Robust Online Control with Model Misspecification

Xinyi Chen [* 1 2] **Udaya Ghai** [* 1 2] **Elad Hazan** [* 1 2] **Alexandre Megretski** [* 3]

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**, which measures how much deviation from the assumed linear approximation can be tolerated while maintaining a bounded ℓ_2-gain.

Some models cannot be stabilized even with perfect knowledge of their coefficients: the robustness is limited by the minimal distance between the assumed dynamics and the set of unstabilizable dynamics. Therefore it is necessary to assume a lower bound on this distance. Under this assumption, and with full observation of the d dimensional state, we describe an efficient controller that attains $\Omega(\frac{1}{\sqrt{d}})$ robustness together with an ℓ_2-gain whose dimension dependence is near optimal. We also give an inefficient algorithm that attains constant robustness independent of the dimension, with a finite but sub-optimal ℓ_2-gain.

1. Introduction

The control of linear dynamical systems is well studied and understood. Classical algorithms such as LQR and LQG are known to be optimal for stochastic control, while robust $\mathcal{H}_\infty$ control is optimal in the worst case, assuming quadratic costs. Recent advancements gave rise to efficient online control methods based on convex relaxation that can minimize regret in the presence of adversarial perturbations. However, the problem of efficient control for general nonlinear systems is intractable.

In this paper we revisit a natural and well studied approach for nonlinear control: that of linear dynamics with model misspecification. The deviation of the nonlinear dynamics from a linear system is captured by an adversarial disturbance term that can scale with the system state history. The amount of such deviation that can be tolerated while maintaining system stability is called the **robustness** of the system.

Our study is motivated by a long standing research direction. The field of adaptive control has addressed the problem of controlling a linear dynamical system with uncertain parameters, providing guarantees of asymptotic optimality of adaptive control algorithms. However, these algorithms were shown to lack robustness under model misspecification (e.g. Rohrs et al. (1982)).

In this paper, we show that a properly designed adaptive control algorithm can exhibit a significant degree of robustness to unmodeled dynamics, even though the associated closed loop ℓ_2 gain grows rapidly[1]. We explore the limits of robust control of a linear dynamical system with adversarial perturbation whose magnitude can depend on the state history. We show that it is indeed possible to achieve constant robustness which depends only on the system dimension, and independent of its other natural parameters.

The controller that achieves this performance is computationally efficient. It 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.

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,

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

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. The perturbation

[*]Equal contribution [1]Department of Computer Science, Princeton University, Princeton, NJ [2]Google AI Princeton, Princeton, NJ [3]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA. Correspondence to: Xinyi Chen <xinyic@princeton.edu>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

[1]The ℓ_2 gain has to grow rapidly regardless of robustness, as per the lower bounds of (Chen & Hazan, 2021)

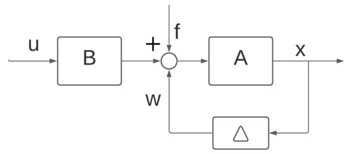

Figure 1. Diagram of the system, where Δ represents model misspecification.

$w_t = \Delta_t(x_{1:t}) : \mathbb{R}^{d\times t} \to \mathbb{R}^d$ represents the deviation of the nonlinear system from the nominal system (A, B), and it crucially satisfies the following assumption:

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

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 tolerated in the system, and our goal is to study the limits of stabilizability of the system with robustness being as large as possible. The measure of stability we use is taken from classical control theory, and is called the ℓ_2-gain of a closed-loop system with control algorithm $\mathcal{A}$ in the feedback loop,

$$\ell_2\text{-gain}(\mathcal{A}) = \max_{f_t} \frac{\|x_{1:T}\|_2}{\|f_{0:T-1}\|_2}, \tag{2}$$

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. B. With this notation, we can formally state our main question:

$\boxed{\text{How large can } h \text{ be for the system to allow a control algorithm which yields a finite } \ell_2\text{-gain?}}$

Our study initiates an answer to this question from both lower and upper bound perspectives. Specifically, for any system with a non-degenerate control matrix B,

- 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, and independently of the other system parameters.

  In addition, we show that this algorithm achieves finite ℓ_2-gain of $2^{\tilde{O}(d \log M)}$, where M is an upper bound on the spectral norm of the system. While the exponential dependence on the dimension may seem daunting, it is known to be necessary as per the lower bound of Chen & Hazan (2021), which is $\Omega(2^d)$.

