

On the Relationship of Optimal State Feedback and Disturbance Response Controllers

Runyu (Cathy) Zhang*, Yang Zheng**, Weiyu Li*, Na Li*

*School of Engineering and Applied Science, Harvard University

**Department of Electrical and Computer Engineering, University of California San Diego

Abstract: This paper studies the relationship between state feedback policies and disturbance response policies for the standard Linear Quadratic Regulator (LQR). For open-loop stable plants, we establish a simple relationship between the optimal state feedback controller $u_t = K_\star x_t$ and the optimal disturbance response controller $u_t = L_{\star,1}^{(H)} w_{t-1} + \dots + L_{\star,H}^{(H)} w_{t-H}$ with H -order. Here x_t, w_t, u_t stands for the state, disturbance, control action of the system, respectively. Our result shows that $L_{\star,1}^{(H)}$ is a good approximation of K_\star and the approximation error $\|K_\star - L_{\star,1}^{(H)}\|$ decays exponentially with H . We further extend this result to LQR for open-loop unstable systems, when a pre-stabilizing controller K_0 is available.

Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Linear Quadratic Regulator, State Feedback Control, Disturbance Response Control

1. INTRODUCTION

Linear quadratic regulator (LQR) is one of the most fundamental optimal control problems (Anderson and Moore, 2007). Its analytic solution and numerical methods have been well-established in the literature. Specifically, the infinite-time horizon LQR in the discrete time domain is formulated as follows:

$$\min_{\{u_t\}_{t=0}^{+\infty}} C := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=0}^{T-1} (x_t^\top Q x_t + u_t^\top R u_t + 2u_t^\top S x_t) \\ \text{subject to } x_{t+1} = Ax_t + Bu_t + w_t, \quad (1)$$

where $x_t \in \mathbb{R}^n$ is the system state, $u_t \in \mathbb{R}^m$ is the control input, $w_t \sim \mathcal{N}(0, I)$ is the external Gaussian process noise, and $Q \succ 0, R \succ 0$ and $S \in \mathbb{R}^{m \times n}$ are performance matrices. Throughout the paper, we make the standard assumption that $\begin{bmatrix} Q & S^\top \\ S & R \end{bmatrix} \succ 0$.

It is well-known that the optimal solution for (1) is a state-feedback controller (or policy) $u_t = K_\star x_t$, and the optimal gain $K_\star \in \mathbb{R}^{m \times n}$ can be computed via solving an algebraic Riccati equation (Anderson and Moore, 2007). The properties of the Riccati equation and its numerical solutions have been extensively studied (Kwakernaak and Sivan, 1969; Kleinman, 1968; Englar and Kalman, 1966). Most of these results are *model-based* and require the knowledge of system matrices A, B and the weight matrices Q, R, S . Motivated by the success of model-free policy optimization in reinforcement learning, many recent studies

(see review Hu et al. (2022) and references therein) have started to directly search an optimal policy by viewing the LQR cost $C(K)$ as a function of the policy parameterization $K \in \mathbb{R}^{m \times n}$. This formulation $C(K)$ is more suitable for *model-free policy optimization* but is generally *non-convex*. Thanks to special properties in the optimization landscape such as gradient dominance (Fazel et al., 2018), these methods can still find the optimal controller for the standard LQR problem despite of the non-convexity. However, these properties often fail to generalize to other linear optimal control problems such as sparse or structured LQR, partially observable systems, (Zheng et al., 2021; Hu et al., 2022), making it still challenging to develop policy optimization methods with provable convergence and optimality guarantees.

To avoid non-convexity, there are many other methods to re-parameterize the control policy such that the cost function becomes convex under new parameters. For the general output-feedback case, the classical approach is the Youla parameterization (Youla et al., 1976), and two recent approaches are system-level synthesis (SLS) (Wang et al., 2019) and input-output parameterization (IOP) (Furieri et al., 2019); also see Zheng et al. (2022) for two new parameterizations. Another specific idea is to parameterize the control policy as a function of the past disturbances w_t , known as the *disturbance response control* (DRC) (Goulart et al., 2006; Agarwal et al., 2019a). In particular, for open-loop stable plants, given a horizon $H \in \mathbb{N}$, we can use a DRC of the form

$$u_t = L_1^{(H)} w_{t-1} + \dots + L_H^{(H)} w_{t-H}, \quad (2)$$

where $L_k^{(H)} \in \mathbb{R}^{m \times n}, k = 1, \dots, H$ are policy parameters, and view the LQR cost in (1) as a function $C(L^{(H)})$ over $L^{(H)} := \{L_1^{(H)}, \dots, L_H^{(H)}\}$. It is not difficult to see that the closed-loop state and input evaluations in (1) become affine

