Automatica 182 (2025) 112504

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Convergence analysis of gradient flow for overparameterized LQR formulations☆

Arthur Castello B. de Oliveira [a,*], Milad Siami [a], Eduardo D. Sontag [a,b]

[a] Department of Electrical and Computer Engineering at Northeastern University, Boston, MA 02115, USA
[b] Department of Bioengineering and affiliated with the departments of Chemical Engineering and Mathematics, at Northeastern University, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 10 May 2024
Received in revised form 27 May 2025
Accepted 23 June 2025

Keywords:
Input-to-state stability
Learning theory
Singularities in optimization
Optimal control theory
Application of nonlinear analysis and design
Stability of nonlinear systems

ABSTRACT

This paper analyzes the intersection between results from gradient methods for the model-free linear quadratic regulator (LQR) problem, and linear feedforward neural networks (LFFNNs). More specifically, it looks into the case where one wants to find an LFFNN feedback that minimizes an LQR cost. It starts by deriving a key conservation law of the system, which is then leveraged to generalize existing results on boundedness and global convergence of solutions, and invariance of the set of stabilizing LFFNNs under the training dynamics (gradient flow). For the single hidden layer LFFNN, the paper proves that the solution converges to the optimal feedback control law for all but a set of Lebesgue measure zero of the initializations. These results are followed by an analysis of a simple version of the problem – the "vector case" – proving the theoretical properties of accelerated convergence and a type of input-to-state stability (ISS) result for this simpler example. Finally, the paper presents numerical evidence of faster convergence of the gradient flow of general LFFNNs when compared to non-overparameterized formulations, showing that the acceleration of the solution is observable even when the gradient is not explicitly computed, but estimated from evaluations of the cost function.

1. Introduction

Neural networks and machine learning (ML) tools are being increasingly used in control design (Alali & Imani, 2023; Motamedi, Behzad, Zandi, Salehinejad, & Siami, 2024; Ravari, Ghoreishi, & Imani, 2022, 2024; Sznaier, Olshevsky, & Sontag, 2022; Wafi & Siami, 2023; Zandi, Salehinejad, Behzad, Motamedi, & Siami, 2023), and are particularly useful in model-free applications, where a model of the system might not be available (Cui, Jiang, & Sontag, 2024; Fazel, Ge, Kakade, & Mesbahi, 2018). In such scenarios, an "oracle" might be queried to estimate the cost associated with a specific control law, as illustrated in Fig. 1. This feedback has adjustable parameters (or "weights"), which are updated through the gradient of the estimated cost, typically employing gradient descent or some other similar numerical optimization method.

Understanding the convergence of such learning techniques is challenging due to inherent nonlinearities. In particular, neural networks leverage both their compositional structure and the nonlinear activation functions of each layer. Previous works on neural networks isolate the compositional structure from the nonlinear activation by studying linear feedforward neural networks (LFFNNs) (Bah, Rauhut, Terstiege, & Westdickenberg, 2022; Chitour, Liao, & Couillet, 2023; de Oliveira, Siami, & Sontag, 2023, 2024b; Eftekhari, 2020; Kawaguchi, 2016; Min, Tarmoun, Vidal, & Mallada, 2021; Min, Vidal, & Mallada, 2023). The results are typically given for solving a static supervised learning problem, *i.e.* a linear regression of labels u on $K_N \dots K_1 y$, where y is an input. Not only powerful "almost everywhere" convergence results have been obtained for the regression problem (Arora, Du, Hu, Li, & Wang, 2019; Bah et al., 2022; Baldi & Hornik, 1989; Chitour et al., 2023; Eftekhari, 2020; Kawaguchi, 2016), and a type of input-to-state stability (ISS) property of an associated problem was characterized (de Oliveira et al., 2023), but, perhaps surprisingly, the optimization on the individual matrices K_i can result in much faster convergence than optimization on a single matrix K (Min et al., 2021, 2023; Tarmoun, Franca, Haeffele, & Vidal, 2021).

Despite the rich literature, current results on LFFNNs cannot be applied out-of-the-box to non-convex problems, even if under some gradient dominance condition (PL-inequality). An extremely popular and well-studied example of such a system

☆ This work was partially supported by ONR, United States Grant N00014-21-1-2431, National Science Foundation, United States Grant 2121121, and AFOSR, United States Grant FA9550-21-1-0289. The material in this paper was partially presented at The 63rd IEEE Conference on Decision and Control (CDC), December 16–19, 2024, Milan, Italy. This paper was recommended for publication in revised form by Associate Editor Marcello Farina under the direction of Editor Florian Dorfler.

* Corresponding author.
E-mail addresses: a.castello@northeastern.edu (A. Castello B. de Oliveira), m.siami@northeastern.edu (M. Siami), e.sontag@northeastern.edu (E.D. Sontag).

https://doi.org/10.1016/j.automatica.2025.112504

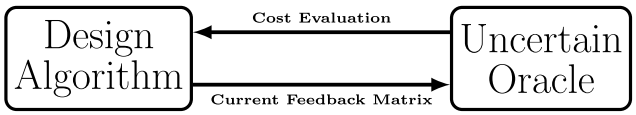

Fig. 1. System overview of the model-free control design process. The design algorithm attempts to find a feedback matrix that minimizes the output of the oracle, which in turn provides a (possibly noisy) estimate of the cost function every time it receives a candidate feedback matrix.

is the linear quadratic regulator (LQR) problem, whose general goal is to minimize a quadratic cost function $J(K)$, where K is a candidate feedback law. When the system dynamics are linear and known, an explicit optimal solution is obtainable by solving a Riccati equation on the system matrices (Sontag, 2013), but such an approach is generally unfit for model-free scenarios, where the system is assumed unknown and only the value of the cost function for different feedback matrices can be queried to some oracle (as illustrated in Fig. 1). This type of scenario can be understood as a "policy optimization" formulation for the LQR problem (poLQR) (Hu et al., 2023), and approximates the LQR problem to reinforcement learning problems, motivating previous works where the optimization is solved by following the negative flow of the gradient $\dot{K} = -\nabla J(K)$, or negative descent direction $K_{n+1} = K_n - h\nabla J(K_n)$ (for some step-size $h > 0$) (Cui et al., 2024; Fazel et al., 2018; Hu et al., 2023; Levine & Athans, 1970; Mohammadi, Zare, Soltanolkotabi, & Jovanovic, 2022; Sontag, 2022). Such "training" is an area of active research to this day due to its non-convex landscape, and traces its origins to pioneering work by Levine and Athans starting in the late 1960s (Levine & Athans, 1970). Recent publications have established global convergence properties (Fazel et al., 2018; Hu et al., 2023; Mohammadi et al., 2022), as well as input-to-state stability (ISS) (Cui et al., 2024; Sontag, 2022) when the computation of the gradient is subject to error or uncertainty. Of special note, in Hu et al. (2023), the authors explore the relationship between LQR (and other classical control problems) and policy optimization, leveraging the explicit expression of the gradient of the linear quadratic cost to prove the convergence of gradient descent methods. Similarly, in Zhang and Başar (2023), the authors also interpret a gradient approach to this problem as policy optimization for the LQR problem, and explore the relationship between the finite and infinite horizon formulations of the LQR problem to propose a strategy that converges even if initialized outside the set of stabilizing controllers. All these results argue for the importance of understanding the behavior of the gradient flow when studying the LQR problem in a model-free context.

In this context, the primary goal of this paper is to study the effects of LFFNNs when applied to the more complex setting of solving a model-free LQR problem (poLQR). Mathematically, the feedback is written as a product $K = K_N \dots K_1$, where K_i represents the weights of the ith layer of the network. In this context, the natural training dynamics take the form $\dot{K}_i = -\nabla_{K_i} J(K_N \dots K_1)$ for $i = 1, \dots, N$, which is a coupled set of gradient flows done on the full set of parameters $(K_1, \dots, K_N)$. Notably, the assumptions of gradient dominance (PL inequality) and coerciveness of the cost function – important for both general non-convex optimization (Agarwal, Kakade, Lee, & Mahajan, 2021; Jin, Ge, Netrapalli, Kakade, & Jordan, 2017; Nesterov & Polyak, 2006; Polyak, 1963) and for gradient methods for solving the poLQR (Fazel et al., 2018; Gravell, Esfahani, & Summers, 2020; Mohammadi, Soltanolkotabi, & Jovanović, 2021) – do not hold when optimizing over layers of an LFFNN. This is due to the introduction of spurious equilibria and multiple non-compact sets of critical points in the gradient dynamics. Despite those issues, we derive convergence properties of the solution of an overparameterized formulation for the poLQR.

Beyond the original goal, literature results on accelerated convergence (Min et al., 2021, 2023; Tarmoun et al., 2021) indicate that even the simpler problem of linear activation functions can be interesting and useful from more than just a theoretical point of view. We demonstrate that this property also holds for the poLQR through numerical simulations, although further discussions on computational and sample complexity are required before it can be determined whether this formulation is inherently useful for practical applications.

In sum, this paper takes steps to blend these two strands of research: gradient methods for the model-free LQR problem; and the analysis of overparameterization in optimization. It looks at the use of overparameterized state feedback for the poLQR, investigating properties that can be derived for its gradient flow.

To accomplish this, the paper starts at Section 2 by presenting a theoretical background of both gradient methods for the LQR problem and overparameterization for linear regression problems. Then, in Section 3 the paper formally defines the overparameterized policy optimization LQR problem and proves that it shares the same convergence properties as the overparameterized linear regression. Then, Section 4 provides a complete characterization of a simplified version of the problem: the single input and single state/output. The center-stable manifold of the spurious equilibria is characterized, and both an ISS-type and accelerated convergence properties are formally proven. The paper then presents numerical simulations to demonstrate the presence of accelerated convergence for the general case in Section 5. The simulations show how initialization affects convergence when compared to the non-overparameterized gradient flow, both when the gradient is perfectly and imperfectly known. Finally, in Section 6 the contributions of this paper are summarized and possible future directions of work are discussed. A preliminary version of this work was previously published (de Oliveira et al., 2024b), but the proofs appear here for the first time, and the discussion is significantly deepened. All proofs of the main results are provided in the appendix for the clarity of the main text, and proofs of minor results can be found in the arXiv version of this paper (de Oliveira, Siami, & Sontag, 2024a).

2. Theoretical background

Throughout this paper, let $\mathbb{R}_+$ and $\mathbb{R}_{++}$ be the set of nonnegative and strictly positive real numbers respectively. For $n \in \mathbb{N}$, let $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ be the set of symmetric positive semi-definite (PSD) and positive definite (PD) n-by-n matrices, respectively. Given a matrix $A \in \mathbb{R}^{n\times n}$, A is said to be Hurwitz if all its eigenvalues have negative real part.

2.1. The LQR problem as policy optimization

We begin by presenting results from (Rautert & Sachs, 1997), which serve as groundwork upon which we derive our new results. We also emphasize that despite the reliance of the following results on the knowledge of the system matrices, the gradient expression derived in this section holds great value for analysis, as demonstrated, for example, in Fazel et al. (2018), Hu et al. (2023), Mohammadi et al. (2022), where it forms the basis for theoretical guarantees regarding convergence rate and accuracy in model-free scenarios.

Consider the following linear system:

$$\Sigma \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{n\times n}$ are the system matrices, with (A, B) assumed controllable and C assumed full rank (Rautert & Sachs, 1997, Assumption 2). Let $\mathcal{K} := \{K \in \mathbb{R}^{m\times n} \mid$

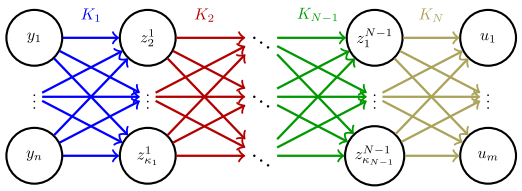

Fig. 2. Graphical representation of a linear feedforward neural network (LFFNN) with an input layer $y \in \mathbb{R}^n$ with n neurons, hidden layers $z^i \in \mathbb{R}^{\kappa_i}$ with κ_i neurons, and output layer $u \in \mathbb{R}^m$ with m neurons. The computation of the network is done for each layer as $z^i = K_i z_{i-1}$, with $z^0 = y$ and $z^N = u$, where the matrices K_i represent, in the figure, the presence and weight of edges between neurons of layer $i-1$ and layer i. The resulting input–output expression for the LFFNN then becomes $u = K_N \dots K_1 y$.

$A+BKC$ is Hurwitz}, then the objective is to determine an output feedback $u = Ky$, $K \in \mathcal{K}$, that minimizes

$$J(K) = \mathbb{E}_{x_0 \sim \mathcal{X}_0}\left[\int_0^\infty x(t)^\top Q x(t) + u(t)^\top R u(t)\, dt\right], \tag{2}$$

with given positive definite cost matrices $R \in \mathbb{S}_{++}^{m\times m}$ and $Q \in \mathbb{S}_{++}^{n\times n}$, and for x_0 sampled from a probability distribution $\mathcal{X}_0$. Throughout this paper we refer to K^* as the unique solution to the LQR problem.

