

Robust Model-Free Learning and Control without Prior Knowledge

Dimitar Ho and John C. Doyle

Abstract—We present a simple model-free control algorithm that is able to robustly learn and stabilize an unknown discrete-time linear system with full control and state feedback subject to arbitrary bounded disturbance and noise sequences. The controller does not require any prior knowledge of the system dynamics, disturbances or noise, yet can guarantee robust stability, uniform asymptotic bounds and uniform worst-case bounds on the state-deviation. Rather than the algorithm itself, we would like to highlight the new approach taken towards robust stability analysis which served as a key enabler in providing the presented stability and performance guarantees. We will conclude with simulation results that show that despite the generality and simplicity, the controller demonstrates good closed-loop performance.

I. INTRODUCTION

Learning to stabilize unknown dynamical systems from online data has been an active research area in the control community since the 1950's [1] and has recently attracted the attention of the machine learning community, foremost in the context of reinforcement learning. Although there has been extensive research on this topic, very few of the developed algorithms have reached the level of adoption in real world applications as one would expect. Particularly in areas where frequent interaction with the physical world is necessary, system failure is very costly and deployment of control algorithms is only possible if the algorithm can guarantee that minimal safety and performance specifications will be met during operation. Although there has been past research [3], [4] and recent research efforts [2], [5], [6], [7], [8], [9] to address this problem, very few algorithms come with the necessary performance and safety guarantees to be deployed in real world applications thus far.

Classical control approaches are found in the literature of adaptive control with [10], [11], [12] focusing on the deterministic and [3] on the stochastic setting. The self-tuning regulator and its variations come with asymptotic optimality [13], yet robust stability guarantees without restrictive assumptions are few and can only be made in the probabilistic sense. On the deterministic side [10], [11] point out that instabilities can occur with traditional adaptive schemes and provide improved version of adaptive controllers that come with robust stability and performance guarantees. Yet, the desired guarantees depend on proper tuning of parameters and other prior knowledge of the system that might not be available in practice. Other challenges associated with classical adaptive control

approaches are discussed in [14], [15]. Methods in safe reinforcement learning [2] like [7], [16], [17] can guarantee robust safety properties for classes of nonlinear systems, yet the synthesis procedures involved are computationally intractable and they require knowledge of an initially robust stabilizing controller. Recent work [6], [18], [5], [19], [8], [9] has made significant progress in providing algorithms with robust finite-time performance guarantees for the adaptive linear quadratic gaussian regulator problem, with [5] being the first tractable, single episodic algorithm that provides sub-linear regret performance guarantees. However, all the methods require that the uncertainty in the system dynamics is small enough to begin with, or that it can be estimated through experiments.

Motivated by the lack of deterministic safety guarantees in current approaches, our work is focused on finding tractable learning and control algorithms that can guarantee robust stability uniformly over an as-large-as-possible class of model uncertainties. Specifically, this paper is concerned with adaptive stabilization of the following linear discrete-time system

$$z_{k+1} = A_0 z_k + u_k + d_k \quad (1)$$

$$x_k = z_k + n_k \quad (2)$$

with state z_k , bounded disturbance d_k , bounded noise n_k and control action u_k that is only allowed to depend on noisy state measurements until time k , i.e: $x_0, \dots, x_k$. Although (1) describes a small set of linear models, due to the assumption of full state feedback and control, still all of the previously mentioned approaches require knowledge of bounds of varying precision on d_k , n_k and/or A_0 in order to state stability and performance guarantees.

In this paper we explore the problem from a convex geometry perspective, which allows us to solve robust adaptive stabilization of (1) in complete generality. To this end, we will present a simple controller that can adaptively stabilize (1) without any further assumptions on disturbance, noise or the system matrix A_0 . The presented algorithm performs tractable computations, does not require tuning and guarantees uniform asymptotic guarantees and worst-case guarantees on state-deviation. An additional surprising feature of the presented algorithm is that it's not based on the well-known certainty-equivalence principle and has a completely model-free formulation.

We would like to emphasize the convex geometry based

Dimitar Ho and John C. Doyle are with the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA. dho@caltech.edu, doyle@caltech.edu

techniques developed for our core stability analysis as they do not rely on traditional Lyapunov-type arguments and could potentially provide new tools for robust design and analysis of learning and control algorithms.

Due to the limited space, most proofs and details will be omitted here and interested readers are referred to the extended version of this work found online.

