2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

Global exponential convergence of gradient methods over the nonconvex landscape of the linear quadratic regulator

Hesameddin Mohammadi, Armin Zare, Mahdi Soltanolkotabi, and Mihailo R. Jovanović

Abstract— **In large-scale and model-free settings, first-order algorithms are often used in an attempt to find the optimal control action without identifying the underlying dynamics. The convergence properties of these algorithms remain poorly understood because of nonconvexity. In this paper, we revisit the *continuous-time* linear quadratic regulator problem and take a step towards demystifying the efficiency of gradient-based strategies. Despite the lack of convexity, we establish a linear rate of convergence to the globally optimal solution for the gradient descent algorithm. The key component of our analysis is that we relate the gradient-flow dynamics associated with the nonconvex formulation to that of a convex reparameterization. This allows us to provide convergence guarantees for the nonconvex approach from its convex counterpart.**

Index Terms— **Linear quadratic regulator, gradient descent, gradient-flow dynamics, model-free control, nonconvex optimization, Polyak-Lojasiewicz inequality.**

I. Introduction

The design of feedback controllers that provide desired performance of engineering systems has been an active area since the 1940's. There have been many developments aimed at broadening the range of applications, improving the speed and scalability of algorithms, and addressing important issues of uncertainty in modeling and data acquisition. In spite of these successes, a significant body of literature focuses on known dynamics and asymptotic analysis. In practice, the plant dynamics are often unknown and only a limited number of input-output measurements may be available. Such challenges have led to the adaptation of Reinforcement Learning (RL) approaches which can be broadly divided into model-based [1], [2] and model-free [3], [4]. While model-based RL relies on an approximation of the underlying dynamics, its model-free counterpart prescribes control action based on estimated values of a cost function without knowledge of the plant. In spite of the impressive empirical success of modern RL in a variety of domains, fundamental questions surrounding algorithmic convergence and sample complexity remain unanswered even for classical control problems, including the Linear Quadratic Regulator (LQR). This is mainly because of nonconvex nature of these algorithms.

The LQR problem is the cornerstone of control theory. The globally optimal solution can be obtained by solving the Riccati equation and efficient numerical schemes with provable convergence guarantees have been developed [5]. However, computing the optimal solution becomes challenging for large-scale problems, when prior knowledge is not available, or in the presence of structural constraints on the controller. This motivates the use of direct search methods for controller synthesis. Unfortunately, the nonconvex nature of this formulation complicates the analysis of first- and second-order optimization algorithms. To make matters worse, structural constraints on the feedback gain matrix may result in a disjoint search landscape limiting the utility of conventional descent-based methods [6].

In this paper we take a step towards providing model-free gradient-based strategies for solving the *continuous-time* LQR problem by directly searching over the parameter space of controllers. Despite the nonconvex nature of LQR formulation, we establish exponential convergence of the gradient-flow dynamics and linear convergence of the gradient descent method. A salient feature of our analysis is that we connect the gradient-flow dynamics of this nonconvex formulation to that of a standard convex reparametrization of the LQR problem [7], [8]. This connection allows us to provide a simple convergence analysis for the nonconvex setting by exploiting properties of its convex reparametrization.

For policy gradient methods applied to the *discrete-time* LQR problem, global convergence guarantees were recently provided for systems with known and unknown dynamics in [9]. While this reference motivates our work, we study the *continuous-time* LQR problem when the plant dynamics are known. In a companion paper (that is currently under preparation), we show how our results enable stronger guarantees in the model-free setting relative to [9].

The paper is structured as follows. In Section II, we revisit the LQR problem and present continuous- and discrete-time variants of the gradient descent algorithm. In Section III, we highlight the main result of the paper. In Section IV, we build on the convex characterization of the $\mathcal{H}_2$ optimal control problem and establish global convergence for gradient-flow dynamics and its discretized variant with small-enough stepsize. In Section V, we extend our analysis over the nonconvex landscape. In Section VI, we provide a computational experiment to illustrate our theoretical developments. Finally, we provide concluding thoughts in Section VII.

Notation: We use $\|\cdot\|_2$ to denote the maximum singular value of linear operators and matrices, $\|M\|_F = \mathrm{trace}\,(M^TM)$ to denote the Frobenius norm, and $\langle X, Y\rangle := \mathrm{trace}\,(X^TY)$ to denote the standard matricial inner product.

Financial support from the National Science Foundation under Award ECCS-1809833 and the Air Force Office of Scientific Research under Award FA9550-16-1-0009 is gratefully acknowledged.

H. Mohammadi, M. Soltanolkotabi, and M. R. Jovanović are with the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089. A. Zare is with the Department of Mechanical Engineering, University of Texas at Dallas, Richardson, TX 75080. E-mails: {hesamedm, soltanol, mihailo}@usc.edu, armin.zare@utdallas.edu.

The smallest eigenvalues of the symmetric matrix M is $\lambda_{\min}(M)$ and we use $\mathbb{E}$ to denote the expected value.

