2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Riemannian Constrained Policy Optimization via Geometric Stability Certificates*

Shahriar Talebi, *Student Member, IEEE,* and Mehran Mesbahi, *Fellow, IEEE*

2022 IEEE 61st Conference on Decision and Control (CDC) | 978-1-6654-6761-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CDC51059.2022.9992877

***Abstract*—In this paper, we consider policy optimization over the Riemannian submanifolds of stabilizing controllers arising from constrained Linear Quadratic Regulators (LQR), including output feedback and structured synthesis. In this direction, we provide a Riemannian Newton-type algorithm that enjoys local convergence guarantees and exploits the inherent geometry of the problem. Instead of relying on the exponential mapping or a global retraction, the proposed algorithm revolves around the developed stability certificate and the constraint structure, utilizing the intrinsic geometry of the synthesis problem. We then showcase the utility of the proposed algorithm through numerical examples.**

Index Terms*— *Structured LQR Control; Output-feedback LQR Control; Policy Optimization on Manifolds

I. Introduction

Policy Optimization (PO) for control, such as LQR synthesis, has recently attracted considerable attention in the literature, as it establishes a direct bridge between control and learning, and in the meantime, puts fundamental questions at the interface of system theory, dynamics, and optimization under the spotlight. PO for *linearly constrained* LQR, e.g., state-feedback Structured Linear Quadratic Regulators (SLQR) and Output-feedback Linear Quadratic Regulators (OLQR), on the other hand, has been explored to a much lesser degree. This is more a reflection of the intricate geometry of such constrained synthesis problems than their importance. For example, although reparameterization of the LQR problem to convex optimization is possible for unconstrained scenarios [2], trivial constraints on policy often lead to nontrivial and non-convex problems after such reparameterizations. Furthermore, the domain of the optimization problems for constrained LQR (and its variants) are generally non-convex [3] and even disconnected [4], with potentially multiple local minima on each component; as such, there are even no guarantees that first-order stationary points are local minima; see [5], [6]. This state of affairs might appear discouraging, as in our view, constrained synthesis problems are often the *main motivation* for control design in practice. Over the past few years, PO for control synthesis has been investigated through the lens of first-order methods for many variants of the LQR problem, such as OLQR [7] and its model-free version [8]. Crucial for such studies has been the so-called gradient dominance property [9], [10], that plays a critical role for showing global convergence of first order methods for control synthesis problems [11]–[13]. This important property however is only valid with respect to the global optimum of the *unconstrained* synthesis problem and not expected for general constrained case, let alone considering the disconnectedness of the respective feasible domains. By merely using first-order information on the cost function, Projected Gradient (PG) techniques—whenever feasible—can be shown to converge sublinearly to first-order stationary points of SLQR and OLQR problems [7], [12]. However, a sublinear rate is generally unfavorable–and theoretically unsatisfying–particularly when second-order information of the synthesis cost and the geometry of the stabilizing gains can be utilized.

The main contribution of this paper is characterizing the notion of *geometric stability certificate* required for developing feasible iterative algorithms for challenging constrained synthesis problems. In particular, this work develops the differential geometric machinery that can be utilized to characterize this certificate in the *absence* of a computationally feasible (global) retraction [14]–[16]. Moreover using this construct, a Newton-type algorithm is developed for linearly constrained synthesis problems with guaranteed local linear convergence rate–that eventually transitions to a quadratic one–even though the feasible domain of such problems may not be connected, and possibly include multiple local minima. Furthermore, we demonstrate how this unified approach applies to a wide range of constrained synthesis problems, including structured state-feedback SLQR and OLQR. The extended version of this work provides a more general differential geometric point of view for system design [1]. The aforementioned manuscript also includes a general extrinsic analysis of smooth synthesis costs over the Riemannian submanifolds of stabilizing controllers.

The rest of the paper is organized as follows. In §II, we introduce the problem setup and discuss its reformulation. In §III, we elaborate on the geometry of stabilizing controllers and the resulting second-order behavior of constrained LQR cost and study the geometry of SLQR and OLQR problems in this context. The proposed algorithm and convergence rate analysis are discussed in §IV. We provide numerical examples in §V, followed by concluding remarks in §VI.

II. Background and Problem Setup

Consider the discrete-time linear system,

$$\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k + B\boldsymbol{u}_k, \quad (1)$$

The authors are with the William E. Boeing Department of Aeronautics and Astronautics, University of Washington, Seattle, WA, USA. S. Talebi is also with the Department of Mathematics at the University of Washington. Emails: *shahriar@uw.edu* and *mesbahi@uw.edu*. The research of the authors has been supported by AFOSR grant FA9550-20-1-0053 and NSF grant ECCS-2149470.

*An extended version of this work is available online [1] and has been accepted for publication in *IEEE Transactions on Automatic Control*.

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are system parameters, i.e., $n \times n$ and $n \times m$ real matrices, $\boldsymbol{x}_k$ and $\boldsymbol{u}_k$ denote the state and input vectors, respectively, and $\boldsymbol{x}_0$ is given. Conventionally, the Linear Quadratic Regulators (LQR) problem is to design a sequence of inputs $\boldsymbol{u} = (\boldsymbol{u}_k)_0^\infty \in \ell_2$ (square summable sequences) that minimizes

$$J_{\boldsymbol{x}_0}(\boldsymbol{u}) = \frac{1}{2}\sum_{k=0}^{\infty} \boldsymbol{x}_k^\intercal Q \boldsymbol{x}_k + \boldsymbol{u}_k^\intercal R \boldsymbol{u}_k, \tag{2}$$

subject to (1), where Q and R are positive semidefinite and positive definite matrices, respectively. It is well known that the optimal solution to this problem reduces to solving the Discrete-time Algebraic Riccati Equation (DARE) for the unknown matrix P_{LQR} that quadratically parameterizes the so-called cost-to-go. Subsequently, one sets $\boldsymbol{u}_k^* = K_{\mathrm{LQR}}\boldsymbol{x}_k$, where the optimal LQR gain (policy) $K_{\mathrm{LQR}} \in \mathbb{R}^{m\times n}$ is given by $K_{\mathrm{LQR}} = -(R + B^\intercal P_{\mathrm{LQR}} B)^{-1} B^\intercal P_{\mathrm{LQR}} A$, and the optimal cost $J_{\boldsymbol{x}_0}(\boldsymbol{u}^*) = \boldsymbol{x}_0^\intercal P_{\mathrm{LQR}} \boldsymbol{x}_0/2$.

