

# Finite Time Adaptive Stabilization of LQ Systems

Mohamad Kazem Shirani Faradonbeh, Ambuj Tewari, and George Michailidis

**Abstract**—Stabilization of linear systems with unknown dynamics is a canonical problem in adaptive control. Since the lack of knowledge of system parameters can cause it to become destabilized, an adaptive stabilization procedure is needed prior to regulation. Therefore, the adaptive stabilization needs to be completed in finite time. In order to achieve this goal, asymptotic approaches are not very helpful. There are only a few existing non-asymptotic results and a full treatment of the problem is not currently available.

In this work, leveraging the novel method of *random* linear feedbacks, we establish high probability guarantees for finite time stabilization. Our results hold for remarkably general settings because we carefully choose a minimal set of assumptions. These include stabilizability of the underlying system and restricting the degree of heaviness of the noise distribution. To derive our results, we also introduce a number of new concepts and technical tools to address *regularity* and instability of the closed-loop matrix.

**Index Terms**—Random Feedbacks, Unstable Estimation, Fast Stabilization, Finite Time Identification, Closed-loop Regularity.

## I. INTRODUCTION

WE consider finite time stabilization of the following linear system. Given the initial state  $x(0) \in \mathbb{R}^p$ , for  $t = 0, 1, \dots$ , we have

$$x(t+1) = A_0x(t) + B_0u(t) + w(t+1), \quad (1)$$

where at time  $t$ , the vector  $x(t) \in \mathbb{R}^p$  corresponds to the state (and output) of the system,  $u(t) \in \mathbb{R}^r$  is the control action, and  $\{w(t)\}_{t=1}^{\infty}$  is a sequence of noise (i.e., random disturbance) vectors. The dynamics of the system, i.e., both the transition matrix  $A_0 \in \mathbb{R}^{p \times p}$ , as well as the input matrix  $B_0 \in \mathbb{R}^{p \times r}$ , are fixed but *unknown*.

In order to stabilize the system, we need to design an adaptive procedure and establish finite time theoretical guarantees of stabilization. Once the system has been stabilized, we can use an adaptive regulation policy to minimize the cost function determined by the application. Hence, the stabilization procedure needs to be completed in a relatively short time period. There is an extensive literature providing *infinite* time analyses to adaptively stabilize a Linear-Quadratic (LQ) system [1], [2], [3], [4], [5], whereas finite time results are scarce and rather incomplete. This work aims to contribute to the limited literature on this topic.

M.K.S.F. acknowledges partial support by the UF Informatics Institute. The work of G.M. was supported in part by the National Science Foundation under grants IIS-1632730, DMS-1821220, and CCF-1442493. A.T. acknowledges the support of NSF CAREER grant IIS-1452099 and of a Sloan Research Fellowship.

M.K. Shirani Faradonbeh and G. Michailidis are with the Department of Statistics, University of Florida, Gainesville, FL, 32611-5585 USA (e-mail: mfaradonbeh@ufl.edu, gmichail@ufl.edu)

Ambuj Tewari is with the Department of Statistics, University of Michigan, Ann Arbor, MI 48109-1107 USA (e-mail: tewaria@umich.edu)

The evolution of LQ systems is governed by linear dynamics, while the operating cost is a quadratic function of the state and the control signal. To deal with the uncertainty about the true matrices guiding the system's dynamics, a standard scheme is Certainty Equivalence (CE) [6]. Its prescription is to assume that the *estimated* parameters coincide with the true dynamics matrices. However, it is shown that inconsistency occurs with positive probability [7], [8], [9], which can lead to instability. This motivated modifying CE to Optimism in the Face of Uncertainty (OFU) [10], [11], [12], [13]. OFU prescribes to act as if an optimistic approximation of the true parameter is the one guiding the evolution of the system.

Recent *finite time* analyses consider a restricted setting [14], [15], where the proposed adaptive stabilization procedure heavily relies on the following strong conditions. First, *controllability* and *observability* of the true dynamics matrices of the system are assumed. Second, the closed-loop transition matrix is required to have *operator norm* less than one. Note that the former does not imply the latter [16]. Third, uncertainty about the true dynamics matrices is restricted to an a priori *known* bounded box in the space of  $p \times (p+r)$  matrices. Finally, the noise vectors are supposed to have a sub-Gaussian distribution with uncorrelated coordinates.

We introduce our stabilization algorithm and establish finite time guarantees for it in Section IV. Leveraging the novel method of *random* linear feedbacks, we address the four aforementioned limitations. The first assumption (which imposes a computationally intractable constraint [17]), as well as the third one (which requires possibly unavailable information) are not needed. Further, we relax the operator norm condition to the minimal assumption of stabilizability. Finally, the noise process is generalized to the much larger class of sub-Weibull distributions which includes some heavy-tailed distributions and allows coordinates to be correlated. Note that unlike the operator norm condition, stability of matrices is not preserved by multiplication (i.e., the product of stable matrices can be unstable). This means that existing theoretical techniques [13], [14] used in addressing the stabilization problem fail to work when the operator norm is not less than one.

To derive finite time guarantees of stabilization, new concepts and technical tools are needed to address the following issues:

- 1) Because of the unbounded growth of the state vectors [18], the classical results of persistent excitation [19] are not applicable.
- 2) Since the system is not fully stabilized yet, the closed-loop matrix can have eigenvalues both inside and outside the unit circle. Thus, the smallest (largest) eigenvalue of the Gram matrix scales linearly (exponentially) with time [20], [21]. This leads to the failure of the existing

approaches which do not need the persistent excitation condition [14], [22].

- 3) For unstable systems, it is shown that the normalized empirical covariance of the state vector is a random matrix [23], [24]. So, in order to obtain reliable identification results, anti-concentration properties of random matrices need to be carefully examined [25].
- 4) For accurate identification, one needs to ensure that the important condition of closed-loop *regularity* holds (see Definition 2). It is a necessary condition on the eigenvalues of magnitude larger than one [26].

The remainder of the paper is organized as follows. The problem is rigorously formulated in Section II. Then, in Section III we study the key identification results for unstable closed-loop dynamics as the cornerstone of the stabilization algorithm presented later on. Subsequently in Section IV, results regarding the properties of random linear feedback are established. Finally, we propose the adaptive stabilization **Algorithm 1**, and show that it is guaranteed to return a high probability stabilizing set.

#### A. Notation

The following notation is used throughout this paper. For matrix  $A \in \mathbb{C}^{p \times q}$ ,  $A'$  is its transpose. When  $p = q$ , the smallest (respectively largest) eigenvalue of  $A$  (in magnitude) is denoted by  $\lambda_{\min}(A)$  (respectively  $\lambda_{\max}(A)$ ) and the trace of  $A$  is denoted by  $\text{tr}(A)$ . For  $v \in \mathbb{C}^q$ , the norm of vector  $v$  is  $\|v\| = \left( \sum_{i=1}^q |v_i|^2 \right)^{1/2}$ . We also use the following notation for the operator norm of matrices:  $\|A\| = \sup_{\|v\|=1} \|Av\|$ . To denote the dimension of manifold  $\mathcal{M}$  over the field  $F$ , we use  $\dim_F(\mathcal{M})$ . The definitions of the symbols  $\theta, K(\theta), L(\theta)$ , and  $\tilde{L}(\theta)$  can be found in Remark 1, equations (3), (4), and Remark 2, respectively.

## II. PROBLEM FORMULATION

We start by discussing the adaptive stabilization problem that constitutes the primary focus of this work. As mentioned above, once the system has stabilized, an adaptive policy for regulating the system (i.e., cost minimization) can be employed. Results of this work can be used in the finite time analysis of adaptive regulation for LQ systems. Further, stabilization of linear systems is intimately related to a Riccati equation for the corresponding LQ system. Therefore, we comprehensively discuss the necessary preliminaries here.