II. Problem formulation and gradient methods

Consider the linear time-invariant system

$$\dot{x} \;=\; A\,x \;+\; B\,u, \quad x(0) \;=\; x_0 \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the control input, and A, B are constant matrices of appropriate dimensions. The LQR problem associated with system (1) is given by

$$\underset{x,\,u}{\text{minimize}} \;\; \mathbb{E}_{x_0 \sim \mathcal{D}} \int_0^\infty (x^T(t) Q\, x(t) + u^T(t) R\, u(t))\, \mathrm{d}t \tag{2}$$

where Q and R are positive definite matrices and x_0 is a random initial condition with distribution $\mathcal{D}$. For a controllable pair (A, B), the solution to the LQR problem is given by

$$u \;=\; -K^\star x \;=\; -R^{-1}B^T P\, x$$

where P is the unique positive definite solution to the algebraic Riccati equation (ARE)

$$A^T P \;+\; P\,A \;+\; Q \;-\; P\,B\,R^{-1}B^T P \;=\; 0.$$

However, conventional approaches for solving ARE are not applicable in the model-free setting. Furthermore, imposing structural constraints (e.g., sparsity) on the feedback gain matrix comes with additional challenges that require developing customized optimization algorithms [10]–[12].

An alternative approach to solving ARE is to search for the optimal solution over the set of stabilizing feedback gains

$$\mathcal{S}_K \;:=\; \{K \in \mathbb{R}^{m\times n} \mid A \,-\, BK \text{ is Hurwitz}\} \tag{3}$$

which is known to be nonconvex [6]. Specifically, we can minimize the LQR cost with respect to the gain matrix K as

$$\underset{K}{\text{minimize}} \quad f(K) \tag{4}$$

where

$$f(K) \;:=\; \begin{cases} \operatorname{trace}\left((Q \,+\, K^T R\, K)X\right), & K \in \mathcal{S}_K \\ \infty, & \text{otherwise.} \end{cases}$$

Here, the matrix X is given by

$$X \;:=\; \int_0^\infty \mathrm{e}^{(A\,-\,BK)t}\, \Omega\, \mathrm{e}^{(A\,-\,BK)^T t}\, \mathrm{d}t \tag{5a}$$

and it can be obtained by solving the Lyapunov equation

$$(A \,-\, BK)X \;+\; X(A \,-\, BK)^T + \;\Omega \;=\; 0 \tag{5b}$$

where $\Omega := \mathbb{E}_{x_0 \sim \mathcal{D}}\, x_0 x_0^T$ is the covariance of the initial condition, which we assume to be positive definite. In (4), K is the optimization variable, and (A, B, Q, R, Ω) are the known problem parameters. We note that $K \in \mathcal{S}_K$ if and only if the solution X to Eq. (5b) is positive definite [13].

The gradient of the function $f(K)$ is given by [14]

$$\nabla f(K) \;=\; 2\,(R\,K - B^T P)\,X \tag{6}$$

where P is the unique positive definite solution to

$$(A \,-\, BK)^T P \;+\; P\,(A \,-\, BK) \;=\; -Q \,-\, K^T R\, K. \tag{7}$$

In this paper, we study the convergence properties of the gradient-flow dynamics associated with problem (4)

$$\dot{K} \;=\; -\nabla f(K), \quad K(0) \,\in\, \mathcal{S}_K. \tag{GF}$$

We also examine the convergence of a discretized version of (GF), namely the gradient descent method

$$K^{k+1} \;:=\; K^k \,-\, \alpha\, \nabla f(K^k), \quad K^0 \,\in\, \mathcal{S}_K \tag{GD}$$

where $\alpha > 0$ is the stepsize.

III. Main results

Our first result shows that (GF) converges exponentially to the LQR solution $K^\star$ for any $K(0) \in \mathcal{S}_K$ despite the nonconvex optimization landscape.

Theorem 1: For any initial stabilizing feedback gain $K(0) \in \mathcal{S}_K$, the solution $K(t)$ to (GF) satisfies

$$f(K(t)) \,-\, f(K^\star) \;\le\; (f(K(0)) \,-\, f(K^\star))\, \mathrm{e}^{-\rho\, t}$$

where the convergence rate ρ depends on $f(K(0))$ and the parameters of optimization problem (4).

The proof of Theorem 1 along with explicit expressions for convergence rate are provided in Section V-A. Moreover, we show that for a sufficiently small stepsize α the discrete analog (GD) also converges over $\mathcal{S}_K$ with a linear rate.

Theorem 2: For any initial stabilizing feedback gain $K^0 \in \mathcal{S}_K$, the iterates of gradient descent (GD) satisfy

$$f(K^k) \,-\, f(K^\star) \;\le\; \gamma^k \left(f(K^0) \,-\, f(K^\star)\right)$$

where the convergence rate γ and the stepsize α depend on $f(K^0)$ and the parameters of optimization problem (4).

We prove Theorem 2 and provide explicit expressions for γ and α in Section V-B.