The PO approach to control design in contrast, starts with the feasible domain of the synthesis problem. First, we define the set of stabilizing controllers by

$$\mathcal{S} \coloneqq \{K \in \mathbb{R}^{m\times n} \mid A + BK \in \mathcal{M}\},$$

where $\mathcal{M}$ denotes a subset of (Schur) stable matrices $\mathcal{M} = \{A \in \mathbb{R}^{n\times n} \mid \rho(A) < 1\}$ and $\rho(\cdot)$ denotes the spectral radius. We also introduce a non-Euclidean geometry over $\mathcal{S}$ by a natural Riemannian metric arising in the context of LQR problem. Specifically, here we are interested in controllers that lie on a relatively "simple" subset $\mathcal{K}$ of $\mathbb{R}^{m\times n}$, such that $\mathcal{K} \cap \mathcal{S}$ is an embedded submanifold of $\mathcal{S}$. Common examples of these subsets are linear subspaces of $\mathbb{R}^{m\times n}$ that characterize a prescribed sparsity pattern in the optimal controller gain, or the output-feedback constraint [1]. Before we proceed with the main analysis, we restate the following lemma from [1] for completeness. Define the "*Lyapunov map*" $\mathbb{L} : \mathcal{M} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ that sends the pair (A, Q) to the solution Y of the following discrete Lyapunov equation,

$$Y = AYA^\intercal + Q. \tag{3}$$

We naturally identify the tangent space $T_{(A,Q)}(\mathcal{M} \times \mathbb{R}^{n\times n}) \cong \mathbb{R}^{n\times n} \oplus \mathbb{R}^{n\times n}$. [1]

Lemma 1 ([1, Lemma III.1])**.** *The subset $\mathcal{M}$ is an open submanifold of $\mathbb{R}^{n\times n}$ and the Lyapunov map $\mathbb{L} : \mathcal{M} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is smooth, and its differential acts as*

$$d\mathbb{L}_{(A,Q)}[E, F] = \mathbb{L}\big(A, E\,\mathbb{L}(A,Q)A^\intercal + A\,\mathbb{L}(A,Q)E^\intercal + F\big),$$

for any $E, F \in \mathbb{R}^{n\times n}$. Furthermore, for any $A \in \mathcal{M}$ and $Q, \Sigma \in \mathbb{R}^{n\times n}$, we have the so-called "Lyapunov-trace" property,

$$\mathrm{tr}\left[\mathbb{L}(A^\intercal, Q)\Sigma\right] = \mathrm{tr}\left[\mathbb{L}(A, \Sigma)Q\right].$$

One can view the LQR cost naturally as a map $K \mapsto J_{\boldsymbol{x}_0}(\boldsymbol{u} = K\boldsymbol{x})$; however, this would depend on $\boldsymbol{x}_0$ and generally, its value can be infinite as K is not necessarily stabilizing, i.e., when $K \notin \mathcal{S}$. This motivates studying the (natural) constrained optimization problem,

$$\text{minimize} \quad f(K) \coloneqq \mathop{\mathbb{E}}_{\boldsymbol{x}_0\sim\mathcal{D}} J_{\boldsymbol{x}_0}(K\boldsymbol{x}) \quad \text{over} \quad K \in \widetilde{\mathcal{S}} \tag{4}$$
$$\text{s.t.} \quad \boldsymbol{x}_{k+1} = (A + BK)\boldsymbol{x}_k, \quad \text{for all } k \geq 0,$$

where $\widetilde{\mathcal{S}}$ is an embedded submanifold of $\mathcal{S}$, and $\mathcal{D}$ denotes a distribution of zero-mean multivariate random variables of dimension n with a positive definite covariance Σ.

We further reformulate the problem in (4) as follows. For each stabilizing controller $K \in \mathcal{S}$, from (1) and (2) we have,

$$J_{\boldsymbol{x}_0}(\boldsymbol{u} = K\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}_0^\intercal \left[\sum_{k=0}^{\infty} (A_{\mathrm{cl}}^k)^\intercal [Q + K^\intercal R K] A_{\mathrm{cl}}^k\right] \boldsymbol{x}_0,$$

where $A_{\mathrm{cl}} \coloneqq A + BK$ is the closed-loop dynamics. Since A_{cl} is a Schur stable matrix, the sum converges to

$$P_K \coloneqq \mathbb{L}(A_{\mathrm{cl}}^\intercal, Q + K^\intercal R K),$$

noting that $A_{\mathrm{cl}} \in \mathcal{M}$ implies $A_{\mathrm{cl}}^\intercal \in \mathcal{M}$. Therefore, for each $K \in \mathcal{S}$, we can compute $f(K)$ as,

$$f(K) = \tfrac{1}{2}\,\mathbb{E}_{\boldsymbol{x}_0\sim\mathcal{D}}\,\mathrm{tr}\left[P_K \boldsymbol{x}_0 \boldsymbol{x}_0^\intercal\right] = \tfrac{1}{2}\mathrm{tr}\left[P_K \Sigma\right].$$

Hence (4) reduces to,

$$\min_K f(K) = \frac{1}{2}\mathrm{tr}\left[P_K\Sigma\right] \quad \text{such that} \quad K \in \widetilde{\mathcal{S}}. \tag{5}$$

If there are no additional constraints on the stabilizing controllers, i.e., $\widetilde{\mathcal{S}} = \mathcal{S}$, then a well-known quasi-Newton algorithm due to Hewer is known to converge to the global optimum (under the usual control-theoretic assumptions) at a Q-quadratic rate [17]. This elegant algorithm has the above favorable convergence properties as, 1) the optimum point of the optimization problem coincides with the stationary point of the gradient field of f with respect to an intrinsic Riemannian metric, 2) the positive definite estimates of the "Hessian" of f are accurate representation of the quadratic behavior of the cost function near the optimum, and 3) the unit step-size–remarkably–keeps the iterates inside $\mathcal{S}$ throughout the updates.

III. Geometry of SLQR and OLQR problems

The motivation behind the geometric analysis of stability certificates is to develop an analogous iterative procedure to that of Hewer's, but for the cost function f restricted to $\widetilde{\mathcal{S}}$. In order to accomplish this, we need to investigate the following questions: 1) What is the "useful" geometry on this constrained submanifold? 2) What is the gradient field and the accurate estimate of the Hessian operator on this submanifold? and finally 3) What choice of step-size keeps the iterates stabilizing throughout the algorithm?