The stochastic evolution of the system is governed by the linear dynamics (1), where  $\{w(t)\}_{t=1}^{\infty}$  are independent mean-zero noise vectors with full rank covariance matrix  $C$ :

$$\mathbb{E}[w(t)] = 0, \quad \mathbb{E}[w(t)w(t)'] = C, \quad |\lambda_{\min}(C)| > 0.$$

Generalizations of the established results to dependent noise vectors (i.e., martingale difference sequences) is rather straightforward. The true dynamics matrices  $A_0, B_0$  are assumed to be stabilizable, as defined below.

**Definition 1** (Stabilizability [16]).  $[A, B]$  is stabilizable if there exists  $L \in \mathbb{R}^{r \times p}$  such that  $|\lambda_{\max}(A + BL)| < 1$ . The linear feedback matrix  $L$  is called a stabilizer for  $[A, B]$ .

**Remark 1.** For notational convenience, henceforth for  $A \in \mathbb{R}^{p \times p}$ ,  $B \in \mathbb{R}^{p \times r}$ , we use  $\theta$  to denote  $[A, B]$ . Clearly,  $\theta \in \mathbb{R}^{p \times q}$ , where  $q = p + r$ .

We assume *perfect observations*, i.e., the operator can fully observe the sequence of state vectors. Next, suppose that  $c_t$  is the quadratic instantaneous cost function at time  $t$ :

$$c_t = x(t)'Qx(t) + u(t)'Ru(t), \quad (2)$$

which is defined according to the known positive definite cost matrices  $Q \in \mathbb{R}^{p \times p}, R \in \mathbb{R}^{r \times r}$ . An adaptive policy is a mapping which designs the control action according to the cost matrices, and the history of the system. That is, for all  $t = 0, 1, \dots$ , the operator needs to determine  $u(t)$  according to  $Q, R, \{x(i)\}_{i=0}^t, \{u(j)\}_{j=0}^{t-1}$ .

The following proposition shows that in order to stabilize a linear system, one can solve a *Riccati* equation. A solution, is a positive semidefinite matrix  $K(\theta)$  satisfying (3). For this purpose, we introduce a notation that simplifies certain expressions throughout this work.

**Remark 2.** For arbitrary stabilizable  $\theta_1, \theta_2 \in \mathbb{R}^{p \times q}$ , let

$$\tilde{L}(\theta_1) = \begin{bmatrix} I_p \\ L(\theta_1) \end{bmatrix} \in \mathbb{R}^{q \times p}.$$

So,  $\theta_2 \tilde{L}(\theta_1) = A_2 + B_2 L(\theta_1)$ .

**Proposition 1.** If  $\theta$  is stabilizable, (3) has a unique solution. Conversely, if (3) has a solution,  $L(\theta)$  defined by (4) is a stabilizer for the dynamics parameter  $\theta$ ; i.e.  $|\lambda_{\max}(\theta \tilde{L}(\theta))| < 1$ .

The proof of Proposition 1 is provided in Appendix A, where the following cost minimization property of Riccati equations (3), (4) is established as well. Assuming the system evolves according to (1), the linear feedback  $u(t) = L(\theta_0)x(t)$  minimizes the expected average cost of the system of dynamics parameter  $\theta_0$ . Namely, letting  $c_t$  be as (2), in general it holds that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \mathbb{E}[c_t] \geq \text{tr}(K(\theta_0)C),$$

where the linear feedback  $u(t) = L(\theta_0)x(t)$  attains the equality. An adaptive stabilization procedure is ignorant about the true parameter  $\theta_0$ , and needs to estimate it. The following lemma addresses the stability when the actual system evolution parameter is  $\theta_0$  but the linear feedback  $L(\theta)$  is designed according to the approximation  $\theta$ . The proof of Lemma 1 can be found in Appendix B.

**Lemma 1** (Stabilizing neighborhood). *There exists  $\epsilon_0 > 0$ , such that for every stabilizable  $\theta$ , if  $\|\theta - \theta_0\| < \epsilon_0$ , then,  $\theta_0 \tilde{L}(\theta)$  is stable.*

## III. CLOSED-LOOP IDENTIFICATION

When applying linear feedback  $L \in \mathbb{R}^{r \times p}$ , the dynamics take the form  $x(t+1) = Dx(t) + w(t+1)$ , where  $D =$

$$K(\theta) = Q + A'K(\theta)A - A'K(\theta)B(B'K(\theta)B + R)^{-1}B'K(\theta)A, \quad (3)$$

$$L(\theta) = -(B'K(\theta)B + R)^{-1}B'K(\theta)A. \quad (4)$$

$A_0 + B_0L$  is the *unstable* closed-loop transition matrix. Subsequently, we present results for the accurate identification of  $D$  through the least-squares estimator. Observing the state vectors  $\{x(t)\}_{t=0}^n$ , for an arbitrary matrix  $E \in \mathbb{R}^{p \times p}$  define the sum-of-squares loss function  $\mathcal{L}_n(E) = \sum_{t=0}^{n-1} \|x(t+1) - Ex(t)\|^2$ . Then, the true closed-loop transition matrix  $D$  is estimated by  $\hat{D}_n$ , which is a minimizer of the loss function;  $\mathcal{L}_n(\hat{D}_n) = \min_{E \in \mathbb{R}^{p \times p}} \mathcal{L}_n(E)$ . To analyze the finite time behavior of the aforementioned identification procedure, the following is assumed for the tail-behavior of every coordinate of the noise vector.

**Assumption 1** (Sub-Weibull distribution [21]). *There are positive reals  $b_1, b_2$ , and  $\alpha$ , such that for all  $t \geq 1; 1 \leq i \leq p; y > 0$ ,*

$$\mathbb{P}(|w_i(t)| > y) \leq b_1 \exp\left(-\frac{y^\alpha}{b_2}\right).$$

Intuitively, smaller values of the exponent  $\alpha$  correspond to heavier tails for the noise distribution, and vice versa. Note that whenever  $\alpha < 1$ , the noise coordinates  $w_i(t)$  do not need to have a moment generating function. Further, the noise coordinates can be either discrete or continuous random variables, and are not assumed to have a probability density function (pdf). Henceforth, the special case of bounded noise can be obtained from the presented results letting  $\alpha \rightarrow \infty$ .

Next, we define an important property of unstable transition matrices which is required in order to obtain accurate estimation results.

**Definition 2** (Regularity [26]).  *$D \in \mathbb{R}^{p \times p}$  is regular if for any eigenvalue  $\lambda$  of  $D$  such that  $|\lambda| > 1$ , the geometric multiplicity of  $\lambda$  is one.*

Regularity implies that the eigenspace corresponding to  $\lambda$  is one dimensional, and vice versa. There are other equivalent formulations for regularity. Indeed,  $D$  is regular if and only if for any eigenvalue  $\lambda$  such that  $|\lambda| > 1$ , in the Jordan decomposition of  $D$  there is only one block corresponding to  $\lambda$ , regardless of its algebraic multiplicity. Another equivalent formulation is that  $D$  is regular, if and only if  $\text{rank}(D - \lambda I_p) \geq p - 1$ , for all  $\lambda \in \mathbb{C}, |\lambda| > 1$ . For example, let  $P_1, P_2 \in \mathbb{C}^{2 \times 2}$  be arbitrary invertible matrices, and assume

$$D_1 = P_1^{-1} \begin{bmatrix} \rho & 1 \\ 0 & \rho \end{bmatrix} P_1, D_2 = P_2^{-1} \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} P_2,$$

are real  $2 \times 2$  matrices, where  $\rho \in \mathbb{C}$  satisfies  $|\rho| > 1$ . Then,  $D_1$  is regular, but  $D_2$  is not.