IV. Convex reparameterization

Because of the nonconvexity of problem (4), it is unclear if gradient-based methods can be used to compute the LQR solution. Indeed, gradient descent may not even converge to local optima for nonconvex problems. Herein, we use a standard change of variables to reparameterize (4) into a convex problem, for which we can provide exponential/linear convergence guarantees for gradient flow/descent. In the next section, we connect the gradient flow on this convex reparameterization to its nonconvex counterpart; this allows us to prove global convergence for gradient flow/descent on (4).

A. Change of variables

The stability constraint on the closed-loop dynamics ($X \succ 0$) in problem (4) allows for a standard change of variables $Y := KX$ to reformulate the LQR synthesis as a convex optimization problem [7], [8]. In particular, for any $K \in \mathcal{S}_K$ and the corresponding X given by (5a), we have

$$f(K) \;=\; h(X, Y)$$

where $h(X, Y) := \operatorname{trace}(Q\,X + Y^T R\, Y X^{-1})$ is a jointly convex function of (X, Y) for $X \succ 0$. In the new set of

variables, the Lyapunov equation (5b) takes the affine form

$$\mathcal{A}(X) \,-\, \mathcal{B}(Y) + \Omega \;=\; 0 \tag{8}$$

where the linear maps $\mathcal{A}$ and $\mathcal{B}$ are given by

$$\mathcal{A}(X) \;:=\; A\,X \,+\, X A^T, \;\; \mathcal{B}(Y) \;:=\; B\,Y \,+\, Y^T B^T.$$

For an invertible map $\mathcal{A}$, we can express the matrix X as an affine function of Y

$$X(Y) \;=\; \mathcal{A}^{-1}(\mathcal{B}(Y) \,-\, \Omega)$$

and bring the LQR problem into the convex form

$$\underset{Y}{\text{minimize}} \quad h(Y).$$

Here,

$$h(Y) \;:=\; \begin{cases} h(X(Y), Y), & Y \in \mathcal{S}_Y \\ \infty, & \text{otherwise} \end{cases}$$

where $\mathcal{S}_Y$ is the set of matrices Y that correspond to stabilizing feedback gains $K = YX^{-1}$,

$$\mathcal{S}_Y \;:=\; \{Y \in \mathbb{R}^{m\times n} \mid \mathcal{A}^{-1}(\mathcal{B}(Y) \,-\, \Omega) \succ 0\}.$$

We note that similar to $\mathcal{S}_K$, the positive definite condition in $\mathcal{S}_Y$ is equivalent to $A - BYX^{-1}(Y)$ being Hurwitz. When the map $\mathcal{A}$ is not invertible, the change of variables $\hat{A} := A - BK^0$, $\hat{K} := K - K^0$, and $\hat{Y} := \hat{K}X$ can be alternatively used without loss of generality; details are omitted for brevity and will be reported elsewhere. Our convergence analysis for gradient descent relies on lower and upper bounds on the second-order approximation to the function $h(Y)$. We next quantify these bounds by showing that $h(Y)$ is strongly convex and smooth over its sub-level sets.

B. Strong convexity and smoothness of $h(Y)$

The gradient of $h(Y)$ is given by [12]

$$\nabla h(Y) \;=\; 2\,RYX^{-1} \,-\, 2\,B^T W \tag{9}$$

where W is the solution to the Lyapunov equation

$$A^T W \,+\, WA \;=\; -Q \,+\, X^{-1}Y^T RYX^{-1}. \tag{10}$$

While the gradient $\nabla h(Y)$ is not Lipschitz continuous over the set $\mathcal{S}_Y$, we show Lipschitz continuity over sublevel sets

$$\mathcal{S}_Y(a) \;:=\; \{Y \in \mathcal{S}_Y \mid h(Y) \;\le\; a\}$$

of the function $h(Y)$. We also show that over any sublevel set $\mathcal{S}_Y(a)$ the function $h(Y)$ is strongly convex. The next lemma is borrowed from [12, Lemma 3] and it provides an expression for the second-order approximation of $h(Y)$.

Lemma 1: The Hessian of the function $h(Y)$ satisfies

$$\left\langle \tilde{Y}, \nabla^2 h(Y;\tilde{Y}) \right\rangle \;=\; 2\,\|R^{\frac{1}{2}}(\tilde{Y} \,-\, K\tilde{X})X^{-\frac{1}{2}}\|_F^2$$

where $\tilde{X}$ is the unique solution to

$$\mathcal{A}(\tilde{X}) \;=\; \mathcal{B}(\tilde{Y}). \tag{11}$$

The following proposition provides expressions for Lipschitz continuity parameter L of $\nabla h(Y)$ and strong convexity module μ of $h(Y)$ over sublevel sets $\mathcal{S}_Y(a)$ in terms of a and parameters of the LQR problem. These are obtained by finding upper and lower bounds on the second-order approximation of $h(Y)$ from Lemma 1.

