Analysis of Policy Gradient Descent for Control: Global Optimality via Convex Parameterization

Yue Sun and Maryam Fazel *

September 30, 2021

Abstract

Common reinforcement learning methods 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 these analyses rely on is that the LQR cost function satisfies the "gradient dominance" property with respect to the policy parameters. Gradient dominance helps guarantee 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 several control problems including continuous/discrete time LQR, maximizing dissipativity, LQR for distributed controller, $\mathcal{H}_2/\mathcal{H}_\infty$ robust control.

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 continuous/discrete time LQR, distributed controller, minimizing the L_2 gain, system-level parameterization [2], $\mathcal{H}_2/\mathcal{H}_\infty$ mixed/robust control, etc. In addition to extending results on convergence to the optimal policy to a far larger class of problems, our unified framework provides insights into the landscape of the loss (or cost) as a function of the policy.

1 Introduction

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

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^p$ is the input, and A, B are constant matrices describing the dynamics. The goal of optimal control is to determine the input series $u(t)$ that minimizes some cost function that typically depends on state and input. In the infinite horizon LQR problem, with constant matrices $Q \in \mathbf{S}^n_{++}, R \in \mathbf{S}^p_{++}$, one minimizes

$$\text{loss}(u(t)) := \mathbf{E}_{x_0} \int_0^\infty (x(t)^\top Q x(t) + u(t)^\top R u(t)) dt \tag{2}$$

It is known that the optimal controller is linear in the state, referred to as static state feedback, and can be described as $u(t) = Kx(t)$ 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 on K, usually used in reinforcement learning settings.

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

In many practical cases, the system dynamics is unknown, and among the optimal control algorithms, there are two major types. The first type is model based methods, when the system is first identified and then a controller is constructed based on the identified system. System identification has a long history, as reviewed [9]. Recently [10] gave sample complexity bounds for state-observed system. [11–14] describe the joint system identification and optimal control approaches.

Another type of method is model free method, when the controller is directly trained by observing the cost (or loss) function, without characterizing the dynamics. Here one does not necessarily estimate the system parameters A, B. [15] is a review of reinforcement learning area and optimal control and studies some fixed point type dynamic programming methods. Q-learning is a typical model free method for reinforcement learning, and it is applied to LQR as in [11, 16, 17].

This paper mainly focus on another model-free method called policy gradient descent. It calls for an estimate of the cost (2) as well as its gradient with respect to controller K when $u = Kx$. One hopes that gradient descent with respect to K converges to the optimal controller K^*. The policy gradient descent is more recently reviewed by [18, 19]. [1] provides a counterexample showing that minimizing the quadratic LQR cost as a function of K is not convex, quasi-convex or star-convex.

Recently people have witnessed the *empirical* success of first order methods in solving nonconvex reinforcement learning problems. [20, Ch. 3] proposes the gradient based method for optimal control and extends to decentralized control. [21] studies feedback control with dynamical controllers, and observes that gradient descent with Youla parameterization is robust within the set of stabilizing controllers while other parameterizations are not. On *theoretical* side, [1] gives the first result by proving the coercivity and gradient dominance property of $\mathcal{L}(K)$ for the discrete time LQR. Based on this, [1] shows the linear convergence of gradient based method. Later [22] shows a similar result for the continuous time case, [23, 24] give a more detailed analysis for both discrete and continuous time LQR. [25] and [26] shows similar results for two settings of zero-sum LQ games. [27] studies the convergence of gradient descent on $\mathcal{H}_2$ control with $\mathcal{H}_\infty$ constraint and shows that gradient descent implicitly makes the controller robust. [28] shows the convergence for finite-horizon distributed control under the quadratic invariance assumption. Those papers all show convergence of policy gradient descent by gradient dominance property, but investigate different control problems and the proofs are given case by case.

Traditionally, convex parameterization (convexification) 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.

Control for nonlinear systems is far more difficult, typically via dynamic programming, solving Bellman equations [29], or recent deep RL that 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 nonconvex policy gradient descent and known convex parameterization methods, which provides insight into why convergence to the optimal solution happens despite nonconvexity in all the problems cited above. We 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 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 other control problems including instances of optimal control, robust control, mixed design and system level synthesis. Thus for all these problems, if one wants to understand whether the (nonconvex) loss with respect to controller parameter K can be minimized by policy gradient descent (first-order optimization methods that update K), one can directly check if it is covered by our theorem, avoiding a case-by-case analysis. Also, as discussed in [1], theoretical guarantees for first-order methods naturally lead to guarantees for the more practical zeroth-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. Sec. 3 presents our main result on the gradient dominance property for the nonconvex loss. Sec. 4 lists more examples of control problems covered by the main theorem. Sec. 5 gives a proof sketch with intuitive connections between the nonconvex and convex formulations.

2 Review of convex parameterization for continuous time LQR

Convexification method (e.g., solving optimal control by linear matrix inequalities (LMI) in [30]) is widely used in optimal control problems, and here we discuss its application for continuous time LQR [22]. Define a continuous time linear time invariant system (1) where x is state and u is input signal, and x_0 is the initial state. We assume that $\mathbf{E}(x_0 x_0^\top) = \Sigma \succ 0$. This is a commonly used setup such as in [23, §3.3], [20, Paper 3].

One can then consider minimizing the linear quadratic (LQ) loss (2) as a function of $u(t)$ where Q, R are positive definite matrices. It is known [3] that, the input signal that minimizes the loss function loss(u) is given by a static state feedback controller, denoted by $u(t) = K^* x(t)$. K^* can be obtained by solving linear equations, called riccati equations. Note that once we know the optimal state feedback controller is static, we can write loss as $\mathcal{L}(K)$ which is a function of K instead, and search only in static state feedback controllers.

An alternative approach is reparameterizing to obtain a convex formulation, as used in [22], which we will review here, starting from the Lyapunov equation. Suppose the initial state satisfies $\mathbf{E}(x_0 x_0^\top) = \Sigma \succ 0$, and $\dot{x} = Ax$. Then with a matrix $P \in \mathbf{S}_{++}^{n\times n}$ (P is a positive definite matrix) as the variable, the Lyapunov equation is written as

$$AP + PA^\top + \Sigma = 0 \tag{3}$$

In our setup (1), we use a state feedback controller $u = Kx$, thus we have $\dot{x} = (A + BK)x$. We denote the set of stabilizing controllers as $\mathcal{S}_{K,\text{sta}}$, which is defined as

$$\mathcal{S}_{K,\text{sta}} = \{K : \ \text{Re}(\lambda_i(A + BK)) < 0, \ i = 1, ..., n\}.$$

If a state feedback controller is applied, the loss is only bounded when $K \in \mathcal{S}_{K,\text{sta}}$ and is coersive in $\mathcal{S}_{K,\text{sta}}$ [24]. Replace A by the closed loop system matrix $A + BK$ in the Lyapunov equation, and let $L = KP \in \mathbb{R}^{p\times n}$, we get

$$AP + PA^\top + BL + L^\top B^\top + \Sigma = 0$$

Let $\mathcal{A}(P) = AP + PA^\top, \ \mathcal{B}(L) = BL + L^\top B^\top$, which are often referred to as Lyapunov maps. Assume $\mathcal{A}$ is invertible, then we have the relation

$$\mathcal{A}(P) + \mathcal{B}(L) + \Sigma = 0. \tag{4}$$

Indeed, once we fix the system and any stabilizing controller A, B, K, the matrices P as well as $L = KP$ are uniquely determined. P is the Grammian matrix

$$P = \int_0^\infty e^{t(A+BK)} \ \Sigma \ e^{t(A+BK)^\top} \ dt. \tag{5}$$

P is positive definite if $\Sigma \succ 0$. We are interested in the loss function $\mathcal{L}(K)$ when[1] $K \in \mathcal{S}_{K,\text{sta}}$, which corresponds to (2) by inserting $u(t) = Kx(t)$.

$$\mathcal{L}(K) = \mathbf{Tr}((Q + K^\top RK)P). \tag{6}$$

One can construct a bijection from P, L to K, and prove that, if we minimize $f(L, P)$ subject to (4), the optimizer P^*, L^* will map to the optimal K^*, and this minimization problem is convex, so we can solve it by convex optimization algorithms.

Convex reparameterization for Continuous time LQR: Suppose the dynamics and loss are (1) and (2), and let $\mathbf{E}(x_0 x_0^\top) = \Sigma \succ 0$. Denote the (static) state feedback controller by K, so that $u(t) = Kx(t)$. The optimal control problem then is

$$\min_K \ \mathcal{L}(K), \quad \text{s.t.} \quad K \in \mathcal{S}_{K,\text{sta}} \tag{7}$$

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

where $\mathcal{L}(K)$ is the cost in (2) with $u = Kx$. This problem can be expressed as the following equivalent convex problem,

$$\min_{L,P,Z} \; f(L,P,Z) := \mathbf{Tr}(QP) + \mathbf{Tr}(ZR) \tag{8a}$$

$$\text{s.t. } \mathcal{A}(P) + \mathcal{B}(L) + \Sigma = 0, \; P \succ 0, \tag{8b}$$

$$\begin{bmatrix} Z & L \\ L^\top & P \end{bmatrix} \succeq 0 \tag{8c}$$

The connection between the two problems is distilled in Sec. 3. For all feasible (L,P,Z) triplets in (8), we can take the first two elements (L,P), and they form a bijection with all stabilizing controllers K in (7). The loss function values are equal under the bijection. So we can solve for L^*, P^*, and $K^* = L^*(P^*)^{-1}$.

