Learning to Stabilize Unknown LTI Systems on a Single Trajectory under Stochastic Noise

Ziyi Zhang[1] **Yorie Nakahira**[1] **Guannan Qu**[1]

[1]Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Abstract

We study the problem of learning to stabilize unknown noisy Linear Time-Invariant (LTI) systems on a single trajectory. The state-of-the-art guarantees that the system is stabilized before the system state reaches $2^{O(k \log n)}$ in L^2-norm, where n is the state dimension, and k is the dimension of the unstable subspace. However, this bound only holds in *noiseless* LTI systems that have a control input dimension at least as large as the dimension of unstable subspace, making it impractical in many real-life scenarios. In noisy systems, unknown noise is not only amplified by unstable system modes but also imposes significant difficulty in estimating the system dynamics or bounding the estimation errors. Furthermore, the aforementioned complexity is only achievable when the system has a number of control inputs that are at least as many as the dimension of the unstable subspace. To address these issues, we develop a novel algorithm with a singular-value-decomposition(SVD)-based analytical framework and show that the system is stabilized with the same complexity guarantee with the state-of-the-art in a noisy environment. With the SVD-based framework, we can bound the error of system identification with Davis-Kahan Theorem and design a controller that does not require the invertibility of the control matrix, making it possible to apply this algorithm in under-actuated settings. To the best of our knowledge, this paper is the first to achieve learning-to-stabilize unknown LTI system without exponential blow-up in noisy and under-actuated systems. We further demonstrate the advantage of the proposed algorithm in under-actuated settings.

1 INTRODUCTION

Driven by the success of machine learning and the practical engineering need in control, there has been a lot of interests in learning-based control of unknown dynamical systems Beard et al. [1997], Li et al. [2022], Bradtke et al. [1994], Krauth et al. [2019], Dean S. Mania [2020]. However, the existing methods commonly rely on the strong assumption of having access to a known stabilizing controller. This motivates the learning-to-stabilize problem, i.e. learning to stabilize an unknown dynamical system, particularly on a single trajectory, which has long been a challenging problem both in theory and for applications such as control of automatic vehicles and unmanned aerial vehicles (UAV).

Although many classical adaptive control approaches can solve the learn-to-stabilize problem and achieve asymptotic stability guarantees [Pasik-Duncan, 1996, Petros A. Ioannou, 2001], it is well known that the learn-to-stabilize problem suffers from an issue known as *exponential blow-up* during transients. As an example, Abbasi-Yadkori and Szepesvári [2011] and Chen and Hazan [2020] presented a model-based approach for learning to stabilize an unknown LTI system $x_{t+1} = Ax_t + Bu_t$. It first excites the system in open loop to learn the dynamics matrices (A, B) and then designs the stabilizer. However, the initial excitation phase needs to run the system in open loop for at least n steps before learning (A, B) where n is the dimension of the state space, because it takes at least n samples to fully explore the n dimensional state space. As a result, the state norm blows up to the order of $2^{\tilde{O}(n)}$ as the system may be unstable in open loop. Such an exponential blow-up can be catastrophic and has been observed in multiple papers [Abbasi-Yadkori and Szepesvári, 2011, Chen and Hazan, 2020, Lale et al., 2020, Perdomo et al., 2021, Tsiamis and Pappas, 2021]. Further, it has also been shown that all general-purpose control algorithms suffer a worst-case regret of $2^{\Omega(n)}$ [Chen and Hazan, 2020].

Despite the exponential blow-up lower bound in Chen and Hazan [2020], it is a worst-case bound and does not rule

out better results for specific systems. This motivates the following question: *is it possible to exploit instance-specific properties to learn to stabilize a noisy LTI system without suffering from the worst-case exponential blow-up in n?* This problem has two challenges. First, in order to avoid the exponential blow-up, one can only collect $o(n)$ samples, based on which we can only get partial information on the dynamics. With only partial information about the system dynamics, it is difficult to stabilize it. Second, the noise in each step of the system is amplified by the open loop unstable system, causing strong statistical dependencies between states, which explode exponentially in a single trajectory.

To solve the first challenge, we take inspiration from the framework proposed in Hu et al. [2022], which gave an algorithm that stabilizes a *deterministic* LTI system with only $\tilde{O}(k)$ state samples along a trajectory, where $k < n$ is the number of unstable eigenvalues of A. Therefore, Hu et al. [2022] offered an algorithm with state norm upper bounded by $2^{\tilde{O}(k)}$, which avoids the exponential blow-up $2^{\tilde{O}(n)}$ [Chen and Hazan, 2020, Tsiamis and Pappas, 2021]. However, Hu et al. [2022] only solves the challenge in the much simplified *noiseless and deterministic* system dynamics, as its methodology has difficulty decoupling the amplified noise from the system dynamics. In addition, Hu et al. [2022] assumes that the control matrix has the same dimension as the instability index k and is invertible. In other words, the system is *fully actuated* when restricted to the unstable subspace. This assumption is also unrealistic in applications, as the dimension of control input is problem-specific and may not be equal to k. Particularly, many real-world systems are under-actuated, meaning that the control dimension can be much less than k. However, Hu et al. [2022] hints at the possibility of stabilizing a general noisy LTI system with fewer data points.

To solve the second challenge and address the limitations in Hu et al. [2022], we need to determine a new method to approximate the unstable part of the system dynamics under stochastic noise and stabilize it with under-actuated control inputs. This is nontrivial as, for example, while some previous works have designed methods to approximate system dynamics from a noisy and blowing-up trajectory[Sarkar and Rakhlin, 2018, Simchowitz et al., 2018], these methods do not study how to separate the unstable part of the dynamics from the stable part and how to stabilize the system. The goal of this paper is to overcome these technical challenges and *to learn-to-stabilize an unknown LTI system without the exponential blow-up state norm in noisy and under-actuated settings.*

Contribution. In this paper, we develop a novel model-based algorithm, LTS_0-N, to stabilize an unknown LTI system. We design a new singular-value-decomposition(SVD)-based subspace estimation technique to estimate the "unstable" part of system dynamics under noise perturbations and stabilize it. Using this new technique, we develop an analytical framework with the Davis-Kahan Theorem to estimate the error of subspace estimation, based on which we show the approach stabilizes the unknown dynamical system with state norm bounded by $2^{O(k\log k+\log(n-k)+m-\log \text{gap})}$, where m is the dimension of control input, and gap is a constant depending on the spectral properties of A. Note that this bound avoids the worst-case exponential blow-up in state dimension $\Theta(2^n)$ and outperforms the state-of-the-art for stabilizing unknown noisy systems Lale et al. [2020], Chen and Hazan [2020]. Further, despite the challenge caused by strong stochastic dependencies, the aforementioned bound achieves a similar guarantee as the norm bound in Hu et al. [2022] for noiseless systems. In addition, as an improvement to Hu et al. [2022], we do not place any requirement on dimensions of system dynamics matrices and maintain the same complexity for under-actuated system dynamics.

Related Work. Our work is mostly related to online learning and adaptation, learn-to-control with known stabilizing controllers, learning-to-stabilize on multiple trajectories, and learn-to-stabilize on a single trajectory. In addition, we will also briefly cover system identification.

Online learning and adaptation. Adaptive control enjoys a long history of study [Pasik-Duncan, 1996, Petros A. Ioannou, 2001, Chen and Astolfi, 2021]. Most classical adaptive control methods focus on asymptotic stability and do not provide finite sample analysis, and therefore do not study the exponential blow-up issue explicitly. The more recent work on non-asymptotic sample complexity of adaptive control has recongnized the exponential blow-up issue when a stabilizing controller is not known a priori [Chen and Hazan, 2020, Faradonbeh, 2017, Lee et al., 2023, Tsiamis and Pappas, 2021, Tu and Recht, 2018]. Specifically, the most typical strategy to stabilize an unknown dynamic system is to use past trajectory to estimate the system dynamics and then design the controller [Berberich et al., 2020, De Persis and Tesi, 2020, Liu et al., 2023]. Therefore, those works need to run in an open loop for at least $O(n)$ steps before stabilizing, resulting in an exponential blow-up in the order of the state space dimension. Compared with those works, we can stabilize the system with fewer samples by identifying and stabilizing only the unstable subspace, thus avoiding the exponential blow-up.

Learn to control with known stabilizing controller. There is abundant literature on stabilizing LTI systems under stochastic noise [Bouazza et al., 2021, Jiang and Wang, 2002, Kusii, 2018, Li et al., 2022]. One line of research uses model-free approaches to learn the optimal controllers [Fazel et al., 2019, Jansch-Porto et al., 2020, Li et al., 2022, Wang et al., 2022, Zhang et al., 2020]. Those algorithms typically require a known stabilization controller as an initialization point for policy search. Another line of research utilizes model-based approaches, which require known stabilizing controllers

to collect data for learning the system dynamics [Cohen et al., 2019, Mania et al., 2019, Plevrakis and Hazan, 2020, Zheng et al., 2020]. Compared with those works, we focus on learn-to-stabilize, and the controller we obtain can serve as the initialization to existing learning-to-control works that require a known stabilizing controller.

Learning-to-stabilize on multiple trajectories. In addition to the aforementioned literatures, some works do not assume open-loop stability and learn the full system dynamics by learning from multiple trajectories. Before designing a stabilizing controller, they require a data complexity of $\widetilde{\Theta}(n)$ [Dean S. Mania, 2020, Tu and Recht, 2018, Zheng and Li, 2020], which is larger than $\widetilde{O}(k)$ of our work. Recently, a model-free approach via the policy gradient method offers a novel perspective with the same complexity [Perdomo et al., 2021]. Those works do not face the same exponential blow-up issue since they allow multiple trajectories, i.e., the state can be "reset" to 0. Compared with their work, we focus on the more challenging setting of stabilizing on a single trajectory.

Learning-to-stabilize on a single trajectory. Learning to stabilize for a linear system in an infinite time horizon has long been studied in traditional control literatures [Lai, 1986, Chen and Zhang, 1989, Lai and Ying, 1991]. There have been algorithms incurring regret of $2^{O(n)}O(\sqrt{T})$ which relies on assumptions of observability and strictly stable transition matrices [Abbasi-Yadkori and Szepesvári, 2011, Ibrahimi et al., 2012]. Some studies have improved the regret to $2^{\tilde{O}(n)} + \tilde{O}(\text{poly}(n)\sqrt{T})$ [Chen and Hazan, 2020, Lale et al., 2020]. Recently, Hu et al. [2022] proposed an algorithm that requires $\tilde{O}(k)$ samples but has assumptions on the dimension of B and does not incorporate noise in the system dynamics. In this work, we propose an algorithm that has the same state norm bound as Hu et al. [2022] in a noisy and potentially under-actuated LTI system.

System identification. Existing literature in system identification focuses on determining system parameters [Oymak and Ozay, 2018, Sarkar and Rakhlin, 2018, Simchowitz et al., 2018, Xing et al., 2022]. Our approach also partially determines the system parameters for the construction of stabilizing controllers. Compared to those works, we do not just conduct the identification but also close the loop by stabilizing the system. Such results require additional characterization of the identification accuracy and its impact on closed-loop response.

2 PROBLEM FORMULATION

Notations. In this paper, we use the L^2-norm as the default norm $\|\cdot\|$. We use M^* to represent the conjugate transpose of M, e_i to denote the unit vector with 1 at the i-th entry and 0 everywhere else, and $\rho(\cdot)$ to denote the spectral radius of a matrix. We provide an indexing of notations at Appendix H.

We consider an LTI system $x_{t+1} = Ax_t + Bu_t + \eta_t$ where $x_t, \eta_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ are the state, noise, and control input at time step t, respectively. The system dynamics determined by A and B are *unknown* to the learner. We further assume $\mathbb{E}[\eta_t] = 0$, and there exists constant $C \in \mathbb{R}^+$ such that $\|\eta_t\| < C$ for all $t \in \mathbb{N}$.[1]

The goal of the learning is to stabilize the system with a learned controller, defined as follows:

Definition 2.1 (Stabilizing controller)**.** Control rule (u_t) is called a **stabilizing controller** if and only if the closed-loop system $x_{t+1} = Ax_t + Bu_t + \eta_t$ is ultimately bounded; i.e. when $\|\eta_t\| \leq C$ for all t, $\limsup_{t\to\infty} \|x_t\| < C_n$ is guaranteed in the closed-loop system for some $C_n \in \mathbb{R}^+$.

The learner is allowed to learn the system by interacting with it on a single trajectory. More specifically, the learner can observe x_t and freely determine u_t. In this paper, we make the standard assumption that (A, B) is controllable. We also assume $x_0 = 0$ for simplicity of proof. Our proof can be easily generalized to nonzero initial conditions.

Exponential blow-up. Although there are many existing works in the learn-to-stabilize problem, including classical adaptive control [Petros A. Ioannou, 2001] or more recent learning-based control papers [Abbasi-Yadkori and Szepesvári, 2011, Chen and Hazan, 2020, Ibrahimi et al., 2012, Lale et al., 2020], it is widely recognized that any generic learn-to-stabilize algorithm inevitably causes exponential blow-up in the state norm as shown by the lower bound in Chen and Hazan [2020] and Tsiamis and Pappas [2021]. This is because $\Theta(n)$ samples are mandatory to sufficiently explore the n-dimensional state space and estimate the system dynamics before designing a stabilizing controller is possible. In contrast to these existing approaches that estimate the full system, our approach breaks the lower-bound by isolating the smaller unstable subspace from the stable subspace, estimating the system dynamics in the unstable subspace under stochastic coupling, and showing that by stabilizing the "smaller" subspace, we can stabilize the entire state space. As such, our approach breaks the exponential blow-up lower-bound in the regime when the unstable subspace has smaller dimension than n.

3 PRELIMINARIES

Our approach uses the decomposition of the state space into stable and unstable subspace (introduced in Hu et al. [2022]), and we only conduct system identification and stabilization

[1]The assumption on boundedness of noise can be loosened to sub-Gaussian random variables at the cost of a slightly more complicated proof. Indeed, in the simulation in Section 6, we show our algorithm stabilizes an LTI system with additive Gaussian noise.

for the unstable subspace. In this section, we provide a review of these concepts.

3.1 DECOMPOSITION OF THE STATE SPACE

Consider the open-loop system $x_{t+1} = Ax_t$, where A is diagonalizable. Let $\lambda_1, \cdots, \lambda_n$ denote the eigenvalues of A such that [2]

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_k| > 1 > |\lambda_{k+1}| > \cdots > |\lambda_n|.$$

We define the unstable subspace E_u as the invariant subspace corresponding to the unstable eigenvalues $\lambda_1, \ldots, \lambda_k$ and the stable subspace E_s as the invariant subspace corresponding to the stable eigenvalues $\lambda_{k+1}, \ldots, \lambda_n$.

The $E_u \oplus E_u^\perp$-decomposition. Let $P_1 \in \mathbb{R}^{n\times k}$ and $P_2 \in \mathbb{R}^{n\times(n-k)}$ denote the orthonormal bases of the unstable subspace E_u and its orthogonal complement $E_u^\perp$, respectively, namely,

$$E_u = \mathrm{col}(P_1), \quad E_u^\perp = \mathrm{col}(P_2).$$

Let $P = [P_1, P_2]$, which is also orthonormal and thus $P^{-1} = P^* = [P_1^*, P_2^*]^*$. Let $\Pi_1 := P_1P_1^*$ and $\Pi_2 := P_2P_2^*$ be the orthogonal projectors onto E_u and $E_u^\perp$, respectively. With the above decomposition, we can transform the matrix A into the two subspaces. Since E_u is an invariant subspace with regard to A, there exists $M_1 \in \mathbb{R}^{k\times k}$, $\Delta \in \mathbb{R}^{k\times(n-k)}$, and $M_2 \in \mathbb{R}^{(n-k)\times(n-k)}$, such that

$$AP = P\begin{bmatrix} M_1 & \Delta \\ & M_2 \end{bmatrix} \Leftrightarrow M := \begin{bmatrix} M_1 & \Delta \\ & M_2 \end{bmatrix} = P^{-1}AP.$$

In the above decomposition, the top-left block $M_1 \in \mathbb{R}^{k\times k}$ acts on the unstable subspace, while M_2 acts on the stable subspace. Consequently, M_1 inherits all the unstable eigenvalues of A, and M_2 inherits all the stable eigenvalues.

Finally, we examine the system dynamics after the above transformation. Let $y = [y_1^*, y_2^*]^*$ represent x in the basis formed by the column vectors of P after coordinate transformation (i.e. $x = Py$). The system dynamics after the transformation can be written as

$$\begin{aligned}\begin{bmatrix} y_{1,t+1} \\ y_{2,t+1} \end{bmatrix} &= P^{-1}AP\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} + P^{-1}Bu_t + \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix}\eta_t \\ &= \begin{bmatrix} M_1 & \Delta \\ & M_2 \end{bmatrix}\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} + \begin{bmatrix} P_1^*B \\ P_2^*B \end{bmatrix}u_t + \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix}\eta_t.\end{aligned} \tag{1}$$

The $E_u \oplus E_s$-decomposition As M is not block diagonal, signified by the top-right Δ block, which represents how much a state shifts from $E_u^\perp$ to E_u in one step, $E_u^\perp$ is in general *not* an invariant subspace with respect to A in the $E_u \oplus E_u^\perp$-decomposition. For convenience of analysis, we introduce another decomposition in the form of $E_u \oplus E_s$, where both E_u and E_s are invariant with respect to A. We also represent $E_u = \mathrm{col}(Q_1)$ and $E_s = \mathrm{col}(Q_2)$ by their *orthonormal* bases, and define $Q := [Q_1 \quad Q_2]$. Since E_u and E_s are generally not orthogonal, we define $R := Q^{-1} = [R_1^*, R_2^*]^*$. The construction detail is further explained in Appendix A.1 of Hu et al. [2022].