Proposition 1: Over any non-empty sublevel set $\mathcal{S}_Y(a)$, the gradient $\nabla h(Y)$ is Lipschitz continuous with parameter

$$L \;=\; \frac{2\,a\,\|R\|_2}{\nu}\left(1 \,+\, \frac{a\,\|\mathcal{A}^{-1}(\mathcal{B})\|_2}{\nu\sqrt{\lambda_{\min}(R)}}\right)^2 \tag{12}$$

and the function $h(Y)$ is μ-strongly convex with

$$\mu \;=\; \frac{2\,\lambda_{\min}(R)\,\lambda_{\min}(Q)}{a\,(1 \,+\, a^2\,\eta)^2} \tag{13}$$

where the constants

$$\eta \;:=\; \frac{\|\mathcal{B}\|_2}{\lambda_{\min}(Q)\,\lambda_{\min}(\Omega)\sqrt{\nu\,\lambda_{\min}(R)}} \tag{14a}$$

$$\nu \;:=\; \frac{\lambda_{\min}^2(\Omega)}{4}\left(\frac{\|A\|_2}{\sqrt{\lambda_{\min}(Q)}} \,+\, \frac{\|B\|_2}{\sqrt{\lambda_{\min}(R)}}\right)^{-2} \tag{14b}$$

only depend on the problem parameters.

Proof: See Appendix A. ∎

C. Exponential stability

The above results can be utilized to establish exponential stability of the gradient-flow dynamics

$$\dot{Y} \;=\; -\nabla h(Y), \quad Y(0) \,\in\, \mathcal{S}_Y \tag{15}$$

and the gradient descent method

$$Y^{k+1} \;:=\; Y^k \,-\, \alpha\,\nabla h(Y^k), \quad Y^0 \,\in\, \mathcal{S}_Y. \tag{16}$$

Proposition 2: The gradient flow dynamics (15) are exponentially stable, i.e.,

$$\|Y(t) \,-\, Y^\star\|_F^2 \;\le\; (L/\mu)\,\|Y(0) \,-\, Y^\star\|_F^2\,\mathrm{e}^{-2\,\mu\,t}$$

for $Y(0) \in \mathcal{S}_Y$, where μ and L are strong convexity and smoothness parameters over the sublevel set $\mathcal{S}_Y(h(Y(0)))$.

Proof: The time-derivative of the Lyapunov function candidate $V(Y) := h(Y) - h(Y^\star)$ satisfies

$$\frac{\dot{V}}{V} \;=\; \frac{-\|\nabla h(Y)\|^2}{h(Y) \,-\, h(Y^\star)} \;\le\; -2\,\mu \tag{17}$$

where $Y^\star$ is the global minimizer. Inequality (17) follows from the strong convexity of $h(Y)$ and it yields [15, Lemma 3.4]

$$V(Y(t)) \;\le\; V(Y(0))\,\mathrm{e}^{-2\,\mu\,t}. \tag{18}$$

Thus, for any $Y(0) \in \mathcal{S}_Y$, the objective function $h(Y(t))$ converges exponentially to $h(Y^\star)$. Moreover, since $h(Y)$ is μ-strongly convex and L-smooth, $V(Y)$ can be upper and lower bounded by quadratic functions. Based on this, the exponential stability of (15) follows from standard Lyapunov theory [15, Theorem 4.10]. ∎

Similarly, we can develop convergence guarantees for the gradient descent method (16) with sufficiently small stepsize.

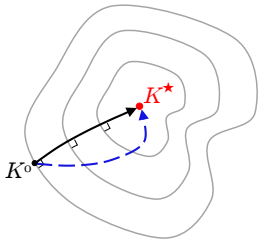

Fig. 1. Flow trajectories of (GF) (solid black) and $K_{\rm ind}$ from Eq. (19) (dashed blue) over sublevel sets $\mathcal{S}_K(a)$ of the function $f(K)$.

In particular, since the function $h(Y)$ is L-smooth over the sublevel set $\mathcal{S}_Y(h(Y^0))$, for any stepsize $\alpha \in [0, 1/L]$, the iterates Y^k remain within $\mathcal{S}_Y(h(Y^0))$. Based on this and the μ-strong convexity of $h(Y)$, we conclude that the iterates Y^k converge to the optimal solution $Y^\star$ at a linear rate $\gamma = 1 - \alpha\,\mu$. We next use this result to prove convergence for gradient flow/descent on the nonconvex formulation.

V. Analysis of the nonconvex formulation

The trajectories $Y(t)$ of (15) defined over the set $\mathcal{S}_Y$ induce the flow

$$K_{\rm ind}(t) \;:=\; Y(t)\,(X(Y(t)))^{-1} \tag{19}$$

over the set of stabilizing feedback gains $\mathcal{S}_K$. The result established in Proposition 2 implies that the objective function $f(K_{\rm ind}(t))$ converges with the exponential rate

$$\frac{f(K_{\rm ind}(t)) \,-\, f(K^\star)}{f(K_{\rm ind}(0)) \,-\, f(K^\star)} \;=\; \frac{h(Y(t)) \,-\, h(Y^\star)}{h(Y(0)) \,-\, h(Y^\star)} \;\leq\; \mathrm{e}^{-2\,\mu\,t}.$$

This inequality follows from (18) where μ denotes the strong-convexity module of the function $h(Y)$ over the sublevel set $\mathcal{S}_Y(h(Y(0)))$; see Proposition 1.

