Robust Online Control with Model Misspecification

Xinyi Chen[1,2] XINYIC@PRINCETON.EDU
Udaya Ghai[1,2] UGHAI@CS.PRINCETON.EDU
Elad Hazan[1,2] EHAZAN@CS.PRINCETON.EDU
Alexandre Megretski[3] AMEG@MIT.EDU
1 *Department of Computer Science, Princeton University*
2 *Google AI Princeton*
3 *Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology*

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

We study online control of an unknown nonlinear dynamical system that is approximated by a time-invariant linear system with model misspecification. Our study focuses on **robustness**, a measure of how much deviation from the assumed linear approximation can be tolerated by a controller while maintaining finite ℓ_2-gain.

A basic methodology to analyze robustness is via the small gain theorem. However, as an implication of recent lower bounds on adaptive control, this method can only yield robustness that is exponentially small in the dimension of the system and its parametric uncertainty. The work of Cusumano and Poolla (1988a) shows that much better robustness can be obtained, but the control algorithm is inefficient, taking exponential time in the worst case.

In this paper we investigate whether there exists an efficient algorithm with provable robustness beyond the small gain theorem. We demonstrate that for a fully actuated system, this is indeed attainable. We give an efficient controller that can tolerate robustness that is polynomial in the dimension and independent of the parametric uncertainty; furthermore, the controller obtains an ℓ_2-gain whose dimension dependence is near optimal.

1. Introduction

The problem of linear control of linear dynamical systems is well studied and understood. Classical algorithms such as $\mathcal{H}_2$ optimization (which includes LQR and LQG) are known to be optimal in appropriate stochastic and worst case settings, while robust $\mathcal{H}_\infty$ control is optimal in the worst case, assuming quadratic costs. Even though these results can be generalized to nonlinear systems, the resulting optimal control synthesis requires solving partial differential equations in high dimensional domains, usually an intractable task. Beyond classical control methods, recent advancements in the machine learning community gave rise to efficient online control methods based on convex relaxations that minimize regret in the presence of adversarial perturbations.

In this paper we revisit a natural and well-studied approach of nonlinear control, where the nonlinear system is approximated by a linear plant with an uncertain (or misspecified) model. We capture the deviation of the plant dynamics from a linear time invariant system with an adversarial disturbance term in the system dynamics that can scale with the system state history. The amount of such deviation that can be tolerated while maintaining system stability constitutes **robustness** of the system under a given controller.

The field of adaptive control addresses the problem of controlling linear (and non-linear) dynamical systems with uncertain parameters. Adaptive control algorithms are frequently challenged on the issues of robustness and transient (finite-time) performance. Here, transient performance is in contrast with asymptotic performance, and as mentioned before, robustness measures the ability to tolerate unmodeled dynamics. A number of papers in the 1980s (e.g. Rohrs et al. (1982)) pointed out a lack of robustness under model misspecification for the classical *model reference adaptive control* (MRAC) approach. One can argue that this is related to the absence of transient behavior guarantees, such as a closed loop ℓ_2-gain bound, with good behavior expected only asymptotically, and this is the motivation for our study.

In this paper, we show that under a fully actuated system, a properly designed adaptive control algorithm can exhibit a significant degree of robustness to unmodeled dynamics and be computationally efficient. This is in contrast to a small gain approach to analyzing robustness, where robustness is guaranteed to be inversely proportional to the ℓ_2-gain of the closed loop system, excluding model misspecification. As recently shown by Chen and Hazan (2021) via a regret lower bound, it is inevitable that the ℓ_2-gain grows *exponentially* with the system dimension, implying a vanishing degree of robustness under the small gain theorem.

We show that it is possible to achieve robustness which depends *inverse polynomialy* on the system dimension, and independent of its parametric uncertainty, while maintaining an ℓ_2-gain that grows as $2^{O(d)}$, consistent with the known lower bounds of $2^{\tilde{\Omega}(d)}$. Previous work by Cusumano and Poolla (1988a) gives a very general, yet inefficient algorithm of adaptive control that achieves constant robustness for both fully actuated and under actuated systems. The algorithm assures finiteness of the close loop ℓ_2-gain, but yields an excessively high ℓ_2-gain bounds (as in having ℓ_2-gain that grows doubly exponentially in the dimension, in the same setting).