Before we proceed, it is worth noting that if we were to use the machinery developed for optimization over manifolds [14], it would be necessary to access a retraction from the tangent bundle $T\widetilde{\mathcal{S}}$ onto $\widetilde{\mathcal{S}}$. Unfortunately, such a mapping is not available in general, due to the intricate geometry of $\mathcal{S}$; however, if $\widetilde{\mathcal{S}}$ is endowed with a linear

[1] The reader is referred to [1] for a more systematic treatment of referenced mathematical constructs and notation.

structure, then one can circumvent this issue by utilizing the tangential projection (with respect to the geometry) onto $T\widetilde{\mathcal{S}}$, a more accessible construct in general.

A. Domain Manifold

In this paper, we focus on $\mathcal{S}$ as a manifold on its own.[2] It is known that $\mathcal{S}$ is contractible, and unbounded when $m \geq 2$ [18]. Furthermore, $\mathcal{S}$ is open in $\mathbb{R}^{m\times n}$ (a consequence of continuity of maximum eigenvalue of a parameterized matrix with smooth entries), and therefore a submanifold without boundary. On the other hand, the constrained submanifold $\widetilde{\mathcal{S}}$ does not even need to be connected and might have exponential number of connected components; first-order stationary points are also not necessarily local minima.

Notation. For manifolds, we follow the notation and results in [19] and [20] unless stated otherwise. At each point $K \in \mathcal{S}$, we identify the tangent space $T_K\mathcal{S}$ canonically with $\mathbb{R}^{m\times n}$. Also, since $\mathcal{S}$ can be covered by a single smooth chart, the tangent bundle of $\mathcal{S}$, denoted by $T\mathcal{S}$, is diffeomorphic to $\mathcal{S} \times \mathbb{R}^{m\times n}$. We refer to this identification as "*the usual identification of the tangent bundle*" (or $T_K\mathcal{S} \cong \mathbb{R}^{m\times n}$ at any point $K \in \mathcal{S}$) if we need to identify an element of $T\mathcal{S}$ (or $T_K\mathcal{S}$). In particular, let us denote the coordinates of the global chart by $(x^{i,j})$, its associated global coordinate frame by $(\frac{\partial}{\partial x^{i,j}})$ or simply $(\partial_{i,j})$, and its dual coframe by $(dx^{i,j})$, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$. We also use the Einstein summation convention as explained in [20] for double indices as, for example, $x^{i,j}\partial_{i,j}$ denotes $\sum_{i=1}^{m}\sum_{j=1}^{n} x^{i,j}\partial_{i,j}$. Lastly, the (i,j)-th element of any matrix $A \in \mathbb{R}^{m\times n}$ is denoted by $[A]_{i,j}$ or $[A]^{i,j}$ depending on viewing A as a point or a tangent vector, respectively. The set of smooth functions on $\mathcal{S}$ is denoted by $C^\infty(\mathcal{S})$. A vector field V on $\mathcal{S}$ is a smooth map $V : \mathcal{S} \to T\mathcal{S}$, usually written as $K \to V_K$, with the property that $V_K \in T_K\mathcal{S}$ for all $K \in \mathcal{S}$. We denote the set of all vector fields over $\mathcal{S}$ by $\mathfrak{X}(\mathcal{S})$. A covariant 2-tensor field is a smooth real-valued multilinear function of two vector fields. Finally, for any general mapping $P : \mathcal{S} \to \star$, we use $P_K, P|_K$ or $P(K)$ to denote the element in $\star$ for which $K \in \mathcal{S}$ has been mapped to. For reasons that become apparent subsequently, in what follows, we define a smoothly varying bilinear function (i.e., a covariant 2-tensor field) on $\mathcal{S}$ which turns out to be a Riemannian metric. Define the map $\langle .,. \rangle_Y : \mathfrak{X}(\mathcal{S}) \times \mathfrak{X}(\mathcal{S}) \to C^\infty(\mathcal{S})$ as follows:[3] for any $V, W \in \mathfrak{X}(\mathcal{S})$,

$$\langle V, W \rangle_Y \big|_K = \operatorname{tr}\left[(V_K)^\intercal \, W_K \, Y_K\right], \quad \forall K \in \mathcal{S}, \tag{6}$$

where $Y_K = \mathbb{L}(A_{\rm cl}, \Sigma) \in \mathbb{R}^{n\times n}$ satisfying

$$Y_K = A_{\rm cl} Y_K A_{\rm cl}^\intercal + \Sigma. \tag{7}$$

Lemma 2. *The map $\langle .,. \rangle_Y : \mathfrak{X}(\mathcal{S}) \times \mathfrak{X}(\mathcal{S}) \to C^\infty(\mathcal{S})$ as in (6) is well-defined and induced by a Riemannian metric g on $\mathcal{S}$. Moreover, with respect to the dual coframe $(dx^{i,j})$, $g = g_{(i,j)(k,\ell)} dx^{i,j} \otimes dx^{k,\ell}$ with $g_{(i,j)(k,\ell)} \in C^\infty(\mathcal{S})$ satisfying*

$$g_{(i,j)(k,\ell)}(K) = \begin{cases} [Y_K]_{\ell,j} & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Note that $\Sigma \succ 0$ in (4); thus, $(A_{\rm cl}, \Sigma^{1/2})$ is controllable and the claim follows by [1, Proposition III.3] with $\Sigma_1 = \Sigma$ and $\Sigma_2 = 0$. ☐

Now, we consider the Riemannian manifold $(\mathcal{S}, g, \nabla)$ where ∇ denotes the associated Riemannian connection (see [1] for computing its associated Christoffel symbols $\Gamma^{i,j}_{(k,\ell)(p,q)}$). This induces a unique geometry on $\widetilde{\mathcal{S}}$ as a Riemannian submanifold $(\widetilde{\mathcal{S}}, \widetilde{g}, \widetilde{\nabla})$, with $\widetilde{g}$ and $\widetilde{\nabla}$ denoting the induced Riemannian metric and connection. Considering any smooth function f on $\mathcal{S}$, recall the gradient of f with respect to the Riemannian metric g, denoted by $\operatorname{grad} f \in \mathfrak{X}(\mathcal{S})$, is the unique vector field satisfying

$$\langle V, \operatorname{grad} f \rangle_Y = V f,$$

for any $V \in \mathfrak{X}(\mathcal{S})$.[4] Also, define the *"Hessian operator"* of $f \in C^\infty(\mathcal{S})$ as the map $\operatorname{Hess} f : \mathfrak{X}(\mathcal{S}) \to \mathfrak{X}(\mathcal{S})$ defined by

$$\operatorname{Hess} f[U] \coloneqq \nabla_U \operatorname{grad} f,$$

for any $U \in \mathfrak{X}(\mathcal{S})$.[5] Note that we use the same notation to denote the gradient and Hessian operators defined on the submanifold $\widetilde{\mathcal{S}}$, even though they are different objects; for further discussions regarding Hessian operator refer to [1]. Finally, we are interested in the restriction $h \coloneqq f|_{\widetilde{\mathcal{S}}}$, and relating its gradient and Hessian to those of f.