Figure 1 illustrates a trajectory of the induced flow $K_{\rm ind}(t)$ and a trajectory $K(t)$ of gradient-flow dynamics (GF) that start from the same initial condition. Although the stable flow $K_{\rm ind}(t)$ traverses a different curve on $\mathcal{S}_K$ than $K(t)$, we establish a relation between them which allows us to show that $K(t)$ also converges to the LQR solution $K^\star$.

A. Gradient flow dynamics: proof of Theorem 1

We start our analysis by relating the convex and nonconvex formulations of the LQR objective function. Specifically, in Lemma 2, we establish a relation between the gradients $\nabla f(K)$ and $\nabla h(Y)$ over the sublevel sets $\mathcal{S}_K(a)$.

Lemma 2: For any stabilizing feedback gain $K \in \mathcal{S}_K(a)$, X given by (5a), and $Y := KX$, we have

$$\|\nabla f(K)\|_F \;\geq\; c\,\|\nabla h(Y)\|_F \tag{20}$$

where the constant c is given by

$$c \;:=\; \frac{\nu\sqrt{\nu\,\lambda_{\min}(R)}}{2\,a^2\,\|\mathcal{A}^{-1}\|_2\,\|B\|_2 \,+\, a\,\sqrt{\nu\,\lambda_{\min}(R)}} \tag{21}$$

and the scalar ν (Eq. (14b)), depends on the problem data.

Proof: See Appendix C. ∎

We next consider the error in the objective value as a Lyapunov function candidate $V(K) := f(K) - f(K^\star)$. For any initial condition $K(0) \in \mathcal{S}_K(a)$, the time-derivative of $V(K)$ along the solutions of (GF) satisfies

$$\frac{\dot V}{V} \;=\; \frac{-\|\nabla f(K)\|_F^2}{f(K) \,-\, f(K^\star)} \;\leq\; \frac{-c^2\,\|\nabla h(Y)\|_F^2}{h(Y) \,-\, h(Y^\star)} \;\leq\; -2\,\mu\,c. \tag{22}$$

Here, the first inequality follows from $f(K) = h(Y)$ combined with (20) and the second follows from (17) (which in turn is a consequence of the strong convexity of $h(Y)$). Following [15, Lemma 3.4], inequality (22) guarantees that system (GF) converges exponentially in the objective value at rate $\rho = 2\,\mu\,c^2$. This completes the proof of Theorem 1.

Remark 1 (Geometric interpretation): For any trajectory $Y(t) \in \mathcal{S}_Y$ of (15), the LQR objective function satisfies

$$h(Y(t)) = f(K_{\rm ind}(t))$$

where $K_{\rm ind}(t) = Y(t)(X(Y(t)))^{-1}$ denotes the trajectory induced by $Y(t)$ over the set $\mathcal{S}_K$. Differentiating both sides of this equality with respect to time t yields

$$-\|\nabla h(Y)\|^2 = \left\langle \nabla f(K_{\rm ind}), \dot K_{\rm ind} \right\rangle. \tag{23}$$

Thus, Eq. (20) in Lemma 2 can be equivalently restated as

$$\|\nabla f(K_{\rm ind})\|_F^2 / \left\langle -\nabla f(K_{\rm ind}), \dot K_{\rm ind} \right\rangle \;\geq\; c^2.$$

In words, the ratio between $\|\nabla f(K_{\rm ind})\|_F$ and the norm of the projection of vector field $\dot K_{\rm ind}$ (associated with flow $K_{\rm ind}(t)$) on vector field $-\nabla f(K_{\rm ind})$ (associated with (GF)) is bounded from below. Due to this feature, we can deduce exponential convergence for the gradient-flow dynamics (GF) from the convergence properties of its convex counterpart.

Remark 2 (Gradient domination): Expression (22) implies that the objective function $f(K)$ over any given sublevel set $\mathcal{S}_K(a)$ satisfies

$$\|\nabla f(K)\|_F^2 \;\geq\; 2\,\mu\,c^2\,(f(K) \,-\, f(K^\star))$$

where the scalars μ and c are functions of a. This condition is known as the Polyak-Łojasiewicz (PL) inequality [16] and it has been recently used to show convergence for gradient-based methods in the case of *discrete-time* LQR [9].

B. Gradient descent dynamics: proof of Theorem 2

The main challenge in analyzing gradient descent (GD) compared to its continuous counterpart is to find a suitable stepsize α that guarantees convergence. Lemma 3 provides a Lipschitz parameter for the gradient $\nabla f(K)$, which is useful in finding such a stepsize. The proof of Lemma 3 relies on the bounds provided in Appendix B and follows a similar line of argument as in [12, Appendix D], but is omitted due to page limitation.

