Algorithms for Learning Switching Models*, vol. 478. Berlin, Germany: Springer, 2018.
[18] F. Lauer, "On the complexity of switching linear regression," *Automatica*, vol. 74, pp. 80–83, 2016.
[19] Z. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Inf. Sci.*, vol. 235, pp. 3–35, 2013.
[20] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, Mar. 2020.
[21] H. J. Van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, "Data informativity: A new perspective on data-driven analysis and control," *IEEE Trans. Autom. Control*, vol. 65, no. 11, pp. 4753–4768, Nov. 2020.
[22] I. Markovsky and F. Dörfler, "Behavioral systems theory in data-driven analysis, signal processing, and control," *Annu. Rev. Control*, vol. 52, pp. 42–64, 2021.
[23] J. Kenanian, A. Balkan, R. M. Jungers, and P. Tabuada, "Data driven stability analysis of black-box switched linear systems," *Automatica*, vol. 109, 2019, Art. no. 108533.
[24] A. Rubbens, Z. Wang, and R. M. Jungers, "Data-driven stability analysis of switched linear systems with sum of squares guarantees," *IFAC-PapersOnLine*, vol. 54, no. 5, pp. 67–72, 2021.
[25] G. O. Berger, R. M. Jungers, and Z. Wang, "Chance-constrained quasi-convex optimization with application to data-driven switched systems control," in *Proc. Learn. Dyn. Control*, 2021, pp. 571–583.
[26] M. Rotulo, C. De Persis, and P. Tesi, "Online learning of data-driven controllers for unknown switched linear systems," *Automatica*, vol. 145, 2022, Art. no. 110519.
[27] G. C. Calafiore, "Random convex programs," *SIAM J. Optim.*, vol. 20, no. 6, pp. 3427–3464, 2010.
[28] M. C. Campi, S. Garatti, and F. A. Ramponi, "A general scenario theory for nonconvex optimization and decision making," *IEEE Trans. Autom. Control*, vol. 63, no. 12, pp. 4067–4078, Dec. 2018.
[29] M. C. Campi and S. Garatti, "Wait-and-judge scenario optimization," *Math. Program.*, vol. 167, no. 1, pp. 155–189, 2018.
[30] P. M. Esfahani, T. Sutter, and J. Lygeros, "Performance bounds for the scenario approach and an extension to a class of non-convex programs," *IEEE Trans. Autom. Control*, vol. 60, no. 1, pp. 46–58, Jan. 2015.
[31] Z. Wang and R. M. Jungers, "On objective function value performance of the scenario approach under regularity conditions," *IEEE Trans. Autom. Control*, early access, Sep. 15, 2023, doi: 10.1109/TAC.2023.3315691.
[32] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[33] Z.. Wang, G. O. Berger, and R. M. Jungers, "Data-driven feedback stabilization of switched linear systems with probabilistic stability guarantees," in *Proc. IEEE 60th Conf. Decis. Control*, 2021, pp. 4400–4405.
[34] R. M. Jungers, *The Joint Spectral Radius: Theory and Applications*, vol. 385. Berlin, Germany: Springer, 2009.
[35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[36] F. Zhang, *The Schur Complement and Its Applications*, vol. 4. Berlin, Germany: Springer, 2006.
[37] J. G. VanAntwerp and R. D. Braatz, "A. tutorial on linear and bilinear matrix inequalities," *J. Process Control*, vol. 10, no. 4, pp. 363–385, 2000.
[38] T. Sarkar, A. Rakhlin, and M. Dahleh, "Nonparametric system identification of stochastic switched linear systems," in *Proc. IEEE 58th Conf. Decis. Control*, 2019, pp. 3623–3628.
[39] H. Modares and F. L. Lewis, "Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning," *IEEE Trans. Autom. Control*, vol. 59, no. 11, pp. 3051–3056, Nov. 2014.
[40] H. Zhang, K. Zhang, Y. Cai, and J. Han, "Adaptive fuzzy fault-tolerant tracking control for partially unknown systems with actuator faults via integral reinforcement learning method," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 10, pp. 1986–1998, Oct. 2019.
[41] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optim.*, vol. 11, no. 3, pp. 796–817, 2001.
[42] D. Henrion and J.-B. Lasserre, "Gloptipoly: Global optimization over polynomials with MATLAB and SeDuMi," *ACM Trans. Math. Softw.*, vol. 29, no. 2, pp. 165–194, 2003.
[43] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. A. Parrilo, "SOSTOOLS: Sum of squares optimization toolbox for MATLAB," 2013, *arXiv:1310.4716v2*.
[44] D. Henrion, J. Lofberg, M. Kocvara, and M. Stingl, "Solving polynomial static output feedback problems with PENBMI," in *Proc. IEEE 44th Conf. Decis. Control*, 2005, pp. 7581–7586.
[45] M. Kheirandishfard, F. Zohrizadeh, and R. Madani, "Convex relaxation of bilinear matrix inequalities Part I: Theoretical results," in *Proc. IEEE 57th Conf. Decis. Control*, 2018, pp. 67–74.
[46] S. Li, "Concise formulas for the area and volume of a hyperspherical cap," *Asian J. Math. Statist.*, vol. 4, no. 1, pp. 66–70, 2011.
[47] Z. Wang and R. M. Jungers, "A data-driven method for computing polyhedral invariant sets of black-box switched linear systems," *IEEE Contr. Syst. Lett.*, vol. 5, no. 5, pp. 1843–1848, Nov. 2021.
[48] A. Banse, Z. Wang, and R. M. Jungers, "Learning stability guarantees for constrained switching linear systems from noisy observations," *Nonlinear Anal.: Hybrid Syst.*, vol. 51, 2024, Art. no. 101425.
[49] N. Buduma and N. Locascio, *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2017.
[50] P. A. Parrilo and A. Jadbabaie, "Approximation of the joint spectral radius using sum of squares," *Linear Algebra Appl.*, vol. 428, no. 10, pp. 2385–2402, 2008.
[51] A. A. Ahmadi, R. M. Jungers, P. A. Parrilo, and M. Roozbehani, "Joint spectral radius and path-complete graph Lyapunov functions," *SIAM J. Control Optim.*, vol. 52, no. 1, pp. 687–717, 2014.
[52] A. Wächter, "Short tutorial: Getting started with ipopt in 90 minutes," in *Proc. Dagstuhl Seminar.*, 2009, vol. 9061.
[53] S. G. Johnson, "The NLopt nonlinear-optimization package, 2007. [Online]. Available: https://github.com/stevengj/nlopt