Our result improves upon previous work in the fully actuated setting, both in terms of computational efficiency and ℓ_2-gain. The controller is based on recent system identification techniques from non-stochastic control whose main component is active large-magnitude deterministic exploration. This technique deviates from one of the classical approaches of using least squares for system estimation and solving for the optimal controller. Our technique demonstrates how carefully chosen exploration for system identification can be used to bound the energy required for exploration and not to activate the system more than necessary, and yet obtain bounded ℓ_2-gain up to the known lower bounds.

1.1. Our contributions

We consider the setting of a linear dynamical system with time-invariant dynamics, together with model misspecification, as illustrated in Fig. 1.

The system evolves according to the following rule,

$$x_{t+1} = Ax_t + Bu_t + \Delta_t(x_{1:t}) + f_t, \tag{1}$$

where $A, B \in \mathbb{R}^{d\times d}$ is the (unknown) linear approximation to the system, $u_t, x_t, f_t \in \mathbb{R}^d$ are the control, state and adversarial perturbation respectively. We refer to an upper bound on the spectral norm of A as the parametric uncertainty. The perturbation $w_t = \Delta_t(x_{1:t})$ represents the deviation of the nonlinear system from the nominal system (A, B). The perturbations w_t crucially must satisfy

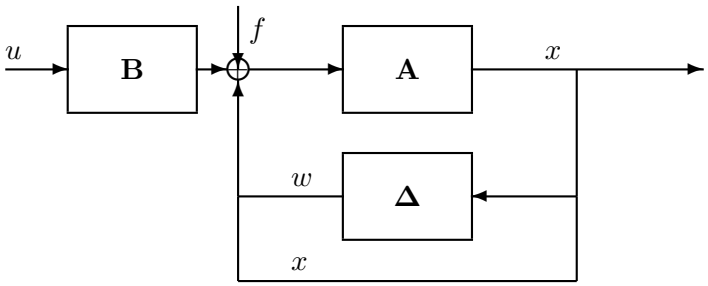

Figure 1: Diagram of the system, where Δ represents model misspecification.

the following assumption:

$$\sum_{s=1}^{t} \|w_s\|_2^2 \;\le\; h^2 (\sum_{s=1}^{t} \|x_s\|_2^2). \tag{2}$$

The parameter h is a measure of the robustness of the system, and is the main object of study. The larger h is, the more model misspecification can be accommodated by the controller. Our goal is to study the limits of robustness with reasonable transient performance. We use ℓ_2-gain, a quantity widely studied in classical control theory, as our performance measure. The ℓ_2-gain of a closed-loop system with control algorithm $\mathcal{A}$ in the feedback loop is defined as

$$\ell_2\text{-gain}(\mathcal{A}) = \max_{f_t} \frac{\|x_{1:T}\|_2}{\|f_{0:T-1}\|_2}, \tag{3}$$

where $x_{1:T}, f_{0:T-1} \in \mathbb{R}^{dT}$ are concatenations of $x_1, \ldots, x_t$, and $f_0, \ldots, f_{T-1}$, respectively. This notion is closely related to the competitive ratio of the control algorithm $\mathcal{A}$, as we show in App. C. With this notation, we can formally state our main question:

Is it possible to design efficient control methods that achieve robustness beyond the small gain theorem, while having ℓ_2-gain with near-optimal [1] dependence on the system dimension?

Our study initiates an answer to this question from both lower and upper bound perspectives. In terms of upper bounds, we consider the case of a fully actuated system, and show that in this important special case, constant robustness and near-optimal ℓ_2-gain are possible. [2]

- We give an efficient algorithm that is able to control the system with robustness $h = \Omega(\frac{1}{\sqrt{d}})$, where d is the system dimension. This is independent of the parametric uncertainty.
- In addition, we show that under parametric uncertainty M, this algorithm achieves finite ℓ_2-gain of $2^{\tilde{O}(d \log M)}$, where the dependence on system dimension is near-optimal given the lower bound of $2^{\Omega(d)}$ in Chen and Hazan (2021).

We also consider the limits of finite ℓ_2-gain and robust control. Clearly, if the system A, B is not stabilizable, then one cannot obtain any lower bound on the robustness regardless of what control

1. Here and elsewhere, near optimal means up to constants in the exponent.
2. Obtaining similar, or even partial, results in the general under-actuated case is an exciting, important, and potentially difficult open problem, see the conclusions section.

method is used. The distance of the system A, B from being stabilizable is thus an upper bound on the robustness, and we provide a proof for completeness in App. B.

For our main results, we use an active explore-then-commit method for system identification and a doubling strategy to handle unknown disturbance levels. As a supplementary result, we also study system identification using the more common online least squares method, and prove that it gives constant robustness and finite ℓ_2-gain bounds for one-dimensional systems in App. D.