II. NOTATION

Concatenation of variables will be extensively used in this work and to shorten the notational burden, we will write $X_{i:j}$ to refer to the matrix $[x_i, x_{i+1}, \dots, x_j]$ and if the second index is dropped it is to be understood that $j = 0$, i.e. $X_i := X_{i:0}$. Moreover, sequences $\{x_k\}$ will be denoted by bold lower case variables $\mathbf{x}$. We will use the sans serif font to stress that a variable S is a set. Furthermore, given a matrix $M \in \mathbb{R}^{n \times H}$ and set S in $\mathbb{R}^n$, we will use the notation $M \subset S$ to express that all columns c_i of the matrix $M = c_{1:k}$ belong to the set S . Notice that this notation is unambiguous, since the expression $\subset$ between matrices and sets is otherwise not defined.

III. PROBLEM STATEMENT AND PREVIEW OF MAIN RESULT

Recall that our goal is to stabilize system (1) without any knowledge of A_0 , d_k , n_k . More precisely, we will assume that $d_k \in \mathbf{S}_w$, $n_k \in \mathbf{S}_v$, for some bounded but unknown sets $\mathbf{S}_d$ and $\mathbf{S}_n$.

In this paper, we will show that the following simple controller

$$u_k = (U_{k-1} - X_{k:1}) \lambda_{k-1}(x_k) \quad (3)$$

where $\lambda_{k-1}(x_k)$ is the solution to the convex optimization problem

$$\begin{aligned} \min_{\lambda} \quad & \|\lambda\|_1 \\ \text{s.t.} \quad & X_{k-1} \lambda = x_k \end{aligned} \quad (4)$$

always ensures a closed loop for which:

- (i) the state $\mathbf{x}$ and input $\mathbf{u}$ are uniformly bounded.
- (ii) an analytic upper-bound on the worst-case state-deviation can be derived
- (iii) eventual exponential convergence to a bounded limit set is guaranteed

Remark III.1. We will define $u_0 := 0$ and if X_{k-1} is rank deficient, we will replace x_k in (4) by its orthogonal projection $x_k^\perp$ onto the column space of X_{k-1} .

We will term the controller (3), (4) the *causal cancellation controller* as it can be interpreted to cancel out the part of the dynamics that can be inferred from all previously collected observations $x_{k:0}$. Next, we will introduce some mathematical tools that are required for the derivations.

IV. MATHEMATICAL PRELIMINARIES

The presented results will make use of the interplay between convex bodies and norms in $\mathbb{R}^n$ and the next section will summarize some of the key mathematical concepts and definitions necessary for the later derivations.

A. Symmetric Convex Bodies and Norms

Symmetric convex bodies and norms are equivalent in $\mathbb{R}^n$, in the sense that a unit norm ball K is a convex symmetric body (Lem.IV.2) and that K induces the same norm through (5) as described in (Lem.IV.1).

Definition IV.1 (Symmetric Convex Body). A set $K \subset \mathbb{R}^n$ is a symmetric convex body if K is a closed, bounded convex set with non-empty interior and $z \in K \Leftrightarrow -z \in K$.

Lemma IV.1. For all convex bodies $K \subset \mathbb{R}^n$, the following function

$$\|x\|_K := \min \{r \geq 0 \mid x \in rK\} \quad (5)$$

is a norm in $\mathbb{R}^n$.

Lemma IV.2. For any norm $\|\cdot\|_A$ in $\mathbb{R}^n$, the corresponding norm ball $B_A = \{x \mid \|x\|_A \leq 1\}$ is a symmetric convex body.

It is a known fact that norms are equivalent in finite dimensional vector spaces and the following definition helps us express this relationship explicitly.

Definition IV.2. Let K_1 and K_2 be symmetric convex bodies in $\mathbb{R}^n$, we will define the quantity $\|K_1\|_{K_2}$ as

$$\|K_1\|_{K_2} := \sup_{z \in K_1} \|z\|_{K_2} = \inf_{K_1 \subset tK_2, t \geq 0} t \quad (6)$$

Lemma IV.3. Let K_1 and K_2 be symmetric convex bodies in $\mathbb{R}^n$, then their respective norms satisfy the following inequality for all z :

$$\frac{1}{\|K_1\|_{K_2}} \|z\|_{K_2} \leq \|z\|_{K_1} \leq \|K_2\|_{K_1} \|z\|_{K_2} \quad (7)$$

Proof. The proof follows by homogeneity property of norms and observing that $z/\|z\|_{K_2} \in K_2$:

$$\|z\|_{K_1} = \|z\|_{K_2} \|z/\|z\|_{K_2}\|_{K_1} \leq \|z\|_{K_2} \sup_{z \in K_2} \|z\|_{K_1} \quad (8)$$

□

Furthermore, we will frequently make use of the following corollary:

Corollary IV.1. $\|K_1\|_{K_2} \leq 1 \Leftrightarrow K_1 \subset K_2 \Rightarrow \|z\|_{K_1} \geq \|z\|_{K_2}$

B. Representation Norm and Distance

The following functional (9) can easily be shown to be a norm and will be useful in our derivations:

Definition IV.3 (Representation Norm). Let $M \in \mathbb{R}^{n \times N}$ be full rank then define $\|x\|_M$ as the representation norm of x w.r.t. to M , where $\|x\|_M$ is the solution to the following optimization problem

$$\|x\|_M := \inf_{\text{s.t. } M\lambda = x} \|\lambda\|_1 \quad (9)$$

We will refer to the corresponding unit ball as $B(M) := \{x \mid \|x\|_M \leq 1\}$ and to $\lambda_M(x)$ as the minimizer of the problem (9).

