Beyond Gradient Dominance of Linear Quadratic Regulator: a Generalized Framework

Yue Sun and Maryam Fazel *

May 7, 2021

Abstract

Model-free methods in reinforcement learning seek optimal controllers for unknown dynamical systems by searching in the "policy" space directly. A recent line of research, starting with [1], aims to provide theoretical guarantees for such direct policy-update methods by exploring their performance in classical control settings, such as the infinite horizon linear quadratic regulator (LQR) problem. A key property that these analyses rely on is that the LQR cost satisfies the "gradient dominance" property with respect to the control gain. Gradient dominance guarantees that the optimal controller can be found by running gradient based algorithms on the LQR cost. The gradient dominance property has so far been verified on a case-by-case basis for LQR and several related control problems. In this paper, we make a connection between this line of work and classical convexification techniques based on linear matrix inequalities (LMIs). Using this, we propose a unified framework for showing that gradient dominance indeed holds for a broad class of control problems, such as problems with different costs and constraints, optimal control using system-level parameterization [2], and minimizing the L_2 gain in a closed loop system.

1 Introduction

Linear quadratic regulator (LQR) is one of the most well studied optimal control problem since being studied decades ago [3]. Consider the continuous time linear time-invariant dynamical system,

$$\dot{x} = Ax + Bu, \quad x(0) = x_0,$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^p$ is the input, and A, B describe the dynamics. The goal of optimal control is to determine a sequence of inputs u that minimizes some cost function that typically depends on state and input. In infinite horizon LQR problem, one minimizes

$$\int_0^\infty (x(t)^\top Qx(t) + u(t)^\top Ru(t))dt. \tag{1}$$

It is known that the optimal controller is linear in the state, called state feedback, where $u(t) = Kx(t)$ for a constant $K \in \mathbb{R}^{p\times n}$ [3]. This can be obtained by solving the algebraic Riccati equation (ARE) [4, 5]. A large number of works have studied the solution of ARE, including approaches based on iterative algorithms [6], algebraic solution methods [7], and semidefinite programming [8]. However, this approach is in sharp contrast to how one would typically minimize a cost function through gradient descent, as in the *policy gradient methods* (named by [9]) used in reinforcement learning settings [1].

In many practical cases, the system dynamics is unknown, and among the optimal control algorithms, there are two major types: **model based methods**, when the system is first identified and then a controller is trained, usually used when we have a good parameterization of the system; or **model free methods**, when the controller is directly trained by observing the cost (or loss) function, without characterizing the dynamics. Here one does not estimate the system parameters A, B, but requires only an estimate of the cost

*Y. Sun (`yuesun@uw.edu`) and M. Fazel (`mfazel@uw.edu`) are with Department of Electrical and Computer Engineering, University of Washington, Seattle, USA.

(1) as well as its gradient with respect to controller K when $u = Kx$. One then hopes that gradient descent with respect to K converges to the optimal controller K^*.

For LQR, model based methods are well studied. System identification dates back to [10, 11] etc., and recently [12, 13] gave sample complexity bounds for state-observed system, and [14, 15] for output-observed system. [16–18] describe a joint system identification and optimal control approach based on the estimate of the system.

The policy gradient method proposed by [9] is a model free method, and is more recently reviewed by [19, 20]. Traditionally, the LQR problem is solved using the associated Riccati equation, using the system parameters A, B (thus falling into the category of model based method). It is generally not recommended to directly compute the loss using this cost function without solving the associated Riccati equation, since it can be shown that minimizing the quadratic LQR cost as a function of K is a nonconvex problem.

Recently people have witnessed the *empirical* success of first order methods in solving nonconvex reinforcement learning problems. LQR is probably one of the most well-studied reinforcement learning problem thus it is of interest whether we can still get a *theoretical* convergence guarantee for gradient descent on the nonconvex objective. [1] proved the convergence for discrete-time LQR and [21] showed a similar result for the continuous-time case. Later [22, 23] gave an alternative analysis by Riccati and Lyapunov equations for both discrete and continuous time LQR, showing they can be solved by gradient descent on K, and [24] extended it to accelerated gradient descent. [25] further showed similar results for zero sum Linear-quadratic games. [26] studied the convergence of gradient descent on $\mathcal{H}_2$ control with $\mathcal{H}_\infty$ constraint, and showed that gradient descent implicitly makes the controller robust. [27] showed the convergence for distributed control, where a convexification is legit with the quadratic invariance (QI) assumption. We will also review another direction of model free method though we don't theoretically study it in this paper: thanks to the development of reinforcement learning, there are works directly solving LQR by dynamical programming type methods. They can be implemented as online algorithms such as [28, 29]. Q-learning is used in [30, 31] to solve LQR, which is also a popular method in reinforcement learning.

Traditionally, convexification approaches such as Youla parametrization, Q-parameterization, or the more recent System Level Synthesis (SLS) have allowed the reformulation of certain control design problems as semidefinite programs. In this paper, we are interested to see if these methods can help us distill the essence of the gradient dominance property of the original control problem that is nonconvex in K.

Note that when the dynamical system is nonlinear, i.e., the dynamics $Ax + Bu$ is replaced by a nonlinear function of state and input, the optimal control problem is much more difficult. The classical solution is via dynamic programming, or solving Bellman equations [32]. Recently, the development of deep learning methods and their use with dynamic programming have led to empirical success in control of complex systems. Yet it is still mysterious how deep learning models work in this context, and recent theoretical studies have focused on linear systems in hope of providing insights into more complex cases.

Contributions: In this paper, we will build a bridge between convexification methods and nonconvex policy gradient methods. We will use a mapping between the landscape of convex and nonconvex objectives, and use this mapping to prove the gradient dominance property of the nonconvex objective under some reasonable assumptions.

Our result is quite general—we show that continuous time LQR is a special case that our theorem applies to, and we generalize the guarantees provided by this method to a range of optimal control problems. Thus for all these problems, if one wants to understand whether the (probably nonconvex) loss with respect to controller parameter K can be minimized by direct policy-update methods (first-order optimization methods that update K), one can directly check if it is covered by our main theorem and does not necessarily need to give a standalone proof. Moreover, the approach is far more transparent and insightful than the existing case-by-case analysis. Also, as discussed in [1], theoretical guarantees for first-order methods naturally lead to guarantees for the more practical zero-order optimization or sampling-based methods, which do not need access to the gradient of the cost with respect to K.

The rest of this paper is structured as follows. Sec. 2 reviews the continuous time LQR problem, and the convexification method for solving it. [21] analyzes the convergence of policy gradient method in this regime, and we briefly review their results. Sec. 3 presents our main result, where we distill the critical property of the continuous time LQR problems and claim the gradient dominance property of the nonconvex loss. Sec. 4 lists more examples of optimal control problems covered by our main result. Sec. 5 proposes the sketch of our proof, mainly the intuition how to build the bridge between two methods. We place the detailed proof in the appendix.

2 Review of convexification method for continuous time LQR

Convexification method (e.g., solving optimal control by linear matrix inequalities (LMI) in [33]) is widely used in optimal control problems, and here we discuss its application for continuous time LQR [21]. Define a continuous time linear time invariant system

$$\dot{x} = Ax + Bu,\ x(0) = x_0, \tag{2}$$

where x is state and u is input signal, x_0 comes from an initial distribution such that $\mathbf{E}(x_0 x_0^\top) = \Sigma \succ 0$ (e.g., average LQR cost over a set of linearly independent initial states [22, Sec 3.3]), one considers minimizing the LQR loss

$$\min_{u(t)} f(u(t)) := \mathbf{E}_{x_0} \int_0^\infty (x(t)^\top Q x(t) + u(t)^\top R u(t)) dt \tag{3}$$

where Q, R are positive definite matrices. It is known that, the input signal that minimizes the loss function $f(u)$ is a state feedback controller

$$u = K^* x = -R^{-1} B P x, \tag{4}$$

$$AP + PA^\top + Q - PBR^{-1}BP = 0. \tag{5}$$

Note that once we know the state feedback controller is static, we can write loss $f(u(t))$ as $\mathcal{L}(K)$ which is a function of K instead, and search only in static state feedback controllers.

