Proceedings of Machine Learning Research vol 144:1–35, 2022

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*

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

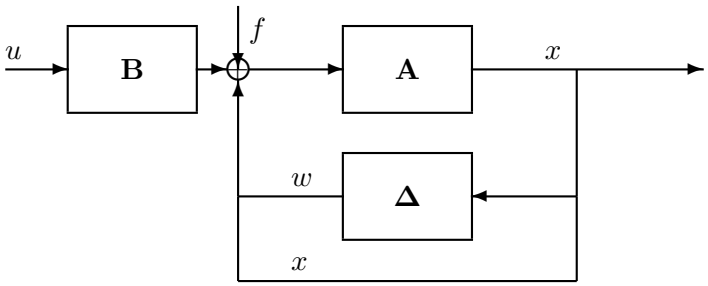

Figure 1: Diagram of the system, where Δ represents model misspecification.

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 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.

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.

Acknowledgments

Appendix A. Small Gain Theorem

The Small Gain Theorem (Zames, 1966) provides a guarantee on the stability on an interconnection of two stable systems, denoted $\mathbf{S}_\Delta$ and depicted in Fig. 2. System S takes as input (f, w) and produces output (x, y) and $\mathbf{\Delta}$ takes as input x and produces output w. The joint system $\mathbf{S}_\Delta$ can be viewed as taking input f and producing output y. If the ℓ_2-gain of S is at most γ, the Small Gain Theorem guarantees stability of $\mathbf{S}_\Delta$ so long as the ℓ_2-gain of $\mathbf{\Delta}$ is upper bounded by $\frac{1}{\gamma}$.

The Small Gain Theorem is an important tool in understanding robustness. For a give closed-loop controlled system $\mathbf{S}$, the coupled system $\mathbf{\Delta}$ can viewed as model misspecification. The Small Gain Theorem gives a prescription for robust control: design a controller such that the closed-loop system $\mathbf{S}$ has small ℓ_2-gain and robustness follows.

While this methodology is appealing, unfortunately in our setting, such an approach yields quite weak bounds. In particular, recent lower bounds for adaptive control without model misspecification scales with the parametric uncertainty as $M^{\Omega(d)}$. As such, the best robustness we can hope for using a small-gain approach is on the order of $\frac{1}{M^d}$, vanishing as the parametric uncertainty grows. In contrast, the algorithms in this work more directly tackle model misspecification, attaining robustness independent of the parametric uncertainty, albeit only for fully actuated systems.

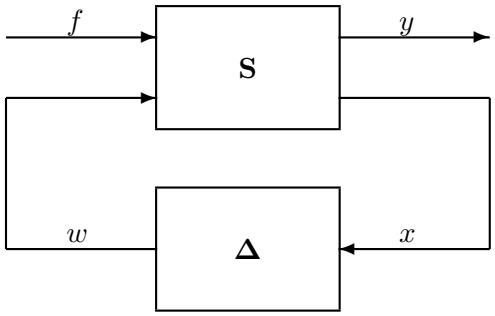

Figure 2: Diagram of an interconnected system $\mathbf{S}_\Delta$

For completeness, we provide a version of the Small Gain Theorem.

Theorem 7 (Small Gain) *If ℓ_2-gain of $\mathbf{S}$ is not larger than γ and ℓ_2-gain of $\mathbf{\Delta}$ is not larger than $\frac{1}{\gamma}$, then the ℓ_2-gain of $\mathbf{S}_\Delta$ is not larger than γ.*

Proof By the provided gain bounds, there exists constants C_1, C_2 such that

$$\sum_{t=1}^{\infty} \left(\gamma^2(\|f_t\|_2^2 + \|w_t\|_2^2) - (\|x_t\|^2 + \|y_t\|_2^2)\right) > C_1 \tag{4}$$

$$\sum_{t=1}^{\infty} \left(\frac{1}{\gamma^2}\|x_t\|_2^2 - \|w_t\|^2\right) > C_2 \tag{5}$$

Scaling (5) by γ^2 and adding to (4), we have

$$\sum_{t=1}^{\infty} \left(\gamma^2 \|f_t\|_2^2 - \|y_t\|_2^2\right) > \gamma^2 C_2 + C_1 \ .$$

Thus, the ℓ_2-gain of $\mathbf{S}_\Delta$ is not larger than γ. ■

Appendix B. Limits on robustness in online control

In this subsection we give a simple example exhibiting the limitation of robustness, and in particular showing that in the case of an unstabilizable system, it is impossible to obtain constant robustness.

Definition 8 (Strong Controllability) *Given a linear time-invariant dynamical system (A, B), let C_k denote*

$$C_k = [B \ AB \ A^2B \ \cdots A^{k-1}B] \in \mathbb{R}^{d \times kd}.$$

Then (A, B) is (k, κ) strongly controllable if C_k has full row-rank, and $\|(C_k C_k^\top)^{-1}\| \leq \kappa$.

Lemma 9 *In general, a system with strong controllability (k, κ) cannot be controlled with robustness larger than $\frac{1}{\sqrt{\kappa}}$.*

Proof

Consider the two dimensional system given by the matrices

$$A_\varepsilon = \begin{bmatrix} 2 & \varepsilon \\ 0 & 2 \end{bmatrix} \ , \ B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The Kalman matrix for this system is given by

$$Q = [B \ \ AB] = \begin{bmatrix} 0 & \varepsilon \\ 1 & 2 \end{bmatrix}$$

For $\varepsilon > 0$, this matrix is full rank, and the system is strongly controllable with parameters $(2, O(\frac{1}{\varepsilon^2}))$. However, for $\varepsilon = 0$, it can be seen that the system becomes uncontrollable even without any noise, since the first coordinate has no control which can cancel it, i.e. $x_{t+1}(1) = 2x_t(1) + z_t(1)$.

For adversarial noise with robustness of ε, we can convert the system A_ε to A_0, rendering it uncontrollable. The noise sequence will simply be

$$w_t = \begin{bmatrix} 0 & -\varepsilon \\ 0 & 0 \end{bmatrix} x_t.$$

This happens with parameter h which is $\varepsilon = \frac{1}{\sqrt{\kappa}}$. ■

Appendix C. Relating competitive ratio to $\ell_2$gain

As discussed, the notion of ℓ_2-gain has a very similiar spirit to a competitive ratio. Here we relate the ℓ_2-gain to the competitive ratio concretely for quadratic costs. We begin with a formal definition.

Definition 10 *(Competitive Ratio) Consider a sequence of cost functions $c_t(x_t, u_t)$. Let $J_T(\mathcal{A}, f_{0:T-1})$ denote the cost of controller $\mathcal{A}$ given the disturbance sequence $f_{0:T-1}$, and let $\mathrm{OPT}(f_{0:T-1})$ denote the cost of the offline optimal controller with full knowledge of $f_{0:T-1}$. Both costs are worst case under any model misspecification that satisfies (2) subject to a fixed $f_{0:T-1}$. The competitive ratio of a control algorithm $\mathcal{A}$, for $w_{1:T-1}$ satisfying Assumption 1 is defined as:*

$$\mathcal{C}(\mathcal{A}) = \max_{f_{0:T-1}} \frac{J_T(\mathcal{A}, f_{0:T-1})}{\mathrm{OPT}(f_{0:T-1})}.$$

The ℓ_2-gain bounds the ratio between $\|x_{1:T}\|_2$ and $\|f_{0:T-1}\|_2$, while under the time-invariant cost function $c_t(x, u) = \|x\|_2^2 + \|u\|_2^2$, the competitive ratio bounds the ratio of $\|x_{1:T}\|_2^2 + \|u_{1:T}\|_2^2$ to $\mathrm{OPT}(f_{0:T-1})$. Here we show that $\mathrm{OPT}(f_{0:T-1}) = \Theta(\|f_{0:T-1}\|_2^2)$, treating M and L as constants. Assuming $\|u_{1:T}\|_2$ is bounded by a constant multiple of $\|x_{1:T}\|_2$, then $\mathcal{C}(\mathcal{A}) = \Theta(\ell_2\text{-gain}(\mathcal{A})^2)$.

Theorem 11 *Under the time-invariant cost function $c_t(x, u) = \|x\|_2^2 + \|u\|_2^2$, for any system satisfying Assumptions 1 and 2, with $h < 1/2$,*

$$\frac{\|f_{0:T-1}\|_2^2}{9M^2} \le \mathrm{OPT}(f_{0:T-1}) \le \frac{8M^2\|f_{0:T-1}\|_2^2}{L^2}\,.$$

Proof We start by bounding $\|f_t\|_2^2$ using $(a+b+c+d)^2 \le 4(a^2+b^2+c^2+d^2)$.

$$\begin{aligned}\|f_t\|_2^2 &= \|x_{t+1} - Ax_t - Bu_t - w_t\|_2^2 \\ &\le 4M^2\|x_t\|_2^2 + 4\|x_{t+1}\|_2^2 + 4M^2\|u_t\|_2^2 + 4\|w_t\|_2^2\end{aligned}$$

Summing over f_t^2, we have

$$\begin{aligned}\|f_{0:T-1}\|_2^2 = \sum_{t=0}^{T-1}\|f_t\|_2^2 &\le 4\sum_{t=0}^{T-1}(M^2\|x_t\|_2^2 + \|x_{t+1}\|_2^2 + M^2\|u_t\|_2^2 + \|w_t\|_2^2) \\ &\le 8M^2(\|x_{1:T}\|_2^2 + \|u_{1:T-1}\|_2^2) + 4\|w_{1:t}\|_2^2 \\ &\le (8M^2 + 4h^2)(\|x_{1:T}\|_2^2 + \|u_{1:T-1}\|_2^2)\,.\end{aligned}$$

The lower bound follows after applying $2h < 1 \le M$.

For the upper bound, consider $u_t = -B^{-1}Ax_t$, which produces closed loop dynamics $x_{t+1} = w_t + f_t$ and hence $\|x_{t+1}\|_2^2 \le 2\|w_t\|_2^2 + 2\|f_t\|_2^2$. Summing over t, we have

$$\|x_{1:T}\|_2^2 \le 2\|f_{0:T-1}\|_2^2 + 2\|w_{0:T-1}\|_2^2 \le 2\|f_{0:T-1}\|_2^2 + 2h^2\|x_{0:T-1}\|_2^2\,.$$

Noting that $x_0 = 0$, we have $\|x_{1:T}\|_2^2 \le \frac{2\|f_{0:T-1}\|_2^2}{(1-2h^2)} \le 4\|f_{0:T-1}\|_2^2$.