In Rautert and Sachs (1997, Theorem 3.2), the authors provide the following expression for the gradient ∇J with respect to the feedback matrix K:

$$\nabla J(K) = 2(B^\top P_K + RKC)L_K C^\top, \tag{3}$$

where for any $K \in \mathcal{K}$, P_K and L_K are the unique positive definite solutions of the following Lyapunov equations

$$P_K(A+BKC) + (A+BKC)^\top P_K + C^\top K^\top R K C + Q = 0 \tag{4}$$

$$L_K(A+BKC)^\top + (A+BKC)L_K + \Sigma_0 = 0, \tag{5}$$

respectively, and the matrix $\Sigma_0 = \mathbb{E}_{x_0\sim\mathcal{X}_0}[x_0 x_0^\top]$ depends on the distribution of initial conditions $\mathcal{X}_0$, and is assumed to be of full rank. From these, we can define the set of desired/optimal values of K as $\mathcal{T} := \{K \in \mathcal{K} \mid \nabla J(K) = 0\}$. With these results established, we next look at key literature results on overparameterization.

2.2. Overparameterization - properties and formulation

The optimization landscape of the gradient flow of a linear neural networks is usually studied in terms of least square/linear regression problems, stated as follows: let $Y = [y_1, y_2, \dots, y_k]$ and $U = [u_1, u_2, \dots, u_k]$ be the column concatenation of (possibly noisy) k input–output pairs sampled from an unknown function $\mathbf{K}$ that one wants to approximate using a linear neural network $\mathbf{K}$. Although arguably a simple formulation, the resulting gradient system is the object of study of many papers in the literature (Bah et al., 2022; Chitour et al., 2023; de Oliveira et al., 2023, 2024b; Eftekhari, 2020; Kawaguchi, 2016; Min et al., 2021, 2023).

For some search space of neural networks $\mathcal{K}$, defined as appropriate to the problem, an optimal neural network $\mathbf{K}^* \in \mathcal{K}$ minimizes $J(\mathbf{K}) = \|U - \mathbf{K}(Y)\|$, where $\mathbf{K}(Y) = [\mathbf{K}(y_1), \dots, \mathbf{K}(y_k)]$, and for some norm $\|\cdot\|$. A linear feedforward neural network (LFFNN) (depicted in Fig. 2) is a feedforward neural network with linear activation functions between layers, and has: an input layer with n neurons; $N-1$ hidden layers, each with $\kappa_i \geq \max(m, n)$ neurons, for $i = 1, \dots, N-1$; and an output layer with m neurons. Then, in the specific case of a LFFNN, and being $K_i \in \mathbb{R}^{\kappa_i \times \kappa_{i-1}}$ the ith layer parameter matrix, the function to be minimized becomes $J(K_1, \dots, K_N) = \|U - K_N \dots K_1 Y\|$.

For this problem, and under some reasonable assumptions on the ranks of Y and U, and on the dimensions of the K_is (see Kawaguchi (2016), and Assumptions 1 and 2 in Chitour et al. (2023) and references therein, or a previous work from the authors (de Oliveira et al., 2023)), the following can be summarized from the literature about the optimization landscape of this problem:

Proposition 1. *Consider a linear regression problem solved with a LFFNN with N layers and trained through gradient flow. Assume U and Y are full column rank and that all hidden layers are wider than the number of inputs and outputs (i.e. all hidden layers have more neurons than the input and output layers), then:*

(1) the problem is generally non-convex and non-concave;
(2) all local minima are global minima;
(3) there are no local maxima;
(4) in the special case where $N = 2$, all critical points are either global minima or strict saddles (i.e. the Hessian at that point has at least one strictly negative eigenvalue);
(5) the solution exists for any initial condition and always converges to a critical point of the dynamics;
(6) if $N = 2$, the solutions converge to a global optimum for all initializations but a set of Lebesgue measure zero.

Proof. Items (1) to (4) are studied in Baldi and Hornik (1989) for the single hidden layer case, and Kawaguchi (2016) generalized these results to the arbitrarily deep case. Properties (5) and (6) are proved in Panageas and Piliouras (2017) for the analogous discretized problem (*i.e.* gradient descent). In Appendix A we will show how to adapt these proofs to the continuous-time (*i.e.* flow) case. An independent proof of (5) and (6) was provided in Chitour et al. (2023) for the specific problem of linear regression and under an additional assumption on the loss function ("distinct critical values").

Furthermore, other works in the literature establish useful properties of overparameterized linear neural networks when compared to equivalent non-overparameterized formulations. In Min et al. (2021), Tarmoun et al. (2021) the authors study the speed of convergence of the gradient flow in overparameterized linear neural networks solving linear regressions, showing that depending on the initialization of the algorithm, the convergence rate can be arbitrarily increased. In Min et al. (2023) the authors extend their results to a more general class of optimization problems, although the required assumption of convexity of the non-overparameterized problem makes it so that their results are not immediately applicable to the LQR problem.

In our previous work (de Oliveira et al., 2023), we provide some insights on the loss of robustness in training overparameterized linear neural networks through gradient flow, and show how judicious restrictions on the set of initializations might circumvent this problem.

Such properties for linear neural networks/ overparameterized linear regressions could be useful if they held in the context of feedback control design. Motivated by these results, the next section looks at how one can extend these important results for the policy optimization LQR problem, and consequently to feedback control design.

3. Feedback control through LFFNNs

Let $\mathbf{K} = (K_1, K_2, \dots, K_n)$ be a LFFNN with $N-1$ hidden layers, an input layer, and an output layer. Let $K_1, K_2, \dots, K_N$ be the

weight matrices of each layer with $K_1 \in \mathbb{R}^{\kappa_1 \times n}$, $K_2 \in \mathbb{R}^{\kappa_2 \times \kappa_1}$ and so forth, with $K_N \in \mathbb{R}^{m \times \kappa_{N-1}}$, where $\kappa_i \in \mathbb{Z}^+$ is the dimension of the ith hidden layer. Furthermore, we are interested in the overparameterized case, *i.e.* $\kappa_i \geq \max(m, n)$ for all $i = 1, \ldots, N-1$. For an input $y \in \mathbb{R}^n$ of the LFFNN, its output $u \in \mathbb{R}^m$ is given by $u = \mathbf{K}(y) = K_N K_{N-1} \cdots K_2 K_1 y$, and its structure is as depicted in Fig. 2. By choosing $\mathbf{K}$ as the output feedback law, the closed-loop dynamics of the LTI system (1) becomes $\dot{x} = Ax + B\mathbf{K}(Cx) = (A + BK_N \cdots K_1 C)x$, and the LQR problem cost becomes

$$J(\mathbf{K}) = \text{trace}\,(P_{\mathbf{K}} \Sigma_0)\,, \tag{6}$$

where for a given $\mathbf{K}$, $P_{\mathbf{K}}$ is the unique solution of the following Lyapunov equation:

$$\begin{aligned} &P_{\mathbf{K}}(A + BK_N \ldots K_1 C) + (A + BK_N \ldots K_1 C)^\top P_{\mathbf{K}} \\ &+ (K_N \ldots K_1 C)^\top R K_N \ldots K_1 C + Q = 0. \end{aligned} \tag{7}$$

The notation $J(\mathbf{K})$ and $J(K_1, K_2, \ldots, K_N)$ are used interchangeably when the goal is to emphasize the dependency on the linear neural network $\mathbf{K}$ or on its parameters $(K_1, \ldots, K_N)$. With this, consider the following problem definition.

Definition 1. Let $\mathbf{K}$ be a LFFNN, and A, B, and C be as in (1). Define $\mathcal{K} := \{\mathbf{K} \mid (A + BK_N \ldots K_1 C)$ is Hurwitz$\}$ and let $R \in \mathbb{S}^{m \times m}_{++}$ and $Q \in \mathbb{S}^{n \times n}_{++}$ be given symmetric positive definite matrices. Solving an overparameterized formulation of the model-free LQR problem consists in finding a $\mathbf{K}^* \in \mathcal{K}$ that solves

$$\begin{aligned} \min_{\mathbf{K} \in \mathcal{K}} \quad & J(\mathbf{K}) := \text{trace}\,(P_{\mathbf{K}} \Sigma_0) \\ \text{s.t.} \quad & (7). \end{aligned}$$

Then, a gradient flow for the overparameterized model-free LQR problem is defined for each $i = 1, \ldots, N$ and any fixed "learning rate" $\eta > 0$ by imposing the following dynamics for the parameter matrices K_i that compose $\mathbf{K}_0$

$$\dot{K}_i = -\eta \frac{\partial J}{\partial K_i}, \tag{8}$$

and a candidate solution to the overparameterized model-free LQR problem is obtained by initializing the gradient flow at some $\mathbf{K}_0 \in \mathcal{K}$ and selecting whichever point the solution converges to (assuming it converges to a point). It is evident that an equilibrium of the gradient flow dynamics (8) is not necessarily the global optimum of the overparameterized poLQR, and a better understanding of the landscape of the problem is required before one can discuss the optimality of a solution obtained in such a manner. Nonetheless, $\dot{K}_i = 0$ for all $i = 1, \ldots, N$ is a necessary condition for global optimality, which makes the equilibria of (8) natural candidates for a optimal solution. Henceforth in this paper, it is assumed $\eta = 1$, although comparisons between the proposed formulation and other formulations that explore variable values for η could prove to be an interesting future direction of work.

Regarding the computation of the gradients of J with respect to the matrices K_i, consider the following result:

Lemma 1. *Let $B_i := BK_N \ldots K_{i+1}$ and $R_i := K_{i+1}^\top \ldots K_N^\top R K_N \ldots K_{i+1}$ for $i \in \{1, \ldots, N-1\}$, $C_i := K_{i-1} \ldots K_1 C$ for $i \in \{2, \ldots, N\}$, $B_N := B$, $C_1 := C$, and $R_N := R$. Then*

$$\nabla_{K_i} J = 2[B_i^\top P_{\mathbf{K}} + R_i K_i C_i] L_{\mathbf{K}} C_i^\top, \tag{9}$$

where $P_{\mathbf{K}}$ is the solution of (7), $L_{\mathbf{K}}$ is the solution of

$$\begin{aligned} &L_{\mathbf{K}}[A + BK_N \ldots K_1 C]^\top \\ &+ [A + BK_N \ldots K_1 C] L_{\mathbf{K}} + \Sigma_0 = 0, \end{aligned} \tag{10}$$

and Σ_0 relates to the distribution of initial conditions, being equal to the covariance matrix if the initialization is random Gaussian with zero mean, or equal to the identity for uniformly sampled unitary vectors.

Notice that we presented the results so far for arbitrary full-rank C to keep the comparison with the results from (Rautert & Sachs, 1997), however moving forward we will assume full state feedback for the system, that is $C = I$, and initializations in the unit sphere, that is $\Sigma_0 = I$. We next look at what can be said regarding convergence guarantees for the proposed problem.

3.1. A conservation law for the overparameterized model-free LQR problem

Notice that, relative to the weight matrix of each hidden layer, the derivative of the cost J relative to each parameter matrix, given by (9) follows an iterative structure that allows the characterization of a conservation law that is satisfied by any solution. Such conservation law follows a very similar structure as the ones characterized for overparameterized linear regression (see for example Lemma 2.3 of Chitour et al. (2023)). This property is given in the following lemma:

Lemma 2. *For a gradient flow dynamics (8) used for solving the overparameterized model-free LQR problem (poLQR) presented in Definition 1, and for any i from 1 to $N-1$, the following quantity is invariant along any solution $(K_1(t), \ldots, K_N(t))$ initialized in $\mathcal{K}$:*

$$\begin{aligned} \mathcal{C}_i :=& K_i K_i^\top - K_{i+1}^\top K_{i+1} \\ =& (K_i K_i^\top - K_{i+1}^\top K_{i+1})_{t=0}, \end{aligned} \tag{11}$$

where $\mathcal{C}_i$ are constant matrices of appropriate dimensions. We refer to the set $(\mathcal{C}_1, \ldots, \mathcal{C}_{N-1})$ as the set of invariants of a given solution.

A similar conservation law is leveraged to prove many of the properties of the overparameterized gradient flow for linear regressions, as can be seen from Lemma 2.3 in Chitour et al. (2023), Lemma 1 in Min et al. (2021), Lemma 2.1 of Bah et al. (2022), and others. The fact that such property also holds for the more general Linear Quadratic cost when overparameterized motivates the search presented in this paper for other useful properties that might hold for this case.

With this, and knowing that the LQR cost function is a rational function (see, for example, a discussion in Sontag (2022), section 4.3) the following result regarding the global convergence of solutions of (8) can be stated:

Theorem 1. *Any solution of the gradient flow (8) initialized in $\mathcal{K}$ (defined as in Definition 1): exists; is precompact; remains in $\mathcal{K}$ for all time; and converges to a critical point of the gradient flow dynamics.*

This result not only guarantees invariance of the set of stabilizing neural networks and global convergence of solutions but also demonstrates how the invariance obtained in Lemma 2 can be used to extend results from the literature on overparameterized linear regressions to the context of the overparameterized model-free LQR problem. We next look at the case with $N = 2$, *i.e.* a single hidden layer, to enunciate an even stronger convergence result.