B. Extrinsic Analysis of the LQR Cost

Herein, we aim to apply the general extrinsic analysis proved in [1, Proposition III.5] to the LQR cost and its restriction to any embedded submanifold of $\mathcal{S}$ (including the ones induced by constrained $\mathcal{K}$). Before applying this abstract extrinsic analysis, we need the geometric notion of Riemannian projection [20]. Denote the (Riemannian) *tangential* and *normal projections* by $\pi^\top : T\mathcal{S}|_{\widetilde{\mathcal{S}}} \to T\widetilde{\mathcal{S}}$ and $\pi^\perp : T\mathcal{S}|_{\widetilde{\mathcal{S}}} \to N\widetilde{\mathcal{S}}$, respectively, with $N\widetilde{\mathcal{S}}$ indicating the normal bundle of $\widetilde{\mathcal{S}}$.

Proposition 1. *Let $f(K) = \frac{1}{2}\operatorname{tr}[P_K \Sigma]$ with $P_K = \mathbb{L}(A_{\rm cl}^\intercal, K^\intercal R K + Q)$. Define $h = f|_{\widetilde{\mathcal{S}}}$ where $\widetilde{\mathcal{S}} \subset \mathcal{S}$ is an embedded Riemannian submanifold with the induced metric and connection. Then, h is smooth and under the usual identification of tangent bundle,*

$$\operatorname{grad} h_K = \pi^\top(RK + B^\intercal P_K A_{\rm cl}).$$

[2] This statement, however, necessitates elaborating on what this approach would entail, justifying our brief exposition of the required differential geometry in §III-A; the reader can refer to this section as needed as we are obliged to provide the key constructs for our subsequent analysis.

[3] The notation $\langle \cdot, \cdot \rangle_Y$ should not be confused with the (ordinary) inner product in inner-product spaces as it is varying over $\mathcal{S}$.

[4] Note that the action of a vector field V on a smooth function f results in a smooth function $Vf \in C^\infty(\mathcal{S})$ where its value at K, i.e., $Vf|_K$, can be viewed as the directional derivative f at K in the direction of $V_K \in T_K\mathcal{S}$.

[5] Note that $\nabla_U \operatorname{grad} f$ is itself a vector field that refers to the covariant derivative of the vector field $\operatorname{grad} f$ in the direction of the vector field U. See [1] and [20, Chp. 4] for more on covariant derivative, its computation and relation to the Christoffel symbols $\Gamma^{i,j}_{(k,\ell)(p,q)}$.

Furthermore, $\operatorname{Hess} h$ *is a self-adjoint operator and can be characterized as follows: for any* $E, F \in T_K\widetilde{\mathcal{S}} \subset T_K\mathcal{S}$,

$$\begin{aligned}\langle \operatorname{Hess} h[E], F\rangle_{Y_K} &= \langle B^\intercal (S_K|_F) A_{\rm cl}, E\rangle_{Y_K} \\ &+ \langle (R + B^\intercal P_K B)E + B^\intercal (S_K|_E) A_{\rm cl}, F\rangle_{Y_K} \\ &- \left\langle \operatorname{grad} h_K, [E]^{k,\ell}[F]^{p,q}\Gamma^{i,j}_{(k,\ell)(p,q)}(K)\partial_{i,j}\right\rangle_{Y_K},\end{aligned}$$

where $\Gamma^{i,j}_{(k,\ell)(p,q)}(K)$ *denotes the Christoffel symbols associated with* g *and*

$$S_K|_E \coloneqq \mathbb{L}(A_{\rm cl}^\intercal, E^\intercal \operatorname{grad} f_K + (\operatorname{grad} f_K)^\intercal E).$$

Remark 1. The second-order behavior of the constrained cost h is captured in the $\operatorname{Hess} h$ operator above which is obtained using the unique Riemannian connection. On the other hand, one could in principle think of capturing this behavior using the so-called Euclidean connection—that corresponds to the connection on $(\mathcal{S}, g)$ with all zero connection coefficient $\Gamma^{i,j}_{(k,\ell)(p,q)}$ in the global coordinate frame. We denote the latter operator by $\overline{\operatorname{Hess}}\, h$ which will behave quite differently than the former one, as will be illustrated through examples. Finally, one could instead incorporate a positive definite estimate of $\operatorname{Hess} h$ which can result in a quasi-Newton variant of the proposed algorithm.

C. Newton Direction on Constrained Stabilizing Policies

In the upcoming sections, we aim to characterize a descent direction using the second-order behavior of the constrained cost h. In particular, we refer to the solution $G \in T_K\widetilde{\mathcal{S}}$ of the following equation as the *Newton direction* on $\widetilde{\mathcal{S}}$:

$$\operatorname{Hess} h_K[G] = -\operatorname{grad} h_K,$$

where $h = f|_{\widetilde{\mathcal{S}}}$; similarly, the analogous construct is referred to as the *Euclidean Newton direction* if $\operatorname{Hess} h$ is replaced by $\overline{\operatorname{Hess}}\, h$. Next, we discuss how SLQR and OLQR problems can be treated similarly using this approach, each corresponding to a different Riemannian submanifold $\widetilde{\mathcal{S}}$ of $\mathcal{S}$. In order to solve for the Newton direction at any $K \in \widetilde{\mathcal{S}}$, suppose that the tuple $(\widetilde{\partial}_{(p,q)}|_{(p,q)\in D})$ denotes a smooth local frame for $\widetilde{\mathcal{S}}$ on a neighborhood of K, where D is a subset of $[m] \times [n]$ depending on the dimension of $\widetilde{\mathcal{S}}$.[6] In fact, for any $G = [G]^{k,\ell}\widetilde{\partial}_{(k,\ell)}|_K \in T_K\widetilde{\mathcal{S}}$ (interpreted as a subspace of $T_K\mathcal{S}$), finding the Newton direction on $\widetilde{\mathcal{S}}$ reduces to solving the following system of $|D|$-linear equations (for each index $(p,q) \in D$):

$$\textstyle\sum_{(k,\ell)\in D}[G]^{k,\ell}\, h_{;(k,\ell)(p,q)}(K) = - \left\langle \pi^\top(\operatorname{grad} f|_K), \widetilde{\partial}_{(p,q)}|_K\right\rangle_{Y_K},$$

with the coefficient $h_{;(k,\ell)(p,q)}$ computed as $h_{;(k,\ell)(p,q)}(K) = \left\langle \operatorname{Hess} h_K[\widetilde{\partial}_{(k,\ell)}|_K], \widetilde{\partial}_{(p,q)}|_K\right\rangle_{Y_K}$; or with $\operatorname{Hess} h$ replaced by $\overline{\operatorname{Hess}}\, h$, depending on the choice of the connection.