1.2. Related work

Adaptive Control. The most relevant field to our work is adaptive control, see for example the book (Ioannou and Sun, 2012) and survey by Tao (2014). This field has addressed the problem of controlling a linear dynamical system with uncertain parameters, providing, in the 70s, guarantees of asymptotic optimality of adaptive control algorithms. However, reports of lack of robustness of such algorithms to *unmodeled dynamics* (as in the Rohrs et al. (1982) example) have emerged. One can argue that this lack of robustness was due to poor *noise rejection* transient performance of such controllers, which can be measured in terms of ℓ_2 *induced norm (gain)* of the overall system. The general task of designing adaptive controllers with finite closed loop ℓ_2-gain was solved by Cusumano and Poolla (1988a), but the ℓ_2-gain bounds obtained there grow very fast with the size of parameter uncertainty, and are therefore only good to guarantee a negligible amount of robustness. It has been confirmed by Megretski and Rantzer (2002/2003) that even in the case of one dimensional linear models, the minimal achievable ℓ_2 gain grows very fast with the size of parameter uncertainty.

Nonlinear Control. Recent research has studied provable guarantees in various complementary (but incomparable) models for nonlinear control. These include planning regret in nonlinear control Agarwal et al. (2021), adaptive nonlinear control under linearly-parameterized uncertainty Boffi et al. (2021), online model-based control with access to non-convex planning oracles Kakade et al. (2020), control with nonlinear observation models Mhammedi et al. (2020), system identification for nonlinear systems Mania et al. (2020) and nonlinear model-predictive control with feedback controllers Sinha et al. (2021).

Robustness and ℓ_2-gain in Control Robust control is concerned with the ability of a controller to tolerate uncertainty in system parameters, including unmodeled dynamics present in nonlinear systems. This field has been studied for many decades, see for example (Zhou et al., 1996) for a survey. One fundamental method for measuring robustness is certifying stability of the closed-loop system under non-parametric uncertainty via the small gain thoerem by Zames (1966), where stability is implied by finite ℓ_2-gain. The achievability of finite ℓ_2-gains for systems with unknown level of disturbance has been studied in control theory. Cusumano and Poolla (1988b) characterize the misspecification, or non-parametric uncertainty, tolerable for finite ℓ_2-gain. Megretski and Rantzer (2002/2003) gives a lower bound on the closed loop ℓ_2-gain of adaptive controllers that achieve finite ℓ_2-gain for all systems with bounded spectral norm. However, the systems studied in this paper do not contain any model misspecification.

Since the small gain theorem is known to be pessimistic, several alternative approaches have been proposed, including positivity theory and other methods of exploiting phase information of the system, constructing parameter-dependent Lyapunov functions, and using other notions of stability such as absolute stability, see Bernstein and Haddad (1992) for a survey.

Competitive Analysis for Control For a given controller, its ℓ_2-gain is closely related to the competitive ratio, which is a quantity more often studied in the computer science community, see next section for details. Yu et al. (2020) gives a control algorithm with constant competitive ratio for the setting of delayed feedback and imperfect future disturbance predictions. Shi et al. (2020) proposes algorithms whose competitive ratios are dimension-free for the setting of optimization with memory, with connections to control under a known, input-disturbed system and adversarial disturbances. More recently, Goel and Hassibi (2021) give an algorithm with optimal competitive ratio for known LTI systems and known quadratic costs, without misspecification.

System Identification for Linear Dynamical Systems. For an LDS with stochastic perturbations, the least squares method can be used to identify the dynamics in the partially observable and fully observable settings (Oymak and Ozay, 2019; Simchowitz et al., 2018; Sarkar and Rakhlin, 2019; Faradonbeh et al., 2019). However, least squares can lead to inconsistent solutions under adversarial disturbances, such as the model misspecification component in the system. The algorithms by Simchowitz et al. (2019) and Ghai et al. (2020) tolerate adversarial disturbances, but the guarantees only hold for stable or marginally stable systems. If the adversarial disturbances are bounded, Hazan et al. (2020) and Chen and Hazan (2021) give system identification algorithms for any unknown system, stable or not, with and without knowledge of a stabilizing controller, respectively. These techniques arose from recent results on nonstochastic control, such as works by Agarwal et al. (2019) and Simchowitz et al. (2020), for a comprehensive survey, see lecture notes by Hazan (2021).