One approach of finding K^* is to solve the Riccati equations (4,5) to get K^*. An alternative approach is reparameterization [21]. The review starts from the Lyapunov equation. Suppose the initial state satisfies $\mathbf{E}(x_0 x_0^\top) = \Sigma \succ 0$, and $\dot{x} = Ax$. Then the Lyapunov equation is written as

$$AG + GA^\top + \Sigma = 0 \tag{6}$$

where $G \in \mathbb{R}^{n\times n}$ is invertible. In our setup (2), we use a state feedback controller $u = Kx$, thus we have $\dot{x} = (A + BK)x$. We denote the set of stabilizing controllers as $\mathcal{S}_{K,\mathrm{sta}}$, which is defined as

$$\mathcal{S}_{K,\mathrm{sta}} = \{K :\ A + BK \text{ is Hurwitz}\}.$$

Denote $L = KG \in \mathbb{R}^{p\times n}$, the Lyapunov equation becomes

$$\begin{aligned} &(A + BK)G + G(A + BK)^\top + \Sigma = 0. \\ \Rightarrow\ & AG + GA^\top + BL + L^\top B^\top + \Sigma = 0 \end{aligned}$$

Let $\mathcal{A}(G) = AG + GA^\top,\ \mathcal{B}(L) = BL + L^\top B^\top$, assume $\mathcal{A}$ is invertible, they have the relation

$$\mathcal{A}(G) + \mathcal{B}(L) + \Sigma = 0 \tag{7}$$

Now we define the Grammian matrix

$$G = \int_0^\infty e^{(A+BK)t} \Sigma e^{(A+BK)^\top t} dt.$$

We are interested in the loss function $\mathcal{L}(K)$ when $K \in \mathcal{S}_K$, which associates with (3) by inserting $u(t) = Kx(t)$. Let P_K solve

$$P_K = A^\top P_K + P_K A + Q + K^\top R K$$

$\mathcal{L}(K)$ can be expressed by $\mathcal{L}(K) = \mathbf{Tr}(P_K \Sigma)$. $\mathcal{L}(K)$ can also be expressed as[1]

$$\mathcal{L}(K) = \mathbf{Tr}((Q + K^\top R K)G).$$

[1] If K is not a stabilizing controller, $\mathcal{L}(K) = +\infty$.

We define $K = LG^{-1}$. The loss function $\mathcal{L}(K)$ can also be expressed as a function of G, L such that

$$\mathcal{L}(K) = f(L, G) = \mathbf{Tr}(QG + L^\top RLG^{-1}). \tag{8}$$

One can construct a bijection from G, L to K, and prove that, if we minimize $f(L, G)$ subject to (7), the optimizer G^*, L^* will map to the optimal K^*, and this minimization problem is convex, so we can solve it by convex optimization algorithms.

Recently, model-free nonconvex optimization algorithms are widely used in machine learning, so it is also of interest whether we can run gradient based algorithm in K space without reparameterization, which means that, if we run gradient based algorithm and end up with $\nabla_K \mathcal{L}(K) = 0$, is K the global optimum K^*? [21] suggests that, if we initialize K in the set of stabilizing controllers, then the gradient flow (continuous time flow) converges to K^* in linear rate. Moreover, they show that gradient descent (discrete iterations) method with proper step size solves the continuous LQR.

This work is a generalization of [21]. We extend the approach to much more general sets of problems which covers the original LQR. We will prove that, for a few optimal control problem satisfying some assumptions, all the stationary points of the nonconvex objective function are global minimum. Based on that, we argue that the nonconvex optimization problem in K space can be solved by a direct descent method applied in the policy space K (also known as policy gradient methods).

Continuous time LQR with convexified formulation: Suppose the dynamics and loss are (2) and (3), and let $\mathbf{E}(x_0 x_0^\top) = \Sigma \succ 0$. Let state feedback controller be K so that $u(t) = Kx(t)$. The optimal control problem is

$$\min_K \ \mathcal{L}(K), \quad \text{s.t.,} \quad K \text{ stabilizes} \tag{9}$$

where $\mathcal{L}(K)$ is the cost in (1) with $u = Kx$. It can be convexified as the following problem where $K = LG^{-1}$.

$$\min_{L,G,Z} \ f(L, G, Z) := \mathbf{Tr}(QG + ZR) \tag{10a}$$

$$\text{s.t., } \mathcal{A}(G) + \mathcal{B}(L) + \Sigma = 0, \ G \succ 0, \tag{10b}$$

$$\begin{bmatrix} Z & L^\top \\ L & G \end{bmatrix} \succeq 0 \tag{10c}$$

3 Main result

We present our main result in Theorem 1. It is described as a pair of problems satisfying Assumptions 1,2,3 and is not necessarily continuous time LQR. In Sec. 4 we will review more examples showing the generality.

Extracting the math properties of continuous time LQR: In the following paragraph, we make the description of the pair of problems (9) and (10) abstract, which will later cover more instances. Suppose we consider the problems

$$\min_K \quad \mathcal{L}(K), \quad \text{s.t., } K \in \mathcal{S}_K \tag{11a}$$

and

$$\min_{L,G,Z} \quad f(L, G, Z), \quad \text{s.t.,} \quad (L, G, Z) \in \mathcal{S} \tag{12}$$

$\mathcal{S}_K, \mathcal{S}$ are two sets. We allow special cases when (12) depends only on L, G,

$$\min_{L,G} \quad f(L, G), \quad \text{s.t.,} \quad (L, G) \in \mathcal{S} \tag{13}$$

We distill three critical properties of the two problems (11) and (12), which results in Theorem 1 and covers more examples in Sec. 4.

Assumption 1. *The feasible set $\mathcal{S}$ is convex in all of the variables (L, G, Z).*

Assumption 2. *The cost function $f(L,G,Z)$ is convex, finite and differentiable in $(L,G,Z) \in \mathcal{S}^2$.*

Assumptions 1 and 2 mean the second problem is convex. Next we extract the property of the connection between (9) and (10), and give an abstract description of the assumptions for (11) and (12).

Assumption 3.

1. *(Bijection between sets from two spaces) Let G be invertible, $K = LG^{-1}$ define a bijection $K \leftrightarrow (L,G)$[3]. For any such bijection $K \leftrightarrow (L,G)$, $\exists Z$, such that $(L,G,Z) \in \mathcal{S}$.*
2. *(Equivalence of functions) Choose a controller $K \in \mathcal{S}_K$ and it maps to $(L,G) \in \mathcal{S}$. Then $\mathcal{L}(K) = \min_Z f(L,G,Z)$ subject to $(L,G,Z) \in \mathcal{S}$.*

We denote $\mathscr{Z}(L,G) \in \mathrm{argmin}_Z f(L,G,Z)$ subject to $(L,G,Z) \in \mathcal{S}$ (if there are multiple minimizers we pick any one).