3.2. Feedback control design with a single hidden layer

Consider now the case where $N = 2$ (single hidden layer). The literature on overparameterized linear regression is rich in results for this case, and this section aims to show that the main ones also hold for the design of optimal state feedback controllers.

We begin by proving that any critical point that is not a global minimum of the problem is necessarily a strict saddle. This result is, then, used to prove almost everywhere convergence to the global minimum of the problem. Then we characterize all critical points to discuss some intuition behind the problematic set of initializations, that is the set of initializations that do not converge to the global minimum. Let $\mathcal{T}$ be defined as in Section 2, then consider the following result:

Theorem 2. *Let (K_1, K_2) be an equilibrium point of the gradient dynamics (8), then either*

- *The point (K_1, K_2) is a global minimum of the system, i.e. $K_2K_1 \in \mathcal{T}$; or*
- *The point (K_1, K_2) is a strict saddle of the dynamics, i.e. the Hessian evaluated at (K_1, K_2) has at least one negative eigenvalue.*

Because Theorem 2 guarantees that the critical points are either strict saddles or global minima, and Theorem 1 guarantees convergence to a critical point, we can apply Corollary 4 provided in Appendix A to get the following Corollary:

Corollary 1. *For all initializations but a set of Lebesgue measure zero, the solution of the overparameterized gradient flow (8) converges to a point (K_1, K_2) such that $K_2K_1 \in \mathcal{T}$, that is, almost all solutions initialized in $\mathcal{K}$ converge to an optimal feedback matrix and minimize (6).*

Notice that Corollary 1 is proven without needing to characterize the set of initializations that converge to a saddle. Such points are hard to characterize for an arbitrary saddle, although we can provide a characterization of the critical points themselves as follows:

Lemma 3. *For the gradient flow (8) with $N = 2$ and $\kappa_1 = \kappa > \max(m, n)$, and for any set of parameter matrices (K_1, K_2) such that $K_2K_1 \in \mathcal{K}$, the following are equivalent:*

(1) The point (K_1, K_2) is an equilibrium of (8), i.e. $\dot{K}_1 = K_2^\top 2[B^\top P_{\mathbf{K}} + RK_2K_1]L_{\mathbf{K}} = 0$, and $\dot{K}_2 = 2[B^\top P_{\mathbf{K}} + RK_2K_1]L_{\mathbf{K}}K_1^\top = 0$.

(2) Let $\nabla_{\mathbf{K}}J := 2[B^\top P_{\mathbf{K}} + RK_2K_1]L_{\mathbf{K}}$, then, there exist an SVD $\nabla_{\mathbf{K}}J(K_2K_1) = \Psi\Sigma\Phi^\top$, and orthogonal matrices Γ_{K_1}, $\Gamma_{K_2} \in \mathbb{R}^{\kappa\times\kappa}$ such that: (a) $K_1 = \Gamma_{K_1}\Sigma_{K_1}\Phi^\top$ and $K_2 = \Psi\Sigma_{K_2}\Gamma_{K_2}^\top$ are SVDs of K_1 and K_2; and (b) $\Sigma\Sigma_{K_1}^\top = 0$ and $\Sigma_{K_2}^\top\Sigma = 0$.

From this Lemma, we can characterize the product K_2K_1 at critical points in terms of low-rank approximations of K^* (the optimal LQR feedback matrix) as in the following corollary

Corollary 2. *Let K^* be the optimal value of $K \in \mathcal{K}$ that minimizes the LQR cost (2). If (K_1, K_2) is a critical point of the gradient flow dynamics (8) with a $N = 2$, then there exists an SVD of K^**

$$K^* = \begin{bmatrix}\Psi_1^*, \Psi_2^*\end{bmatrix}\begin{bmatrix}\Sigma_1^* & 0\\ 0 & \Sigma_2^*\end{bmatrix}\begin{bmatrix}(\Phi_1^*)^\top\\ (\Phi_2^*)^\top\end{bmatrix},$$

with its singular values not necessarily in any order, such that

$$K_2K_1 = \begin{bmatrix}\Psi_1^*, \Psi_2^*\end{bmatrix}\begin{bmatrix}\Sigma_1^* & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}(\Phi_1^*)^\top\\ (\Phi_2^*)^\top\end{bmatrix}.$$

As an immediate consequence of Lemma 3 and the consequent Corollary 2, one can notice that there is a finite number of values that the cost function (6) can have at any critical point of the gradient flow dynamics. This characterizes a finite number of sets of critical points, as a function of the number of possible low-rank factorization of K^*. Despite that characterization, however, it is still hard to compute the center-stable manifold of the saddles, as we hope to illustrate next.

For some $p < \min(m, n)$, let K_p^* denote a rank-p factorization of K^* and let the set of all (K_1, K_2) such that $K_2K_1 = K_p^*$ be given by $\mathcal{T}_p := \{(K_1, K_2) \mid K_2K_1 = K_p^*\}$. It is evident that for any $(K_1, K_2) \in \mathcal{T}_p$, $(K_1\mu, K_2(1/\mu)) \in \mathcal{T}_p$ as well for any $\mu \neq 0$, and therefore $\mathcal{T}_p$ is continuous and unbounded. However, it is also easy to see that there exist two $(\bar{K}_1, \bar{K}_2) \in \mathcal{T}_p$ and $(\tilde{K}_1, \tilde{K}_2) \in \mathcal{T}_p$ for which there exist no $\mu \neq 0$ such that $(\bar{K}_1, \bar{K}_2) = (\mu\tilde{K}_1, (1/\mu)\tilde{K}_2)$.

The degrees of freedom for points in $\mathcal{T}_p$ come from the fact that multiple different values of Σ_{K_1}, Σ_{K_2}, Γ_{K_1} and Γ_{K_2} exist such that $K_2K_1 = K_p^*$. However, necessary and sufficient conditions on these matrices for the equality to hold do not exist to the authors' knowledge, which makes an analytic characterization of all points in a given set $\mathcal{T}_p$ difficult. This difficulty also explains why characterizing the center-stable manifold of the saddles is hard. Assume that for a given $(K_1, K_2) \in \mathcal{T}_p$ the center-stable manifold of that point is known, then to extend it to a "neighborhood" of the point in $\mathcal{T}_p$, one would need to be able to: first characterize all points arbitrarily close to (K_1, K_2); and second derive how that characterization reflects in the characterization of the center-stable manifold of a point in $\mathcal{T}_p$.

In this section, we have collected powerful results about the convergence of the gradient flow solution for the general problem and the single hidden layer case. These results provide some guarantee to the behavior of the solution but also illustrate some of the fundamental challenges of understanding deep and wide optimization formulations. We will follow up in the next section with a complete analysis of a simpler version of the problem, in the hopes of illustrating better some of the intuition derived from the results from this section.

4. Analysis of the single-input/single-state case with one hidden-layer

To provide a better intuition behind the results given in the previous section, we now study a simple example of the considered problem. Assume $N = 2$, $n = m = 1$, but $\kappa_1 =: \kappa$ arbitrary. The case where the parameters take these values is referred to as "the vector case", and if $\kappa = 1$ then it is referred to as "the scalar case".

For the vector case, the system in consideration is of the form of (1) with $A, B \in \mathbb{R}$ and $x, u : \mathbb{R}_+ \to \mathbb{R}$. Without loss of generality, assume $x(0) = 1$, $B = 1$, and denote $A = a$ to emphasize its scalar nature. Furthermore, assume the scalar weights for the cost (6) are given by $Q = q > 0$ and $R = r > 0$, and the parameters to be optimized by $K_1 = k_1 \in \mathbb{R}^{\kappa\times 1}$ and $K_2 = k_2 \in \mathbb{R}^{1\times\kappa}$. Furthermore, the valid parameter space is defined as $\mathcal{K} := \{(k_1, k_2) \in \mathbb{R}^{\kappa\times 1}\times\mathbb{R}^{1\times\kappa} \mid a + k_2k_1 < 0\}$. Assuming a feedback of the form $u = k_2k_1x$, with $(k_1, k_2) \in \mathcal{K}$ results in

$$\begin{aligned} J(k_1, k_2) &= \mathbb{E}_{x_0\in\mathcal{X}_0}\left[\int_0^\infty x(t)^2q + u(t)^2r dt\right]\\ &= \mathbb{E}_{x_0\in\mathcal{X}_0}\left[\int_0^\infty x(t)^2(q + (k_2k_1)^2r)dt\right]\\ &= \mathbb{E}_{x_0\in\mathcal{X}_0}\left[x(0)^2\right](q + (k_2k_1)^2r) \qquad (12)\\ &\quad\times\int_0^\infty \mathrm{e}^{2(a+k_2k_1)t}dt\\ &= -\frac{(q + (k_2k_1)^2r)}{2(a + k_2k_1)}. \qquad (13)\end{aligned}$$

Taking the gradient with respect to k_1 and k_2 gives

$$\nabla_{k_1}J(k_1, k_2) = f(k_1, k_2)k_2^\top \qquad (14)$$

$$\nabla_{k_2}J(k_1, k_2) = f(k_1, k_2)k_1^\top, \qquad (15)$$

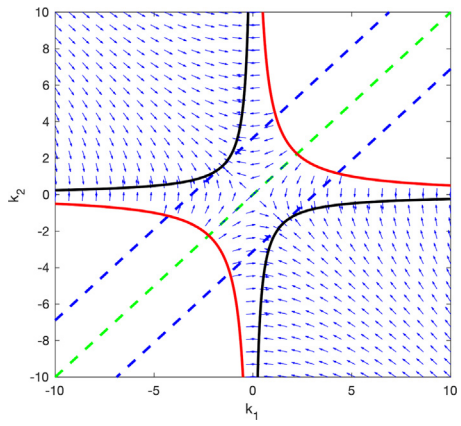

Fig. 3. Phase Plane for the gradient flow dynamics for the scalar case described in Section 4, drawn for a stable A. The blue arrows depict the vector field at different points of the state space. The black hyperbolas are the new equilibria introduced by the condition $f(k_1, k_2) = 0$, with $f(\cdot)$ as in (14) and (15). The red hyperbolas are the borders of the set of (k_1, k_2) such that $a + k_2k_1 < 0$, that is, such that the closed loop is stable. The blue dashed lines are composed of the points that satisfy $d(k_1, k_2) = 2\sqrt{|k^*_-|}$, while the green dashed line is the set for which $d(k_1, k_2) = 0$ where $d(k_1, k_2)$ is as defined in Proposition 2.

where

$$f(k_1, k_2) = -\frac{r(k_2k_1)^2 + 2ark_2k_1 - q}{2(a + k_2k_1)^2},$$

which, in turn, results in the following dynamics for the parameters

$$\dot{k}_1 = -f(k_1, k_2)k_2^\top \tag{16}$$

$$\dot{k}_2 = -f(k_1, k_2)k_1^\top. \tag{17}$$

Notice that, similar to the observation made in de Oliveira et al. (2023) for the vector case in linear regression, the vector dynamics of this problem is a simple nonlinear reparameterization of a linear dynamics. This means that inside $\mathcal{K}$, the phase plane should be that of a saddle with an inversion in the direction of the flow whenever $f < 0$, and an extra equilibrium set given by $\{(k_1, k_2) \in \mathcal{K} \mid f(k_1, k_2) = 0\}$. This can be observed graphically for the scalar case in the plot given by Fig. 3.

The new equilibrium set given by $f(k_1, k_2) = 0$ can be studied explicitly, this condition is satisfied for any $(k_1, k_2) \in \mathcal{K}$ that solves $r(k_2k_1)^2 + 2ark_2k_1 - q = 0$. The solutions to this quadratic equation are

$$k_2k_1 = -a + \sqrt{a^2 + q/r} =: k^*_+ \tag{18}$$

$$k_2k_1 = -a - \sqrt{a^2 + q/r} =: k^*_-, \tag{19}$$

with $k^*_+ \notin \mathcal{K}$ leaving k^*_- as the only viable solution, which coincides with the optimal solution of the LQR problem for the scalar system, since from the theory on this problem one can write

$$k^*_{\text{LQR}} = -R^{-1}B^\top P,$$

where P is the solution of

$$A^\top P + PA - PBR^{-1}B^\top P + Q = 0,$$

which results in $K^*_{\text{LQR}} = k^*_-$ since $P > 0$.