- We give an (inefficient) control algorithm with a finite ℓ_2-gain and constant robustness $h = \Omega(1)$, independent of the other system parameters.

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. 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. A.

For our main results, we use an active explore-then-commit method, and we use a doubling strategy to handle unknown disturbance levels. We also study system identification using online least squares, and prove that it gives constant robustness and finite ℓ_2-gain bounds for one dimensional systems in App. C. We explain why this methodology is hard to generalize to higher dimensions, and motivate our use of the active exploration technique.

1.2. Related work

Adaptive Control. The most relevant field to our work is adaptive control, see for example a survey by Tao (2014). This field has addressed the problem of controlling a liner 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, in abstract, by Cusumano & Poolla (1988b), 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 & 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., 2020), 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 The achievability of finite ℓ_2-gains for systems with unknown level of disturbance has been studied in control theory. Cusumano & Poolla (1988a) gives a claim on the level of disturbance needed for finite ℓ_2-gain. Megretski & 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.

Competitive Analysis for Control 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. (2021) 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.

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 & Ozay, 2019; Simchowitz et al., 2018; Sarkar & Rakhlin, 2019; Faradonbeh et al., 2019). However, least squares can lead to inconsistent solutions under adversarial disturbances. 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 & Hazan (2021) give system identification algorithms for any unknown system, stable or not, with and without knowledge of a stabilizing controller, respectively.

2. Definitions and 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 denote an ε-net as $\mathcal{N}_{\varepsilon,d}$, defined as:

Definition 1. *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 $\|x - x'\|_2 \le \varepsilon$.*

Goal. Given access to a black box LDS as in Section 1.1 satisfying the assumptions below, and **without** the ability to restart the system, obtain the best possible ℓ_2-gain. First we make the assumption on 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:*[2]

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

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 the control matrix is invertible.

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 App.B). We choose to present our results in terms of ℓ_2-gain for simplicity.

3. 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 discovers that q is not sufficiently large and increases the upper bound. The key to this doubling strategy is identifying when and how much the upper bound should increase.
The algorithm uses an exploration set for system identification, and then executes the stabilizing controller of the estimated system. If the upper bound q indeed exceeds $\|f_{0:T-1}\|_2$, the algorithm is guaranteed to find a stabilizing controller. The efficient version of the algorithm uses the standard basis vectors as the exploration set, but attains robustness depending on $\sqrt{d}$. The 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 theorems below present the main guarantees of our algorithm.

Theorem 1. *For $h \le \frac{1}{12\sqrt{d}}$, $V = \{e_1, \ldots, e_d\}$, there exists ε, α such that Alg. 1 has ℓ_2-gain$(\mathcal{A}) \le (\frac{Md}{L})^{O(d)}$.*

[2]Notice that w_t can depend on the actual trajectory of states, and not only their magnitude. This is important to capture misspecification of the dynamics.

Theorem 2. *For $h \leq \frac{1}{15}$, $V = \mathcal{N}_{1/2,d}$, there exists ε, α such that Alg. 1 has ℓ_2-gain$(\mathcal{A}) \leq \left(\frac{Md}{L}\right)^{2^{O(d)}}$.*

Remark 1. *We note that when V is the standard basis, $\hat{A}$ has a closed form. In particular, the unconstrained solution[3] of Line 16 in Alg. 2 has $\Phi(\hat{A}) = 0$, where $\hat{A} = [\frac{x_{d+3}}{\xi_0} \cdots \frac{x_{d+2(d-1)+3}}{\xi_{d-1}}]$. When V is an ε-net, Φ is a maximum of convex functions, and hence a convex function.*

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, and exploration set $V \subseteq \mathbb{S}^{d-1}$.
Set $q = 0, K = 0$.
while $t \leq T$ **do**
    Observe x_t.
    if $\|x_{1:t}\|_2 > \alpha q$ **then**
        Update $q = \|x_{1:t}\|_2$.
        Call Alg 2 with parameters $(q, M, L, \varepsilon, \alpha, V)$, obtain updated K and budget q.
    else
        Execute $u_t = -Kx_t$.
        $t \leftarrow t + 1$
    end if
end while

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 if
    play $u_{t+i} = \lambda_i e_{i+1}$, $\lambda_i = \frac{4^{2i} M^{2i+1} q}{\varepsilon^{i+1}}$.
end for
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 if
Return $q, \hat{B}$