In order to examine the accuracy of the least-squares estimation, we leverage existing finite time identification results for unstable dynamics [21]. First, if the empirical covariance matrix  $V_n = \sum_{t=0}^{n-1} x(t)x(t)'$  is non-singular, one can write

$\hat{D}_n = \sum_{t=0}^{n-1} x(t+1)x(t)'V_n^{-1}$ . Hence, the behavior of  $V_n$  governs the estimation accuracy. For unstable  $D$ , an appropriately normalized  $V_n$  is shown to be a random matrix [21]. Thus, letting  $\tilde{V}_n$  denote the normalized matrix, the accuracy of  $\hat{D}_n$  depends on the stochastic lower bounds of  $\tilde{V}_n$ . Let  $\psi(\delta)$  be the high probability lower bound of  $\tilde{V}_n$ ; i.e. it is sufficiently small to satisfy  $\mathbb{P}\left(|\lambda_{\min}(\tilde{V}_n)| < \psi(\delta)\right) < \delta$ . The following statement studies  $\psi(\delta)$  based on anti-concentration results for sequences of random matrices [25].

**Proposition 2.** [21] *Suppose that  $D$  is regular. In general,  $\delta > 0$  implies that  $\psi(\delta) > 0$ . Further, if  $w(t_0)$  has a bounded pdf for some  $t_0 \geq 1$ , then for all  $\delta > 0$  we have  $\psi(\delta) \geq \psi_0\delta$ , where  $\psi_0 > 0$  is a fixed constant.*

Theorem 1 determines the time length the user should interact with the system, in order to collect sufficiently many observations for accurate identification of the unstable matrix  $D$ . The sample size is based on the constant  $\rho$ , for which the exact dependence on the noise parameters  $b_1, b_2, \alpha, |\lambda_{\min}(C)|, |\lambda_{\max}(C)|$ , and the closed-loop matrix  $D$  is available [21]. Moreover, let  $\bar{\lambda}_1, \dots, \bar{\lambda}_k$  (respectively  $\underline{\lambda}_1, \dots, \underline{\lambda}_\ell$ ) be the distinct eigenvalues of  $D$  outside (respectively inside) the unit circle. Then,  $\psi(\delta)$  depends on [21]:

$$\begin{aligned} & |\lambda_{\min}(C)|, \quad \min_{1 \leq i \leq \ell} 1 - |\underline{\lambda}_i|, \\ & \min_{1 \leq i \leq k} \log |\bar{\lambda}_i|, \quad \min_{1 \leq i < j \leq k} \log |\bar{\lambda}_i - \bar{\lambda}_j|. \end{aligned}$$

The constant  $\psi_0$  depends on the upper bound of the pdf of  $w(t_0)$  as well. The explicit specification of these dependencies is fully presented in [21] and hence omitted. Next, let  $N(\epsilon, \delta)$  be large enough, such that  $n \geq N(\epsilon, \delta)$  implies

$$\frac{n}{(\log n)^{4/\alpha}} \geq \frac{\rho}{\epsilon^2} \left( (-\log \delta)^{1+4/\alpha} - \log \psi(\delta) \right). \quad (5)$$

**Theorem 1** (Unstable identification [21]). *Suppose that  $D$  is regular, and has no eigenvalue of unit size. As long as  $n \geq N(\epsilon, \delta)$ , we have*

$$\mathbb{P}\left(\left\| \hat{D}_n - D \right\| \leq \epsilon\right) \geq 1 - \delta.$$

Hence, by (5), the probability  $\delta$  of having an identification error of magnitude  $\epsilon$ , decays exponentially fast when  $n$  grows. In the next section, we show that one can satisfy the assumptions of Theorem 1 by applying random linear feedbacks to a stabilizable system with unknown dynamics parameters.

#### IV. STABILIZATION ALGORITHM

Although the true parameter  $\theta_0$  is unknown, according to Lemma 1, a stabilizing linear feedback  $L(\theta)$  can be designed, if one can find a stabilizing neighborhood  $\Omega^{(0)}$ , such that

$$\Omega^{(0)} \subset \{\theta \in \mathbb{R}^{p \times q} : \|\theta - \theta_0\| \leq \epsilon_0\}. \quad (6)$$

Using Theorem 1, we establish that  $\Omega^{(0)}$  can be estimated if one applies a random linear feedback to the system. Since in Theorem 1 the closed-loop transition matrix needs to be regular with no eigenvalue of unit size, first we need to show that these conditions can be satisfied. Lemma 2, and Lemma 3 accomplish this, with no knowledge beyond stabilizability of  $[A_0, B_0]$ . Based on the properties of the distribution of a random linear feedback matrix  $L$ , the above lemmas provide general statements, which hold almost surely. Then, we present a finite time stabilizing algorithm, and prove that it will provide us the desired stabilizing neighborhood. To proceed, we define the following classes of probability distributions over real valued vectors and matrices.

**Definition 3** (Full rank distributions). *Let  $X$  be a random vector in  $\mathbb{R}^m$ .  $X$  has a linearly full rank distribution if for any arbitrary hyperplane  $\mathcal{P} \subset \mathbb{R}^m$ , it holds that  $\mathbb{P}(X \in \mathcal{P}) = 0$ . Further,  $X$  has a general full rank distribution, if for every manifold  $\mathcal{M} \subset \mathbb{R}^m$  such that  $\dim_{\mathbb{R}}(\mathcal{M}) \leq m - 1$ , it holds that  $\mathbb{P}(X \in \mathcal{M}) = 0$ .*

The following example illustrates the difference between the two types of full rank distributions defined above.

**Example 1.** *Let  $Z \in \mathbb{R}^p$  be normally distributed,  $Z \sim \mathcal{N}(\mu, \Sigma)$ , with arbitrary mean  $\mu \in \mathbb{R}^p$ , and positive definite covariance matrix  $\Sigma \in \mathbb{R}^{p \times p}$ . Then,  $Z$  has a general full rank distribution. Letting  $Y = (Z/\|Z\|) \mathbf{1}_{\{Z \neq 0\}}$ , the random vector  $Y$  has a linearly full rank distribution, but since it lives on the unit sphere,  $Y$  does not have a general full rank distribution.*

Random linear feedbacks with full rank distributions induce the desired properties to the closed-loop transition matrix, as we rigorously establish below.

**Lemma 2** (Closed-loop regularity). *Assume  $[A_0, B_0]$  is stabilizable. Let the columns of  $L \in \mathbb{R}^{r \times p}$  be independent (but not necessarily identically distributed), with linearly full rank distributions. The matrix  $A_0 + B_0L$  is regular, with probability one.*

**Proof of Lemma 2.** Let the event  $\mathcal{G}$  be that  $D = A_0 + B_0L$  is irregular. We prove that for all  $\lambda \in \mathbb{C}$ ,  $|\lambda| \geq 1$ , with probability one,  $\text{rank}(D - \lambda I_p) \geq p - 1$ . Note that according to the discussion after Definition 2, this implies  $\mathbb{P}(\mathcal{G}) = 0$ .