Furthermore, notice that $f(k_1, k_2) > 0$ for all $(k_1, k_2) \in \mathcal{K}$ such that $k_2k_1 > k^*_-$, since the positive root $k^*_+ = -a + \sqrt{a^2 + q/r} > 0$ is such that $a + k^*_+ > 0$, and the concavity of the parabola is negative. Also notice that if $k_2k_1 < k^*$ then $f(k_1, k_2) < 0$ by a similar argument. However, notice that there is another equilibrium to this dynamics, given by $(k_1, k_2) = (0, 0)$. For this equilibrium, $k_2k_1 = 0$ which is not the optimal solution of the LQR problem. Such equilibrium is referred to as a spurious equilibrium of the system and is only in $\mathcal{K}$ if $a < 0$. Still, it is convenient to characterize a condition for which convergence to a global minimum is guaranteed. To do so, we adapt a result from (de Oliveira et al., 2023) to the design of feedback controllers:

Proposition 2. *For the overparameterized poLQR given by Definition 1 with $n = m = 1$ and $N = 2$ (i.e. the vector case), the gradient flow solution converges to the global optimal value of the cost function (2) if and only if the gradient flow is initialized such that*

$$d(k_1, k_2) := \|k_1 - k_2^\top\|_2^2 > 0.$$

Proposition 2 gives a necessary and sufficient condition for the convergence of a solution to the target set $\mathcal{T} := \{(k_1, k_2) \in \mathcal{K} \mid k_2k_1 = k^*\}$. For any point in $\mathcal{T}$, the value of the cost function $J(k_1, k_2)$ is the same, but that does not mean that all initializations that converge to $\mathcal{T}$ are equivalent. It was shown in Lemma 2 that different values for the conservation law are invariant along trajectories, so we show next how the values of this conservation law influence the convergence through the following definition and proposition.

Definition 2. For the overparameterized poLQR given by Definition 1 with $n = m = 1$ and $N = 2$ (*i.e.* the vector case), denote by $\mathcal{C} := \mathcal{C}_1 = k_1k_1^\top - k_2^\top k_2$, that is, the value of the invariant. Then, define the level of imbalance of a given solution as $c := 2\,\text{trace}\left(\mathcal{C}^2\right) - \text{trace}\,(\mathcal{C})^2$.

Proposition 3. *For the overparameterized poLQR given by Definition 1 with $n = m = 1$ and $N = 2$ (i.e. the vector case), let $(\phi_{k_1}(t, (k_1, k_2)), \phi_{k_2}(t, (k_1, k_2)))$ be the solution to the gradient flow (8) initialized at (k_1, k_2) and let $\phi_J(t, (k_1, k_2)) = J(\phi_{k_1}(t, (k_1, k_2)), \phi_{k_2}(t, (k_1, k_2)))$ be the trajectory of the cost function (6) along a solution. For two distinct initializations $(\tilde{k}_1, \tilde{k}_2)$ and $(\bar{k}_1, \bar{k}_2)$ with levels of imbalance given by $\tilde{c}$ and $\bar{c}$ respectively, let*

- $J(\tilde{k}_1, \tilde{k}_2) = J(\bar{k}_1, \bar{k}_2)$; *and*
- $|\tilde{c}| > |\bar{c}| \geq 0$, *with* $\bar{k}_1 \neq \bar{k}_2^\top$.

Then, for all time $t > 0$ it follows that $\phi_J(t, (\tilde{k}_1, \tilde{k}_2)) < \phi_J(t, (\bar{k}_1, \bar{k}_2))$. In other words, the cost converges faster to the minimum value for solutions initialized with a larger level of imbalance.

Proposition 3 proves an increase in the rate of convergence for different solutions of the system, however, it provides no quantitative result, *i.e.* it does not prove that the acceleration is unbounded. To further study advantages and trade-offs between different initializations, we next characterize the robustness of the solutions, *i.e.* how the solutions can be expected to behave when the gradient is computed with an associated level of additive uncertainty.

Some intuition regarding the behavior of the solution under disturbance can be obtained from analyzing the scalar case. One can notice graphically from Fig. 3 that as c increases, the associated equilibrium gets closer to the border of the set of stabilizing controllers, *i.e.* the red and black hyperbolas in the figure "meet at infinity". At first sight, this can be a problematic observation when considering disturbances, as points in the target set can be arbitrarily close to the border of instability. However, this does not mean that any disturbance during the training can take the feedback matrix to instability. In fact, let $\delta\mathcal{K}$ be the border of $\mathcal{K}$ (*i.e.* the red hyperbolas), and notice from (14) and (15) that in general, as $(k_1, k_2) \to \delta\mathcal{K}$, $|f(k_1, k_2)| \to \infty$, with its direction being away from the border. This means that only a

disturbance of infinite magnitude on the training dynamics could take a solution initialized in $\mathcal{K}$ away from it.

To formalize this intuition, we prove the following "ISS-type" result regarding solutions of the overparameterized poLQR in the vector case when subject to additive uncertainties.

Proposition 4. *For the overparameterized poLQR given by Definition 1 with $n = m = 1$ and $N = 2$ (i.e. the vector case), consider solutions initialized in $\mathcal{K}$ and such that $\|k_1 - k_2^\top\|_2\big|_{t=0} > 2\sqrt{a_+}$, where $a_+ = \max(0, a)$. Furthermore, let the dynamics be disturbed in the following form*

$$\dot{k}_{1,2} = -\nabla_{k_{1,2}} J + u_{1,2}, \tag{20}$$

where $u_1, u_2^\top : \mathbb{R}^+ \to \mathbb{R}^\kappa$. Then for every $\epsilon > 0$, there exists a $\delta > 0$ such that if $\|u_1\|_\infty + \|u_2^\top\|_\infty \leq \delta$ then $\limsup_{t\to\infty} J(k_2(t)k_1(t)) - J(k_-^*) \leq \epsilon$, where $\|\cdot\|_\infty$ is the infinity norm of a function.

Notice that the property characterized in Proposition 4 is not input-to-state stability as it is usually defined, and is more akin to a "input-to-cost" stability. Furthermore, due to the non-compactness of the sets of critical points, one can even prove that for an arbitrarily small disturbance, the state will diverge, but will do so along a trajectory that will keep the value of the cost bounded. Nonetheless, in some sense this still guarantees that the solution remains "close" in the sense of the cost $J(\cdot)$ to the target set, even when subject to additive disturbances.

Through this simple example, one can see how interesting and rich the problem discussed in this paper can be, as well as capture some of its intuition in a simpler context. The next section investigates numerically whether the increased speed of convergence, proven for the vector case here, might still hold for the general problem.

5. Numerical results

In this section, we investigate empirical distinctions between overparameterized and regular model-free LQR problems. The simulations were done using Matlab, and all code is available online in a repository (de Oliveira, 2025). The selected A and B for the simulations are

$$A = -\begin{bmatrix} 5.2373 & 0.3452 & 0.6653 & 0.6715 & 0.3288 \\ 0.3452 & 5.4889 & 0.8060 & 0.3889 & 0.5584 \\ 0.6653 & 0.8060 & 5.0377 & 0.5735 & 0.5100 \\ 0.6715 & 0.3889 & 0.5735 & 5.3354 & 0.6667 \\ 0.3288 & 0.5584 & 0.5100 & 0.6667 & 5.4942 \end{bmatrix} \tag{21}$$

$$B^\top = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{22}$$

All simulations are done for a 10-neuron single hidden layer neural network, since a single hidden layer is enough to observe overparameterization and has better convergence guarantees. The choice of 10 hidden neurons was arbitrary.

For the gradient flow solution to be well-defined, a stabilizing initialization is required. Although this is a common necessary condition (Fazel et al., 2018; Mohammadi et al., 2022), as mentioned in the introduction, recent works in the literature (Zhang & Başar, 2023) explore the finite horizon formulation for the LQR problem to allow for arbitrary initializations. However, while studying the effects of overparameterization on such formulations could prove interesting, it is not in the scope of this paper.

Therefore, to generate the synthetic results that illustrate the distinct behaviors of an overparameterized formulation over the non-overparameterized formulation for the poLQR, we must first discuss the difference in the behavior of the solution based on the initialization.

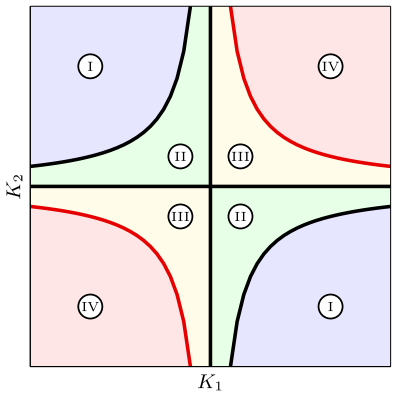

Fig. 4. Depiction of the four different regions of the state-space based on the expected behavior of the solution.

5.1. On the choice of initialization

Consider the phase plane of the scalar case depicted in Fig. 3 for reference. The first clear segmentation of the state space is the one done by the red hyperbolas, *i.e.* between the values of (K_1, K_2) such that $A + BK$ is Hurwitz or not.

Another similar segmentation is done by the black hyperbolas in the same figure. Notice that any solution initialized in between the two hyperbolas will never cross either hyperbola, and vice-versa. This happens because at the black hyperbolas both gradients $\nabla_{K_1} J(K_1, K_2) = 0$ and $\nabla_{K_2} J(K_1, K_2) = 0$, and by continuity of the solution, it cannot cross over.

Finally, the quadrants also separate the state space in two, where any initialization in the second and fourth quadrants always converges to the global optimum (black hyperbola), while initializations in the first and third quadrants can converge to the saddle at the origin.

From this informal analysis, one can draw Fig. 4, which can be expected to describe the behavior of the solution to some degree, even if not extensively. We will perform the simulations for initializations $(K_1(0), K_2(0)) = (K_{10}, K_{20})$ such that $K_{20}K_{10} = \eta K^*$ where K^* is the optimal feedback matrix and η is a scalar. Intuitively, one would expect that if $\eta > 1$, then the system would be initialized in a region of the state-space analogous to Ⓘ in Fig. 4, where η can be arbitrarily large and the solution still converges to the target set. Similarly, if $1 > \eta > 0$, the solution is in a region analogous to Ⓘ, and the closer η is to 0, the longer the initialization should take to converge to the target set. Finally, if $\eta < 0$ then it is in a region analogous to Ⓘ, or if $|\eta|$ is too large, then the solution does not exist.

To be more specific, for any given desired η we compute K^* first, then compute a SVD for it as $K^* = \Psi\Sigma\Phi$ and a random orthogonal 10×10 matrix Γ. Then, we define $K_{10} = \mathbf{sign}(\eta)\sqrt{|\eta|}\mu\Gamma\Sigma^{1/2}\Phi^\top$ and $K_{20} = (\sqrt{|\eta|}/\mu)\Psi\Sigma^{1/2}\Gamma^\top$ for μ varying from 1 to 100 defining more or less imbalanced initializations for the same η.

As mentioned before, this does not encompass all possible behaviors for the solution of the general case with a single hidden layer. To illustrate this fact, we will perform simulations for all three cases described above ($\eta > 1$, $1 > \eta > 0$ and $\eta < 0$) and a final simulation for an initialization selected specifically to not lie in any of the regions described by the different values of η.

After an overview of the behavior of the solution is provided, we will investigate how overparameterization affects the convergence in a scenario where the gradient is numerically estimated from evaluations of the cost function, resulting in imprecise approximations and introducing uncertainty to the dynamics.

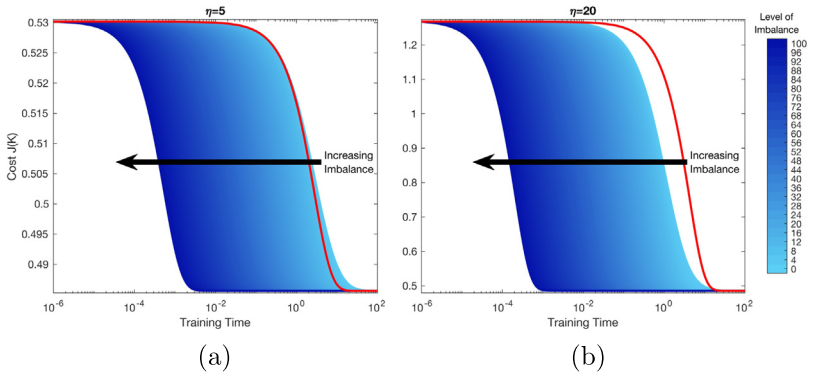

Fig. 5. Simulations done for initializations with $\eta > 1$. Solutions were initialized with $\eta = 5$ in (a) and with $\eta = 20$ in (b). Solutions from light to dark blue depict overparameterized solutions with different levels of imbalance $\mu \in [1, 100]$, and the red curve shows the non-overparameterized solution.

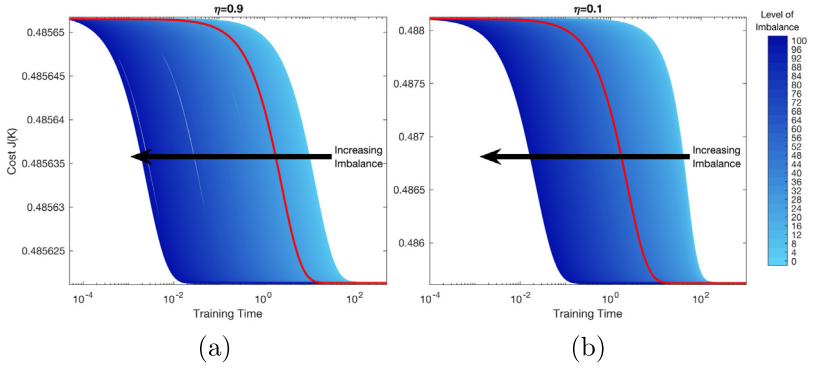

Fig. 6. Simulations done for initializations with $1 > \eta > 0$. Solutions were initialized with $\eta = 0.9$ in (a) and $\eta = 0.1$ in (b). Solutions from light to dark blue depict overparameterized solutions with different levels of imbalance $\mu \in [1, 100]$, and the red curve shows the non-overparameterized solution.