[3]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 2 Adversarial System ID on Budget

1: 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 $V \subseteq \mathbb{S}^{d-1}$.
2: Define $N = |V| \geq d$ with $V = (v_0, v_1, \ldots, v_{N-1})$.
3: Call Alg. 3 with parameters $(q, M, L, \varepsilon, \alpha)$, obtain estimator $\hat{B}$ and updated budget q. Suppose the system evolves to time $t' = t + d$.
4: Set $q' = 4^{2d} M^{2d} \varepsilon^{-d} q$.
5: **for** i = 0, 1, …, $2N - 1$ **do**
6:     observe $x_{t'+i}$.
7:     **if** $\|x_{1:t'+i}\|_2 > \alpha q$ **then**
8:         Restart SysID from Line 2 with $q = \|x_{1:t'+i}\|_2$.
9:     **end if**
10:     **if** i is even **then**
11:         play $u_{t'+i} = \xi_{i/2} \hat{B}^{-1} v_{i/2}$, $\xi_{i/2} = \frac{4^{3i/2} M^{3i/2+2} q'}{\varepsilon^{i/2+1}}$.
12:     **else**
13:         Play $u_{t'+i} = 0$.
14:     **end if**
15: **end for**
16: Observe $x_{t'+2N}$, compute

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

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

3.1. Proof sketch

The algorithm has three components: exploration to estimate B, exploration to estimate A, and controlling the system with linear controller $K = \hat{B}^{-1}\hat{A}$. We first sketch out the analysis if the upper bound on the disturbance magnitude is correct and $\|f_{0:T-1}\|_2 \leq q$. In this case, the algorithm will not start a new epoch 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.

Identifying B (see App. D.2). Alg. 3 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.

Identifying A (see App. D.3). Once we have an accurate estimate $\hat{B}$, identification of A in Alg. 2 works by applying controls $u_t = \xi \hat{B}^{-1} v_t$ every other iteration, where $\|v_t\|_2 = 1$ 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, $\|z_{t+1}\|_2 \leq h\|x_{1:t+1}\|_2 + \|f_{0:t+1}\|_2 = O(h\xi + q)$. As a result, we have $\|\frac{x_{t+2}}{\xi} - Av_t\|_2 = O(h)$. By definition of $\hat{A}$

in Line 16, 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 precondtioned with $\hat{B}^{-1}$ to acheive robustness independent of $\sigma_{\min}(B)$.

Exploration on the standard basis and on an ε-net (see Lem. 14 and Lem. 15). If we explore with the standard basis, then we assure that each row of $\hat{A}$ is accurate to $O(h)$, so $\|A - \hat{A}\|_2 \le \|A - \hat{A}\|_F \le h\sqrt{d}$. Because we use a Frobenius norm analysis, we only produce an accurate estimate of A for $h = \Omega(1/\sqrt{d})$. Exploration using an ε-net guarantees $\|(A-\hat{A})v\|_2 = O(h)$ in all directions, providing an accurate estimate $\hat{A}$ for $h = \Omega(1)$.

Stabilizing the system (see Lem. 13). Once exploration is complete, the system is stabilized by linear controller K. By controlling the accuracy of $\hat{A}$ and $\hat{B}$, we guarantee the closed loop system satisfies $\|A - BK\|_2 < \frac{1}{2}$. We can obtain an end-to-end ℓ_2-gain bound by bounding $\|x_{1:t}\|_2^2$ in terms of $\|f_{0:t-1}\|_2^2$ and using our exploration analysis.

Handling changing disturbance budget (see App. D.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.

4. Conclusions

We have shown, contrary to common wisdom in control theory, that it is possible to control a misspecified LDS with robustness that is independent of the system magnitude. In addition, our control algorithm has near-optimal dimension dependence in terms of ℓ_2-gain. The most immediate open question is whether 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

Agarwal, N., Hazan, E., Majumdar, A., and Singh, K. A regret minimization approach to iterative learning control. *arXiv preprint arXiv:2102.13478*, 2021.

Boffi, N. M., Tu, S., and Slotine, J.-J. E. Regret bounds for adaptive nonlinear control, 2020.

Chen, X. and Hazan, E. Black-box control for linear dynamical systems, 2021.