* The work is supported by NSF AI institute: 2112085, ONR YIP: N00014-19-1-2217, NSF CNS: 2003111. The work of Yang Zheng is supported by NSF ECCS-2154650. Contact: runyuzhang@fas.harvard.edu, zhengy@eng.ucsd.edu, weiyuli@g.harvard.edu, nali@seas.harvard.edu. Corresponding author: Runyu (Cathy) Zhang.

in $L^{(H)}$, and the LQR cost $C(L^{(H)})$ is thus convex in terms of $L^{(H)}$. Thanks to the convexity, disturbance-based policy parameterizations appear to be easier and more suitable for model-free and online learning setups, which have indeed received increasing attention in online learning and control communities; see e.g., Simchowitz et al. (2020); Li et al. (2021); Agarwal et al. (2019a,b); Goulart et al. (2006). It is known that DRC-type controllers are closely related with other convex parameterizations such as the aforementioned Youla, SLS, and IOP.¹

Our contribution. In this paper, we study the relationship between the optimal state feedback policy K_* and the optimal DRC policy $L_*^{(H)}$ in (2). For open-loop stable plants (i.e., A in (1) is stable), it is not surprising that as the horizon H increases, the optimal performance $C(L_*^{(H)})$ will improve and converge to the optimal LQR performance $C(K_*)$. Similar analysis has appeared in Agarwal et al. (2019a) but in a slightly different online learning setting. Our paper presents an interesting and not obvious relationship: the first element $L_{*,1}^{(H)}$ in $L_*^{(H)}$ is a good approximation of K_* and the approximation error decays exponentially with increasing H (Theorem 1). Our result points out a possibly simple way of converting disturbance feedback controllers to state feedback controllers – instead of obtaining a state feedback controller using transfer functions², we can simply extract $L_{*,1}^{(H)}$ which is already a near optimal state feedback control gain. We further generalize the result to the LQR with an unstable open-loop system through considering DRC with a fixed pre-stabilizing controller K_0 (Corollary 1).

The proofs of our results are based on two intuitions: i) the optimal infinite disturbance response $L_*^{(\infty)}(z) = \sum_{k=1}^{+\infty} L_{*,k}^{(\infty)} z^{-k}$ induced by the optimal state feedback K_* has the exact equivalence $L_{*,1}^{(\infty)} = K_*$ (see (14)); ii) as the horizon $H \rightarrow +\infty$, the optimal $L_{*,1}^{(H)}$ should converge to $L_{*,1}^{(\infty)} = K_*$. In particular, proving (ii) is more technically involved, where we first derive a system of linear equations satisfied by $L_*^{(H)}$ (Lemma 1) and then show that $L_*^{(\infty)}$ is an approximate solution of the linear equations (Lemma 2 and Corollary 2).

Related Work. Some previous studies have built certain relationship between the state representation and other convex parameterizations, e.g., (Goulart et al., 2006; Nett et al., 1984). The setting that is most similar to our paper is Goulart et al. (2006), where the authors established an equivalence between the affine state feedback controllers and the affine disturbance feedback controllers. However, Goulart et al. (2006) only considered the finite time horizon problem and dynamic state feedback controllers, which is different from our setting in the infinite-time horizon. The relationship established in Goulart et al. (2006) is very different from our results, and the techniques involved in the proofs are different as well.

¹ For example, interested readers can find some explicit connections in the note: <https://zhengy09.github.io/course/notes/L3.pdf>.

² That is, solving the transfer function from state x to control u when controller u is in the DRC form (2).

Due to space limit, the detailed proofs and supplementary materials of this paper is deferred to the ArXive version Zhang et al. (2023)

2. PRELIMINARIES AND PROBLEM SETUP

In this paper, we consider the infinite-time horizon, time-invariant, discrete time LQR problem as defined in (1). Throughout this paper, we use $\|\cdot\|$ to denote the matrix ℓ_2 norm, and $\lambda_{\min}(X)$ to denote the smallest eigenvalue for a symmetric matrix X .

2.1 State-feedback controllers

When the plant (A, B) is stabilizable, the optimal solution to problem (1) is a linear state feedback controller $u_t = K_* x_t$ with (c.f. Anderson and Moore (2007))

$$K_* = -(R + B^\top P B)^{-1} (B^\top P A + S), \quad (3)$$

where the cost-to-go matrix P satisfies the algebraic Riccati equation

$$P = A^\top P A - (A^\top P B + S^\top)(R + B^\top P B)^{-1}(B^\top P A + S) + Q.$$

Thus, one natural perspective is to parameterize the policy using a single feedback matrix $K \in \mathbb{R}^{m \times n}$, i.e., $u_t = K x_t$, which we call as the state feedback representation. As stated in the introduction, this state feedback controller is easy to implement, yet it has one drawback that the LQR cost $C(K)$ becomes non-convex with respect to K .