[54] E. D. Andersen and K. D. Andersen, "The Mosek interior point optimizer for linear programming: An implementation of the homogeneous algorithm," in *High Performance Optimization*, Boston, MA, USA: Springer, 2000, pp. 197–232.

[55] B. Lincoln and A. Rantzer, "Relaxing dynamic programming," *IEEE Trans. Autom. Control*, vol. 51, no. 8, pp. 1249–1260, 2006.

[56] A. Rantzer, "Relaxed dynamic programming in switching systems," *IEE Proc.-Control Theory Appl.*, vol. 153, no. 5, pp. 567–574, Aug. 2006.

[57] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. Hoboken, NJ, USA: Wiley, 2012.

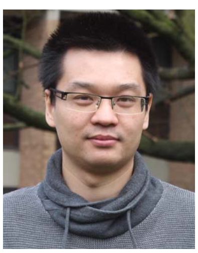
[58] M. V. Kothare, V. Balakrishnan, and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, vol. 32, no. 10, pp. 1361–1379, 1996.

[59] G. Vankeerberghen, J. Hendrickx, and R. M. Jungers, "JSR: A toolbox to compute the joint spectral radius," in *Proc. 17th Int. Conf. Hybrid Syst.: Comput. Control*, 2014, pp. 151–156.

[60] P. Fazenda, P. Lima, and P. Carreira, "Context-based thermodynamic modeling of buildings spaces," *Energy Build.*, vol. 124, pp. 164–177, 2016.

[61] B. Ajib, S. Lefteriu, A. Caucheteux, and S. Lecoeuche, "Building thermal modeling using a hybrid system approach," in *Proc. 20th World Congr. Int. Federation Autom. Control*, 2017, pp. 10716–10720.

[62] A. Kelman, Y. Ma, and F. Borrelli, "Analysis of local optima in predictive control for energy efficient buildings," *J. Build. Perform. Simul.*, vol. 6, no. 3, pp. 236–255, 2013.



Zheming Wang (Member, IEEE) received his B.S. degree in mechanical engineering and automation (Mechatronics) from the School of Mechanical Engineering of Shanghai Jiao Tong University, in 2012 and Ph.D. degree in control and mechatronics from the Department of Mechanical Engineering of National University of Singapore, in 2016.

He is currently a Professor with the Department of Automation, at Zhejiang University of Technology, Hangzhou, China. His main research interests include model predictive control, set theoretic methods, and data-driven control.

Dr. Wang serves as an Associate Editor for the IEEE CSS Conference Editorial Board, the IEEE CSS Technology Conference Editorial Board, and the IFAC/Elsevier journal Nonlinear Analysis: Hybrid Systems.

Guillaume O. Berger (Member, IEEE) received the M.Sc. and Ph.D. degrees in mathematical engineering from UCLouvain, Ottignies-Louvain-la-Neuve, Belgium, in 2017 and 2022, respectively.

He is currently an FNRS Postdoctoral Researcher at UCLouvain. His research interests include the interaction of Machine Learning and Control, with a focus on data-driven formal methods for cyber-physical systems.

Dr. Berger was the recipient of the ACM SIGBED Best Paper Award in the ACM International Conference on Hybrid Systems: Computation and Control (HSCC) in 2020. He is a BAEF Fellow.

Raphaël M. Jungers (Senior Member, IEEE) received the first M.Sc. degree in applied mathematics from the Ecole Centrale Paris, Gif-sur-Yvette, France, in 2004, and the second M.Sc. degree in applied mathematics from UCLouvain, Ottignies-Louvain-la-Neuve, Belgium, in 2005, and the Ph.D. degree in mathematical engineering from UCLouvain, in 2008.

He has held various invited positions, with the Universit Libre de Bruxelles (2008–2009), Laboratory for Information and Decision Systems of the Massachusetts Institute of Technology (2009–2010), University of L'Aquila (2011, 2013, and 2016), and University of California Los Angeles (2016–2017). He is currently a Professor with UCLouvain. His main research interests include in the fields of computer science, graph theory, optimization, and control.

Dr. Jungers is a FNRS, BAEF, and Fulbright Fellow. He has been an Associate Editor for the IEEE CSS Conference Editorial Board, and the journals *NAHS, Systems, and Control Letters*, and IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He was the recipient of the IBM Belgium 2009 award and a finalist of the ERCIM Cor Baayen award 2011. He was the co-recipient of the SICON Best Paper award 2013–2014.