Noting that $\|u_t\|_2 \le \frac{M}{L}\|x_t\|_2$, we have

$$\|x_{1:T}\|_2^2 + \|u_{1:T-1}\|_2^2 \le \frac{2M^2\|x_{1:T}\|_2^2}{L^2} \le \frac{8M^2\|f_{0:T-1}\|_2^2}{L^2}\,.$$

■

Remark 12 *Dependence on M^2 is required in Theorem 11. Consider the system $x_{t+1} = Mx_t + u_t + f_t$ with $x_1 = 1$, $u_t = 0$ for all t and f_t alternates between $-M$ and 1. As a result, x_t oscillates between 1 and 0 for an average cost of $\frac{1}{2}$, while f_t^2 is on average $\frac{M^2+1}{2}$.*

Appendix D. One Dimensional Analysis

In this section,we show that for a simple one dimensional system constant robustness can be achieved using certainty equivalence control (i.e online least squares system identification) with tight dependence on the system uncertainty radius. We consider the following system

$$x_{t+1} = ax_t + u_t + w_t + f_t, \quad |w_{0:t}| \le h|x_{0:t}| \quad \forall\, t \in \mathbb{Z}_+, \quad x_0 = 0, \tag{6}$$

where w_t and f_t are similarly defined as in Section 1.1.

In this section, we use the notation $\gamma : [-M, M] \to [1, \infty)$ for some function (preferably, as small as possible), such that the following inequalities hold

$$|x_{:t}| \le \gamma(a)|f_{:t}| \quad \forall\, t \in \mathbb{Z}_+ \tag{7}$$

for an upper bound on the ℓ_2-gain.

D.1. Lower Bound

First, we formulate (and prove) a stronger version of the result of Megretski and Rantzer (2002/2003), for the case $h = 0$ (which means that $w \equiv 0$ in (6)).

Theorem 13 *If a control algorithm $\mathcal{A}$ has finite ℓ_2-gain bounds specified by $\gamma : [-M, M] \to [0, \infty)$ with $h = 0$, for all $a \in [-M, M]$ and $f_t, w_t, x_t \in \mathbb{R}$ satisfying (6), then $\gamma(a) \ge \max\{|a - M|, |a + M|\}/8 \ge M/8$ for all $a \in [-M, M]$, and therefore ℓ_2-gain$(\mathcal{A}) \ge M/8$.*

The result of Thm. 13 suggests that the closed loop ℓ_2-gain, no matter which adaptive controller is used, if it obtains finite ℓ_2-gain for all systems satisfying the parametric uncertainty, must grow linearly with the parametric uncertainty size for all values of the uncertain parameter. In particular, this makes it impossible to sacrifice performance at some values of a to gain much improvement at other values of a. Given signals x, f satisfying equations (6) with $w_t \equiv 0$, define ξ_t, p_t, q_t, r_t by

$$\xi_t = x_{t+1} - u_t, \qquad p_t = \sum_{s=0}^{t-1} x_s^2, \qquad q_t = \sum_{s=0}^{t-1} x_s(x_{s+1} - u_s), \qquad r_t = \sum_{s=0}^{t-1} \xi_s^2.$$

The interaction between control (u_t) and noise (f_t) can be viewed as a game (between u and f), in which u_t is decided, based on knowing all x_s with $s \le t$ (and, of course, all u_s with $s < t$), to keep the inequality

$$\gamma(a)^2 \left[p_t a^2 - 2q_t a + r_t\right] \ge p_t \tag{8}$$

satisfied for all $a \in [-M, M]$ and all $t \in \mathbb{Z}_+$, while f_t is decided based on knowing all u_s and x_s with $s \le t$ (and all f_s with $s < t$), in an effort to violate the inequality from (8) at some time $t \in \mathbb{Z}_+$ and some $a \in [-M, M]$.

In this proof, we work with the *normalized* versions δ_t, ν_t, z_t of u_t, ξ_t, x_t, as well as additional signals β_t, θ_t. Let t_0 be the smallest $t \in \mathbb{Z}_+$ such that $p_t > 0$. For $t \geq t_0$ let

$$\beta_t = \frac{q_t}{p_t}, \qquad \theta_t = \frac{r_t p_t - q_t^2}{p_t^2}, \qquad z_t = \frac{x_t}{\sqrt{p_t}}, \qquad \delta_t = \frac{u_t + \beta_t x_t}{\sqrt{p_t + x_t^2}}, \qquad \nu_t = \frac{\xi_t - \beta_t x_t}{\sqrt{p_t + x_t^2}}.$$

Since $x_t, \xi_t, u_t, p_t, q_t, r_t$ satisfy equations

$$x_{t+1} = \xi_t + u_t,\ p_{t+1} = p_t + x_t^2,\ q_{t+1} = q_t + x_t\xi_t,\ r_{t+1} = r_t + \xi_t^2,$$

δ_t, ν_t, z_t, β_t, θ_t, for $t \geq t_0$, satisfy

$$z_{t+1} = \nu_t + \delta_t, \qquad \beta_{t+1} = \beta_t + \frac{z_t\nu_t}{\sqrt{1+z_t^2}}, \qquad \theta_{t+1} = \frac{\theta_t + \nu_t^2}{1+z_t^2}. \tag{9}$$

The game between u and f can now be interpreted as the game between δ and ν, defined by the dynamical equations (9) with the state $y_t = (z_t, \beta_t, \theta_t)$, where the normalized control effort δ_t, for $t \geq t_0$, is best constructed as a function $\delta_t = S_t(y_t)$ of the current state, to keep the inequality

$$\gamma(a)^2\left[(a-\beta_t)^2 + \theta_t\right] \geq 1 \tag{10}$$

satisfied for all $a \in [-M, M]$ and all $t \geq t_0$, while ν_t, for $t \geq t_0$, is best constructed as a function $\nu_t = D_t(y_t, \delta_t)$ of current normalized control effort and the current state, in an effort to violate the inequality from (10) at some time $t \geq t_0$ and some $a \in [-M, M]$.

Claim 1: *If a control algorithm $\delta_t = S_t(y_t)$ maintains (10) for $a = a_0$ and for all $t \geq t_0$ then it also satisfies*

$$\gamma(a_0)^2\left[(a_0-\beta_t)^2 + \theta_t\right] \geq 1 + z_t^2 \tag{11}$$

for all $t \geq t_0$. Indeed, with $\nu_t = \frac{(a_0-\beta_t)z_t}{\sqrt{1+z_t^2}}$ we have

$$\gamma(a_0)^2\left[(a_0-\beta_{t+1})^2 + \theta_{t+1}\right] = \gamma(a_0)^2\left[\left|a_0 - \beta_t - \frac{z_t\nu_t}{\sqrt{1+z_t^2}}\right|^2 + \frac{\theta_t + \nu_t^2}{1+z_t^2}\right] = \gamma(a_0)^2\frac{(a_0-\beta_t)^2 + \theta_t}{1+z_t^2},$$

hence (11) must be satisfied to maintain (10).

Claim 2: *If a control algorithm $\delta_t = S_t(y_t)$ maintains (11) for all $t \geq t_0$ then it also satisfies*

$$\gamma(a_0)^2\left[\frac{\theta_t}{1+z_t^2} + \frac{(a_0-\beta_t)^2}{z_t^2}\right] - 1 \geq \left|\delta_t + \frac{(a_0-\beta_t)\sqrt{1+z_t^2}}{z_t}\right|^2 \tag{12}$$

whenever $z_t \neq 0$, for all $t \geq t_0$. Indeed, when $z_t \neq 0$, with $\nu_t = \frac{(a_0-\beta_t)\sqrt{1+z_t^2}}{z_t}$ we have

$$\gamma(a_0)^2\left[(a_0-\beta_{t+1})^2 + \theta_{t+1}\right] - 1 - z_{t+1}^2 = \gamma(a_0)^2\left[\frac{\theta_t}{1+z_t^2} + \frac{(a_0-\beta_t)^2}{z_t^2}\right] - 1 - \left|\delta_t + \frac{(a_0-\beta_t)\sqrt{1+z_t^2}}{z_t}\right|^2,$$

hence (12) must be satisfied to maintain (11).

To continue the proof, take any point $a_0 \in [-M, M]$, and let $\gamma_0 = \gamma(a_0) \geq 1$. We aim to show that, for a sufficiently small $\mu > 0$,

(A) an appropriate adversary strategy $\nu_t = D_t(y_t, \delta_t)$, assures that $\theta_t \to 0$, while β_t stays (for sufficiently large t) within the interval $[a_0 + (4+\mu)\gamma_0, a_0 + (8+3\mu)\gamma_0]$, no matter which normalized control algorithm $\delta_t = S_t(y_t)$ is used.

(B) an appropriate adversary strategy $\nu_t = D_t(y_t, \delta_t)$, assures that $\theta_t \to 0$, while β_t stays (for sufficiently large t) within the interval $[a_0 - (8+3\mu)\gamma_0, a_0 - (4+\mu)\gamma_0]$, no matter which normalized control algorithm $\delta_t = S_t(y_t)$ is used.

Combining assertions (A) and (B) (and making $\mu > 0$ sufficiently small) assures that, as long as $\gamma(a)$ is finite for all $a \in [-M, M]$, the value of $8\gamma(a)$ cannot be smaller than the distance from a to either end of the $[-M, M]$ interval, implying the desired lower bound for $\gamma(a)$.

We are only presenting the strategy for selecting ν_t in (A), as the construction for (B) is symmetric. Let I_0 denote the interval $[a_0 + (4+\mu)\gamma_0, a_0 + (8+3\mu)\gamma_0]$. For $t \geq t_0$, define ν_t by

$$\nu_t = 0, \qquad \text{when } |\delta_t| \geq 2\gamma_0,\ \beta_t \in I_0, \tag{13}$$

$$\nu_t = -\mu\gamma_0 \text{sign}[z_t(\beta_t - a_0 - (6+2\mu)\gamma_0)], \qquad \text{when } |\delta_t| \geq (2+\mu)\gamma_0,\ \beta_t \notin I_0, \tag{14}$$

$$\nu_t = -(4+\mu)\gamma_0 \text{sign}[z_t(\beta_t - a_0 - (6+2\mu)\gamma_0)], \qquad \text{otherwise}, \tag{15}$$

where the "sign" function is defined by $\text{sign}(x) = 1$ for $x \geq 0$, $\text{sign}(x) = -1$ for $x < 0$, i.e., only takes values 1 or -1. Intuitively, the adversary strategy (13)-(15) pursues the following three objectives:

(a) keep $|z_t| \geq 2\gamma_0$ at all times;

(b) force $\beta_t \in I_0$, eventually;

(c) make $\nu_t = 0$ when objectives (a) and (b) are satisfied.