First, let  $Y_i \in \mathbb{R}^m, i = 1, \dots, m$  have linearly full rank distributions. Define  $Y = [Y_1, \dots, Y_m]$ , and let  $M(\lambda)$  be a  $m \times m$  matrix, with all coordinates being real polynomials of  $\lambda$ . Let  $f(\lambda)$  be a real polynomial of  $\lambda$  as well. We show that 
$$\mathbb{P}\left(\exists \lambda \in \mathbb{C}, f(\lambda) \neq 0 : \text{rank}\left(Y - \frac{M(\lambda)}{f(\lambda)}\right) < m - 1\right) = 0. \quad (7)$$

If  $\text{rank}(Y - M(\lambda_0)/f(\lambda_0)) < m - 1$ , denoting

$$\frac{M(\lambda_0)}{f(\lambda_0)} = [M_1, \dots, M_m],$$

two of the vectors  $Y_i - M_i, i = 1, \dots, m$ , such as  $Y_{m-1} - M_{m-1}, Y_m - M_m$ , can be written as linear combinations of the others. There are finitely many values of  $\lambda_0$  for which  $Y_{m-1} - M_{m-1}$  is a linear combination of  $Y_1 - M_1, \dots, Y_{m-2} - M_{m-2}$ ,

since for every such a  $\lambda_0$ ,  $\det(\tilde{Y}) = 0$ , where  $\tilde{Y}$  is the square matrix whose columns are  $Y_1 - M_1, \dots, Y_{m-1} - M_{m-1}$ , removing an arbitrary row. Note that  $\det(\tilde{Y})$  is a polynomial of  $\lambda_0$ , divided by  $f(\lambda_0)$ , and  $f(\lambda_0) \neq 0$ .

Note that  $\lambda_0$  is a deterministic function of  $Y_1, \dots, Y_m$ . For every such  $\lambda_0$ , the dimension of the subspace  $\mathcal{P}$  spanned by  $Y_1 - M_1, \dots, Y_{m-2} - M_{m-2}, M_m$  is at most  $m - 1$ . Because  $Y_m$  is independent of  $Y_1, \dots, Y_{m-1}$ , and  $Y_m$  has a linearly full rank distribution,  $\mathbb{P}(Y_m \in \mathcal{P}) = 0$ ; i.e. (7) holds.

Now, let  $m = \text{rank}(B_0)$ . If  $m = p$ , applying the above argument to  $Y = D, M(\lambda) = \lambda I_p, f(\lambda) = 1$ , we have  $\mathbb{P}(\mathcal{G}) = 0$ , since full rankness of  $B_0$  implies linearly full rank distributions for all columns of  $B_0L$ . If  $m < p$ , there is a  $p \times p$  permutation matrix  $J$ , and  $K \in \mathbb{R}^{(p-m) \times m}$ , such that

$$JB_0 = \begin{bmatrix} \tilde{B} \\ K\tilde{B} \end{bmatrix} = \begin{bmatrix} I_m \\ K \end{bmatrix} \tilde{B},$$

where  $\tilde{B} \in \mathbb{R}^{m \times r}$  is full rank. Let  $L_0$  be a stabilizer,  $D_0 = A_0 + B_0L_0$ , and  $JD_0 = [D'_1, D'_2]'$ ,  $D_1 \in \mathbb{R}^{m \times p}, D_2 \in \mathbb{R}^{(p-m) \times p}$ , to get

$$J(A_0 + B_0L) = \begin{bmatrix} D_1 + \tilde{B}(L - L_0) \\ D_2 + K\tilde{B}(L - L_0) \end{bmatrix}.$$

Writing  $J = [J'_1, J'_2]'$ ,  $J_1 \in \mathbb{R}^{m \times p}, J_2 \in \mathbb{R}^{(p-m) \times p}$ , we have

$$\begin{aligned} & \text{rank}(A_0 + B_0L - \lambda I_p) \\ &= \text{rank}\left(\begin{bmatrix} I_m & 0_{m \times (p-m)} \\ -K & I_{p-m} \end{bmatrix} (J(A_0 + B_0L) - \lambda J)\right) \\ &= \text{rank}\left(\begin{bmatrix} D_1 + \tilde{B}(L - L_0) - \lambda J_1 \\ [-K, I_{p-m}]J(D_0 - \lambda I_p) \end{bmatrix}\right). \end{aligned}$$

Denote the last matrix above by  $\tilde{X}$ . Since  $|\lambda_{\max}(D_0)| < 1$ , for  $|\lambda| \geq 1$  the matrix  $D_0 - \lambda I_p$  is full rank. Therefore, because of  $\text{rank}([-K, I_{p-m}]) = p - m$ , we have  $\text{rank}([-K, I_{p-m}]J(D_0 - \lambda I_p)) = p - m$ .

Rearrange the columns of matrix  $\tilde{X}$  to get

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix},$$

such that  $X_{11} \in \mathbb{C}^{m \times m}, X_{22} \in \mathbb{C}^{(p-m) \times (p-m)}, \text{rank}(X_{22}) = p - m$ . In other words,  $p - m$  linearly independent columns of  $[-K, I_{p-m}]J(D_0 - \lambda I_p)$  have been put together to form  $X_{22}$ . If  $D$  is not regular,

$$\begin{aligned} p - 2 &\geq \text{rank}(\tilde{X}) = \text{rank}(X) \\ &= \text{rank}\left(X \begin{bmatrix} I_m & 0_{m \times (p-m)} \\ -X_{22}^{-1}X_{21} & I_{p-m} \end{bmatrix}\right) \\ &= \text{rank}\left(\begin{bmatrix} X_{11} - X_{12}X_{22}^{-1}X_{21} & X_{12} \\ 0_{(p-m) \times m} & X_{22} \end{bmatrix}\right). \end{aligned}$$

Hence,  $\text{rank}(X_{11} - X_{12}X_{22}^{-1}X_{21}) \leq m - 2$ . Recall that columns of  $[X_{11}, X_{12}]$  are exactly the same as  $D_1 + \tilde{B}(L - L_0) - \lambda J_1$ , and all coordinates of  $\det(X_{22})X_{12}X_{22}^{-1}X_{21}$  are polynomials of  $\lambda$  (since all coordinates of  $\det(X_{22})X_{22}^{-1}$  are polynomials of the coordinates of  $X_{22}$ ). Taking  $f(\lambda) = \det(X_{22})$ , by (7), since full rankness of  $\tilde{B}$  implies linearly full rank

distributions for all columns of  $\tilde{B}(L - L_0)$ , we have  $\mathbb{P}(\text{rank}(X_{11} - X_{12}X_{22}^{-1}X_{21}) \leq m - 2) = 0$ , which is the desired result since  $\text{rank}(X_{22}) = p - m$ .  $\square$

If the distribution of linear feedback  $L$  is generally full rank, the following results shows that  $A_0 + B_0L$  has no eigenvalue on the unit circle of the complex plane.

**Lemma 3** (Closed-loop eigenvalues). *Assume  $[A_0, B_0]$  is stabilizable. Let  $L \in \mathbb{R}^{r \times p}$  have a general full rank distribution over  $\mathbb{R}^{r \times p}$ . With probability one,  $A_0 + B_0L$  has no unit size eigenvalue.*

**Proof of Lemma 3.** Assume  $D = A_0 + B_0L$  has a unit-root eigenvalue, denoted by  $\lambda \in \mathbb{C}, |\lambda| = 1$ . Further, assume that  $m = \text{rank}(B_0)$ , and let the permutation matrix  $J$  and the matrix  $K \in \mathbb{R}^{(p-m) \times m}$  be such that

$$JB_0 = \begin{bmatrix} \tilde{B} \\ K\tilde{B} \end{bmatrix} = \begin{bmatrix} I_m \\ K \end{bmatrix} \tilde{B},$$

where  $\tilde{B} \in \mathbb{R}^{m \times r}$  is full rank. Letting  $L_0$  be a stabilizer,  $D_0 = A_0 + B_0L_0$ , and  $X = \tilde{B}(L - L_0) \in \mathbb{R}^{m \times p}$ , note that  $X$  has a general full rank distribution, thanks to the full-rankness of  $\tilde{B}$ . Since  $D_0$  is stable,  $\det(D_0 - \lambda I_p) \neq 0$ , and

$$\begin{aligned} 0 &= \det(A_0 + B_0L - \lambda I_p) \\ &= \det\left(JD_0 + \begin{bmatrix} I_m \\ K \end{bmatrix} X - \lambda J\right) \\ &= \det\left((D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix} X + I_p\right) \\ &= \det\left(X(D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix} + I_m\right), \end{aligned}$$

where the last equality above is implied by Sylvester's determinant identity. Denote the complex conjugate of  $\lambda$  by  $\bar{\lambda}$ , and define the real matrix

$$M(\lambda) = M(\bar{\lambda}) = (D_0 - \bar{\lambda}I_p)^{-1} (D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix}.$$