5.2. Results with the exact gradient

In this section, we will discuss the numerical simulation results for the gradient flow for the case when the gradient is perfectly known. The simulations are done for two different initializations with $\eta > 1$ (Fig. 5), two with $1 > \eta > 0$ (Fig. 6), and two with $\eta < 0$ (Fig. 7). Finally, two other simulations are done with a different initialization to illustrate a distinct behavior of the solution (Fig. 9).

For the simulations with $\eta > 1$ in Fig. 6, notice how the overparameterized solutions (shades of blue) converge relative to the non-overparameterized solution (red) depending on how far from the optimum the solution is initialized. For $\eta = 5$, the slowest of the overparameterized solutions (lightest blue) converges almost as quickly as the non-overparameterized solution but is overtaken as the solutions get closer to the optimum. Nonetheless, with an arguably small value for the imbalance term μ, it is verifiable that an overparameterized solution will converge more quickly than the non-overparameterized one. This becomes even more evident for the solutions initialized with $\eta = 20$, where all overparameterized solutions converge to the optimum more rapidly than the non-overparameterized solution. Furthermore, notice how the solution to the overparameterized gradient flow has a different profile than the non-overparameterized solution, indicating that the overparameterized formulation did not simply accelerate the convergence, but changed the behavior of the solution.

For the simulations with $1 > \eta > 0$ in Fig. 6, the variation in the values of the cost function is limited by the values of $J(K)$ for $K = K^*$ and $K = 0$. Furthermore, notice that the solutions initialized with $\eta = 0.9$ converge generally faster than the ones initialized with $\eta = 0.1$. This happens because as $\eta \to 0$, the initialization approaches the saddle, exhibiting reduced convergence rates. Despite that, however, a big enough imbalance can always be imposed to generate a solution that converges more rapidly than the non-overparameterized solution.

Next, for the simulations with $\eta < 0$ depicted in Fig. 7, before we can discuss the simulation results we first need to argue that theoretically for any initialization with $\eta < 0$ and in $\mathcal{K}$, if $\mu = 1$ then the resulting solution should converge to the saddle-point at the origin. To show this, first notice that for $\mu = 1$, $\mathcal{C} = 0$ by construction of the initialization. Then, notice that if $\eta < 0$, then $J(K_{20}K_{10}) > J(0)$. This can be shown theoretically, but for the simplicity of this analysis, this was verified numerically for this specific example. Next, since $J(K_{20}K_{10}) > J(0) > J(K^*)$, by continuity any solution initialized at (K_{10}, K_{20}) must pass through a point such that $K_{20}K_{10} = 0$ before it can reach the target set $\mathcal{T}$. Finally, notice that the only point such that $\mathcal{C} = 0$ and that $K_2K_1 = 0$ is the origin, which is a saddle of the dynamics.

This explains the strange behavior of the solutions initialized with $\mu = 1$ and $\eta = -0.1$ in Fig. 7, where the solution looks like it is converging to a suboptimal value for the cost function. However, despite theoretically converging to the saddle at 0, the simulation solution eventually escapes it due to accumulated errors in the numerical simulation, and reaches the global minimum. When looking at solutions initialized at $\eta = -20$ one might think a priori that the same phenomenon observed when $\eta = -0.1$ does not happen, however, if one looks at the zoomed graph in Fig. 8(a), one can see clearly that the solution initialized with $\eta = -20$ and $\mu = 1$ is affected by the proximity to the saddle, although less than when initialized with $\eta = -0.1$. This effect is even more evident if we look at Fig. 8(b), which depicts the Frobenius norm of $K_2(t)K_1(t)$ along the solution with $\mu = 1$. Notice that the norm of the matrix product approaches zero,

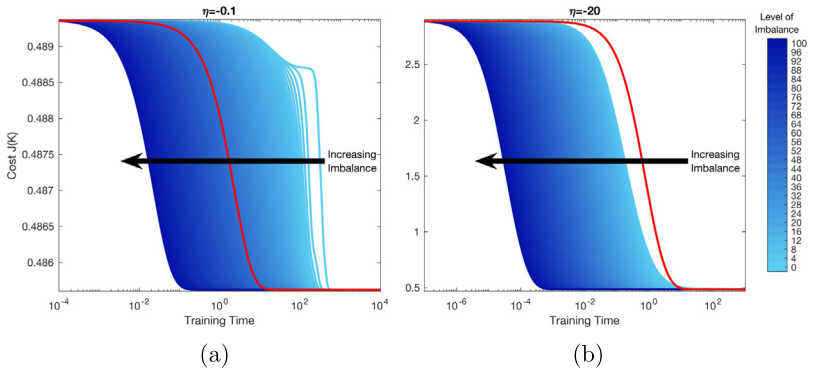

Fig. 7. Simulations done for initializations with $\eta < 0$. Solutions were initialized with $\eta = -0.1$ in (a) and with $\eta = -20$ in (b). Solutions from light to dark blue depict overparameterized solutions with different levels of imbalance $\mu \in [1, 100]$, and the red curve shows the non-overparameterized solution.

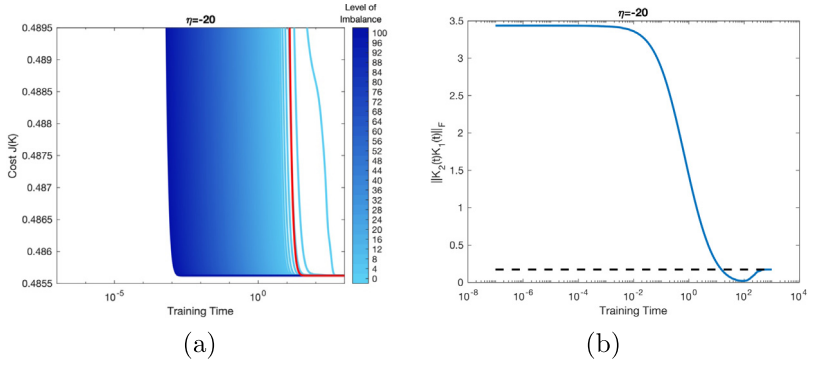

Fig. 8. Simulations done for initializations with $\eta = -20$. In (a) we have a zoomed version of the right graph in Fig. 7, where the influence of the saddle in the imbalanced solution becomes more evident. In (b) we have a plot of the Frobenius norm of the product $K_2(t)K_1(t)$, showing that the solution comes very close to $K_2K_1 = 0$, but then converges to the dashed line, which is the Frobenius norm of K^*.

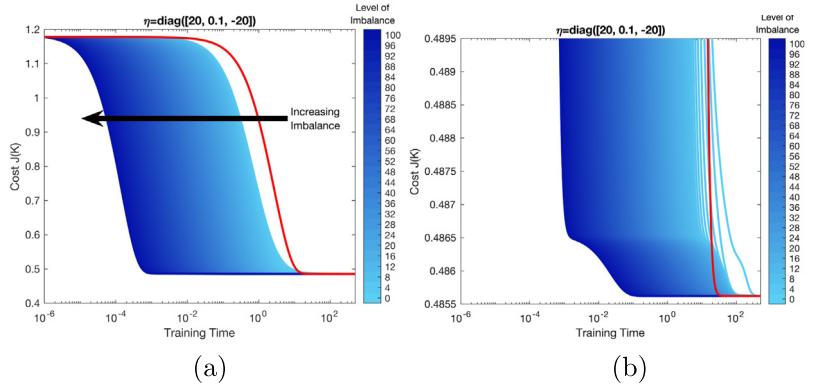

Fig. 9. Simulations done for initializations with $\eta = \text{diag}([20, 0.1, -20])$. On (a) we have the entire trajectory for the solutions and on (b) we have a zoomed version of the plot. Notice that despite this initialization not lying in any of the pre-identified regions of the state space, many of the qualitative observations we made for the behavior of the solution still hold.

but eventually escapes the saddle due to accumulated numerical errors.

Finally, we present a set of simulations selected specifically to not fit in any of the previously discussed cases. To do that, notice that K^* has three singular values, so instead of multiplying all three by the same η, we multiply the first one by 20, the second by 0.1, and the third by -20. The resulting solutions are shown in Fig. 9. Notice that despite this initialization not lying in any of the pre-identified regions of the state space, many of the qualitative observations made for the behavior of the solution still hold. Furthermore, the saddle that the solutions approach in this case is not the origin (which is an isolated critical point), but a non-compact set of saddles, which explains why the effect of the proximity to the saddle affects all solutions, regardless of the level of imbalance.

We conclude this section of simulations with exact knowledge of the value of the gradient with a final observation regarding the level of imbalance. Theoretically, there is no limit to how imbalanced one can make an initialization, however, in practice, the more imbalanced an initialization, the stiffer the resulting ODE, making it harder for numerical solvers for ordinary differential equations to simulate the system. Therefore, although the gradient flow converges "more quickly" in simulation time, the stiff ODE starts to take longer to solve in practice if the initialization is chosen to be too imbalanced. This poses a real-life trade-off on how imbalanced one can make the initialization.

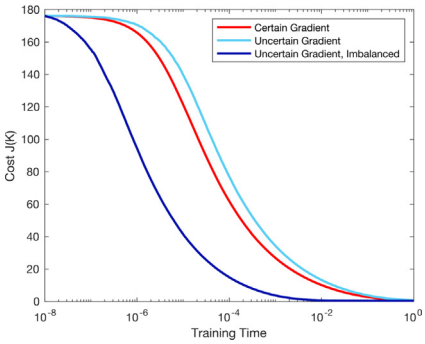

Fig. 10. Simulation results for uncertain oracle. For some randomly picked initialization, the red curve shows the time evolution of the cost function for the vector field generated when the gradient if perfectly computed through its closed-form expression. In light and dark blue, the gradient is estimated numerically through evaluations of the cost function, with the light blue trajectory being the one initialized at the same point as the red trajectory, and the dark blue being the one initialized at the same point as the other two, except for an imbalance factor of 10, as described in Section 5.

5.3. Results with uncertain gradient

We now look at the case where the exact value of the gradient is unknown, and the algorithm samples the value of the cost function at different directions around the current point to estimate it numerically. The code for this set of simulations is also available at de Oliveira (2025).

The gradient is estimated by disturbing the cost at the current value of (K_1, K_2) in the direction of 20 different elementary matrices, *i.e.* in the direction of 20 different entries of (K_1, K_2). The resulting estimated gradient can be viewed as the true gradient plus a noise term. All simulations are done for the same initialization, picked randomly in a distribution around zero — this works for our example because A was specifically selected to be stable.

The resulting solutions are displayed in Fig. 10. Notice that the solution computed with perfect knowledge of the gradient and no enforced imbalance (in red) converges faster than the balanced initialization with the estimated gradient (light blue). However, once we increase the imbalance of the initialization by a factor of 10, the resulting solution (dark blue) converges much quicker than even the solution without uncertainty. This indicates that the disturbance caused by the uncertainty in the dynamics can be overcome by the acceleration brought by imbalanced initializations.

6. Conclusions

This paper investigated the use of linear feedforward neural networks (LFFNNs) for computing the optimal solution of the LQR problem. The theoretical exploration conducted yielded several important results, as summarized below.

In Section 2 we revised key literature results on both gradient methods for the LQR problem and for overparameterized linear regressions, both areas that compose the main contributions of this paper. Then, in Section 3 we introduced the overparameterized policy-optimization LQR problem (poLQR) and proved the main theoretical results of the paper regarding the convergence of the solutions in Theorem 1. Also in this section, we deepened our analysis of the case with a single hidden layer, proving almost everywhere convergence to the optimal feedback matrix in Theorem 2 and Corollary 1, and characterizing all saddles in Lemma 3. We believe these results serve as a strong basis from which to derive an intuitive understanding of the behavior of the solutions of overparameterized formulations. To better develop such intuition, we proceeded in Section 4 with analyzing the vector case, whose simpler setup allows for explicit computation of convergence conditions to the different critical points of the problem. Then, in Section 5, we performed a comprehensive numerical analysis of the problem, showing how different initializations affect the convergence when compared to a non-overparameterized poLQR formulation. The simulations illustrate the distinct behavior the solution can present depending on its initialization and show how the overparameterized formulation can accelerate or decelerate the convergence of the solution to the optimal solution of the poLQR. The simulations indicate that a solution can be arbitrarily accelerated by increasing levels of imbalance for the initialization, however, the stiffness of the resulting ODE provides a practical trade-off to the acceleration.

Many open problems related to the work in this paper remain. A natural follow-up question is how general an optimization problem can be for an overparameterized formulation to hold the properties characterized in this paper. Alternatively, one might be interested in possible practical applications of properties observed in this work, in which case sample and computational complexity analysis are essential to rigorously establishing the trade-offs of adopting such an approach in practice. A more specific open problem lies in the characterization of the center-stable manifold of the saddles of the overparameterized gradient flow. In this paper, we indicated what we believe are the main obstacles to doing so; however, if that were to be done in a future work, it could be leveraged to state formal robustness results for the general case and improve the general understanding of the behavior of the solution.