3.2 τ-HOP CONTROL

A τ-hop controller only inputs non-zero control u_t for once every τ steps, i.e. when $t = s\tau$, $s \in \mathbb{N}$. We inherit the τ-hop mechanism introduced in Hu et al. [2022] but change the stopping time mechanism. Let $\tilde{x}_s := x_{s\tau}$ and $\tilde{u}_s := u_{s\tau}$ denote state and control action τ time steps apart. We can then write the dynamics of the τ-hop control system as:

$$\tilde{x}_{s+1} = A^\tau\tilde{x}_s + A^{\tau-1}B\tilde{u}_s + \sum_{i=0}^{\tau-1} A^i\eta_{s\tau+i}. \tag{2}$$

Let $\tilde{y}_s$ denote the state under $E_u \oplus E_u^\perp$-decomposition, i.e. $\tilde{y}_s = P^*\tilde{x}_s$. The state evolution becomes

$$\begin{aligned}&\begin{bmatrix} \tilde{y}_{1,s+1} \\ \tilde{y}_{2,s+1} \end{bmatrix} \\ =&P^{-1}A^\tau P\begin{bmatrix} \tilde{y}_{1,s} \\ \tilde{y}_{2,s} \end{bmatrix} + P^{-1}A^{\tau-1}B\tilde{u}_s + \sum_{i=0}^{\tau-1} P^{-1}A^i\eta_{s\tau+i} \\ =&M^\tau\begin{bmatrix} \tilde{y}_{1,s} \\ \tilde{y}_{2,s} \end{bmatrix} + \begin{bmatrix} P_1^*A^{\tau-1}B \\ P_2^*A^{\tau-1}B \end{bmatrix}\tilde{u}_s + \sum_{i=0}^{\tau-1}\begin{bmatrix} P_1^*A^i \\ P_2^*A^i \end{bmatrix}\eta_{s\tau+i}.\end{aligned} \tag{3}$$

We shall denote $B_\tau := P_1^*A^{\tau-1}B$ for simplicity, and

$$\begin{aligned}M^\tau &= \left(\begin{bmatrix} M_1 & \\ & M_2 \end{bmatrix} + \begin{bmatrix} 0 & \Delta \\ & 0 \end{bmatrix}\right)^\tau \\ &= \begin{bmatrix} M_1^\tau & \sum_{i=1}^{\tau-1} M_1^i\Delta M_2^{\tau-1-i} \\ & M_2^\tau \end{bmatrix} := \begin{bmatrix} M_1^\tau & \Delta_\tau \\ & M_2^\tau \end{bmatrix}.\end{aligned}$$

Now we use a state feedback controller $\tilde{u}_s = K_1\tilde{y}_{1,s}$ in the τ-hop control system to stabilize the system by acting on the unstable component $\tilde{y}_{1,s}$. The closed-loop dynamics can be written as

$$\tilde{y}_{s+1} = \begin{bmatrix} M_1^\tau + P_1^*A^{\tau-1}BK_1 & \Delta_\tau \\ P_2^*A^{\tau-1}BK_1 & M_2^\tau \end{bmatrix}\tilde{y}_s + \sum_{i=0}^{\tau-1} P^{-1}A^iP\eta_{s\tau+i}. \tag{4}$$

[2] In practice, if A does have the same eigenvalues, a slight perturbation will make A have distinct eigenvalues, to which our method will apply. Further, a light perturbation will only introduce a log factor, as our dependence on the eigenvalue-related "gap" constant is only logarithmic, as shown in Theorem 4.2.

Algorithm 1 LTS$_0$-N: learning a τ-hop stablilzing controller

1: **Stage 1: learning the unstable subspace of A.**
2: Run the system in open loop for T steps and let $D \leftarrow [x_1, \cdots, x_T]$.
3: Compute the singular value decomposition of $D = U\Sigma V^*$. Let $\hat{P}_1 \leftarrow U^{(k)}$ be the top k columns of U.
4: Calculate $\hat{\Pi}_1 \leftarrow \hat{P}_1\hat{P}_1^*$.
5: **Stage 2: approximate M_1 on the unstable subspace.**
6: Solve the least square problem $\hat{M}_1 \leftarrow \arg\min_{M_1 \in \mathbb{R}^{k\times k}} \mathcal{L}(M_1) := \sum_{t=0}^{T} \left\| \hat{P}_1^* x_{t+1} - M_1 \hat{P}_1^* x_t \right\|^2$.
7: **Stage 3: restore B_τ for τ-hop control.**
8: **for** $i = 1, \cdots, m$ **do**
9: Let the system run in open loops for ω_i steps until $\frac{\|(I-\hat{\Pi}_1)x_{t_i}\|}{\|x_{t_i}\|} < (1-\epsilon)\gamma$ and $\frac{C}{\|x_{t_i}\|} < \delta$.
10: Run for τ more steps with initial $u_{t_i} = \alpha \|x_{t_i}\| e_i$, where $t_i = T + \sum_{j=1}^{i} \omega_j + (i-1)\tau$.
11: **end for**
12: Let $\hat{B}_\tau \leftarrow [\hat{b}_1, \cdots, \hat{b}_m]$, where the i-th column $\hat{b}_i \leftarrow \frac{1}{\alpha\|x_{t_i}\|}\left(\hat{P}_1^* x_{t_i+\tau} - \hat{M}_1^\tau \hat{P}_1^* x_{t_i}\right)$.
13: **Stage 4: construct a τ-hop stabilizing controller K.**
14: Construct the τ-hop stabilizing controller $\hat{K}_1$ from $\hat{M}_1^\tau$ and $\hat{B}_\tau$.

4 MAIN RESULTS

4.1 ALGORITHM

In this section, we propose Learning to Stabilize from Zero with Noise (LTS$_0$-N). The algorithm is divided into 4 stages: (i) learn an orthonormal basis P_1 of the unstable subspace E_u (Stage 1); (ii) learn M_1, the restriction of A onto the subspace E_u (Stage 2); (iii) learn $B_\tau = P_1^* A^{\tau-1} B$ (Stage 3); and (iv) design a controller that seeks to stabilize the "unstable" E_u subspace (Stage 4). This is formally described in Algorithm 1. We provide detailed descriptions of the four stages in LTS$_0$-N.

Stage 1: Learning the unstable subspace of A. We let the system run in open-loop (with control input $u_t \equiv 0$) for T time steps. Per the stable/unstable decomposition, the ratio between the norms of the state components in the unstable and stable subspace increases exponentially, and, very quickly, the state will lie "almost" in E_u. Consequently, the subspace spanned by the T states, i.e. the column space of $D := [x_1, \cdots, x_T]$, is very close to E_u. Thus, we use the top k left singular vectors of D (the top k eigenvectors of DD^*), denoted as $U^{(k)}$, as an estimate of the basis of the unstable subspace $\hat{P}_1$. In other words, we set $\hat{P}_1 = U^{(k)}$ and use it to construct the orthogonal projector onto E_u, namely $\hat{\Pi}_1 = U^{(k)}(U^{(k)})^*$, as an estimation of the projector $\Pi_1 = P_1 P_1^*$ onto E_u.

Stage 2: Learn M_1 on the unstable subspace. Recall that M_1 is the system dynamics matrix for the subspace E_u under $E_u \oplus E_u^\perp$-decomposition. Therefore, to estimate M_1, we first compute the projection of states $x_{1:T}$ on subspace E_u, i.e. $\hat{y}_{1,t} = \hat{P}_1^* x_{1,t}$ for $t = 1, \cdots, T$. Then we use least squares to estimate M_1, i.e. find $\hat{M}_1$ that minimizes the square loss:

$$\mathcal{L}(\hat{M}_1) := \sum_{t=0}^{T} \left\| \hat{y}_{1,t+1} - \hat{M}_1 \hat{y}_{1,t} \right\|^2. \tag{5}$$

Stage 3: Learn B_τ for τ-hop control. In this stage, we estimate B_τ, which quantifies the effect of control input on states in the unstable subspace E_u (as discussed in Section 3.2). Note that (3) shows

$$\begin{aligned} y_{1,t_i+\tau} =& M^\tau y_{1,t_i} + \Delta_\tau y_{2,t_i} + B_\tau u_{t_i} \\ &+ \sum_{j=1}^{\tau-1} M^{\tau-j} \eta_{1,t_i+j} + \Delta_{\tau-j} \eta_{2,t_i+j}. \end{aligned} \tag{6}$$

We estimate the columns of B_τ one by one. Specifically, we use a scaled unit vector e_i as control input at time t_i, run the system in open loop for τ steps, and use (6) but simply ignore the Δ_τ related terms to estimate b_i, the i-th column of B_τ, as

$$\hat{b}_i = \frac{1}{\|u_{t_i}\|}\left(\hat{P}_1^* x_{t_i+\tau} - \hat{M}_1^\tau \hat{P}_1^* x_{t_i}\right), \tag{7}$$

where u_{t_i} is parallel to e_i with magnitude $\alpha \|x_{t_i}\|$ for normalization. Here, α is an adjustable constant to guarantee that the E_s-component does not increase too much to blur our estimation after injecting u_{t_i}. Since we ignored the Δ_τ related terms in the estimation of b_i, to ensure that those terms do not cause much error in our estimation of B_τ, we let the system run in open loop for ω_i time steps before the estimation of b_i starts. Here, ω_i is a stopping time (cf. Line 9 in Algorithm 1). The purpose of the stepping time is to reduce the estimation error caused by the Δ_τ. For more details, see Proposition E.5 in the proof.

Stage 4: Construct a τ-hop stabilizing controller K. With the estimated M_1^τ and B_τ from the last stage, denoted as $\hat{M}_1^\tau$ and $\hat{B}_\tau$, the learner can choose any stabilization algorithm to find $\hat{K}_1$ by stabilizing the linear system

$$\hat{\tilde{y}}_{i+1} = \hat{M}_1^\tau \hat{\tilde{y}}_i + \hat{B}_\tau \tilde{u}_i, \qquad \tilde{u}_i = \hat{K}_1 \hat{\tilde{y}}_i,$$

where the tilde in $\hat{\tilde{y}}$ emphasizes the use of τ-hop control and the hat emphasizes the use of estimated projector $\hat{P}_1$, which introduces an extra estimation error to the final closed-loop dynamics. As $\hat{K}_1$ is chosen by the learner, we denote $\mathcal{K}$ to be a constant such that $\left\|\hat{K}_1\right\| < \mathcal{K}$. Furthermore, by Proposition F.1, there exists a positive definite matrix $\bar{U}$ such that $\left\|\hat{M}_1^\tau - \hat{B}_\tau \hat{K}_1\right\|_{\bar{U}} := \mathcal{U} < 1$, where $\|\cdot\|_{\bar{U}}$ denotes the

weighted norm induced by $\bar{U}$. These user-defined constants are used in the proof of Theorem 4.2.

To sum up, Algorithm 1 terminates in $T+\sum_{i=1}^{m}(1+\omega_i+\tau)$ time steps, where ω_i is the stopping time for the system to satisfy $\frac{\|(I-\hat{\Pi}_1)x_{t_i}\|}{\|x_{t_i}\|} < (1-\epsilon)\gamma$ and $\frac{C}{\|x_{t_i}\|} < \delta$.

Remark 4.1. Our algorithm is different from the algorithm proposed in Hu et al. [2022] in three aspects. Firstly, to account for the noise, we do not directly use the span of consecutive k vectors as the estimator for the unstable subspace. Instead, to identify the unstable subspace under noise, we utilize the singular value decomposition to identify the dominating state space in the trajectory and use that space as an estimation of P_1. Such an estimator requires a much more delicate analysis framework to bound the error based on Davis-Kahan Theorem, which we elaborate in Appendix A. Secondly, the above algorithm generalizes the problem to an under-actuated setting, where the control matrix $B \in \mathbb{R}^{n\times m}$ with $m \neq k$. To achieve this, unlike Hu et al. [2022] we no longer try to cancel out the unstable matrix M_1, but rather allow the learner to choose the stabilization controller. We show in Section 6 that our algorithm outperforms Hu et al. [2022] in an under-actuated setting in simulation. Thirdly, we use a stopping time to monitor the state norm in estimating B_τ, so that our algorithm always terminates at the earliest possible time.

4.2 STABILITY GUARANTEE

In this section, we formally state the assumptions and show our approach finds a stabilizing controller without suffering from exponential blow-up in n. Our first assumption is regarding the spectral properties of A, which requires distinct eigenvalues with specified eigengap.

Assumption 1 (Spectral Property)**.** A is diagonalizable with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_k| > 1 > |\lambda_{k+1}| > \cdots > |\lambda_n|$.

We assume the learner knows the value of k. However, we point out that our algorithm works as long as the learner picks a value $\hat{k}$ at least as large as k. In order to provide guarantee to the estimation of the open-loop unstable system dynamics, we also need an assumption on the distribution of noise η.

Assumption 2 (pdf of η)**.** Let $M_1 := \bar{P}^{-1}J\bar{P}$ denote the Jordan normal form of M_1, and $\bar{P} := [\bar{P}_1, \bar{P}_2, \cdots, \bar{P}_k]^*$. There exists $C_z \in \mathbb{R}$, such that the supremum of the probability distribution function (pdf) of $\left|\bar{P}_i^* \sum_{j=1}^{t} M_1^{-j} P_1^* \eta_j\right|$ is upper bounded almost everywhere,.i.e. $\operatorname{ess\,sup}\operatorname{pdf}\left(\left|\bar{P}_i^* \sum_{j=1}^{t} M_1^{-j} P_1^* \eta_j\right|\right) < C_z$, for all $i \in \{1, \ldots, k\}$ and $t \in \mathbb{N}$.

Assumption 2 holds for most common noise distributions, including bounded uniform distribution and Gaussian distributions(Lemma C.3). We further discuss this assumption in Appendix B and C.

With the above assumptions, our main result is as follows.

Theorem 4.2. *Given a noisy LTI system $x_{t+1} = Ax_t + Bu_t + \eta_t$ subject to Assumption 1, Assumption 2, and additionally, $|\lambda_1||\lambda_{k+1}| < 1$. Further, denote gap $:= \left|\prod_{\substack{m_1\neq m_2,\\ m_1,m_2\in\{1,\ldots,k\}}} (\lambda_{m_1}^{-1} - \lambda_{m_2}^{-1})\right|$. By running Algorithm 1 with parameters $\gamma = O(1)$, $\delta = O(m^{-\frac{1}{2}})$, $\tau = O(1)$, $\alpha = O(1)$, and $T = O\left(k\log k + \log(n-k) + \log m - \log gap\right)$, the controller returned by Algorithm 1 is a stabilizing controller. Further, Algorithm 1 guarantees that*

$$\|x_t\| < \exp\left(O\left(k\log k + \log(n-k) + m - \log gap\right)\right),$$

before termination. Here the big-O notation only shows dependence on k, m and n, while omitting dependence on $C, C_z, |\lambda_1|, |\lambda_k|, |\lambda_{k+1}|, \theta, \mathcal{K}$, and $\mathcal{U}$.

The precise bound given for each constant can be found at (65),(66),(67), and (68) in the Appendix, and the bound for T is given in Theorem 5.1. Despite the more challenging setting with noises and potentially underactuated systems, Theorem 4.2 achieves a similar guarantee as Hu et al. [2022]. Specifically, in the regime of $m = O(k)$,[3] the above Theorem shows that LTS_0-N finds a stabilizing controller with an upper bound on state norm at $2^{\tilde{O}(k)}$, which is better than the state-of-the-art $2^{\Theta(n)}$ complexity in the noisy settings. Therefore, our approach leverages instance specific properties (the dimension of unstable subspace k) to *break the exponential lower bound [Chen and Hazan, 2020] and learns to stabilize without the exponential blow-up in n in noisy and under-actuated settings*.

We also point out that constant gap is also k-dependent. In the worst case, the gap has an order of $2^{O(k^2)}$. This is still independent of n. We note that Hu et al. [2022] did not show explicit dependence on this constant. We leave it as future work whether this additional constant is essential or is an artifact of the proof. Moreover, our assumption that $|\lambda_1||\lambda_{k+1}| < 1$ is weaker than the assumption in Hu et al. [2022], which requires $|\lambda_1|^2|\lambda_{k+1}| < |\lambda_k|$. If the system can not satisfy this assumption, than the learner can pick $\hat{k} > k$, such that $|\lambda_1||\lambda_{\hat{k}+1}| < 1$ for Algorithm 1. In the worst case, LTS_0-N has the same complexity bound as the algorithm proposed in Chen and Hazan [2020]. Overall, we can achieve a better bound than the previous works for the LTI system with additive noise using singular value

[3] We note that the regime of $m = O(k)$ is the most interesting regime as it covers the under-actuated setting, which is known to be more challenging.

decomposition to isolate the unstable component. The key steps will be shown in Theorem 5.1 in Section 5.

We demonstrate the effectiveness of our algorithm in simulation in Section 6, showing our algorithm's state norm does not blow-up with n and also outperforms other benchmarks.

5 PROOF OUTLINE

In this section, we will give a high-level overview of the key proof ideas for the main theorem. The full proof details can be found in Appendix F.

Proof Structure. The proof is largely divided into four steps. In Step 1, we examine how accurately the learner estimates the unstable subspace E_u in Stage 1. We will show that Π_1, P_1 can be estimated up to an error of ϵ, δ respectively within $T = O\left(k \log k + \log(n-k) - \log \epsilon - \log \text{gap}\right)$ steps, where $\delta := \sqrt{2k}\epsilon$. In Step 2, we examine how accurately the learner estimates M_1. We show that M_1 can be estimated up to an error of $3\|A\|\delta$. In Step 3, we examine the estimation error of B_τ in Stage 3. Lastly, in Step 4, we eventually show that the τ-hop controller output by Algorithm 1 makes the system stable.