Further, define the space of eigenvectors in  $\mathbb{C}^m$  as follows. First, consider the relation  $\sim$  on  $\mathbb{C}^m$ , defined as

$$x \sim y, \text{ if } x = cy \text{ for some } c \in \mathbb{C}, c \neq 0.$$

Since  $\sim$  is an equivalence relation, for the set of equivalence classes denoted by  $S = \frac{\mathbb{C}^m}{\sim}$  (which is the direction space in  $\mathbb{C}^m$ ) we have  $\dim_{\mathbb{C}}(S) = m - 1$ ; i.e.  $\dim_{\mathbb{R}}(S) = 2m - 2$ .

Note that for every matrix  $Y \in \mathbb{C}^{m \times m}$  and every vector  $v \in \mathbb{C}^m$ ,  $Yv = 0$  if and only if  $Y\tilde{v} = 0$  for every  $\tilde{v} \sim v$ . Thus, according to

$$\det\left(X(D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix} + I_m\right) = 0,$$

there is  $v \in S, v \neq 0$ , such that

$$(X(D_0 - \bar{\lambda}I_p)M(\lambda) + I_m)v = 0 \quad (8)$$

Denote the set of all matrices  $X$  satisfying (8) by  $\mathcal{X}(\lambda, v) \subset \mathbb{R}^{m \times p}$ . Separating the real ( $\Re$ ) and imaginary ( $\Im$ ) parts, we get  $Xa(v) = \Re(v)$ ,  $Xb(v) = \Im(v)$ , where for  $v \in S$ , the

vectors  $a(v), b(v) \in \mathbb{R}^p$  are defined as

$$\begin{aligned} a(v) &= M(\lambda)\Re(\bar{\lambda}v) - D_0M(\lambda)\Re(v), \\ b(v) &= M(\lambda)\Im(\bar{\lambda}v) - D_0M(\lambda)\Im(v). \end{aligned}$$

Next, we partition  $S$  to  $S_1, S_2$ ; i.e.  $S = S_1 \cup S_2, S_1 \cap S_2 = \emptyset$ , where

$$\begin{aligned} S_1 &= \{v \in S : a(v), b(v) \text{ are in-line}\}, \\ S_2 &= \{v \in S : a(v), b(v) \text{ are not in-line}\}. \end{aligned}$$

Whenever  $v \in S_2$ , for  $j = 1, \dots, m$ , the  $j$ -th row of  $X$  needs to be in the intersection of two nonparallel hyperplanes  $\mathcal{P}_1, \mathcal{P}_2 \subset \mathbb{R}^p$ , where

$$\begin{aligned} \mathcal{P}_1 &= \{y \in \mathbb{R}^p : y'a(v) = \Re(v_j)\}, \\ \mathcal{P}_2 &= \{y \in \mathbb{R}^p : y'b(v) = \Im(v_j)\}. \end{aligned}$$

Since  $\dim_{\mathbb{R}}(\mathcal{P}_1) \leq p - 1$ ,  $\dim_{\mathbb{R}}(\mathcal{P}_2) \leq p - 1$ , and  $v \in S_2$ , we have  $\dim_{\mathbb{R}}(\mathcal{P}_1 \cap \mathcal{P}_2) \leq p - 2$ . Therefore, for  $v \in S_2$ , we have  $\dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \leq m(p - 2)$ . Since  $\dim_{\mathbb{R}}(|\lambda| = 1) = 1$ , using  $\dim_{\mathbb{R}}(S_2) \leq 2m - 2$  we have

$$\dim_{\mathbb{R}}(\mathcal{Z}_1) \leq 1 + 2m - 2 + m(p - 2) = mp - 1, \quad (9)$$

where  $\mathcal{Z}_1 = \bigcup_{|\lambda|=1, v \in S_2} \mathcal{X}(\lambda, v)$ .

On the other hand, for  $v \in S_1$ , there is a real number, say  $\varphi(v)$ , such that  $b(v) = \varphi(v)a(v)$ . Then,

$$\Im(v) = Xb(v) = \varphi(v)Xa(v) = \varphi(v)\Re(v), \quad (10)$$

i.e. whenever  $v \in S_1$ , the vectors  $\Re(v), \Im(v)$  are in-line. So,  $\dim_{\mathbb{R}}(S_1) = m - 1$ , and for  $v \in S_1$ , we have  $\mathcal{P}_1 = \mathcal{P}_2$ , i.e.  $\dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \leq m(p - 1)$ . After doing some algebra, we obtain

$$\begin{aligned} 0 &= \varphi(v)a(v) - b(v) \\ &= \varphi(v)(\Re(\lambda)I_p + \varphi(v)\Im(\lambda)I_p - D_0)M(\lambda)\Re(v) \\ &\quad - (\varphi(v)\Re(\lambda)I_p - \Im(\lambda)I_p - \varphi(v)D_0)M(\lambda)\Re(v) \\ &= (1 + \varphi(v)^2)\Im(\lambda)M(\lambda)\Re(v), \end{aligned}$$

i.e. either  $\Im(\lambda) = 0$ , or  $M(\lambda)\Re(v) = 0$ . According to the definition of  $M(\lambda)$ , the latter case implies  $\Re(v) = 0$ , which due to (10) leads to  $v = 0$ , and is impossible. So, by  $\dim_{\mathbb{R}}(|\lambda| = 1, \Im(\lambda) = 0) = 0$ , we have

$$\dim_{\mathbb{R}}(\mathcal{Z}_2) \leq m - 1 + m(p - 1) = mp - 1, \quad (11)$$

where  $\mathcal{Z}_2 = \bigcup_{|\lambda|=1, \Im(\lambda)=0} \mathcal{X}(\lambda, v)$ . Writing  $\mathcal{X} =$

$\bigcup_{|\lambda|=1, v \in S} \mathcal{X}(\lambda, v) \subset \mathcal{Z}_1 \cup \mathcal{Z}_2$ , according to (9), (11) we have  $\dim_{\mathbb{R}}(\mathcal{X}) \leq mp - 1$ , and by general full-rankness of the distribution of  $X$ , the desired result holds:  $\mathbb{P}(\mathcal{X}) = 0$ .  $\square$

Subsequently, an algorithmic procedure to find a stabilizing neighborhood will be presented based on random linear feedbacks discussed above. First, letting  $k = 1 + \lceil r/p \rceil$ , draw the columns of  $L_1, \dots, L_k \in \mathbb{R}^{r \times p}$  from independent standard Gaussian distributions  $\mathcal{N}(0, I_r)$ . Note that because of independence, for all  $i = 1, \dots, k$ , the random feedback  $L_i$  has a general full rank distribution  $\mathbb{R}^{r \times p}$ . Lemma 2 and Lemma 3 show that the conditions of Theorem 1 hold. Therefore,

every closed-loop transition matrix  $D^{(i)} = A_0 + B_0 L_i$  can be estimated arbitrarily accurate. We show how to find a high probability confidence set for  $\theta_0$ , using the accurate estimates of  $D^{(1)}, \dots, D^{(k)}$ .