[6] Each $T_K\widetilde{\mathcal{S}}$ can be viewed as a subspace of $T_K\mathcal{S}$ as $\widetilde{\mathcal{S}} \subset \mathcal{S}$ is embedded; and, the specific choice of the local frame for $\widetilde{\mathcal{S}}$ depends on the application.

D. State-feedback SLQR

Any desired sparsity pattern on the policy, i.e., controller gain K, leads to a linearly constrained set, denoted by $\mathcal{K}_D$, indicating a linear subspace of $\mathbb{R}^{m\times n}$ with nonzero entries only for a prescribed subset D of entries, i.e.,

$$\mathcal{K}_D \coloneqq \left\{K \in \mathbb{R}^{m\times n} \mid [K]_{i,j} = 0 \text{ whenever } (i,j) \notin D\right\}.$$

One can show that $\widetilde{\mathcal{S}} = \mathcal{S} \cap \mathcal{K}_D$ is a properly embedded submanifold of $\mathcal{S}$ with dimension $|D|$. Furthermore, at any point $K \in \widetilde{\mathcal{S}}$ and for any tangent vector $E \in T_K\mathcal{S}$, the tangential projection $\pi^\top : T_K\mathcal{S} \to T_K\widetilde{\mathcal{S}}$ sends E to $\widetilde{E}$ where it must satisfy $E - \widetilde{E} \perp \mathcal{K}_D$ with respect to the Riemannian metric at K; or equivalently,

$$\operatorname{Proj}_{\mathcal{K}_D}\left[(E - \widetilde{E})Y_K\right] = 0,$$

with $\operatorname{Proj}_{\mathcal{K}_D}$ denoting the Euclidean projection onto the sparsity pattern $\mathcal{K}_D$. Note that at each $K \in \widetilde{\mathcal{S}}$, the last equality consists of $|D|$ nontrivial linear equations involving $|D|$ unknowns (as the nonzero entries of $\widetilde{E}$).

Finally, if $\widetilde{\partial}_{(i,j)}$ (as described in §III-C) is taken to be $\widetilde{\partial}_{(i,j)} = \partial_{(i,j)}$ for $(i,j) \in D$, then $(\widetilde{\partial}_{(i,j)}|_{(i,j)\in D})$ forms a global smooth frame for $T\widetilde{\mathcal{S}}$. Thus, for each $(k,\ell), (p,q) \in D$, the coordinates $h_{;(k,\ell)(p,q)}(K)$ simplifies to

$$\begin{aligned}h_{;(k,\ell)(p,q)}(K) &= \left\langle B^\intercal (S_K|_{\partial_{(p,q)}}) A_{\rm cl}, \partial_{(k,\ell)}\right\rangle_{Y_K} \\ &+ \left\langle (R + B^\intercal P_K B)\partial_{(k,\ell)} + B^\intercal (S_K|_{\partial_{(k,\ell)}}) A_{\rm cl}, \partial_{(p,q)}\right\rangle_{Y_K} \\ &- \left\langle \pi^\top \operatorname{grad} f_K, \Gamma^{i,j}_{(k,\ell)(p,q)}(K)\partial_{i,j}\right\rangle_{Y_K}.\end{aligned}$$

E. Output-feedback LQR

The OLQR problem can be formulated as the optimization problem in (5) with the submanifold $\widetilde{\mathcal{S}} = \mathcal{S} \cap \mathcal{K}_C$ with the constraint set $\mathcal{K}_C$ defined as

$$\mathcal{K}_C \coloneqq \left\{K \in \mathbb{R}^{m\times n} \mid K = LC,\ L \in \mathbb{R}^{m\times d}\right\},$$

where $C \in \mathbb{R}^{d\times n}$ is the prescribed output matrix. Note that $\mathcal{K}_C$ is a linear subspace of $\mathbb{R}^{m\times n}$ whose dimension depends on the rank of C. For simplicity of presentation, we suppose C has full rank equal to $d \leq n$. Then, one can show that $\widetilde{\mathcal{S}}$ is a properly embedded submanifold of $\mathcal{S}$ with dimension md. Furthermore, at each $K \in \widetilde{\mathcal{S}}$, we can canonically identify the tangent space at K with $T_K\widetilde{\mathcal{S}} \cong \mathcal{K}_C$.

Similar to the SLQR case, at any $K \in \widetilde{\mathcal{S}}$ and for any $E \in T_K\mathcal{S}$, the tangential projection of E, denoted by $\widetilde{E} = \pi^\top E$, must satisfy $E - \widetilde{E} \perp \mathcal{K}_C$ (with respect to the Riemannian metric), or equivalently,

$$\operatorname{tr}\left[C^\intercal L^\intercal (E - \widetilde{E})Y_K\right] = 0, \quad \forall L \in \mathbb{R}^{m\times d}.$$

Since C is assumed to be full-rank, we conclude that $\pi^\top E = L^* C$ where $L^* \in \mathbb{R}^{m\times d}$ is the unique solution of the following linear equation

$$L^* C Y_K C^\top = E Y_K C^\top.$$

Finally, let $D \coloneqq [m] \times [d]$ (the set of ordered pairs each from the corresponding index set) and consider the identification $T_K\widetilde{\mathcal{S}} \cong \mathcal{K}_C$. Then, at each point $K \in \widetilde{\mathcal{S}}$, we can choose $\widetilde{\partial}_{(i,j)} = \partial_{(i,j)}C$ with $(i,j) \in D$ as a global smooth frame for $T\widetilde{\mathcal{S}}$. But then, the coordinates of the covariant Hessian $h_{;(k,\ell)(p,q)}(K)$ with respect to this frame can be computed by substituting $E = \partial_{(k,\ell)}C$ and $F = \partial_{(p,q)}C$ in Proposition 1 for each $(k,\ell),(p,q) \in D$—similar to the SLQR case. It is worth noting that the sparsity pattern in E, F and Christoffel symbols can simplify the computations; we will not further delve into such considerations for brevity.