Remark IV.1. The minimizer of (9) might not be unique, but one can always define a continuous function $\lambda_M(x)$ of minimizers.

(Coro.IV.2) and (Coro.IV.3) show that representation norms are merely a more convenient way to work with norms which are induced by symmetric polytopes.

Corollary IV.2. If P is a symmetric polytope of full volume, then we can equivalently write $\|\cdot\|_P$ as the representation norm $\|\cdot\|_{\mathcal{V}(P)}$, where $\mathcal{V}(P)$ is a matrix $[v_1, \dots, v_N]$ with v_i being all extreme points of P .

Remark IV.2. An example in which (Coro.IV.2) applies, are the $\|\cdot\|_1$ and $\|\cdot\|_\infty$ norm.

Corollary IV.3. The unit norm ball B_M corresponding to a representation norm $\|x\|_M$ with $M = [m_1, \dots, m_N]$ is the convex hull of the columns of M and $-M$, i.e.

$$B_M = \text{ConvexHull}(\{\pm m_1, \dots, \pm m_N\}) \quad (10)$$

Furthermore, (Def. IV.4) will be of use to compare different symmetric convex bodies

Definition IV.4 (Representation Distance). Let P_1 and P_2 be symmetric convex bodies in $\mathbb{R}^n$ of dimension n . Then define the representation distance $d_R(P_1, P_2)$ as

$$d_R(P_1, P_2) = \max\{\|P_1\|_{P_2}, \|P_2\|_{P_1}\}$$

and induces a metric over symmetric convex bodies in $\mathbb{R}^n$:

Lemma IV.4. $\log(d_R(P_1, P_2))$ is a metric over $\mathcal{B}_n^s$, the space of symmetric convex bodies in $\mathbb{R}^n$.

Given a fix M , and some x and y in $\mathbb{R}^n$, we can use (Lem.IV.4) to compare $B([M, x])$ and $B([M, y])$, which induces the following pseudo metric on $\mathbb{R}^n$

Corollary IV.4 (Representation Metric). Let $M \in \mathbb{R}^{n \times N}$ be full rank and define $d_M : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ as

$$d_M(x, y) = d_R(B([M, x]), B([M, y])), \quad (11)$$

then $\log d_M(x, y)$ is a pseudo-metric on $\mathbb{R}^n$.

C. Metric Entropy and Packing Numbers of Metric Spaces

Assume we have a metric space (M, d) and let $B_r(x)$ denote the ball of radius r centered at x .

Definition IV.5 (Packing and Packing Number). A subset $P \subset K$ is called packing of K , if $\{B_r(x)\}_{x \in P}$ are pair-wise disjoint. The packing number of K , denoted as $N^{\text{pack}}(K, r)$, is the maximum cardinality of any packing of K .

Definition IV.6 (Separated Sets and Metric Entropy). A subset $S \subset K$ is called r -separated if for each pair of points $x, y \in S$ holds $d(x, y) > r$. The metric entropy of K , denoted as $N^{\text{ent}}(K, r)$, is the maximum cardinality of any r -separated subset of K .

Lemma IV.5. The following relationship holds $\forall r > 0$:

$$N^{\text{ent}}(K, 2r) \leq N^{\text{pack}}(K, r) \leq N^{\text{ent}}(K, r)$$

V. MODEL-FREE CLOSED LOOP EQUATION

Notice that the system (1) can be rewritten in terms of the available measurements x_k to obtain the form:

$$x_{k+1} = A_0 x_k + u_k + w_k \quad (12)$$

$$w_k := d_k + n_{k+1} - A_0 n_k. \quad (13)$$

It is clear that the extra disturbance term w_k is bounded and for ease of exposition we will define B to be the following upperbounding superset, which will represent all lumped disturbances in our closed loop.

Definition V.1. Define B to be a symmetric convex body such that $d + n' - A_0 n \in B$ for all $d \in S_d$, $n, n' \in S_n$.