---

**Algorithm 1 : Adaptive Stabilization**


---

Output: Stabilizing Set  $\Omega^{(0)}$

---

Let  $k = 1 + \lceil \frac{r}{p} \rceil, \tau_0 = 0$

**for**  $i = 1, \dots, k$  **do**

**for**  $j = 1, \dots, p$  **do**

        Draw column  $j$  of  $L_i$  from  $\mathcal{N}(0, I_r)$ , independently

**end for**

**end for**

Define  $M, \tilde{\epsilon}$  according to (12), (13), respectively

**for**  $i = 1, \dots, k$  **do**

    Define  $\tau_i$  by (14)

**while**  $t < \tau_i$  **do**

        Apply control action  $u(t) = L_i x(t)$

**end while**

    Estimate  $\hat{D}^{(i)}$  by (15)

    Construct  $\Omega^{(i)}$  by (16)

**end for**

**return**  $\Omega^{(0)} = \bigcap_{i=1}^k \Omega^{(i)}$

---

Letting  $\epsilon_0$  be as Lemma 1, define the precision  $\tilde{\epsilon}$  and the matrix  $M$  containing all matrices  $L_1, \dots, L_k$  by

$$M = \begin{bmatrix} I_p & \cdots & I_p \\ L_1 & \cdots & L_k \end{bmatrix} \in \mathbb{R}^{q \times kp}, \quad (12)$$

$$\tilde{\epsilon} = \frac{\epsilon_0}{2k} \inf \left\{ \frac{\| \theta M \|}{\| \theta \|} : \theta \in \mathbb{R}^{p \times q} \right\}. \quad (13)$$

As a matter of fact, since  $kp \geq q$  and  $L_1, \dots, L_k$  are independent, we have  $\text{rank}(M) = q$ , almost surely; i.e.  $\mathbb{P}(\tilde{\epsilon} > 0) = 1$ . Further, if  $\tilde{\epsilon}$  became too small once  $L_1, \dots, L_k$  are drawn, one can repeatedly draw the random feedbacks to avoid pathologically small values of  $\tilde{\epsilon}$ .

Then, let  $\tau_0 = 0$ , and for  $i = 1, \dots, k$  define the following:

$$\tau_i = \tau_{i-1} + N \left( \tilde{\epsilon}, \frac{\delta}{k} \right), \quad (14)$$

$$\hat{D}^{(i)} = \arg \min_{E \in \mathbb{R}^{p \times p}} \sum_{t=\tau_{i-1}}^{\tau_i-1} \|x(t+1) - Ex(t)\|^2, \quad (15)$$

$$\Omega^{(i)} = \left\{ \theta \in \mathbb{R}^{p \times q} : \left\| \theta \begin{bmatrix} I_p \\ L_i \end{bmatrix} - \hat{D}^{(i)} \right\| \leq \tilde{\epsilon} \right\}, \quad (16)$$

where the sample size  $N(\cdot, \cdot)$  is given in (5). Conceptually,  $\tau_i$  is the time point when the control action changes,  $\hat{D}^{(i)}$  is the least-squares estimate, and  $\Omega^{(i)}$  is a confidence set for  $\theta_0$ . In fact, for each  $i = 1, \dots, k$ , Algorithm 1 applies the linear feedback  $u(t) = L_i x(t)$  during the time period  $\tau_{i-1} \leq t < \tau_i$ . Then, observing  $\{x(t)\}_{t=\tau_{i-1}}^{\tau_i-1}$ , the algorithm uses  $\hat{D}^{(i)}$  to estimate the true closed-loop matrix

$$D^{(i)} = A_0 + B_0 L_i = \theta_0 \begin{bmatrix} I_p \\ L_i \end{bmatrix}.$$

Finally, the high probability confidence set  $\Omega^{(0)}$  is constructed

for the true parameter  $\theta_0$ , according to  $L_i$ . Iterating the above procedure for all  $1 \leq i \leq k$ , the algorithm constructs  $\Omega^{(1)}, \dots, \Omega^{(k)}$ , and returns  $\Omega^{(0)} = \bigcap_{i=1}^k \Omega^{(i)}$  as a stabilizing set. Below, we show that it satisfies (6).

By Theorem 1, (14) implies that  $\left\| \hat{D}^{(i)} - D^{(i)} \right\| \leq \tilde{\epsilon}$ , with probability at least  $1 - \delta/k$ . So, by (16), we have  $\mathbb{P}(\theta_0 \notin \Omega^{(i)}) \leq \delta/k$ ; i.e.  $\mathbb{P}(\theta_0 \notin \Omega^{(0)}) \leq \delta$ . To show that  $\Omega^{(0)}$  is a stabilizing set, let  $\theta_1 \in \Omega^{(0)}$  be arbitrary. On the event  $\theta_0 \in \Omega^{(0)}$ , for all  $i = 1, \dots, k$  we have

$$\left\| (\theta_1 - \theta_0) \begin{bmatrix} I_p \\ L_i \end{bmatrix} \right\| \leq 2\tilde{\epsilon}.$$

Using the definition of  $M$  in (12), the latter result leads to  $\|(\theta_1 - \theta_0) M\| \leq 2k\tilde{\epsilon}$ . Thus, (13) implies the following for  $\theta_1 \neq \theta_0$ :

$$\frac{2k\tilde{\epsilon}}{\| \theta_1 - \theta_0 \|} \geq \frac{\| (\theta_1 - \theta_0) M \|}{\| \theta_1 - \theta_0 \|} \geq \frac{2k\tilde{\epsilon}}{\epsilon_0},$$

or equivalently  $\| \theta_1 - \theta_0 \| \leq \epsilon_0$ , which is the desired inequality of (6). Note that since  $\theta_0 \in \Omega^{(0)}$ , with probability at least  $1 - \delta$ , the failure probability of Algorithm 1 is at most  $\delta$ . This completes the proof of the following result.

**Theorem 2 (Stabilization).** *Let  $\Omega^{(0)}$  be the stabilizing set provided by Algorithm 1. For arbitrary  $\theta \in \Omega^{(0)}$ , we have*

$$\mathbb{P} \left( \left| \lambda_{\max} \left( \theta_0 \tilde{L}(\theta) \right) \right| < 1 \right) \geq 1 - \delta.$$

In other words, the probability of failing to stabilize the system decays exponentially when the time of interaction with the system grows (see (5)). Obviously, the normal distribution  $\mathcal{N}(0, I_r)$  used in Algorithm 1 is not unique, and can be substituted by any general full rank distribution over  $\mathbb{R}^r$ .

## V. CONCLUSION

We studied an adaptive stabilization scheme for linear dynamical systems, focusing on finite time analysis. Tailoring a novel procedure based on random linear feedbacks, we established non-asymptotic results under mild assumptions, namely those of system stabilizability and a fairly general noise process that encompasses heavy-tailed distributions.

There are a number of interesting extensions of the current work. First, finite time analysis of stabilization given *noisy observations* of the state vector is an interesting topic for future investigation. Second, studying the stabilization problem in a high-dimensional setting (assuming sparsity or some other low dimensional structure) is also an interesting subject to be addressed in the future.