Claim 3: *Subject to (9) and (13)-(15), condition $|z_t| \geq 2\gamma_0$ will be satisfied for all $t > t_0$.* Indeed,

(3a): in (13), $|\delta_t| \geq 2\gamma_0$ and $\nu_t = 0$, hence $|z_{t+1}| = |\delta_t + \nu_t| = |\delta_t| \geq 2\gamma_0$.

(3b): in (14), $|\delta_t| \geq (2+\mu)\gamma_0$ and $|\nu_t| = \mu\gamma_0$, hence $|z_{t+1}| = |\delta_t + \nu_t| \geq |\delta_t| - |\nu_t| \geq 2\gamma_0$.

(3c): in (15), $|\delta_t| \leq (2+\mu)\gamma_0$ and $|\nu_t| = (4+\mu)\gamma_0$, hence $|z_{t+1}| = |\delta_t + \nu_t| \geq |\delta_t| - |\nu_t| \geq 2\gamma_0$.

Claim 4: *Subject to (9), (13)-(15), and assuming $\mu \in (0, 4)$, there exists $t_1 > t_0$ such that $\theta_t < 4 + 3\mu$ for all $t > t_1$.* Indeed, since $|\nu_t| \leq (4+\mu)\gamma_0$ and $|z_t| \geq 2\gamma_0$ for all $t > t_0$, we have

$$\theta_{t+1} - \frac{(4+\mu)^2}{4} = \frac{\theta_t + \nu_t^2}{1+z_t^2} - \frac{(4+\mu)^2}{4} \leq \frac{\theta_t + (4+\mu)^2\gamma_0^2}{1+4\gamma_0^2} - \frac{(4+\mu)^2}{4} = \frac{1}{1+4\gamma_0^2}\left(\theta_t - \frac{(4+\mu)^2}{4}\right)$$

for all $t > t_0$, which leads to the conclusion, since $4 + 3\mu > (4+\mu)^2/4$ for all $\mu \in (0, 4)$.

Claim 5: *If $0 < \mu < \gamma_0^{-2}/3$, and $\beta_t \in I_0$ for some $t > t_1$ then $|\delta_t| \geq 2\gamma_0$ (and therefore $\nu_t = 0$, $\beta_{t+1} = \beta_t$).* Indeed, combining $|z_t| \geq 2\gamma_0$ with $\theta_t < 4 + 3\mu$ shows that

$$\frac{\gamma_0^2 \theta_t}{1 + z_t^2} - 1 \leq \frac{\gamma_0^2(4 + 3\mu)}{1 + 4\gamma_0^2} - 1 = \frac{3\mu\gamma_0^2 - 1}{1 + 4\gamma_0^2} < 0,$$

hence, using (12),

$$\begin{aligned} |\delta_t| &\geq \left| \frac{(\beta_t - a_0)\sqrt{1 + z_t^2}}{z_t} \right| - \left| \delta_t - \frac{(\beta_t - a_0)\sqrt{1 + z_t^2}}{z_t} \right| \\ &\geq \frac{|\beta_t - a_0|\sqrt{1 + z_t^2}}{|z_t|} - \frac{|\beta_t - a_0|\gamma_0}{|z_t|} = |\beta_t - a_0| \frac{\sqrt{1 + z_t^2} - \gamma_0}{|z_t|}. \end{aligned}$$

Since, for $\gamma_0 \geq 1$, the function $\phi : (0, \infty) \to \mathbb{R}$ defined by $\phi(z) = \frac{\sqrt{1+z^2}-\gamma_0}{z}$ has positive derivative

$$\dot{\phi}(z) = \frac{1}{z^2} \left(\gamma_0 - \frac{1}{\sqrt{1 + z^2}} \right) > 0,$$

it is monotonically increasing on $(0, \infty)$, which transofrms the lower bound for $|\delta_t|$ into

$$|\delta_t| \geq |\beta_t - a_0| \frac{\sqrt{1 + z_t^2} - \gamma_0}{|z_t|} \geq (4 + \mu)\gamma_0 \frac{\sqrt{1 + 4\gamma_0^2} - \gamma_0}{2\gamma_0} > (2 + \mu/2)\gamma_0 > 2\gamma_0.$$

Claim 5 establishes that, once $\beta_\tau \in I_0$ for some $\tau > t_1$, the equalities $\beta_t = \beta_\tau \in I_0$ and $\nu_t = 0$ will hold for all $t \geq \tau$, which will guarantee that $\theta_{t+1} = \theta_t/(1 + z_t^2) \leq \theta_t/(1 + 4\gamma_0^2) \leq \theta_t/5$ will converge to zero as $t \to +\infty$, thus preventing (10) with $a = \beta_\tau$ from being satisfied when t is large enough, no matter how large $\gamma(\beta_\tau)$ is.

To finish the proof, we need to show that condition $\beta_t \in I_0$ will be satisfied for some $t > t_1$. Indeed, with $\beta_t \notin I_0$, the value of ν_t will be defined by either (14) or (15). Taking into accout that, for $|z_t| \geq 2\gamma_0 \geq 2$, the value of $|z_t|/\sqrt{1 + z_t^2}$ is between 0.5 and 1,

- **Case 1:** according to (14) and (9), β_{t+1} results from moving β_t, which is at least $(2 + \mu)\gamma_0$ away from the center $c_0 = a_0 + (6 + 2\mu)\gamma_0$ of I_0, by a distance between $0.5\mu\gamma_0$ and $\mu\gamma_0$ in the direction "to the center", which ensures that $|\beta_{t+1} - c_0| \leq |\beta_t - c_0| - \mu/2$.

- **Case 2:** according to (15) and (9), β_{t+1} results from moving β_t, which is at least $(2 + \mu)\gamma_0$ away from c_0, by a distance between $(2 + \mu/2)\gamma_0$ and $(4 + \mu)\gamma_0$ in the direction "to the center", which ensures that $|\beta_{t+1} - c_0| \leq |\beta_t - c_0| - \mu$.

Therefore, as long as $t > t_1$ and $\beta_t \notin I_0$, the value of $|\beta_t - c_0|$ decreases by at least $\mu/2$ at each step, thus guaranteeing that condition $\beta_t \in I_0$ will be satisfied for some $t > t_1$.

D.2. Upper Bound

In this section, we show that, for $h < 1/2$, the so-called *certainty equivalence controller* achieves an upper bound for the closed loop ℓ_2 gain which grows linearly with M. This is remarkable, as

a large ℓ_2-gain is usually associated with low robustness to uncertain dynamical feedback, with an expectation that a system with ℓ_2-gain γ can be destabilized by a feedback with ℓ_2-gain of $1/\gamma$ (which is certainly true for *linear time invariant* systems). Since, with $h = 0$, ℓ_2-gain of the closed loop is shown to be at least $M/8$, one would expect that some uncertainty $\mathbf{\Delta}$ of ℓ_2-gain $h = 8M^{-1}$ would destabilize the system, but this is evidently not happening: according to Thm. 14 below, all uncertainty ℓ_2-gains below $h = 0.5$ are well tolerated by the certainty equivalence controller (Alg. 4).

Algorithm 4: Certainty Equivalence Control

Input: Time horizon T, system upper bound parameter M.

Initialize $x_0, u_0 = 0$

for $t = 1 \dots T$ **do**

  Observe x_t and define $\tilde{z}_{t-1}(\hat{a}) = x_t - \hat{a}x_{t-1} - u_{t-1}$.

  Compute $\bar{a}_t = \arg\min_{\hat{a}} \sum_{s=0}^{t-1} \tilde{z}_s^2(\hat{a})$

  Compute $\hat{a}_t = \text{clip}_{[-M,M]}(\bar{a}_t)$

  Execute $u_t = -\hat{a}_t x_t$.

end

Theorem 14 *For $M \geq 1/4$, with a system satisfying condition (6) with $h < 1/2$, Alg. 4 has*

$$\ell_2\text{-gain}(\mathcal{A}) \leq \frac{(64M^2 - 8h)^{1/2}}{(1-2h)^{3/2}}\,.$$

We provide an intuitive explanation why certainty equivalence can provide an ℓ_2-gain bound. If the algorithm estimates $\hat{a}$ inaccurately, strong convexity of the one dimensional least squares objective implies that the magnitude of the disturbances is a nontrivial fraction of the magnitude of the states up to that point. On the other hand, if $\hat{a}_{t+1}$ is an accurate estimate of a, we can bound $\|x_{1:t}\|_2^2$ using the stability of the closed loop dynamics. An ℓ_2-gain bound follows from stitching these regimes together. While we would like to extend these ideas to high dimensions, we note that the least squares objective is no longer strongly convex in such a setting. In particular, $\|A - \hat{A}_t\|_2$ can be large in a direction where disturbances are small relative to the magnitude of the state. A more technical approach that yields the tighter bound of Thm. 14 can be found in App. D.2. We restate Alg. 4 in closed form:

$$\tilde{a}_t = \text{clip}_{[-M,M]}(\hat{a}_t), \quad \hat{a}_t = \begin{cases} Q_t/Y_t, & Y_t \neq 0, \\ 0, & X_t = 0, \end{cases} \qquad X_t = \sum_{s=0}^{t-1} x_s^2, \quad Q_t = \sum_{s=0}^{t-1} (x_{s+1} - u_s)x_s, \tag{16}$$

Note that $\hat{a}_t$ is the argument of minimum (with respect to $a \in \mathbb{R}$) of the "equation error" functional

$$V_t(a) = \sum_{s=0}^{t-1} |x_{t+1} - ax_t - u_t|^2 = \sum_{s=0}^{t-1} |w_s + f_s|^2, \tag{17}$$

computed with the data available to the controller at time t.

Let $f, w, x \in \ell$ be some signals satisfying (6), where, $M \geq 1/4$, and $h < 1/2$. Let $v = f + w$. In addition to variables x_t, Q_t defined in (16), define $V_t = V_t(a)$ as in (17), and let

$$H_t = x_t^2, \qquad Z_t = Z_t(a) = \max_{\tau \leq t} \{V_\tau(a) - hx_\tau/2\}, \qquad R_t = \sum_{s=0}^{t-1} f_s^2, \qquad W_t = \sum_{s=0}^{t-1} w_s^2.$$

In contrast with x_t, Q_t, and H_t, the values $V_t = V_t(a)$ and $Z_t = Z_t(a)$ are not available to the controller, as they depend not only on the past observations $x_{0:t}$, but on the unknown parameter $a \in [-M, M]$, and variables R_t and W_t depend not only on the past observations $x_{0:t}$ and $a \in [-M, M]$, but also on the unknown dynamics of $\mathbf{\Delta}$.