3 Main result

Motivated by methods that use gradient descent in the policy space, we ask whether running a gradient-based algorithm and getting $\nabla_K \mathcal{L}(K) = 0$ for some K in fact gives the globally optimum K^*. [1, 22] show the coercivity and gradient dominance property of $\mathcal{L}(K)$ for the discrete time and continuous time LQR respectively. In this paper, we generalize these results from the special case of continuous-time LQR to a much broader set of control problems, showing the gradient dominance property of the nonconvex losses as functions of policy.

We present our main result in Theorem 1. It is described as a pair of problems satisfying Assumptions 1, 3, which covers problems extending beyond continuous time LQR. In Sec. 4 we will review more examples showing the generality of this result.

We begin by considering an abstract description of the pair of problems (7) and (8). These problem descriptions cover LQR as discussed in the last section, as well as more problems discussed in Sec. 4. Consider the problems

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

and

$$\min_{L,P,Z} \quad f(L,P,Z), \quad \text{s.t.} \quad (L,P,Z) \in \mathcal{S}, \tag{10}$$

where the sets $\mathcal{S}_K, \mathcal{S}$ capture the control constraints. They are defined differently for each specific example in Sec. 4. For example, for continuous time LQR, $\mathcal{S}_K$ is the set of all stabilizing controllers (7) and $\mathcal{S}$ is the intersection of (8b) & (8c). We allow special cases when (10) depends only on L, P,

$$\min_{L,P} \quad f(L,P), \quad \text{s.t.} \quad (L,P) \in \mathcal{S} \tag{11}$$

We distill three properties of the two problems (9) and (10) that will be critical for Theorem 1, and allow us to cover more problems as discussed in Sec. 4.

Assumption 1. *The feasible set $\mathcal{S}$ is convex in (L,P,Z). The cost function $f(L,P,Z)$ is convex, bounded, and differentiable in $(L,P,Z) \in \mathcal{S}$.*

Assumption 1 imply the second problem is convex. Next, we extract the property of the connection between (7) and (8), and give an abstract description of the assumptions for (9) and (10).

Assumption 2. *Let P be always invertible*[2] *in $\mathcal{S}$. Assume we can express $\mathcal{L}(K)$ as:*

$$\mathcal{L}(K) = \min_{L,P,Z} \; f(L,P,Z)$$
$$\textit{s.t. } (L,P,Z) \in \mathcal{S}, \; LP^{-1} = K.$$

[2]The invertibility of P guarantees a well defined map between L, P and K, which is usually true, e.g., for the instances in Sec. 4.

With the assumptions above, we will present the main theorem.

Theorem 1. *We consider the problems* (9) *and* (10)*, and we require Assumptions 1,2. Let K^* denote the global minimizer of $\mathcal{L}(K)$ in S_K. Then there exist constants $C_1, C_2 > 0$ independent of K,*

1. *if f is convex, the gradient of $\mathcal{L}$ satisfies*[3]

$$\|\nabla \mathcal{L}(K)\|_F \geq C_1(\mathcal{L}(K) - \mathcal{L}(K^*)). \tag{12}$$

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

$$\|\nabla \mathcal{L}(K)\|_F \geq C_2(\mu(\mathcal{L}(K) - \mathcal{L}(K^*)))^{1/2}. \tag{13}$$

The constants C_1, C_2 are discussed below and in appendix.

Remark 1. *The constants are case by case. We show that, for continuous time LQR, in the sublevel set where $\mathcal{L}(K) \leq a$, we define*

$$\begin{aligned}
\nu &= 4a\left(\sigma_{\max}(A)\lambda_{\min}^{-1/2}(Q) + \sigma_{\max}(B)\lambda_{\min}^{-1/2}(R)\right)^2, \\
C_{1,1} &= 2a\nu\lambda_{\min}^{-1}(\Sigma)\lambda_{\min}^{-1/2}(Q)\lambda_{\min}^{-1/2}(R), \\
C_{1,2} &= 2a^2\nu^2\lambda_{\min}^{-2}(\Sigma)\lambda_{\min}^{-3/2}(Q)\lambda_{\min}^{-1/2}(R),
\end{aligned}$$

Then $C_1 = (\max\{C_{1,1}, C_{1,2}\})^{-1}$. [22] gives another convex formulation with strong convexity and we can get C_2 for that form, the details are in appendix.

Our lower bound of the gradient, $\|\nabla \mathcal{L}(K)\|_F \gtrsim (\mathcal{L}(K) - \mathcal{L}(K^*))^\alpha$, is known as Lojasiewicz inequality [31]. When $\alpha = 1/2$, it is also called the *gradient dominance* property. If Lojasiewicz inequality holds, all local minima of the objective function are global minima, then an iterative method with $\|\nabla \mathcal{L}(K)\|_F \to 0$ makes the iterates converge to the global minimum. Nonconvex functions that satisfy Lojasiewicz inequality are easily optimized, compared to those with spurious local minimums. Gradient dominance is usually true about the neighborhood of a local minimum and typically used in optimization as a tool for local convergence analysis, whereas it is rare that we show that it holds for $\mathcal{L}(K)$ globally.

Assumption 2 is a rather weak assumption. Assumption 3 is a stronger one covered by Assumption 2 that, we assume that there is a bijection between K and (L, P). This is true for many control problems such as continuous time LQR. Theorem 1 also holds with Assumptions 1,2. We emphasize the special case for Assumptions 1,3 since it is to illustrate in Sec. 5.

Assumption 3.

1. *(Bijection between two feasible sets) Let P be invertible, $K = LP^{-1}$ define a bijection*[4] *$K \leftrightarrow (L, P)$. For any such bijection $K \leftrightarrow (L, P)$, $\exists Z$, such that $(L, P, Z) \in \mathcal{S}$.*
2. *(Equivalence of functions) Choose a controller $K \in \mathcal{S}_K$ with corresponding $(L, P) \in \mathcal{S}$. Then $\mathcal{L}(K) = \min_Z f(L, P, Z)$ subject to $(L, P, Z) \in \mathcal{S}$.*

Our main theorem suggests that, when the original nonconvex optimization problem can be mapped to a convex optimization problem that satisfies Assumptions 1,2 or 1,3, 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 (9), its convexified form (10) and a few assumptions. This is an abstract and general description that does not need the exact continuous

[3] We always consider the directional derivative of a feasible direction within descent cone.

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

time LQR formulation in Sec. 2. We can easily check that the continuous time LQR satisfies the Assumptions 1,3, thus we can directly apply Theorem 1 to argue that the continuous time LQR cost $\mathcal{L}(K)$ satisfies (12).

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 1,3.

1. Continuous time LQR problem: See (7), (8).

2. Budget on total energy of input, which is the following constraint

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

   It can be added to (8) 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,

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

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

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

   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 Discrete time LQR

We consider a discrete time linear system

$$x(t+1) = Ax(t) + Bu(t),\ x(0) = x_0, \tag{14}$$

The goal is to find a state feedback controller K such that the loss function

$$\mathcal{L}(K) = \mathbf{E}_{x_0} \sum_{i=0}^\infty x(t)^\top Q x(t) + u(t)^\top R u(t),\ u = Kx$$

is minimized. In other words, we will solve

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

Similar to the continuous time system, one can choose the same parameterization P, L, Z and another PSD matrix $G \in \mathbb{R}^{n\times n} \succeq 0$ and solve the following problem

$$\min_{L,P,Z,G} \; f(L,P,Z,G) := \mathbf{Tr}(QP) + \mathbf{Tr}(ZR) \tag{15a}$$

$$\text{s.t. } P \succ 0,\; G - P + \Sigma = 0, \tag{15b}$$

$$\begin{bmatrix} Z & L \\ L^\top & P \end{bmatrix} \succeq 0,\; \begin{bmatrix} G & AP+BL \\ (AP+BL)^\top & P \end{bmatrix} \succeq 0 \tag{15c}$$

The goal is to argue that $\mathcal{L}(K)$ and (15) has the connection such that Theorem 1 applies, so that the stationary point of $\mathcal{L}(K)$ has to be the global optimum.

Lemma 1. *The LQR problem* $\min\ \mathcal{L}(K)$ *with stabilizing* K*, and problem* (15)*, satisfy Assumption 1, 2.*

Proof. (15) is a convex optimization problem. Now we prove Assumption 2, i.e., we prove that $L(K)$ equals the minimum of the problem (15) with an extra constraint $K = LP^{-1}$.