IV. Newton-type algorithm

We first show how the perspective pursued in this work is related to the well-known Hewer algorithm [17]; this was first pointed out in [12]. Consider the Hewer's algorithm updating the stabilizing controller as,

$$K^+ = -(R + B^\intercal P_K B)^{-1} B^\intercal P_K A,$$

that can be reformulated as,

$$\begin{aligned} K^+ &= K - (R + B^\intercal P_K B)^{-1}((R + B^\intercal P_K B)K + B^\intercal P_K A) \\ &= K - (R + B^\intercal P_K B)^{-1}(RK + B^\intercal P_K A_{\mathrm{cl}}). \end{aligned}$$

The algorithm is thereby equivalent to,

$$K^+ = K + G,$$

where G is the quasi-Newton direction obtained by solving,

$$(R + B^\intercal P_K B)G = -\operatorname{grad} f_K.$$

One can in fact argue that $(R+B^\intercal P_K B)$ is a "good" positive definite approximation of the Riemannian Hessian operator $\operatorname{Hess} f_K$ of the (unstructured) LQR problem! The reason becomes clear if we view the LQR cost as discussed in Proposition 1 with no constraints (i.e., when $\widetilde{\mathcal{S}} = \mathcal{S}$). In particular, at optimality, we have $\operatorname{grad} f|_{K_{\mathrm{LQR}}} = RK_{\mathrm{LQR}} + B^\intercal P_{K_{\mathrm{LQR}}} A_{\mathrm{cl}} = 0$ and thus by Proposition 1, we have $S_{K_{\mathrm{LQR}}}|_E = S_{K_{\mathrm{LQR}}}|_F = 0$ for any vector E and F. Therefore, the Hessian operator as in Proposition 1 reduces to,

$$\langle \operatorname{Hess} f_K[E], F\rangle_{Y_{K_{\mathrm{LQR}}}} = \left\langle (R + B^\intercal P_{K_{\mathrm{LQR}}} B)E, F\right\rangle_{Y_{K_{\mathrm{LQR}}}},$$

for any tangent vectors E, F. The connection between Hewer's algorithm as a Riemannian quasi-Newton method for the unconstrained LQR problem is now clear. The rest of this section illustrates how we can extend this perspective to propose a Newton-type algorithm for *constrained* LQR with convergence guarantees. We claim that, starting close enough to a local minimum (when one exits), a Newton-type method using Riemannian metric and the Euclidean/Riemannian connection must converge quadratically if one could have used stepsize $\eta = 1$. This is in fact due to the exponential mapping with respect to the Euclidean connection that serves as a retraction with the desirable properties. However, unlike the case with Hewer's algorithm, the stability constraint suggests that at least away from the local minimum, it might not be possible to use such a large stepsize. Therefore, a stepsize rule has to be deduced—that in turn, hinges upon the notion of stability certificate; since deriving this certiciate is intimately connected with the Riemannian geometry of the synthesis problem, we refer to it as the *geometric stability certificate*.

A. Geometric Stability Certificate and Step-size Selection

Our road map for choosing a step-size is to come up with a stability certificate s_t at each iteration t that can be combined with the geometric tangential projection as described in §III-B. Later on, we show how to use this certificate to choose step-sizes in practice with theoretical guarantees, and more importantly, how it ensures existence of neighborhoods each containing a local minimum on which our algorithm achieves a quadratic rate of convergence.

For brevity and simplifying the presentation, here we assume that $Q \succ 0$; when $Q \succeq 0$, one can leverage the observability of the pair $(A, Q^{1/2})$ to obtain analogous results. The following lemma can be used for analyzing generic iterative PO algorithms for structured and unstructured synthesis with stability requirements.

Lemma 3. *Given any direction $G \in T_K\mathcal{S} \cong \mathbb{R}^{m\times n}$, if $K \in \mathcal{S}$ and step-size η is such that,*

$$0 \le \eta \le s_K := \underline{\lambda}(K^\intercal RK + Q)\big/(2\,\overline{\lambda}(P_K)\|BG\|_2),$$

then $K^+ = K + \eta G \in \mathcal{S}$. Furthermore,

$$s_K \ge \underline{\lambda}(Q)\,\underline{\lambda}(\Sigma)\big/(4f(K)\|BG\|_2).$$

Here, $\|.\|_2$ refers to the spectral norm, and $\overline{\lambda}$ (or $\underline{\lambda}$) denotes the maximum (or minimum) eigenvalue of a matrix. We refer to s_K as the geometric "*stability certificate at K*".

Proof. We can choose a mapping $\mathcal{Q} : K \to \mathcal{Q}_K$ to be $\mathcal{Q}_K = Q + K^\top RK \succ 0$ as $Q \succ 0$. Then, the stability certificate s_K as defined in Lemma 3 follows using [1, Lemma IV.1]. Next, since also $R \succ 0$, $f(K) \ge \frac{1}{2}\,\underline{\lambda}(\Sigma)\,\overline{\lambda}(P_K)$, where we have used the trace inequality as both $P_K, \Sigma \succ 0$. The claimed lower-bound on the stability certificate now follows. ☐

Algorithm 1: Riemannian Newton-type Policy Optimization (RNPO) for Constrained Problems on $\mathcal{S}$

1: **Initialization:** Problem parameters (A, B, Q, R), the linear constraint $\mathcal{K}$ and an initial feasible stabilizing controller $K_0 \in \widetilde{\mathcal{S}} = \mathcal{S} \cap \mathcal{K}$; set $t \leftarrow 0$
2: **Until stopping criteria are met, do**
3: Find the Newton direction G_t on $\widetilde{\mathcal{S}}$ satisfying

$$\operatorname{Hess} h_{K_t}[G_t] = -\operatorname{grad} h_{K_t}$$

4: Obtain a stability certificate s_{K_t} from Lemma 3
5: Compute the step-size $\eta_t \leftarrow \min\{s_{K_t}, 1\}$
6: Update $K_{t+1} \leftarrow K_t + \eta_t G_t$
7: $t \leftarrow t + 1$

In Line 4, $\operatorname{Hess} h$ can be replaced by its Euclidean counterpart $\overline{\operatorname{Hess}}\, h$.