Theorem 1. *We consider the problems* (11) *and* (12)*, and we require Assumptions 1,2,3. Define K^* as the global minimum of $\mathcal{L}(K)$ in feasible region. Then we have*

1. *if f is convex, the gradient satisfies*[4][5][6]

$$\|\nabla\mathcal{L}(K)\|_F \gtrsim \mathcal{L}(K) - \mathcal{L}(K^*). \tag{14}$$

2. *if f is μ-strongly convex, the gradient satisfies*

$$\|\nabla\mathcal{L}(K)\|_F \gtrsim (\mu(\mathcal{L}(K) - \mathcal{L}(K^*)))^{1/2}. \tag{15}$$

The hidden constants in right hand side are in Appendix I.

In Assumption 3, we assume that there is a bijection between K and (L,G). The following assumption is a relaxed version that allows multiple (L,G) mapping to K. With this mapping, we need to find an alternative description for "equivalence of functions" in Assumption 2. In this case, we assume $\mathcal{L}(K)$ is exactly the minimum of for all $f(L,G,Z)$ such that $K \leftrightarrow (L,G)$.

Assumption 4. *Let G be invertible in $\mathcal{S}$. Assume we can express $\mathcal{L}(K)$ as:*

$$\begin{aligned}\mathcal{L}(K) = \min_{L,G,Z}\ & f(L,G,Z)\\ s.t.,\ & (L,G,Z) \in \mathcal{S},\ LG^{-1} = K.\end{aligned}$$

Theorem 1 also holds with Assumptions 1,2,4.[7]

Our main theorem suggests that, when the original nonconvex optimization problem can be mapped to a convex optimization problem that satisfies Assumptions 1,2,3 or 1,2,4, all stationary points of the nonconvex objective are global minima. So if we can evaluate the gradient of nonconvex objective and run gradient descent algorithm, the iterates converge to the optimal controller.

4 Optimal control problems covered by main theorem

In order to reach the conclusion, Theorem 1 requires an optimal control problem (11), its convexified form (12) and a few assumptions. This is an abstract and general description that does not need the exact continuous time LQR formulation in Sec. 2. We can easily check that the continuous time LQR satisfies the Assumptions 1,2,3, thus we can directly apply Theorem 1 to argue that the continuous time LQR cost $\mathcal{L}(K)$ satisfies (14).

[2] We can extend our result to the case where $f(L,G,Z)$ has well defined one-sided directional derivative, then the left hand side of (14) and (15) becomes norm of one-sided directional derivative in direction $K - K^*$. With an abuse of notation, we can denote it as $|\nabla\mathcal{L}(K)[\frac{K-K^*}{\|K-K^*\|_F}]|$ which is one-sided.

[3] Note that generally $K = LG^{-1}$ cannot guarantee a bijection. However bijection is possible with extra constraint $(L,G) \in \mathcal{S}$.

[4] $a \gtrsim b$ means there exists constant C that $a \geq Cb$.

[5] Known as Lojasiewicz inequality [34].

[6] We always consider the gradient restricted to the feasible region.

[7] We emphasize the special case for Assumptions 1,2,3 since it is true for the important example – continuous LQR, and easy to be illustrated in Sec. 5.

Below, we will start from continuous time LQR and then propose more examples, showing that Theorem 1 covers a wide range of optimal control problems. This shows the **generality** of Theorem 1. If one encounters optimal control problems from the following set, and moreover, if one encounters a new problem that can be convexified, we hope that one can check if the problem satisfies the assumptions in Theorem 1. If so, one can directly claim that the original loss function has no spurious local minimum and can be optimized by policy gradient method., which can be applied in model free setup.

4.1 Continuous time LQR with its variants.

We will first present the examples related to continuous time LQR, which is the first motivation of our paper. They satisfy Assumption 1,2,3.

1. Continuous time LQR problem: See (9), (10).
2. Budget on total energy of input, which is the following constraint

$$\mathbf{E}_{x_0} \int_0^\infty \|u(t)\|^2 dt = \mathbf{Tr}(Z) \le e_0.$$

   It can be added to (10) as an extra constraint.
3. Different loss: Define positive numbers e_i where we assume the energy of the ith channel of the input is less than e_i, in other words,

$$e_i \ge \int_0^\infty u_i^2(t) dt$$

   Instead of the LQR cost with constraints, we can use a barrier function for input energy

$$\begin{aligned}\mathcal{L}(K) = \mathbf{E}_{x_0} \Bigg(& \int_0^\infty (x(t)^\top Q x(t) + u(t)^\top R u(t)) dt \\ & - \sum_{i=1}^p \lambda_i \log(e_i - \int_0^\infty u_i^2(t) dt) \Bigg) \\ f(L, G, Z) = \mathbf{Tr}(QG + ZR) & - \sum_{i=1}^p \lambda_i \log(e_i - Z_{ii}),\end{aligned}$$

   where the log barrier function can be used to enforce the constraints (as is typically used in iterative methods for solving constrained optimization problems). In this case, with the extra $\sum_{i=1}^p \lambda_i \log(e_i - Z_{ii})$ term, we enforce that $e_i - Z_{ii} > 0$.

4.2 Minimizing L_2 gain

We quote from [33] the problem of minimizing the L_2 gain with static state feedback controller K. As discussed in [33, 6.3.2], this problem has an associated convex optimization problem and we can show it satisfies Assumption 1,2,4.

We consider minimizing the L_2 gain of a closed loop system. The continuous time linear dynamical system is

$$\dot{x} = Ax + Bu + B_w w, \ y = Cx + Du \tag{16}$$

For any signal z, denote

$$\|z\|_2 := \left(\int_0^\infty \|z(t)\|_2^2 dt \right)^{1/2}$$

Suppose we use a state feedback controller $u = Kx$, and aim to find the optimal controller K^* that minimizes the L_2 gain. We minimize the squared L_2 gain as

$$\min_K \ \mathcal{L}(K) := (\sup_{\|w\|_2=1} \|y\|_2)^2.$$

This problem can be further reformulated as [33, Sec 7.5.1]

$$\begin{aligned}&\min_{L,G,\gamma} \ f(L,G,\gamma) := \gamma, \ \text{s.t.,}\\ &\begin{bmatrix} AG + GA^\top + BL + L^\top B^\top + B_w B_w^\top & (CG+DL)^\top \\ CG + DL & -\gamma I \end{bmatrix}\\ &:= M(L,G,\gamma) \preceq 0. \end{aligned} \tag{17}$$

The minimum L_2 gain is $\sqrt{\gamma^*}$ and $K^* = L^* G^{*-1}$. We will show in the appendix that the parameters K and (L, G, γ), with loss $\mathcal{L}(K)$ and $f(L, G, \gamma)$, satisfy Assumptions 1,2,4. Thus we can claim that all stationary points of $\mathcal{L}(K)$ are global minimum.

[33, 6.3.2] suggests that L_2 gain is also the $\mathcal{H}_\infty$ norm of transfer function, so it covers the instances in [26]. We discuss it in Append. B.1.

4.3 Dissipativity

We use the dynamical system defined in (16). The notion of dissipativity can be found in [33, Sec. 6.3.3, 7.5.2]. Our goal is to maximize the dissipativity, which is defined and formulated as with a parameterization that convexified in [33, Sec. 6.3.3, 7.5.2].

The dissipativity is defined as

$$\int_0^T w^\top y - \eta w^\top w dt \geq 0, \ \forall T > 0.$$

Same as the last example, we use a state feedback controller K and let it be factorized as LG^{-1}. We can maximize the dissipativity η as a function of K. From [33, Sec. 7.5.2], we maximize η subject to the dissipativity constraint (18),

$$\begin{bmatrix} AG + GA^\top + BL + L^\top B^\top & B_w - GC^\top \\ B_w^\top - CG & 2\eta I - (D + D^\top) \end{bmatrix} \preceq 0. \tag{18}$$

With similar reasoning as in Sec. 4.3, the assumptions for Theorem 1 holds. Thus we can claim that all stationary points of $\mathcal{L}(K)$ are global minimum.