2.2 Disturbance response controllers

Another approach to solve the LQR problem (1) is from a disturbance response perspective, which converts the problem to a convex optimization.

Open-loop stable systems. For open-loop asymptotically stable systems, i.e., the spectral radius of A is smaller than 1, we can consider a disturbance response controller (DRC) of the form (Agarwal et al., 2019a; Simchowitz et al., 2020):

$$u_t = L_1^{(H)} w_{t-1} + \cdots + L_H^{(H)} w_{t-H},$$

where $w_s = 0$ for $s < 0$. We can view the LQR cost as a function $C(L^{(H)})$ of the DRC matrices

$$L^{(H)} := \{L_1^{(H)}, \dots, L_H^{(H)}\}.$$

We then solve the following optimization problem to get the optimal DRC controller:

$$\begin{aligned} \min_{L^{(H)}} \quad & C(L^{(H)}) := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=0}^{T-1} x_t^\top Q x_t + u_t^\top R u_t + 2 u_t^\top S x_t \\ \text{subject to} \quad & x_{t+1} = Ax_t + Bu_t + w_t, \\ & u_t = L_1^{(H)} w_{t-1} + \cdots + L_H^{(H)} w_{t-H}. \end{aligned} \quad (4)$$

It is not difficult to see that (4) is a convex problem over $L^{(H)}$ since the closed-loop state x_t and input u_t all become affine in $L^{(H)}$.

In this paper, we are interested in establishing the relationship between the optimal state feedback policy K_* from (3) and the optimal DRC policy $L_*^{(H)}$ from (4). First of all, it is not surprising that as the horizon H increases, the optimal performance $C(L_*^{(H)})$ will improve and converge to the optimal LQR performance $C(K_*)$. Similar analysis has appeared in Agarwal et al. (2019a) but in a slightly

different online learning setting. For the self-completeness, we provide our own analysis for the LQR problem on how $C(L_\star^{(H)})$ approximates $C(K_\star)$ as H increases in Appendix B in Zhang et al. (2023). In addition to this relationship between $L_\star^{(H)}$ and K_\star , we will establish an interesting and not obvious relationship which directly connects the first element $L_{\star;1}^{(H)}$ in $L_\star^{(H)}$ with K_\star , which will be presented in Theorem 1.

Open-loop unstable systems. The above open-loop stability assumption is common for DRC type of controllers, e.g., (Agarwal et al., 2019b; Simchowitz et al., 2020). Our results can be easily extended to the unstable case by adding a fixed pre-stabilizing state feedback gain K_0 to the DRC, as presented below. For unstable system, instead of considering a DRC as in (2), we consider the following modified DRC with a fixed pre-stabilizing state feedback control gain K_0 :

$$u_t = K_0 x_t + L_1^{(H)} w_{t-1} + \cdots + L_H^{(H)} w_{t-H}. \quad (5)$$

Note that K_0 in (5) is a pre-fixed matrix and does not change when optimizing $C(L^{(H)})$. Given that K_0 stabilizes the system, i.e., $A + BK_0$ is stable, we could re-formulate equation (1) by defining an auxiliary variable

$$\bar{u}_t := u_t - K_0 x_t,$$

then the LQR problem could be re-formulated as:

$$\begin{aligned} \min_{L^{(H)}} C(L^{(H)}) &:= \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=0}^{T-1} x_t^\top \bar{Q} x_t + \bar{u}_t^\top R \bar{u}_t + 2\bar{u}_t^\top \bar{S} x_t \\ \text{s.t. } x_{t+1} &= \bar{A} x_t + \bar{B} \bar{u}_t + w_t, \\ \bar{u}_t &= L_1^{(H)} w_{t-1} + \cdots + L_H^{(H)} w_{t-H}, \end{aligned} \quad (6)$$

where

$$\begin{aligned} \bar{A} &:= A + BK_0, \quad \bar{S} := RK + S \\ \bar{Q} &:= Q + K_0^\top S + S^\top K_0 + K_0^\top R K_0. \end{aligned} \quad (7)$$

Note that $\bar{A} = A + BK_0$ is now a stable matrix. It can also be shown that $\lambda_{\min}(R - \bar{S} \bar{Q}^{-1} \bar{S}^\top) > 0$; see Lemma 7 in the Appendix in Zhang et al. (2023). Furthermore, it is not hard to verify that the optimal \bar{u}_t should satisfy $\bar{u}_t = \bar{K}_\star x_t$, where

$$\bar{K}_\star = K_\star - K_0.$$

Thus by considering the DRC with a pre-stabilizing K_0 , we could transform the LQR problem (1) with an unstable A to an LQR problem (6) with a stable \bar{A} .