Writing out (12) for all measurements up until k gives us the relation

$$X_{k:1} = A_0 X_{k-1} + U_{k-1} + W_{k-1} \quad (14)$$

$$\Leftrightarrow A_0 X_{k-1} = X_{k:1} - U_{k-1} - W_{k-1} \quad (15)$$

Recalling the definition (3) and (4) of our proposed causal cancellation controller, we notice that by design we always satisfy the relations

$$x_k = X_{k-1} \lambda_{k-1}(x_k) \quad (16)$$

$$u_k = (U_{k-1} - X_{k:1}) \lambda_{k-1}(x_k). \quad (17)$$

By substituting (16), (17) and (15) into (12) we obtain the following model-free form of the closed loop dynamics

$$x_{k+1} = -W_{k-1} \lambda_{k-1}(x_k) + w_k. \quad (18)$$

Assuming $x_0 = 0$, it can be verified that the closed loop map from disturbance to state is homogenous of degree one:

Corollary V.1. If (x, w) is a solution to (18), then for any $c > 0$, (cx, cw) is a solution as well.

Furthermore, the following inequality will be used frequently throughout the derivations:

Lemma V.1. Let x be a trajectory of (18), then

$$\|x_{k+1}\|_B \leq \|x_k\|_{X_{k-1}} + 1 \quad (19)$$

VI. SEPARABILITY OF UNSTABLE EVENTS AND FINITE OCCURRENCE

In this section, we will present the core theorem that enables the later analysis of the closed loop. Rather than searching for a Lyapunov-type of argument, we will take an alternative approach to stability analysis inspired by convex geometry. First, we will show that a certain kind of unstable event can only occur a finite number of times in the trajectories of the closed loop system (18). In the later section we will show how this conclusion implies that all trajectories have to be bounded.

In particular, we will show that if the unstable event occurred in the closed loop at time k_i and state x_{k_i} , then the state

x_{k_i} has to be significantly different from states $x_{k_{i-1}}$, $x_{k_{i-2}}$ where the unstable event has occurred previously. In a mathematical sense, we will express this by showing that the set of states at which the unstable event occurred $\{x_{k_i}, x_{k_{i-1}}, \dots, x_{k_0}\}$ has to be a separated set w.r.t. to a certain metric. By using the concept of metric entropy, we can then establish that the cardinality of this set has to be bounded, i.e. the total number of occurrence of this event in the trajectory has to be finite.

A. Finite Occurrence of Unstable Events

For a fix $0 < \mu < 1$, let us define the occurrence of an unstable event $\mathcal{U}_\mu$ at time t :

Definition VI.1 (Unstable Event). Given a trajectory $\mathbf{x}$ and we will say that the unstable event $\mathcal{U}_\mu$ occurred at time t in trajectory $\mathbf{x}$, if

$$\|x_{t+1}\|_B \geq \max\{r(\mu), \mu \|x_t\|_B + 1\} \quad (20)$$

where $r(\mu) := \frac{1}{1-\mu}$.

Correspondingly, define $\mathcal{T}_\mu$ to be the set of times t at which the unstable event $\mathcal{U}_\mu$ occurred:

Definition VI.2. Given a trajectory $\mathbf{x}$ and $0 < \mu < 1$, define $\mathcal{T}_\mu \subset \mathbb{N}$ to be the set of all time-steps $\mathcal{T}_\mu(\mathbf{x}) = \{t \mid \text{s.t. } x_{t+1}, x_t \text{ satisfy (20)}\}$ and define $\mathcal{X}_\mu(\mathbf{x})$ to be the set of states x_t at which the unstable transition $\mathcal{U}_\mu$ occurred, i.e. $\mathcal{X}_\mu(\mathbf{x}) := \{x_t \mid t \in \mathcal{T}_\mu(\mathbf{x})\}$.

Lemma VI.1. Any $x_t \in \mathcal{X}_\mu(\mathbf{x})$ satisfies the inequalities:

$$\|x_t\|_B \geq \frac{\mu r(\mu)}{\|B\|_{X_{t-1}}}, \quad \left\| \frac{x_t}{\|x_t\|_B} \right\|_{X_{t-1}} \geq \mu \quad (21)$$

Proof. First, from (Lem.V.1), (Lem.IV.3) we obtain

$$\|x_{t+1}\|_B \leq \|x_t\|_{X_{t-1}} + 1 \quad (22)$$

$$\Rightarrow \|x_{t+1}\|_B \leq \|B\|_{X_{t-1}} \|x_t\|_B + 1 \quad (23)$$

and by combining them with (20) from the definition of the unstable event yields the result, where we used the fact $r(\mu) - 1 = \mu r(\mu)$. $\square$

The following result shows that the states $\mathcal{X}_\mu(\mathbf{x})$ corresponding to the occurrences of the unstable event $\mathcal{U}_\mu$ in a trajectory $\mathbf{x}$ have to satisfy the following separation condition:

Lemma VI.2. Consider an arbitrary closed loop trajectory $\mathbf{x}$ and any t_0 such that X_{t_0} is full column rank. Then every pair of states $x_{t_1}, x_{t_2} \in \mathcal{X}_\mu(\mathbf{x})$ with distinct times $t_1 \neq t_2 \in \mathcal{T}_\mu(\mathbf{x})$ and $t_1, t_2 > t_0$ satisfy the separation condition

$$d_B \left(\mu r(\mu) \frac{x_{t_1}}{\|x_{t_1}\|_B}, \mu r(\mu) \frac{x_{t_2}}{\|x_{t_2}\|_B} \right) \geq \gamma \mu, \gamma := \frac{\mu r(\mu)}{\|B\|_{X_{t_0}}} \quad (24)$$

with respect to the representation distance $d_B(\cdot, \cdot)$.

Proof. Without loss of generality, assume $t_2 > t_1$, then from (Lem.VI.1) we obtain

$$\left\| \frac{x_{t_2}}{\|x_{t_2}\|_B} \right\|_{X_{t_2-1}} \geq \mu \quad (25)$$

Now observe that since $t_2 > t_1 > t_0$ and inequality (21), we have that for all $z \in \mathbb{R}^n$ holds

$$\|z\|_{X_{t_2-1}} \leq \|z\|_{\left[X_{t_0}, \gamma \frac{x_{t_1}}{\|x_{t_1}\|_B} \right]}, \quad \gamma := \frac{\mu r(\mu)}{\|B\|_{X_{t_0}}} \quad (26)$$

since X_{t_0} and $\gamma x_{t_1} / \|x_{t_1}\|_B$ are both in the norm ball of $B(X_{t_2-1})$. To see the later, notice that x_{t_1} is a column of X_{t_2-1} which implies $\|x_{t_1}\|_{X_{t_2-1}} \leq 1$. Moreover from (21) and since $\|B\|_{X_{t_0}} \geq \|B\|_{X_{t_1-1}}$, we obtain

$$\begin{aligned} \frac{\gamma}{\|x_{t_1}\|_B} &= \frac{\mu r(\mu)}{\|B\|_{X_{t_0}} \|x_{t_1}\|_B} \\ &\leq \frac{\mu r(\mu)}{\|B\|_{X_{t_1-1}} \|x_{t_1}\|_B} \leq 1 \end{aligned} \quad (27)$$

which gives us by convexity and symmetry of $B(X_{t_2-1})$ the inequality:

$$\left\| \gamma \frac{x_{t_1}}{\|x_{t_1}\|_B} \right\|_{X_{t_2-1}} \leq 1. \quad (28)$$

Finally, using inequality (26) to upperbound (25) and multiplying by γ gives us the relation

$$\left\| \gamma \frac{x_{t_2}}{\|x_{t_2}\|_B} \right\|_{\left[X_{t_0}, \gamma \frac{x_{t_1}}{\|x_{t_1}\|_B} \right]} \geq \gamma \mu = \frac{\mu^2 r(\mu)}{\|B\|_{X_{t_0}}} \quad (29)$$

Moreover, using the fact that $\|A\|_B = \|cA\|_{cB}$ for any $c > 0$ and that $B \subset \|B\|_{X_{t_0}} X_{t_0}$ we can further upperbound (29) as

$$\left\| \mu r(\mu) \frac{x_{t_2}}{\|x_{t_2}\|_B} \right\|_{\left[B, \mu r(\mu) \frac{x_{t_1}}{\|x_{t_1}\|_B} \right]} \geq \gamma \mu = \frac{\mu^2 r(\mu)}{\|B\|_{X_{t_0}}}$$

Finally since trivially $B_{X_{t_0}} \subset B \left(\left[X_{t_0}, \mu r(\mu) \frac{x_{t_1}}{\|x_{t_1}\|_B} \right] \right)$ and referring to the definition of the representation distance (Coro.IV.4), we obtain the desired separation condition (24). $\square$

Recalling that $\log d_{B_w}(\cdot, \cdot)$ defines a metric, Lemma (VI.2) enforces a separation condition between states at which the event $\mathcal{U}_\mu$ can occur in the trajectory $\mathbf{x}$. In particular, consider the set $\bar{\mathcal{X}}_\mu(\mathbf{x})$ of normalized states corresponding to the times $\mathcal{T}_\mu(\mathbf{x})$ of unstable events:

$$\bar{\mathcal{X}}_\mu(\mathbf{x}) = \left\{ \mu r(\mu) \frac{x_t}{\|x_t\|_B} \mid t \in \mathcal{T}_\mu(\mathbf{x}) \right\}. \quad (30)$$