1.3. Structure of the paper

In the next section we give a few preliminaries and definitions to precisely define our setting and problem. In Sec. 3 we give our main result: an efficient method with $\Omega(\frac{1}{\sqrt{d}})$ robustness and ℓ_2-gain of $2^{\tilde{O}(d \log M)}$ under unknown disturbance levels. We sketch out the analysis in Sec. 4.

Due to space constraints, significant technical material appears in the appendix, which can be found at (Chen et al., 2021). App. A provides additional background on the small gain approach to robust control. In App. B and App. C, we explore the limits of robustness of any controller and clarify the relationship between the performance metric ℓ_2-gain and the competitive ratio, respectively. In App. D we give an optimal result limited to the one-dimensional setting, where the ℓ_2-gain bounds are tight in the parametric uncertainty. In App. E we include proofs for Sec. 3, and in App. F we provide a complete analysis of an algorithm analogous to that of Cusumano and Poolla (1988a).

2. Preliminaries

Notation. We use the $\tilde{O}$ notation to hide constant and logarithmic terms in the relevant parameters. We use $\|\cdot\|_2$ to denote the spectral norm for matrices, and the Euclidean norm for vectors. We use $x_{s:t} \in \mathbb{R}^{d(t-s+1)}$ to denote the concatenation of $x_s, x_{s+1}, \ldots, x_t$, and similar notations are used for f, w, z.

We make the assumptions on the model misspecification component and the disturbances in Section 1.1 formal.

Assumption 1 *We treat the model misspecification component of the system, w_s, as an adversarial disturbance sequence. They are arbitrary functions of past states such that for all t:*[3]

$$\|w_{1:t}\|_2 \;\le\; h\|x_{1:t}\|_2.$$

The disturbance f_t in the system is arbitrary, and let $z_t = w_t + f_t$. Without loss of generality, let $w_0 = x_0 = u_0 = 0$.

Further, we assume the system is bounded and fully actuated.

Assumption 2 *The magnitude of the dynamics A, B are bounded by a known constant $\|A\|_2, \|B\|_2 \le M$, where $M \ge 1$. B's minimum singular value is also lower bounded as $\sigma_{\min}(B) > L$, where $0 < L \le 1$.*

ℓ_2-gain and Competitive Ratio. The competitive ratio of a controller is a concept that is closely related to ℓ_2-gain, but is more widely studied in the machine learning community. Informally, for any sequence of cost functions, the competitive ratio is the ratio between the cost of a given controller and the cost of the optimal controller, which has access to the disturbances $f_{0:T-1}$ a priori. Importantly, the notion of competitive ratio is counterfactual: it allows for different state trajectories $x_{1:T}$ as a function of the control inputs. Under some assumptions that our algorithm satisfies, ℓ_2-gain bounds can be converted to competitive ratio bounds (see Sec. C). We choose to present our results in terms of ℓ_2-gain for simplicity.

3. Main Algorithm and Results

In this section we describe our algorithm. The main algorithm, Alg.1, is run in epochs, each with a proposed upper bound q on the disturbance magnitude $\|f_{0:T-1}\|_2$. A new epoch starts whenever the controller implicitly discovers that q is not sufficiently large and increases the upper bound. While the disturbances f_t are not directly observed, with a valid upper bound q, the algorithm guarantees a bounded state expansion and bounded estimates of (A, B). When these conditions are broken, we deduce that the bound on $\|f_{0:T-1}\|_2$ was incorrect and restart the system identification procedure, appropriately scaling up our upper bound q .

The algorithm explores with large controls along the standard basis. If the upper bound q indeed exceeds $\|f_{0:T-1}\|_2$, the algorithm is guaranteed to find a stabilizing controller. By using the standard basis vectors as the exploration set, the algorithm attains robustness depending on $\sqrt{d}$ using $O(d)$ controls. In contrast, an inefficient version of the algorithm achieves dimension-free robustness, but uses an ϵ-net for exploration, resulting in an exponential number of large controls for system estimation. The alternate variant and analysis can be found in App. E.

The theorem below presents the main guarantee of our algorithm.

Theorem 1 *For $h \le \frac{1}{12\sqrt{d}}$, there exists ε, α such that Alg. 1 has ℓ_2-gain$(\mathcal{A}) \le (\frac{Md}{L})^{O(d)}$.*

3. Notice that w_t can depend on the actual trajectory of states, and not only their magnitude. This is important to capture miss-specification of the dynamics.

Algorithm 1: ℓ_2-gain algorithm

Input: System upper bound M, control matrix singular value lower bound L, system identification parameter ε, threshold parameter α.