- We first minimize over Z, the minimizer is $Z = LP^{-1}L^\top$. Now replace L by KP and the loss becomes $\mathbf{Tr}((Q + K^\top RK)P)$.
- Eliminate G by

  $$G - P + \Sigma = 0,\; \begin{bmatrix} G & AP+BL \\ (AP+BL)^\top & P \end{bmatrix} \succeq 0$$

  Using Schur complement, it is equivalent to

  $$(AP+BL)P^{-1}(AP+BL)^\top - P + \Sigma \preceq 0$$

  Plug in $L = KP$, we have

  $$(A+BK)P(A+BK)^\top - P + \Sigma \preceq 0.$$

  The loss does not involve G so it does not change.
- Now, we need to prove that $\mathcal{L}(K)$ is equal to

  $$\begin{aligned} &\min_P \; \mathbf{Tr}((Q + K^\top RK)P) \\ &\text{s.t. } (A+BK)P(A+BK)^\top - P + \Sigma \preceq 0. \end{aligned} \tag{16}$$

  The last constraint can be written as

  $$(A+BK)P(A+BK)^\top - P + \Theta = 0,\; \Theta \succeq \Sigma.$$

- Denote the solution to $(A+BK)P(A+BK)^\top - P + \Theta = 0$ as $P(\Theta)$. $P(\Theta)$ for all $\Theta \succeq \Sigma$ covers the feasible points of (16). $P(\Theta)$ is expressed as:

  $$P(\Theta) = \sum_{t=0}^{\infty} (A+BK)^t \Theta ((A+BK)^\top)^t$$

  So $P(\Theta) \succeq P(\Sigma)$, $\forall \Theta \succeq \Sigma$. Since Q and $K^\top RK$ are positive semidefinite, $\mathbf{Tr}((Q + K^\top RK)P)$ achieves the minimum at $P = P(\Sigma)$.
- At the end, $P(\Sigma)$ is the Grammian $\mathbf{E}\sum_{t=0}^{\infty} x(t)x(t)^\top$ when $\mathbf{E}x(0)x(0)^\top = \Sigma$. We studied the connection between continous time Grammian (5) and the loss (6), a similar result holds for discrete time LQR:

  $$\mathbf{Tr}((Q + K^\top RK)P(\Sigma)) = \mathcal{L}(K).$$

We build the connection between minimizing $\mathcal{L}(K)$, and the convex optimization (15). We argued this pair of problems satisfies the assumptions of Theorem 1. Theorem 1 suggests that $\mathcal{L}(K)$ is gradient dominant, so we can approach K^* by gradient descent on K. This is essentially the conclusion of [1, 23]. Note that the proof of discrete time LQR [1, 23] and continuous time LQR [22, 24] cannot trivially extend to each other.

4.3 Minimizing L_2 gain

We quote from [30] the problem of minimizing the L_2 gain with static state feedback controller K. As discussed in [30, §6.3.2], this problem has an associated convex optimization problem and we can show it satisfies Assumption 1,2.

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

$$\dot{x} = Ax + Bu + B_w w,\ y = Cx + Du \tag{17}$$

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 [30, Sec 7.5.1]

$$\begin{aligned}
&\min_{L,P,\gamma}\ f(L,P,\gamma) := \gamma,\ \text{s.t.}\\
&\begin{bmatrix} AP + PA^\top + BL + L^\top B^\top + B_w B_w^\top & (CP + DL)^\top \\ CP + DL & -\gamma I \end{bmatrix}\\
&:= M(L,P,\gamma) \preceq 0.
\end{aligned} \tag{18}$$

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

[30, §6.3.2] suggests that L_2 gain is also the $\mathcal{H}_\infty$ norm of transfer function, so it covers the instances in [27]. We discuss this further in Appendix D.2.1.

4.4 Dissipativity

We use the dynamical system

$$\dot{x} = Ax + Bu + B_w w,\ y = Cx + Du + D_w w \tag{19}$$

The notion of dissipativity can be found in [30, §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 [30, Sec. 6.3.3, 7.5.2].

The dissipativity is defined as all $\eta > 0$ (we usually take the maximum one) that satisfy

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

Same as the last example, we use a state feedback controller K, and the goal is to find K^* that maximizes the dissipativity η. Same as before, let K be factorized as LP^{-1}. We can maximize the dissipativity η as a function of K. From [30, §7.5.2], we maximize η subject to the dissipativity constraint (21),

$$\max_{\eta,L,P}\ \eta, \tag{20}$$

$$\text{s.t.} \begin{bmatrix} AP + PA^\top + BL + L^\top B^\top & B_w - PC^\top - (DL)^\top \\ B_w^\top - CP - DL & 2\eta I - (D + D^\top) \end{bmatrix} \preceq 0. \tag{21}$$

With similar reasoning as in Sec. 4.4, the assumptions for Theorem 1 holds. Thus we can claim that all stationary points of $\mathcal{L}(K)$ are global minimum.

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

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

$$x(t+1) = A(t)x(t) + B(t)u(t) + w(t) \tag{22}$$

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

$$X = \begin{bmatrix} x(0) \\ ... \\ x(T) \end{bmatrix},\ U = \begin{bmatrix} u(0) \\ ... \\ u(T) \end{bmatrix},$$

$$W = \begin{bmatrix} x(0) \\ w(0) \\ ... \\ w(T-1) \end{bmatrix}, Z = \begin{bmatrix} 0 & 0 & ... & 0 & 0 \\ I & 0 & ... & 0 & 0 \\ 0 & I & ... & 0 & 0 \\ ... & & & & \\ 0 & 0 & ... & I & 0 \end{bmatrix},$$

$$\mathcal{A} = \operatorname{diag}(A(0), ..., A(T-1), 0),$$
$$\mathcal{B} = \operatorname{diag}(B(0), ..., B(T-1), 0)$$

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

$$u(t) = \sum_{i=0}^{t} K(t, t-i)x(i) \tag{23}$$

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 (23) and define (More examples of nonquadratic cost in [2, Sec 2.2])

$$\mathcal{L}(\mathcal{K}) = \sum_{t=0}^{T} x(t)^\top Q(t)x(t) + u(t)^\top R(t)u(t), \tag{24}$$

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

Convexification: The dynamics (22) 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:

$$\begin{bmatrix} I - Z\mathcal{A} & -Z\mathcal{B} \end{bmatrix} \begin{bmatrix} \Phi_X \\ \Phi_U \end{bmatrix} = I. \tag{25}$$

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 (25).

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

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

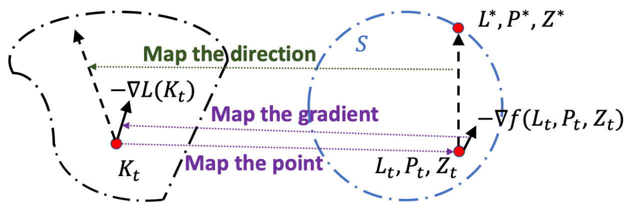

Figure 1: Mapping between nonconvex and convex landscapes. Suppose we run gradient descent at iteration t, for any controller K, we can map it to L, P, Z in the other parameterized space. and then we map the direction $(L^*, P^*, Z^*) - (L, P, Z)$ and the gradient $\nabla f(L, P, Z)$ back to the original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

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

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

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 (25):

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

This problem is convex and satisfy Assumption 1. [2, 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.

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 parameterized space with L, P, Z where the loss is convex.

For simplicity, we sketch the proof using Assumptions 1,3. For any point K, we can find a point (L, P, Z) in the parameterized space. If it is not the optimizer, we can find the line segment linking (L, P, Z) and the optimizer (L^*, P^*, Z^*). Note that the optimization problem is convex in this space so that $\langle \nabla f(L, P, Z), (L^*, P^*, Z^*) - (L, P, Z) \rangle$ is upper bounded by $f(L^*, P^*, Z^*) - f(L, P, Z)$. 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.

Before concluding, we remark that the assumptions in Theorem 1 come from an optimization theory perspective, and we do not dive into the control theoretic interpretations of the constants and assumptions. Our approach has the benefit that it unifies the analysis of many control problems in a single abstract result. We leave it to future work to refine the analysis to obtain the best case-specific convergence rates, and to provide an interpretation of the associated constants in terms of control theoretic notions.

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.

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

[15] 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.

[16] 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.

[17] 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.

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

[19] 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.

[20] K. Mårtensson, "Gradient methods for large-scale and distributed linear quadratic control," Ph.D. dissertation, Lund University, 2012.

[21] J. W. Roberts, I. R. Manchester, and R. Tedrake, "Feedback controller parameterizations for reinforcement learning," in *2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL).* IEEE, 2011, pp. 310–317.

[22] H. Mohammadi, A. Zare, M. Soltanolkotabi, and M. R. Jovanović, "Global exponential convergence of gradient methods over the nonconvex landscape of the linear quadratic regulator," in *2019 IEEE 58th Conference on Decision and Control (CDC).* IEEE, 2019, pp. 7474–7479.

[23] 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.

[24] J. Bu, A. Mesbahi, and M. Mesbahi, "Policy gradient-based algorithms for continuous-time linear quadratic control," *arXiv preprint arXiv:2006.09178*, 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, Z. Yang, and T. Basar, "Policy optimization provably converges to nash equilibria in zero-sum linear quadratic games," *Advances in Neural Information Processing Systems*, vol. 32, pp. 11 602–11 614, 2019.

[27] 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.

[28] 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.

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

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

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

[32] A. Zare, H. Mohammadi, N. K. Dhingra, T. T. Georgiou, and M. R. Jovanović, "Proximal algorithms for large-scale statistical modeling and sensor/actuator selection," *IEEE Transactions on Automatic Control*, vol. 65, no. 8, pp. 3441–3456, 2019.