Appendix A. Systems with strict saddles

We state and prove a few more general results about the convergence of nonlinear systems with multiple equilibria.

In this section, we consider a general differential equation

$$\dot{x} = f(x) \tag{A.1}$$

evolving on an open subset $\mathbb{X} \subseteq \mathbb{R}^n$. We assume that $f : \mathbb{X} \to \mathbb{R}^n$ is continuously differentiable. The solution $x(t) = \phi(t, \xi)$ of (A.1) with initial state $\xi \in \mathbb{X}$ is defined (and in $\mathbb{X}$) on a maximal interval $t \in (T_\xi^{\min}, T_\xi^{\max})$, where $-\infty \leq T_\xi^{\min} < 0 < T_\xi^{\max} \leq +\infty$. The $n \times n$ Jacobian matrix of f evaluated at a point $x \in \mathbb{X}$ is denoted by $J_f(x)$.

For any subset $S \subseteq \mathbb{X}$ define the finite-time domain of attraction $\mathcal{D}_F(S)$ of S as the set of all $\xi \in \mathbb{X}$ such that $T_\xi^{\max} = +\infty$ and there is some $\tau_\xi \geq 0$ such that $\phi(t, \xi) \in S$ for all $t \geq \tau_\xi$.

We say that $\bar{x} \in \mathbb{X}$ is a *strict saddle equilibrium* of (A.1) if

(1) $f(\bar{x}) = 0$ and
(2) $J_f(\bar{x})$ has at least one eigenvalue with positive real part and at least one eigenvalue with non-positive real part.

The following theorem and corollary generalize results for discrete-time gradient iterations that were given in Panageas and Piliouras (2017), which in turn generalized a result from (Lee, Simchowitz, Jordan, & Recht, 2016) that restricted to discrete sets of strict saddles.

Theorem 3. *Suppose that $\bar{x} \in \mathbb{X}$ is a strict saddle equilibrium of (A.1). Then there exists an open neighborhood $B \subseteq \mathbb{X}$ of $\bar{x}$ such that $\mathcal{D}_F(B)$ has Lebesgue measure zero.*

Proof. Pick any equilibrium point $\bar{x} \in \mathbb{X}$. Next modify the vector field f to a vector field g so that g coincides with f on an open neighborhood of U of $\bar{x}$ and g vanishes outside a compact set $K \subseteq \mathbb{X}$. Since g has compact support, solutions are defined for all $t \in \mathbb{R}$, and the map $G : x \mapsto \gamma(1, x)$ (time-1 map for g, where

γ is the flow of g) is a C^1 diffeomorphism. Since $\gamma(1,\bar{x}) = \bar{x}$, it follows that $G(\bar{x}) = \bar{x}$, and since G is a diffeomorphism, there is some neighborhood V of $\bar{x}$ in which $G = F$, where F is the time-1 map for f. The Center-Stable Manifold Theorem as, for example, stated in Shub (2013), Theorem III.7, applied G restricted to V, gives the existence of an open subset B of V and a local center stable manifold W of dimension equal to the number of eigenvalues with nonpositive real part, with the property that for any $x \in B$ such that $G^\ell(x) \in V$ for all $\ell \in \mathbb{Z}_+$ necessarily $x \in W$. Since $F = G$ on V, the same property is true for F.

Pick any point $\xi \in \mathcal{D}_{\mathrm{F}}(B)$ and pick $k = \tau_\xi \geq 0$, without loss of generality a positive integer, such that $\phi(t,\xi) \in B$ for all $t \geq k$. Let $x = \phi(k,\xi)$. Then $F^\ell(x) = \phi(k+\ell,\xi) \in B$ for all $\ell \in \mathbb{Z}_+$, and therefore necessarily $x \in W$. We have established that for each $\xi \in \mathcal{D}_{\mathrm{F}}(B)$ there is some k such that F^k, the time-k map of the flow f, is defined at ξ and satisfies $F^k(\xi) \in W$. It follows that $\mathcal{D}_{\mathrm{F}}(B)$ is the union of the (countably many) sets S_k consisting of those points $x \in \mathbb{X}$ such that $F^k(x) \in W$. Thus it will suffice to show that each set S_k has Lebesgue measure zero. Note that F^k is a local diffeomorphism, it being a time-k map for a differentiable vector field. (It is not necessarily a global diffeomorphism, so we cannot argue that $(F^k)^{-1}(W)$ is diffeomorphic to W. In fact, preimages may not even belong to $\mathbb{X}$.) Thus, there is an open neighborhood N_ξ of ξ in $\mathbb{X}$ that maps diffeomorphically by F^k into an open neighborhood M_ξ of $F^k(\xi)$. By uniqueness of solutions in time $-k$, the preimage of M_ξ is exactly N_ξ. Note that S_k is included in the union N_k over $\xi \in \mathbb{X}$ of the sets N_ξ. Also, for each ξ, $N_\xi \cap S_k$ maps diffeomorphically onto $M_\xi \cap W$, and therefore $N_\xi \cap S_k$ has Lebesgue measure zero (because W has measure zero and diffeomorphisms transform null sets into null sets). Recall that Lindelöf's Lemma (see e.g (Kelley, 1975)) insures that every open cover of any subset S of $\mathbb{R}^n$ (or more generally, of any second-countable space) admits a countable subcover. Applied to N_k, we have a countable subcover by sets N_{ξ_k}, and for each of these $N_{\xi_k} \cap S_k$ has measure zero, so $N_k \cap S_k = S_k$ has measure zero as well. ■

Corollary 3. *Suppose that $E \subseteq \mathbb{X}$ is a set consisting of strict saddle equilibria of (A.1). Then the set $\mathcal{C}_E$ of points $\xi \in \mathbb{X}$ whose trajectories converge to points in E has measure zero.*

Proof. For each $\bar{x} \in E$, we may pick by Theorem 3 an open neighborhood $B_{\bar{x}} \subseteq \mathbb{X}$ of $\bar{x}$ such that $\mathcal{D}_{\mathrm{F}}(B_{\bar{x}})$ has measure zero. The union of the sets $B_{\bar{x}}$ covers E. By Lindelöf's Lemma applied to $S = E$, we conclude that there is a countable subset of balls $\{B_{\bar{x}_k}, k \in \mathbb{Z}_+\}$ which covers E. We claim that $\mathcal{C}_E \subseteq \bigcup_k \mathcal{D}_{\mathrm{F}}(B_{\bar{x}_k})$. Since a union of measure zero sets has measure zero, this will establish the claim. So pick any $\xi \in \mathcal{C}_E$. Thus, $\phi(t,\xi) \to \bar{x}$ for some $\bar{x} \in E$. Since $E \subseteq \bigcup_k B_{\bar{x}_k}$, it follows that $\bar{x} \in B_{\bar{x}_k}$ for some k. Since $B_{\bar{x}_k}$ is a neighborhood of $\bar{x}$, this means that there is some $\tau_\xi \geq 0$ such that $\phi(t,\xi) \in B_{\bar{x}_k}$ for all $t \geq \tau_\xi$. Therefore $\xi \in \mathcal{D}_{\mathrm{F}}(B_{\bar{x}_k})$. This completes the proof. ■

Corollary 4. *Let $\mathcal{L}$ be a real-analytic (loss) function from $\mathbb{X}$ into $\mathbb{R}_+$. Let f be the gradient of $\mathcal{L}$ and let the set Z of points where $f(x) = 0$ be the union of two sets $Z = M \cup S$, where M is the set of points at which $\mathcal{L}$ is minimized, and S consists of strict saddles for the gradient flow dynamics $\dot{x} = -f(x)$. Assume in addition that every trajectory of the gradient flow dynamics is pre-compact. Then, except for a set of measure zero, all trajectories converge to M.*

Proof. Lojasiewcz's Theorem states that every pre-compact trajectory of a real analytic gradient system converges to a unique equilibrium. This theorem is given (Łojasiewicz, 1984) and an excellent exposition is given in Colding and Minicozzi II (2014). From this, one can apply Corollary 3 and conclude that the set of initializations for which the trajectories of the system converge to S must have measure zero. ■

Appendix B. Proofs of main results

Proof of Lemma 1. For any i between 1 and N, notice that under closed loop with $u = K_N \dots K_i \dots K_1 Cx$ the dynamics of the system become

$$\begin{aligned}\dot{x} &= (A + BK_n \dots K_i \dots K_1 C)x \\ &= (A + B_i K_i C_i)x,\end{aligned}$$

which are equivalent to a system under simple output feedback K_i, input matrix B_i and output matrix C_i. When computing the partial derivatives in (8), one fixes the value of all K_j for $j \neq i$, therefore computing $\partial J/\partial K_i$ is equivalent to computing the partial derivative for the linear system with parameter matrices A, B_i, C_i and single feedback matrix K_i, as done in (3).

From here, the remainder of the proof is obtained by applying Theorem 3.2 of Rautert and Sachs (1997) to the system (A, B_i, C_i, K_i). We omit it due to space limitations, but it is reproduced in full in the arXiv version of this paper (de Oliveira et al., 2024a). ■

Proof of Lemma 2. First, notice from the definitions in Lemma 1 that for all i between 1 and $N-1$

$$B_i = BK_N \dots K_{i+1} = B_{i+1}K_{i+1} \tag{B.1}$$

$$C_{i+1} = K_i \dots K_1 = K_i C_i \tag{B.2}$$

$$\begin{aligned}R_i &= K_{i+1}^\top \dots K_N^\top R K_N \dots K_{i+1} \\ &= K_{i+1}^\top R_{i+1} K_{i+1}.\end{aligned} \tag{B.3}$$

Then notice that $\dot{K}_i K_i^\top = K_{i+1}^\top \dot{K}_{i+1}$, which after substituting into $\frac{d}{dt}(K_i K_i^\top) = \dot{K}_i K_i^\top + K_i \dot{K}_i^\top$, results in $\frac{d}{dt}(K_i K_i^\top - K_{i+1}^\top K_{i+1}) = 0$, proving the theorem. ■

Proof of Theorem 1. We begin the proof by showing that $\overline{K}(t) := K_N(t) \dots K_1(t)$ is precompact. To do that, notice that $J(\overline{K})$ is non-increasing along any trajectory, since $\dot{J}(\overline{K}) = \dot{J}(K_N \dots K_1) = -\sum_{i=1}^N \|\nabla_{K_i} J(K_N \dots K_1)\|_F^2 \leq 0$. Then, because $J(\overline{K})$ is a proper function, the level sets of J in the space of $\overline{K}$ are compact, which means that $\overline{K}(t)$ lies within a compact set along any trajectory (otherwise $J(\overline{K})$ would increase), proving precompactness of $\overline{K}(t)$.

The next step is to show that $\overline{K}$ being bounded implies that all K_is are also bounded. This part of the argument follows closely the proof of Proposition 1 in Chitour et al. (2023), and of Theorem 3.2 of Bah et al. (2022), albeit done for the LQ cost rather than linear regression.

To show that $\overline{K}$ bounded implies K_i bounded for all i, notice that a consequence of Lemma 2 is that for any i, j between 1 and N, there exists some constant c_{ij} such that

$$\|K_i\|_F^2 = \|K_j\|_F^2 + \operatorname{trace}\left(W_{ij}\right). \tag{B.4}$$

where $W_{ij} = \sum_{k=i}^j C_k$ if $i < j$, $W_{ij} = -\sum_{k=j}^i C_k$ if $i > j$ and $W_{ij} = 0$ if $i = j$. This is easily verified by taking the trace on both sides of (11) and concatenating the resulting equations from i to j. With this established, following relationship between $\overline{K}$ and K_i can be used:

$$\|K_i\|_F \leq \eta_i \|\overline{K}\|_F^{1/N} + \xi_i, \tag{B.5}$$

where η_i and ξ_i depend only on the initialization of the parameter matrices. The derivation of (B.5) can be obtained in the same way as equation (3.1) of Bah et al. (2022) once (B.4) is established, regardless of the different cost functions. Due to limited space, we refer the reader to that paper for this part of the proof, or to the arXiv version of the paper for the complete proof done for the LQR cost function (de Oliveira et al., 2024a).