Step 1: Show that the relation between H_t, x_t, V_t, Z_t, v_t, and H_{t+1}, x_{t+1}, V_{t+1}, Z_{t+1} can be described, for all $t \in \mathbb{Z}_+$, by

$$\begin{aligned}
H_{t+1} &\leq \min\left\{8M^2, 2V_t/x_t\right\} H_t + 2v_t^2, \quad H_0 = 0, && (18)\\
x_{t+1} &= x_t + H_t, \quad x_0 = 0, && (19)\\
V_{t+1} &= V_t + v_t^2, \quad V_0 = 0, && (20)\\
Z_{t+1} &= \max\{Z_t,\ V_{t+1} - hx_{t+1}/2\}, \quad Z_0 = 0. && (21)
\end{aligned}$$

Indeed, equations (19)-(21) are evident. To prove (18), note first that, by the definition of $\hat{a}_t$,

$$V_t(a) = (a - \hat{a}_t)^2 x_t + \min_{b \in \mathbb{R}} V_t(b) \qquad \text{for all} \quad a \in \mathbb{R},$$

hence $V_t \geq (a - \hat{a}_t)^2 x_t$. Since $|a - \text{clip}_{-M,M]}(b)| \leq |a - b|$ for every $a \in [-M, M]$ and $b \in \mathbb{R}$, applying this to $b = \hat{a}$ yields $V_t \geq (a - \tilde{a}_t)^2 x_t$. Also, since both a and $\tilde{a}_t$ are in $[-M, M]$, we have $|a - \tilde{a}_t| \leq 2M$. Since the quadratic form $\sigma(p, q) = 2p^2 + 2q^2 - (p + q)^2$ is positive semidefinite, we have

$$x_{t+1}^2 = [(a - \tilde{a}_t)x_t + v_t]^2 \leq 2(a - \tilde{a}_t)^2 x_t^2 + 2v_t^2 \leq 2\min\left\{4M^2, V_t/x_t\right\} x_t^2 + 2v_t^2.$$

Step 2: Show that

$$Z_t \leq \frac{1}{1 - 2h} R_t \qquad \text{for all} \qquad t \in \mathbb{Z}_+. \tag{22}$$

Indeed, for $h = 0$ we have $w_t \equiv 0$, hence $V_t = R_t$, and the inequality (22) holds. For $h \in (0, 1/2)$, the quadratic form

$$\sigma(f, w) = \frac{1}{2h} w^2 + \frac{1}{1 - 2h} f^2 - (f + w)^2$$

is positive semidefinite. Hence, realizing that the L2 gain bound $|w_{0:t}| \leq h|x_{0:t}|$ means $h^{-2} W_{t+1} \leq x_{t+1}$, we get

$$V_t - 0.5hx_t \leq V_t - \frac{1}{2h} W_t = \sum_{s=0}^{t-1} \left\{ (f_s + w_s)^2 - \frac{1}{2h} w_s^2 \right\} \leq \frac{1}{1 - 2h} \sum_{s=0}^{t-1} f_s^2 = \frac{1}{1 - 2h} R_t.$$

Finally, since R_t is monotonically non-decreasing as t increases,

$$Z_t = \max_{\tau \leq t} \{V_\tau - hx_\tau/2\} \leq \max_{\tau \leq t} \frac{1}{1 - 2h} R_\tau = \frac{1}{1 - 2h} R_t.$$

Step 3: Use (18)-(20) to show that

$$D_t \stackrel{\text{def}}{=} H_t - 2V_t - (8M^2-1)x_t \le 0 \qquad \text{for all} \quad t \in \mathbb{Z}_+. \tag{23}$$

Indeed, $D_0 = H_0 - 2V_0 - (8M^2-1)x_0 = 0 \le 0$, and bounding D_{t+1} in terms of D_t yields

$$H_{t+1}-2V_{t+1}-(8M^2-1)x_{t+1} \le 8M^2H_t+2v_t^2-2(V_t+v_t^2)-(8M^2-1)(x_t+H_t) = H_t-2V_t-(8M^2-1)x_t.$$

Step 4: Use (18)-(21) to show that

$$C_t \stackrel{\text{def}}{=} \rho Z_t - 2H_t - x_t + 4V_t \ge 0 \quad \text{for all} \quad t \in \mathbb{Z}_+, \qquad \text{where} \qquad \rho = \frac{64M^2-4}{1-2h}. \tag{24}$$

Indeed, the inequality is evidently satisfied for $t=0$. Assuming that $\rho Z_t - 2H_t - x_t + 4V_t \ge 0$:

Case 4a: If $x_t \le M^{-2}V_t/4$ then $Z_t \ge V_t - hx_t/2 \ge (1-hM^{-2}/8)V_t \ge (64M^2-4)\rho^{-1}V_t$, hence

$$\begin{aligned} C_{t+1} &\ge \rho Z_t - 2(8M^2H_t + 2v_t^2) - (x_t+H_t) + 4(V_t+v_t^2) \\ &= \rho Z_t - (16M^2+1)H_t - x_t + 4V_t \\ &\ge \rho Z_t - (16M^2+1)[2V_t + V_t(8M^2-1)M^{-2}/4] - V_tM^{-2}/4 + 4V_t \\ &= \rho Z_t - (64M^2-4)V_t \ge 0. \end{aligned}$$

Case 4b: If $M^{-2}V_t/4 \le x_t \le 4$ then $Z_t \ge V_t - hx_t/2 \ge (1-2h)V_t = (64M^2-4)\rho^{-1}V_t$, hence

$$\begin{aligned} C_{t+1} &\ge \rho Z_t - 2[2H_tV_t/x_t + 2v_t^2] - (x_t+H_t) + 4(V_t+v_t^2) \\ &\ge \rho Z_t - [2V_t + (8M^2-1)x_t](1+4V_t/x_t) - x_t + 4V_t \\ &= \rho Z_t - (32M^2-6)V_t - 8M^2x_t - 8V_t^2/x_t. \end{aligned}$$

The last expression is a concave function of x_t, hence its values, with x_t ranging over the interval $[M^{-2}V_t/4,\ 4]$, are not smaller than its values at the ends of the interval, which are both equal to $\rho Z_t - (64M^2-4)V_t \ge 0$

Case 4c: if $x_t \ge 4$ then

$$C_{t+1} \ge \rho Z_t - 2[2H_tV_t/(4V_t) + 2v_t^2] - (x_t+H_t) + 4(V_t+v_t^2) = C_t \ge 0.$$

Step 5: using the inequalities $\rho Z_t \ge 2H_t + x_t - 4V_t$ and $Z_t \ge V_t - hx_t/2$, for $\delta = (1-2h)\rho/(4+\rho)$ we have either

Case 5a: $(1-\delta)x_t \ge 4V_t$, in which case $Z_t \ge \delta\rho^{-1}x_t = \frac{1-2h}{4+\rho}x_t$, or

Case 5b: $(1-\delta)x_t \le 4V_t$, in which case $Z_t \ge V_t - hx_t/2 \ge \frac{1-\delta-2h}{4}x_t = \frac{1-2h}{4+\rho}x_t$.

Combining these two observations together, we have

$$x_t \le \frac{4+\rho}{1-2h}Z_t \le \frac{4+\rho}{(1-2h)^2}R_t = \frac{64M^2-8h}{(1-2h)^3}R_t,$$

which proves that the closed loop L2 gain from f to y is not larger than $\frac{(64M^2-8h)^{1/2}}{(1-2h)^{3/2}}$.

Appendix E. Full Analysis

In this section we provide a complete analysis of our main algorithm. We present the algorithm including a generic exploration set and provide an alternate analysis when all control directions in an ε-net are explored. We denote an ε-net as $\mathcal{N}_{\varepsilon,d}$, defined as:

Definition 15 *We define $\mathcal{N}_{\varepsilon,d} \subseteq \mathbb{R}^d$ to be an ε-net of $\mathbb{S}^{d-1}$, the unit sphere with the euclidean metric, if for any $x \in \mathbb{S}^{d-1}$, we have $x' \in \mathcal{N}_{\varepsilon,d}$ such that $\|x - x'\|_2 \le \varepsilon$.*

Algorithm 5: ℓ_2-gain algorithm

Input: System upper bound M, control matrix singular value lower bound L, system identification parameter ε, threshold parameter α, and exploration set $V \subseteq \mathbb{S}^{d-1}$.

```
Set q = 0, K = 0.
while t ≤ T do
    Observe x_t.
    if ‖x_{1:t}‖_2 > αq then
        Update q = ‖x_{1:t}‖_2.
        Call Alg. 6 with parameters (q, M, L, ε, α, V), obtain updated K and budget q.
    else
        Execute u_t = −K x_t.
        t ← t + 1
    end
end
```

Remark 16 *We note that when V is the standard basis, $\hat{A}$ has the closed form used in Alg.2. In particular, the unconstrained solution[4] of Line 16 in Alg. 6 has $\Phi(\hat{A}) = 0$, where $\hat{A} = [\frac{x_{t'+2}}{\xi_0} \cdots \frac{x_{t'+2d}}{\xi_{d-1}}]$. When V is an ε-net, Φ is a maximum of convex functions, and hence a convex function.*

E.1. Epoch Notation.

We define epochs in terms of rounds of system identification. In particular, for the kth epoch s_k is t on the kth call to Alg. 6. and $e_k = \min(s_{k+1} - 1, T)$. 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. Correspondingly, we denote the value of q' in the kth epoch as q'_k.

E.2. Estimation of the Control Matrix

Lemma 17 *Suppose $\|f_{0:T-1}\|_2 \le q_k$ and $\alpha \ge 4^{2d} M^{2d} \varepsilon^{-d}$, then in Alg. 7, we have $\|x_{1:s_k+i}\|_2 \le 4^{2i} M^{2i} q_k \varepsilon^{-i}$, for $0 \le i \le d$.*

Proof We prove the lemma by induction. Note that if the lemma was true, no new epoch will start because $\|x_{1:t+i}\|_2 > \alpha q$ for any i. Now for the base case, note that for $i = 0$, the inequality holds

4. With small modifications to analysis, the constrained optimization can be replaced by a failure check if $\|\hat{A}\|_2 > 2M$ as this would indicate our disturbance budget is too small.

Algorithm 6: 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, and exploration set $V \subseteq \mathbb{S}^{d-1}$.

1 Define $N = |V| \geq d$ with $V = (v_0, v_1, \ldots, v_{N-1})$.

2 Call Alg. 7 with parameters $(q, M, L, \varepsilon, \alpha)$, obtain estimator $\hat{B}$ and updated budget q. Suppose the system evolves to time $t' = t + d$.