4.4 System level synthesis (SLS) for finite horizon time varying discrete LQR

This problem and its convexified form are introduced in [2]. It satisfies Assumption 1,2,3. We consider the following linear dynamical system

$$x(t+1) = A(t)x(t) + B(t)u(t) + w(t) \tag{19}$$

over a finite horizon $0, \ldots T$. Let the state be x and the input be u. Define

$$\begin{aligned} &X = \begin{bmatrix} x(0) \\ \ldots \\ x(T) \end{bmatrix}, \ U = \begin{bmatrix} u(0) \\ \ldots \\ u(T) \end{bmatrix}, \\ &W = \begin{bmatrix} x(0) \\ w(0) \\ \ldots \\ w(T-1) \end{bmatrix}, Z = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ I & 0 & \ldots & 0 & 0 \\ 0 & I & \ldots & 0 & 0 \\ \ldots & & & & \\ 0 & 0 & \ldots & I & 0 \end{bmatrix}, \\ &\mathcal{A} = \text{diag}(A(0), \ldots, A(T-1), 0), \\ &\mathcal{B} = \text{diag}(B(0), \ldots, B(T-1), 0) \end{aligned}$$

Now we consider the time varying controller K that links state and input as

$$u(t) = \sum_{i=0}^{t} K(t, t-i)x(i) \tag{20}$$

and let

$$\mathcal{K} = \begin{bmatrix} K(0,0) & 0 & ... & 0 \\ K(1,1) & K(1,0) & ... & 0 \\ ... & & & \\ K(T,T) & K(T,T-1) & ... & K(T,0) \end{bmatrix}$$

We will minimize some loss function with the constraint. For example, in the discrete time LQR regime, let the input be (20) and define (More examples of nonquadratic cost in [2, Sec 2.2])

$$\mathcal{L}(\mathcal{K}) = \sum_{t=0}^{T} x(t)^\top Q(t)x(t) + u(t)^\top R(t)u(t), \tag{21}$$

here $Q(t), R(t) \succeq 0$. We will minimize $\mathcal{L}(\mathcal{K})$ where $\mathcal{K}$ is the variable.

Convexification: The dynamics (19) can be written as

$$X = Z\mathcal{A}X + Z\mathcal{B}U + W = Z(\mathcal{A} + \mathcal{B}\mathcal{K})X + W$$

We define the mapping from W to X, U by

$$\begin{bmatrix} X \\ U \end{bmatrix} = \begin{bmatrix} \Phi_X \\ \Phi_U \end{bmatrix} W.$$

where Φ_X, Φ_U are block lower triangular. There is a constraint on Φ_X, Φ_U:

$$\begin{bmatrix} I - Z\mathcal{A} & -Z\mathcal{B} \end{bmatrix} \begin{bmatrix} \Phi_X \\ \Phi_U \end{bmatrix} = I. \tag{22}$$

It is proven in [2, Thm 2.1] that $\mathcal{K} = \Phi_U \Phi_X^{-1}$. $\mathcal{K}$ and Φ_X, Φ_U is a bijection given Φ_X, Φ_U satisfying (22).

Let $\mathcal{Q} = \mathrm{diag}(Q(0), ..., Q(T))$, $\mathcal{R} = \mathrm{diag}(R(0), ..., R(T))$, the LQR loss with $x(0) \sim \mathcal{N}(0, \Sigma)$ and no noise is

$$f(\Phi_X, \Phi_U) = \left\| \mathrm{diag}(\mathcal{Q}^{1/2}, \mathcal{R}^{1/2}) \begin{bmatrix} \Phi_X(:,0) \\ \Phi_U(:,0) \end{bmatrix} \Sigma^{1/2} \right\|_F^2$$

the LQR loss with $x(0), w(t)$ being i.i.d from $\mathcal{N}(0, \Sigma)$ is

$$f(\Phi_X, \Phi_U) = \left\| \mathrm{diag}(\mathcal{Q}^{1/2}, \mathcal{R}^{1/2}) \begin{bmatrix} \Phi_X \\ \Phi_U \end{bmatrix} \Sigma^{1/2} \right\|_F^2.$$

If we do

$$\min_{\mathcal{K}} \ \mathcal{L}(\mathcal{K}), \ \mathcal{K} \text{ is lower left triangular}$$

with the above two models of $w(t)$, both can be minimized with constraint (22):

$$\min_{\Phi_X, \Phi_U} \ f(\Phi_X, \Phi_U), \text{ s.t., } \begin{bmatrix} I - Z\mathcal{A} & -Z\mathcal{B} \end{bmatrix} \begin{bmatrix} \Phi_X \\ \Phi_U \end{bmatrix} = I,$$
$$\Phi_X, \Phi_U \text{ are lower left triangular}$$

This problem is convex and satisfy Assumption 1, 2. [2, Thm 2.1] suggests the relation between $\mathcal{L}$ and f satisfying the Assumption 3 for Theorem 1. With Theorem 1, we can argue that all stationary points of $\mathcal{L}(\mathcal{K})$ are global minimum.

Generally SLS can be applied to infinite horizon where the variables K, Φ_X, Φ_U are transfer functions. We discuss it in Appendix III.

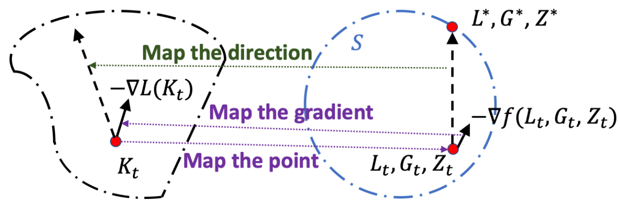

Figure 1: Mapping between nonconvex and convex landscapes. Suppose we run gradient descent at iteration t, for any controller K, we can map it to L, G, Z in the other parameterized space. and then we map the direction $(L^*, G^*, Z^*) - (L, G, Z)$ and the gradient $\nabla f(L, G, Z)$ back to the original K space. Since in (L, G, Z) space the loss is convex, then $\langle \nabla f(L, G, Z), (L^*, G^*, Z^*) - (L, G, Z) \rangle < 0$. We prove that similar correlation holds for the nonconvex objective.

5 Proof Sketch

We put the full proof of Theorem 1 in the appendix, and give a sketch of the proof in this section. We illustrate the idea in Figure 1, which, on the high level, maps the loss function in original space of controller K where the loss is nonconvex, and the paraterized space with L, G, Z where the loss is convex.

For simplicity, we sketch the proof using Assumptions 1,2,3. Assumptions 4 is weaker and requires a little more work, but the proof is similar. For any point K, we can find a point (L, G, Z) in the parameterized space. If it is not the optimizer, we can find the line segment linking (L, G, Z) and the optimizer (L^*, G^*, Z^*). Note that the optimization problem is convex in this space so that $\langle \nabla f(L, G, Z), (L^*, G^*, Z^*) - (L, G, Z) \rangle < 0$, in other words, the directional derivative of f is not 0[8]. Then with the help of our assumptions, we can map the directional derivative back to the original K space, and show that the directional derivative in $\mathcal{L}(K)$ is not 0.

We emphasize that, although Theorem 1 requires some assumptions, they are in the perspective of optimization theory and we do not dive into the specific property of control aspect (e.g., the traditional control theory). On the contrary, as it's a simple approach that only uses optimization aspect, it has its own restrictions such as suboptimality in convergence rate, and not being able to generalize to a few other optimal control problems. As a future direction, we hope that the insight of bridging the nonconvex and convexified landscapes can stimulate more rigorous analysis of a variety of algorithms and their convergence rates. Another future direction is to extend to more optimal control problems, such as infinite horizon design of dynamic controllers. The optimization is over the (infinite dimensional) transfer function characterizing the controller, and for implementation it can be further approximated by finite-dimensional representations such as FIR.