Lemma 3: Over any non-empty sublevel set $\mathcal{S}_K(a)$, the gradient $\nabla f(K)$ is Lipschitz continuous with the parameter

$L_f = L_{f1} + L_{f2}$ where $L_{f1} := a\|R\|_2/\lambda_{\min}(Q)$,

$$L_{f2} := \frac{4a^3}{\lambda_{\min}^2(Q)\lambda_{\min}(\Omega)}\left(\frac{\|B\|_2^2}{\lambda_{\min}(\Omega)} + \frac{\|B\|_2\|R\|_2}{\sqrt{\nu\lambda_{\min}(R)}}\right),$$

and the constant ν (Eq. (14b)) depends on problem data.

For any line segment in $\mathcal{S}_K(a)$ with endpoints K and $K+\alpha\tilde{K}$, the L_f-smoothness of the function $f(K)$ implies

$$f(K+\alpha\,\tilde{K}) \;-\; f(K) \;\le\; \alpha\left\langle \nabla f(K), \tilde{K}\right\rangle + \frac{\alpha^2 L_f}{2}\,\|\tilde{K}\|_F^2. \tag{24}$$

Let $\tilde{K} \in \mathbb{R}^{m\times n}$ be a decent direction of the function $f(K)$ for some $K \in \mathcal{S}_K(a)$, i.e., $f(K+\alpha\hat{K}) - f(K) < 0$ for small enough $\alpha > 0$. If the right-hand side of (24) is negative for all $\alpha \in (0, b]$ (for some scalar b), then inequality (24) follows from the continuity of $f(K)$. The negative gradient update in (GD) is clearly a descent direction of the function $f(K)$. Now, let L_f be the Lipschitz parameter of $\nabla f(K)$ over the sublevel set $\mathcal{S}_K(f(K^0))$. It is easy to verify that the right-hand side of (24) with $K := K^k$ and $\tilde{K} := -\nabla f(K^k)$ is negative for all $\alpha \in (0, 1/L_f]$. Therefore, from (24) it follows that the iterates of gradient descent (GD) satisfy

$$f(K^{k+1}) \;-\; f(K^k) \;\le\; -\frac{2\,\alpha - L_f\,\alpha^2}{2}\,\|\nabla f(K^k)\|_F^2.$$

This inequality in conjunctions with the PL condition

$$\|\nabla f(K^k)\|_F^2 \;\ge\; 2\,\mu\,c^2\,\big(f(K^k) \;-\; f(K^\star)\big)$$

established in (22) guarantees convergence for gradient descent (GD) with the linear rate $\gamma \le 1 - \alpha\,\mu\,c^2$ for all $\alpha \in (0, 1/L_f]$. This concludes the proof of Theorem 2.

VI. An example

We use a mass-spring-damper system with N masses to compare the performance of gradient descent on K given by (GD) and gradient descent on Y given by (16). We set all spring and damping constants as well as masses to unity. In state-space representation (1), the state vector $x = [\,p^T\ v^T\,]^T$ contains the position and velocity of masses and the dynamic and input matrices are given by

$$A \;=\; \begin{bmatrix} 0 & I \\ -T & -I \end{bmatrix}, \quad B \;=\; \begin{bmatrix} 0 \\ I \end{bmatrix}$$

where O and I are zero and identity matrices of suitable size, and T is a Toeplitz matrix with 2 on the main diagonal and -1 on the first super and sub-diagonals.

We solve the LQR problem with $Q = I + 100\,\mathrm{e}_1\mathrm{e}_1^T$, $R = I + 100\,\mathrm{e}_4\mathrm{e}_4^T$, and $\Omega = I$ for $N = 10$ and 50 masses (i.e., $n = 2N$ states) where e_i is the ith unit vector in the standard bases of $\mathbb{R}^n$. The algorithms were initialized with $Y^0 = K^0 = 0$. Figure 2 illustrates the convergence curves of both algorithms with a stepsize selected using a backtracking procedure that guarantees stability of the feedback loop. We observe that the asymptotic rates of convergence for gradient descent on $\mathcal{S}_K$ and $\mathcal{S}_Y$ demonstrate similar trends.

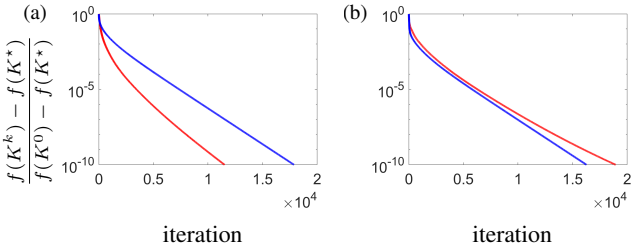

Fig. 2. Convergence curves for gradient descent (blue) over the set $\mathcal{S}_K$, and gradient descent (red) over the set $\mathcal{S}_Y$. (a) and (b) correspond to $N = 10$ and $N = 50$ masses, respectively.