By applying (Lem.VI.2) pairwise, we can conclude that $\bar{\mathcal{X}}_\mu(\mathbf{x})$ is a $\log(\gamma\mu)$ -separated set in the set $\mu r(\mu)B$. Appealing to the notion of metric entropy as defined in (Def. IV.6), we can conclude that the cardinality of the set $\mathcal{X}_\mu(\mathbf{x})$ has to be upperbounded by the metric entropy $N_B^{ent}(\mu r(\mu)B, \log(\gamma\mu))$ of the set $\mu r(\mu)B$ w.r.t to the pseudo-metric $\log d_B(\cdot, \cdot)$ and the radius $\log(\gamma\mu)$. We will denote this number by $\bar{N}(\mu)$

Definition VI.3. Define $\bar{N}(\mu)$ as the number

$$\bar{N}(\mu) := N_B^{ent}(\mu r(\mu)B, \log(\gamma\mu)) \quad (31)$$

and an analytic bound for the quantity can be established by a volume comparison argument and using the equivalence relationship between the metric entropy and the packing number (Lem.IV.5):

Lemma VI.3.

$$\bar{N}(\mu) \leq 1 + \left(\frac{1}{2}\right)^n \left(\frac{\mu^2 r(\mu)^{3/2}}{\mu \sqrt{r(\mu)} - \sqrt{\|B\|_{X_{t_0}}}} \right)^n \quad (32)$$

if μ satisfies

$$\frac{\mu^2}{1 - \mu} > \|B\|_{X_{t_0}} \quad (33)$$

Proof. The proof is technical and long and can not be presented in the scope of this paper. $\square$

Condition (33) can be equivalently written as

$$\frac{1}{2} \left(\sqrt{\|B\|_{X_{t_0}}^2 + 4\|B\|_{X_{t_0}}} - \|B\|_{X_{t_0}} \right) < \mu < 1 \quad (34)$$

Overall, the result (Lem.VI.3) shows us that if μ is chosen in the range (34), then the cardinality of the sets $|\mathcal{X}_\mu(\mathbf{x})| = |\mathcal{T}_\mu(\mathbf{x})|$ are bounded above by the quantity (32). This lets us conclude that for any closed loop trajectory $\mathbf{x}$ the occurrence of the corresponding events $\mathcal{U}_\mu$ is finite. We will summarize our findings with the following theorem:

Theorem VI.4. *Let $\mathbf{x}$ be closed loop trajectory of (18), the number of occurrences $|\mathcal{T}_\mu(\mathbf{x})|$ of the unstable event $\mathcal{U}_\mu$ is bounded above by $\bar{N}(\mu) < \infty$.*

Corollary VI.1. $\bar{N}(\mu)$ is independent of x_0 , A_0 and w_k .

Proof. This follows directly by noticing that the metric entropy bound (32) does not depend on x_0 , A_0 and w_k . $\square$

Remark VI.1. The bound (32) depends on $\|B\|_{X_{t_0}}$ which can be interpreted as a signal-to-noise ratio type quantity, as it compares the size of the disturbance polytope B to the polytope $B(X_{t_0})$ of prior data collected up until time t_0 .

VII. ROBUSTNESS GUARANTEES OF THE CLOSED LOOP

Based on the results derived in the last section, the robustness guarantees of the closed loop (18) follow directly. For ease of exposition we will phrase the following guarantees w.r.t. an arbitrary time-step t_0 for which X_{t_0} is guaranteed to be full column rank. Furthermore, it can be shown that $\lambda_{X_{t-1}}(x_t)$ are continuous maps and boundedness of $\mathbf{u}$ is directly implied by boundedness of the state $\mathbf{x}$.

Theorem VII.1. *Consider a trajectory $\mathbf{x}$ of the closed loop (18) for $t > t_0$, where t_0 is such that X_{t_0} is full rank. Then the following holds:*

- (i) $\mathbf{x}$ is bounded
- (ii) $\limsup_{t \rightarrow \infty} \|x_t\|_B \leq m(\|B\|_{X_{t_0}})$, with the bound $m(\|B\|_{X_{t_0}})$ defined as

$$m := \frac{1}{2} \left(\sqrt{\|B\|_{X_{t_0}}^2 + 4\|B\|_{X_{t_0}}} + \|B\|_{X_{t_0}} + 2 \right) \quad (35)$$

and $\|B\|_{X_{t_0}} \leq m(\|B\|_{X_{t_0}}) \leq \|B\|_{X_{t_0}} + 2$.