References

[1] M. Fazel, R. Ge, S. M. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for linearized control problems," *arXiv preprint arXiv:1801.05039*, 2018.

[2] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, "System level synthesis," *Annual Reviews in Control*, vol. 47, pp. 364–393, 2019.

[3] R. E. Kalman *et al.*, "Contributions to the theory of optimal control," *Bol. soc. mat. mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[4] R. F. Stengel, *Optimal control and estimation.* Courier Corporation, 1994.

[5] G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach.* Springer Science & Business Media, 2013, vol. 36.

[6] G. Hewer, "An iterative technique for the computation of the steady state gains for the discrete optimal regulator," *IEEE Transactions on Automatic Control*, vol. 16, no. 4, pp. 382–384, 1971.

[8] We can actually give an upper bound which is negative and depends on $f(L^*, G^*, Z^*) - f(L, G, Z)$. For simplicity of illustration we omit the exact value and readers can find details in the proof.

[7] P. Lancaster and L. Rodman, *Algebraic riccati equations.* Clarendon press, 1995.

[8] V. Balakrishnan and L. Vandenberghe, "Semidefinite programming duality and linear time-invariant systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 1, pp. 30–41, 2003.

[9] S. J. Bradtke, B. E. Ydstie, and A. G. Barto, "Adaptive linear quadratic control using policy iteration," in *Proceedings of 1994 American Control Conference-ACC'94*, vol. 3. IEEE, 1994, pp. 3475–3479.

[10] J. A. Cadzow, "Signal enhancement-a composite property mapping algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 1, pp. 49–62, 1988.

[11] L. Ljung, "System identification: theory for the user," *PTR Prentice Hall, Upper Saddle River, NJ*, pp. 1–14, 1999.

[12] M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," in *Conference On Learning Theory*, 2018, pp. 439–473.

[13] T. Sarkar and A. Rakhlin, "Near optimal finite time identification of arbitrary linear dynamical systems," *arXiv preprint arXiv:1812.01251*, 2019.

[14] S. Oymak and N. Ozay, "Non-asymptotic identification of lti systems from a single trajectory," *arXiv preprint arXiv:1806.05722*, 2018.

[15] T. Sarkar, A. Rakhlin, and M. A. Dahleh, "Finite-time system identification for partially observed lti systems of unknown order," *arXiv preprint arXiv:1902.01848*, 2019.

[16] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *arXiv preprint arXiv:1710.01688*, 2017.

[17] S. Dean, S. Tu, N. Matni, and B. Recht, "Safely learning to control the constrained linear quadratic regulator," in *2019 American Control Conference (ACC).* IEEE, 2019, pp. 5582–5588.

[18] H. Mania, S. Tu, and B. Recht, "Certainty equivalent control of lqr is efficient," *arXiv preprint arXiv:1902.07826*, 2019.

[19] S. M. Kakade, "A natural policy gradient," in *Advances in neural information processing systems*, 2002, pp. 1531–1538.

[20] A. Rajeswaran, K. Lowrey, E. V. Todorov, and S. M. Kakade, "Towards generalization and simplicity in continuous control," in *Advances in Neural Information Processing Systems*, 2017, pp. 6550–6561.

[21] H. Mohammadi, A. Zare, M. Soltanolkotabi, and M. R. Jovanović, "Convergence and sample complexity of gradient methods for the model-free linear quadratic regulator problem," *arXiv preprint arXiv:1912.11899*, 2019.

[22] J. Bu, A. Mesbahi, M. Fazel, and M. Mesbahi, "Lqr through the lens of first order methods: Discrete-time case," *arXiv preprint arXiv:1907.08921*, 2019.

[23] J. Bu, A. Mesbahi, and M. Mesbahi, "Policy gradient-based algorithms for continuous-time linear quadratic control," *arXiv preprint arXiv:2006.09178*, 2020.

[24] J. Bu and M. Mesbahi, "A note on nesterov's accelerated method in nonconvex optimization: a weak estimate sequence approach," *arXiv preprint arXiv:2006.08548*, 2020.

[25] J. Bu, L. J. Ratliff, and M. Mesbahi, "Global convergence of policy gradient for sequential zero-sum linear quadratic dynamic games," *arXiv preprint arXiv:1911.04672*, 2019.

[26] K. Zhang, B. Hu, and T. Basar, "Policy optimization for $\mathcal{H}_2$ linear control with $\mathcal{H}_\infty$ robustness guarantee: Implicit regularization and global convergence," in *Learning for Dynamics and Control*, 2020, pp. 179–190.

[27] L. Furieri, Y. Zheng, and M. Kamgarpour, "Learning the globally optimal distributed lq regulator," in *Learning for Dynamics and Control.* PMLR, 2020, pp. 287–297.

[28] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE circuits and systems magazine*, vol. 9, no. 3, pp. 32–50, 2009.

[29] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[30] J. Y. Lee, J. B. Park, and Y. H. Choi, "Integral q-learning and explorized policy iteration for adaptive optimal control of continuous-time linear systems," *Automatica*, vol. 48, no. 11, pp. 2850–2859, 2012.

[31] D. Lee and J. Hu, "Primal-dual q-learning framework for lqr design," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3756–3763, 2018.

[32] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.

[33] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory.* SIAM, 1994.

[34] S. Lojasiewicz, "A topological property of real analytic subsets," *Coll. du CNRS, Les équations aux dérivées partielles*, vol. 117, pp. 87–89, 1963.

[35] G. Zhang, S. Jiao, X. Xu, and L. Wang, "Compressed sensing and reconstruction with bernoulli matrices," in *Information and Automation (ICIA), 2010 IEEE International Conference on.* IEEE, 2010, pp. 455–460.

A Proof of the main theorem

Theorem 1. *We consider the problems*

$$\min_{K} \quad \mathcal{L}(K), \quad s.t., \quad K \in \mathcal{S}_K \tag{23}$$

$$\min_{L,G,Z} \quad f(L,G,Z), \quad s.t., \quad (L,G,Z) \in \mathcal{S} \tag{24}$$

Then we require assumptions either 1,2,3 or 1,2,4. Define K^* as the global minimum of $\mathcal{L}(K)$ in feasible region. Then we have (exact values in (30),(31)*)*

1. *if f is convex, the gradient satisfies*

$$\|\nabla \mathcal{L}(K)\|_F \geq C_1(L,G)(\mathcal{L}(K) - \mathcal{L}(K^*)). \tag{25}$$

2. *if f is μ-strongly convex, the gradient satisfies*

$$\|\nabla \mathcal{L}(K)\|_F \geq C_2(L,G)(\mu(\mathcal{L}(K) - \mathcal{L}(K^*)))^{1/2}. \tag{26}$$

where (the constants can be bounded with simple constraints bounding norms of L, G or K)

$$\begin{aligned} C_1(L,G) &= (\max\{\sigma_{\max}(L-L^*)\sigma_{min}^{-1}(G), \\ &\quad \sigma_{\max}(G-G^*)\sigma_{min}^{-2}(G)\sigma_{\max}(L)\})^{-1} \\ C_2(L,G) &= 2(\max\{\sigma_{min}^{-1}(G), \sigma_{min}^{-2}(G)\sigma_{\max}(L)\})^{-1} \end{aligned}$$

Proof of Theorem 1. First, for any convex function $f(x)$, let the minimum be x^*, and $x - x^* = \Delta$. Let $\nabla f(x) = g$. For any non-stationary point, $f(x) \leq f(x^*) + g^\top \Delta$. Since $\mathcal{S}$ is a convex set, $-\Delta$ belongs to the horizon of $\mathcal{S}$ at x, so there is a direction $\frac{\Delta}{\|\Delta\|}$ such that $f(x) - f(x - t\frac{\Delta}{\|\Delta\|}) > tg^\top \frac{\Delta}{\|\Delta\|}$, $t \to 0$, so the norm of projected gradient $\|\mathcal{P}_{\mathcal{S}}(\nabla f(x))\| \geq g^\top \frac{\Delta}{\|\Delta\|} = \frac{f(x)-f(x^*)}{\|x-x^*\|}$.