VII. Concluding remarks

We prove exponential/linear convergence of gradient flow/descent algorithms for solving the continuous-time LQR problem based on a nonconvex formulation that directly searches for the controller. A salient feature of our analysis is that we relate the gradient-flow dynamics associated with this nonconvex formulation to that of a convex reparameterization. This allows us to deduce convergence of the nonconvex approach from its convex reparameterization. While in this paper we focus on known dynamics, in a companion paper we extend our results to the model-free setting with unknown A and B. Our efforts serve as a first step towards providing a general sample-based framework for the learning and control of large-scale dynamical systems. Some future directions include: (i) developing data-driven synthesis with convergence guarantees that involves finite-time stochastic approximation of the objective and its gradient; and (ii) providing theoretical guarantees for the convergence of gradient-based methods for structured control synthesis.

Appendix

A. Proof of Proposition 1

Here we only show that the function $h(Y)$ is μ-strongly convex. See [12, Appendix D] for a proof of smoothness. The proof relies on the bounds provided in Appendix B and Lemma 4 that provides an upper bound on the norm of the inverse Lyapunov operator for stable systems. The proof of Lemma 4 is omitted due to page limitations.

Lemma 4: For any Hurwitz matrix $F \in \mathbb{R}^{n\times n}$, the linear map $\mathcal{F}: S^n \to S^n$

$$\mathcal{F}(W) \;:=\; \int_0^\infty \mathrm{e}^{F\,t}\,W\,\mathrm{e}^{F^T t}\,\mathrm{d}t \tag{25}$$

is well defined and for any $\Omega \succ 0$,

$$\|\mathcal{F}\|_2 \;\le\; \mathrm{trace}\,(\mathcal{F}(I)) \;\le\; \mathrm{trace}\,(\mathcal{F}(\Omega))/\lambda_{\min}(\Omega). \tag{26}$$

We show that for any $\tilde{Y} \in \mathbb{R}^{m\times n}$ and $Y \in \mathcal{S}_Y(a)$, the Hessian of $h(Y)$ satisfies $\left\langle \tilde{Y}, \nabla^2 h(Y;\tilde{Y})\right\rangle \;\ge\; \mu\,\|\tilde{Y}\|_F^2$ where μ is given by (13). Using Lemma 1, we can write

$$\left\langle \tilde{Y}, \nabla^2 h(Y;\tilde{Y})\right\rangle = 2\,\|R^{\frac{1}{2}}\,H\,X^{-\frac{1}{2}}\|_F^2 \;\ge\; \frac{2\lambda_{\min}(R)}{\|X\|_2}\,\|H\|_F^2 \tag{27}$$

where $H := \tilde{Y} - K\,\tilde{X}$. Next, we show that

$$\|H\|_F/\|\tilde{X}\|_F \;\geq\; \lambda_{\min}(\Omega)/\mathrm{trace}(X)\,\|\mathcal{B}\|_2. \tag{28}$$

To do so, we substitute $H + K\,\tilde{X}$ for $\tilde{Y}$ in (11), which yields

$$\Gamma \;=\; B\,H \,+\, H^T B^T, \tag{29}$$

where $\Gamma := (A - B\,K)\tilde{X} + \tilde{X}(A - B\,K)^T$. Equation (29) allows us to lower bound the norm of H as

$$\|H\|_F \;\geq\; \|\Gamma\|_F/\|\mathcal{B}\|_2. \tag{30}$$

From the stability of the closed loop system, we have

$$\tilde{X} \;=\; -\int_0^\infty \mathrm{e}^{(A-BK)t}\,\Gamma\,\mathrm{e}^{(A-BK)^T t}\,\mathrm{d}t.$$

Now, we use Lemma 4 with $F := A - BK$ to lower bound the norm of Γ as follows

$$\|\Gamma\|_F \;\geq\; \frac{\|\tilde{X}\|_F}{\|\mathcal{F}\|_2} \;\geq\; \frac{\lambda_{\min}(\Omega)\|\tilde{X}\|_F}{\mathrm{trace}(\mathcal{F}(\Omega))} \;=\; \frac{\lambda_{\min}(\Omega)\,\|\tilde{X}\|_F}{\mathrm{trace}(X)} \tag{31}$$

where the linear map $\mathcal{F}$ is defined in (25). Inequality (28) follows from combining (30) and (31).

An upper bound on $\|\tilde{Y}\|_F$ can thus be established as

$$\begin{aligned}\|\tilde{Y}\|_F \;&=\; \|H \,+\, K\,\tilde{X}\|_F \;\leq\; \|H\|_F \,+\, \|K\|_F\,\|\tilde{X}\|_F \\ &\leq\; \|H\|_F\left(1 \,+\, \frac{a\,\mathrm{trace}(X)\,\|\mathcal{B}\|_2}{\lambda_{\min}(\Omega)\sqrt{\nu\,\lambda_{\min}(R)}}\right) \\ &\leq\; \|H\|_F\left(1 \,+\, a^2\,\eta\right)\end{aligned} \tag{32}$$

where η is given by (14a). Here, the second inequality follows from (34c) and (28) and the last inequality follows from (34a). Finally, inequalities (27) and (32) yield

$$\begin{aligned}\frac{\left\langle \tilde{Y}, \nabla^2 f(Y;\tilde{Y})\right\rangle}{\|\tilde{Y}\|_F^2} \;&\geq\; \frac{2\,\lambda_{\min}(R)\,\|H\|_F^2}{\|X\|_2\,\|\tilde{Y}\|_F^2} \\ &\geq\; \frac{2\,\lambda_{\min}(R)}{\|X\|_2(1+\eta)^2} \;\geq\; \frac{2\,\lambda_{\min}(R)\lambda_{\min}(Q)}{a\,(1 \,+\, a^2\,\eta)^2} \;=\; \mu\end{aligned} \tag{33}$$

where the last inequality follows from (34a). This completes the proof.