- (iii) for any μ in the range (34), there exists $t' > t_0$ for which trajectory $\|x_t\|_B$ converges exponentially ($\sim \mu^{t-t'}$) to the set $r(\mu)B$.
- (iv) $\sup_t \|x_t\|_B \leq h_{N^*}$, where $N^* := \inf_{\mu \text{ s.t. (34)}} \bar{N}(\mu) = \bar{N}(\bar{\mu})$ and the bound h_{N^*} is computed through the recursion

$$h_0 = \max\{\|x_{t_0}\|_B, \frac{1}{1 - \bar{\mu}}\} \quad (36)$$

$$h_{j+1} = \|B\|_{X_{t_0}} h_j + 1 \quad (37)$$

Proof. Consider an arbitrary trajectory $\mathbf{x}$. For any μ that lies in the range of (34), we know due to (Thm.VI.4) that $|\mathcal{T}_\mu(\mathbf{x})|$ is finite and therefore there exists a time $\bar{t}$ such that the following inequality

$$\|x_{t+1}\|_B \leq \max\{r(\mu), \mu \|x_t\|_B + 1\} \quad (38)$$

holds true for all $t > \bar{t}$. Furthermore, since $\mu < 1$ and by comparison Lemma we can conclude that $\|x_t\|_B$ converges exponentially with the rate ($\sim \mu^{t-\bar{t}}$) to the set $\frac{1}{1-\mu}B$. This establishes (iii) and (i). Furthermore, this also implies the asymptotic bound $\limsup_{t \rightarrow \infty} \|x_t\|_B \leq \frac{1}{1-\mu}$ for every μ that satisfies (34). We obtain the smallest such bound by taking the limit of μ to

$$\mu^* = \frac{1}{2} \left(\sqrt{\|B\|_{X_{t_0}}^2 + 4\|B\|_{X_{t_0}}} - \|B\|_{X_{t_0}} \right) \quad (39)$$

and obtain for $\frac{1}{1-\mu^*}$

$$\limsup_{t \rightarrow \infty} \|x_t\|_B \leq \frac{1}{1 - \mu^*} \quad (40)$$

$$= \frac{1}{2} \left(\sqrt{\|B\|_{X_{t_0}}^2 + 4\|B\|_{X_{t_0}}} + \|B\|_{X_{t_0}} + 2 \right) \quad (41)$$

$$\leq \|B\|_{X_{t_0}} + 2 \quad (42)$$

Finally, it can be shown that $\bar{N}(\mu)$ has a minimum and based on (Thm.VI.4) we can formulate a uniform worst-case bound on the trajectory $\|x_t\|_B$ by upperbounding the worst case N^* -steps unstable transitions, which gives us (iv). $\square$

Remark VII.1. The assumption of requiring X_{t_0} to be full rank is not essential for the results. Assume that $\text{rank}(X_t) = n - 1$ for times $t < t_0$. Then this means that all signals $\mathbf{u}$, $\mathbf{x}$ and $\mathbf{w}$ are part of a $n - 1$ -dimensional subspace of $\mathbb{R}^n$ and all the same arguments can be repeated for this lower dimensional space. This way, inductively we can see that the essence of the results apply for $t_0 = 0$ and arbitrary initial conditions and disturbances.

VIII. SIMULATION

(Fig. 1) shows simulation results of the controller (3) on the three-dimensional linear system

$$x_{k+1} = A_0 x_k + u_k + w_k, \quad w_k \in B = B_\infty \quad (43)$$

with A_0 and eigenvalues $\lambda(A_0)$:

$$A_0 = \begin{bmatrix} 1.4 & 0.2 & 1 \\ 0.2 & 1.3 & 1 \\ 0.5 & 0.3 & 2 \end{bmatrix} \quad \lambda(A_0) = \begin{bmatrix} 2.7 \\ 1.13 \\ 0.86 \end{bmatrix}. \quad (44)$$

(Fig. 1) displays $\|x_t\|_B$, $\|u_t\|_B$ along with the improvement of the asymptotic bound $m(\|B\|_{X_{t-1}})$ over time. The dotted part of the curve $m(\|B\|_{X_{t-1}})$ indicates that X_{t-1} is not full rank and therefore the quantity $m(\|B\|_{X_{t-1}})$ is computed w.r.t. to the projection of B onto the column space of X_{t-1} . Moreover, this explains the initial increase of $m(\|B\|_{X_{t-1}})$ for $t < 3$.