Overview of Step 1. To upper bound the estimation errors in Stage 1, we use SVD to isolate the unstable subspace and use the Davis-Kahan Theorem to decouple the system dynamics from the noise perturbation. The bounds on $\left\|\Pi_1 - \hat{\Pi}_1\right\|$ is shown in Theorem 5.1.

Theorem 5.1. *For a linear dynamic system with noise $x_{t+1} = Ax_t + \eta_t$ satisfying Assumption 1 and Assumption 2, let E_u be the unstable subspace of A, $k = \dim E_u$ be the instability index of the system and Π_1 be the orthogonal projector onto subspace E_u. Then for any $\epsilon > 0$, by running Stage 1 of Algorithm 1 for T time steps, where*

$$T = O\left(k \log k + \log(n-k) - \log \epsilon - \log gap\right),$$

we get an estimation $\hat{\Pi}_1 = U^{(k)}(U^{(k)})^*$ with error $\left\|\hat{\Pi}_1 - \Pi_1\right\| < \epsilon$. Here, the big-O notation only shows dependence on k, n and ϵ, while omitting dependence on $C, C_z, |\lambda_1|, |\lambda_k|, |\lambda_{k+1}|$, and θ.

The proof of Theorem 5.1 is deferred to Appendix A. Overall, Theorem 5.1 gives a non-asymptotic bound on the speed in which the last $n-k$ singular values of D decay. Therefore, even if the learner has no information on the exact value of k, the learner will find a large gap between the k-th and $(k+1)$-th singular value of D as the first k singular values grow exponentially and the last $(n-k)$ singular value decays exponentially, from which the learner can infer the value of k.

Overview of Step 2. To upper bound the error in Stage 2, We upper bound the error in $\arg\min_{M_1} \sum_{t=0}^{T} \left\|(U^{(k)})^* x_{t+1} - M_1 (U^{(k)})^* x_t\right\|^2$ and obtain the following proposition.

Proposition 5.2. *Under the premise of Theorem 4.2, we have*

$$\left\|\hat{M}_1^\tau - M_1^\tau\right\| \leq 3\tau \|A\| \zeta_{\epsilon_1}(A)^2 (|\lambda_1| + \epsilon_1)^{\tau-1} \delta,$$

where $\zeta_{\epsilon_1}(A)$ is constant for Gelfand's formula defined in Lemma G.2, and we recall δ is the estimation error for P_1.

The proof in this step and the related lemmas and propositions are deferred to Appendix D.

Overview of Step 3. To bound the error in Stage 3, we upper bound the error in each column of B_τ. In particular, we show that (7) generates an estimation of B_τ with an error in the same order as δ. The detail is left to Proposition E.5 in Appendix E.

Overview of Step 4. To analyze the stability of the closed-loop system, we shall first write out the closed-loop dynamics under the τ-hop controller. Recall in Section 3.2, we have defined $\tilde{u}_s, \tilde{x}_s, \tilde{y}_s$ to be the control input, state in x-coordinates, and state in y-coordinates in the τ-hop control system, respectively. Using those notations, the learned controller is obtained from the estimation of M_1^τ and B_τ by the learner with any stabilization algorithm (e.g. LQR, pole-placement).

Therefore, the closed-loop, the closed-loop τ-hop dynamics should be

$$\begin{aligned} \tilde{y}_{s+1} &= \hat{L} \begin{bmatrix} \tilde{y}_{1,s} \\ \tilde{y}_{2,s} \end{bmatrix} + \sum_{i=0}^{\tau-1} P^{-1} A^i P \eta_{s\tau+i} \\ &:= \hat{L}\tilde{y}_s + \sum_{i=0}^{\tau-1} \begin{bmatrix} P_1^* A^i \\ P_2^* A^i \end{bmatrix} \eta_{s\tau+i}, \end{aligned} \tag{8}$$

where

$$\begin{aligned} \hat{L} &:= \begin{bmatrix} M_1^\tau + P_1^* A^{\tau-1} B \hat{K}_1 \hat{P}_1^* P_1 & \Delta_\tau + P_1^* A^{\tau-1} B \hat{K}_1 \hat{P}_1^* P_2 \\ P_2^* A^{\tau-1} B \hat{K}_1 \hat{P}_1^* P_1 & M_2^\tau + P_2^* A^{\tau-1} B \hat{K}_1 \hat{P}_1^* P_2 \end{bmatrix} \\ &:= \begin{bmatrix} \hat{L}_{1,1} & \hat{L}_{1,2} \\ \hat{L}_{2,1} & \hat{L}_{2,2} \end{bmatrix}. \end{aligned} \tag{9}$$

We will show the above system to be ultimately bounded (i.e. $\rho(\hat{L}_\tau) < 1$). Note that $\hat{L}_\tau$ is given by a 2-by-2 block form, and we can utilize the following lemma for the spectral analysis of block matrices.

Lemma 5.3. *For block matrices $A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$, $E = \begin{bmatrix} 0 & E_{12} \\ E_{21} & 0 \end{bmatrix}$, the spectral radii of A and $A + E$ differ by at*

most $|\rho(A+E)-\rho(A)| \leq \chi(A+E)\,\|E_{12}\|\,\|E_{21}\|$, *where* $\chi(A+E)$ *is a constant.*

The proof of the lemma can be found in existing literature such as Nakatsukasa [2017]. Therefore, we need to ensure the stability of the diagonal blocks of $\hat{L}$ and upper-bound the norms of the off-diagonal blocks via estimation of factors appearing in these blocks. Complete proofs can be found in Appendices F.

6 NUMERICAL SIMULATION

Lastly, we include numerical simulations to demonstrate the performance of our algorithm. We consider an LTI system with additive noise

$$x_{t+1} = Ax_t + Bu_t + \eta_t, \quad \text{where } \eta_t \sim \mathcal{N}(0, \sigma^2 I),$$

where σ^2 is the variance of the additive Gaussian noise at each step. Note we use unbounded Gaussian noise here, and noise with bounded uniform distribution would generate similar results. The dynamics matrix B is generated randomly. Matrix A is generated by $A = V\Lambda V^{-1}$, where V is a randomly generated matrix, and Λ is a diagonal matrix of eigenvalues generated uniformly at random from the interval that satisfies $|\lambda_1||\lambda_{k+1}| < 1$.

In our first experiment, we compare the performance of LTS_0-N in different settings (with different n, σ). In each setting, we conduct 200 trials and record the minimal time steps it takes to stabilize the system, and the results are in Figure 1. In our second experiment, we compare our proposed algorithm to three different algorithms: a classical self-tunning regulator in Pasik-Duncan [1996], black-box control proposed in Chen and Hazan [2020], and the LTS_0 algorithm proposed in Hu et al. [2022] and the results are in Figure 2.

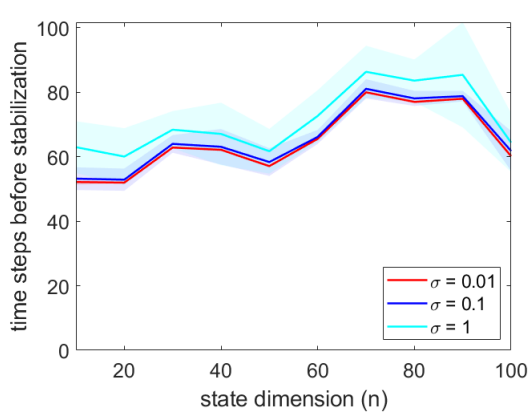

Figure 1: **Running steps of LTS_0-N.** The line shows the average steps it takes to stabilize the system, and the shadow area shows the standard deviation.

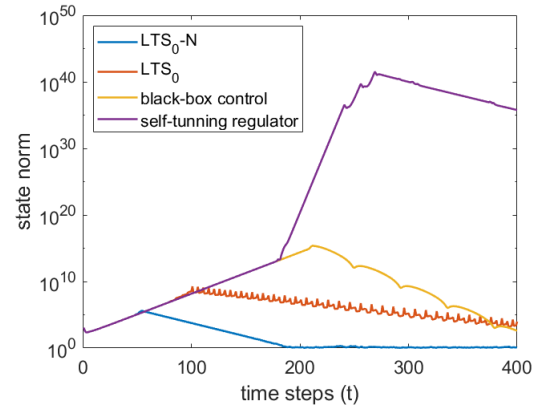

Figure 2: **State norm along one trajectory.** The trajectory of our algorithm, the algorithm in Hu et al. [2022], the black-box controller in Chen and Hazan [2020] and a self-tunning regulator in Pasik-Duncan [1996] are compared in a randomly generated LTI system with $n = 128$, $k = 4$, $m = 3$, and $\sigma = 0.01$.

Performance difference under different n and σ. Figure 1 shows the relationship between the number of steps between running LTS_0-N and the dimension of states. It is evident that the increase in the number of steps is at most linear in $\log(n)$, as proven in Theorem 4.2. As we used the same randomly generated matrices for each $(n, \sigma)-$pair, all three curves in Figure 1 have a similar trend at each node. This observation verifies that the number of steps needed for stabilization also depends on the eigenvalue distribution of the system dynamics matrices, as we showed in the proof. Moreover, we see that an increase in noise slightly increases the number of steps for stabilization, as shown in the proof of Theorem 5.1. As expected, an increase in noise also increases the standard deviation of the number of steps before stabilization.

Difference in performance in single trajectory Figure 2 shows a typical trajectory of our LTS_0-N algorithm. It is evident that our algorithm takes significantly fewer steps than adaptive control algorithms (self-tuning regulator and black-box control) and also fewer steps than the LTS_0 algorithm proposed Hu et al. [2022]. This is because the self-tuning regulator and the black-box control algorithm cannot take stabilzing control actions before the system runs for at least n steps and learns the system dynamics. Moreover, due to the stochastic coupling of the system, estimation of system dynamics becomes much more difficult, and the adaptive control methods need a relatively large state to overcome the disturbance of noise in system identification. In comparison to LTS_0, note that in this simulation, we chose $m < k$ to demonstrate the advantage of our algorithm in an under-actuated system. We see that our algorithm incurred less zig-zagging than LTS_0, since we can stabilize directly on the existing state space, and LTS_0 has to stabilize on a compos-

ite state space, the details of which can be seen at Appendix C of Hu et al. [2022].

7 ACKNOWLEDGEMENT

This work is supported by NSF 2154171, NSF CAREER 2339112, CMU CyLab seed funding, Department of the Navy, Office of Naval Research, no. N00014-23-1-2252 Japan Science and Technology Agency, no. JPMJPR2136. The views expressed are those of the authors and do not reflect the official policy or position of the US Navy, Department of Defense or the US Government.

8 CONCLUSION

This paper designs a novel algorithm for stabilizing the unstable subspace in LTI problems with additive noise. We showed that our algorithm finds a stabilizing controller with a guarantee of state norm below $2^{\tilde{O}(k)}$, which significantly improves upon the previous literature with a state norm of $2^{\Theta(n)}$. Future directions include dropping the assumption of diagonalizable A matrix and other assumptions on the spectral properties of A.

References

Yasin Abbasi-Yadkori and Csaba Szepesvári. Regret bounds for the adaptive control of linear quadratic systems. In Sham M. Kakade and Ulrike von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 1–26, Budapest, Hungary, 09–11 Jun 2011. PMLR.

Randal W. Beard, George N. Saridis, and John T. Wen. Galerkin approximations of the generalized hamilton-jacobi-bellman equation. *Automatica*, 33(12):2159–2177, 1997. ISSN 0005-1098. doi: https://doi.org/10.1016/S0005-1098(97)00128-3.

Julian Berberich, Johannes Köhler, Matthias Muller, and Frank Allgöwer. Data-driven model predictive control with stability and robustness guarantees. *IEEE Transactions on Automatic Control*, PP:1–1, 06 2020. doi: 10.1109/TAC.2020.3000182.

Leila Bouazza, Benjamin Mourllion, Abdenacer Makhlouf, and Abderazik Birouche. Controllability and observability of formal perturbed linear time invariant systems. *International Journal of Dynamics and Control*, 9, 12 2021. doi: 10.1007/s40435-021-00786-4.

Steven Bradtke, B.Erik Ydstie, and Andrew Barto. Adaptive linear quadratic control using policy iteration. *Proceedings of the American Control Conference*, 3, 09 1994. doi: 10.1109/ACC.1994.735224.

Yueqi Cao. Davis-kahan theorem, 2021. `https://www.citedrive.com/overleaf`.

Ioannis Chatzigeorgiou. Bounds on the lambert function and their application to the outage analysis of user cooperation. *Communications Letters, IEEE*, 17:1505–1508, 08 2013. doi: 10.1109/LCOMM.2013.070113.130972.

Han-Fu Chen and Ji-Feng Zhang. Convergence rates in stochastic adaptive tracking. *International Journal of Control*, 49(6):1915–1935, 1989. doi: 10.1080/00207178908559752.

Kaiwen Chen and Alessandro Astolfi. Adaptive control for systems with time-varying parameters. *IEEE Transactions on Automatic Control*, 66(5):1986–2001, 2021. doi: 10.1109/TAC.2020.3046141.

Xinyi Chen and Elad Hazan. Black-box control for linear dynamical systems. *CoRR*, abs/2007.06650, 2020.

Alon Cohen, Tomer Koren, and Yishay Mansour. Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 1300–1309. PMLR, 09–15 Jun 2019.

Claudio De Persis and Pietro Tesi. Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, 65(3):909–924, 2020. doi: 10.1109/TAC.2019.2959924.

N. et al. Dean S. Mania, H. Matni. On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*, 20:633–679, 2020. doi: https://doi.org/10.1007/s10208-019-09426-y.

Mohamad Kazem Shirani Faradonbeh. *Non-Asymptotic Adaptive Control of Linear-Quadratic Systems*. PhD thesis, University of Michigan, 2017.

Maryam Fazel, Rong Ge, Sham M. Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator, 2019.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, USA, 2nd edition, 2012. ISBN 0521548233.

Yang Hu, Adam Wierman, and Guannan Qu. On the sample complexity of stabilizing lti systems on a single trajectory. In *Neurips*, pages 1–2, 2022. doi: 10.1109/Allerton49937.2022.9929403.

Morteza Ibrahimi, Adel Javanmard, and Benjamin Van Roy. Efficient reinforcement learning for high dimensional linear quadratic systems. In *Neural Information Processing Systems*, 2012.

Joao Paulo Jansch-Porto, Bin Hu, and Geir E. Dullerud. Policy learning of mdps with mixed continuous/discrete variables: A case study on model-free control of markovian jump systems. *CoRR*, abs/2006.03116, 2020. URL `https://arxiv.org/abs/2006.03116`.

Zhong-Ping Jiang and Yuan Wang. A converse lyapunov theorem for discrete-time systems with disturbances. *Systems & Control Letters*, 45(1):49–58, 2002. ISSN 0167-6911. doi: https://doi.org/10.1016/S0167-6911(01)00164-5.

Karl Krauth, Stephen Tu, and Benjamin Recht. Finite-time analysis of approximate policy iteration for the linear quadratic regulator. *Advances in Neural Information Processing Systems*, 32, 2019.

S.M. Kusii. Stabilization and attenuation of bounded perturbations in discrete control systems. *Journal of Mathematical Sciences*, 2018. doi: https://doi.org/10.1007/s10958-018-3848-3.

T.L Lai. Asymptotically efficient adaptive control in stochastic regression models. *Advances in Applied Mathematics*, 7(1):23–45, 1986. ISSN 0196-8858. doi: https://doi.org/10.1016/0196-8858(86)90004-7.

Tze Leung Lai and Zhiliang Ying. Parallel recursive algorithms in asymptotically efficient 496 adaptive control of linear stochastic systems. *SIAM Journal on Control and Optimization*, 29(5):1091–1127, 1991.

Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, and Anima Anandkumar. Explore more and improve regret in linear quadratic regulators. *CoRR*, abs/2007.12291, 2020. URL `https://arxiv.org/abs/2007.12291`.

Bruce D. Lee, Anders Rantzer, and Nikolai Matni. Nonasymptotic regret analysis of adaptive linear quadratic control with model misspecification, 2023.

Yingying Li, Yujie Tang, Runyu Zhang, and Na Li. Distributed reinforcement learning for decentralized linear quadratic control: A derivative-free policy optimization approach. *IEEE Transactions on Automatic Control*, 67 (12):6429–6444, 2022. doi: 10.1109/TAC.2021.3128592.

Wenjie Liu, Yifei Li, Jian Sun, Gang Wang, and Jie Chen. Robust control of unknown switched linear systems from noisy data. *ArXiv*, abs/2311.11300, 2023.

Horia Mania, Stephen Tu, and Benjamin Recht. Certainty equivalence is efficient for linear quadratic control. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 2019.

Yuji Nakatsukasa. Off-diagonal perturbation, first-order approximation and quadratic residual bounds for matrix eigenvalue problems. In Tetsuya Sakurai, Shao-Liang Zhang, Toshiyuki Imamura, Yusaku Yamamoto, Yoshinobu Kuramashi, and Takeo Hoshi, editors, *Eigenvalue Problems: Algorithms, Software and Applications in Petascale Computing*, pages 233–249, Cham, 2017. Springer International Publishing. ISBN 978-3-319-62426-6.

Samet Oymak and Necmiye Ozay. Non-asymptotic identification of lti systems from a single trajectory. *2019 American Control Conference (ACC)*, pages 5655–5661, 2018.

B. Pasik-Duncan. Adaptive control [second edition, by karl j. astrom and bjorn wittenmark, addison wesley (1995)]. *IEEE Control Systems Magazine*, 16(2):87–, 1996.

Juan C. Perdomo, Jack Umenberger, and Max Simchowitz. Stabilizing dynamical systems via policy gradient methods. *ArXiv*, abs/2110.06418, 2021.