3. MAIN RESULTS

In this section, we present our main results on the relationship between state feedback policies and disturbance response policies for LQR.

To characterize the stability degree, we introduce the following definition of exponential stability.

Definition 1. $((\tau, e^{-\rho})$ -stability). For $\tau \geq 1, \rho > 0$, we call a matrix A $(\tau, e^{-\rho})$ -stable if $\|A^k\| \leq \tau e^{-\rho k}$.

Note that for any open-loop asymptotically stable system, there exist some $\tau \geq 1, \rho > 0$ such that both A and $A - BK_\star$ are $(\tau, e^{-\rho})$ -stable, i.e.,

$$\|A^k\| \leq \tau e^{-\rho k}, \quad \|(A - BK_\star)^k\| \leq \tau e^{-\rho k}. \quad (8)$$

We will use τ, ρ later in our main result.

3.1 Open-loop stable systems

Our main result in this paper establishes a simple relationship between the optimal control gain K_\star from the algebraic Riccati equation (3) and the optimal $L_\star^{(H)}$ from (4). In particular, we can prove that $L_{\star;1}^{(H)}$ is a good approximation of K_\star , which is summarized in the theorem below.

Theorem 1. (Main Result). For open loop asymptotically stable systems, let K_\star be the optimal feedback gain in (3), and $L_\star^{(H)}$ be the optimal solution of (4). Then, we have

$$\|K_\star - L_{\star;1}^{(H)}\| \leq \frac{2\tau^3(\|B\|^2\|K_\star\|\|Q\| + \|B\|\|K_\star\|\|S\|)}{\lambda_{\min}(R - SQ^{-1}S^\top)(1 - e^{-2\rho})^{5/2}} e^{-H\rho},$$

where $L_{\star;1}^{(H)}$ denotes the first element in $L_\star^{(H)}$. Here τ, ρ are given in (8).

Theorem 1 suggests that as long as H is large enough, $L_{\star;1}^{(H)}$ is a good approximation of K_\star and the approximation error decays exponentially w.r.t H . Applying certain perturbation analysis arguments (Lemma B.1 in Tu and Recht (2019) and Lemma C.3 in Krauth et al. (2019)) suggests that for H that is sufficiently large, the system is also closed-loop stable and achieves near-optimal LQR cost with state feedback gain $L_{\star;1}^{(H)}$. Thus instead of implementing the disturbance feedback as

$$u_t = L_{\star;1}^{(H)} w_{t-1} + \cdots + L_{\star;H}^{(H)} w_{t-H}$$

(which is hard to implement because it needs computation and storage of history disturbances w_{t-k}), we could simply design a state feedback with gain $L_{\star;1}^{(H)}$, which is much easier to implement and still guarantees near-optimal performance. However, we would also like to point out that Theorem 1 heavily relies on the fact that the problem is unconstrained. It would be an interesting future direction to figure out whether a similar relationship still holds for constrained or distributed LQ control settings.

Remark 1. (Discussion on the stability assumption). We would like to emphasize that Theorem 1 only holds under the open-loop stability assumption, i.e., the spectral radius of A is smaller than 1. Specifically in the proof, one major lemma (Lemma 1) will not hold if A is not stable (see more discussion in Remark 2 after the lemma). In fact, without the stability assumption, for H that is not large enough, it can be shown that there is no H -order DRC that stabilizes the system (see Lemma 5 in the Appendix in Zhang et al. (2023)). Theorem 1 also suggests that the approximation error depends on the stability factors τ, ρ , the more stable the system is, the better the approximation error will be.

3.2 Extension to unstable systems

As discussed in the end of Section 2, we can transform the LQR problem (1) with an unstable A to an LQR problem (6) with a stable \bar{A} by considering the DRC with a pre-stabilizing K_0 :

$$u_t = K_0 x_t + L_1^{(H)} w_{t-1} + \cdots + L_H^{(H)} w_{t-H}.$$

Therefore, we can easily extend Theorem 1 to the unstable systems, as shown below,

Corollary 1. (Extension to the unstable case). Let K_\star be the optimal feedback gain in (3), and $\bar{K}_\star := K_\star - K_0$, If

both $\|A + BK_0\|$ and $\|A + BK_\star\|$ are $(\tau, e^{-\rho})$ -stable, the optimal solution $L_\star^{(H)}$ from (6) satisfies

$$\|\bar{K}_\star - L_{\star;1}^{(H)}\| \leq \frac{2\tau^3(\|B\|^2\|\bar{K}_\star\|\|\bar{Q}\| + \|B\|\|\bar{K}_\star\|\|\bar{S}\|)}{\lambda_{\min}(R - \bar{S}\bar{Q}^{-1}\bar{S}^\top)(1 - e^{-2\rho})^{5/2}} e^{-H\rho},$$

where \bar{S}, \bar{Q} are defined as in (7).