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

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

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

Ghai, U., Lee, H., Singh, K., Zhang, C., and Zhang, Y. No-regret prediction in marginally stable systems. In Abernethy, J. and Agarwal, S. (eds.), *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pp. 1714–1757. PMLR, 09–12 Jul 2020.

Hazan, E., Kakade, S., and Singh, K. The nonstochastic control problem. In *Algorithmic Learning Theory*, pp. 408–421. PMLR, 2020.

Kakade, S., Krishnamurthy, A., Lowrey, K., Ohnishi, M., and Sun, W. Information theoretic regret bounds for online nonlinear control, 2020.

Mania, H., Jordan, M. I., and Recht, B. Active learning for nonlinear system identification with guarantees, 2020.

Megretski, A. and Rantzer, A. 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.

Mhammedi, Z., Foster, D. J., Simchowitz, M., Misra, D., Sun, W., Krishnamurthy, A., Rakhlin, A., and Langford, J. Learning the linear quadratic regulator from nonlinear observations. *arXiv preprint arXiv:2010.03799*, 2020.

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

Rohrs, C. E., Valavani, L., Athans, M., and Stein, G. Robustness of adaptive control algorithms in the presence of unmodeled dynamics, 1982.

Sarkar, T. and Rakhlin, A. Near optimal finite time identification of arbitrary linear dynamical systems. In Chaudhuri, K. and Salakhutdinov, R. (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 5610–5618, Long Beach, California, USA, 09–15 Jun 2019. PMLR.

Shi, G., Lin, Y., Chung, S.-J., Yue, Y., and Wierman, A. Online optimization with memory and competitive control, 2021.

Simchowitz, M., Mania, H., Tu, S., Jordan, M. I., and Recht, B. Learning without mixing: Towards a sharp analysis of linear system identification. In Bubeck, S., Perchet, V., and Rigollet, P. (eds.), *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pp. 439–473. PMLR, 06–09 Jul 2018.

Simchowitz, M., Boczar, R., and Recht, B. Learning linear dynamical systems with semi-parametric least squares. In Beygelzimer, A. and Hsu, D. (eds.), *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, pp. 2714–2802, Phoenix, USA, 25–28 Jun 2019. PMLR.

Sinha, R., Harrison, J., Richards, S. M., and Pavone, M. Adaptive robust model predictive control with matched and unmatched uncertainty, 2021.

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

Vershynin, R. Introduction to the non-asymptotic analysis of random matrices, 2011.

Yu, C., Shi, G., Chung, S.-J., Yue, Y., and Wierman, A. Competitive control with delayed imperfect information, 2020.

A. 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 2 (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 $\|(C_k C_k^\top)^{-1}\| \le \kappa$.

Lemma 3. *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 $\varepsilon = \frac{1}{\sqrt{\kappa}}$. □

B. Relating competitive ratio to ℓ_2-gain

Here we relate the ℓ_2-gain to the competitive ratio. We begin with a formal definition.

Definition 3. *(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* (1) *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 4. *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 2. *Dependence on M^2 is required in Theorem 4. 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 $\frac{M^2+1}{2}$.*

C. Online Linear Regression

In this section, we provide an algorithm with bounded ℓ_2-gain for *any* disturbance sequence $z_{0:T-1}$ that satisfies Assumption 1 for $h < 1/2$, for the 1-d system

$$x_{t+1} = ax_t + u_t + z_t.$$

C.1. Algorithm and Analysis

Algorithm 4 Online Least Squares Control

1: Input: time horizon T, system upper bound parameter M.
2: Initialize $x_0, u_0 = 0$
3: **for** $t = 1 \dots T$ **do**
4: Observe x_t and define $\tilde{z}_{t-1}(\hat{a}) = x_t - \hat{a}x_{t-1} - u_{t-1}$.
5: Compute $\bar{a}_t = \arg\min_{\hat{a}} \sum_{s=0}^{t-1} \tilde{z}_s^2(\hat{a})$
6: Compute $\hat{a}_t = \text{clip}_{[-M,M]}(\bar{a}_t)$
7: Execute $u_t = -\hat{a}_t x_t$.
8: **end for**

Proof Sketch The key idea to the analysis is that 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 (see (3)). 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. The result follows from stitching these regimes together. While we would like to extend this analysis to high dimensions, we note that (3) does not have a natural high dimensional extension. In particular, $\|A - \hat{A}_t\|_2$ can be large in a direction where disturbances are small relative to the magnitude of the state.