Jing Sun Petros A. Ioannou. Robust adaptive control. *Autom.*, 37(5):793–795, 2001.

Orestis Plevrakis and Elad Hazan. Geometric exploration for online control. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS'20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Tuhin Sarkar and Alexander Rakhlin. Near optimal finite time identification of arbitrary linear dynamical systems. In *International Conference on Machine Learning*, 2018.

Max Simchowitz, Horia Mania, Stephen Tu, Michael I. Jordan, and Benjamin Recht. Learning without mixing: Towards a sharp analysis of linear system identification. *Annual Conference Computational Learning Theory*, 2018.

Anastasios Tsiamis and George J. Pappas. Linear systems can be hard to learn. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 2903–2910, 2021. doi: 10.1109/CDC45484.2021.9682778.

Stephen Tu and Benjamin Recht. The gap between model-based and model-free methods on the linear quadratic regulator: An asymptotic viewpoint. In *Annual Conference Computational Learning Theory*, 2018.

Gabriel H. Tucci and Philip A. Whiting. Asymptotic results on generalized vandermonde matrices and their extreme eigenvalues. In *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1816–1823, 2011. doi: 10.1109/Allerton.2011.6120389.

Wei Wang, Xiangpeng Xie, and Changyang Feng. Model-free finite-horizon optimal tracking control of discrete-time linear systems. *Appl. Math. Comput.*, 433(C), nov 2022. ISSN 0096-3003. doi: 10.1016/j.amc.2022.127400.

Yu Xing, Benjamin Gravell, Xingkang He, Karl Henrik Johansson, and Tyler H. Summers. Identification of linear systems with multiplicative noise from multiple trajectory data. *Automatica*, 144:110486, 2022. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2022.110486.

Kaiqing Zhang, Bin Hu, and Tamer Basar. Policy optimization for $\mathcal{H}_2$ linear control with $\mathcal{H}_\infty$ robustness guarantee: Implicit regularization and global convergence. In Alexandre M. Bayen, Ali Jadbabaie, George Pappas, Pablo A. Parrilo, Benjamin Recht, Claire Tomlin, and Melanie Zeilinger, editors, *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, volume 120 of *Proceedings of Machine Learning Research*, pages 179–190. PMLR, 10–11 Jun 2020.

Yang Zheng and Na Li. Non-asymptotic identification of linear dynamical systems using multiple trajectories. *IEEE Control Systems Letters*, PP:1–1, 12 2020. doi: 10.1109/LCSYS.2020.3042924.

Yang Zheng, Luca Furieri, Maryam Kamgarpour, and N. Li. Sample complexity of linear quadratic gaussian (lqg) control for output feedback systems. *Conference on Learning for Dynamics & Control*, 2020.

Learning to Stabilize Unknown LTI Systems on a Single Trajectory under Stochastic Noise (Supplementary Material)

Ziyi Zhang[1] **Yorie Nakahira**[1] **Guannan Qu**[1]

[1]Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

A PROOF OF THEOREM 5.1

One of the key innovations of this work is the SVD-based framework we use to decouple the unstable subspace from the rest of the system. Therefore, we prove Theorem 5.1 here. After the system runs for time T, we record the state space in a $n \times T$ matrix D whose t-th column is defined as:

$$D(t) = x_t = Ax_{t-1} + \eta_t.$$

We decompose A based on $E_u \oplus E_s$-decomposition. Suppose E_u and E_s are represented by their orthonormal bases $Q_1 \in \mathbb{R}^{n\times k}$ and $Q_2 \in \mathbb{R}^{n\times(n-k)}$, respectively, i.e. $E_u = \text{col}(Q_1), E_s = \text{col}(Q_2)$. Let $Q = [Q_1, Q_2]$ (which is invertible as A is diagonalizable), and let $R = [R_1^*, R_2^*]^* := Q^{-1}$. Since E_u and E_s are both invariant with regard to A, we know there exists $N_1 \in \mathbb{R}^{k\times k}$, $N_2 \in \mathbb{R}^{(n-k)\times(n-k)}$, s.t.

$$AQ = Q \begin{bmatrix} N_1 & \\ & N_2 \end{bmatrix} \Leftrightarrow N := \begin{bmatrix} N_1 & \\ & N_2 \end{bmatrix} = RAQ.$$

We are now ready to prove Theorem 5.1.

Proof: Let $D = U\Sigma V^*$ denote the compressed singular value decomposition of D and $\sigma_1 > \cdots > \sigma_n$ denote its singular values. In this case, we have $U \in \mathbb{R}^{n\times\min\{n,T\}}$, $\Sigma \in \mathbb{R}^{\min\{n,T\}\times\min\{n,T\}}$, and $V \in \mathbb{R}^{T\times\min\{n,T\}}$. Moreover, denote $U = [u_1, \ldots, u_n]$ and $V = [v_1, \ldots, v_n]$.

Furthermore, we have the following equalities

$$D = QRD = Q\begin{bmatrix} R_1 D \\ R_2 D \end{bmatrix} = Q\begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}\begin{bmatrix} D_1 \\ 0 \end{bmatrix} + \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}\begin{bmatrix} 0 \\ D_2 \end{bmatrix} = Q_1 D_1 + Q_2 D_2.$$

Let

$$\mathcal{D} = \begin{bmatrix} 0 & (Q_1 D_1)^* \\ Q_1 D_1 & 0 \end{bmatrix}, \quad J = \begin{bmatrix} 0 & (Q_2 D_2)^* \\ Q_2 D_2 & 0 \end{bmatrix}, \quad \mathcal{D} + J = \begin{bmatrix} 0 & D^* \\ D & 0 \end{bmatrix}.$$

We can decompose $\mathcal{D} + J$ in the following form

$$\mathcal{D} + J = \begin{bmatrix} 0 & V\Sigma U^* \\ U\Sigma V^* & 0 \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} V \\ U \end{bmatrix}\Sigma\begin{bmatrix} V \\ U \end{bmatrix}^* - \begin{bmatrix} V \\ -U \end{bmatrix}\Sigma\begin{bmatrix} V \\ -U \end{bmatrix}^*\right).$$

Therefore, we see that the eigenvalues of $\mathcal{D} + J$ are exactly $\{\pm\sigma_i\}$ with eigenvectors $[v_i^*, \pm u_i^*]^*$, respectively. Correspondingly, the top k largest eigenvalues of $\mathcal{D} + J$ are the top k largest singular values of D, or the square root of top k largest eigenvalues of DD^*.

Similarly, we use compressed singular value composition on $D_1 = U_1\Sigma_1 V_1^*$, where $U_1 \in \mathbb{R}^{k\times k}, \Sigma_1 \in \mathbb{R}^{k\times k}, V_1 \in \mathbb{R}^{T\times k}$,

and decompose $\mathcal{D}$ as follows:

$$\begin{aligned}
\mathcal{D} =& \begin{bmatrix} 0 & V_1\Sigma_1 U_1^* Q_1^* \\ Q_1 U_1 \Sigma_1 V_1^* & 0 \end{bmatrix} \\
=& \frac{1}{2}\left(\begin{bmatrix} V_1\Sigma_1 V_1^* & V_1\Sigma_1 U_1^* Q_1^* \\ Q_1 U_1 \Sigma_1 V_1^* & Q_1 U_1 \Sigma_1 U_1^* Q_1^* \end{bmatrix} - \begin{bmatrix} V_1\Sigma_1 V_1^* & -V_1\Sigma_1 U_1^* Q_1^* \\ -Q_1 U_1 \Sigma_1 V_1^* & Q_1 U_1 \Sigma_1 U^* Q_1^* \end{bmatrix} \right) \\
=& \frac{1}{2}\left(\begin{bmatrix} V_1\Sigma_1 \\ Q_1 U_1 \Sigma_1 \end{bmatrix} \begin{bmatrix} V_1^* & U_1^* Q_1^* \end{bmatrix} - \begin{bmatrix} V_1\Sigma_1 \\ -Q_1 U_1 \Sigma_1 \end{bmatrix} \begin{bmatrix} V_1^* & -U_1^* Q_1^* \end{bmatrix} \right) \\
=& \frac{1}{2}\left(\begin{bmatrix} V_1 \\ Q_1 U_1 \end{bmatrix} \Sigma_1 \begin{bmatrix} V_1 \\ Q_1 U_1 \end{bmatrix}^* - \begin{bmatrix} V_1 \\ -Q_1 U_1 \end{bmatrix} \Sigma_1 \begin{bmatrix} V_1 \\ -Q_1 U_1 \end{bmatrix}^* \right).
\end{aligned}$$

We see that the top k largest eigenvalues of $\mathcal{D}$ are the top k largest singular values of D_1, denoted as $\hat{\sigma}_1, \ldots, \hat{\sigma}_k$.

Let $U^{(k)}$ and $V^{(k)}$ denote the submatrices containing the first k columns of U and V, respectively. Let Π and Π' denote the projection onto the eigenspaces of the largest k eigenvectors of $\mathcal{D} + J$ and $\mathcal{D}$, respectively.

It is clear that

$$\begin{aligned}
\Pi &= \frac{1}{2} \begin{bmatrix} V^{(k)} \\ U^{(k)} \end{bmatrix} \begin{bmatrix} (V^{(k)})^* & (U^{(k)})^* \end{bmatrix} = \frac{1}{2} \begin{bmatrix} V^{(k)}(V^{(k)})^* & V^{(k)}(U^{(k)})^* \\ U^{(k)}(V^{(k)})^* & U^{(k)}(U^{(k)})^* \end{bmatrix}, \\
\Pi' &= \frac{1}{2} \begin{bmatrix} V_1 \\ Q_1 U_1 \end{bmatrix} \begin{bmatrix} V_1^* & U_1^* Q_1^* \end{bmatrix} = \frac{1}{2} \begin{bmatrix} V_1 V_1^* & V_1 U_1^* Q_1^* \\ Q_1 U_1 V_1^* & Q_1 U_1 U_1^* Q_1^* \end{bmatrix} = \frac{1}{2} \begin{bmatrix} V_1 V_1^* & V_1 U_1^* Q_1^* \\ Q_1 U_1 V_1^* & Q_1 Q_1^* \end{bmatrix}.
\end{aligned}$$

By Davis-Kahan Theorem (see Cao [2021] and Appendix G), we have

$$\|\Pi - \Pi'\| \leq \frac{1}{2} \frac{\sqrt{2k}\,\|J\|_2}{\hat{\sigma}_k - \sigma_{k+1}} = \frac{\sqrt{2k}\,\|Q_2 D_2\|}{\hat{\sigma}_k - \sigma_{k+1}} \leq \frac{\sqrt{2k}\,\|Q_2\|\,\|D_2\|}{\hat{\sigma}_k - \sigma_{k+1}} = \frac{\sqrt{2k}\,\|D_2\|}{\hat{\sigma}_k - \sigma_{k+1}}.$$

Since $\widehat{\Pi}_1 = U^{(k)}(U^{(k)})^*, \Pi_1 = Q_1 Q_1^*$, we have

$$\left\| \widehat{\Pi}_1 - \Pi_1 \right\| \leq \|\Pi - \Pi'\| \leq \frac{\sqrt{2k}\,\|D_2\|}{\hat{\sigma}_k - \sigma_{k+1}}.$$

We next show that $\hat{\sigma}_k = \Omega(|\lambda_k|^T)$, $\sigma_{k+1} = O(T)$ and $\|D_2\| = O(T)$, based on which $\left\| \widehat{\Pi}_1 - \Pi_1 \right\| \leq \frac{O(T)}{\Omega(\lambda_k^T - T)} \to 0$. More formally, we have the following.

Lemma A.1. *If*

$$T > \Theta\left(\frac{\log k - 2\log\left(\frac{gap}{k^{\frac{k}{2}+3}}\right) - 3\log\theta}{\log|\lambda_k|} \right) \tag{10}$$

is satisfied, with probability at least $1 - 4\theta$*,*

$$D_1 D_1^* \succeq \frac{\pi|\lambda_k|^{2T}\theta^2}{4} \frac{gap^2}{k^{k+6}} \frac{|\lambda_1|^2}{|\lambda_1|^2 - 1},$$

where we recall gap $= \left| \prod_{\substack{m_1 \neq m_2, \\ m_1, m_2 \in \{1,\ldots,k\}}} (\lambda_{m_1}^{-1} - \lambda_{m_2}^{-1}) \right|$.

The proof of Lemma A.1 is delayed to Appendix B.

For D_2, we have the following inequalities

$$\|D_2\|_2 \leq \sqrt{T}\,\|D_2\|_1 \leq \sqrt{T} \sum_{i=k+1}^{n} \left(\sum_{j=1}^{T} \lambda_i^j C \right) \leq \sqrt{T}(n-k)\left(\frac{C}{1 - |\lambda_{k+1}|} \right). \tag{11}$$

By Lemma A.1 and (11), in order to have $\left\|\widehat{\Pi}_1 - \Pi_1\right\| < \epsilon$, we need

$$
\begin{aligned}
&\left\|\widehat{\Pi}_1 - \Pi_1\right\| < \epsilon \\
\Leftarrow &\frac{\sqrt{2k}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\frac{\sqrt{\pi}|\lambda_k|^T\theta}{2}\frac{\text{gap}}{k^{\frac{k}{2}+3}}\sqrt{\frac{|\lambda_1|^2}{|\lambda_1|^2-1}} - 2\sqrt{2k}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)} < \epsilon \\
\Leftarrow &\frac{2\sqrt{2k}k^{\frac{k}{2}+3}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\sqrt{\pi}|\lambda_k|^T\theta\text{gap} - 4\sqrt{2k}k^{\frac{k}{2}+3}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)} < \epsilon \\
\Leftarrow &\frac{2\sqrt{2}k^{\frac{k+7}{2}}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\frac{1}{2}\sqrt{\pi}|\lambda_k|^T\theta\text{gap}} < \epsilon \qquad (12) \\
\Leftarrow &4\sqrt{2}k^{\frac{k+7}{2}}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right) < \sqrt{\pi}|\lambda_k|^T\theta\text{gap}\epsilon \\
\Leftarrow &\frac{1}{2}\log T + \log\left(4\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)\right) < T\log|\lambda_k| + \log\left(\sqrt{\pi}\theta\text{gap}\epsilon\right) \\
\Leftarrow &\frac{1}{2}T\log|\lambda_k| > \log\left(\frac{4\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\sqrt{\pi}\theta\text{gap}\epsilon}\right) \qquad (13) \\
\Leftarrow &T > \frac{2\log\left(\frac{4\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\sqrt{\pi}\theta\text{gap}\epsilon}\right)}{\log|\lambda_k|} \qquad (14)
\end{aligned}
$$

where in (12), we require

$$
\begin{aligned}
&4\sqrt{2k}k^{\frac{k}{2}+3}\sqrt{T}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right) < \frac{1}{2}\sqrt{\pi}|\lambda_k|^T\theta\text{gap} \\
\Leftarrow &\frac{1}{2}\log T + \log\left(4\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)\right) < T\log|\lambda_k| + \log(\frac{1}{2}\sqrt{\pi}\theta\text{gap}) \\
\Leftarrow &\frac{1}{2}T\log|\lambda_k| > \log\left(4\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)\right) - \log(\frac{1}{2}\sqrt{\pi}\theta\text{gap}) \qquad (15) \\
\Leftarrow &T > \frac{2\log\left(\frac{8\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\sqrt{\pi}\theta\text{gap}}\right)}{\log|\lambda_k|} \qquad (16)
\end{aligned}
$$

where in (13) and (15), we need $T\log|\lambda_k| > \log T$. In order to have $T\log|\lambda_k| > \log T$, define

$$f(T) := T\log|\lambda_k| - \log T.$$

When $T > \log|\lambda_k|$, we have $f(T) = (\log|\lambda_k|)^2 - \log\log|\lambda_k| > 0$ and $f'(T) = \log|\lambda_k| - \frac{1}{T} > 0$.

Therefore, when $T > \log|\lambda_k|$, we have $T\log|\lambda_k| > \log T$.

Combining (14), (16), and $T > \log|\lambda_k|$ required above, we get

$$T > \max\left\{\frac{2\log\left(\frac{8\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\sqrt{\pi}\theta\text{gap}}\right)}{\log|\lambda_k|}, \frac{2\log\left(\frac{4\sqrt{2}k^{\frac{k+7}{2}}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)}{\sqrt{\pi}\theta\text{gap}\epsilon}\right)}{\log|\lambda_k|}, \log|\lambda_k|\right\}. \qquad (17)$$

Treating the eigenvalue terms and θ to be constants as stated in the theorem, for $\left\|\widehat{\Pi}_1 - \Pi_1\right\| < \epsilon$ to hold, we need

$$T > \Theta((k\log k + \log(n-k) - \log\epsilon - \log\text{gap}). \qquad (18)$$

This concludes the proof. □

B AUXILLARY LEMMAS FOR STEP 1

We derive a lower bound on $D_1 D_1^*$ from Appendix 11 of Sarkar and Rakhlin [2018], which requires two additional functions $\phi_{\min}(M_1, T)$ and $\psi(M_1, T)$:

For the space $\mathbb{R}^d$, define the $a-$outbox, $S_d(a)$, as the following set

$$S_d(a) = \{v \in \mathbb{R}^d | \min_{1 \le i \le d} |v(i)| > a\},$$

which is used to quantify the following norm–like quantities of a matrix:

$$\phi_{\min}(M_1, T) = \sqrt{\inf_{v \in S_k(1)} \sigma_{\min}\left(\sum_{i=1}^{T} J^{-i+1} v v^* (J^{-i+1})^*\right)}, \tag{19}$$

where $M_1 = \bar{P}^{-1} J \bar{P}$ is the Jordan normal form of M_1. Here, Since A is diagonalizable (so M_1 is diagonalizable), J is the diagonal matrix of $\lambda_1, \cdots, \lambda_k$.

$$\psi(M_1, T) = \frac{1}{2k \sup_{1 \le i \le k} C_{|\bar{P}_i^* z_T|}}, \tag{20}$$

where C_X is the essential supremum of the probability distribution function (pdf) of X, $\bar{P} = [\bar{P}_1, \bar{P}_2, \cdots, \bar{P}_k]^*$, and

$$z_t := M_1^{-t} P_1^* x_t = \sum_{j=1}^{t} M_1^{-j} P_1^* \eta_j. \tag{21}$$

The following lemma is adapted from Appendix 11 of Sarkar and Rakhlin [2018].