[33] J. P. Jansch-Porto, B. Hu, and G. E. Dullerud, "Convergence guarantees of policy optimization methods for markovian jump linear systems," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 2882–2887.

[34] C. Scherer and S. Weiland, "Linear matrix inequalities in control," *Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands*, vol. 3, no. 2, 2000.

A Proof of the main theorem

Theorem 1. *We consider the problems*

$$\min_{K} \quad \mathcal{L}(K), \quad s.t. \quad K \in \mathcal{S}_K \tag{26}$$

$$\min_{L,P,Z} \quad f(L,P,Z), \quad s.t. \quad (L,P,Z) \in \mathcal{S} \tag{27}$$

Then we require assumptions either 1,3 or 1,2. Define K^* as the global minimum of $\mathcal{L}(K)$ in feasible region. Then we have (exact values in (33),(34)*)*

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

$$\|\nabla\mathcal{L}(K)\|_F \geq C_1(L,P)(\mathcal{L}(K) - \mathcal{L}(K^*)). \tag{28}$$

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

$$\|\nabla\mathcal{L}(K)\|_F \geq C_2(L,P)(\mu(\mathcal{L}(K) - \mathcal{L}(K^*)))^{1/2}. \tag{29}$$

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

$$C_1(L,P) = (\max\{\sigma_{\max}(L-L^*)\sigma_{\min}^{-1}(P), \sigma_{\max}(P-P^*)\sigma_{\min}^{-2}(P)\sigma_{\max}(L)\})^{-1}$$
$$C_2(L,P) = 2(\max\{\sigma_{\min}^{-1}(P), \sigma_{\min}^{-2}(P)\sigma_{\max}(L)\})^{-1}$$

Proof of Theorem 1. Denote $\mathcal{P}_{\mathcal{S}}(\nabla f(x))$ as the projection of $\nabla f(x)$ onto the descent cone of $\mathcal{S}$ at x. 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^*, P^*) be the optimal point in the parameterized space. We have $\mathcal{L}(K^*) = f(Z^*, L^*, P^*)$. Denote V as any matrix in K space, $\mathcal{P}_V$ is projection of a vector onto direction V, then

$$\nabla\mathcal{L}(K)^\top \nabla\mathcal{L}(K) \geq (\mathcal{P}_V \nabla\mathcal{L}(K))^\top \mathcal{P}_V \nabla\mathcal{L}(K) \tag{30}$$
$$= \left(\frac{\nabla\mathcal{L}(K)[V]}{\|V\|_F}\right)^2. \tag{31}$$

We look at any feasible controller K. We denote $\mathscr{Z}(L,P) \in \operatorname{argmin}_Z f(L,P,Z)$ subject to $(L,P,Z) \in \mathcal{S}$ (if there are multiple minimizers we pick any one). With either Assumption 3 or 2, we can define the mapping from K to (L,P,Z) respectively in one of the following ways:

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

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

Note f is convex, so

$$\begin{aligned}&\nabla f(L,P,Z)[(L,P,Z) - (L^*,P^*,Z^*)]\\ &\geq f(L,P,Z) - f(L^*,P^*,Z^*)\\ &= f(\mathscr{Z}(L,P),L,P) - f(\mathscr{Z}(L^*,P^*),L^*,P^*)\\ &= \mathcal{L}(K) - \mathcal{L}(K^*).\end{aligned} \tag{32}$$

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 $\Delta L = L^* - L$, $\Delta P = P^* - P$, and $V = \Delta L P^{-1} - LP^{-1}\Delta P P^{-1}$. Then

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

With Assumption 3, we continue with

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

With assumption 2, we continue with

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

and then

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

The inequality results in that $(L+t\Delta L, P+t\Delta P, Z+t\Delta Z)$ is feasible so the value is bigger than the optimal value. So the final inequality holds either by Assumption 3 or 2. So

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

Using (30) and (32), we have

$$\|\nabla\mathcal{L}(K)\|_F^2 \ge \frac{1}{\|V\|_F^2}(\mathcal{L}(K)-\mathcal{L}(K^*))^2 \tag{33}$$

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

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

then we have that

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

then

$$\begin{aligned}&\|\nabla\mathcal{L}(K)\|_F^2\\ &\ge \frac{1}{\|V\|_F^2}(\nabla f(L,P,Z)[(L,P,Z)-(L^*,P^*,Z^*)])^2\\ &\ge \frac{\mu\|(L,P,Z)-(L^*,P^*,Z^*)\|^2}{\|V\|_F^2}\\ &\quad \cdot (f(L,P,Z)-f(L^*,P^*,Z^*))\\ &= \frac{\mu(\|L^*-L\|^2+\|P^*-P\|^2+\|Z^*-Z\|^2)}{\|(L^*-L)P^{-1}-LP^{-1}(P^*-P)P^{-1}\|_F^2}\\ &\quad \cdot (f(L,P,Z)-f(L^*,P^*,Z^*))\\ &\ge \frac{\mu(\|L^*-L\|^2+\|P^*-P\|^2)}{\|(L^*-L)P^{-1}-LP^{-1}(P^*-P)P^{-1}\|_F^2}\\ &\quad \cdot (f(L,P,Z)-f(L^*,P^*,Z^*))\\ &\ge \frac{4\mu(f(L,P,Z)-f(L^*,P^*,Z^*))}{(\max\{\sigma_{\min}^{-1}(P), \sigma_{\min}^{-2}(P)\sigma_{\max}(L)\})^2}.\end{aligned} \tag{34}$$

B Constant for continuous time LQR

Thm. 1 asks for two constants C_1, C_2. They are bounded differently for different examples. As an instance, we will calculate the constant C_1 for continuous time LQR, quoted from [22] Appendix B. First $P \succ 0$, so we replace singular value by eigenvalue with P

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

We need upper bounds for P, L and a lower bound for $\lambda_{\min}(P)$ to guarantee C_1, C_2 finite. We will show the bounds within the sublevel set that $\{K : \mathcal{L}(K) \le a\}$. Since we can randomly initialize a feasible K_0 and run (projected) gradient descent method with respect to K if $\mathcal{L}(K)$ is gradient dominant, it is reasonable to assume that during all iterations of the optimization algorithm, the function value is always upper bounded by $\mathcal{L}(K_0)$, or some values related to $\mathcal{L}(K_0)$. So our derivation with a finite sublevel set is reasonable.
Suppose the matrices $Q, R \succ 0$, and we consider the sublevel set when $\mathcal{L}(K) \le a$. The sublevel set gives $\mathbf{Tr}(QP) + \mathbf{Tr}(LP^{-1}L^\top R) \le a$, so

$$\begin{aligned}
\lambda_{\min}(R)\lambda_{\max}^{-1}(P)\|L\|_F^2 &\le \lambda_{\min}(R)\|LP^{-1/2}\|_F^2 \\
&\le \mathbf{Tr}(LP^{-1}L^\top R) \\
&\le \mathbf{Tr}(QP) + \mathbf{Tr}(LP^{-1}L^\top R) \le a
\end{aligned}$$

So $\|L\|_F \le a(\lambda_{\max}(P)\lambda_{\min}^{-1}(R))^{1/2}$, and we know from [22] eq(34) $\mathbf{Tr}(P) \le a\lambda_{\min}^{-1}(Q)$. So we can bound P, L

$$\begin{aligned}
&\mathbf{Tr}(P) \le a\lambda_{\min}^{-1}(Q), \\
&\|L\|_F \le a(\lambda_{\min}(Q)\lambda_{\min}(R))^{-1/2}
\end{aligned}$$

Define

$$\nu = 4a\left(\sigma_{\max}(A)\lambda_{\min}^{-1/2}(Q) + \sigma_{\max}(B)\lambda_{\min}^{-1/2}(R)\right)^2$$

[32] eq(38,40) suggests $\lambda_{\min}(P) \ge \lambda_{\min}(\Sigma)/\nu$. In summary, we upper bounded L, and upper and lower bounded P in the sublevel set $\mathcal{L}(K) \le a$, and those bounds are also true for L^*, P^*. We can complete the calculation by inserting the bounds into $C_1(L,P)$.

$$\begin{aligned}
\nu &= 4a\left(\sigma_{\max}(A)\lambda_{\min}^{-1/2}(Q) + \sigma_{\max}(B)\lambda_{\min}^{-1/2}(R)\right)^2, \\
C_1(L,P) &= (\max\{2a\nu\lambda_{\min}^{-1}(\Sigma)\lambda_{\min}^{-1/2}(Q)\lambda_{\min}^{-1/2}(R), 2a^2\nu^2\lambda_{\min}^{-3/2}(Q)\lambda_{\min}^{-1/2}(R)\lambda_{\min}^{-2}(\Sigma)\})^{-1}.
\end{aligned}$$

$C_2(L,P)$ is calculated similarly with upper bound on P, L, P^{-1}.

C Strongly convex parameter of continuous time LQR

In our previous convexification of continuous time LQR (8), we translate the objective function as a linear function in the new variables L, P, Z. The problem (8) can be slightly reformulated as

$$\min_{L,P} \; f(L,P) := \mathbf{Tr}(QP) + \mathbf{Tr}(LP^{-1}L^\top R) \tag{35a}$$

$$\text{s.t. } \mathcal{A}(P) + \mathcal{B}(L) + \Sigma = 0, \; P \succ 0. \tag{35b}$$

Compared with (8), (35) does not contain the variable Z. Below, we will prove that the new objective function $f(L,P)$, restricted within the feasible set, is a strongly convex function, which is not the case for the linear objective (8). In Thm. 1, there is another result with strongly convex f and the gradient domminance parameter depends on the strongly convex parameter μ. We also calculate μ of $f(L,P)$ below.