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

4 **for** *i = 0, 1, …, $2N-1$* **do**

5 Observe $x_{t'+i}$.

6 **if** $\|x_{1:t'+i}\|_2 > \alpha q$ **then**

7 Restart SysID from Line 2 with $q = \|x_{1:t'+i}\|_2$.

8 **end**

9 **if** *i is even* **then**

10 Play $u_{t'+i} = \xi_{i/2} \hat{B}^{-1} v_{i/2}$, $\xi_{i/2} = \frac{4^{3i/2} M^{3i/2+2} q'}{\varepsilon^{i/2+1}}$.

11 **else**

12 Play $u_{t'+i} = 0$.

13 **end**

14 **end**

15 Observe $x_{t'+2N}$, compute

$$\hat{A} \in \operatorname*{arg\,min}_{\tilde{A}: \|\tilde{A}\|_2 < M} \Phi(\tilde{A}) := \max_{i \in [0,N)} \|\tilde{A} v_i - \frac{x_{t'+2i+2}}{\xi_i}\|_2 .$$

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

trivially. Suppose the condition holds for i. For $i+1$, we have

$$\begin{aligned}
\|x_{s_k+i+1}\|_2 &= \|A x_{s_k+i} + B u_{s_k+i} + z_{s_k+i}\|_2 \\
&\leq M \|x_{s_k+i}\|_2 + M\lambda_i + h \|x_{1:s_k+i}\|_2 + q_k \\
&\leq 4^{2i} M^{2i+1} q_k \varepsilon^{-i} + 4^{2i} M^{2i+2} q_k \varepsilon^{-(i+1)} + h 4^{2i} M^{2i} q_k \varepsilon^{-i} + q_k \\
&\leq 4^{2i+1} M^{2i+2} q_k \varepsilon^{-(i+1)}
\end{aligned}$$

Adding previous iterations, we have

$$\|x_{1:s_k+i+1}\|_2 \leq 4^{2i+1} M^{2i+2} q_k \varepsilon^{-(i+1)} + 4^{2i} M^{2i} q_k \varepsilon^{-i} \leq 4^{2(i+1)} M^{2(i+1)} q_k \varepsilon^{-(i+1)}.$$

■

Lemma 18 *Suppose $\|f_{0:T-1}\|_2 \leq q_k$ and $\alpha \geq 4^{2d} M^{2d} \varepsilon^{-d}$, then running Alg. 7 with $\varepsilon \leq \frac{L}{12\sqrt{d}}$ produces $\hat{B}$ such that $\|\hat{B} - B\|_2 \leq 3\varepsilon\sqrt{d}$ and $\|B\hat{B}^{-1} - I\|_2 \leq \frac{6\varepsilon\sqrt{d}}{L} \leq \frac{1}{2}$, with $\|x_{1:s_k+d}\|_2 \leq 4^{2d} M^{2d} q_k \varepsilon^{-d}$.*

Algorithm 7: 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.

1 **for** *i = 0, 1, …,* $d-1$ **do**
2 Observe x_{t+i}.
3 **if** $\|x_{1:t+i}\|_2 > \alpha q$ **then**
4 Restart SysID with $q = \|x_{1:t+i}\|_2$.
5 **end**
6 Play $u_{t+i} = \lambda_i e_{i+1}$, $\lambda_i = \frac{4^{2i}M^{2i+1}q}{\varepsilon^{i+1}}$.
7 **end**
8 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**
9 Restart SysID with $q = \|x_{1:t+d}\|_2$.
10 **end**
11 Return $q, \hat{B}$

Proof First note that as in Lem. 17, no new epoch will start because $\|x_{1:t+i}\|_2 > \alpha q$ for any i. Let $i \in [0, d)$. Consider the estimation error of the $i+1$-th column of B:

$$\|\frac{x_{s_k+i+1}}{\lambda_i} - Be_{i+1}\|_2 = \frac{1}{\lambda_i}\|Ax_{s_k+i} + z_{s_k+i}\|_2 \le \frac{M}{\lambda_i}\|x_{s_k+i}\|_2 + \frac{1}{\lambda_i}\|z_{s_k+i}\|_2.$$

By Lem. 17, we have $\|x_{s_k+i}\|_2, \|w_{s_k+i}\|_2 \le 4^{2i}M^{2i}q_k\varepsilon^{-i}$. Therefore we have

$$\|\frac{x_{s_k+i+1}}{\lambda_i} - Be_{i+1}\|_2 \le \frac{M}{\lambda_i}\|x_{s_k+i}\|_2 + \frac{1}{\lambda_i}\|z_{s_k+i}\|_2 \le 3\varepsilon.$$

Concatenating the column estimates, we upper bound the Frobenius norm of $B - \hat{B}$,

$$\|B - \hat{B}\|_F^2 = \sum_{i=0}^{d-1} \|\frac{x_{s_k+i+1}}{\lambda_i} - Be_{i+1}\|_2^2 \le 9d\varepsilon^2.$$

We conclude that $\|B - \hat{B}\|_2 \le \|B - \hat{B}\|_F \le 3\varepsilon\sqrt{d}$. Moreover, with our choice of ε, we have $\|B - \hat{B}\|_2 \le \frac{L}{4}$, so by Ky Fan singular value inequalities, we have $\sigma_{\min}(B) \le \sigma_{\min}(\hat{B}) + \frac{L}{4}$, and hence $\sigma_{\min}(\hat{B}) \ge \frac{L}{2}$, and the condition in Line 10 will not be triggered.

Now, we can write $B = \hat{B} + 3\varepsilon\sqrt{d}C$ for some $C \in \mathbb{R}^{d\times d}$, $\|C\| \le 1$. Then we have

$$\|B\hat{B}^{-1} - I\| = 3\varepsilon\sqrt{d}\|C\hat{B}^{-1}\| \le \frac{3\sqrt{d}\varepsilon}{\sigma_{\min}(\hat{B})} \le \frac{6\varepsilon\sqrt{d}}{L}.$$

■

E.3. Estimation of the System

Lemma 19 *Suppose $\|f_{0:T-1}\|_2 \le q_k$, $\alpha \ge 4^{2d} M^{2d} \varepsilon^{-d}$, and Alg. 6 produces $\hat{A}$ such that $\|A - \hat{A}\|_2 \le \varepsilon_A$ then the resultant controller K satisfies $\|A - BK\| \le \varepsilon_A + \frac{6\varepsilon M\sqrt{d}}{L}$.*

Proof By Lem. 18, the algorithm will not start a new epoch with the choice of α, and we have $\|B\hat{B}^{-1} - I\|_2 \le \frac{6\varepsilon\sqrt{d}}{L}$, so we have

$$BK = B\hat{B}^{-1}\hat{A} = \hat{A} + \frac{6\varepsilon\sqrt{d}}{L} C\hat{A}$$

for C with $\|C\|_2 \le 1$. Thus, we have

$$\|A - BK\|_2 \le \|A - \hat{A}\|_2 + \frac{6\varepsilon\sqrt{d}}{L}\|C\|\|\hat{A}\| \le \varepsilon_A + \frac{6\varepsilon M\sqrt{d}}{L} .$$

■

Lemma 20 *Suppose $\|f_{0:T-1}\|_2 \le q_k$ and $\alpha > R = (4M)^{5N}\varepsilon^{-2N}$, then Alg. 6 produces $\hat{A}$ such that*

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

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

Proof We first note by choice of α, the SysID will not be restarted. We first upper bound $\Phi(A)$. We also have $\Phi(\hat{A}) \le \Phi(A)$ by optimality of $\hat{A}$.

Let $i \in [0, N)$. Consider the estimation error of Av_i:

$$\begin{aligned}\|\frac{x_{t'+2i+2}}{\xi_i} - AB\hat{B}^{-1}v_i\|_2 &= \frac{1}{\xi_i}\|A^2 x_{t'+2i} + Az_{t'+2i} + z_{t'+2i+1}\|_2 \\ &\le \frac{M^2}{\xi_i}\|x_{t'+2i}\|_2 + \frac{M}{\xi_i}\|z_{t'+2i}\|_2 + \frac{1}{\xi_i}\|z_{t'+2i+1}\|_2 .\end{aligned}$$

By Lemma 21, we have $\|x_{t'+2i}\|_2, \|w_{t'+2i}\|_2 \le 4^{3i}M^{3i}q_k'\varepsilon^{-i}$. Therefore for the first two terms we have,

$$\frac{M^2}{\xi_i}\|x_{t'+2i}\|_2 + \frac{M}{\xi_i}\|z_{t'+2i}\|_2 \le \frac{M^2}{\xi_i}\|x_{t'+2i}\|_2 + \frac{M}{\xi_i}(\|w_{t'+2i}\|_2 + \|f_{t'+2i}\|_2) \le 3\varepsilon.$$

For the trajectory-dependent noise at time $t' + 2i + 1$, we have

$$\begin{aligned}\frac{1}{\xi_i}\|w_{t'+2i+1}\|_2 \le \frac{h}{\xi_i}\|x_{1:t'+2i+1}\|_2 &\le \frac{h}{\xi_i}(\|x_{1:t'+2i}\|_2 + \|x_{t'+2i+1}\|_2) \\ &\le \frac{h}{\xi_i}(4^{3i}M^{3i}q_k'\varepsilon^{-i} + \|Ax_{t'+2i} + \xi_i B\hat{B}^{-1}v_i + z_{t'+2i}\|_2) \\ &\le h\varepsilon + \frac{hM}{\xi_i}\|x_{t'+2i}\|_2 + h\|B\hat{B}^{-1}\| + \frac{h}{\xi_i}\|z_{t'+2i}\|_2 \\ &\le 4\varepsilon + h\|B\hat{B}^{-1}\| \le 4\varepsilon + h(1 + \frac{1}{2}) .\end{aligned}$$

The last inequality holds, via Lem. 18. Therefore we have

$$\|\frac{x_{t'+2i+2}}{\xi_i} - AB\hat{B}^{-1}v_i\|_2 \le 8\varepsilon + \frac{3h}{2}.$$

Adding the error induced by the bias of $\hat{B}$,

$$\begin{aligned}\|\frac{x_{t'+2i+2}}{\xi_i} - Av_i\|_2 &\le \|\frac{x_{t'+2i+2}}{\xi_i} - AB\hat{B}^{-1}v_i\|_2 + \|AB\hat{B}^{-1}v_i - Av_i\|_2 \\ &\le 8\varepsilon + \frac{3h}{2} + \|A\|\|B\hat{B}^{-1} - I\| \\ &\le 8\varepsilon + \frac{3h}{2} + \frac{6M\sqrt{d}\varepsilon}{L} \le \frac{14\varepsilon M\sqrt{d}}{L} + \frac{3h}{2}\,.\end{aligned}$$