Let K^* be the optimal K and (Z^*, L^*, G^*) be the optimal point in the parameterized space. We have $\mathcal{L}(K^*) = f(Z^*, L^*, G^*)$. Denote V as any matrix in K space, $\mathcal{P}_V$ is projection of a vector onto direction V, then

$$\nabla\mathcal{L}(K)^\top \nabla\mathcal{L}(K) \geq (\mathcal{P}_V \nabla\mathcal{L}(K))^\top \mathcal{P}_V \nabla\mathcal{L}(K) \tag{27}$$

$$= \left(\frac{\nabla\mathcal{L}(K)[V]}{\|V\|_F}\right)^2. \tag{28}$$

We look at any feasible controller K. With either Assumption 3 or 4, we can define the mapping from K to (L, G, Z) respectively in one of the following ways:

1. (Assumption 3) let K map to (L, G) and $Z = \mathscr{Z}(L, G)$.
2. (Assumption 4) let

$$\begin{aligned}&(L, G, Z) = \operatorname{argmin}_{L', G', Z'}\ f(L', G', Z')\\ &\text{s.t., } (L', G', Z') \in \mathcal{S},\ G' \succ 0,\ L'{G'}^{-1} = K.\end{aligned}$$

Note f is convex, so

$$\begin{aligned}&\nabla f(L, G, Z)[(L, G, Z) - (L^*, G^*, Z^*)]\\ &\geq f(L, G, Z) - f(L^*, G^*, Z^*)\\ &= f(\mathscr{Z}(L, G), L, G) - f(\mathscr{Z}(L^*, G^*), L^*, G^*)\\ &= \mathcal{L}(K) - \mathcal{L}(K^*).\end{aligned} \tag{29}$$

Now we consider the directional derivative in K space. By definition,

$$\nabla\mathcal{L}(K)[V] = \lim_{t\to 0^+} (\mathcal{L}(K + tV) - \mathcal{L}(K))/t.$$

Let $\Delta L = L^* - L$, $\Delta G = G^* - G$, and $V = \Delta L G^{-1} - LG^{-1}\Delta G G^{-1}$. Then

$$\begin{aligned}&\nabla\mathcal{L}(K)[V]\\ &= \lim_{t\to 0^+} (\mathcal{L}(K + tV) - \mathcal{L}(K))/t\\ &= \lim_{t\to 0^+} (\mathcal{L}(LG^{-1} + t(\Delta L G^{-1} - LG^{-1}\Delta G G^{-1}))\\ &\quad - \mathcal{L}(LG^{-1}))/t\\ &= \lim_{t\to 0^+} (\mathcal{L}((L + t\Delta L)(G + t\Delta G)^{-1}) - \mathcal{L}(LG^{-1}))/t\end{aligned}$$

With assumption 3, we continue with

$$\begin{aligned}&\nabla\mathcal{L}(K)[V]\\ &= \lim_{t\to 0^+} (f(L + t\Delta L, G + t\Delta G, \mathscr{Z}(L + t\Delta L, G + t\Delta G))\\ &\quad - f(L, G, \mathscr{Z}(L, G)))/t\\ &\leq \lim_{t\to 0^+} (f(L + t\Delta L, G + t\Delta G, Z + t\Delta Z)\\ &\quad - f(L, G, \mathscr{Z}(L, G)))/t\\ &= \nabla f(L, G, Z)[(L^*, G^*, Z^*) - (L, G, Z)]\end{aligned}$$

With assumption 4, we continue with

$$\begin{aligned}\nabla\mathcal{L}(K)[V] = \lim_{t\to 0^+} \min_{L,G,Z}\ & f(L, G, Z) - f(L, G, Z)\\ \text{s.t., } & (L', G', Z') \in \mathcal{S},\ G' \succ 0,\\ & L'{G'}^{-1} = (L + t\Delta L)(G + t\Delta G)^{-1}.\end{aligned}$$

and then

$$\begin{aligned}\nabla\mathcal{L}(K)[V] &\leq \lim_{t\to 0^+}(f(L+t\Delta L, G+t\Delta G, Z+t\Delta Z)\\ &\quad - f(L,G,\mathscr{Z}(L,G)))/t\\ &= \nabla f(L,G,Z)[(L^*,G^*,Z^*)-(L,G,Z)]\end{aligned}$$

The inequality results in that $(L+t\Delta L, G+t\Delta G, Z+t\Delta Z)$ is feasible so the value is bigger than the optimal value. So the final inequality holds either by Assumption 3 or 4. So

$$\begin{aligned}&\nabla\mathcal{L}(K)[-V]\\ &\geq \nabla f(L,G,Z)[(L,G,Z)-(L^*,G^*,Z^*)] > 0.\end{aligned}$$

Using (27) and (29), we have

$$\|\nabla\mathcal{L}(K)\|_F^2 \geq \frac{1}{\|V\|_F^2}(\mathcal{L}(K)-\mathcal{L}(K^*))^2 \tag{30}$$

If $f(L,G,Z)$ is μ strongly convex, then we can restrict f in the line segment $(L,G,Z)-(L^*,G^*,Z^*)$ and get

$$\begin{aligned}&\|\mathcal{P}_{(L,G,Z)-(L^*,G^*,Z^*)}\nabla f(L,G,Z)\|\\ &\geq \mu^{1/2}(f(L,G,Z)-f(L^*,G^*,Z^*))^{1/2}\end{aligned}$$

then we have that

$$\begin{aligned}&\nabla f(L,G,Z)[(L,G,Z)-(L^*,G^*,Z^*)]\\ &= \|\mathcal{P}_{(L,G,Z)-(L^*,G^*,Z^*)}\nabla f(L,G,Z)\|\\ &\quad\cdot\|(L,G,Z)-(L^*,G^*,Z^*)\|\\ &\geq \mu^{1/2}(f(L,G,Z)-f(L^*,G^*,Z^*))^{1/2}\\ &\quad\cdot\|(L,G,Z)-(L^*,G^*,Z^*)\|.\end{aligned}$$

then

$$\begin{aligned}&\|\nabla\mathcal{L}(K)\|_F^2\\ &\geq \frac{1}{\|V\|_F^2}(\nabla f(L,G,Z)[(L,G,Z)-(L^*,G^*,Z^*)])^2\\ &\geq \frac{\mu\|(L,G,Z)-(L^*,G^*,Z^*)\|^2}{\|V\|_F^2}\\ &\quad\cdot(f(L,G,Z)-f(L^*,G^*,Z^*))\\ &= \frac{\mu(\|L^*-L\|^2+\|G^*-G\|^2+\|Z^*-Z\|^2)}{\|(L^*-L)G^{-1}-LG^{-1}(G^*-G)G^{-1}\|_F^2}\\ &\quad\cdot(f(L,G,Z)-f(L^*,G^*,Z^*))\\ &\geq \frac{\mu(\|L^*-L\|^2+\|G^*-G\|^2)}{\|(L^*-L)G^{-1}-LG^{-1}(G^*-G)G^{-1}\|_F^2}\\ &\quad\cdot(f(L,G,Z)-f(L^*,G^*,Z^*))\\ &\geq \frac{4\mu(f(L,G,Z)-f(L^*,G^*,Z^*))}{(\max\{\sigma_{min}^{-1}(G),\sigma_{min}^{-2}(G)\sigma_{\max}(L)\})^2}.\end{aligned}\tag{31}$$

B Proof for Sec. 4.2

In this section, we will show that Sec. 4.2 satisfy Assumption 1,2,4. Thus Theorem 1 applies to this problem, which means all stationary points of $\mathcal{L}(K)$ are global minimum.