Set $q = 0, K = 0$.

while $t \le T$ **do**
  Observe x_t.
  if $\|x_{1:t}\|_2 > \alpha q$ **then**
    Update $q = \|x_{1:t}\|_2$.
    Call Alg. 2 with parameters $(q, M, L, \varepsilon, \alpha)$, obtain updated K and budget q.
  else
    Execute $u_t = -Kx_t$.
    $t \leftarrow t + 1$
  end
end

Algorithm 2: Adversarial System ID on Budget

Input: Disturbance budget q, system upper bound M, control matrix singular value lower bound L, system identification parameter ε, threshold parameter α.

Call Alg. 3 with parameters $(q, M, L, \varepsilon, \alpha)$, obtain estimator $\hat{B}$ and updated budget q. Suppose the system evolves to time $t' = t + d$.

Set $q' = 4^{2d} M^{2d} \varepsilon^{-d} q$.

for *i = 0, 1, …,* $2d - 1$ **do**
  Observe $x_{t'+i}$.
  if $\|x_{1:t'+i}\|_2 > \alpha q$ **then**
    Restart SysID from Line 2 with $q = \|x_{1:t'+i}\|_2$.
  end
  if *i is even* **then**
    Play $u_{t'+i} = \xi_{i/2} \hat{B}^{-1} e_{i/2+1}$, $\xi_{i/2} = \frac{4^{3i/2} M^{3i/2+2} q'}{\varepsilon^{i/2+1}}$.
  else
    Play $u_{t'+i} = 0$.
  end
end

Observe $x_{t'+2d}$, compute

$$\hat{A} = [\frac{x_{t'+2}}{\xi_0} \cdots \frac{x_{t'+2d}}{\xi_{d-1}}] .$$

if $\|\hat{A}\|_2 > 2M$ **then**
  Restart SysID from Line 2 with $q = \|x_{1:t'+2d}\|_2$.
end

Return $q, K = \hat{B}^{-1} \hat{A}$

Algorithm 3: Adversarial Control Matrix ID on Budget

Input: Disturbance budget q, system upper bound M, control matrix singular value lower bound L, system identification parameter ε, threshold parameter α.

for *i = 0, 1, …, $d-1$* **do**
  Observe x_{t+i}.
  if $\|x_{1:t+i}\|_2 > \alpha q$ **then**
    Restart SysID with $q = \|x_{1:t+i}\|_2$.
  end
  Play $u_{t+i} = \lambda_i e_{i+1}$, $\lambda_i = \frac{4^{2i}M^{2i+1}q}{\varepsilon^{i+1}}$.
end
Observe x_{t+d}, compute

$$\hat{B} = [\frac{x_{t+1}}{\lambda_0} \cdots \frac{x_{t+d}}{\lambda_{d-1}}].$$

if $\|x_{1:t+d}\|_2 > \alpha q_k$ *or* $\sigma_{\min}(\hat{B}) < L/2$ **then**
  Restart SysID with $q = \|x_{1:t+d}\|_2$.
end
Return $q, \hat{B}$

4. Analysis

The algorithm has three components: exploration to estimate B, exploration to estimate A, and controlling the system with linear controller $K = \hat{B}^{-1}\hat{A}$. The parameter α serves as a relative upper bound, where the state energy $\|x_{1:T}\|_2$ is guaranteed not to surpass αq if q is a true upper bound on $\|f_{0:T-1}\|_2$. We first analyze the case if the upper bound on the disturbance magnitude is correct and $\|f_{0:T-1}\|_2 \le q$. In this case, the algorithm is designed with a suitable threshold α such that a new epoch will not be started and we are guaranteed to obtain a stabilizing controller. Note that in both exploration stages, the state can grow exponentially, so exploratory controls must also grow to keep up.

Epoch Notation. We define epochs in terms of rounds of system identification. In particular, for the kth epoch, s_k denotes the iteration number t on the kth call to the system identification procedure Alg. 2, and $e_k = \min(s_{k+1} - 1, T)$ is the iteration number of the end of the epoch. As such, within an epoch, q is fixed, so we denote $q_k = \|x_{1:s_k}\|_2$ the value of q within epoch k.

Identifying B (see App. E.2). The first step involves identifying the control matrix using Alg. 3. The following lemma shows that the control identification process will produce an accurate estimate of B in the spectral norm with singly-exponential growth in the state energy. Because our final controller is $K = \hat{B}^{-1}\hat{A}$, we also bound the distance of $B\hat{B}^{-1}$ from identity in order to properly stabilize the system.