Lemma 2. *Define a sublevel set of of f at level a, consisting of all L, P such that $f(L,P) \leq a$. Define*

$$
\begin{aligned}
\nu &= \frac{\lambda_{\min}(\Sigma)}{4}\left(\|A\|\lambda_{\min}^{-1/2}(Q) + \|B\|\lambda_{\min}^{-1/2}(R)\right), \\
\eta &= \left(\nu^{1/2}\|\mathcal{B}\|\lambda_{\min}(\Sigma)\lambda_{\min}(Q)\lambda_{\min}^{1/2}(R)\right)^{-1}, \\
\mu_0 &= \frac{2\lambda_{\min}(Q)\lambda_{\min}(R)}{a(1+a^2\eta)^2}.
\end{aligned}
$$

The strongly convex parameter μ of the function $f(L,P)$ restricted within the feasible set in (35) *satisfies*

$$\mu \geq (\|\mathcal{A}^{-1} \circ \mathcal{B}\| + 1)^{-1}\mu_0.$$

Proof. Denote $\mathcal{A}^{-1}$ as the inverse of $\mathcal{A}$, a linear operator such that $\mathcal{A}^{-1}(\mathcal{A}(P)) = P$. [22, Proposition 1] concludes that the following function $h(\cdot)$ is μ_0 strongly convex.

$$h(L) = f(L, -\mathcal{A}^{-1}(\mathcal{B}(L) + \Sigma)).$$

Define a perturbation direction $(\tilde{L}, \tilde{P})$. Any feasible perturbation at the point L, P will satisfy $\mathcal{A}(\tilde{P}) + \mathcal{B}(\tilde{L}) = 0$, so $\tilde{P} = -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L}))$.

Let the strongly convex parameter of f in the feasible directions be μ, we will show its connection with μ_0.

Let L be perturbed by $\tilde{L}$.

$$\nabla^2 h(L)[\tilde{L}, \tilde{L}] = \nabla^2 f(L,P)[(\tilde{L}, -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L}))), (\tilde{L}, -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L})))] \tag{36}$$

Due to the strong convexity of h,

$$\nabla^2 h(L)[\tilde{L}, \tilde{L}] \geq \frac{\mu_0\|\tilde{L}\|_F^2}{2}.$$

We perturb f at (L,P) in direction $(\tilde{L}, \tilde{P}) = (\tilde{L}, -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L})))$. The strongly convex parameter of f in feasible directions is defined as the positive number mu such that

$$\nabla^2 f(L,P)[(\tilde{L}, \tilde{P}), (\tilde{L}, \tilde{P})] \geq \frac{\mu(\|\tilde{P}\|_F^2 + \|\tilde{L}\|_F^2)}{2}$$

for all $(\tilde{L}, \tilde{P})$ such that $\tilde{P} = -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L}))$. The directional Hessian is

$$\nabla^2 f(L,P)[(\tilde{L}, \tilde{P}), (\tilde{L}, \tilde{P})] = \nabla^2 f(L,P)[(\tilde{L}, -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L}))), (\tilde{L}, -\mathcal{A}^{-1}(\mathcal{B}(\tilde{L})))] \tag{37}$$

Indeed (37) is same as (36). So that

$$
\begin{aligned}
&\nabla^2 f(L,P)[(\tilde{L}, \tilde{P}), (\tilde{L}, \tilde{P})] \geq \frac{\mu_0\|\tilde{L}\|_F^2}{2} \\
&= \frac{\|\tilde{P}\|_F^2 + \|\tilde{L}\|_F^2}{2} \cdot \frac{\mu_0\|\tilde{L}\|_F^2}{\|\tilde{P}\|_F^2 + \|\tilde{L}\|_F^2} \\
&= \frac{\|\tilde{P}\|_F^2 + \|\tilde{L}\|_F^2}{2} \cdot \frac{\mu_0\|\tilde{L}\|_F^2}{\|\mathcal{A}^{-1}(\mathcal{B}(\tilde{L}))\|_F^2 + \|\tilde{L}\|_F^2}
\end{aligned}
$$

So

$$\mu \geq (\|\mathcal{A}^{-1} \circ \mathcal{B}\| + 1)^{-1}\mu_0.$$

D Checking the assumptions for examples

D.1 Markov jump linear system

We generalize the discrete time linear system in this part. Suppose there are N linear systems, the i-th one being

$$x(t+1) = A_i x(t) + B_i u(t)$$

Now we study the Markov jump linear system [33]. At each time t, the dynamics linking $x(t+1)$ and the past state and input $x(t), u(t)$ is given by

$$x(t+1) = A_{w(t)} x(t) + B_{w(t)} u(t),\ w(t) \in [N] := \{1, ..., N\}.$$

So at any time t, the system is chosen from the N candidate linear systems. The transition of the linear systems, or the transition of $w(t)$, follows a probabilistic model

$$\mathbf{Pr}(w(t+1) = j | w(t) = i) = \rho_{ij} \in [0,1],\ \forall t \geq 0.$$

For the i-th system, we will use a state feedback controller K_i. Let $K = [K_1, ..., K_N]$. Define the loss as

$$\mathcal{L}(K) = \mathbf{E}_{w,x_0} \sum_{t=0}^{\infty} x(t)^\top Q x(t) + u(t)^\top R u(t),\ u(t) = K_{w(t)} x(t)$$

We propose the following convexification: Denote $P_0, P_1, ..., P_N \in \mathbb{R}^{n\times n} \succ 0$, $G_1, ..., G_N \in \mathbb{R}^{n\times n} \succeq 0$, $L_1, ..., L_N \in \mathbb{R}^{p\times n}$, $Z_0, Z_1, ..., Z_N \in \mathbb{R}^{p\times p} \succeq 0$. Suppose $\mathbf{Pr}(w(0) = i) = p_i$.

Then

$$\min\ \mathbf{Tr}(QP_0) + \mathbf{Tr}(Z_0 R) \tag{38a}$$

$$\text{s.t. } P_0 = \sum_{i=1}^{N} p_i P_i,\ Z_0 = \sum_{i=1}^{N} p_i Z_i, \tag{38b}$$

$$P_i \succ 0,\ G_i - P_i + \Sigma = 0,\ P^i = \sum_{j=1}^{N} \rho_{ij} P_j \tag{38c}$$

$$\begin{bmatrix} Z_i & L_i \\ L_i^\top & P_i \end{bmatrix} \succeq 0,\ \begin{bmatrix} G_i & A_i P^i + B_i L_i \\ (A_i P^i + B_i L_i)^\top & P^i \end{bmatrix} \succeq 0,\ \forall i \in [N] \tag{38d}$$

The variables are P^i, P_i, G_i, L_i, Z_i for all $i \in [N]$ and P_0, Z_0. The mapping between the controller K_i and the new variables is $K_i = L_i (P^i)^{-1}$.

Now we will show the connection between $\mathcal{L}(K)$ and the convexification. With (38), if we fix the other variables and minimize over Z_i, then $Z_i = L_i P_i L_i^\top$. Later we will fix all $Z_i = L_i P_i L_i^\top$.

Then, we minimize over G_i's. We fix the stabilizing controllers K, in other words we fix the closed loop systems. Denote $\mathfrak{A}_i(P) : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n} = (A_i + BK_i) P (A_i + BK_i)^\top$. $\mathfrak{A}_i$'s are positive definite operators and the operator norms of $\mathfrak{A}_i$ are all less than 1. Denote

$$\mathfrak{A} : \mathbb{R}^{Nn\times n} \to \mathbb{R}^{Nn\times n} = \text{diag}(\mathfrak{A}_1, ..., \mathfrak{A}_N).$$

Denote $\Omega \in \mathbb{R}^{Nn\times Nn}$ as a block matrix, each block being $\mathbb{R}^{n\times n}$, and the i,j-th block is $\rho_{ij} I_n$.

Because of

$$G_i - P_i + \Sigma = 0,\ P^i = \sum_{j=1}^{N} \rho_{ij} P_j,\ \begin{bmatrix} G_i & A_i P^i + B_i L_i \\ (A_i P^i + B_i L_i)^\top & P^i \end{bmatrix} \succeq 0$$

denote $P = \text{diag}(P_1, ..., P_N)$, and $\Sigma^N =\in \mathbb{R}^{Nn\times Nn} = \text{diag}(\Sigma, ..., \Sigma)$ then we have

$$\mathfrak{A}(\Omega P) - P + \Sigma^N \preceq 0 \tag{39}$$

There exists feasible G_i's such that for all $i = 1, ..., N$, $G_i = \mathfrak{A}_i(P)$ satisfying the constraints in (38), which maps to the true dynamic of the system and P_i's are the Grammians of the states starting from system i. Denote the solution P satisfying LHS of (39) being equal to 0 as $\tilde{P}$. The value of f at $\tilde{P}$ (with other variables properly chosen, such as $L_i = K_i\tilde{P}^i$, etc.) equals $\mathcal{L}(K)$.