Example 1. *(Assumption 1,2,4) We consider minimizing the L_2 gain of a closed loop system. The input output system is*

$$\dot{x} = Ax + Bu + B_w w, \;\; y = Cx + Du \tag{32}$$

and we use the state feedback controller $u = Kx$, and let.

$$\mathcal{L}(K) := (\sup_{\|w\|_2=1} \|y\|_2)^2.$$

If we minimize the loss $\mathcal{L}(K)$, this problem can be reformulated as

$$\begin{aligned}
&\min_{L,G,\gamma} \;\; f(L, G, \gamma) := \gamma \\
&s.t., \;\; \begin{bmatrix} AG + GA^\top + BL + L^\top B^\top + B_w B_w^\top & (CG + DL)^\top \\ CG + DL & -\gamma I \end{bmatrix} \\
&:= M(L, G, \gamma) \preceq 0.
\end{aligned}$$

And $K^* = L^* {G^*}^{-1}$. This pair of problems satisfy Assumptions 1,2,4.

Proof. We have $\mathcal{L}(K^*) = f(L^*, G^*, \gamma^*)$ from [33]. Note that, K^* can map to different pairs (L, G) whenever $LG^{-1} = K^*$, and γ^* associates to one pair among them. We can equivalently formulate

$$\begin{aligned}
\mathcal{L}(K^*) = &\min_{L,G,\gamma} \gamma \\
&\text{s.t., } M(L, G, \gamma) \preceq 0, \;\; LG^{-1} = K^*.
\end{aligned}$$

The question is, can we establish the same connection at any stabilizing controller K, say,

$$\mathcal{L}(K) = \min_{L,G,\gamma} \gamma \tag{33a}$$

$$\text{s.t., } M(L, G, \gamma) \preceq 0, \;\; LG^{-1} = K. \tag{33b}$$

Note that, the intermediate step [33, Sec 7.5.1] is

$$\mathcal{L}(K) = \min_{G,\gamma} \gamma, \quad \text{s.t.,} \tag{34a}$$

$$\begin{bmatrix} (A + BK)G + G(A + BK)^\top + B_w B_w^\top & G^\top (C + DK)^\top \\ (C + DK)G & -\gamma I \end{bmatrix} \preceq 0. \tag{34b}$$

Denote the optimizer of (33) by $\hat{L}, \hat{G}, \hat{\gamma}$, and the optimizer of (34) by $\check{G}, \check{\gamma}$.

Note $\hat{\gamma} \leq \check{\gamma}$. If (33) is not true, $\hat{\gamma} < \check{\gamma}$, we can replace $\check{G}, \check{\gamma}$ with $\hat{G}, \hat{\gamma}$ in (34) and it's still feasible. Thus the optimality condition of $\check{G}, \check{\gamma}$ in (34) is violated, which contradicts the assumption that (33) is not true. Then we claim that (33) is true.

B.1 Connection with $\mathcal{H}_\infty$ robustness

In this part, we will explain the connection between the above example with [26]. [26] suggests[9] an $\mathcal{H}_\infty$ bound constraint, where the $\mathcal{H}_\infty$ norm of the transfer function of y to w, is upper bounded by a number γ. [33, 6.3.2] suggests that the L_2 gain is also the $\mathcal{H}_\infty$ norm of transfer function. The time domain $\mathcal{H}_\infty$ norm, is exactly defined as

$$\sup_{\|w\|_2=1} \|y\|_2, \text{ s.t., dynamics (32)}$$

[9]To be consistent with the notations, [26] uses $\sqrt{W}$ for B_w, $\sqrt{Q}$ for C and $\sqrt{R}$ for D as in our formulation, and $u = Kx$ where K is a static state feedback controller.

The frequency domain objective is $\sup_{\|w\|_{\mathcal{H}_2}=1} \|y\|_{\mathcal{H}_2}$, and can be explicitly expressed with system parameters as

$$\sup_{\theta} \ \sigma_{\max}\left[(C+DK)(j\theta I-(A+BK))^{-1}B_w\right].$$

The time and frequency domain definition are equivalent. With these definitions, if we enforce an upper bound of squared $\mathcal{H}_\infty$ norm being γ, we just need to set $M(L,G,\gamma) \preceq 0$ as a constraint. The objective (loss) functions discussed in [35] have P as variable (we use G above) and they are convex, thus these instances has convex formulation and can be covered by our result.

C A more general description of Assumption 4

Note that, in Assumption 4, we still require $K = LG^{-1}$, which is still a specific mapping. Here we will propose the following assumption which replaces the mapping by a general notation Φ.

Suppose we consider the problems

$$\min_{K} \quad \mathcal{L}(K), \quad \text{s.t., } K \in \mathcal{S}_K \tag{35}$$

and

$$\min_{G_1,...,G_m} \quad f(G_1,...,G_m), \quad \text{s.t.,} \quad (G_1,...,G_m) \in \mathcal{S} \tag{36}$$

We denote $G = (G_1,...,G_m)$. With

Assumption 5. *The feasible set $\mathcal{S}$ is convex in G, i.e., in all of the variables $G_1,...,G_m$.*

Assumption 6. *The cost function $f(G) := f(G_1,...,G_m)$ is convex, finite and differentiable in $G \in \mathcal{S}$.*

Assumption 7. *Assume we can express $\mathcal{L}(K)$ as:*

$$\begin{aligned}\mathcal{L}(K) = &\min_{G} \ f(G)\\ &\quad s.t., \ G \in \mathcal{S}, \ K = \bar{\Phi}(G) := \Phi(G_1,...,G_m).\end{aligned}$$

And we assume the first order Taylor expansion of the mapping Φ can be written as

$$\bar{\Phi}(G+\mathrm{d}G) = \bar{\Phi}(G) + \bar{\Psi}(G)[\mathrm{d}G] + o(\mathrm{d}G)$$

for any $G \in \mathcal{S}$ and any legit perturbation $\mathrm{d}G$. $\bar{\Psi}$ is a linear and finite operator, estimated at point G and acting on perturbation $\mathrm{d}G$. Or

$$\begin{aligned}&\Phi(G_1+\mathrm{d}G_1,...,G_m+\mathrm{d}G_m)\\ &= \Phi(G_1,...,G_m) + \Psi(G_1,...,G_m)[\mathrm{d}G_1,...,\mathrm{d}G_m]\\ &\quad + o(\mathrm{d}G_1,...,\mathrm{d}G_m)\\ &= \Phi(G_1,...,G_m) + \Psi_1(G_1,...,G_m)[\mathrm{d}G_1]\\ &\quad + ... + \Psi_m(G_1,...,G_m)[\mathrm{d}G_m] + o(\mathrm{d}G_1,...,\mathrm{d}G_m)\end{aligned}$$

where Ψ and Ψ_i for $i \in [m]$ are linear and finite operators.