Theorem 5. *Given Assumptions 1 and 2, for $h < 1/2$, Algorithm 4 has ℓ_2-gain bounded by $O(M^2(1-4h^2)^{-2.5})$.*

Proof. Suppose $|\hat{a}_t - a| \geq \frac{1}{2}$, then $|\bar{a}_t - a| \geq \frac{1}{2}$. By definition, $\bar{a}_t$ is the unconstrained minimizer of $\tilde{Z}_{t-1}(\hat{a}) = \sum_{s=0}^{t-1} \tilde{z}_s^2(\hat{a})$, where $\tilde{z}_0(\hat{a}) = x_1$. Furthermore, since $\tilde{z}_t(a) = x_{t+1} - ax_t - u_t = z_t$, we have

$$\|z_{0:t-1}\|_2^2 = \tilde{Z}_{t-1}(a) = \tilde{Z}_{t-1}(\bar{a}_t) + (\bar{a}_t - a)^2 \|x_{1:t-1}\|_2^2 \geq \frac{\|x_{1:t-1}\|_2^2}{4} \, . \tag{3}$$

Now suppose $t^* = \min(\inf\{t \leq T : \forall s \geq t, |\hat{a}_s - a| \leq 1/2\}, T+1)$, is the first time such that the dynamics are stable for the remainder of the time horizon. If $t^* = T+1$ or if $t^* \geq 2$, then $|\hat{a}_{t^*-1} - a| > \frac{1}{2}$. Using Assumption 1, we have

$$\begin{aligned}
\|z_{0:t^*-2}\|_2^2 &= \sum_{s=0}^{t^*-2} z_s^2 = \sum_{s=0}^{t^*-2} (w_s + f_s)^2 \\
&\leq \sum_{s=0}^{t^*-2} (1 + 1 - 4h^2) w_s^2 + (1 + \frac{1}{1-4h^2}) f_s^2 \\
&\leq 2(1-2h^2)h^2 \|x_{1:t^*-2}\|_2^2 + \frac{2}{1-4h^2} \|f_{0:t^*-2}\|_2^2.
\end{aligned}$$

Applying (3), we have

$$\|x_{1:t^*-2}\|_2^2 \leq \frac{8\|f_{0:t^*-2}\|_2^2}{(1-4h^2)(1-8h^2(1-2h^2))} = \frac{8\|f_{0:t^*-2}\|_2^2}{(1-4h^2)^3}.$$

Beyond t^*, we can bound the states using stability of the dynamics, but we first need to bound the cost from $\|x_{1:t^*-2}\|_2^2$ to $\|x_{1:t^*}\|_2^2$. Note, if $t^* \leq 2$, we do not need to use (3) and $\|x_{1:t^*}\|_2^2$ is appropriately bounded by unrolling dynamics via Lem. 6. Applying Lem. 6, we have

$$\|x_{1:t^*}\|_2^2 \leq 100M^4(\|x_{1:t^*-2}\|_2^2 + \|f_{0:t^*-1}\|_2^2) \leq \frac{800M^4\|f_{0:t^*-2}\|_2^2}{(1-4h^2)^3} + 100M^4\|f_{0:t^*-1}\|_2^2 \leq \frac{1000M^4\|f_{0:t^*-1}\|_2^2}{(1-4h^2)^3} \, .$$

We complete our bound using Lem. 7, yielding

$$\|x_{1:T}\|_2^2 \leq \frac{2(1-4h^2)\|x_{1:t^*}\|_2^2 + 8\|f_{t^*:T-1}\|_2^2}{(1-4h^2)^3} \leq \frac{2000M^4\|f_{0:T-1}\|_2^2}{(1-4h^2)^5} \, .$$

□

Lemma 6. *If Assumptions 1 and 2 hold, then for any sequence of f_t's and $h < 1/2$, x_t produced by Algorithm 4 satisfies*

$$\|x_{1:t+2}\|_2^2 \leq 100M^4(\|x_{1:t}\|_2^2 + \|f_{0:t+1}\|_2^2) \, .$$

Proof. We first note that $|a - \hat{a}_t| \leq 2M$ so unrolling the dynamics once we have

$$\|x_{1:t+2}\|_2^2 \leq 2 \cdot 4M^2 \|x_{1:t+1}\|_2^2 + 2\|z_{0:t+1}\|_2^2 = 8M^2\|x_{1:t+1}\|_2^2 + 2\|z_{0:t+1}\|_2^2 \, .$$