Proof of Theorem 2. Consider the Taylor expansion of the LQR cost as follows

$$\begin{aligned}J((K_2+\mathrm{d}K_2)(K_1+\mathrm{d}K_1)) = \\ J(K_2K_1)+J'_{\mathrm{d}K_1}(K_2K_1)+J'_{\mathrm{d}K_2}(K_2K_1) \\ +\frac{1}{2}J''_{\mathrm{d}K_1^2}(K_2K_1)+\frac{1}{2}J''_{\mathrm{d}K_2^2}(K_2K_1) \\ +J''_{\mathrm{d}K_1\mathrm{d}K_2}(K_2K_1)+\text{h.o.t.}\end{aligned}$$

The expression for $J''_{\mathrm{d}K_1^2}(K_2K_1)$, $J''_{\mathrm{d}K_2^2}$, and $J''_{\mathrm{d}K_1\mathrm{d}K_2}(K_2K_1)$ can be computed by

$$\begin{aligned}J''_{\mathrm{d}K_1^2}(K_2K_1) &= 4\,\mathrm{trace}\left(\mathrm{d}K_1^\top K_2^\top BP'_{\mathrm{d}K_1}L_{\mathbf{K}}\right) \\ &\quad +2\,\mathrm{trace}\left(\mathrm{d}K_1^\top K_2^\top RK_2\mathrm{d}K_1L_{\mathbf{K}}\right),\end{aligned}$$

$$\begin{aligned}J''_{\mathrm{d}K_2^2}(K_2K_1) &= 4\,\mathrm{trace}\left(K_1^\top \mathrm{d}K_2^\top BP'_{\mathrm{d}K_2}L_{\mathbf{K}}\right) \\ &\quad +2\,\mathrm{trace}\left(K_1^\top dK_2^\top RK_2K_1L_{\mathbf{K}}\right),\end{aligned}$$

and

$$\begin{aligned}&J''_{\mathrm{d}K_2\mathrm{d}K_1}(K_2K_1) = \\ &+\mathrm{trace}\left(2[B^\top P'_{\mathrm{d}K_2}+R\mathrm{d}K_2K_1]L_{\mathbf{K}}\mathrm{d}K_1^\top K_2^\top\right) \\ &+\mathrm{trace}\left(2[B^\top P_{\mathbf{K}}+RK_2K_1]L'_{\mathrm{d}K_2}\mathrm{d}K_1^\top K_2^\top\right) \\ &+\mathrm{trace}\left(2[B^\top P_{\mathbf{K}}+RK_2K_1]L_{\mathbf{K}}\mathrm{d}K_1^\top \mathrm{d}K_2^\top\right),\end{aligned}$$

where $P'_{\mathrm{d}K_1}$ are the solution of the following Lyapunov equation

$$\begin{aligned}&P'_{\mathrm{d}K_1}[A+BK_2K_1]+[A+BK_2K_1]^\top P'_{\mathrm{d}K_1} = \\ &-\mathrm{d}K_1^\top K_2^\top[B^\top P_{\mathbf{K}}+RK_2K_1] \\ &-[B^\top P_{\mathbf{K}}+RK_2K_1]^\top K_2\mathrm{d}K_1\end{aligned} \tag{B.6}$$

with an analogous one existing for $P'_{\mathrm{d}K_2}$. Due to space limitations, we refer the reader to the arXiv version of this paper (de Oliveira et al., 2024a) for the explicit derivation of the expressions above. With this, define the Hessian function as

$$\begin{aligned}H(K_1,K_2,\mathrm{d}K_1,\mathrm{d}K_2) &:= \frac{1}{2}J''_{\mathrm{d}K_1^2}(K_1,K_2) \\ &+\frac{1}{2}J''_{\mathrm{d}K_2^2}(K_1,K_2)+J''_{\mathrm{d}K_1\mathrm{d}K_2}(K_1,K_2).\end{aligned} \tag{B.7}$$

With this established, let (K_1, K_2) be a critical point of the gradient flow, *i.e.* $\dot{K}_1 = \nabla_K J(K_2K_1)K_2^\top = 0$ and $\dot{K}_2 = K_1^\top\nabla_K J(K_2K_1) = 0$ where $\nabla_K J(K_2K_1)$ is the expression in (3) computed for $K = K_2K_1$. For the overparameterized poLQR, the conditions above hold for any K_1, K_2 orthogonal to $\nabla_K J(K_2K_1)$, even if $\nabla_K J(K_2K_1)\neq 0$.

This orthogonality implies that there must exist two unitary vectors ψ and ϕ such that $K_1^\top\phi = 0$, $\psi^\top K_2^\top = 0$, $\phi^\top\nabla_K J(K_2K_1) = \lambda\psi$ and $\nabla_K J(K_2K_1)\psi = \lambda\phi^\top$ for some $\lambda < 0$, for if no such vectors existed and $\nabla_K J(K_2K_1)\neq 0$ then $\dot{K}_1 = 0$ and $\dot{K}_2 = 0$ could never hold. To prove this by contradiction, assume no such ψ and ϕ exist, and since $\nabla_K J(K_2, K_1) \neq 0$ by assumption, let ψ and ϕ be any unitary vectors such that $\nabla_K J(K_2K_1)\psi = \lambda\phi^\top$ for some λ. Then notice that $\dot{K}_2\psi = K_1^\top\nabla_K J(K_2K_1)\psi = K_1^\top\phi\lambda \neq 0$ and $\phi^\top\dot{K}_1 = \phi^\top\nabla_K J(K_2K_1)K_2^\top = \lambda\psi^\top K_2^\top \neq 0$, reaching contradiction.

With this, let γ_1 and γ_2 be any two unitary vectors such that $\gamma_1^\top K_1^\top = 0$, $K_2^\top\gamma_2 = 0$, and $\gamma_1^\top\gamma_2 > 0$ (if $\gamma_1^\top\gamma_2 < 0$, simply pick $-\gamma_1$ instead). Define $\mathrm{d}K_1 = \psi\gamma_2^\top$ and $\mathrm{d}K_2 = \phi\gamma_1^\top$. For this choice of $\mathrm{d}K_1$ and $\mathrm{d}K_2$ notice that $\mathrm{d}K_2K_1 = 0$ and $K_2\mathrm{d}K_1 = 0$, which implies that $J''_{dK_1^2} = 0, J''_{dK_2^2} = 0$, and

$$\begin{aligned}&J''_{dK_1dK_2}(K_2K_1) = \\ &\mathrm{trace}\left(\mathrm{d}K_2^\top\nabla_K J(K_2K_1)\mathrm{d}K_1^\top\right) \le \lambda < 0,\end{aligned}$$

proving that the Hessian has at least one negative eigenvalue, which implies that the spurious equilibria is a strict saddle of the gradient flow. ■

Before proceeding to prove Lemma 3, following the order in the paper, we must first introduce and prove an auxiliary lemma that will help us prove Lemma 3.

Lemma 4. *Given two matrices $A \in \mathbb{R}^{p\times o}$ and $B \in \mathbb{R}^{q\times o}$ for $p, q, o \in \mathbb{N}$ with $q \ge o$, the following two statements are equivalent*

(1) $AB^\top = 0$;
(2) There exist orthogonal matrices Ψ_A, Φ, and Ψ_B, and rectangular diagonal matrices with non-negative diagonal elements Σ_A and Σ_B, such that $A = \Psi_A\Sigma_A\Phi^\top$, and $B = \Psi_B\Sigma_B\Phi^\top$ are SVDs of A and B, and $\Sigma_A\Sigma_B^\top = 0$.

Furthermore, in (2) we can write Σ_A and Σ_B as

$$\Sigma_A = \begin{bmatrix}\bar{\Sigma}_A & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},\quad \Sigma_B = \begin{bmatrix}0 & 0 & 0\\ 0 & \bar{\Sigma}_B & 0\\ 0 & 0 & 0\end{bmatrix}$$

where $\bar{\Sigma}_A$ and $\bar{\Sigma}_B$ are diagonal matrices whose main diagonal elements are the nonzero singular values of A and B respectively.

Due to space limitations, we refer the reader to the arXiv version of this paper (de Oliveira et al., 2024a) for the proof. With this auxiliary lemma established, one can prove the statement in Lemma 3 as follows.

Proof of Lemma 3. To prove that (2)⇒(1) we simply compute $[\dot{K}_1^\top; \dot{K}_2]$ for a (K_1, K_2) that satisfies the properties in (2) and verify that it is equal to zero. To prove that (1)⇒(2), apply Lemma 4 with $A = \nabla_{\mathbf{K}}J(K_2K_1)^\top$ and $B = K_2^\top$, which implies that $o = n \le k = q$, since $\dot{K}_1 = -\nabla_{\mathbf{K}}J(K_2K_1)^\top K_2 = 0$, which allows us to write K_2 and $\nabla_{\mathbf{K}}J(K_2K_1)$ as follows:

$$\begin{aligned}\nabla_{\mathbf{K}}J(K_2K_1) &= \Psi\begin{bmatrix}\bar{\Sigma}_2 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\Phi_{K_2}^\top = \Psi\Sigma_2\Phi_{K_2}^\top \\ K_2 &= \Psi\begin{bmatrix}0 & 0 & 0\\ 0 & \bar{\Sigma}_{K_2} & 0\\ 0 & 0 & 0\end{bmatrix}\Gamma_2 = \Psi\Sigma_{K_2}\Gamma_2.\end{aligned} \tag{B.8}$$

Similarly, applying Lemma 4 with $A = \nabla_{\mathbf{K}}J(K_2K_1)$ and $B = K_1$ gives

$$\begin{aligned}\nabla_{\mathbf{K}}J(K_2K_1) &= \Psi_{K_1}\Sigma_1\Phi^\top \\ K_1 &= \Gamma_{K_1}\Sigma_{K_1}\Phi^\top.\end{aligned} \tag{B.9}$$

where Σ_1 and Σ_{K_1} have the same structure as Σ_2 and Σ_{K_2}. Notice that even if $\bar{\Sigma}_1 \neq \bar{\Sigma}_2$, they must still have the same diagonal elements, albeit possibly in a different order. Changing the order of the elements of $\bar{\Sigma}_1$ and $\bar{\Sigma}_2$ so they match means swapping the columns of Ψ, Φ, Ψ_{K_1} and Φ_{K_2}, but we can also swap the singular vectors corresponding to the kernels of K_2 and $K_1^\top$ such that the results from Lemma 4 still hold. As such we can assume without loss of generality that $\bar{\Sigma}_1 = \bar{\Sigma}_2 = \bar{\Sigma}$. Notice that this is enough to prove that $1 \to 2b$, since $\nabla_{\mathbf{K}}J(K_2K-1)K_1^\top = \Psi_{K_1}\Sigma\Phi^\top\Phi\Sigma_{K_1}^\top\Gamma_1 = 0 \iff \Sigma\Sigma_{K_1}^\top = 0$ (and similarly for $\Sigma^\top\Sigma_{K_2} = 0$). Next, notice that

$$\Psi_{K_1}\Sigma_1\Phi^\top = \Psi\Sigma_2\Phi_{K_2} \tag{B.10}$$

implies that

$$\Psi_{1,K_1}\bar{\Sigma}\Phi_1^\top = \Psi_1\bar{\Sigma}\Phi_{1,K_2}^\top. \tag{B.11}$$

If we impose, for example, $\Psi_{1,K_1} = \Psi_1$, then we must also impose $\Phi = \Phi_{1,K_2}$. This leaves the SVD of K_2 intact, but changes part of the SVD of K_1. To show it still satisfies Lemma 4, consider

$$K_1^\top = \begin{bmatrix}\Phi_{1,K_2} & \Phi_2 & \Phi_3\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & \bar{\Sigma}_{K_1} & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\Gamma_{1,P}^\top\\ \Gamma_{2,P}^\top\\ \Gamma_{3,P}^\top\end{bmatrix} = \Phi_2\bar{\Sigma}_{K_1}\Gamma_{2,P}^\top$$

which shows that the above is still a valid SVD of $K_1^\top$ as long as the span of the columns of $[\Phi_{1,K_2}, \Phi_3]$ is equal to the Kernel of K_1. Indeed, we know that colspace $([\Phi_1, \Phi_3])$ = kernel (K_1), and since (B.11) shows that colspace (Φ_1) = colspace $\left(\Phi_{1,K_2}\right)$ (by contradiction) then we can conclude that colspace $\left([\Phi_{1,K_2}, \Phi_3]\right)$ = kernel (K_1). We have, therefore, established that we can always match the singular vectors associated with the nonzero singular values of $\nabla_{\mathbf{K}} J(K_2K_1)$ for the SVDs in (B.8) and (B.9) and still satisfy both conditions on Lemma 4.

Next, notice that we can simply pick $\Psi_{3,K_1} = \Psi_3$, $\Phi_{3,K_2} = \Phi_3$, since all matrices are related to the intersection of the kernels and their choice is arbitrary as long as they compose an orthonormal basis of kernel $(\nabla_{\mathbf{K}} J(K_2K_1)) \cap$ kernel (K_1) and of kernel $\left(\nabla_{\mathbf{K}} J(K_2K_1)^\top\right) \cap$ kernel $\left(K_2^\top\right)$ respectively.

For the remaining matrices, Ψ_2 and Φ_2 are imposed by the SVDs of K_2 and K_1 respectively, and as such cannot be changed arbitrarily. We can, however, freely change the columns of Ψ_{2,K_1} (resp. Φ_{2,K_2}) as long as when composed with the columns of Ψ_3 (resp. Φ_3) they form a basis of the kernel of $\nabla_{\mathbf{K}} J(K_2K_1)^\top$ (resp. $\nabla_{\mathbf{K}} J(K_2K_1)$). Therefore we can select Ψ_{2,K_1} (resp. Φ_{2,K_2}) to be equal to Ψ_2 (resp Φ_2) without any loss of generality, completing the proof.