We take any other feasible P. Let $\Delta = P - \tilde{P}$. We have

$$\Delta - \mathfrak{A}(\Omega\Delta) \succeq 0$$

We will prove $\Delta \succeq 0$. Note that, the operator $I - \mathfrak{A} \circ \Omega$ is a block M-operator, whose inverse is a block positive definite operator. Denote this operator as $(I - \mathfrak{A} \circ \Omega)^{-1}$. Then, block positive definite operator preserves the inequality, meaning

$$\begin{aligned} \Delta - \mathfrak{A}(\Omega\Delta) &\succeq 0 \\ \Rightarrow (I - \mathfrak{A} \circ \Omega)(\Delta) &\succeq 0 \\ \Rightarrow (I - \mathfrak{A} \circ \Omega)^{-1}(I - \mathfrak{A} \circ \Omega)(\Delta) &\succeq 0 \\ \Rightarrow \Delta &\succeq 0 \end{aligned}$$

This means, if we fix Z and K, any feasible point P will satisfy $P \succeq \tilde{P}$. Then with the similar argument as single linear system case, we know the minimum of f will be same as the value of f at $\tilde{P}$, which equals $\mathcal{L}(K)$.

That says, if we fix K and minimize (38) with extra constraints $K_i = L_i(P^i)^{-1}$, then the optimal value is $\mathcal{L}(K)$. This verifies Assumption 2, and overall, we conclude that $\mathcal{L}(K)$ is gradient dominant, satisfying our main theorem.

D.2 Minimizing L_2 gain

In this section, we will show that Sec. 4.3 satisfy Assumption 1,2. Thus Theorem 1 applies to this problem, which means all stationary points of $\mathcal{L}(K)$ are global minimum.

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

$$\dot{x} = Ax + Bu + B_w w, \; y = Cx + Du \tag{40}$$

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

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

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

Proof. We have $\mathcal{L}(K^*) = f(L^*, P^*, \gamma^*)$ from [30]. Note that, K^* can map to different pairs (L, P) whenever $LP^{-1} = K^*$, and γ^* associates to one pair among them. We can equivalently formulate

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

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

$$\mathcal{L}(K) = \min_{L,P,\gamma} \gamma \tag{41a}$$

$$\text{s.t. } M(L, P, \gamma) \preceq 0, \; LP^{-1} = K. \tag{41b}$$

Note that, the intermediate step [30, Sec 7.5.1] is

$$\mathcal{L}(K) = \min_{P,\gamma} \gamma, \quad \text{s.t.} \tag{42a}$$

$$\begin{bmatrix} (A+BK)P + P(A+BK)^\top + B_w B_w^\top & P^\top (C+DK)^\top \\ (C+DK)P & -\gamma I \end{bmatrix} \preceq 0. \tag{42b}$$

Denote the optimizer of (41) by $\hat{L}, \hat{P}, \hat{\gamma}$, and the optimizer of (42) by $\check{P}, \check{\gamma}$.

Note $\hat{\gamma} \leq \check{\gamma}$. If (41) is not true, $\hat{\gamma} < \check{\gamma}$, we can replace $\check{P}, \check{\gamma}$ with $\hat{P}, \hat{\gamma}$ in (42) and it's still feasible. Thus the optimality condition of $\check{P}, \check{\gamma}$ in (42) is violated, which contradicts the assumption that (41) is not true. Then we claim that (41) is true.

D.2.1 Connection with $\mathcal{H}_\infty$ robustness

In this part, we will explain the connection between the above example with [27]. [27] suggestsan $\mathcal{H}_\infty$ bound constraint, where the $\mathcal{H}_\infty$ norm of the transfer function of y to w, is upper bounded by a number γ. [30, 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

$$\sup_{\|w\|_2=1} \|y\|_2, \text{ s.t., dynamics (40)}$$

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 $M(L, P, \gamma) \preceq 0$ as a constraint. The objective (loss) functions discussed in [27] have P as variable (we use P above) and they are convex, thus these instances has convex formulation and can be covered by our result.

E A more general description of Assumption 2

Note that, in Assumption 2, we require $K = LP^{-1}$, which is still a specific mapping. We choosed this mapping because this is frequently used in convexifying the optimal control problem. For example, with the continuous time LQR problem motivated in Sec. 2, the mapping between K and L, P is (almost) the only widely used convexification method. If we choose another change of variable, the resulting objective function is usually not convex in the new variables. It is not easy to find other mappings suitable for convexifying these problems.

On the other hand, although the mapping $K = LP^{-1}$ is studied, we can generalize Thm. 1 with arbitrary mappings – of course, the new mappings have to satisfy a few assumptions to preserve gradient dominance.

Here we will propose the following assumptions which replace the mapping by a general notation Φ. In the following proposal, we will study the original optimization problem (43), and map it to a convex optimization problem (44) where the mapping between K and the variables of the other problem $P_1, ..., P_m$ is abstractly denoted by $K = \Phi(P)$ in (45).

Suppose we consider the problems

$$\min_K \quad \mathcal{L}(K), \quad \text{s.t. } K \in \mathcal{S}_K \tag{43}$$

and

$$\min_P \quad f(P), \quad \text{s.t.} \quad P \in \mathcal{S} \tag{44}$$

P can be a shorthand as concatenation of many variables, just as an expression. For example, P represent P, L, Z of continuous LQR. With

Assumption 4. *The feasible set $\mathcal{S}$ is convex in P. The cost function $f(P)$ is convex, finite and differentiable in $P \in \mathcal{S}$. $\mathcal{L}(K)$ is Lipschitz in K.*

Assumption 5. *Assume we can express $\mathcal{L}(K)$ as:*

$$\mathcal{L}(K) = \min_{P} \; f(P), \; s.t. \; P \in \mathcal{S}, \; K = \Phi(P). \tag{45}$$

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

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

for any $P \in \mathcal{S}$ and any perturbation $\mathrm{d}P$ such that $\mathrm{d}P$ is in the descent cone of $\mathcal{S}$ at P.

Thm. 1 holds with the above Assumptions 4, 5, which means the loss $\mathcal{L}(K)$ is gradient dominant in K given the above assumptions. It generalizes beyond the specific mapping $\Phi(P, L) = LP^{-1}$ to a more general definition, and we propose some instances of convexification with different Φ in the subsections.

Remark 2. *Denote $\Delta K = \Psi(P)[P^* - P]$. Let $\nabla\mathcal{L}(K)[\Delta K]$ be the directional derivative of $\mathcal{L}(K)$ in direction ΔK. Then with Assumptions 4, 5 we have*

$$\nabla\mathcal{L}(K)[\Delta K] \leq \mathcal{L}(K^*) - \mathcal{L}(K)$$

Proof. Suppose $f(P)$ is convex in P, and the optimizer of (44) is P^*. Denote

$$P = \mathrm{argmin}_{P'} \; f(P'), \; \text{s.t. } P' \in \mathcal{S}, \; K = \Phi(P')$$

and

$$\Delta P = P^* - P, \; \Delta K = \Psi(P)[\Delta P]$$

We take the directional derivative and get

$$\begin{aligned}
\nabla\mathcal{L}(K)[\Delta K] &= \lim_{t\to 0} \frac{\mathcal{L}(K + t\Delta K) - \mathcal{L}(K)}{t} \\
&= \lim_{t\to 0} \frac{\mathcal{L}(K + t\Psi(P)[\Delta P]) - f(P)}{t} \qquad (46) \\
&= \lim_{t\to 0} \frac{\mathcal{L}(\Phi(P) + t\Psi(P)[\Delta P]) - f(P)}{t} \qquad (47) \\
&= \lim_{t\to 0} \frac{\mathcal{L}(\Phi(P + t\Delta P) - o(t)) - f(P)}{t} \qquad (48) \\
&= \lim_{t\to 0} \frac{\mathcal{L}(\Phi(P + t\Delta P)) - f(P)}{t} \\
&= \lim_{t\to 0} \frac{\min_{P'\in\mathcal{S},\; \Phi(P+t\Delta P)=\Phi(P')} \; f(P') - f(P)}{t} \qquad (49) \\
&= \lim_{t\to 0} \frac{\min_{P'\in\mathcal{S},\; \Phi(P+t\Delta P)=\Phi(P')} \; f(P') - f(P + t\Delta P) + f(P + t\Delta P) - f(P)}{t} \\
&= \lim_{t\to 0} \frac{\min_{P'\in\mathcal{S},\; \Phi(P+t\Delta P)=\Phi(P')} \; f(P') - f(P + t\Delta P)}{t} + \nabla f(P)[\Delta P] \qquad (50)
\end{aligned}$$

(46) and (47) replaces ΔK and K with expressions in P and ΔP. (48) applies the Taylor expansion of Φ:

$$\Phi(P + t\Delta P) - (\Phi(P) + t\Psi(P)[\Delta P]) = o(t)$$

(49) applies Assumption 5, and we plug in $K = \Phi(P + t\Delta P)$. (50) applied the definition of directional derivative

$$\nabla f(P)[\Delta P] = \lim_{t\to 0} \frac{f(P + t\Delta P) - f(P)}{t}.$$

Now we bound the first term of (50). Note that, since $P + t\Delta P$ for $t > 0$ and $t \to 0$ belongs to the line segment from P to P^*. Since $\mathcal{S}$ is a convex set, we know that the line segment between to feasible points P^* and P is in $\mathcal{S}$. then

$$P + t\Delta P \in \{P' \mid P' \in \mathcal{S},\ \Phi(P + t\Delta P) = \Phi(P')\}$$

so that $f(P + t\Delta P)$ is no less than the minimum of the optimization problem (45),

$$\lim_{t\to 0} \frac{\min_{P' \in \mathcal{S},\ \Phi(P+t\Delta P)=\Phi(P')}\ f(P') - f(P + t\Delta P)}{t} \leq 0.$$

$\nabla f(P)[\Delta P]$ is the directional derivative of $f(P)$ in the direction of $P^* - P$, for a convex function f, if P is not an optimizer, $\nabla f(P)[\Delta P]$ is upper bounded by $f(P^*) - f(P) = \mathcal{L}(K^*) - \mathcal{L}(K) < 0$.