Thm. 1 holds with the above Assumptions 5, 6, 7. We can apply the proof in Append. A with the matrix V defined as $\bar{\Psi}(G)[\mathrm{d}G]$.

Example 2. *(Distributed LQ regulator [27]) We consider the time varying linear system*

$$\begin{aligned}x(t+1) &= A(t)x(t) + B(t)u(t) + w(t),\\ y(t) &= C(t)x(t) + v(t).\end{aligned}$$

This is in finite time horizon $t = 0, ..., T$. The state evolution is same as the setup in our SLS example, and we can use the same notation $X, U, W, Z, \mathcal{A}, \mathcal{B}$. We further define

$$Y = \begin{bmatrix} y(0) \\ ... \\ y(T) \end{bmatrix},\ V = \begin{bmatrix} v(0) \\ ... \\ v(T) \end{bmatrix},$$
$$\mathcal{C} = \mathrm{diag}(C(0), ..., C(T-1), 0), .$$

Now we will consider

$$u(t) = \sum_{i=0}^{t} K(t, t-i) y(i) \tag{37}$$

The $\mathcal{K}$ matrix is same as SLS there. In [27, Sec. 3], there is an extra assumption that ensures quadratic invariance (QI) for the sake of distributed control, and QI leads to a legit convexification. We refer readers to the corresponding section of that paper for details. The loss function is same as (21)*. It is proven that, define*

$$P_{12} = (I - Z\mathcal{A})^{-1} Z\mathcal{B},$$
$$\Phi(\mathcal{G}) = (I + \mathcal{G}\mathcal{C}P_{12})^{-1}\mathcal{G}.$$

Then we can get a new variable $\mathcal{G}$ and a function Φ. With $\mathcal{K} = \Phi(\mathcal{G})$, the loss can be proven to be convex in $\mathcal{G}$. Indeed, the mapping satisfies the assumptions 5, 6, 7, and the exact formulation of the two optimization problems are described in [27, Append. A, Lem. 5]. Thus we can also claim via our Thm. 1 that, such distributed LQ regulator problem with $\mathcal{K}$ as variable has no spurious local minimum.

D System Level Synthesis with Infinite Horizon

In this work, we studied the landscape of the optimal control problem where the variables are matrices (which are finite dimensional), and SLS is an example. Generally, SLS also works with the infinite horizon problem. In this regime, the variables are *transfer functions* and they are infinite dimensional. In practice, when the problem is convexified, one can parameterize the transfer function (say as finite impulse response) and minimize the loss with respect to the finite dimensional parameters. However, our theorem does not apply to the infinite dimensional optimization problems, and it is not obvious that the parameterization satisfies the assumptions for our main theorem. We review the infinite horizon SLS here. A future direction is to judge whether the gradient dominance holds in the space of transfer function or its parameterized form, and how to analyze it using SLS.

Example 3. *(System level synthesis with infinite horizon) [2] Suppose one has a discrete time dynamical system with*

$$x(t+1) = Ax(t) + Bu(t) + w(t)$$

One can apply a dynamic controller $K(z)$. The goal is to find the optimial controller which minimizes the LQR cost where $u(z) = K(z)x(z)$

$$\mathcal{L}(K) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} x(t)^\top Q x(t) + u(t)^\top R u(t)$$

Suppose x_0, w_t are i.i.d. from $\mathcal{N}(0, \Sigma)$. The SLS convexification is to define two transfer functions $\Phi_X(z), \Phi_U(z)$, and solve the following convex optimization problem

$$\min_{\Phi_X(z), \Phi_U(z)} \left\| \begin{bmatrix} Q^{1/2}\Phi_X(z) \\ R^{1/2}\Phi_U(z) \end{bmatrix} \Sigma^{1/2} \right\|_{\mathcal{H}_2},$$
$$s.t.,\ \begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \Phi_X(z) \\ \Phi_U(z) \end{bmatrix} = I,$$
$$\Phi_X(z), \Phi_U(z) \in \frac{1}{z}\mathcal{RH}_\infty.$$

Let the optimizer be $\Phi_U^*(z), \Phi_X^*(z)$. The optimal controller is $K^*(z) = \Phi_U^*(z)(\Phi_X^*(z))^{-1}$.

E Conditions of convexifiable nonconvex loss

We consider the pair of problems in Theorem 1, and ask the question: what property of the nonconvex loss function $\mathcal{L}(K)$ allows us to reformulate the problem (11) as a *convex* optimization problem (12) (which corresponds to Assumption 2)? In this section we propose the following lemma.

Lemma 1. *Suppose Assumptions 1, 3 hold, and $\mathcal{L}(LG^{-1})$ as a function of L, G is differentiable*[10]. *We define the notation $\nabla^2_{L,G}\mathcal{L}(LG^{-1})[\Gamma_1, \Gamma_2]$ as in* (38). *If $\nabla^2_{L,G}\mathcal{L}(LG^{-1})[\Gamma_1, \Gamma_2] > 0$ for all $(L, G) \in \mathcal{S}$ and all (Γ_1, Γ_2) such that $\mathcal{A}(\Gamma_2) + \mathcal{B}(\Gamma_1) = 0$, then we can define a convex function $f(L, G)$ so that Assumption 2 holds. We can apply Theorem 1 so that all stationary points of $\mathcal{L}(K)$ are global minimum.*

Proof. Suppose we observe the simple version (13). We know from Assumption 3 that, $f(L, G) = \mathcal{L}(K) = \mathcal{L}(LG^{-1})$ is convex in L, G. We take the Hessian and ask for

$$\nabla \begin{bmatrix} \nabla\mathcal{L}(LG^{-1})G^{-1} \\ -G^{-1}L^\top\nabla\mathcal{L}(LG^{-1})G^{-1} \end{bmatrix} \succ 0.$$

Note that this is a tensor and it is positive definite. For simplicity, we analyze the directional Hessian as the following. We expand the left hand side of the inequality above and define $\nabla^2_{L,G}\mathcal{L}(LG^{-1})[\Gamma_1, \Gamma_2]$ as

$$\begin{aligned} &\nabla^2_{L,G}\mathcal{L}(LG^{-1})[\Gamma_1, \Gamma_2] \\ :=\ &\nabla^2\mathcal{L}(LG^{-1})[\Gamma_1 G^{-2}, \Gamma_1] \\ &- 2\nabla^2\mathcal{L}(LG^{-1})[\Gamma_1, LG^{-3}\Gamma_2] \\ &- 2\langle\Gamma_1, \nabla\mathcal{L}(LG^{-1})G^{-1}\Gamma_2 G^{-1}\rangle \\ &+ 2\langle\Gamma_2, LG^{-1}\Gamma_2 G^{-1}\nabla\mathcal{L}(LG^{-1})G^{-1}\rangle \\ &+ \nabla^2\mathcal{L}(LG^{-1})[LG^{-2}\Gamma_2, LG^{-2}\Gamma_2] \end{aligned} \tag{38}$$

This is the directional Hessian of $\mathcal{L}$ with respect to (L, G) in direction (Γ_1, Γ_2). Thus, if $\nabla^2_{L,G}\mathcal{L}(LG^{-1})[\Gamma_1, \Gamma_2] > 0$ for all $(L, G) \in \mathcal{S}$ and all (Γ_1, Γ_2) such that $\mathcal{A}(\Gamma_2) + \mathcal{B}(\Gamma_1) = 0$ (which is a condition on nonconvex loss $\mathcal{L}$), we know that $f(L, G)$ is convex in L, G and the convexification can be made.

[10] We also allow the relaxation where it has well defined one sided directional derivative, as discussed in the footnote of Assumption 2.