Applying Assumption 1 with $h < \frac{1}{2}$ we have

$$\|z_{0:t+1}\|_2^2 \leq 2\|f_{0:t+1}\|_2^2 + 2h^2\|x_{1:t+1}\|_2^2 \leq 2\|f_{0:t+1}\|_2^2 + \frac{\|x_{1:t+1}\|_2^2}{2} \, .$$

Combining, we have

$$\|x_{1:t+2}\|_2^2 \leq (8M^2+1)\|x_{1:t+1}\|_2^2 + 4\|f_{0:t+1}\|_2^2 \leq 9M^2\|x_{1:t+1}\|_2^2 + 4\|f_{0:t+1}\|_2^2 \, .$$

Unrolling, one more time, yields

$$\begin{aligned}
\|x_{1:t+2}\|_2^2 &\leq 9M^2\|x_{1:t+1}\|_2^2 + 10\|f_{0:t+1}\|_2^2 \\
&\leq 9M^2(9M^2\|x_{1:t}\|_2^2 + 10\|f_{0:t}\|_2^2) + 10\|f_{0:t+1}\|_2^2 \\
&\leq 100M^4(\|x_{1:t}\|_2^2 + \|f_{0:t+1}\|_2^2) \, .
\end{aligned}$$

□

Lemma 7. *Suppose for all $t \in [t^*, T]$, Alg. 4 produces $\hat{a}_t$ such that $|a - \hat{a}_t| \le 1/2$. Then for any sequence of f_t's and $h < \frac{1}{2}$:*

$$\|x_{1:t}\|_2^2 \le \frac{2(1-4h^2)\|x_{1:t^*}\|_2^2 + 8\|f_{t^*:t-1}\|_2^2}{(1-4h^2)^3} ,$$

where $\|f_{t^*:T-1}\|_2^2 = 0$ if $t^* \ge T$ for time horizon T.

Proof. For $t \ge t^*$, we will first prove $\|x_{t^*:t}\|_2^2 \le 4\|z_{t^*:t-1}\|_2^2 + 2x_{t^*}^2$ by induction. For the base case, we have $\|x_{t^*:t^*}\| \le 2x_{t^*}^2$ and $\|z_{t^*:t^*-1}\|_2^2 = 0$. Now note that

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

Applying the inductive hypothesis, we have

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

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, we have

$$\begin{aligned}
\|z_{t^*:t-1}\|_2^2 = \sum_{s=t^*}^{t-1} z_s^2 &= \sum_{s=t^*}^{t-1} (w_s + f_s)^2 \\
&\le \sum_{s=t^*}^{t-1} (1 + 1 - 4h^2)w_s^2 + (1 + \frac{1}{1-4h^2})f_s^2 \\
&\le 2(1-2h^2)h^2 \sum_{s=1}^{t-1} x_s^2 + \frac{2}{1-4h^2}\sum_{s=t^*}^{t-1} f_s^2 \\
&= 2(1-2h^2)h^2\|x_{1:t-1}\|_2^2 + \frac{2}{1-4h^2}\|f_{t^*:t-1}\|_2^2.
\end{aligned}$$

Using this bound, we have

$$\begin{aligned}
\|x_{1:t}\|_2^2 &\le 2\|x_{1:t^*}\|_2^2 + 4\|z_{t^*:t-1}\|_2^2 \\
&\le 2\|x_{1:t^*}\|_2^2 + 8(1-2h^2)h^2\|x_{1:t}\|_2^2 + \frac{8}{1-4h^2}\|f_{t^*:t-1}\|_2^2.
\end{aligned}$$

Rearranging,

$$\|x_{1:t}\|_2^2 \le \frac{2(1-4h^2)\|x_{1:t^*}\|_2^2 + 8\|f_{t^*:t-1}\|_2^2}{(1-8h^2(1-2h^2))(1-4h^2)} = \frac{2(1-4h^2)\|x_{1:t^*}\|_2^2 + 8\|f_{t^*:t-1}\|_2^2}{(1-4h^2)^3}.$$

□

D. Full Analysis

D.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. 2. 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.