4. PROOF SKETCHES

In this section, we present the proof ideas for our main result in Theorem 1 by a thorough investigation of problem (4). We first introduce a result from Zhang et al. (2022) which shows that the solution to (4) can be explicitly expressed as the solution to a system of linear equations (Lemma 1). We next demonstrate that the disturbance response induced by the optimal control gain K_\star is an approximate solution to the linear equations (Lemma 2). Combining these two lemmas leads to the final result in Theorem 1.

4.1 Explicit solution for problem (4)

It is not difficult to see that problem (4) is an unconstrained quadratic optimization problem w.r.t. the variables $L_1^{(H)}, \dots, L_H^{(H)}$. Thus, it is expected that the optimal solution comes from a system of linear equations. Indeed, Zhang et al. (2022) has identified these equations, which are formally stated in the following lemma.

Lemma 1. (Zhang et al. (2022)) For open-loop asymptotically stable systems, the optimal $L^{(H)}$ of problem (4) satisfies a set of linear equations

$$M^{(H)}L^{(H)} + J^{(H)} = 0, \quad (9)$$

where $M^{(H)} \in \mathbb{R}^{Hn_u \times Hn_u}$ and $J^{(H)} \in \mathbb{R}^{Hn_u \times n_x}$ are

$$M^{(H)} := \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1H} \\ M_{21} & M_{22} & \cdots & M_{2H} \\ \vdots & \vdots & & \\ M_{H1} & M_{H2} & \cdots & M_{HH} \end{bmatrix}, \quad J^{(H)} := \begin{bmatrix} J_1 \\ \vdots \\ J_H \end{bmatrix}, \quad (10)$$

with submatrices $M_{km} \in \mathbb{R}^{n_u \times n_u}$, $J_k \in \mathbb{R}^{n_u \times n_x}$ defined as:

$$M_{km} := \begin{cases} B^\top GB + R, & k = m \\ B^\top GA^{k-m}B + SA^{k-m-1}B, & k > m, \\ B^\top (A^{m-k})^\top GB + B^\top (A^{m-k-1})^\top S^\top, & k < m \end{cases}, \quad (11)$$

$$J_k := B^\top GA^k + SA^{k-1}.$$

Here $G \in \mathbb{R}^{n_x \times n_x}$ is defined as:

$$G := \sum_{t=0}^{\infty} (A^t)^\top QA^t. \quad (12)$$

Remark 2. Note that Lemma 1 requires A to be exponentially stable; otherwise the matrix G in (12) is undefined. Since (4) is a quadratic problem with respect to $L^{(H)}$, it can be expected that the proof of Lemma 1 (see Zhang et al. (2022)) can be obtained by purely linear algebraic manipulation that writes out $C(L^{(H)})$ explicitly. In the process, there is one step that uses the Taylor series:

$$(I - z^{-1}A)^{-1} = \sum_{k=0}^{+\infty} z^{-k} A^k,$$

which only holds true if A is exponentially stable.

4.2 Relationship to optimal state feedback control gain

We first consider the following disturbance response controller induced by the optimal state feedback gain K_\star , which we denoted as $L_\star^{(\infty)}(z)$. That is, $L_\star^{(\infty)}(z)$ is the transfer function from the disturbance signal ω to the control u when the controller is $u(t) = K_\star x(t)$. It is straightforward to obtain the formulation of $L_\star^{(\infty)}(z)$,

$$\begin{aligned} L_\star^{(\infty)}(z) &= z^{-1}K_\star(I - z^{-1}(A + BK_\star))^{-1} \\ &= \sum_{k=1}^{+\infty} L_{\star;k}^{(\infty)} z^{-k}, \end{aligned} \quad (13)$$

where

$$L_{\star;k}^{(\infty)} := K_\star(A + BK_\star)^{k-1}, \quad k \geq 1. \quad (14)$$

Note that implementing the disturbance response controller with transfer function $L_\star^{(\infty)}(z)$ is equivalent to implementing the state feedback controller with control gain K_\star . To study the relationship of $L_{\star;1}^{(H)}$ and K_\star , it is natural to first study the relationship of $L_\star^{(H)}$ and $L_\star^{(\infty)}$. We establish the relationship by showing that $L_\star^{(\infty)}$ solves an ‘infinite dimension’ version of equation (9) that is satisfied by $L_\star^{(H)}$:

Lemma 2. The matrices $L_{\star;k}^{(\infty)}$ defined in (14) satisfy

$$\sum_{m=1}^{+\infty} M_{km} L_{\star;m}^{(\infty)} + J_k = 0, \quad \forall k \geq 1, \quad (15)$$

where M_{km}, J_k are defined in (11).