Lemma 2 *Suppose $\|f_{0:T-1}\|_2 \le q_k$ and $\alpha \ge 4^{2d}M^{2d}\varepsilon^{-d}$, then running Alg. 3 with $\varepsilon \le \frac{L}{12\sqrt{d}}$ produces $\hat{B}$ such that $\|\hat{B} - B\|_2 \le 3\varepsilon\sqrt{d}$ and $\|B\hat{B}^{-1} - I\|_2 \le \frac{1}{2}$, with $\|x_{1:s_k+d}\|_2 \le 4^{2d}M^{2d}q_k\varepsilon^{-d}$.*

The algorithm works by probing the system with scaled standard basis vectors. With sufficiently large scaling, $x_{t+1} = Ax_t + Bu_t + z_t \approx Bu_t$. This allows us to estimate B one column at a

time. Arbitrarily large probing controls can yield an arbitrarily accurate estimate of B, though the magnitude of such controls will factor into the resultant ℓ_2-gain. This accuracy-gain trade off is balanced deeper in the analysis.

Identifying A (see App. E.3). Once we have an accurate estimate of B, we use Alg. 2 to produce an estimate $\hat{A}$ that is $O(h)$ accurate in each of the standard basis directions, again with a singly exponential state energy growth.

Lemma 3 *Suppose $\|f_{0:T-1}\|_2 \le q_k$ and $\alpha > R = (4M)^{5d}\varepsilon^{-2d}$, then Alg. 2 produces $\hat{A}$ such that*

$$\max_{i\in[d]} \|(A-\hat{A})e_i\|_2 \le \frac{28\varepsilon M\sqrt{d}}{L} + 3h\,,$$

with $\|x_{1:t'+2N}\|_2 \le Rq_k$.

Identification of A in Alg. 2 works by applying controls $u_t = \xi\hat{B}^{-1}v_t$ every other iteration, where v_t is a standard basis vector and ξ is a large constant such that $x_{t+1} \approx Ax_t + \xi v_t + z_t \approx \xi v_t$. One more time evolution with zero control gives $x_{t+2} = Ax_{t+1} + z_{t+1} \approx \xi A v_t + z_{t+1}$. By Assumption 1, $\|z_{t+1}\|_2 \le h\|x_{1:t+1}\|_2 + \|f_{0:t+1}\|_2 = O(h\xi + q)$. As a result, we have $\|\frac{x_{t+2}}{\xi} - Av_t\|_2 = O(h)$. By definition of $\hat{A}$ in Line 14, we also have $\|\frac{x_{t+2}}{\xi} - \hat{A}v_t\|_2 = O(h)$, so $\|(A-\hat{A})v_t\|_2 = O(h)$. Exploratory controls are preconditioned with $\hat{B}^{-1}$ to achieve robustness independent of $\sigma_{\min}(B)$.

By exploring with the standard basis, we assure that each row of $\hat{A}$ is accurate to $O(h)$, so $\|A-\hat{A}\|_2 \le \|A-\hat{A}\|_F \le h\sqrt{d}$. By bounding the spectral norm of the estimation error loosely through a bound on the Frobenius norm, we only produce an accurate estimate of A for $h = \Omega(1/\sqrt{d})$. With exploration complete, we shift to stabilizing the system.

Stabilizing the system (see App. E.4). The system is subsequently stabilized by linear controller $K = \hat{B}^{-1}\hat{A}$. By controlling the accuracy of $\hat{A}$ and $\hat{B}$, we guarantee the closed loop system satisfies $\|A - BK\|_2 < \frac{1}{2}$ via the following simple technical lemma:

Lemma 4 *Suppose $\|f_{0:T-1}\|_2 \le q_k$, $\alpha \ge 4^{2d}M^{2d}\varepsilon^{-d}$, with appropriate choice of ε the resultant controller K satisfies $\|A - BK\|_2 \le \frac{1}{2}$.*

Now, with a stable linear system, we can bound the remaining cost of using this stabilizing controller. In the below theorem t^* represents a time such that the controller plays a stabilizing linear controller for the remainder of the time horizon. In particular, we can view t^* as the last iteration of exploration.

Lemma 5 *If $\|f_{0:T-1}\|_2 \le q_k$, and let t^* be such that $u_t = -Kx_t$ for $t \ge t^* \ge s_k$, with $\|A - BK\|_2 \le 1/2$, then for $h \le \frac{1}{6}$,*

$$\|x_{1:e_k}\|_2^2 \le \frac{18\|x_{1:t^*}\|_2^2 + 72q_k^2}{7}.$$

This follows via induction arguments involving unrolling the linear dynamics. We can then obtain the following end-to-end bound by bounding $\|x_{1:t}\|_2^2$ in terms of q_k, plugging in $\|x_{1:t^*}\|_2 \le Rq_k$ via the exploration analysis of Lem. 3.