Remark 2 (Conditioning and Choice of Mapping $\mathcal{Q}$)**.** In general, the convergence rate of RNPO is directly related to the stability certificate s_K that depends on the "conditioning" of the Lyapunov mapping at $(A_{\mathrm{cl}}^\intercal, \mathcal{Q}_K)$), defined as,

$$\kappa_K := \overline{\lambda}(\,\mathbb{L}(A_{\mathrm{cl}}^\intercal, \mathcal{Q}_K))\big/\underline{\lambda}(\mathcal{Q}_K),$$

for any $\mathcal{Q}_K \succ 0$. This may affect the algorithm performance by constant factors, and also the regions on which the convergence rate remains linear. The natural choice of the mapping $\mathcal{Q}$ is $K \to Q + K^\intercal RK$ relates this conditioning to the value of $f(K)$ as shown in Lemma 3.

Remark 3 (Stopping Criteria)**.** Given $\epsilon > 0$, a simple choice for stopping criteria may be $\langle \operatorname{grad} h_{K_t}, \operatorname{grad} h_{K_t} \rangle_{Y_{K_t}} < \epsilon$. Furthermore, one may derive a more elaborate stopping criteria based on the sample complexity of the algorithm and its linear-quadratic convergence rate as follows:

Stopping criteria: Initialize $i = 0$ at $t = 0$ and check $(\forall t > 0)$:

```
If η_t < 1 and t < O(ln(1/ϵ))
    i ← i + 1
    Continue
else-if η_t = 1 and t − i < O(ln ln(1/ϵ))
    Continue
else
    Stop
```

The big O notation hides constant factors that depend on the local *condition number* and *variation bound* of $\operatorname{Hess} h$.[7]

Next, we present the local linear-quadratic convergence of RNPO algorithm on the submanifold $\widetilde{\mathcal{S}}$; the complete proofs are deferred to [1] which treats a more general setup. To proceed, we say that K^* is a critical point of h if $\operatorname{grad} h_{K^*} = 0$, and additionally, it is "nondegenerate" if $\operatorname{Hess} h_{K^*}$ is nondegenerate, i.e.,

$$\langle \operatorname{Hess} h_{K^*}[G_1], G_2 \rangle_{Y_{K^*}} = 0, \forall G_2 \in T_{K^*}\widetilde{\mathcal{S}} \implies G_1 = 0.$$

Lemma 4. *Suppose K^* is a nondegenerate local minimum of $h := f|_{\widetilde{\mathcal{S}}}$. Then, it is isolated,* $\operatorname{grad} h_{K^*} = 0$ *and there exists a neighborhood of K^* on which* $\operatorname{Hess} h$ *is positive definite. Furthermore,* $\operatorname{Hess} h_{K^*} = \overline{\operatorname{Hess}}\, h_{K^*}$.

Theorem 1. *Suppose K^* is a nondegenerate local minimum of $h := f|_{\widetilde{\mathcal{S}}}$ over the submanifold $\widetilde{\mathcal{S}} = \mathcal{S} \cap \mathcal{K}$ for some linear constraint $\mathcal{K}$. Then, there exists a neighborhood $\mathcal{U}^* \subset \widetilde{\mathcal{S}}$ of K^* with the following property: whenever $K_0 \in \mathcal{U}^*$, the sequence $\{K_t\}$ generated by RNPO remains in $\mathcal{U}^*$ (and therefore, stabilizing), and it converges to K^* at least at a linear rate– and eventually–with a quadratic one.*

Remark 4 (Basin of attraction)**.** The proposed algorithm is guaranteed to converge in a neighborhood $\mathcal{U}^*$ of each nondegenerate local minimum where the size of this neighborhood indeed depends on the local *condition number* and *variation bound* of $\operatorname{Hess} h$.[7]

V. Numerical Examples

We now provide examples for optimizing the LQR cost over submanifolds induced by SLQR and OLQR.

Example 1: First, we consider a system with $n = 6$ number of states and $m = 3$ number of inputs, and simulate the behavior of RNPO and PG for 100 randomly sampled system parameters. We adopt a sampling approach where parameters (A, B) are sampled from a zero-mean unit-variance normal distribution, and the system matrix A is scaled so that the open-loop system is stable, i.e., $K_0 = 0$ is stabilizing, and the pair is controllable. This provides consistency in the choice of the initial controller as well as bypassing finding the initial stabilizing controller that is not the subject of the present contribution. Furthermore, we choose $Q = \Sigma = I_n$ and $R = I_m$ in order to consistently compare the convergence behaviors across different samples as $\underline{\lambda}(Q)$ and $\underline{\lambda}(\Sigma)$ may affect the constants in the convergence bound. For each problem, we have simulated three different algorithms; the first two are variants of Algorithm 1, where we use Riemannian connection or Euclidean connection to compute $\operatorname{Hess} h$ or $\overline{\operatorname{Hess}}\, h$, respectively. Note that we expect $\operatorname{Hess} h$ and $\overline{\operatorname{Hess}}\, h$ to have distinct information on neighborhoods of isolated local minima that directly influences the performance of RNPO as will be discussed below. The third algorithm is the Projected Gradient (PG) as studied in [12]. PG is feasible for constraints in our examples as, under relevant assumptions, one is able to perform PG updates by having access to merely the projection onto linear subspace of matrices—see for example [12, Theorem 7.1]. Here, the step size for PG is a constant value that guarantees the iterates stay stabilizing as suggested therein. For the SLQR problem, we randomly sample for the sparsity pattern D so that at least half of the entries are zero and the algorithm has converged from $K_0 = 0$ for all instances in less than 30 iterations. For the OLQR problem, we also randomly sample the output matrix C with $d = 2$, where 98% and 92% of the corresponding iterates have converged from $K_0 = 0$ in less than 50 iterations using Hess and $\overline{\operatorname{Hess}}$, respectively. The minimum, maximum and median progress of the three algorithms for both SLQR and OLQR problems are illustrated in Figure 1a and Figure 1b, respectively. As guaranteed by Theorem 1, the linear-quadratic convergence behavior of RNPO is observed in these problems. In both cases, Algorithm 1 (blue curves) built upon the Riemannian connection has a superior convergence compared with the case of using the Euclidean connection (orange curves); this is expected as the Riemannian connection is compatible with the metric induced by the geometry inherent to the cost function itself.