Lemma 2 is the key enabler of proving Theorem 1. For structural clearness, we defer the proof of Lemma 2 to Appendix A in Zhang et al. (2023). We would like to emphasize that the proof of Lemma 2 is technically involved and may be of independent interest.

Here we give an intuitive explanation of this lemma. The key insight is that $L_\star^{(\infty)}$ should satisfy an ‘infinite dimension’ version of equation (9) (i.e., $H \rightarrow +\infty$), which is exactly (15). Since $u_t = K_\star x_t$ is globally optimal among all control policies, it is expected that its induced disturbance response $L_\star^{(\infty)}$ solves the optimization problem (4) for $H \rightarrow +\infty$. Thus intuitively, it is expected that if we let the horizon H goes to infinity, the solution $L_\star^{(H)}$ for (9) will converge to the optimal $L_\star^{(\infty)}$. This is the reason we expect $L_\star^{(\infty)}$ to satisfy (15), which is an ‘infinite dimension’ version of (9). The detailed proof is provided in Appendix A in Zhang et al. (2023).

Lemma 2 immediately results in the following corollary.

Corollary 2. Define $L_{\star;1:H}^{(\infty)}$ as

$$L_{\star;1:H}^{(\infty)} = \left[L_{\star;1}^{(\infty)} \dots L_{\star;H}^{(\infty)} \right], \quad (16)$$

then we have

$$M^{(H)}L_{\star;1:H}^{(\infty)} + J^{(H)} = \mathcal{E},$$

where for all $1 \leq k \leq H$,

$$[\mathcal{E}]_k = \sum_{m=H+1}^{+\infty} B^\top (A^\top)^{m-k-1} (A^\top GB + S^\top) K_\star (A + BK_\star)^{m-1}.$$

Proof. From Lemma 2, we know that

$$\begin{aligned} & \sum_{m=1}^{+\infty} M_{km} L_{\star;m}^{(\infty)} + J_k = 0, \\ \implies & \sum_{m=1}^H M_{km} L_{\star;m}^{(\infty)} + J_k = - \sum_{m=H+1}^{+\infty} M_{km} L_{\star;m}^{(\infty)} = [\mathcal{E}]_k, \end{aligned}$$

which completes the proof.

4.3 Proof of Theorem 1

Proof. [of Theorem 1] From Lemma 1 and Corollary 2, we have

$$M^{(H)} L_{\star}^{(H)} + J^{(H)} = 0, \quad M^{(H)} L_{\star;1:H}^{(\infty)} + J^{(H)} = \mathcal{E}.$$

Subtracting these two equations leads to

$$M^{(H)} (L_{\star}^{(H)} - L_{\star;1:H}^{(\infty)}) = -\mathcal{E}.$$

Then, it is not difficult to see that

$$\begin{aligned} & (L_{\star}^{(H)} - L_{\star;1:H}^{(\infty)})^{\top} (L_{\star}^{(H)} - L_{\star;1:H}^{(\infty)}) = \mathcal{E}^{\top} (M^{(H)})^{-2} \mathcal{E} \\ & \preceq \frac{1}{\lambda_{\min}(M^{(H)})^2} \sum_{k=1}^H [\mathcal{E}]_k^{\top} [\mathcal{E}]_k \\ & \preceq \frac{1}{\lambda_{\min}(R - SQ^{-1}S^{\top})^2} \sum_{k=1}^H \|[\mathcal{E}]_k\|^2 I, \end{aligned}$$

where the last inequality uses the result: $\lambda_{\min}(M^{(H)}) \geq \lambda_{\min}(R - SQ^{-1}S^{\top})$, which can be found in Zhang et al. (2022) (Lemma 9).