Lemma 6 *Suppose* $h \le \frac{1}{12\sqrt{d}}$, *and* $\varepsilon = \frac{L}{150Md}$, *then if* $\|f_{0:T-1}\|_2 \le q_k$ *and* $\alpha = \left(\frac{4^{14}M^8d^2}{L^2}\right)^d$, *the running Alg. 1 has states bounded by*

$$\|x_{1:e_k}\|_2 \le \alpha q_k \,.$$

The restart mechanism of the algorithm eventually assures us that $q_k \approx \|f_{1:T-1}\|_2$ up to a multiplicative factor, providing an ℓ_2-gain bound.

Handling changing disturbance budget (see App. E.7). We now sketch out the extension to unknown disturbance magnitude. In Alg 1, q is the proposed upper bound on $\|f_{0:T-1}\|_2$. There are a variety of conditions for failure in the algorithms (i.e. where we have proof that q was not a valid upper bound) which trigger re-exploration and the start of a new epoch. If q is indeed an upper bound, the above steps all will work without triggering a failure and we have $\|x_{1:T}\|_2 \le \alpha q$ for some constant α. On the other hand, when a failure is detected, it is proof that $\|f_{0:T-1}\|_2 > q$. We can relate the penultimate budget q' to the final budget q by bounding the state growth from a single time evolution where budget is exceeded. Combining the upper bound of $\|x_{1:T}\|_2$ and lower bound on $\|f_{0:T-1}\|_2$ produces an ℓ_2-gain bound.

5. Conclusions

We have shown that for fully actuated systems, it is possible to control a misspecified LDS with robustness that is independent of the system magnitude, going beyond the small gain theorem, with an efficient algorithm. In addition, our control algorithm has near-optimal dimension dependence in terms of ℓ_2-gain, improving upon the classical algorithm of Cusumano and Poolla (1988b).

The most important open question is to continue this investigation to the much more general case of underactuated systems. Are efficient and optimally-robust algorithms possible? Can an efficient algorithm can be derived to obtain constant robustness, independent of the dimension, and with a tighter bound on ℓ_2-gain in terms of the system magnitude?

Other future directions include systems with partial observability and degenerate control matrices. It is also interesting to explore whether the same result can be obtained when the system inputs, not only the states, are subject to noise and misspecification.

References

Naman Agarwal, Brian Bullins, Elad Hazan, Sham Kakade, and Karan Singh. Online control with adversarial disturbances. In *International Conference on Machine Learning*, pages 111–119, 2019.

Naman Agarwal, Elad Hazan, Anirudha Majumdar, and Karan Singh. A regret minimization approach to iterative learning control. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 100–109. PMLR, 18–24 Jul 2021.

Dennis S. Bernstein and Wassim M. Haddad. Is there more to robust control theory than small gain? In *1992 American Control Conference*, pages 83–84, 1992. doi: 10.23919/ACC.1992.4792025.

Nicholas M. Boffi, Stephen Tu, and Jean-Jacques E. Slotine. Regret bounds for adaptive nonlinear control. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pages 471–483. PMLR, 07 – 08 June 2021.

Xinyi Chen and Elad Hazan. Black-box control for linear dynamical systems. In *Conference on Learning Theory*, pages 1114–1143. PMLR, 2021.

Xinyi Chen, Udaya Ghai, Elad Hazan, and Alexandre Megretski. Robust online control with model misspecification. *arXiv preprint arXiv:2107.07732*, 2021.

Alon Cohen, Avinatan Hasidim, Tomer Koren, Nevena Lazic, Yishay Mansour, and Kunal Talwar. Online linear quadratic control. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1029–1038. PMLR, 10–15 Jul 2018.

S. J. Cusumano and K. Poolla. Adaptive control of uncertain systems: A new approach. In *Proceedings of the American Automatic Control Conference*, pages 355–359, June 1988a.

S.J. Cusumano and K. Poolla. Nonlinear feedback vs. linear feedback for robust stabilization. In *Proceedings of the 27th IEEE Conference on Decision and Control*, pages 1776–1780 vol.3, 1988b. doi: 10.1109/CDC.1988.194633.

M. K. S. Faradonbeh, A. Tewari, and G. Michailidis. Finite-time adaptive stabilization of linear systems. *IEEE Transactions on Automatic Control*, 64(8):3498–3505, 2019.