## APPENDIX A

**Proof of Proposition 1.** For convenience, let  $K_0 = K(\theta_0)$ , and  $L_0 = L(\theta_0)$ . First, assume  $[A_0, B_0]$  is stabilizable,  $L$  is a stabilizer,  $D = A_0 + B_0 L$ , and  $|\lambda_{\max}(D)| < 1$ . For arbitrary fixed PSD matrix  $P_0$ , define  $P_t(P_0), t = 1, \dots, T$  recursively,

$$P_t(P_0) = Q + A_0' P_{t-1}(P_0) A_0 - A_0' P_{t-1}(P_0) B_0 (B_0' P_{t-1}(P_0) B_0 + R)^{-1} B_0' P_{t-1}(P_0) A_0.$$

Letting  $c_t$  be as defined in (2), the optimal control policy for minimizing the finite horizon cumulative cost  $\mathcal{J}_T = \sum_{t=0}^{T-1} \mathbb{E}[c_t] + \mathbb{E}[x(T)'P_0x(T)]$ , is  $u(t) = L_t x(t)$ ,  $t = 0, \dots, T-1$ , [16], where

$$L_t = -(B_0'P_{T-t-1}(P_0)B_0 + R)^{-1} B_0'P_{T-t-1}(P_0)A_0. \quad (17)$$

Moreover, this optimal policy yields the optimal cost

$$\min \mathcal{J}_T = x(0)'P_T(P_0)x(0) + \sum_{t=0}^{T-1} \text{tr}(CP_t(P_0)). \quad (18)$$

On the other hand, applying the control policy  $u(t) = L_t x(t)$ ,  $0 \leq t \leq T-1$ , we have

$$\begin{aligned} \mathbb{E}[x(T)'P_0x(T)|x(T-1)] &= x(T-1)'D'P_0Dx(T-1) \\ &\quad + \text{tr}(CP_0), \\ \mathbb{E}[c_{t+1}|x(t)] &= x(t)'D'(Q + L'RL)Dx(t) \\ &\quad + \text{tr}(C(Q + L'RL)), \end{aligned}$$

for  $t = 0, \dots, T-2$ . Hence, the finite horizon cost becomes

$$\mathcal{J}_T = x(0)'\tilde{P}_T(P_0)x(0) + \sum_{t=0}^{T-1} \text{tr}(C\tilde{P}_t(P_0)), \quad (19)$$

where  $\tilde{P}_t(P_0)$ ,  $t = 1, \dots, T$  are defined recursively as

$$\tilde{P}_0(P_0) = P_0, \quad (20)$$

$$\tilde{P}_t(P_0) = Q + L'RL + D'\tilde{P}_{t-1}(P_0)D. \quad (21)$$

Since  $|\lambda_{\max}(D)| < 1$ ,  $\lim_{T \rightarrow \infty} \tilde{P}_T(P_0) = P_\infty$  for a PSD matrix  $P_\infty$ . Letting  $C \rightarrow 0$ , by (18), (19) we have

$$x(0)'P_T(P_0)x(0) \leq x(0)'\tilde{P}_T(P_0)x(0),$$

i.e.  $x(0)'P_T(P_0)x(0)$ ,  $T = 1, 2, \dots$  is bounded. If  $P_0 = 0$ , this sequence is nondecreasing, because minimizing both sides of

$$\sum_{t=0}^{T-1} c_t \leq \sum_{t=0}^T c_t$$

subject to

$$x(t+1) = A_0x(t) + B_0u(t),$$

we get

$$x(0)'P_T(0)x(0) \leq x(0)'P_{T+1}(0)x(0).$$

Therefore, the nondecreasing bounded sequence  $\{x(0)'P_T(0)x(0)\}_{T=1}^{\infty}$  converges. Since  $x(0)$  is arbitrary,  $\{P_T(0)\}_{T=1}^{\infty}$  itself converges:

$$\lim_{T \rightarrow \infty} P_T(0) = P_\infty(0).$$

According to the recursive definition of  $P_t(0)$  in (17),  $P_\infty(0)$  is a solution of (3). This shows the existence of a solution, while uniqueness will be established later.

Next, since  $\lim_{T \rightarrow \infty} P_T(0) = P_\infty(0)$ , (18) implies  $\lim_{t \rightarrow \infty} \text{tr}(CP_t(0)) = \text{tr}(CP_\infty(0))$ . So, the Cesaro mean also converges to this limit, i.e.

$$\mathcal{J}^*(\theta_0) = \text{tr}(CP_\infty(0)).$$

Optimality of the linear feedback  $u(t) = L_0x(t)$ , is then established through (17). Now, we are ready to show that  $L_0$  is a stabilizer. Letting

$$D = A_0 + B_0L_0, C \rightarrow 0, K_0 = P_\infty(0),$$

we show that for arbitrary  $x(0)$ ,  $x(t) = D^t x(0)$  vanishes as  $t$  grows. First, note that by (3), (4),

$$\begin{aligned} (B_0'K_0B_0 + R)L_0 &= -B_0'K_0A_0, \\ L_0'(B_0'K_0B_0 + R)L_0 &= A_0'K_0B_0(B_0'K_0B_0 + R)^{-1}B_0'K_0A_0. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} &Q + L_0'RL_0 + D'K_0D \\ &= Q + A_0'K_0A_0 + L_0'(B_0'K_0A_0 + R)L_0 \\ &\quad + A_0'K_0B_0L_0 + L_0'B_0'K_0A_0 \\ &= Q + A_0'K_0A_0 - A_0'K_0B_0(B_0'K_0B_0 + R)^{-1}B_0'K_0A_0 \\ &\quad + [L_0'(B_0'K_0B_0 + R) + A_0'K_0B_0]L_0 \\ &\quad + L_0'[(B_0'K_0B_0 + R)L_0 + B_0'K_0A_0] = K_0, \end{aligned}$$

that is,

$$K_0 - D'K_0D = Q + L_0'RL_0. \quad (22)$$

So,

$$x(t+1)'K_0x(t+1) - x(t)'K_0x(t) = -x(t)'(Q + L_0'RL_0)x(t). \quad (23)$$

Adding up both sides of (23), because  $K_0$  is PSD we get

$$\begin{aligned} -x(0)'K_0x(0) &\leq x(t+1)'K_0x(t+1) - x(0)'K_0x(0) \\ &= -\sum_{i=0}^t x(i)'(Q + L_0'RL_0)x(i). \quad (24) \end{aligned}$$

In other words,

$$\lim_{t \rightarrow \infty} x(t)'(Q + L_0'RL_0)x(t) = 0.$$

Thus, since  $Q$  is positive definite,  $\lim_{t \rightarrow \infty} x(t) = 0$ , i.e.  $L_0$  is a stabilizer. Back to the proof of the existence of a solution  $K_0$ , we show that for arbitrary PSD  $P_0$ , it holds that  $\lim_{T \rightarrow \infty} P_T(P_0) = P_\infty(0)$ . To do so, minimizing both sides of

$$\sum_{t=0}^{T-1} c_t \leq \sum_{t=0}^{T-1} c_t + x(T)'P_0x(T),$$

subject to

$$x(t+1) = A_0x(t) + B_0u(t),$$

we get

$$x(0)'P_T(0)x(0) \leq x(0)'P_T(P_0)x(0). \quad (25)$$

On the other hand, applying controller  $u(t) = L_0x(t)$ , the cost  $\sum_{t=0}^{T-1} c_t + x(T)'P_0x(T)$  becomes

$$\sum_{t=0}^{T-1} x(0)'D'^t(Q + L_0'RL_0)D^t x(0) + x(0)'D'^T P_0 D^T x(0). \quad (26)$$

Note that because of stability  $|\lambda_{\max}(D)| < 1$ , we have

$$\lim_{T \rightarrow \infty} x(0)' D'^T P_0 D^T x(0) = 0.$$

Therefore, by combining (25), (26), and (22),

$$\begin{aligned} & x(0)' P_\infty(0) x(0) \\ = & \lim_{T \rightarrow \infty} x(0)' P_T(0) x(0) \leq \lim_{T \rightarrow \infty} x(0)' P_T(P_0) x(0) \\ \leq & \lim_{T \rightarrow \infty} \sum_{t=0}^{T-1} x(0)' D'^t (Q + L'_0 R L_0) D^t x(0) \\ + & x(0)' D'^T P_0 D^T x(0) \\ = & \lim_{T \rightarrow \infty} \sum_{t=0}^{T-1} x(0)' D'^t (K_0 - D' K_0 D) D^t x(0) \\ = & x(0)' K_0 x(0), \end{aligned}$$

i.e. for an arbitrary  $P_0$ ,

$$\lim_{T \rightarrow \infty} P_T(P_0) = P_\infty(0).$$

Using this, we show that  $K_0$  is the unique solution of (3). If  $P_*$  is another solution, let  $P_0 = P_*$ , which plugging in (17) implies that  $P_t(P_*) = P_*$ , for all  $t = 1, 2, \dots$ . Hence  $P_* = \lim_{T \rightarrow \infty} P_T(P_*) = P_\infty(0)$ , i.e. the solution  $K_0$  of (3) exists, and is unique.