Therefore, we have $\Phi(\hat{A}) \le \Phi(A) \le \frac{14\varepsilon M\sqrt{d}}{L} + \frac{3h}{2}$ and it follows that

$$\begin{aligned}\max_{v\in V}\|(A-\hat{A})v\|_2 &= \max_{i\in[0,N)}\|(A-\hat{A})v_i\|_2 \\ &\le \max_{i\in[0,N)}\left(\|Av_i - \frac{x_{t'+2i+2}}{\xi_i}\|_2 + \|\hat{A}v_i - \frac{x_{t'+2i+2}}{\xi_i}\|_2\right) \\ &\le \Phi(A) + \Phi(\hat{A}) \le \frac{28\varepsilon M\sqrt{d}}{L} + 3h\,.\end{aligned}$$

Finally, for the state magnitude at the final iteration, by Lem. 21

$$\begin{aligned}\|x_{t'+2N}\|_2 &= \|Ax_{t'+2N-1} + w_{t'+2N-1} + f_{t'+2N-1}\|_2 \\ &\le M\|x_{t'+2N-1}\|_2 + h\|x_{1:t'+2N-1}\|_2 + q_k \\ &\le 4^{3N-1}M^{3N}q_k'\varepsilon^{-N} + h4^{3N-1}M^{3N-1}q_k'\varepsilon^{-N} + q_k \\ &\le 3\cdot 4^{3N-1}M^{3N}q_k'\varepsilon^{-N}.\end{aligned}$$

Adding previous iterations, we have $\|x_{1:t'+2N}\|_2 \le 4^{3N}M^{3N}q_k'\varepsilon^{-N} \le 4^{5N}M^{5N}\varepsilon^{-2N}q_k$. ■

Lemma 21 *Suppose $\|f_{0:T-1}\|_2 \le q_k$ and $\alpha > R = (4M)^{5N}\varepsilon^{-2N}$, then in Alg. 6, for odd iterations after t', we have $\|x_{1:t'+2i+1}\|_2 \le 4^{3i+2}M^{3i+2}q_k'\varepsilon^{-(i+1)}$, and for even iterations we have $\|x_{1:t'+2i}\|_2 \le 4^{3i}M^{3i}q_k'\varepsilon^{-i}$, for $0 \le i < N$.*

Proof We prove this by induction. Note, by the condition on α, SysID will not be restarted as long as our bounds on $\|x_{1:t'+j}\|_2$ hold. For the base case, note that for $i = 0$, the even case holds because by Lemma 17, $\|x_{1:s_k+d}\|_2 \le q_k'$. For the odd case, we have

$$\begin{aligned}\|x_{t'+1}\|_2 &\le \|Ax_{t'}\|_2 + \xi_0\|B\hat{B}^{-1}\|_2 + \|z_{t'}\|_2 \\ &\le M\|x_{t'}\|_2 + 2\xi_0 + hq_k' + q_k \\ &\le 3Mq_k' + \frac{2M^2q_k'}{\varepsilon} \le \frac{5M^2q_k'}{\varepsilon},\end{aligned}$$

where the first inequality holds by Lemma 18. Adding the previous iterations, we have $\|x_{1:t'+1}\|_2 \leq 6M^2 q'_k \varepsilon^{-1} \leq 4^2 M^2 q'_k \varepsilon^{-1}$. Now, suppose the conditions hold for both even and odd iterations for i. For $i+1$, for the even iteration,

$$\begin{aligned}
\|x_{t'+2(i+1)}\|_2 &= \|Ax_{t'+2i+1} + w_{t'+2i+1} + f_{t'+2i+1}\|_2 \\
&\leq M\|x_{t'+2i+1}\|_2 + h\|x_{1:t'+2i+1}\|_2 + q_k \\
&\leq 4^{3i+2} M^{3(i+1)} q'_k \varepsilon^{-(i+1)} + h 4^{3i+2} M^{3i+2} q'_k \varepsilon^{-(i+1)} + q_k \\
&\leq 3 \cdot 4^{3i+2} M^{3(i+1)} q'_k \varepsilon^{-(i+1)}.
\end{aligned}$$

Adding previous iterations, we have

$$\|x_{1:t'+2(i+1)}\|_2 \leq 4^{3(i+1)} M^{3(i+1)} q'_k \varepsilon^{-(i+1)}.$$

For the odd iteration,

$$\begin{aligned}
\|x_{t'+2(i+1)+1}\|_2 &= \|Ax_{t'+2(i+1)} + Bu_{t'+2(i+1)} + w_{t'+2(i+1)} + f_{t'+2(i+1)}\|_2 \\
&\leq M\|x_{t'+2(i+1)}\|_2 + 2\xi_{i+1} + h\|x_{1:t'+2(i+1)}\|_2 + q_k \\
&\leq 4^{3i+3} M^{3i+4} q'_k \varepsilon^{-(i+1)} + 2 \cdot 4^{3i+3} M^{3i+5} q'_k \varepsilon^{-(i+2)} + h 4^{3i+3} M^{3i+3} q'_k \varepsilon^{-(i+1)} + q_k \\
&\leq 5 \cdot 4^{3i+3} M^{3i+5} q'_k \varepsilon^{-(i+2)}.
\end{aligned}$$

Adding the previous iterations, we have

$$\|x_{1:t'+2(i+1)+1}\|_2 \leq 4^{3(i+1)+2} M^{3(i+1)+2} q'_k \varepsilon^{-(i+2)} .$$

■

E.4. Cost of linear control

Lemma 22 *If $\|f_{0:T-1}\|_2 \leq q_k$, and $u_t = -Kx_t$ for $t \geq t^* \geq s_k$, with $\|A - BK\|_2 \leq 1/2$ then for $h \leq \frac{1}{6}$,*

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

Proof We first prove that $\|x_{t^*:t}\|_2^2 \leq 4\|z_{t^*:t-1}\|_2^2 + 2\|x_{t^*}\|_2^2$ by induction on $t \geq t^*$. For the base case, we have $\|x_{t^*:t^*}\|_2^2 \leq 2\|x_{t^*}\|_2^2$. Now note that

$$\begin{aligned}
\|x_{t^*:t+1}\|_2^2 = \sum_{s=t^*}^{t+1} \|x_s\|_2^2 &= \|x_{t^*}\|_2^2 + \sum_{s=t^*}^{t} \|x_{s+1}\|_2^2 \\
&= \|x_{t^*}\|_2^2 + \sum_{s=t^*}^{t} \|(A - BK)x_s + z_s\|_2^2 \\
&\leq \|x_{t^*}\|_2^2 + 2\sum_{s=t^*}^{t} \|(A - BK)\|_2^2 \|x_s\|_2^2 + \|z_s\|_2^2 \\
&\leq \|x_{t^*}\|_2^2 + \frac{1}{2}\sum_{s=t^*}^{t} \|x_s\|_2^2 + 2\sum_{s=t^*}^{t} \|z_s\|_2^2 \\
&= \|x_{t^*}\|_2^2 + \frac{\|x_{t^*:t}\|_2^2}{2} + 2\|z_{t^*:t}\|_2^2.
\end{aligned}$$

Applying the inductive hypothesis, we have

$$\begin{aligned}\|x_{t^*:t+1}\|_2^2 \le \|x_{t^*}\|_2^2 + \frac{\|x_{t^*:t}\|_2^2}{2} + 2\|z_{t^*:t}\|_2^2 &\le \|x_{t^*}\|_2^2 + \frac{4\|z_{t^*:t-1}\|_2^2 + 2\|x_{t^*}\|_2^2}{2} + 2\|z_{t^*:t}\|_2^2 \\ &\le 2\|x_{t^*}\|_2^2 + \frac{4\|z_{t^*:t}\|_2^2}{2} + 2\|z_{t^*:t}\|_2^2 \\ &\le 4\|z_{t^*:t}\|_2^2 + 2\|x_{t^*}\|_2^2.\end{aligned}$$

Adding $\|x_{1:t^*-1}\|_2^2$ to both sides, we have, for $t \ge t^*$, $\|x_{1:t}\|_2^2 \le 2\|x_{1:t^*}\|_2^2 + 4\|z_{t^*:t-1}\|_2^2$. Using Assumption 1 and using the shorthand w_s for $w_s(x_{1:s})$, we have

$$\begin{aligned}\|z_{t^*:t-1}\|_2^2 = \sum_{s=t^*}^{t-1} \|z_s\|_2^2 &= \sum_{s=t^*}^{t-1} \|w_s + f_s\|_2^2 \\ &\le 2\sum_{s=t^*}^{t-1} \|w_s\|_2^2 + \|f_s\|_2^2 \\ &\le 2h^2\|x_{1:t-1}\|_2^2 + 2\|f_{0:t-1}\|_2^2.\end{aligned}$$

Using this bound, we have

$$\|x_{1:t}\|_2^2 \le 2\|x_{1:t^*}\|_2^2 + 8h^2\|x_{1:t-1}\|_2^2 + 8\|f_{0:t-1}\|_2^2 \le 2\|x_{1:t^*}\|_2^2 + 8h^2\|x_{1:t}\|_2^2 + 8\|f_{0:t-1}\|_2^2 \,.$$

Rearranging and bounding using $h = \frac{1}{6}$, we have

$$\|x_{1:t}\|_2^2 \le \frac{2\|x_{1:t^*}\|_2^2 + 8\|f_{0:t-1}\|_2^2}{1 - 8h^2} \le \frac{18\|x_{1:t^*}\|_2^2 + 72\|f_{0:t-1}\|_2^2}{7} \,.$$

The result follows using $t = e_k$ and using $\|f_{0:T-1}\| \le q_k$. ∎

E.5. Exploration on Standard Basis

We consider the case where $V = \{e_1, e_2, \ldots, e_d\}$.

Lemma 23 *Suppose $h \le \frac{1}{12\sqrt{d}}$, $V = \{e_1, e_2, \ldots, e_d\}$, 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. 5 has states bounded by*

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

Proof We first note that α is sufficiently large such that Lem. 20 holds, and we have for each i, $\|(A - \hat{A})e_i\|_2 \le \frac{28\varepsilon M\sqrt{d}}{L} + 3h$. Now we note,

$$\|A - \hat{A}\|_2 \le \|A - \hat{A}\|_F = \sqrt{\sum_{i=1}^{d} \|Ae_i - \hat{A}e_i\|_2^2} \le \sqrt{d}\Big(\frac{28\varepsilon M\sqrt{d}}{L} + 3h\Big).$$

Applying, Lem. 19 and plugging in bounds on ε and h, we have

$$\|A - BK\|_2 \le \|A - \hat{A}\|_2 + \frac{6\varepsilon M\sqrt{d}}{L} \le \frac{34\varepsilon M d}{L} + 3h\sqrt{d} \le \frac{34}{150} + \frac{1}{4} < \frac{1}{2}\,.$$

Now applying, Lem. 22 along with the state bound $\|x_{1:t'+2d}\| \le (4M)^{5d}\varepsilon^{-2d}q_k$ from Lem. 20, we have

$$\|x_{1:e_k}\|_2^2 \le \frac{18((4M)^{5d}\varepsilon^{-2d}q_k)^2 + 72q_k^2}{7} \le ((4M)^{6d}\varepsilon^{-2d}q_k)^2\,.$$

Noting that $\varepsilon > \frac{L}{4^4 M d}$, we get our result by bounding $(4M)^{6d}\varepsilon^{-2d}$. ■

E.6. Exploration on ε-net

We consider the case where V is an ε-net of the unit sphere.