Udaya Ghai, Holden Lee, Karan Singh, Cyril Zhang, and Yi Zhang. No-regret prediction in marginally stable systems. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1714–1757. PMLR, 09–12 Jul 2020.

Gautam Goel and Babak Hassibi. Competitive control. *arXiv preprint arXiv:2107.13657*, 2021.

Elad Hazan. Lecture notes on online and nonstochastic control theory, 2021.

Elad Hazan, Sham Kakade, and Karan Singh. The nonstochastic control problem. In *Algorithmic Learning Theory*, pages 408–421. PMLR, 2020.

Petros A Ioannou and Jing Sun. *Robust adaptive control*. Courier Corporation, 2012.

Sham Kakade, Akshay Krishnamurthy, Kendall Lowrey, Motoya Ohnishi, and Wen Sun. Information theoretic regret bounds for online nonlinear control. In *Advances in Neural Information Processing Systems*, volume 33, pages 15312–15325. Curran Associates, Inc., 2020.

Horia Mania, Michael I. Jordan, and Benjamin Recht. Active learning for nonlinear system identification with guarantees. *arXiv preprint arXiv:2006.10277*, 2020.

Alexandre Megretski and Anders Rantzer. Lower and upper bounds for optimal l2 gain nonlinear robust control of first order linear system. Technical Report No. 41, Institut Mittag-Leffler, 2002/2003.

Zakaria Mhammedi, Dylan J Foster, Max Simchowitz, Dipendra Misra, Wen Sun, Akshay Krishnamurthy, Alexander Rakhlin, and John Langford. Learning the linear quadratic regulator from nonlinear observations. In *Advances in Neural Information Processing Systems*, volume 33, pages 14532–14543. Curran Associates, Inc., 2020.

S. Oymak and N. Ozay. Non-asymptotic identification of lti systems from a single trajectory. In *2019 American Control Conference (ACC)*, pages 5655–5661, 2019.

Charles E. Rohrs, Lena Valavani, Michael Athans, and Gunter Stein. Robustness of adaptive control algorithms in the presence of unmodeled dynamics. In *1982 21st IEEE Conference on Decision and Control*, pages 3–11, 1982. doi: 10.1109/CDC.1982.268392.

Tuhin Sarkar and Alexander Rakhlin. Near optimal finite time identification of arbitrary linear dynamical systems. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 5610–5618, Long Beach, California, USA, 09–15 Jun 2019. PMLR.

Guanya Shi, Yiheng Lin, Soon-Jo Chung, Yisong Yue, and Adam Wierman. Online optimization with memory and competitive control. In *Advances in Neural Information Processing Systems*, volume 33, pages 20636–20647. Curran Associates, Inc., 2020.

Max Simchowitz, Horia Mania, Stephen Tu, Michael I. Jordan, and Benjamin Recht. Learning without mixing: Towards a sharp analysis of linear system identification. In *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 439–473. PMLR, 06–09 Jul 2018.

Max Simchowitz, Ross Boczar, and Benjamin Recht. Learning linear dynamical systems with semi-parametric least squares. In *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, pages 2714–2802, Phoenix, USA, 25–28 Jun 2019. PMLR.

Max Simchowitz, Karan Singh, and Elad Hazan. Improper learning for non-stochastic control. In *Conference on Learning Theory*, pages 3320–3436. PMLR, 2020.

Rohan Sinha, James Harrison, Spencer M. Richards, and Marco Pavone. Adaptive robust model predictive control with matched and unmatched uncertainty. *arXiv preprint arXiv:2104.08261*, 2021.

Gang Tao. Multivariable adaptive control: A survey. *Automatica*, 50:2737–2764, 11 2014.

Roman Vershynin. *Introduction to the non-asymptotic analysis of random matrices*, page 210–268. Cambridge University Press, 2012. doi: 10.1017/CBO9780511794308.006.

Chenkai Yu, Guanya Shi, Soon-Jo Chung, Yisong Yue, and Adam Wierman. Competitive control with delayed imperfect information. *arXiv preprint arXiv:2010.11637*, 2020.

G. Zames. On the input-output stability of time-varying nonlinear feedback systems part one: Conditions derived using concepts of loop gain, conicity, and positivity. *IEEE Transactions on Automatic Control*, 11(2):228–238, 1966. doi: 10.1109/TAC.1966.1098316.

Kemin Zhou, John C. Doyle, and Keith Glover. *Robust and Optimal Control*. Prentice-Hall, Inc., USA, 1996. ISBN 0134565673.