B. Bounds on optimization variables

The following bounds on the variables X and K hold [12], [14]. Over a sublevel set $\mathcal{S}_K(a)$, we have

$$\mathrm{trace}(X) \;\leq\; \frac{a}{\lambda_{\min}(Q)} \tag{34a}$$

$$\frac{\nu}{a} \;\leq\; \lambda_{\min}(X) \tag{34b}$$

$$\|K\|_F \;\leq\; \frac{a}{\sqrt{\nu\,\lambda_{\min}(R)}} \tag{34c}$$

where the constant ν is given by (14b).

C. Proof of Lemma 2

The gradients can be written as $\nabla f(K) = EX$ and $\nabla h(Y) = E + 2\,(B^T(P - W))$, where $E := 2(R\,K - B^T P)$, and the matrices P and W are given by Eqs. (7) and (10), respectively. Subtracting (10) from (7) yields

$$A^T(P - W) + (P - W)\,A \;=\; -\frac{1}{2}\left(K^T E \,+\, E^T K\right).$$

This equation gives us

$$\|P - W\|_F \;\leq\; \|\mathcal{A}^{-1}\|_2\,\|K\|_F\,\|E\|_F \;\leq\; \frac{a\|\mathcal{A}^{-1}\|_2\,\|E\|_F}{\sqrt{\nu\,\lambda_{\min}(R)}}$$

where the second inequality follows from (34c) in Appendix B. We thus have

$$\|\nabla h(Y)\|_F/\|E\|_F \;\leq\; 1 \,+\, 2\,a\|\mathcal{A}^{-1}\|_2\,\|B\|_2/\sqrt{\nu\,\lambda_{\min}(R)}. \tag{35}$$

On the other hand, using the lower bound (34b) on $\lambda_{\min}(X)$, it follows that

$$\|\nabla f(K)\|_F \;=\; \|E\,X\|_F \;\geq\; \frac{\nu}{a}\,\|E\|_F.$$

Combining this inequality and (35) completes the proof.

References

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

[2] M. Simchowitz, H. Mania, S. Tu, M. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," 2018, arXiv:1802.08334.

[3] D. Bertsekas, "Approximate policy iteration: A survey and some new methods," *J. Control Theory Appl.*, vol. 9, no. 3, pp. 310–335, 2011.

[4] Y. Abbasi-Yadkori, N. Lazic, and C. Szepesvári, "Model-free linear quadratic control via reduction to expert prediction," 2018, arXiv:1804.06021.

[5] B. Anderson and J. Moore, *Optimal Control; Linear Quadratic Methods*. New York, NY: Prentice Hall, 1990.

[6] J. Ackermann, "Parameter space design of robust control systems," *IEEE Trans. Automat. Control*, vol. 25, no. 6, pp. 1058–1072, 1980.

[7] E. Feron, V. Balakrishnan, S. Boyd, and L. El Ghaoui, "Numerical methods for H_2 related problems," in *Proceedings of the 1992 American Control Conference*, 1992, pp. 2921–2922.

[8] B. Polyak, M. Khlebnikov, and P. Shcherbakov, "An LMI approach to structured sparse feedback design in linear control systems," in *Proceedings of the 2013 European Control Conference*, 2013, pp. 833–838.

[9] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," 2018, arXiv:1801.05039v2.

[10] F. Lin, M. Fardad, and M. R. Jovanović, "Augmented Lagrangian approach to design of structured optimal state feedback gains," *IEEE Trans. Automat. Control*, vol. 56, no. 12, pp. 2923–2929, 2011.

[11] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Trans. Automat. Control*, vol. 58, no. 9, pp. 2426–2431, 2013.

[12] A. Zare, H. Mohammadi, N. K. Dhingra, M. R. Jovanović, and T. T. Georgiou, "Proximal algorithms for large-scale statistical modeling and optimal sensor/actuator selection," 2018, arXiv:1807.01739.

[13] H. Kwakernaak and R. Sivan, *Linear optimal control systems*. Wiley-interscience New York, 1972, vol. 1.

[14] H. T. Toivonen, "A globally convergent algorithm for the optimal constant output feedback problem," *Int. J. Control*, vol. 41, no. 6, pp. 1589–1599, 1985.

[15] H. K. Khalil, *Nonlinear Systems*. New York: Prentice Hall, 1996.

[16] H. Karimi, J. Nutini, and M. Schmidt, "Linear convergence of gradient and proximal-gradient methods under the Polyak-łojasiewicz condition," in *In European Conference on Machine Learning*, 2016, pp. 795–811.