From Lemma 5.3 of Vershynin (2012), there exists an ε-net for the unit sphere of size $\left(1 + \frac{2}{\varepsilon}\right)^d$. We consider $V = \mathcal{N}_{1/2,d}$ such that $N = |V| = 5^d$.

Lemma 24 *Suppose $h \le \frac{1}{15}$, $V = \mathcal{N}_{1/2,d}$, and $\varepsilon = \frac{L}{1000M\sqrt{d}}$, then if $\|f_{0:T-1}\|_2 \le q_k$ and $\alpha = (\frac{4^{16}M^8 d}{L^2})^{5^d}$, the running Alg. 5 has states bounded by*

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

Proof By Lem. 20, we have for each $v \in \mathcal{N}_{1/2,d}$, $\|(A - \hat{A})v\|_2 \le \frac{28\varepsilon M\sqrt{d}}{L} + 3h$. Now, we note that $\|A - \hat{A}\|_2 \le (1 - 1/2)^{-1}\max_{v\in\mathcal{N}_{1/2,d}}\|(A - \hat{A})v\|_2$ by a triangle inequality argument (see Lemma 5.4 of Vershynin (2012)), so we have $\|A - \hat{A}\|_2 \le \frac{56\varepsilon M\sqrt{d}}{L} + 6h$. Applying, Lem. 19 and plugging in bounds on ε and h, we have

$$\|A - BK\|_2 \le \|A - \hat{A}\|_2 + \frac{6\varepsilon M\sqrt{d}}{L} \le \frac{62\varepsilon M\sqrt{d}}{L} + 6h \le \frac{62}{1000} + \frac{2}{5} < \frac{1}{2}\,.$$

Now applying, Lem. 22 along with the state bound $\|x_{1:t'+2N}\| \le (4M)^{5N}\varepsilon^{-2N}q_k$ from Lem. 20 with $N = 5^d$, we have

$$\|x_{1:e_k}\|_2^2 \le \frac{18((4M)^{5N}\varepsilon^{-2N}q_k)^2 + 72q_k^2}{7} \le ((4M)^{6N}\varepsilon^{-2N}q_k)^2\,.$$

Noting that $\varepsilon > \frac{L}{4^5 M\sqrt{d}}$, we get our result by bounding $(4M)^{6N}\varepsilon^{-2N}$. ■

E.7. Final ℓ_2-gain bounds

Theorem 25 *Suppose* $h \le \frac{1}{12\sqrt{d}}$, $V = \{e_1, e_2, \ldots, e_d\}$, *and* $\varepsilon = \frac{L}{150Md}$ *and* $\alpha = \left(\frac{4^{14}M^8d^2}{L^2}\right)^d$, *then Alg. 5 has* ℓ_2 *gain bounded by* $\frac{10M^2\alpha^2}{L} < \left(\frac{4^{15}M^{10}d^2}{L^3}\right)^{2d}$.

Proof First, observe that $\|x_{1:T}\|_2 \le \alpha q_k$, where k is the final epoch. Indeed, if $s_k < T$, then by the design of the algorithm this condition is satisfied. Otherwise, we take $q_k = \|x_{1:T}\|_2$, and the algorithm stops before entering the system identification subroutine. Now we will show that running Alg. 5, $\|f_{0:T-1}\|_2 \ge \frac{q_k L}{10M^2\alpha}$, so

$$\frac{\|x_{1:T}\|_2}{\|f_{0:T-1}\|_2} \le \frac{10M^2\alpha^2}{L}\,.$$

We break into three cases:

1. No failure occurred.
2. $\sigma_{\min}(\hat{B}) < \frac{L}{2}$ in Alg. 7 (line 10).
3. Failure check $\|x_{1:s_k}\|_2 > \alpha q_{k-1}$ occurs in Alg. 5 (line 5), Alg. 7 (line 4), or Alg. 6 (line 7),or Alg. 7 (line 10).

We first note that $q_k = \|x_{1:s_k}\|_2$ by definition. We also note that if $k > 1$ (Cases 2 and 3), $\|f_{0:T-1}\|_2 > q_{k-1}$. Suppose $\|f_{0:T-1}\|_2 \le q_{k-1}$, then by Lem. 23 and choice of α, the epoch $k-1$ would never have ended. We now analyze each case separately.

Case 1: Failure never occurs Here we must have $\|x_{1:T}\|_2 = 0$ because q is initialized at 0. K is initialized to 0, so $\|u_{1:T-1}\|_2 = 0$ and $\|f_{0:T-1}\|_2 = 0 = q$.

Case 2: Failure occurs in Alg. 7 (line 10) second condition
We know $\sigma_{\min}(B) > L$, so we must have $\|\hat{B} - B\| > \frac{L}{2}$. By Lemma 18, if $\|f_{0:T-1}\| \le q_{k-1}$, $\|\hat{B} - B\| \le 3\sqrt{d}\varepsilon \le \frac{L}{2}$, so by contradiction we must have $\|f_{0:T-1}\| > q_{k-1}$. We now note that $q_k \le \alpha q_{k-1}$, otherwise, we would have failed the other condition of the if-statement. Combining, we have $\|f_{0:T-1}\| > \frac{q_k}{\alpha}$.

Case 3: Failure occurs in Alg. 5 (line 5), in Alg. 7 (line 4), or Alg. 6 (line 7), or the first condition of Alg. 7 (line 10)
There are three possibilities for the control in the previous iteration: $u_{s_k-1} = -Kx_{s_k-1}$, $u_{s_k-1} = 0$, or u_{s_k-1} is from Alg. 7 (line 5) or Alg. 6 (line 11) and is a fixed control such that $\|u_{s_k-1}\|_2 < \alpha q_{k-1}$. To see this, we note that exploration controls are progressively increasing so we just need to look at the last large control played by Alg. 6. Thus, $\|u_{s_k-1}\|_2 \le \|\hat{B}^{-1}\|\xi_N \le \frac{2\xi_N}{L} \le \alpha q_{k-1}$.

For the first case, we note that

$$\|K\|_2 = \|\hat{B}^{-1}\hat{A}\|_2 \le \|\hat{B}^{-1}\|_2\|\hat{A}\|_2 \le \frac{2M}{L}\,.$$

Above, we use the fact that Alg. 7 always produces a $\hat{B}$ with $\sigma_{\min}(\hat{B}) \ge \frac{L}{2}$ and $\|\hat{A}\|_2 < M$. Noting that $\|x_{1:s_k-1}\|_2 \le \alpha q_{k-1}$, because otherwise the epoch would have ended on the previous iteration , we have $\|u_{s_k-1}\|_2 \le \frac{2M\alpha q_{k-1}}{L}$ in all cases.

We now bound $\|x_{s_k}\|_2$ by applying the triangle inequality and system bounds:

$$\begin{aligned}\|x_{s_k}\|_2 &= \|Ax_{s_k-1} + Bu_{s_k-1} + w_{s_k-1} + f_{s_k-1}\|_2 \\ &\le M\|x_{s_k-1}\|_2 + M\|u_{s_k-1}\|_2 + \|w_{s_k-1}\|_2 + \|f_{s_k-1}\|_2 \\ &\le M\|x_{1:s_k-1}\|_2 + M\|u_{s_k-1}\|_2 + h\|x_{1:s_k-1}\|_2 + \|f_{s_k-1}\|_2 \\ &\le \frac{4M^2\alpha}{L} q_{k-1} + \|f_{s_k-1}\|_2\end{aligned}$$

Adding the previous iterations, we have

$$\begin{aligned}q_k = \|x_{1:s_k}\|_2 &\le \|x_{1:s_k-1}\|_2 + \frac{4M^2\alpha}{L} q_{k-1} + \|f_{s_k-1}\|_2 \\ &\le \alpha q_{k-1} + \frac{4M^2\alpha}{L} q_{k-1} + \|f_{s_k-1}\|_2 \le \frac{5M^2\alpha}{L} q_{k-1} + \|f_{s_k-1}\|_2 .\end{aligned}$$

Suppose $\|f_{s_k-1}\|_2 > \frac{5M^2\alpha}{L} q_{k-1}$, then we immediately have $\|f_{0:T-1}\|_2 \ge \frac{q_k}{2}$. Alternatively, we have $q_k \le \frac{10M^2\alpha q_{k-1}}{L}$. Now since $\|f_{0:T-1}\|_2 > q_{k-1}$, we have $\|f_{0:T-1}\|_2 > \frac{Lq_k}{10M^2\alpha}$. ■

Theorem 26 *Suppose $h \le \frac{1}{15}$, $V = \mathcal{N}_{1/2,d}$, $\varepsilon = \frac{L}{1000M\sqrt{d}}$, and $\alpha = (\frac{4^{16}M^8 d}{L^2})^{5^d}$, then Alg. 5 has ℓ_2 gain bounded by $(\frac{4^{17}M^{10}d}{L^3})^{2\cdot 5^d}$.*

Proof This follows exactly as Theorem 25 using Lem. 24 in place of Lem. 23. ■

Appendix F. Cusumano-Poolla Algorithm

While Thm. 1 and Thm. 26 achieve robustness independent of any system parameters with exponential and doubly-exponential ℓ_2-gain respectively, these algorithms are only applicable in the limiting settings where the control input matrix B is full rank. In contrast, the general task of designing adaptive controllers with finite closed loop ℓ_2-gain has been solved by Cusumano and Poolla (1988a). The Cusumano-Poolla algorithm works by iteratively testing all controllers, switching to a new controller when the gain exceeds some proposed bound. As long as their is a candidate controller that can satisfy the proposed bound, this algorithm will eventually converge. For completeness, we will analyze this algorithm, in the case that that our nominal linear dynamical system is strongly stabilizable (Cohen et al., 2018), a quantitative notion of stabilizability. Unlike our main setting, here we assume $B \in \mathbb{R}^{d\times p}$ so controls $u_t \in \mathbb{R}^p$.