E.1 Distributed finite horizon LQR

[28] 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 (Sec. 4.5), and we can use the same notation $X, U, W, Z, \mathcal{A}, \mathcal{B}$. We further define

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

Now we will consider

$$u(t) = \sum_{i=0}^{t} K(t, t-i)y(i) \tag{51}$$

The $\mathcal{K}$ matrix is same as SLS there. [28] studies the problem under the context of distributed control. One searches for the controller $K \in \mathcal{S}_K$ where $\mathcal{S}_K$ a subset of controllers. In distributed control, there is a graph model for controllers such that the i-th controller might not be able to access the state j for (i, j) in a set of indices $\mathcal{S}_{\text{idx}}$. In this case, $K_{i,j} = 0$ is an extra constraint for the control problem. Therefore, if one searches for the optimal controller in $\mathcal{S}_K$, we can define the subspace

$$\mathcal{S}_K := \{K \mid K_{i,j} = 0,\ \forall (i, j) \in \mathcal{S}_{\text{idx}}\}$$

The extra constraint is not always easily handled, but [28, Sec. 3] proposes an extra assumption, called quadratic invariance (QI) that makes the problem convexifiable.

QI means that, for all $K \in \mathcal{S}_K$, $KCP_{12}K \in \mathcal{S}_K$. QI leads to a convexification. The loss function is same as (24):

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

It is proven that, define

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

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 $\mathcal{G}$. $\mathcal{G}$ is in the same subspace as $\mathcal{K}$ determined by $\mathcal{S}_K$. Indeed, the mapping satisfies the Assumptions 4, 5, and the exact formulation of the two optimization problems are described in [28, 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.

E.2 Multi-objective and mixed controller design

In this part, we study a few synthesis problems with dynamical controllers, where the objectives are transfer functions of the closed form system. We study the dynamical system with state, disturbance, input, output, and controller's input x, w, u, z, y with the following dynamics

$$\begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix} = \begin{bmatrix} A & B_w & B \\ C_z & D & E \\ C & F & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix} \tag{52}$$

The dynamical controller follows

$$\begin{bmatrix} \dot{x}_c \\ u \end{bmatrix} = \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} x_c \\ y \end{bmatrix} \tag{53}$$

and we will design the controller $\begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix}$ to minimize certain cost objectives in the next few subsections. [34, eq(4.2.15)] studies the parameterization of the problem, by introducing a list of variables $v = [X, Y, K, L, M, N]$, and invertible matrices U, V such that $UV^\top = I - XY$. The Lyapunov matrix P of the closed loop system and A_c, B_c, C_c, D_c are the unique solution of

$$\begin{bmatrix} Y & V \\ I & 0 \end{bmatrix} P = \begin{bmatrix} I & 0 \\ X & U \end{bmatrix} \tag{54}$$

$$\begin{bmatrix} K & L \\ M & N \end{bmatrix} = \begin{bmatrix} U & XB \\ 0 & I \end{bmatrix} \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} V^\top & 0 \\ CY & I \end{bmatrix} + \begin{bmatrix} XAY & 0 \\ 0 & 0 \end{bmatrix} \tag{55}$$

The change of variable enables us to make some control problems as convex optimization, listed below. For simplicity of notation, let

$$\mathscr{X} = \begin{bmatrix} Y & I \\ I & X \end{bmatrix}, \ \mathscr{A} = \begin{bmatrix} AY + BM & A + BNC \\ K & AX + LC \end{bmatrix} \tag{56}$$

$$\mathscr{B} = \begin{bmatrix} B_w + BNF \\ XB_w + LF \end{bmatrix}, \ \mathscr{C} = \begin{bmatrix} C_zY + EM & C_z + ENC \end{bmatrix}, \ \mathscr{D} = D + ENF \tag{57}$$

Remark 3. *The mapping in* (54), (55) *can be written as*

$$\begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} = \Phi(v)$$

where Φ plays the role in (45)*. We propose a few control problems with convexification in the next few subsections. All these objective functions with respect to $v = [X, Y, K, L, M, N]$ being convex, and the original control problem with the respective loss functions and the convexified forms satisfy Assumptions 4, 5. Thus the loss functions with respect to matrix $\begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix}$ are gradient dominant.*

In the following subsections, we refer to the result of [34] that, the optimal $\mathcal{H}_\infty$ design, $\mathcal{H}_2$ design and the so-called multi-objective design that is a combination of the previous two, can be convexified with our proposed way. Thus, the three objectives with respect to controller K are all gradient dominant.

E.2.1 $\mathcal{H}_\infty$ design

[34, Sec. 4.2.3] The transfer function of the system is

$$\mathscr{T} = \mathscr{C}(sI - \mathscr{A})^{-1}\mathscr{B} + \mathscr{D} \tag{58}$$

The goal in this part is to minimize the $\mathcal{H}_\infty$ norm of the transfer function by designing the optimal controller. The problem with its raw form is to minimize the $\mathcal{H}_\infty$ norm over A_c, B_c, C_c, D_c, and we will propose the convexification – the change of variable trick such that the argument we minimize over becomes v. We solve

$$\min\ \gamma, \tag{59}$$

$$\text{s.t.} \quad \mathscr{X} \succeq 0, \begin{bmatrix} \mathscr{A}^\top + \mathscr{A} & \mathscr{B} & \mathscr{C}^\top \\ \mathscr{B}^\top & -\gamma I & \mathscr{D}^\top \\ \mathscr{C} & \mathscr{D} & -\gamma I \end{bmatrix} \preceq 0 \tag{60}$$

If we fix all other parameters and optimize over γ, then $\gamma^*(v)$ is the $\mathcal{H}_\infty$ value of the closed loop system with the mapping from v to controller by (54), (55). If we minimize over γ and v, then we can get optimal $\mathcal{H}_\infty$ design.

E.2.2 $\mathcal{H}_2$ design

[34, Sec. 4.2.5] This part is similar to $\mathcal{H}_\infty$ design. Suppose the goal is to minimize $\|\mathscr{T}\|_{\mathcal{H}_2}$, one can alternatively solve

$$\min\ \gamma, \tag{61}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathscr{A}^\top + \mathscr{A} & \mathscr{B} \\ \mathscr{B}^\top & -\gamma I \end{bmatrix} \preceq 0,\ \mathscr{D} = 0 \tag{62}$$

$$\begin{bmatrix} \mathscr{X} & \mathscr{C}^\top \\ \mathscr{C} & Z \end{bmatrix} \succeq 0,\ \mathbf{Tr}(Z) \leq \gamma \tag{63}$$

If we fix all other parameters and optimize over γ, Z, then $\gamma^*(v)$ is the $\mathcal{H}_2$ value of the closed loop system with the mapping from v to controller by (54), (55). If we minimize over γ, Z and v, then we can get optimal $\mathcal{H}_2$ design.

E.2.3 Multi-objective

[34, Sec. 4.3] Let the system be

$$\begin{bmatrix} \dot{x} \\ z_1 \\ z_2 \\ y \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 & B \\ C_1 & D_1 & D_{12} & E_1 \\ C_2 & D_{21} & D_2 & E_2 \\ C & F_1 & F_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ w_1 \\ w_2 \\ u \end{bmatrix} \tag{64}$$

Now we study the mixed design for $\mathcal{H}_\infty$ design from z_1 to w_1 and $\mathcal{H}_2$ design from z_2 to w_2. We keep the mappings in (54), (55) and the change of parameter (56), but replace (57) by

$$\mathscr{B}_i = \begin{bmatrix} B_i + BNF_i \\ XB_i + LF_i \end{bmatrix},\ \mathscr{C}_i = \begin{bmatrix} C_iY + E_iM & C_i + E_iNC \end{bmatrix},\ \mathscr{D}_i = D_i + E_iNF_i \tag{65}$$

for $i = 1, 2$. Suppose we are given a positive number λ and hope to study $\|\mathscr{T}_1\|_{\mathcal{H}_\infty} + \lambda\|\mathscr{T}_2\|_{\mathcal{H}_2}$ where $\mathscr{T}_i$ is the transfer function of the i-th system (z_1 to w_1, z_2 to w_2), then we can write

$$\min\ \gamma_1 + \lambda\gamma_2, \tag{66}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathscr{A}^\top + \mathscr{A} & \mathscr{B}_1 & \mathscr{C}_1^\top \\ \mathscr{B}_1^\top & -\gamma_1 I & \mathscr{D}_1^\top \\ \mathscr{C}_1 & \mathscr{D}_1 & -\gamma_1 I \end{bmatrix} \preceq 0 \tag{67}$$

$$\begin{bmatrix} \mathscr{A}^\top + \mathscr{A} & \mathscr{B}_2 \\ \mathscr{B}_2^\top & -\gamma_2 I \end{bmatrix} \preceq 0,\ \mathscr{D}_2 = 0 \tag{68}$$

$$\begin{bmatrix} \mathscr{X} & \mathscr{C}_2^\top \\ \mathscr{C}_2 & Z \end{bmatrix} \succeq 0,\ \mathbf{Tr}(Z) \leq \gamma_2 \tag{69}$$

If we fix all other parameters and optimize over γ_1, γ_2, Z, then the function value is the mixed $\mathcal{H}_\infty/\mathcal{H}_2$ value of the closed loop system with the mapping from v to controller by (54), (55). If we minimize over γ_1, γ_2, Z and v, then we can get the optimal mixed $\mathcal{H}_\infty/\mathcal{H}_2$ design.