We can upper bound the norm of $[\mathcal{E}]_k$ by

$$\begin{aligned} \|[\mathcal{E}]_k\| & \leq \sum_{m=H+1}^{+\infty} \|B\|^2 \|K_{\star}\| \|GA^{m-k}\| \|(A + BK_{\star})^{m-1}\| \\ & \quad + \|B\| \|K_{\star}\| \|S\| \|A^{m-k-1}\| \|(A + BK_{\star})^{m-1}\| \\ & \leq \frac{\|B\|^2 \|K_{\star}\| \|Q\| \tau^3}{1 - e^{-2\rho}} \sum_{m=H+1}^{+\infty} e^{-(2m-k-1)\rho} \\ & \quad + \tau^2 \|B\| \|K_{\star}\| \|S\| \sum_{m=H+1}^{+\infty} e^{-(2m-k-2)\rho} \\ & = \frac{\|B\|^2 \|K_{\star}\| \|Q\| \tau^3}{(1 - e^{-2\rho})^2} e^{-(2H-k+1)\rho} \\ & \quad + \frac{\|B\| \|K_{\star}\| \|S\| \tau^2}{1 - e^{-2\rho}} e^{-(2H-k)\rho}, \end{aligned}$$

where the second inequality uses the result (in Zhang et al. (2023) Appendix D)

$$\|GA^m\| \leq \frac{\tau^2 \|Q\| e^{-\rho m}}{1 - e^{-2\rho}}.$$

Thus

$$\begin{aligned} \sum_{k=1}^H \|[\mathcal{E}]_k\|^2 & \leq 2 \left(\frac{\|B\|^2 \|K_{\star}\| \|Q\| \tau^3}{(1 - e^{-2\rho})^2} \right)^2 \sum_{k=1}^H e^{-2(2H-k+1)\rho} \\ & \quad + 2 \left(\frac{\|B\| \|K_{\star}\| \|S\| \tau^2}{1 - e^{-2\rho}} \right)^2 \sum_{k=1}^H e^{-2(2H-k)\rho} \\ & \leq 2 \left(\frac{\|B\|^2 \|K_{\star}\| \|Q\| \tau^3}{(1 - e^{-2\rho})^2} \right)^2 \frac{1}{1 - e^{-4\rho}} e^{-2(H+1)\rho} \\ & \quad + 2 \left(\frac{\|B\| \|K_{\star}\| \|S\| \tau^2}{(1 - e^{-2\rho})^{3/2}} \right)^2 e^{-2H\rho} \\ & \leq 2 \left(\frac{\tau^3 (\|B\|^2 \|K_{\star}\| \|Q\| + \|B\| \|K_{\star}\| \|S\|)}{(1 - e^{-2\rho})^{5/2}} \right)^2 e^{-2H\rho}. \end{aligned}$$

Finally, we have

$$\begin{aligned} & (L_{\star;1}^{(H)} - K_{\star})^{\top} (L_{\star;1}^{(H)} - K_{\star}) = (L_{\star;1}^{(H)} - L_{\star;1}^{(\infty)})^{\top} (L_{\star;1}^{(H)} - L_{\star;1}^{(\infty)}) \\ & \preceq (L_{\star}^{(H)} - L_{\star;1:H}^{(\infty)})^{\top} (L_{\star}^{(H)} - L_{\star;1:H}^{(\infty)}) \\ & \preceq \frac{2}{\lambda_{\min}(R - SQ^{-1}S^{\top})^2} \\ & \quad \times \left(\frac{\tau^3 (\|B\|^2 \|K_{\star}\| \|Q\| + \|B\| \|K_{\star}\| \|S\|)}{(1 - e^{-2\rho})^{5/2}} \right)^2 e^{-2H\rho} I. \end{aligned}$$

This leads to

$$\|L_{\star;1}^{(H)} - K_{\star}\| \leq \frac{2\tau^3 (\|B\|^2 \|K_{\star}\| \|Q\| + \|B\| \|K_{\star}\| \|S\|)}{\lambda_{\min}(R - SQ^{-1}S^{\top}) (1 - e^{-2\rho})^{5/2}} e^{-H\rho},$$

which completes the proof.

5. NUMERICAL EXAMPLES

We consider the following randomly generated set of system matrices A, B, Q, R, S :

$$\begin{aligned} A &= \begin{bmatrix} -0.584 & 0.351 & 0.398 \\ -0.366 & -0.739 & 0.401 \\ 0.512 & 0.187 & -0.761 \end{bmatrix}, \quad B = \begin{bmatrix} -0.1659 \\ 1.7690 \\ -0.1603 \end{bmatrix}, \\ Q &= \begin{bmatrix} 9.549 & -2.660 & 6.993 \\ -2.660 & 2.702 & -1.599 \\ 6.993 & -1.599 & 8.282 \end{bmatrix}, \quad R = 2.593, \quad S = \begin{bmatrix} 0.043 \\ 0.206 \\ -1.964 \end{bmatrix}. \end{aligned}$$

Given the system matrices, we directly call builtin function `d1qr` in MATLAB System Control Toolbox to solve the optimal state feedback gain K_{\star} . The optimal DRC $L_{\star}^{(H)}$ is solved using (9). Figure 1 plots the approximation error $\|L_{\star;1}^{(H)} - K_{\star}\|$ as well as the cost difference $C(L_{\star;1}^{(H)}) - C(K_{\star})$ decays exponentially as H grows larger, which corroborates our theoretical finding in Theorem 1.