Definition 27 *A LDS (A, B) is κ-strongly stabilizable if there exists a linear controller K with $\|K\|_2 \le \kappa$ such that $A+BK = HLH^{-1}$, where H, L are matrices satisfying $\|H\|_2, \|H^{-1}\|_2, \|H\|_2\|H^{-1}\|_2 \le \kappa$ and $\|L\|_2 \le 1 - \frac{1}{\kappa}$.*

We also will need to use an ε-net over candidate controllers for this algorithm.

Algorithm 8: Cusumano-Poolla algorithm

Input: System upper bound M, strong controllability parameter κ, disturbance bound F.

Set $\alpha = 27\kappa^4, \mathcal{K}' = \mathcal{N}^*_{\frac{1}{2M\kappa^2},\mathcal{K}}$
Initialize $q = F$, pick arbitrary controller K from $\mathcal{K}'$.
for $t = 1 \dots T$ **do**
  Observe x_t.
  if $\|x_{1:t}\|_2 > \alpha q$ **then**
    Update $q = \|x_{1:t}\|_2$.
    Update candidate controllers: $\mathcal{K}' = \mathcal{K}' \setminus \{K\}$.
    Pick any K from $\mathcal{K}$.
  else
    Execute $u_t = -Kx_t$.
  end
end

Definition 28 *Let $\mathcal{K} = \{K \in \mathbb{R}^{d\times p} : \|K\|_2 \le \kappa\}$ be the spectral norm ball of radius κ. We define $\mathcal{N}^*_{\varepsilon,\mathcal{K}} \subseteq \mathcal{K}$ to be an ε-net of $\mathcal{K}$, with the metric $d(X,Y) = \|X - Y\|_2$ if for any $K \in \mathcal{K}$, we have $K' \in \mathcal{N}^*_{\varepsilon,\kappa,d\times p}$ such that $\|K - K'\|_2 \le \varepsilon$.*

Theorem 29 *If $F = \|f_{0:T-1}\|_2$, and $h < \frac{1}{5\kappa^4}$, then Alg. 8 has ℓ_2-gain bounded by $\left(135\kappa^5 M\right)^{(4M\kappa^3\sqrt{dp})^{dp}}$.*

Proof We first note that there exists a controller $K \in \mathcal{K}'$ such that Lem. 31 holds, in which case the controller will never switch. To get an ℓ_2-gain bound, we need to bound the state expansion that occurs using any other $K \in \mathcal{K}'$. We note that $\|A + BK\|_2 \le 2\kappa M$, when a controller pushes $\|x_{1:t}\|_2$ above αq, we have

$$\|x_t\|_2 \le 2\kappa M\|x_{t-1}\|_2 + \|w_{t-1}\|_2 + \|f_{t-1}\| \le (2\kappa\alpha M + h + 1)q\,.$$

Adding in $\|x_{1:t-1}\|_2 < \alpha q$, we have $\|x_{1:t}\|_2 \le 5\kappa\alpha M q$. This state expansion can occur at most once per controller in $\mathcal{K}'$. By Lem. 30, $|\mathcal{K}'| < (4M\kappa^3\sqrt{dp})^{dp}$ and hence,

$$\|x_{1:T}\|_2 \le \left(5\kappa\alpha M\right)^{(2M\kappa^3\sqrt{dp})^{dp}} F\,.$$

Plugging in α, yields the result. ■

Lemma 30 *There exists an ε-net for $\mathcal{K}$ such that $|\mathcal{N}^*_{\varepsilon,\mathcal{K}}| \le \left(\frac{2\kappa\sqrt{dp}}{\varepsilon}\right)^{dp}$.*

Proof This follows by choosing a grid with granularity $\frac{\varepsilon}{\sqrt{pd}}$ in each coordinate, assuring that for any K there is a K' in the net at most $\frac{\varepsilon}{\sqrt{pd}}$ away in each coordinate. This guarantees that $\|K - K'\|_F \le \varepsilon$ and so $\|K - K'\|_2 \le \varepsilon$. ■

Lemma 31 *If (A, B) is κ-strongly stabilizable, and $h < \frac{1}{5\kappa^4}$ then there exists a linear controller $K \in \mathcal{N}^*_{\frac{1}{2M\kappa^2},\mathcal{K}}$ such that if the controller is played from iteration t^* to iteration T, the states are bounded by*

$$\|x_{1:T}\|_2 \le \sqrt{100\kappa^4\|x_{1:t^*}\|_2^2 + 625\kappa^7\|f_{0:T-1}\|_2^2} \le 27\kappa^4 \max(\|x_{1:t^*}\|_2, \|f_{0:T-1}\|_2) \,.$$

Proof From strong stablizability, for some $K^* \in \mathcal{K}$ we can write $A + BK^* = HLH^{-1}$ with $\|H\|_2\|H^{-1}\|_2 \le \kappa$ and $\|L\|_2 \le 1 - \frac{1}{\kappa}$. We define $y_t = H^{-1}x_t$. Let K be a controller in the ε-net such that $\|K - K^*\| \le \frac{1}{M\kappa^2}$.

We first prove that $\|y_{t^*:t}\|_2^2 \le 12\kappa^5\|z_{t^*:t-1}\|_2^2 + 4\kappa^2\|y_{t^*}\|_2^2$ by induction on $t \ge t^*$. For the base case, we have $\|y_{t^*:t^*}\|_2^2 \le 4\kappa^2\|y_{t^*}\|_2^2$ since $\kappa \ge 1$. Now note that

$$\begin{aligned}
\|y_{t^*:t+1}\|_2^2 = \sum_{s=t^*}^{t+1}\|y_s\|_2^2 &= \|y_{t^*}\|_2^2 + \sum_{s=t^*}^{t}\|y_{s+1}\|_2^2 \\
&= \|y_{t^*}\|_2^2 + \sum_{s=t^*}^{t}\|H^{-1}((A-BK)x_s + z_s)\|_2^2 \\
&= \|y_{t^*}\|_2^2 + \sum_{s=t^*}^{t}\|H^{-1}((A-BK^*+B(K-K^*)))x_s + z_s)\|_2^2 \\
&= \|y_{t^*}\|_2^2 + \sum_{s=t^*}^{t}\|(L + H^{-1}B(K-K^*)H)y_t + H^{-1}z_s\|_2^2 \\
&\le \|y_{t^*}\|_2^2 + (1+\frac{1}{2\kappa})\sum_{s=t^*}^{t}\|L + H^{-1}B(K-K^*)H\|_2^2\|y_s\|_2^2 + (1+2\kappa)\sum_{s=t^*}^{t}\kappa^2\|z_s\|_2^2 \\
&\le \|y_{t^*}\|_2^2 + (1-\frac{1}{4\kappa^2})\sum_{s=t^*}^{t}\|y_s\|_2^2 + 3\kappa^3\sum_{s=t^*}^{t}\|z_s\|_2^2 \\
&= \|y_{t^*}\|_2^2 + (1-\frac{1}{4\kappa^2})\|y_{t^*:t}\|_2^2 + 3\kappa^3\|z_{t^*:t}\|_2^2 \,.
\end{aligned}$$

Above, we use the fact that

$$\|L + H^{-1}B(K-K^*)H\|_2 \le \|L\|_2 + \|H\|_2\|H^{-1}\|_2\|B\|_2\|K-K^*\|_2 \le (1-\frac{1}{\kappa}) + \kappa M \cdot \frac{1}{2M\kappa^2} \le 1 - \frac{1}{2\kappa} \,.$$

Applying the inductive hypothesis, we have

$$\begin{aligned}
\|y_{t^*:t+1}\|_2^2 &\le \|y_{t^*}\|_2^2 + (1-\frac{1}{4\kappa^2})\|y_{t^*:t}\|_2^2 + 3\kappa^3\|z_{t^*:t}\|_2^2 \\
&\le \|y_{t^*}\|_2^2 + (1-\frac{1}{4\kappa^2})(12\kappa^5\|z_{t^*:t-1}\|_2^2 + 4\kappa^2\|y_{t^*}\|_2^2) + 3\kappa^3\|z_{t^*:t}\|_2^2 \\
&\le 4\kappa^2\|y_{t^*}\|_2^2 + 12\kappa^5\|z_{t^*:t}\|_2^2 \,.
\end{aligned}$$

Adding $\|y_{1:t^*-1}\|_2^2$ to both sides, we have, for $t \ge t^*$, $\|y_{1:t}\|_2^2 \le 4\kappa^2\|y_{1:t^*}\|_2^2 + 12\kappa^5\|z_{t^*:t-1}\|_2^2$. Now, by definition of y_t and $\|H\|_2\|H^{-1}\|_2 \le \kappa$, we note that $\|x_{1:t}\|_2^2 \le 4\kappa^4\|x_{1:t^*}\|_2^2 + 12\kappa^7\|z_{t^*:t-1}\|_2^2$.

Using Assumption 1 and using the shorthand w_s for $w_s(x_{1:s})$, we have

$$\begin{aligned}\|z_{t^*:t-1}\|_2^2 = \sum_{s=t^*}^{t-1} \|z_s\|_2^2 &= \sum_{s=t^*}^{t-1} \|w_s + f_s\|_2^2 \\ &\le 2\sum_{s=t^*}^{t-1} \|w_s\|_2^2 + \|f_s\|_2^2 \\ &\le 2h^2\|x_{1:t-1}\|_2^2 + 2\|f_{0:t-1}\|_2^2.\end{aligned}$$

Using this bound, we have

$$\|x_{1:t}\|_2^2 \le 4\kappa^4\|x_{1:t^*}\|_2^2 + 24\kappa^7h^2\|x_{1:t-1}\|_2^2 + 24\kappa^7\|f_{0:t-1}\|_2^2 \le 4\kappa^4\|x_{1:t^*}\|_2^2 + 24\kappa^7h^2\|x_{1:t}\|_2^2 + 24\kappa^7\|f_{0:t-1}\|_2^2 .$$

Rearranging and plugging in the bound on h, we have

$$\|x_{1:t}\|_2^2 \le \frac{4\kappa^4\|x_{1:t^*}\|_2^2 + 24\kappa^7\|f_{0:t-1}\|_2^2}{1 - 24\kappa^7h^2} \le 100\kappa^4\|x_{1:t^*}\|_2^2 + 625\kappa^7\|f_{0:t-1}\|_2^2 .$$

■

Remark 32 *For simplicity, Alg. 8 assumes we know the total disturbance magnitude $\|f_{1:T}\|_2$ exactly. The same doubling scheme from Alg. 5 can be adapted if the disturbance magnitude is not known by looping through the full ε-net of controllers in epochs until an appropriate F is found.*