D.2. Estimation of the Control Matrix

Lemma 8. *Suppose $\|f_{0:T-1}\|_2 \leq q_k$ and $\alpha \geq 4^{2d} M^{2d} \varepsilon^{-d}$, then in Alg. 3, we have $\|x_{1:s_k+i}\|_2 \leq 4^{2i} M^{2i} q_k \varepsilon^{-i}$, for $0 \leq i \leq d$.*

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

$$\begin{aligned}
\|x_{s_k+i+1}\|_2 &= \|Ax_{s_k+i} + Bu_{s_k+i} + z_{s_k+i}\|_2 \\
&\leq M\|x_{s_k+i}\|_2 + M\lambda_i + h\|x_{1:s_k+i}\|_2 + q_k \\
&\leq 4^{2i} M^{2i+1} q_k \varepsilon^{-i} + 4^{2i} M^{2i+2} q_k \varepsilon^{-(i+1)} + h4^{2i} M^{2i} q_k \varepsilon^{-i} + q_k \\
&\leq 4^{2i+1} M^{2i+2} q_k \varepsilon^{-(i+1)}
\end{aligned}$$

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 9. *Suppose $\|f_{0:T-1}\|_2 \leq q_k$ and $\alpha \geq 4^{2d} M^{2d} \varepsilon^{-d}$, then running Alg. 3 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 $\|x_{1:s_k+d}\|_2 \leq 4^{2d} M^{2d} q_k \varepsilon^{-d}$.*

Proof. First note that as in Lem. 8, 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:

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

By Lem. 8, we have $\|x_{s_k+i}\|_2, \|w_{s_k+i}\|_2 \leq 4^{2i} M^{2i} q_k \varepsilon^{-i}$. Therefore we have

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

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

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

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

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

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

□

D.3. Estimation of the System

Lemma 10. *Suppose* $\|f_{0:T-1}\|_2 \le q_k$, $\alpha \ge 4^{2d}M^{2d}\varepsilon^{-d}$, *and Alg. 2 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. 9, 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 11. *Suppose* $\|f_{0:T-1}\|_2 \le q_k$ *and* $\alpha > R = (4M)^{5N}\varepsilon^{-2N}$, *then Alg. 2 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:

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

By Lemma 12, 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. 9. 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 $\hat{B}$,

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

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

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

Finally, for the state magnitude at the final iteration, by Lem. 12

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

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

Lemma 12. *Suppose $\|f_{0:T-1}\|_2 \leq q_k$ and $\alpha > R = (4M)^{5N}\varepsilon^{-2N}$, then in Alg. 2, for odd iterations after t', we have $\|x_{1:t'+2i+1}\|_2 \leq 4^{3i+2}M^{3i+2}q'_k\varepsilon^{-(i+1)}$, and for even iterations we have $\|x_{1:t'+2i}\|_2 \leq 4^{3i}M^{3i}q'_k\varepsilon^{-i}$, for $0 \leq i < N$.*

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 8, $\|x_{1:s_k+d}\|_2 \leq q'_k$. For the odd case, we have

$$
\begin{aligned}
\|x_{t'+1}\|_2 &\leq \|Ax_{t'}\|_2 + \xi_0\|B\hat{B}^{-1}\|_2 + \|z_{t'}\|_2 \\
&\leq M\|x_{t'}\|_2 + 2\xi_0 + hq'_k + q_k \\
&\leq 3Mq'_k + \frac{2M^2q'_k}{\varepsilon} \leq \frac{5M^2q'_k}{\varepsilon},
\end{aligned}
$$

where the first inequality holds by Lemma 9. Adding the previous iterations, we have $\|x_{1:t'+1}\|_2 \leq 6M^2q'_k\varepsilon^{-1} \leq 4^2M^2q'_k\varepsilon^{-1}$. Now, suppose the conditions hold for both even and odd iterations for i. For $i+1$, for the even iteration,

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

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,

$$
\begin{aligned}
\|x_{t'+2(i+1)+1}\|_2 &= \|Ax_{t'+2(i+1)} + Bu_{t'+2(i+1)} + w_{t'+2(i+1)} + f_{t'+2(i+1)}\|_2 \\
&\le M\|x_{t'+2(i+1)}\|_2 + 2\xi_{i+1} + h\|x_{1:t'+2(i+1)}\|_2 + q_k \\
&\le 4^{3i+3}M^{3i+4}q_k'\varepsilon^{-(i+1)} + 2\cdot 4^{3i+3}M^{3i+5}q_k'\varepsilon^{-(i+2)} + h4^{3i+3}M^{3i+3}q_k'\varepsilon^{-(i+1)} + q_k \\
&\le 5\cdot 4^{3i+3}M^{3i+5}q_k'\varepsilon^{-(i+2)}.
\end{aligned}
$$