6. CONCLUSION

This paper has established a simple relationship between the optimal state feedback gain K_{\star} and the optimal disturbance response controller $u_t = L_{\star;1}^{(H)} w_{t-1} + \dots + L_{\star;H}^{(H)} w_{t-H}$. The result shows that $L_{\star;1}^{(H)}$ well approximates K_{\star} and the approximation error decays exponentially with H , which points out a possibly simpler way of converting disturbance feedback controllers to state feedback controllers.

REFERENCES

Agarwal, N., Bullins, B., Hazan, E., Kakade, S., and Singh, K. (2019a). Online control with adversarial disturbances.

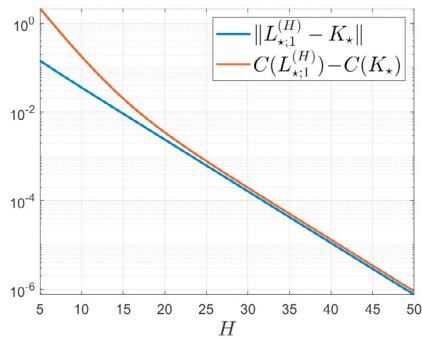


Fig. 1. Approximation error $\|L_{\star;1}^{(H)} - K_\star\|$ and performance difference $C(L_{\star;1}^{(H)}) - C(K_\star)$ decays exponentially with H

In *International Conference on Machine Learning*, 111–119. PMLR.

Agarwal, N., Hazan, E., and Singh, K. (2019b). Logarithmic regret for online control. *Advances in Neural Information Processing Systems*, 32.

Anderson, B.D. and Moore, J.B. (2007). *Optimal control: linear quadratic methods*. Courier Corporation.

Englar, T. and Kalman, R. (1966). A user’s manual for the automatic synthesis program/program c. Technical report.

Fazel, M., Ge, R., Kakade, S., and Mesbahi, M. (2018). Global convergence of policy gradient methods for the linear quadratic regulator. In *International Conference on Machine Learning*, 1467–1476. PMLR.

Furieri, L., Zheng, Y., Papachristodoulou, A., and Kamgarpour, M. (2019). An input–output parametrization of stabilizing controllers: Amidst youla and system level synthesis. *IEEE Control Systems Letters*, 3(4), 1014–1019.

Goulart, P.J., Kerrigan, E.C., and Maciejowski, J.M. (2006). Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4), 523–533.

Hu, B., Zhang, K., Li, N., Mesbahi, M., Fazel, M., and Başar, T. (2022). Towards a theoretical foundation of policy optimization for learning control policies. *arXiv preprint arXiv:2210.04810*.

Kleinman, D.Z. (1968). On an iterative technique for riccati equation computations. *IEEE Transactions on Automatic Control*, 13, 114–115.

Krauth, K., Tu, S., and Recht, B. (2019). Finite-time analysis of approximate policy iteration for the linear quadratic regulator. *Advances in Neural Information Processing Systems*, 32.

Kwakernaak, H. and Sivan, R. (1969). *Linear optimal control systems*, volume 1072. Wiley-interscience.

Li, Y., Das, S., Shamma, J., and Li, N. (2021). Safe adaptive learning-based control for constrained linear quadratic regulators with regret guarantees. *arXiv preprint arXiv:2111.00411*.

Nett, C., Jacobson, C., and Balas, M. (1984). A connection between state-space and doubly coprime fractional representations. *IEEE Transactions on Automatic Control*, 29(9), 831–832.

Simchowitz, M., Singh, K., and Hazan, E. (2020). Improper learning for non-stochastic control. In *Conference on Learning Theory*, 3320–3436. PMLR.

Zheng, Y., Furieri, L., Kamgarpour, M., and Li, N. (2022). System-level, input–output and new parameterizations of stabilizing controllers, and their numerical computation. *Automatica*, 140, 110211.

Zheng, Y., Tang, Y., and Li, N. (2021). Analysis of the optimization landscape of linear quadratic gaussian (lqg) control. *arXiv preprint arXiv:2102.04393*.