Proof of Corollary 2. Let (K_1, K_2) be a saddle point of the gradient flow dynamics (8) with $N = 2$ and assume rank $(K_2K_1) = p < \min(m, n)$. Also, let K^* be such that $\nabla_K J(K^*) = 0$. Then, let v be any vector such that $v^\top \nabla_{\mathbf{K}} J(K_2K_1) = 0$ and notice that

$$v^\top \nabla_K J(K_2K_1) = v^\top \nabla_K J(K^*)$$
$$v^\top R K_2 K_1 = v^\top R K^*.$$

Then, let $[v_1, \ldots, v_p]$ be p linearly independent (LI) vectors such that $v_i^\top \nabla_K J(K_2K_1) = 0$, a set which must exist because from Lemma 3 the left kernel of $\nabla_K J(K_2K_1)$ has dimension p. Then, notice that for two LI vectors u and v and full rank matrix R, $u^\top R$ and $v^\top R$ must also be LI. Finally, let Ψ_1^* be the matrix whose columns are vectors composing an orthonormal base of span$(v_1, \ldots, v_p)$ and notice that

$$(\Psi_1^*)^\top K_2 K_1 = (\Psi_1^*)^\top K^*,$$

however, since K_2K_1 is rank p, one can always pick the orthonormal basis that compose the columns of Ψ_1^* to be the left singular vectors of K_2K_1, implying that there exist a Φ_1^* whose columns are orthonormal vectors such that

$$(\Psi_1^*)^\top K_2 K_1 \Phi_1^* = \Sigma_1^* = (\Psi_1^*)^\top K^* \Phi_1^*.$$

For the remaining components of the SVD of K^* we can pick whichever eigenvectors are left since they are all a basis for the left kernel of K_2K_1. ■

Appendix C. Proofs for the simple example

Due to space limitations, we refer the reader to the arXiv version of this paper (de Oliveira et al., 2024a) for the proofs of the results for the simple example.

References

Agarwal, A., Kakade, S. M., Lee, J. D., & Mahajan, G. (2021). On the theory of policy gradient methods: Optimality, approximation, and distribution shift. *Journal of Machine Learning Research*, *22*(98), 1–76.

Alali, M., & Imani, M. (2023). Reinforcement learning data-acquiring for causal inference of regulatory networks. In *American control conference*. IEEE.

Arora, S., Du, S., Hu, W., Li, Z., & Wang, R. (2019). Fine-grained analysis of optimization and generalization for overparameterized two-layer neural networks. In *International conference on machine learning* (pp. 322–332). PMLR.

Bah, B., Rauhut, H., Terstiege, U., & Westdickenberg, M. (2022). Learning deep linear neural networks: Riemannian gradient flows and convergence to global minimizers. *Information and Inference: A Journal of the IMA*, *11*(1), 307–353.

Baldi, P., & Hornik, K. (1989). Neural networks and principal component analysis: Learning from examples without local minima. *Neural Networks*, *2*(1), 53–58.

Chitour, Y., Liao, Z., & Couillet, R. (2023). A geometric approach of gradient descent algorithms in linear neural networks. *Mathematical Control and Related Fields*, *13*(3), 918–945.

Colding, T. H., & Minicozzi II, W. P. (2014). Lojasiewicz inequalities and applications. In H.-D. Cao, R. Schoen, & S.-T. Yau (Eds.), *XIX of surveys in differential geometry*. Boston: International Press of Boston.

Cui, L., Jiang, Z.-P., & Sontag, E. D. (2024). Small-disturbance input-to-state stability of perturbed gradient flows: Applications to LQR problem. *Systems & Control Letters*, *188*, Article 105804.

de Oliveira, A. C. B. (2025). OVP_ModelFree_LQR_Automatica2025. https://github.com/ArthurCBO/OVP_ModelFree_LQR_Automatica2025.git,

de Oliveira, A. C. B., Siami, M., & Sontag, E. D. (2023). Dynamics and perturbations of overparameterized linear neural networks. In *2023 62nd IEEE conference on decision and control* (pp. 7356–7361). IEEE.

de Oliveira, A. C. B., Siami, M., & Sontag, E. D. (2024a). Convergence analysis of overparametrized LQR formulations. arXiv preprint arXiv:2408.15456.

de Oliveira, A. C. B., Siami, M., & Sontag, E. D. (2024b). Remarks on the gradient training of linear neural network based feedback for the LQR problem. In *2024 63rd IEEE conference on decision and control* (pp. 7846–7852). IEEE.

Eftekhari, A. (2020). Training linear neural networks: Non-local convergence and complexity results. In *Proceedings of the 37th international conference on machine learning* (pp. 2836–2847). PMLR, ISSN: 2640-3498.

Fazel, M., Ge, R., Kakade, S., & Mesbahi, M. (2018). Global convergence of policy gradient methods for the linear quadratic regulator. In *International conference on machine learning* (pp. 1467–1476). PMLR.

Gravell, B., Esfahani, P. M., & Summers, T. (2020). Learning optimal controllers for linear systems with multiplicative noise via policy gradient. *IEEE Transactions on Automatic Control*, *66*(11), 5283–5298.

Hu, B., Zhang, K., Li, N., Mesbahi, M., Fazel, M., & Başar, T. (2023). Toward a theoretical foundation of policy optimization for learning control policies. *Annual Review of Control, Robotics, and Autonomous Systems*, *6*(1), 123–158.

Jin, C., Ge, R., Netrapalli, P., Kakade, S. M., & Jordan, M. I. (2017). How to escape saddle points efficiently. In *International conference on machine learning* (pp. 1724–1732). PMLR.

Kawaguchi, K. (2016). Deep learning without poor local minima. In *Advances in neural information processing systems*: *vol. 29*, Curran Associates, Inc..

Kelley, J. (1975). General topology. *Graduate texts in mathematics*, Springer New York.

Lee, J. D., Simchowitz, M., Jordan, M. I., & Recht, B. (2016). Gradient Descent Only Converges to Minimizers. In V. Feldman, A. Rakhlin, & O. Shamir (Eds.), *Proceedings of machine learning research*: *vol. 49*, *29th annual conference on learning theory* (pp. 1246–1257). Columbia University, New York, New York, USA: PMLR.

Levine, W., & Athans, M. (1970). On the determination of the optimal constant output feedback gains for linear multivariable systems. *IEEE Transactions on Automatic Control*, *15*(1), 44–48.

Łojasiewicz, S. (1984). Sur les trajectoires du gradient d'une fonction analytique. (Trajectories of the gradient of an analytic function). *Semin. Geom., Univ. Studi Bologna*, *1982/1983*, 115–117.

Min, H., Tarmoun, S., Vidal, R., & Mallada, E. (2021). On the Explicit Role of Initialization on the Convergence and Implicit Bias of Overparametrized Linear Networks. In *Proceedings of the 38th international conference on machine learning* (pp. 7760–7768). PMLR, ISSN: 2640-3498.

Min, H., Vidal, R., & Mallada, E. (2023). On the convergence of gradient flow on multi-layer linear models. In *International conference on machine learning* (pp. 24850–24887). PMLR.

Mohammadi, H., Soltanolkotabi, M., & Jovanović, M. R. (2021). On the lack of gradient domination for linear quadratic Gaussian problems with incomplete state information. In *2021 60th IEEE conference on decision and control* (pp. 1120–1124). IEEE.

Mohammadi, H., Zare, A., Soltanolkotabi, M., & Jovanovic, M. R. (2022). Convergence and Sample Complexity of Gradient Methods for the Model-Free Linear–Quadratic Regulator Problem. *IEEE Transactions on Automatic Control*, *67*(5), 2435–2450.

Motamedi, E., Behzad, K., Zandi, R., Salehinejad, H., & Siami, M. (2024). Robustness evaluation of machine learning models for robot arm action recognition in noisy environments. In *ICASSP 2024-2024 IEEE international conference on acoustics, speech and signal processing* (pp. 6215–6219). IEEE.

Nesterov, Y., & Polyak, B. T. (2006). Cubic regularization of Newton method and its global performance. *Mathematical Programming*, *108*(1), 177–205.

Panageas, I., & Piliouras, G. (2017). Gradient Descent Only Converges to Minimizers: Non-Isolated Critical Points and Invariant Regions. In C. H. Papadimitriou (Ed.), *Leibniz international proceedings in informatics (lIPics)*: *vol. 67*, *8th innovations in theoretical computer science conference* (pp. 2:1–2:12). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

Polyak, B. T. (1963). Gradient methods for the minimisation of functionals. *USSR Computational Mathematics and Mathematical Physics*, *3*(4), 864–878.
Rautert, T., & Sachs, E. W. (1997). Computational Design of Optimal Output Feedback Controllers. *SIAM Journal on Optimization*, *7*(3), 837–852.
Ravari, A., Ghoreishi, S. F., & Imani, M. (2022). Optimal Recursive Expert-Enabled Inference in Regulatory Networks. *IEEE Control Systems Letters*, *7*, 1027–1032.
Ravari, A., Ghoreishi, S. F., & Imani, M. (2024). Optimal Inference of Hidden Markov Models Through Expert-Acquired Data. *IEEE Transactions on Artificial Intelligence*.
Shub, M. (2013). *Global stability of dynamical systems*. Springer.
Sontag, E. D. (2013). *Mathematical control theory: deterministic finite dimensional systems*: *vol. 6*, Springer Science & Business Media.
Sontag, E. D. (2022). Remarks on input to state stability of perturbed gradient flows, motivated by model-free feedback control learning. *Systems & Control Letters*, *161*, Article 105138.
Sznaier, M., Olshevsky, A., & Sontag, E. D. (2022). The role of systems theory in control oriented learning. In *25th international symposium on mathematical theory of networks and systems*.
Tarmoun, S., Franca, G., Haeffele, B. D., & Vidal, R. (2021). Understanding the Dynamics of Gradient Flow in Overparameterized Linear models. In *Proceedings of the 38th international conference on machine learning* (pp. 10153–10161). PMLR, ISSN: 2640-3498.
Wafi, M. K., & Siami, M. (2023). A Comparative Analysis of Reinforcement Learning and Adaptive Control Techniques for Linear Uncertain Systems. In *2023 proceedings of the conference on control and its applications* (pp. 25–32). SIAM.
Zandi, R., Salehinejad, H., Behzad, K., Motamedi, E., & Siami, M. (2023). Robot motion prediction by channel state information. In *2023 IEEE 33rd international workshop on machine learning for signal processing* (pp. 1–6). IEEE.
Zhang, X., & Başar, T. (2023). Revisiting LQR control from the perspective of receding-horizon policy gradient. *IEEE Control Systems Letters*, *7*, 1664–1669.

Arthur Castello B. de Oliveira received his B.Sc. in electrical engineering with emphasis in control and M.Sc. in control engineering from the University of São Paulo in 2016 and 2019 respectively. He received his Ph.D. degree in electrical engineering from Northeastern University in 2024. He is currently a postdoctoral research associate at Northeastern University, Boston, MA. His research interests include nonlinear systems and control, nonlinear optimization and optimal control, gradient systems, artificial intelligence techniques and machine learning.

Milad Siami received his dual B.Sc. degrees in electrical engineering and pure mathematics from Sharif University of Technology in 2009, and his M.Sc. degree in electrical engineering from Sharif University of Technology in 2011. He received his M.Sc. and Ph.D. degrees in mechanical engineering from Lehigh University in 2014 and 2017, respectively. From 2017 to 2019, he was a postdoctoral associate at MIT Institute for Data, Systems, and Society. He is currently an Associate Professor in the Department of Electrical & Computer Engineering at Northeastern University, Boston, MA. His research interests include distributed control systems, distributed optimization, and sparse sensing, which are applied in robotics and cyber–physical systems.

Eduardo D. Sontag received his Licenciado in Mathematics at the University of Buenos Aires (1972) and a Ph.D. in Mathematics (1977) under Rudolf E. Kalman at the University of Florida. From 1977 to 2017, he was at Rutgers University, where he was a Distinguished Professor of Mathematics and a Member of the Graduate Faculty of the Departments of Computer Science of Electrical and Computer Engineering and the Cancer Institute of NJ. He directed the undergraduate Biomathematics Interdisciplinary Major and the Center for Quantitative Biology, and was Graduate Director at the Institute for Quantitative Biomedicine. In January 2018, Dr. Sontag became a University Distinguished Professor in the Departments of Electrical and Computer Engineering and of BioEngineering at Northeastern University, where he is also affiliated with the Mathematics and the Chemical Engineering departments. Since 2006, he has been a Research Affiliate at the Laboratory for Information and Decision Systems, MIT, and since 2018 he has been the Faculty Member in the Program in Therapeutic Science at Harvard Medical School. His major current research interests lie in several areas of machine learning, control and dynamical systems theory, systems molecular biology, cancer and immunology, and computational biology. Sontag is a Fellow of IEEE, AMS, SIAM, and IFAC. He was awarded the 2001 Reid Prize in Mathematics from SIAM, the 2002 Hendrik W. Bode Lecture Prize and the 2011 Control Systems Field Award from the IEEE, the 2022 Richard E. Bellman Control Heritage Award, the 2023 IFAC Triennial Award on Nonlinear Control, the 2002 Board of Trustees Award for Excellence in Research from Rutgers, and the 2005 Teacher/Scholar Award from Rutgers. In 2024, he was elected to the American Academy of Arts and Sciences and in 2025 to the US National academy of Sciences.