Adding the previous iterations, we have

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

□

D.4. Cost of linear control

Lemma 13. *If $\|f_{0:T-1}\|_2 \le q_k$, and $u_t = -Kx_t$ for $t \ge t^* \ge s_k$, with $\|A-BK\|_2 \le 1/2$ then for $h \le \frac{1}{6}$,*

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

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

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

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 $\|f_{0:T-1}\| \le q_k$. □

D.5. Exploration on Standard Basis

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

Lemma 14. *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. 1 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. 11 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,

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

Applying, Lem. 10 and plugging in bounds on ε and h, we have

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

Now applying, Lem. 13 along with the state bound $\|x_{1:t'+2d}\| \le (4M)^{5d}\varepsilon^{-2d}q_k$ from Lem. 11, 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 $\varepsilon > \frac{L}{4^4Md}$, we get our result by bounding $(4M)^{6d}\varepsilon^{-2d}$. □

D.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, 2011), 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 15. *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 $\alpha = \left(\frac{4^{16}M^8d}{L^2}\right)^{5^d}$, the running Alg. 1 has states bounded by*

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

Proof. By Lem. 11, 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, 2011)), so we have $\|A-\hat{A}\|_2 \le \frac{56\varepsilon M\sqrt{d}}{L} + 6h$. Applying, Lem. 10 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. 13 along with the state bound $\|x_{1:t'+2N}\| \leq (4M)^{5N}\varepsilon^{-2N}q_k$ from Lem. 11 with $N = 5^d$, we have

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

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

D.7. Final ℓ_2 gain bounds

Theorem 16. *Suppose $h \leq \frac{1}{12\sqrt{d}}$, $V = \{e_1, e_2, \ldots, e_d\}$, and $\varepsilon = \frac{L}{150Md}$ and $\alpha = \left(\frac{4^{14}M^8d^2}{L^2}\right)^d$, then Alg. 1 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 $\|x_{1:T}\|_2 \leq \alpha q_k$, 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. 1, $\|f_{0:T-1}\|_2 \geq \frac{q_k L}{10M^2\alpha}$, so

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

We break into three cases:

1. No failure occurred.
2. $\sigma_{\min}(\hat{B}) < \frac{L}{2}$ in Alg. 3 (line 10).
3. Failure check $\|x_{1:s_k}\|_2 > \alpha q_{k-1}$ occurs in Alg. 1 (line 5), Alg. 3 (line 4), or Alg. 2 (line 7),or Alg. 3 (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 $\|f_{0:T-1}\|_2 \leq q_{k-1}$, then by Lem. 14 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. 3 (line 10) second condition
We know $\sigma_{\min}(B) > L$, so we must have $\|\hat{B} - B\| > \frac{L}{2}$. By Lemma 9, if $\|f_{0:T-1}\| \leq q_{k-1}$, $\|\hat{B} - B\| \leq 3\sqrt{d}\varepsilon \leq \frac{L}{2}$, so by contradiction we must have $\|f_{0:T-1}\| > q_{k-1}$. We now note that $q_k \leq \alpha q_{k-1}$, 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. 1 (line 5), in Alg. 3 (line 4), or Alg. 2 (line 7), or the first condition of Alg. 3 (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. 3 (line 5) or Alg. 2 (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. 2. Thus, $\|u_{s_k-1}\|_2 \leq \|\hat{B}^{-1}\|\xi_N \leq \frac{2\xi_N}{L} \leq \alpha q_{k-1}$.

For the first case, we note that

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

Above, we use the fact that Alg. 3 always produces a $\hat{B}$ with $\sigma_{\min}(\hat{B}) \geq \frac{L}{2}$ and $\|\hat{A}\|_2 < M$. Noting that $\|x_{1:s_k-1}\|_2 \leq \alpha q_{k-1}$, because otherwise the epoch would have ended on the previous iteration , we have $\|u_{s_k-1}\|_2 \leq \frac{2M\alpha q_{k-1}}{L}$ in all cases.

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

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

Adding the previous iterations, we have

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

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

$\square$

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

Proof. This follows exactly as Theorem 16 using Lem. 15 in place of Lem. 14.

$\square$