Lemma B.1. *With probability at least* $1 - 4\theta$*,*

$$D_1 D_1^* \succeq \frac{1}{2} \phi_{\min}(M_1, T)^2 \psi(M_1, T)^2 \theta^2 M_1^T (M_1^T)^*,$$

whenever

$$\begin{aligned} &\left(4T^3 \lambda_k^{-2(T+1)\upsilon} k + \frac{T^2 k \sum_{i=1}^{k} \lambda_i^{-2(T+1)}}{\theta}\right) \\ &\le \frac{\phi_{\min}(M_1, T)^2 \psi(M_1, T)^2 \theta^2}{2}, \end{aligned} \tag{22}$$

and

$$T > \max\left\{\frac{C}{1 - |\lambda_{k+1}|}, \frac{C}{|\lambda_k| - 1}\right\}. \tag{23}$$

for some υ *such that* $(T+1)\upsilon = \lfloor \frac{T+1}{2} \rfloor$.

Note that in (22), we select T such that $\sum_{i=1}^{k} \sum_{t=1}^{T} |\lambda_i|^{-t} < kT$, or $T > \frac{1}{k} \sum_{i=1}^{k} \frac{\lambda_i}{\lambda_i - 1}$.

In Section C, we further prove the bounds on $\phi_{\min}$ and ψ in Lemma C.2 and Lemma C.4, which, combining with Lemma B.1 leads to the result in Lemma A.1 directly. It is clear that the bound in (10) under Lemma A.1 satisfies (23) in Lemma B.1 trivially. Therefore, to prove Lemma A.1, we just need to show that under (10), (22) in Lemma B.1 is satisfied.

Proof: [proof of Lemma A.1] To satisfy (22), we need

$$T^3 \lambda_k^{-2(T+1)\upsilon} k \le \frac{\phi_{\min}(M_1, T)^2 \psi(M_1, T)^2 \theta^2}{16}, \tag{24}$$

and

$$\frac{T^2 k \sum_{i=1}^{k} |\lambda_i|^{-2(T+1)}}{\theta} \le \frac{\phi_{\min}(M_1, T)^2 \psi(M_1, T)^2 \theta^2}{4}. \tag{25}$$

We then separately evaluate the conditions that would guarantee the satisfaction of the above inequities.

Condition (24): Taking the log, we have

$$\begin{aligned}
&3\log T - 2(T+1)\upsilon\log|\lambda_k| + \log k \le 2\log\left(\phi_{\min}(M_1,T)\psi(M_1,T)\theta\right) - \log 16\\
\overset{(a)}{\Leftarrow}&3T\upsilon\log|\lambda_k| - 2(T+1)\upsilon\log|\lambda_k| + \log k \le 2\log\left(\phi_{\min}(M_1,T)\psi(M_1,T)\theta\right) - \log 16\\
\Leftarrow&-(3T+2)\upsilon\log|\lambda_k| \le 2\log\left(\phi_{\min}(M_1,T)\psi(M_1,T)\theta\right) - \log 16 - \log k\\
\Leftarrow&T \ge \frac{\log 16 + \log k - 2\log\left(\phi_{\min}(M_1,T)\psi(M_1,T)\theta\right)}{3\upsilon\log|\lambda_k|} - 2.
\end{aligned}$$

where the step (a) uses the following: $T\upsilon\log|\lambda_k| > \log T$, which we show now. Define

$$f(T) := T\log|\lambda_k|^{\upsilon} - \log T.$$

When $T = \log|\lambda_k|^{\upsilon}$, we have $f(T) = \left(\log|\lambda_k|^{\upsilon}\right)^2 - \log\log|\lambda_k|^{\upsilon} > 0$. When $T \ge \log|\lambda_k|^{\upsilon}$, we have $f'(T) = \log|\lambda_k|^{\upsilon} - \frac{1}{T} > 0$.

Therefore, when $T > \log|\lambda_k|^{\upsilon}$, we have $T\upsilon\log|\lambda_k| > \log T$.

Condition (25): Since $|\lambda_1| > \ldots > |\lambda_k|$, to meet (25), it suffices to show:

$$\begin{aligned}
&T^2k^2|\lambda_k|^{-2(T+1)} \le \frac{\phi_{\min}(M_1,T)^2\psi(M_1,T)^2\theta^3}{4}\\
\Leftarrow&T^2|\lambda_k|^{-2(T+1)} \le \frac{\phi_{\min}(M_1,T)^2\psi(M_1,T)^2\theta^3}{4k^2}\\
\Leftarrow&2\log T - 2(T+1)\log|\lambda_k| \le \log\frac{\phi_{\min}(M_1,T)^2\psi(M_1,T)^2\theta^3}{4k^2}\\
\Leftarrow&T\log|\lambda_k| - 2(T+1)\log|\lambda_k| \le \log\frac{\phi_{\min}(M_1,T)^2\psi(M_1,T)^2\theta^3}{4k^2}\\
\Leftarrow&T \ge -\frac{\log\frac{\phi_{\min}(M_1,T)^2\psi(M_1,T)^2\theta^3}{4k^2}}{\log|\lambda_k|} + 2.
\end{aligned}$$

Similar to the derivation of (24), in order to get $T\log|\lambda_k| > 2\log T$, we need $T > 2\log|\lambda_k|$.

Combining the above and applying Lemma C.2 and Lemma C.4, we get the condition for T as in (10).

This concludes the proof of Lemma A.1. □

The following Corollary directly follows from Theorem 5.1.

Corollary B.2. *Under the premise of Theorem 5.1, for any orthonormal basis $\hat{P}_1$ of col$(\widehat{\Pi}_1)$ (where $\widehat{\Pi}_1$ is obtained by Algorithm 1), there exists a corresponding orthonormal basis P_1 of col(Π_1), such that*

$$\left\|P_1 - \hat{P}_1\right\| < \sqrt{2k}\epsilon := \delta.$$

The proof structure of Corollary B.2 is identical to the proof of Corollary 5.2 of Hu et al. [2022].

C PROOF OF AUXILIARY LEMMAS FOR APPENDIX B

In this section, we prove a few Lemmas that is used to bound $D_1D_1^*$ in Appendix B.

Lemma C.1. *Given a $k \times k$ Vandermonde Matrix Λ*

$$\Lambda = \begin{bmatrix} 1 & \lambda_1^{-1} & \cdots & \lambda_1^{-k+1}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & \lambda_k^{-1} & \cdots & \lambda_k^{-k+1}\end{bmatrix}, \tag{26}$$

and $\lambda_1, \ldots, \lambda_k \neq 0$*, then* $\left\|\Lambda^{-1}\right\| \leq \frac{k^{\frac{k}{2}+1}}{gap}$*, where*

$$gap = \left| \prod_{m_1 \neq m_2} (\lambda_{m_1}^{-1} - \lambda_{m_2}^{-1}) \right|. \tag{27}$$

Proof: From Theorem 1 of Tucci and Whiting [2011], we have

$$\Lambda^{-1}(i,j) = \frac{(-1)^{k-i}\sigma_{k-i}^{j}}{\prod_{m_1 \neq m_2}(\lambda_{m_1}^{-1} - \lambda_{m_2}^{-1})}, \tag{28}$$

where $\sigma_{k-i}^{j} := \sum_{\rho_{k-i}^{j}} \prod_{\ell \in \rho_{k-i}^{j}} \lambda_{\ell}^{-1}$ and ρ_{k-i}^{j} goes through all subsets of $\{\lambda_1^{-1}, \ldots, \lambda_{j-1}^{-1}, \lambda_{j+1}^{-1}, \ldots, \lambda_k^{-1}\}$ with cardinality $k-i$. In the above expression, the quantity σ_{k-i}^{j} can be bounded as:

$$\sigma_{k-i}^{j} \leq \binom{k}{k-i} \left(\frac{1}{\lambda_k}\right)^{k-i}. \tag{29}$$

Plugging (29) into (28) gives a bound for $|\Lambda^{-1}(i,j)|$ as follows:

$$\left\|\Lambda^{-1}(i,j)\right\| \leq \frac{\binom{k}{k-i}\left(\frac{1}{\lambda_k}\right)^{k-i}}{\text{gap}}. \tag{30}$$

Moreover, we have the following well-known inequality (see, for example, Horn and Johnson [1985])

$$\frac{1}{\sqrt{k}} \left\|\Lambda^{-1}\right\|_1 \leq \left\|\Lambda^{-1}\right\|_2 \leq \sqrt{k} \left\|\Lambda^{-1}\right\|_1. \tag{31}$$

Combining the above, we get

$$\left\|\Lambda^{-1}\right\| \leq \max_i \left\{ \sum_j \left|\Lambda^{-1}(i,j)\right| \right\} \leq \frac{k^{\frac{k}{2}+\frac{3}{2}}}{\text{gap}}. \tag{32}$$

where we have used the Sterling's formula for bounding $\binom{k}{k-i}$ in the summation. □

Lemma C.2. *Under the premise of Theorem 5.1, given* $\phi_{\min}$ *as defined in* (19)*, we have*

$$\phi_{\min}(M_1, T) \geq \frac{gap}{k^{\frac{k}{2}+2}}.$$

Proof:

Let $h_i(v) = \begin{bmatrix} \lambda_1^{-i+1} v(1) \\ \lambda_2^{-i+1} v(2) \\ \vdots \\ \lambda_k^{-i+1} v(k) \end{bmatrix} \in \mathbb{R}^k$, and $H(v) = \begin{pmatrix} h_1(v) & h_2(v) & \ldots & h_T(v) \end{pmatrix}$. Then we have

$$\begin{aligned} \phi_{\min}(M_1, T) &= \sqrt{\inf_{v \in S_d(1)} \sigma_{\min}\left(\sum_{i=1}^{T} h_i(v) h_i^*(v)\right)} \\ &= \sqrt{\inf_{v \in S_d(1)} \sigma_{\min}\left(H(v) H^*(v)\right)} \\ &= \sqrt{\inf_{v \in S_d(1)} \frac{1}{\|H^{-1}(v)\|^2}} \\ &= \inf_{v \in S_d(1)} \frac{1}{\|H^{-1}(v)\|} \end{aligned}$$

and we can decompose $H(v)$ as follows

$$\begin{aligned} H(v) &= \text{diag}\left(v(1), \ldots, v(k)\right) \begin{bmatrix} 1 & \lambda_1^{-1} & \ldots & \lambda_1^{-T+1} \\ 1 & \lambda_2^{-1} & \ldots & \lambda_2^{-T+1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_k^{-1} & \ldots & \lambda_k^{-T+1} \end{bmatrix} \\ &:= \text{diag}(v)\tilde{H}. \end{aligned}$$

Therefore,

$$\left\|H^{-1}(v)\right\| = \left\|\tilde{H}^{-1}(\text{diag}(v))^{-1}\right\| \leq \left\|\tilde{H}^{-1}\right\| \left\|(\text{diag}(v))^{-1}\right\|. \tag{33}$$

By Lemma C.1, we get

$$\left\|\tilde{H}^{-1}\right\| \leq \frac{k^{\frac{k}{2}+\frac{3}{2}}}{\text{gap}}.$$

Plugging the above inequality into (33) gives

$$\left\|H^{-1}(v)\right\| \leq \frac{k^{\frac{k}{2}+2}}{\text{gap}},$$

and

$$\phi_{\min}(M_1, T) \geq \frac{\text{gap}}{k^{\frac{k}{2}+2}}.$$

□

We also need a similar bound for $\psi(M_1, T)$. However, since we do not have an explicit formula for the pdf of noise η, it is difficult to evaluate $\sup_{1\leq i\leq k} C_{|\bar{P}_i^* z_T|}$ in (20) explicitly. However, it is intuitively clear that $\sup_{1\leq i\leq k} C_{|\bar{P}_i^* z_T|}$ is upper bounded by a constant, as z_T in (21) converges in distribution as $T \to \infty$. Therefore, the probability distribution function of $\bar{P}_i^* z_T$ also converges.

To demonstrate this more concretely, we explicitly compute the bound when $\eta_t \sim N(0, 1)$ follows the standard normal distribution:

Lemma C.3. *If η_t follows the standard normal distribution for all t, then*

$$C_{|\bar{P}_i^* z_T|} < \sqrt{\frac{2}{\pi}}\sqrt{\frac{|\lambda_i|^2 - 1}{|\lambda_i|^2}}$$

Proof: The j-th entry of $\bar{P}_i^* z_T$ can be represented as

$$\bar{P}_i^* z_T(j) = \sum_{t=1}^{T} v_i \cdot \left(M_1^{-t} P_1^* \eta_t\right) \sim N\left(0, \sum_{t=1}^{T} \left(|\lambda_j|^{-t}\right)^2\right)$$

so

$$\text{pdf}_{\bar{P}_i^* Z_T}(y) = \frac{1}{\sqrt{2\pi \sum_{t=1}^{T} \left(|\lambda_i|^{-t}\right)^2}} e^{-\frac{y^2}{2\sum_{t=1}^{T}\left(|\lambda_i|^{-t}\right)^2}}, \qquad y \in \mathbb{R}.$$

With some algebra, we get

$$\text{pdf}_{|\bar{P}_i^* z_T|}(y) = \frac{\sqrt{2}}{\sqrt{\pi \sum_{t=1}^{T} \left(|\lambda_i|^{-t}\right)^2}} e^{-\frac{y^2}{2\sum_{t=1}^{T}\left(|\lambda_i|^{-t}\right)^2}}, \qquad y \in \mathbb{R}^+.$$

Therefore, $C_{|\bar{P}_i^* z_T|} \leq \frac{\sqrt{2}}{\sqrt{\pi \sum_{t=1}^{T}\left(|\lambda_i|^{-t}\right)^2}} \leq \sqrt{\frac{2}{\pi}}\sqrt{\frac{|\lambda_i|^2-1}{|\lambda_i|^2}}$. □

In the rest of the paper, we will assume $C_{|\bar{P}_i' z_T|}$ is bounded and take

$$C_{|\bar{P}_i^* z_T|} < C_z, \tag{34}$$

for some constant C_z, as in Assumption 2. Therefore, the following result directly follows:

Lemma C.4. *Under the premise of Theorem 5.1, given ψ as defined in* (20)*, we have*

$$\psi(M_1, T) \geq \frac{1}{2kC_z}.$$

D SOLUTION TO THE LEAST SQUARE PROBLEM IN STAGE 2

Lemma D.1 gives the explicit form for the solution to the least squares problem in Algorithm 1

Lemma D.1. *Given $D := [x_0, \cdots, x_T]$ and $\hat{\Pi}_1 = U^{(k)}(U^{(k)})^*$, the solution to*

$$\hat{M}_1 = \arg\min_{M_1} \sum_{t=0}^{T} \left\| (U^{(k)})^* x_{t+1} - M_1 (U^{(k)})^* x_t \right\|^2$$

is uniquely given by $\hat{M}_1 = (U^{(k)})^* A U^{(k)} + \varpi$, where $\varpi = \left(\sum_t (U^{(k)})^* \eta_t x_t^* U^{(k)}\right) ((\Sigma^{(k)})^2)^{-1}$.