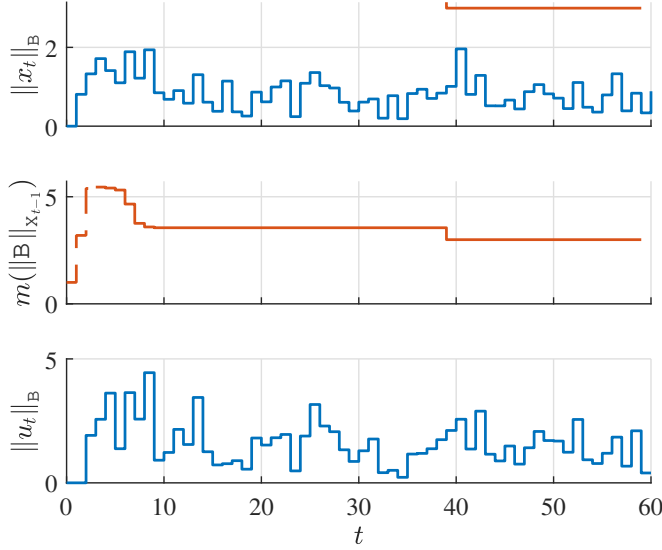


Fig. 1: $\|x_t\|_B$ and $\|u_t\|_B$ trajectories for closed loop with uniform disturbance $w_{i,t}$ in $[-1, 1]$ and $x_0 = [0, 0, 0]^T$.

IX. CONCLUSION

In this paper we derive a simple model-free controller that can adaptively and robustly stabilize a linear system with full actuation without any additional knowledge on disturbance, noise or parameter bounds. The controller comes with uniform asymptotic and worst-case guarantees on the state-deviation. The control design and stability analysis is enabled by a novel approach inspired by convex geometry and simulations show that the controller is able to simultaneously learn and control the system in an efficient manner, even when applied to an open loop system with large unstable eigenvalues. Future work will further explore how this new perspective to adaptive control can provide more learning and control algorithms with robustness guarantees and nonrestrictive assumptions in a more general setting. In addition, we will investigate how the presented ideas can help in providing robustness and performance bounds for present methods in adaptive control and reinforcement learning.

REFERENCES

- [1] A. M. J. A. Aseltine and C.W. Sartune, "A survey of adaptive control systems," *IEEE Transactions on Automatic Control*, vol. 3, pp. 102–108, 1958.
- [2] J. Garcia and F. Fernández, "A comprehensive survey on safe reinforcement learning," *Journal of Machine Learning Research*, vol. 16, no. 1, pp. 1437–1480, 2015.
- [3] B. W. Karl J. Astrom, *Adaptive control*, 2nd ed. Dover Publications, 1994.

- [4] K. J. Åström and B. Wittenmark, "On self tuning regulators," *Automatica*, vol. 9, no. 2, pp. 185–199, 1973.
- [5] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "Regret bounds for robust adaptive control of the linear quadratic regulator," *arXiv preprint arXiv:1805.09388*, 2018.
- [6] Y. Abbasi-Yadkori, N. Lazic, and C. Szepesvari, "Regret bounds for model-free linear quadratic control," *arXiv preprint arXiv:1804.06021*, 2018.
- [7] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," in *Advances in neural information processing systems*, 2017, pp. 908–918.
- [8] A. Cohen, T. Koren, and Y. Mansour, "Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret," *arXiv preprint arXiv:1902.06223*, 2019.
- [9] S. Dean, S. Tu, N. Matni, and B. Recht, "Safely learning to control the constrained linear quadratic regulator," in *2019 American Control Conference (ACC)*, July 2019, pp. 5582–5588.
- [10] P. A. Ioannou and J. Sun, *Robust adaptive control*. PTR Prentice-Hall Upper Saddle River, NJ, 1996, vol. 1.
- [11] P. Ioannou, *Adaptive Control Tutorial (Advances in Design and Control)*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2006.
- [12] S. Sastry and M. Bodson, *Adaptive control: stability, convergence and robustness*. Courier Corporation, 2011.
- [13] L. Guo, "Convergence and logarithm laws of self-tuning regulators," *Automatica*, vol. 31, no. 3, pp. 435–450, 1995.
- [14] B. D. Anderson and A. Dehghani, "Challenges of adaptive control—past, permanent and future," *Annual reviews in control*, vol. 32, no. 2, pp. 123–135, 2008.
- [15] B. D. Anderson *et al.*, "Failures of adaptive control theory and their resolution," *Communications in Information & Systems*, vol. 5, no. 1, pp. 1–20, 2005.
- [16] A. K. Akametalu, J. F. Fisac, J. H. Gillula, S. Kaynama, M. N. Zeilinger, and C. J. Tomlin, "Reachability-based safe learning with gaussian processes," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 1424–1431.
- [17] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Transactions on Automatic Control*, 2018.
- [18] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *arXiv preprint arXiv:1710.01688*, 2017.
- [19] H. Mania, S. Tu, and B. Recht, "Certainty equivalent control of lqr is efficient," *arXiv preprint arXiv:1902.07826*, 2019.