E.2.4 Robust state feedback control

[34, Sec. 8.1.2] We study the robust state feedback control problem. We apply the system model (64). "State feedback" means that $C = I$ and $F_1, F_2 = 0$. The connection between w_1 and z_1 is an uncertain channel

$$w_1(t) = \Delta(t) z_1(t)$$

for any

$$\Delta(t) \in \Delta_c := \mathrm{conv}\{0, \Delta_1, ..., \Delta_{N_\Delta}\}$$

The goal is to minimize a certain norm of the transfer function from z_2 to w_2, which can be found in the previous part. We consider minimizing the norm under an extra constraint when the closed loop system achieves robust stability with z_1 to w_1 (z_1 with finite norm) and robust quadratic performance with z_2 to w_2 via a matrix P_p. The robust quadratic performance is defined as: there exists a matrix

$$P_p = \begin{bmatrix} \tilde{Q}_p & \tilde{S}_p \\ \tilde{S}_p^\top & \tilde{R}_p \end{bmatrix}, \; P_p^{-1} = \begin{bmatrix} Q_p & S_p \\ S_p^\top & R_p \end{bmatrix}$$

such that $\tilde{R}_p \succ 0, Q_p \prec 0$, and

$$\int_0^\infty \begin{bmatrix} w_2(t) \\ z_2(t) \end{bmatrix}^\top P_p \begin{bmatrix} w_2(t) \\ z_2(t) \end{bmatrix} dt \le \epsilon \|w_2\|_{\mathcal{H}_2}^2$$

for some $\epsilon > 0$.

[34, Sec. 8.1.2] proposes a convexification of the extra constraint. Define new variables Q, S, R in addition to $v = [X, Y, K, L, M, N]$, and let $\mathscr{M}$ replace

$$\begin{bmatrix} -(AY + BM)^\top & -(C_1 Y + E_1 M)^\top & -(C_2 Y + E_2 M)^\top \\ I & 0 & 0 \\ -B_1^\top & -D_1^\top & -D_{21}^\top \\ 0 & I & 0 \\ -B_2^\top & -D_{12}^\top & -D_2^\top \\ 0 & 0 & I \end{bmatrix}.$$

The constraints, which is proven to be convex [34, Sec. 8.1.2] can be written as

$$R \succ 0, \; Q \prec 0, \; \begin{bmatrix} I \\ -\Delta_j \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} I \\ -\Delta_j \end{bmatrix} \prec 0, \; \forall j \in [N_\Delta]$$

$$Y \succ 0, \; \mathscr{M}^\top \begin{bmatrix} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^\top & R & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^\top & R_p \end{bmatrix} \mathscr{M} \succ 0.$$

For example, if we aim to minimize the $\mathcal{H}_2$ norm of the transfer function from z_2 to w_2, then we can minimize γ_2 subject to (68), (69) and the constraints above. The main theorem of this paper suggests that, with the convexification, if we apply policy gradient descent with respect to $\mathcal{H}_2$ norm of the transfer function from z_2 to w_2 with robust stability (z_1 with finite norm) and robust quadratic performance constraints, then policy gradient descent converges to globally optimal controller.

E.2.5 Discrete time system

[34, Sec. 4.6] Suppose we study the discrete time system, and we define the system in a similar way of defining the continuous time system:

$$\begin{bmatrix} x(t+1) \\ z_1(t) \\ z_2(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 & B \\ C_1 & D_1 & D_{12} & E_1 \\ C_2 & D_{21} & D_2 & E_2 \\ C & F_1 & F_2 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ w_1(t) \\ w_2(t) \\ u(t) \end{bmatrix}, \; \begin{bmatrix} x_c(t+1) \\ u(t) \end{bmatrix} = \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} x_c(t) \\ y(t) \end{bmatrix} \tag{70}$$

Now we study the mixed design for $\mathcal{H}_\infty$ design from z_1 to w_1 and $\mathcal{H}_2$ design from z_2 to w_2. Suppose we are given a positive number λ and hope to study $\|\mathscr{T}_1\|_{\mathcal{H}_\infty} + \lambda\|\mathscr{T}_2\|_{\mathcal{H}_2}$ where $\mathscr{T}_i$ is the transfer function of the i-th system (z_1 to w_1, z_2 to w_2), then we can write

$$\min \ \gamma_1 + \lambda\gamma_2, \tag{71}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathscr{X} & 0 & \mathscr{A}^\top & \mathscr{C}_1^\top \\ 0 & \gamma_1 I & \mathscr{B}_1^\top & \mathscr{D}_1 \\ \mathscr{A} & \mathscr{B}_1 & \mathscr{X} & 0 \\ \mathscr{C}_1 & \mathscr{D}_1 & 0 & \gamma_1 I \end{bmatrix} \succ 0, \ \mathbf{Tr}(Z) \le \gamma_2, \tag{72}$$

$$\begin{bmatrix} \mathscr{X} & \mathscr{A} & \mathscr{B}_2 \\ \mathscr{A}^\top & \mathscr{X} & 0 \\ \mathscr{B}_2^\top & 0 & \gamma_2 I \end{bmatrix} \succ 0, \quad \begin{bmatrix} \mathscr{X} & 0 & \mathscr{C}_2 \\ 0 & \mathscr{X} & \mathscr{D}_2 \\ \mathscr{C}_2 & \mathscr{D}_2 & Z \end{bmatrix} \succ 0 \tag{73}$$

If we fix all other parameters and optimize over γ_1, γ_2, Z, then the function value is the mixed $\mathcal{H}_\infty/\mathcal{H}_2$ value of the closed loop system with the mapping from v. If we minimize over γ_1, γ_2, Z and v, then we can get the optimal mixed $\mathcal{H}_\infty/\mathcal{H}_2$ design.

F 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 2. *(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

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

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

G 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 (9) as a *convex* optimization problem (10)? In this section we propose the following lemma.

Lemma 3. *Suppose Assumptions 1, 3 hold, and $\mathcal{L}(LP^{-1})$ as a function of L, P is differentiable. We define the notation $\nabla^2_{L,P}\mathcal{L}(LP^{-1})[\Gamma_1, \Gamma_2]$ as in (74). If $\nabla^2_{L,P}\mathcal{L}(LP^{-1})[\Gamma_1, \Gamma_2] > 0$ for all $(L, P) \in \mathcal{S}$ and all (Γ_1, Γ_2) such that $\mathcal{A}(\Gamma_2) + \mathcal{B}(\Gamma_1) = 0$, then we can define a convex function $f(L, P)$ so that Assumption 1 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 (11). We know from Assumption 3 that, $f(L, P) = \mathcal{L}(K) = \mathcal{L}(LP^{-1})$ is convex in L, P. We take the Hessian and ask for

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

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 $\nabla^2_{L,P}\mathcal{L}(LP^{-1})[\Gamma_1, \Gamma_2]$ as

$$\begin{aligned} &\nabla^2_{L,P}\mathcal{L}(LP^{-1})[\Gamma_1, \Gamma_2] \\ :=\ &\nabla^2\mathcal{L}(LP^{-1})[\Gamma_1 G^{-2}, \Gamma_1] \\ &- 2\nabla^2\mathcal{L}(LP^{-1})[\Gamma_1, LP^{-3}\Gamma_2] \\ &- 2\langle \Gamma_1, \nabla\mathcal{L}(LP^{-1})P^{-1}\Gamma_2 P^{-1}\rangle \\ &+ 2\langle \Gamma_2, LP^{-1}\Gamma_2 P^{-1}\nabla\mathcal{L}(LP^{-1})P^{-1}\rangle \\ &+ \nabla^2\mathcal{L}(LP^{-1})[LP^{-2}\Gamma_2, LP^{-2}\Gamma_2] \end{aligned} \tag{74}$$

This is the directional Hessian of $\mathcal{L}$ with respect to (L, P) in direction (Γ_1, Γ_2). Thus, if $\nabla^2_{L,P}\mathcal{L}(LP^{-1})[\Gamma_1, \Gamma_2] > 0$ for all $(L, P) \in \mathcal{S}$ 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 $f(L, P)$ is convex in L, P and the convexification can be made.