Proof: Sincec M_1 is a stationary point of $\mathcal{L}$, for any Δ in the neighborhood of O, we have

$$\begin{aligned}
0 \leq & \mathcal{L}(M_1 + \Delta) - \mathcal{L}(M_1) \\
= & \sum_t \|\hat{y}_{1,t+1} - M_1 \hat{y}_{1,t} - \Delta \hat{y}_{1,t}\|^2 - \sum_t \|\hat{y}_{1,t+1} - M_1 \hat{y}_{1,t}\|^2 \\
= & \sum_t \langle \Delta \hat{y}_{1,t}, \hat{y}_{1,t+1} - M_1 \hat{y}_{1,t} \rangle + O(\|\Delta\|^2) \\
= & \sum_t \operatorname{tr}\left(\hat{y}_{1,t}^* \Delta^* (\hat{y}_{1,t+1} - M_1 \hat{y}_{1,t})\right) + O(\|\Delta\|^2) \\
= & \sum_t \operatorname{tr}\left(\Delta^* (\hat{y}_{1,t+1} - M_1 \hat{y}_{1,t}) \hat{y}_{1,t}^*\right) + O(\|\Delta\|^2) \\
= & \operatorname{tr}\left(\Delta^* \sum_t (\hat{y}_{1,t+1} - M_1 \hat{y}_{1,t}) \hat{y}_{1,t}^*\right) + O(\|\Delta\|^2).
\end{aligned}$$

Since the above holds for all Δ, we get

$$\sum_t (\hat{y}_{1,t+1} - M_1 \hat{y}_{1,t}) \hat{y}_{1,t}^* \Leftrightarrow M_1 \sum_t \hat{y}_{1,t} \hat{y}_{1,t}^* = \sum_t \hat{y}_{1,t+1} \hat{y}_{1,t}^*.$$

Plugging in $\hat{y}_{1,t} = (U^{(k)})^* x_t$ and $\hat{y}_{1,t+1} = (U^{(k)})^*(A x_t + \eta_t)$, we have

$$\begin{aligned}
M_1 (U^{(k)})^* D D^* U^{(k)} &= M_1 \sum_t (U^{(k)})^* x_t x_t^* U^{(k)} \\
&= \sum_t (U^{(k)})^* (A x_t + \eta_t) x_t^* U^{(k)} \\
&= (U^{(k)})^* A D D^* U^{(k)} + \sum_t (U^{(k)})^* \eta_t x_t^* U^{(k)}.
\end{aligned}$$

Since $U^{(k)}$ are the first k singular vectors of D, we have the following equalities:

$$(U^{(k)})^* D D^* U^{(k)} = (U^{(k)})^* U \Sigma V^* V \Sigma^* U^* U^{(k)} = \begin{bmatrix} I^{(k)} & 0 \end{bmatrix} \Sigma^2 \begin{bmatrix} I^{(k)} \\ 0 \end{bmatrix} = (\Sigma^{(k)})^2, \tag{35}$$

which is invertible, and $\hat{M}_1$ is explicitly given by

$$\hat{M}_1 = \left((U^{(k)})^* A D D^* U^{(k)} + \sum_t (U^{(k)})^* \eta_t x_t^* U^{(k)} \right) (\Sigma^{(k)})^{-2}. \tag{36}$$

Moreover, we have

$$\begin{aligned}
& U^{(k)} (U^{(k)})^* D D^* U^{(k)} = U^{(k)} (\Sigma^{(k)})^2 \\
= & \begin{bmatrix} U^{(k)} & 0 \end{bmatrix} \begin{bmatrix} (\Sigma^{(k)})^2 \\ 0 \end{bmatrix} = U \begin{bmatrix} (\Sigma^{(k)})^2 \\ 0 \end{bmatrix} \\
= & U \Sigma^2 \begin{bmatrix} I^{(k)} \\ 0 \end{bmatrix} = U \Sigma^2 U^* U^{(k)} = D D^* U^{(k)},
\end{aligned}$$

where the first equality is obtained by using (35). Substituting the above in (36) yields

$$\begin{aligned}\hat{M}_1 &= \left((U^{(k)})^* A (U^{(k)}(U^{(k)})^* DD^*) U^{(k)}\right)(\Sigma^{(k)})^{-2} + \varpi \\ &= \left((U^{(k)})^* A U^{(k)} (U^{(k)})^*\right)\left(DD^* U^{(k)}\right)(\Sigma^{(k)})^{-2} + \varpi \\ &= (U^{(k)})^* A U^{(k)} + \varpi,\end{aligned}$$

where $\varpi = \left(\sum_t (U^{(k)})^* \eta_t x_t^* U^{(k)}\right)(\Sigma^{(k)})^{-2}$. □

We want to show $(U^{(k)})^* A U^{(k)}$ is the dominating term of the above expression, as we will bound ϖ in the following lemma.

Lemma D.2. *Under the premise of Theorem 5.1,*

$$\left\| M_1 - \hat{M}_1 \right\| < 3 \|A\| \delta$$

for any $\delta > 0$ whenever

$$T \geq \frac{\log\left(\frac{4C}{\pi\theta^2\|A\|\delta}\frac{k^{k+6}}{gap^2}\right)}{\log|\lambda_k|}.$$

Proof: First, we prove that $\varpi \leq \delta$. Let $H = [\eta_1, \ldots, \eta_T]$, then we have

$$\begin{aligned}\varpi &= (U^{(k)})^* H D^* U^{(k)} (\Sigma^{(k)})^{-2} \\ &= (U^{(k)})^* H V \Sigma^* U^* U^{(k)} (\Sigma^{(k)})^{-2} \\ &= (U^{(k)})^* H V \Sigma^* \begin{bmatrix} I^{(k)} \\ 0 \end{bmatrix} (\Sigma^{(k)})^{-2} \\ &= (U^{(k)})^* H V \begin{bmatrix} \Sigma^{(k)} \\ 0 \end{bmatrix} (\Sigma^{(k)})^{-2} \\ &= (U^{(k)})^* H V \begin{bmatrix} (\Sigma^{(k)})^{-1} \\ 0 \end{bmatrix}.\end{aligned}$$

Therefore,

$$\begin{aligned}&\|\varpi\| \leq \|A\| \delta \\ \Leftarrow &\|H\| \left\|(\Sigma^{(k)})^{-1}\right\| \leq \|A\| \delta \\ \Leftarrow &\sqrt{T} C \frac{2}{\sqrt{\pi}|\lambda_k|^T \theta} \frac{k^{\frac{k}{2}+3}}{\text{gap}} \sqrt{\frac{|\lambda_1|^2 - 1}{|\lambda_1|^2}} \leq \|A\| \delta \qquad (37) \\ \Leftarrow &\frac{|\lambda_k|^T}{\sqrt{T}} \geq \frac{2C}{\sqrt{\pi}\theta \|A\| \delta} \frac{k^{\frac{k}{2}+3}}{\text{gap}} \\ \Leftarrow &T \log|\lambda_k| - \frac{1}{2}\log T \geq \log\left(\frac{2C}{\pi\theta\|A\|\delta}\frac{k^{\frac{k}{2}+3}}{\text{gap}}\right) \\ \Leftarrow &\frac{1}{2} T \log|\lambda_k| \geq \log\left(\frac{2C}{\pi\theta\|A\|\delta}\frac{k^{\frac{k}{2}+3}}{\text{gap}}\right) \qquad (38) \\ \Leftarrow &T \geq \frac{2\log\left(\frac{2C}{\pi\theta\|A\|\delta}\frac{k^{\frac{k}{2}+3}}{\text{gap}}\right)}{\log|\lambda_k|}, \qquad (39)\end{aligned}$$

where (37) used Lemma A.1 and that for a $n \times T$ matrix H, $\|H\|_2 \leq \sqrt{T}\|H\|_1$, and (38) requires $\log T < T \log|\lambda_k|$, which is satisfied when we derived (13) and (15). We can use Lemma A.1 to bound $\left\|(\Sigma^{(k)})^{-1}\right\|$ is a direct result of Cauchy Interlacing Theorem. We further observe that (39) does not change the criteria obtained in (18).

Recall that $U^{(k)} = \hat{P}_1$. We obtain

$$\begin{aligned}
\left\| M_1 - \hat{M}_1 \right\| &= P_1^* A P_1 - \left((U^{(k)})^* A U^{(k)} + \varpi \right) \\
&\leq \left\| P_1^* A P_1 - P_1^* A \hat{P}_1^* \right\| + \left\| P_1^* A \hat{P}_1 - \hat{P}_1^* A \hat{P}_1^* \right\| + \|\varpi\| \\
&\leq \|A\| \left\| P_1 - \hat{P}_1 \right\| + \|A\| \left\| P_1 - \hat{P}_1 \right\| + \|\varpi\| \\
&\leq 3 \|A\| \delta.
\end{aligned}$$

where in the last inequality, we used Corollary B.2. □

With Lemma D.2, we are ready to prove Proposition 5.2.

Proof: [Proof of Proposition 5.2] By Lemma D.2, we get $\left\| M_1 - \hat{M}_1 \right\| < 3 \|A\| \delta$. Moreover, by Gelfand's formula, we have

$$\begin{aligned}
\left\| M_1^t \right\| &= \left\| P_1^* A^t P_1 \right\| \leq \left\| A^t \right\| \leq \zeta_{\epsilon_1}(A)(|\lambda_1| + \epsilon_1)^t, \\
\left\| \hat{M}_1^t \right\| &= \left\| \hat{P}_1^* A^t \hat{P}_1 \right\| \leq \left\| A^t \right\| \leq \zeta_{\epsilon_1}(A)(|\lambda_1| + \epsilon_1)^t,
\end{aligned}$$

Therefore, by telescoping, we get

$$\begin{aligned}
\left\| M_1^\tau - \hat{M}_1^\tau \right\| &= \left\| \sum_{i=1}^{\tau} (M_1^i \hat{M}_1^{\tau-i} - M_1^{i-1} \hat{M}_1^{\tau-i+1}) \right\| \\
&\leq \left\| M_1^{i-1} \right\| \left\| M_1^{\tau-i} \right\| \left\| M_1 - \hat{M}_1 \right\| \\
&< \tau \cdot \zeta_{\epsilon_1}(A)^2 (|\lambda_1| + \epsilon_1)^{\tau-1} \cdot 3 \|A\| \delta \\
&= 3\tau \|A\| \zeta_{\epsilon_1}(A)^2 (|\lambda_1| + \epsilon_1)^{\tau-1} \delta.
\end{aligned}$$

□

With Proposition 5.2, the following corollary easily follows:

Corollary D.3. *Under the premise of Theorem 4.2, when $\delta < \frac{1}{\tau}$,*

$$\left\| \hat{M}_1^\tau \right\| < \left(\zeta_{\epsilon_1}(M_1)(|\lambda_1| + \epsilon_1) + 3 \|A\| \zeta_{\epsilon_1}(A) \right) (|\lambda_1| + \epsilon_1)^{\tau-1}.$$

Proof: By Gelfand's formula and Proposition 5.2,

$$\begin{aligned}
\left\| \hat{M}_1^\tau \right\| &\leq \|M_1^\tau\| + \left\| \hat{M}_1^\tau - M_1^\tau \right\| \\
&\leq \zeta_{\epsilon_1}(A)(\lambda_1 + \epsilon_1)^\tau + 3\tau \|A\| \zeta_{\epsilon_1}(A)^2 (|\lambda_1| + \epsilon_1)^{\tau-1} \delta \\
&< \left(\zeta_{\epsilon_1}(M_1)(|\lambda_1| + \epsilon_1) + 3 \|A\| \zeta_{\epsilon_1}(A) \right) (|\lambda_1| + \epsilon_1)^{\tau-1}.
\end{aligned}$$

where the last inequality requires $\delta < \frac{1}{\tau}$. □

E BOUNDING $\left\| \hat{B}_\tau - B_\tau \right\|$

Lemma E.1. *For any $\gamma > \epsilon$, the following implication holds:*

$$\frac{\|R_2 x\|}{\|x\|} \leq \gamma - \epsilon := \gamma' \quad \Rightarrow \quad \frac{\left\| (I - \hat{\Pi}_1) x \right\|}{\|x\|} \leq \gamma$$

Proof:

$$\begin{aligned}\frac{\left\|(I-\hat{\Pi}_1)x\right\|}{\|x\|} &= \frac{\left\|(I-\hat{\Pi}_1+\Pi_1-\Pi_1)x\right\|}{\|x\|} \\ &\leq \frac{\|(I-\Pi_1)x\|}{\|x\|} + \frac{\left\|\hat{\Pi}_1-\Pi_1\right\|\|x\|}{\|x\|} \\ &\leq \frac{\|\Pi_2 x\|}{\|x\|} + \epsilon \\ &= \frac{\|\Pi_2\Pi_s x\|}{\|x\|} + \epsilon \qquad (40) \\ &\leq \frac{\|\Pi_s x\|}{\|x\|} + \epsilon \\ &\leq \gamma \qquad (41)\end{aligned}$$

where (40) holds because E_2 is orthogonal to E_1, therefore $\Pi_2\Pi_u = 0$, as $P_2P_2^*Q_1R_1 = 0$ by orthogonality of P_2 and $Q_1 = P_1$. □

In the following propositions, we show that the stopping time ω_i defined in Algorithm 1 guarantees a bound on $\|x_t\|$.

Proposition E.2. *Under the premise of Theorem 4.2, for any constant $\gamma > \epsilon$, if in the open loop system,*

$$\frac{\left\|(I-\hat{\Pi}_1)x_t\right\|}{\|x_t\|} > \gamma,$$

then, exists $C_\gamma \in \mathbb{R}^+$ such that $\|x_t\| < C_\gamma$.

Proof: Since we have that $x_t = \sum_{j=0}^{t} A^{t-j}\eta_j$, we have

$$Rx_t = \begin{bmatrix} R_1 x_t \\ R_2 x_t \end{bmatrix} = \begin{bmatrix} R_1 \sum_{j=0}^{t} A^{t-j}\eta_j \\ R_2 \sum_{j=0}^{t} A^{t-j}\eta_j \end{bmatrix} = \begin{bmatrix} \sum_{j=0}^{t} N_1^{t-j} R_1 \eta_j \\ \sum_{j=0}^{t} N_2^{t-j} R_2 \eta_j \end{bmatrix}$$

Therefore, we have that

$$\begin{aligned}\|R_2 x\| &\leq \sum_{j=0}^{t} \left\|N_2^j\right\| \|R_2\| C \leq \sum_{j=0}^{t} \zeta_{\epsilon_4}(N_2)(\lambda_{k+1}+\epsilon_4)^j \|R_2\| C \\ &\leq \frac{\zeta_{\epsilon_4}(N_2)C}{1-\xi} \frac{1}{1-(\lambda_{k+1}+\epsilon_4)}\end{aligned}$$

where we used Lemma A.1 of Hu et al. [2022]. As $\|R_2 x_j\|$ is bounded above by a constant, so is $\|\Pi_s x_t\| = \|Q_2 R_2 x_t\|$.

Since $\frac{\left\|(I-\hat{\Pi}_1)x_t\right\|}{\|x_t\|} > \gamma$, by Lemma E.1, $\frac{\|R_2 x_t\|}{\|x_t\|} > \gamma'$. Correspondingly, we have

$$\gamma' < \frac{\|R_2 x_t\|}{\|x_t\|},$$

which implies

$$\|x_t\| < \frac{\zeta_{\epsilon_4}(N_2)C}{\gamma'(1-\xi)} \frac{1}{1-(|\lambda_{k+1}|+\epsilon_4)} := C_\gamma. \qquad (42)$$

□

Proposition E.3. *Under the premise of Theorem 4.2, for any constant $\gamma > \epsilon$, consider the initial state x_i such that $\frac{\|P_2^* x_i\|}{\|x_i\|} > \gamma$. Moreover, $x_{i+1} = Ax_i + Bu + \eta_i$, i.e. we insert control right after the initial state and let the system run in open-loop thereafter. If for $t \in \mathbb{Z}^+$ such that*

$$\frac{\left\|(I-\hat{\Pi}_1)x_{i+t}\right\|}{\|x_{i+t}\|} > \gamma,$$

then, for all $\alpha < \frac{1}{\|B\|}$,

$$\|x_{i+t}\| < \frac{1}{\gamma'}\left(\frac{2\zeta_{\epsilon_4}(N_2)}{1-\xi}\|x_i\| + C_\gamma\right).$$

Proof:

$$\begin{aligned}\|R_2 x_{i+t}\| \leq & \left\|N_2^t R_2 x_i + N_2^{t-1} R_2 B u\right\| + \sum_{j=0}^{t}\left\|N_2^j\right\| \|R_2\| C \\ \leq & \frac{\zeta_{\epsilon_4}(N_2)}{1-\xi}(|\lambda_{k+1}| + \epsilon_4)^{t-1}((1+\alpha\|B\|)\|x_i\|) + C_\gamma \\ \leq & \frac{2\zeta_{\epsilon_4}(N_2)}{1-\xi}(|\lambda_{k+1}| + \epsilon_4)^{t-1}\|x_i\| + C_\gamma. \end{aligned} \tag{43}$$

Since $\frac{\|(I-\hat{\Pi}_1)x_{i+t}\|}{\|x_{i+t}\|} > \gamma$, by Lemma E.1, we have that

$$\gamma' < \frac{\|R_2 x_{i+t}\|}{\|x_{i+t}\|}.$$

Substitute the above in (43) finishes the proof. □

Proposition E.4. *Under the premise of Theorem 4.2, for any constant* $\gamma > \epsilon$ *and stopping time* ω_i *such that:*

$$\omega_i = \min\left\{t > t_{i-1} : \frac{\left\|(I-\hat{\Pi}_1)x_t\right\|}{\|x_t\|} \leq \gamma \wedge \|x_t\| > \frac{C}{\delta}\right\},$$

where we assume $t_0 = T$. *Then, Algorithm 1 guarantees that*

$$\frac{\|P_2^* x_{t_i}\|}{\|x_{t_i}\|} < \gamma + \epsilon, \qquad \forall i \in \{1, \ldots, m\},$$

while maintaining

$$\|x_{t_1}\| \leq \max\left\{\|A\|\frac{C}{\delta} + C, \|A\| C_\gamma + C, \|x_T\|\right\},$$

$$\|x_t\| < \max\left\{\|A\|\frac{C}{\delta} + C, \left(\frac{\|A\|}{\gamma'}\frac{2\zeta_{\epsilon_4}(N_2)}{1-\xi}\right)^i \|x_{t_1}\| + \sum_{j=1}^{i-1}\left(\frac{\|A\|}{\gamma'}\frac{2\zeta_{\epsilon_4}(N_2)}{1-\xi}\right)^j\left(\frac{\|A\|}{\gamma'}C_\gamma + C\right)\right\}, \quad \forall t_i \leq t \leq t_{i+1}.$$

Proof: Similar to the steps in proof of Lemma E.1, we obtain that

$$\frac{\|P_2^* x_{t_i}\|}{\|x_{t_i}\|} = \frac{\|\Pi_2(\Pi_u + \Pi_s)x_{t_i}\|}{\|x_{t_i}\|} = \frac{\|\Pi_2\Pi_s x_{t_i}\|}{\|x_{t_i}\|} \leq \frac{\|\Pi_2 x_{t_i}\|}{\|x_{t_i}\|} = \frac{\left\|(I - \hat{\Pi}_1 + \hat{\Pi}_1 - \Pi_1)x_{t_i}\right\|}{\|x_{t_i}\|} \leq \gamma + \epsilon,$$

which shows the first part of the result.