Conversely, if  $K_0$  is a solution of (3), define  $L_0$  as (4) and  $D = A_0 + B_0 L_0$ . Note that  $K_0$  is positive semidefinite, and let  $P_0 = K_0$ . Defining  $P_t$  by (17), we obtain  $P_t = K_0$ , for all  $t = 0, 1, \dots$ . As before, (3), (4) imply (22). Similarly, (23), (24) hold, i.e.  $\lim_{t \rightarrow \infty} D^t x(0) = 0$  for arbitrary  $x(0)$ , which implies that  $L_0$  defined in (4) is a stabilizer.  $\square$

## APPENDIX B

**Proof of Lemma 1.** Since  $\theta$  is stabilizable, according to Proposition 1,  $\theta \tilde{L}(\theta)$  is stable;

$$\left| \lambda_{\max}(\theta \tilde{L}(\theta)) \right| \leq 1 - 2\rho,$$

for some  $\rho > 0$ . For arbitrary fixed  $1 \leq i \leq p, 1 \leq j \leq q$ , let all entries of the matrix  $X_{ij} \in \mathbb{R}^{p \times q}$  be zero, except the  $ij$ -th entry, which is one. Then, for  $\varphi \in \mathbb{R}$ , consider the polynomial

$$f_\varphi(\lambda) = \det\left((\theta + \varphi X_{ij}) \tilde{L}(\theta) - \lambda I_p\right).$$

All coefficients of  $f_\varphi(\lambda)$  are linear functions of  $\varphi$ . Further, the magnitudes of the roots of  $f_\varphi(\lambda)$  are continuous with respect to the coefficients, and so, are also continuous with respect to  $\varphi$ . Since all roots of  $f_0(\lambda)$  are in magnitude at most  $1 - 2\rho$ , there exists  $\epsilon_{ij} > 0$ , such that  $|\varphi| < \epsilon_{ij}$  implies that all roots of  $f_\varphi(\lambda)$  are in magnitude at most  $1 - (2 - 1/(pq))\rho$ . Taking  $\epsilon_0 = \min_{i,j} \epsilon_{ij}$ , by  $\|\theta - \theta_0\| < \epsilon_0$ ,  $\theta_0$  can be written in the

form of  $\theta_0 = \theta + \sum_{i=1}^p \sum_{j=1}^q \varphi_{ij} X_{ij}$ , where  $|\varphi_{ij}| < \epsilon_{ij}$ , for all  $i, j$ . Therefore, all roots of

$$f(\lambda) = \det\left(\theta_0 \tilde{L}(\theta) - \lambda I_p\right)$$

are in magnitude at most  $1 - \rho$ , which is the desired result.  $\square$

## REFERENCES

- [1] T. Lai, "Asymptotically efficient adaptive control in stochastic regression models," *Advances in Applied Mathematics*, vol. 7, no. 1, pp. 23–45, 1986.
- [2] L. Guo and H. Chen, "Convergence rate of els based adaptive tracker," *Syst. Sci & Math. Sci.*, vol. 1, pp. 131–138, 1988.
- [3] H.-F. Chen and J.-F. Zhang, "Convergence rates in stochastic adaptive tracking," *International Journal of Control*, vol. 49, no. 6, pp. 1915–1935, 1989.
- [4] T. L. Lai and Z. Ying, "Parallel recursive algorithms in asymptotically efficient adaptive control of linear stochastic systems," *SIAM journal on control and optimization*, vol. 29, no. 5, pp. 1091–1127, 1991.
- [5] L. Guo and H.-F. Chen, "The Åström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers," *IEEE Transactions on Automatic Control*, vol. 36, no. 7, pp. 802–812, 1991.
- [6] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp. 494–500, 1974.
- [7] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *The Annals of Statistics*, pp. 154–166, 1982.
- [8] A. Becker, P. Kumar, and C.-Z. Wei, "Adaptive control with the stochastic approximation algorithm: Geometry and convergence," *IEEE Transactions on Automatic Control*, vol. 30, no. 4, pp. 330–338, 1985.
- [9] P. Kumar, "Convergence of adaptive control schemes using least-squares parameter estimates," *IEEE Transactions on Automatic Control*, vol. 35, no. 4, pp. 416–424, 1990.
- [10] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, 1985.
- [11] M. C. Campi, "Achieving optimality in adaptive control: the 'bet on the best' approach," in *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, vol. 5. IEEE, 1997, pp. 4671–4676.
- [12] M. C. Campi and P. Kumar, "Adaptive linear quadratic gaussian control: the cost-biased approach revisited," *SIAM Journal on Control and Optimization*, vol. 36, no. 6, pp. 1890–1907, 1998.
- [13] S. Bittanti and M. C. Campi, "Adaptive control of linear time invariant systems: the bet on the best principle," *Communications in Information & Systems*, vol. 6, no. 4, pp. 299–320, 2006.
- [14] Y. Abbasi-Yadkori and C. Szepesvári, "Regret bounds for the adaptive control of linear quadratic systems," in *COLT*, 2011, pp. 1–26.
- [15] M. Ibrahim, A. Javanmard, and B. V. Roy, "Efficient reinforcement learning for high dimensional linear quadratic systems," in *Advances in Neural Information Processing Systems*, 2012, pp. 2636–2644.
- [16] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, MA, 1995, vol. 1, no. 2.
- [17] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, "Optimality of fast matching algorithms for random networks with applications to structural controllability," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 770–780, 2017.
- [18] T. Lai and C. Wei, "Asymptotic properties of multivariate weighted sums with applications to stochastic regression in linear dynamic systems," *Multivariate Analysis VI*, pp. 375–393, 1985.
- [19] M. Green and J. B. Moore, "Persistence of excitation in linear systems," in *American Control Conference, 1985*. IEEE, 1985, pp. 412–417.
- [20] T. Lai and C. Wei, "Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters," *Journal of Multivariate Analysis*, vol. 13, no. 1, pp. 1–23, 1983.
- [21] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, "Finite time identification in unstable linear systems," *Automatica*, vol. 96, pp. 342–353, 2018.
- [22] T. Lai and C.-Z. Wei, "Extended least squares and their applications to adaptive control and prediction in linear systems," *IEEE Transactions on Automatic Control*, vol. 31, no. 10, pp. 898–906, 1986.
- [23] B. Nielsen, "Strong consistency results for least squares estimators in general vector autoregressions with deterministic terms," *Econometric Theory*, pp. 534–561, 2005.
- [24] —, "Order determination in general vector autoregressions," in *Time Series and Related Topics*. Institute of Mathematical Statistics, 2006, pp. 93–112.
- [25] T. Lai and C. Wei, "A note on martingale difference sequences satisfying the local Marcinkiewicz-Zygmund condition," *Bulletin of the Institute of Mathematics, Academia Sinica*, vol. 11, no. 1, p. 13, 1983.
- [26] B. Nielsen, "Singular vector autoregressions with deterministic terms: Strong consistency and lag order determination," *Discussion paper, Nuffield College*, 2009.