We now focus on the second part (bounding $\|x_t\|$). For the base case, We either have $t_1 = T$, thus $x_{t_1} = x_T$, in which case the stopping time criteria is already met after Stage 1 of algorithm 1, or, if $t_1 > T$, there are two scenarios depending which of the two stopping criteria is violated at time $t_1 - 1$. If $\frac{\|(I-\hat{\Pi}_1)x_{t_1-1}\|}{\|x_{t_1-1}\|} > \gamma$, by Proposition E.2, we have $\|x_{t_1-1}\| < C_\gamma$, where C_γ is defined in (42), in which case, we have

$$\|x_{t_1}\| = \|Ax_{t_1-1} + \eta_{t_1-1}\| \leq \|A\| C_\gamma + C. \tag{44}$$

In the second case, $\|x_{t_1-1}\| \leq \frac{C}{\delta}$, so we have

$$\|x_{t_1}\| < \|A\|\frac{C}{\delta} + C.$$

Therefore, to sum up the base case, we have

$$\|x_{t_1}\| \leq \max\left\{\|A\|\frac{C}{\delta}+C, \|A\|C_\gamma + C, \|x_T\|\right\}$$

For the induction case, given $\|x_{t_i}\|$, there are again two cases depending on which criterion is violated at time $t_{i+1}-1$. If $\left\|x_{t_{i+1}-1}\right\| \leq \frac{C}{\delta}$, we have

$$\left\|x_{t_{i+1}}\right\| < \|A\|\frac{C}{\delta}+C.$$

Otherwise, if $\frac{\left\|(I-\hat{\Pi}_1)x_{t_{i+1}-1}\right\|}{\left\|x_{t_{i+1}-1}\right\|} > \gamma$, by Proposition E.3, we obtain that

$$\left\|x_{t_{i+1}-1}\right\| < \frac{1}{\gamma'}\left(\frac{2\zeta_{\epsilon_4}(N_2)}{1-\xi}\|x_i\| + C_\gamma\right), \tag{45}$$

where γ' is defined in Lemma E.1.

By the definition of ω_i, the maximum of the above inequalities also holds for all x_t such that $t < t_{i+1}$. Therefore,

$$\left\|x_{t_{i+1}}\right\| < \max\left\{\|A\|\frac{C}{\delta}+C, \frac{\|A\|}{\gamma'}\left(\frac{2\zeta_{\epsilon_4}(N_2)}{1-\xi}\|x_i\| + C_\gamma\right)+C\right\},$$

as required. Note that the same bound above also holds for all $t_i < t < t_{i+1}$. Hence we get the desired result after a simple recursive expansion. □

We are now ready to bound $\left\|\hat{B}_\tau - B_\tau\right\|$.

Proposition E.5. *Under the premise of Theorem 4.2,*

$$\left\|\hat{B}_\tau - B_\tau\right\| < C_B\left(|\lambda_1|+\epsilon_1\right)^{\tau-1}\delta,$$

where $C_B := (\zeta_{\epsilon_1}^2(A)(3\tau\|A\| + \|B\| + \tau C + 1) + (\tau+1)C_\Delta)\frac{\sqrt{m}}{\alpha}$.

Proof: We have

$$\begin{aligned}
\left\|b_i - \hat{b}_i\right\| =& \frac{1}{\alpha\|x_{t_i}\|}\Bigg\|P_1^* x_{t_i+\tau} - M_1^\tau P_1^* x_{t_i} - \Delta_\tau P_2^* x_{t_i} - \sum_{j=1}^{\tau-1}(M_1^{\tau-j}P_1^*\eta_{t_i+j} - \Delta_{\tau-j}P_2^*\eta_{t_i+j}) \\
&- \left(\hat{P}_1^* x_{t_i+\tau} - \hat{M}_1^\tau \hat{P}_1^* x_{t_i}\right)\Bigg\| \\
\leq& \frac{1}{\alpha\|x_{t_i}\|}\Bigg(\left\|(P_1-\hat{P}_1)^*\left(A^\tau x_{t_i} + B_\tau u_{t_i}\right)\right\| + \left\|\sum_{j=1}^{\tau-1} M_1^{\tau-j}(P_1-\hat{P}_1)^*\eta_{t_i+j}\right\| + \left\|M_1^\tau P_1^* x_{t_i} - \hat{M}_1^\tau\hat{P}_1^* x_{t_i}\right\| \\
&+ \|\Delta_\tau P_2^* x_{t_i}\| + \sum_{j=1}^{\tau-1}\left\|M_1^{\tau-j}P_1^*\eta_{t_i+j}\right\| + \sum_{j=1}^{\tau-1}\|\Delta_{\tau-j}P_2^*\eta_{t_i+j}\|\Bigg).
\end{aligned}$$

Here, the first term is bounded by

$$\begin{aligned}
\left\|(P_1-\hat{P}_1)^*\left(A^\tau x_{t_i} + B_\tau u_{t_i}\right)\right\| \leq& \left\|P_1 - \hat{P}_1\right\|\left(\|A^\tau\| + \left\|A^{\tau-1}B\right\|\right)\|x_{t_i}\| \\
\leq& \|x_{t_i}\|\,\zeta_{\epsilon_1}(A)\left(|\lambda_1|+\epsilon_1\right)^{\tau-1}(\|A\|+\|B\|)\delta,
\end{aligned}$$

where in the last inequality we applied Corollary B.2 and Gelfand's formula; the second term is bounded by

$$\begin{aligned}
\left\|\sum_{j=1}^{\tau-1} M_1^{\tau-j}(P_1-\hat{P}_1)^*\eta_{t_i+j}\right\| \leq& \sum_{j=1}^{\tau-1}\zeta_{\epsilon_1}(A)\left(|\lambda_1|+\epsilon_1\right)^{\tau-j}C\delta \\
<& \tau\zeta_{\epsilon_1}(A)\left(|\lambda_1|+\epsilon_1\right)^{\tau-1}C\delta,
\end{aligned}$$

where we used Corollary B.2 and Gelfand's formula.

The third term is bounded above by

$$\begin{aligned}\left\|M_1^\tau P_1^* x_{t_i} - \hat{M}_1^\tau \hat{P}_1^* x_{t_i}\right\| \leq & \left(\left\|M_1^\tau (P_1 - \hat{P}_1)^*\right\| + \left\|(M_1^\tau - \hat{M}_1^\tau)\hat{P}_1^*)\right\|\right) \|x_{t_i}\| \\ <& \left(\zeta_{\epsilon_1}(A)\left(|\lambda_1| + \epsilon_1\right)^{\tau-1} \|A\| \delta + 3\tau \|A\| \zeta_{\epsilon_1}(A)\left(|\lambda_1| + \epsilon_1\right)^{\tau-1} \delta\right) \|x_{t_i}\| \\ \leq & \|x_{t_i}\| \zeta_{\epsilon_1}(A)^2 \left(|\lambda_1| + \epsilon_1\right)^{\tau-1} (3\tau + 1) \|A\| \delta,\end{aligned}$$

where we applied Gelfand's formula and Proposition 5.2. The fourth term is bounded by

$$\frac{\|\Delta_\tau\| \|P_2^* x_{t_i}\|}{\|x_{t_i}\|} \leq C_\Delta (|\lambda_1| + \epsilon_1)^\tau (\gamma + \epsilon) \tag{46}$$

$$\leq C_\Delta (|\lambda_1| + \epsilon_1)^\tau \delta, \tag{47}$$

where in (46), we used Proposition G.1 of Hu et al. [2022] and Proposition E.4, while and (47) we need to pick stopping time ω defined by γ:

$$\gamma \leq \delta - \epsilon = (\sqrt{2k} - 1)\epsilon. \tag{48}$$

For the second to last and the last term,

$$\begin{aligned}\frac{1}{\|x_{t_i}\|} \sum_{j=1}^{\tau-1} \left\|M_1^{\tau-j} P_1^* \eta_{t_i+j}\right\| \leq & \frac{1}{\|x_{t_i}\|} \sum_{j=1}^{\tau-1} \zeta_{\epsilon_1}(A)\left(|\lambda_1| + \epsilon_1\right)^{\tau-j} C \\ < & \frac{1}{\|x_{t_i}\|} \tau \zeta_{\epsilon_1}(A)\left(|\lambda_1| + \epsilon_1\right)^{\tau-1} C \\ < & \tau \zeta_{\epsilon_1}(A)\left(|\lambda_1| + \epsilon_1\right)^{\tau-1} \delta,\end{aligned} \tag{49}$$

$$\begin{aligned}\frac{1}{\|x_{t_i}\|} \sum_{j=1}^{\tau-1} \|\Delta_{\tau-j} P_2^* \eta_{t_i+j}\| \leq & \frac{1}{\|x_{t_i}\|} \tau C_\Delta (|\lambda_1| + \epsilon_1)^\tau C \\ \leq & \tau C_\Delta (|\lambda_1| + \epsilon_1)^\tau \delta,\end{aligned} \tag{50}$$

where in (49) and (50), we need

$$\frac{C}{\|x_{t_i}\|} < \delta. \tag{51}$$

We notice that (51) happens with high probability since the system runs mostly in open loop. If the above inequality is not satisfied, we can keep the system running in open loop until it is. If the above is never satisfied, then the system is stable. More formally, as the first stopping time t_1 stated in Proposition E.4 is never reached, the bound for $\|x_{t_1}\|$ holds for all x_t.

Finally, to bound the error of the whole matrix, we simply apply the definition

$$\begin{aligned}\left\|\hat{B}_\tau - B_\tau\right\| = & \max_{\|u\|=1} \left\|(\hat{B}_\tau - B_\tau)u\right\| \leq \max_{\|u\|=1} \sum_{i=1}^m |u_i| \left\|\hat{b}_i - b_i\right\| \\ < & (\zeta_{\epsilon_1}^2(A)(3\tau \|A\| + \|B\| + \tau C + 1) + (\tau + 1)C_\Delta)\left(|\lambda_1| + \epsilon_1\right)^{\tau-1} \delta \frac{\sqrt{m}}{\alpha}.\end{aligned}$$

□

F PROOF OF MAIN THEOREM

We assumed the system (A, B) is controllable. As we are stabilizing the system in (M^τ, B_τ), we need to first show that (M^τ, B_τ) is stabilizable.

Proposition F.1. *If (A, B) is controllable, then $(\hat{M}_1^\tau, R_1 \hat{B}_\tau)$ is stabilizable.*

Proof: Since (A, B) is controllable, by the PBH test criteria, there exists b, such that for all unit left eigenvector $\bar{w}$ of A, $\|\bar{w}^* B\| > b$.

Let w^* denote an arbitrary unit left eigenvector of N_1 with eigenvalue λ, so

$$w^* N_1 = \lambda w^* \quad \Rightarrow \quad (R_1^* w)^* A = w^* R_1 Q_1 N_1 R_1 = \lambda (R_1^* w)^*.$$

Therefore, $R_1^* w$ is a left eigenvector of A, which leads to

$$\|w^* R_1 B\| = \|(R_1^* w)^* B\| > \|R_1^* w\| \, b.$$

By the construction of R_1, as R is invertible, we see that all singular values of R_1 are nonzero. Therefore, $\|R_1^* w\| \, b > 0$. Correspondingly, $(N_1, R_1 B)$ is controllable.

We then consider the system under τ-hop control. Since w is the left eigenvector of N_1, it is also the left eigenvector of N_1^τ. In particular, $w^* N^{\tau-1}$ is a left eigenvector of N_1. Since N_1 is the expanding portion of A, we derive the following lower bound:

$$\left\| w^* \left(N_1^{\tau-1} R_1 B \right) \right\| = \left\| \left(w^* N_1^{\tau-1} \right) R_1 B \right\| \geq \lambda_k^{\tau-1} \|R_1^* w\| \, b.$$

Recall that $B_\tau = P_1^* A^{\tau-1} B$.

$$\begin{aligned} B_\tau &= P_1^* \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} N_1^{\tau-1} & \\ & N_2^{\tau-1} \end{bmatrix} \begin{bmatrix} R_1 B \\ R_2 B \end{bmatrix} \\ &= \begin{bmatrix} P_1^* Q_1 & P_1^* Q_2 \end{bmatrix} \begin{bmatrix} N_1^{\tau-1} R_1 B \\ N_2^{\tau-1} R_2 B \end{bmatrix} \\ &= N_1^{\tau-1} R_1 B + P_1^* Q_2 N_2^{\tau-1} R_2 B. \end{aligned}$$

By Gelfand's Formula, $\left\| N_2^{\tau-1} \right\| \leq \zeta_{\epsilon_4}(N_2) \left(\lambda_{k+1} + \epsilon_4 \right)^{\tau-1}$. Moreover, since $E_u^\perp$ and E_s are ξ-close, by Lemma A.1 of Hu et al. [2022], $P_1^* Q_2 \leq \sqrt{2\xi}$.

Therefore, we know that

$$\begin{aligned} \|w^* B\| = &\left\| w^* \left(N_1^{\tau-1} R_1 B + P_1^* Q_2 N_2^{\tau-1} R_2 B \right) \right\| \\ \geq &|\lambda_k|^{\tau-1} \|R_1^* w\| \, b \\ &- \sqrt{2\xi} \, \|Q_2\| \, \|R_2\| \, \|B\| \, \zeta_{\epsilon_4}(N_2) \left(|\lambda_{k+1}| + \epsilon_4 \right)^{\tau-1} \\ > &\frac{1}{2} \|R_1^* w\| \, b, \end{aligned}$$

where the last inequality requires $\epsilon_4 < 1 - \lambda_{k+1}$, and

$$\tau \geq \frac{\log \frac{\|R_1^* w\| b}{2\sqrt{2\xi} \|Q_2\| \|R_2\| \|B\| \zeta_{\epsilon_4}(N_2)}}{\log \frac{|\lambda_k|}{|\lambda_{k+1}| + \epsilon_4}}. \tag{52}$$

Therefore, we conclude $(M_1^\tau, R_1 B_\tau)$ is also controllable, as $M_1 = N_1$.

Lastly, we prove $\left(\hat{M}_1^\tau, R_1 \hat{B}_\tau \right)$ is stabilizable. Denote $\mathcal{A} := M_1^\tau - R_1 B_\tau K_1$. Since $(M_1^\tau, R_1 B_\tau)$ is controllable, we know there exists K_1 such that $\rho(\mathcal{A}) < 1$. Since an asymptotically stable linear system is also exponentially stable, by the Lyapunov equation, for every $k \times k$ matrix $G > 0$, the following discrete Lyapunov equation has a unique solution $H = H^* > 0$.

$$\mathcal{A}^* H \mathcal{A} + G - H = 0$$

In particular, we pick G such that $\sigma_{\min}(G) > 2$ and $W(v) := \frac{1}{\min\{1, \sigma_{\min}(H)\}} v^* H v$ is a Lyapunov function of $\mathcal{A}$. Moreover, $W(v)$ satisfies the following criteria regarding $\|v\|$ and forward difference with respect to $\mathcal{A}$:

$$\|v\|^2 \leq W(v) \leq \kappa(H) \|v\|^2,$$

$$\begin{aligned} W(\mathcal{A}v) - W(v) &= \frac{v^* \mathcal{A}^* H \mathcal{A} v - v^* H v}{\min\{1, \sigma_{\min}(H)\}} \\ &\leq -v^* G v \\ &< -2 \|v\|^2, \end{aligned}$$

where $\kappa(H)$ is the condition number of H.

We now consider the forward difference with respect to $\hat{\mathcal{A}} = \hat{M}_1^\tau - R_1\hat{B}_\tau K_1$, as a consequence of Jensen's inequality, for any $\iota > 0$,

$$\begin{aligned} W\left(\hat{\mathcal{A}}v\right) =& W\left(\mathcal{A}v + \left(\hat{\mathcal{A}} - \mathcal{A}\right)v\right) \\ \leq& (1+\iota^2)W(\mathcal{A}v) + \left(1 + \frac{1}{\iota^2}\right) W\left(\left(\hat{\mathcal{A}} - \mathcal{A}\right)v\right), \end{aligned}$$

and

$$\begin{aligned} & W\left(\hat{\mathcal{A}}v\right) - W(v) \\ =& W(\mathcal{A}v) - W(v) + W\left(\hat{\mathcal{A}}v\right) - W(\mathcal{A}v) \\ \leq& W(\mathcal{A}v) - W(v) + \iota^2 W(\mathcal{A}v) + \left(1 + \frac{1}{\iota^2}\right) W\left(\left(\hat{\mathcal{A}} - \mathcal{A}\right)v\right) \\ <& -2\left\|v\right\|^2 + \iota^2\kappa(H)\left\|v\right\|^2 + \left(1 + \frac{1}{\iota^2}\right)\left\|\hat{\mathcal{A}} - \mathcal{A}\right\|^2 \left\|v\right\|^2 \\ \leq& -\left\|v\right\|^2, \end{aligned}$$

The last inequality requires

$$\iota^2 < \frac{1}{2\kappa(H)}, \qquad \left\|\hat{\mathcal{A}} - \mathcal{A}\right\|^2 < \frac{1}{2}\frac{\iota^2}{1+\iota^2}.$$

By Proposition 5.2 and E.5, we get

$$\begin{aligned} \left\|\hat{M}_1^\tau - M_1^\tau\right\| &< 3\tau\left\|A\right\|\zeta_{\epsilon_1}(A)^2\left(|\lambda_1| + \epsilon_1\right)^{\tau-1}\delta, \\ \left\|\hat{B}_\tau - B_\tau\right\| &< C_B(|\lambda_1| + \epsilon_1)^{\tau-1}\delta. \end{aligned}$$

So we require

$$\delta < \frac{\frac{1}{6}\frac{\iota^2}{1+\iota^2}}{\tau\left\|A\right\|\zeta_{\epsilon_1}(A)^2\left(|\lambda_1| + \epsilon_1\right)^{\tau-1} + \left\|K_1\right\| C_B(|\lambda_1| + \epsilon_1)^{\tau-1}}. \tag{53}$$

When all requirements above are satisfied, by Theorem 2 of Jiang and Wang [2002], we conclude $\left(\hat{M}_1^\tau, R_1\hat{B}_\tau\right)$ is stabilizable. □

As the control matrix $\hat{K}_1$ is obtained by the learner, we denote constant $\mathcal{K}$ such that $\left\|\hat{K}_1\right\| < \mathcal{K}$ to be a user-defined constant.

After the proof of the stabilizability of the system after transformation, we are now ready to prove the main theorem.

Proof: [proof of Theorem 4.2] We shall bound each of the four terms in $\hat{L}$ defined in (9). We first guarantee that the diagonal blocks are stable. For the top-left block, by Proposition F.1, there exists positive-definite matrix $\bar{U}$ such that $\left\|\hat{M}_1^\tau - \hat{B}_\tau\hat{K}_1\right\|_{\bar{U}} = \mathcal{U} < 1$, where $\|\cdot\|_{\bar{U}}$ denotes the weighted norm induced by $\bar{U}$. Therefore,

$$\begin{aligned} \rho(\hat{L}_{1,1}) \leq& \left\|M_1^\tau + P_1^* A^{\tau-1}B\hat{K}_1\hat{P}_1^* P_1\right\|_{\bar{U}} && (54) \\ \leq& \left\|M_1^\tau - \hat{M}_1^\tau\right\|_{\bar{U}} + \left\|\hat{M}_1^\tau - \hat{B}_\tau\hat{K}_1\right\|_{\bar{U}} + \left\|(B_\tau - \hat{B}_\tau)\hat{K}_1\right\|_{\bar{U}} + \left\|B_\tau\hat{K}_1(I - \hat{P}_1^* P_1)\right\|_{\bar{U}} \\ \leq& \kappa(\bar{U})^{\frac{1}{2}}\left(\left\|M_1^\tau - \hat{M}_1^\tau\right\| + \left\|B_\tau - \hat{B}_\tau\right\|\left\|\hat{K}_1\right\| + \left\|B_\tau\right\|\left\|\hat{K}_1\right\|\left\|I - \hat{P}_1^* P_1\right\|\right) + \mathcal{U} \\ \leq& 3\kappa(\bar{U})^{\frac{1}{2}}\tau\left\|A\right\|\zeta_{\epsilon_1}(A)^2(|\lambda_1| + \epsilon_1)^{\tau-1}\delta + \kappa(\bar{U})^{\frac{1}{2}}C_B\mathcal{K}(|\lambda_1| + \epsilon_1)^{\tau-1}\delta \\ &+ \kappa(\bar{U})^{\frac{1}{2}}\zeta_{\epsilon_1}(A)(|\lambda_1| + \epsilon_1)^{\tau-1}\left\|B\right\|\mathcal{K}\delta + \mathcal{U} && (55) \\ <& \kappa(\bar{U})^{\frac{1}{2}}(C_B\mathcal{K} + \zeta_{\epsilon_1}(A)\left\|B\right\|\mathcal{K} + 1)(\lambda_1| + \epsilon)^{\tau-1}\delta + \mathcal{U} && (56) \\ <& \frac{1}{2} + \frac{\mathcal{U}}{2}, && (57) \end{aligned}$$

where in (55) we apply proposition E.1 of Hu et al. [2022] and Proposition 5.2 and Proposition E.5; In (56), we require

$$\frac{1}{\tau}(|\lambda_1|+\epsilon_1)^{\tau-1} > 3\,\|A\|\,\zeta_{\epsilon_1}(A)^2. \tag{58}$$

In (57), we require

$$\delta < \frac{(1-\mathcal{U})(\lambda_1|+\epsilon)^{-(\tau-1)}}{2\kappa(\bar{U})^{\frac{1}{2}}(C_B\mathcal{K}+\zeta_{\epsilon_1}(A)\,\|B\|\,\mathcal{K}+1)}. \tag{59}$$

For the bottom-right block, it is straightforward to see that

$$\begin{aligned}\rho(\hat{L}_{2,2}) \leq& \|M_2^\tau\| + \left\|P_2^* A^{\tau-1}\right\| \|B\| \left\|\hat{K}_1\right\| \left\|\hat{P}_1^* P_2\right\| \\ \leq& \zeta_{\epsilon_2}(M_2)(|\lambda_{k+1}|+\epsilon_2)^\tau + \zeta_{\epsilon_2}(M_2)\,\|B\|\,\mathcal{K}(|\lambda_{k+1}|+\epsilon_2)^{\tau-1}\delta \\ <& \frac{1}{2},\end{aligned}$$

where the last inequality requires

$$\tau > \frac{\log 1/(4\zeta_{\epsilon_2}(M_2))}{\log(|\lambda_{k+1}|+\epsilon_2)}, \tag{60}$$

$$\delta < \frac{1}{4\zeta_{\epsilon_2}(M_2)\,\|B\|\,\mathcal{K}}(|\lambda_{k+1}|+\epsilon_2)^{-(\tau-1)}. \tag{61}$$

Now it suffices to bound the spectral norms of off-diagonal blocks. Note that, by applying Proposition G.1 of Hu et al. [2022], the top right block is bounded as

$$\begin{aligned}\rho(\hat{L}_{2,1}) \leq& \|\Delta_\tau\| + \|B_\tau\| \left\|\hat{K}_1\right\| \left\|\hat{P}_1^* P_2\right\| \\ <& C_\Delta(|\lambda_1|+\epsilon_1)^\tau + \zeta_{\epsilon_1}(A)\,\|B\|\,\mathcal{K}(|\lambda_1|+\epsilon_1)^{\tau-1}\delta \\ <& (C_\Delta+1)(|\lambda_1|+\epsilon_1)^\tau,\end{aligned}$$

where the last inequality requires

$$\delta < \frac{1}{\zeta_{\epsilon_1}(A)\,\|B\|\,\mathcal{K}}(|\lambda_1|+\epsilon_1)^{-(\tau-1)}. \tag{62}$$

The bottom-left block is bounded as

$$\begin{aligned}\rho(\hat{L}_{1,2}) \leq& \left\|P_2^* A^{\tau-1}\right\| \|B\| \left\|\hat{K}_1\right\| \\ <& \zeta_{\epsilon_2}(M_2)\,\|B\|\,\mathcal{K}(|\lambda_{k+1}|+\epsilon_2)^{\tau-1}.\end{aligned}$$

By Lemma 5.3 of Hu et al. [2022], we can guarantee that

$$\rho\left(\hat{L}_\tau\right) \leq \frac{1}{2} + \frac{\mathcal{U}}{2} + \chi\left(\hat{L}_\tau\right)\frac{(C_\Delta+1)\zeta_{\epsilon_2}(M_2)\,\|B\|\,\mathcal{K}}{|\lambda_1|+\epsilon_1}\left((|\lambda_1|+\epsilon_1)(|\lambda_{k+1}|+\epsilon_2)\right)^{\tau-1} < 1, \tag{63}$$

which requires

$$\tau > \frac{\log\frac{(1-\mathcal{U})(|\lambda_1|+\epsilon_1)(|\lambda_{k+1}|+\epsilon_2)}{2\chi(\hat{L}_\tau)(C_\Delta+1)\zeta_{\epsilon_2}(M_2)\|B\|\mathcal{K}}}{\log((|\lambda_1|+\epsilon_1)(|\lambda_{k+1}|+\epsilon_2))}. \tag{64}$$

Note that the above constraints make sense only if $|\lambda_1||\lambda_{k+1}| < 1$. Therefore, when all constraints above are satisfied, system (8) is ultimately bounded, and so is system (1).

We will then collect all the constraints. Combining (58) (60) and (64), we obtain

$$\begin{aligned}\tau > \max\Bigg\{&\frac{\log 1/(4\zeta_{\epsilon_2}(M_2))}{\log(|\lambda_{k+1}|+\epsilon_2)}, \frac{\log\frac{(\mathcal{U}+1)(|\lambda_1|+\epsilon_1)(|\lambda_{k+1}|+\epsilon_2)}{2\chi(\hat{L}_\tau)(C_\Delta+1)\zeta_{\epsilon_2}(M_2)\|B\|\mathcal{K}}}{\log((|\lambda_1|+\epsilon_1)(|\lambda_{k+1}|+\epsilon_2))}, \\ &-\frac{1}{\log(|\lambda_1|+\epsilon_1)}W_{-1}\left(-\frac{\log(|\lambda_1|+\epsilon_1)}{3\,\|A\|\,\zeta_{\epsilon_1}(A)^2(|\lambda_1|+\epsilon_1)}\right)\Bigg\},\end{aligned}$$

where W_{-1} denotes the non-principle branch of the Lambert-W function. Here we utilize the fact that, for $x > \frac{1}{\log a}$, $y = \frac{a^*}{x}$ is monotone increasing with inverse function $x = -\frac{1}{\log a} W_{-1}\left(-\frac{\log a}{y}\right)$, which can be upper bounded by Theorem 1 in Chatzigeorgiou [2013] as

$$\begin{aligned} \tau &> \frac{\log \frac{\sqrt{\xi}}{1-\xi} + \log \frac{1}{c} + \log \chi\left(\hat{L}_\tau\right) + 5\log \bar{\zeta} + \log \frac{\|A\|}{|\lambda_1| - |\lambda_{k+1}|} + C_\tau}{\log |\lambda_1|} \\ &= O(1), \end{aligned} \tag{65}$$

where $\bar{\zeta} := \max\left\{\zeta_{\epsilon_1}(A), \zeta_{\epsilon_2}(M_2), \zeta_{\epsilon_2}(N_2), \zeta_{\epsilon_3}(N_1^{-1})\right\}$, and C_τ is a numerical constant.

We then collect all the bound on γ, α, δ as follows:

$$\gamma > \epsilon, \tag{66}$$

$$\alpha < \frac{1}{\|B\|} = O(1). \tag{67}$$

Combining (53), (59), (61), (62) yields the following bound on δ:

$$\begin{aligned} \delta < \max\Bigg\{ & \frac{\frac{1}{6}\frac{\iota^2}{1+\iota^2}}{\tau \|A\| \zeta_{\epsilon_1}(A)^2 (|\lambda_1| + \epsilon_1)^{\tau-1} + \|K_1\| C_B (|\lambda_1| + \epsilon_1)^{\tau-1}}, \frac{(1-\mathcal{U})(\lambda_1| + \epsilon)^{-(\tau-1)}}{2\kappa(\bar{U})^{\frac{1}{2}}(C_B\mathcal{K} + \zeta_{\epsilon_1}(A)\|B\|\mathcal{K} + 1)}, \\ & \frac{1}{4\zeta_{\epsilon_2}(M_2)\|B\|\mathcal{K}}(|\lambda_{k+1}| + \epsilon_2)^{-(\tau-1)}, \frac{1}{\zeta_{\epsilon_1}(A)\|B\|\mathcal{K}}(|\lambda_1| + \epsilon_1)^{-(\tau-1)} \Bigg\}. \end{aligned}$$

which can be simplified to

$$\delta < \frac{C_\delta}{\sqrt{m}\bar{\zeta}^3(\|A\| + \|B\|)}|\lambda_1|^{-2\tau} = O(m^{-1/2}|\lambda_1|^{-2\tau}), \tag{68}$$

where C_δ is a constant collecting minor factors. Recall that $\delta = \sqrt{2k}\epsilon$. Substitute the above in (17) transfers the bound on δ into a bound on T:

$$T > \frac{2\log\left(\frac{8k^{\frac{k}{2}+4}(n-k)\left(\frac{C}{1-|\lambda_{k+1}|}\right)\left(\frac{\sqrt{m}\bar{\zeta}^3(\|A\|+\|B\|)}{C_\delta|\lambda_1|^{-2\tau}}\right)}{\sqrt{\pi}\theta\text{gap}\epsilon}\right)}{\log|\lambda_k|} = O\left(k\log k + \log(n-k) + \log m - \log\text{gap}\right) \tag{69}$$

Different from Hu et al. [2022], we do not explicitly choose ω but let $(\omega_i)_{i\in\{1,\ldots,m\}}$ be the stopping time defined in Proposition E.4.

Combining the above constant with Theorem 5.1, we conclude that Algorithm 1 controls x with the following bound:

$$\begin{aligned} \|x\| &\leq \exp\left(O\left(T + \sum_{i=1}^m \omega_i + \tau m\right)\right) \\ &\leq \exp\Bigg(O\Bigg(\frac{1}{\log|\lambda_k|}\Bigg(-\log\text{gap} + k\log k - \log\theta + \log(n-k) \\ &\quad + \log|\lambda_1| + \log C - \log(1-|\lambda_{k+1}|) + (1+\log|\lambda_1|)m\Bigg)\Bigg) \end{aligned}$$

Assuming that the eigenvalue-related terms are constants, the algorithm achieves $\exp(O(k\log k + \log(n-k) + m - \log\text{gap}))$ space complexity for $\|x\|$.

This finishes the proof of Theorem 4.2. □

G ADDITIONAL MATHEMATICAL BACKGROUND

In this section, we introduce some relevant math background used in this paper. The notation of this section is independent of the rest of the paper.

Theorem G.1 (Davis-Kahan)**.** *Let A be an $n \times n$ Hermitian matrix, and suppose we have the following spectral decomposition for A*

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^*,$$

where λ_i's are the eigenvalues of A such that $\lambda_1 > \cdots > \lambda_n$, and u_i's are corresponding eigenvectors. Let H be another $n \times n$ perturbation matrix, and the spectral decomposition of $A + H$ is

$$A + H = \sum_{i=1}^{n} \mu_i v_i v_i^*.$$

Define

$$P = \sum_{i=1}^{k} u_i u_i^* := UU^*$$

to be the orthogonal projection operator to the k-dimensional eigenspace spanned by $u_1 \ldots, u_k$. Similarly, define $Q = \sum_{i=1}^{k} v_i v_i^* := VV^*$.

Suppose there exists $\delta > 0$, such that $|\lambda_i - \mu_j| > \delta$ for all $i \in \{1, \ldots, k\}, j \in \{k+1, \ldots, n\}$, then the operator norm of $\|P - Q\|_{op}$ satisfy

$$\|P - Q\|_{op} \leq \|P - Q\|_F \leq \frac{\sqrt{2k}\,\|H\|_{op}}{\delta},$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

This is a relatively common theorem, and the proof detail can be found at, for instance, Cao [2021].

Lemma G.2 (Gelfand's formula)**.** *For any square matrix X, we have*

$$\rho(X) = \lim_{t \to \infty} \left\|X^t\right\|^{1/t}.$$

In other words, for any $\epsilon > 0$, there exists a constant $\zeta_\epsilon(X)$ such that

$$\sigma_{\max}(X^t) = \|X\| \leq \zeta_\epsilon(X)(\rho(X) + \epsilon)^t.$$

Further, if X is invertible, let $\lambda_{\min}(X)$ denote the eigenvalue of X with minimum modulus, then

$$\sigma_{\min}(X^t) \geq \frac{1}{\zeta_\epsilon(X^{-1})} \left(\frac{|\lambda_{\min}(X)|}{1 + \epsilon|\lambda_{\min}(X)|}\right)^t.$$

The proof can be found in existing literatures (e.g. Horn and Johnson [2012].

H INDEXING

For the convenience of readers, we provide a table summarizing all constants appearing in the bounds.

Table 1: Lists of parameters and constants appearing in the bound.

Constant	Appearance	Explanation
T	Stage 1	T initialization steps to separate unstable components.
ω_i	Stage 3	Stopping time in each iteration to learn B_τ.
α	Stage 3	$u_{t_i} = \alpha \left\| x_{t_i} \right\| e_i$ to estimate columns of B_τ.
τ	Stage 3	τ-steps between consecutive control inputs are injected.

Table 2: System parameters.

Constant	Appearance	Explanation
C	Section 2	Upper bound the magnitude of noise.
λ_i	Section 3.1	(Complex) eigenvalue of A with i-th largest modulus.
ξ	Definition 3.1 of Hu et al. [2022]	$E_u^\perp$ and E_s are ξ-close subspaces, i.e. $\sigma_{\min} P_2^* Q_1 > 1 - \xi$.
$\zeta_\epsilon(\cdot)$	Lemma G.2	Gelfand constant for the norm of matrix exponents

Table 3: Shorthand notations (introduced in proofs).

Constant	Appearance	Explanation
C_Δ	Proposition G.1 of Hu et al. [2022]	$C_\Delta := \zeta_{\epsilon_1}(M_1)\zeta_{\epsilon_2}(M_2)\frac{(2-\xi)\sqrt{2\xi}\|A\|}{1-\xi}\frac{2\lvert\lambda_{k+1}\rvert}{\lvert\lambda_1+\epsilon_1-\lvert\lambda_{k+1}\rvert-\epsilon_2}$.
C_γ	(42) in the proof of Proposition E.2	$C_\gamma := \frac{\zeta_{\epsilon_4}(N_2)C}{\gamma'(1-\xi)}\frac{1}{1-(\lvert\lambda_{k+1}\rvert+\epsilon_4)}$.
C_B	Proposition E.5	$(\zeta_{\epsilon_1}^2(A)(\|A\| + \|B\| + (C+2)\tau + 1) + (\tau+1)C_\Delta)\frac{\sqrt{m}}{\alpha}$.
$\mathcal{K}$	Stage 4	Upper bounding $\left\|\hat{K}_1\right\|$ chosen by the user.
$\mathcal{U}$	Stage 4	Upper bounding $\left\|\hat{M}_1^\tau - \hat{B}_\tau \hat{K}_1\right\|_{\bar{U}}$.
gap	Theorem 4.2	gap $:= \left\lvert\prod_{m_1 \neq m_2}(\lambda_{m_1}^{-1} - \lambda_{m_2}^{-1})\right\rvert, m_1, m_2 \in \{1, \ldots, k\}$.