Example 2: Next, we consider optimal control of diffusion dynamics on random networks. In particular, we consider a random-3-regular graph $\mathcal{G}$ on 14 nodes (Figure 1c), containing two specific control nodes (also chosen at random). While the rest of the network has adopted consensus dynamics, the control nodes act as "leaders" where we can measure their states and inject control signals in order to derive the dynamics of the entire network. We consider the Laplacian $L(G)$ with node labels such that the control nodes appear as the last two label indices. Next, we partition $L(G) = \begin{pmatrix} A_f & B_f \\ B_f^\intercal & * \end{pmatrix}$, where $*$ hides a 2-by-2 matrix. Then, the dynamics of the scalar states of the nodes 1 through

[7] See the proof of Theorem 1 in [1] for a specific description of the local condition number and variation bound appearing as the ratios of M/m and L/m, respectively.

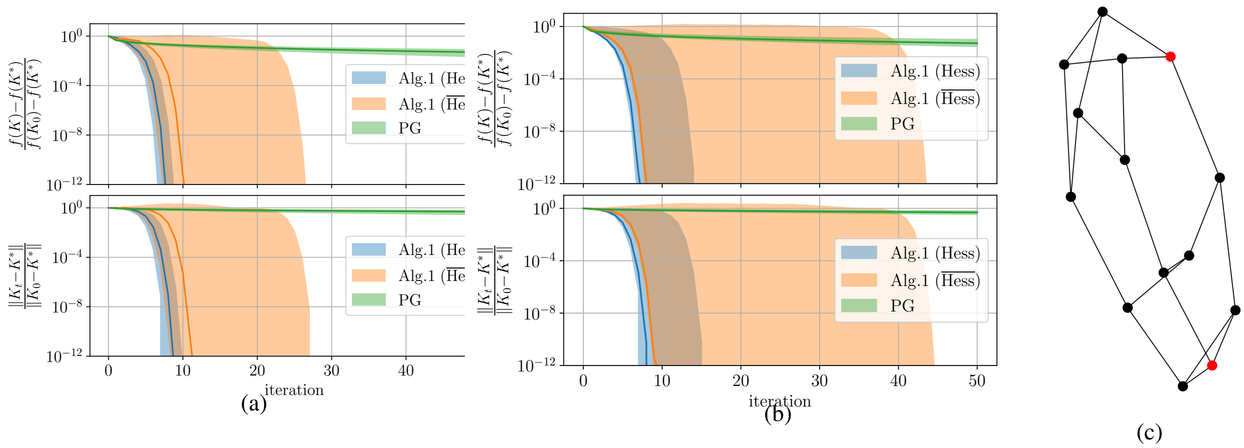

Fig. 1: The min, max and median progress of normalized error of iterates and cost values at each iteration of Algorithm 1, for the (a) SLQR and (b) OLQR with 100 different randomly sampled system parameters, sparsity patterns and output matrices. (c) The random-3-regular graph in Example 2 with the two control nodes in red.

12 are governed by the continuous-time LTI dynamics with system parameters $A = -A_f$, $B = -B_f$, and $C = -B_f^\intercal$; see [21, Chapter 10] discussing this network control problem. Finally, we discretize the continuous dynamics with 0.1 time-step and consider the OLQR problem with the rest of the problem parameters Q, R, Σ as set in the first simulation. Regulating the entire network is possible with the natural choice of zero input (i.e., with $K_0 = 0 \in \mathcal{K}_C \cap \mathcal{S}$ as $-A_f$ is Schur stable) which amounts to the LQR cost of 22.02 units. On the other hand, using RNPO, the locally optimal output-feedback LQR policy is,

$$L^* = \begin{pmatrix} -0.23486552 & 0.01978442 \\ 0.02255697 & -0.25106536 \end{pmatrix},$$

incurring 17.41 units of cost, a 26% improvement.

VI. Conclusions

We considered minimizing the LQR cost over submanifolds of stabilizing policies. In order to treat this problem in a natural geometric framework, we studied the first and second-order behavior of the synthesis cost when constrained to an embedded submanifold of stabilizing policies. Then, by leveraging on the second-order behavior of the restricted LQR cost and the developed geometric stability certificate, we proposed a Riemannian Newton-type algorithm with guaranteed convergence to a local minima at a linear-quadratic rate. The proposed algorithm was tailored to linear constraints on stabilizing controllers arising from SLQR and OLQR; however, the machinery is rather general and can handle more elaborate submanifolds in PO problems.

References

[1] S. Talebi and M. Mesbahi, "Policy optimization over submanifolds for constrained feedback synthesis," *IEEE Transactions on Automatic Control (accepted), arXiv preprint arXiv:2201.11157*, 2022.

[2] H. Mohammadi, A. Zare, M. Soltanolkotabi, and M. R. Jovanović, "Convergence and sample complexity of gradient methods for the model-free linear–quadratic regulator problem," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2435–2450, 2021.

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

[4] H. Feng and J. Lavaei, "On the exponential number of connected components for the feasible set of optimal decentralized control problems," in *2019 American Control Conference (ACC)*, pp. 1430–1437, 2019.

[5] H. T. Toivonen and P. M. Makila, "Newton's method for solving parametric linear quadratic control problems," *International Journal of Control*, vol. 46, no. 3, pp. 897–911, 1987.

[6] P. M. Makila and H. T. Toivonen, "Computational methods for parametric LQ problems–a survey," *IEEE Transactions on Automatic Control*, vol. 32, no. 8, pp. 658–671, 1987.

[7] I. Fatkhullin and B. Polyak, "Optimizing static linear feedback: Gradient method," *SIAM Journal on Control and Optimization*, vol. 59, no. 5, pp. 3887–3911, 2021.

[8] H. Mohammadi, M. Soltanolkotabi, and M. R. Jovanovic, "On the linear convergence of random search for discrete-time LQR," *IEEE Control Systems Letters*, vol. 5, no. 3, pp. 989–994, 2021.

[9] B. T. Polyak, "Gradient methods for the minimisation of functionals," *USSR Computational Mathematics and Mathematical Physics*, vol. 3, pp. 864–878, Jan. 1963.

[10] H. Karimi, J. Nutini, and M. Schmidt, "Linear convergence of gradient and Proximal-Gradient methods under the Polyak-Łojasiewicz condition," *arXiv:1608.04636*, Aug. 2016.

[11] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, pp. 1467–1476, 2018.

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

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

[14] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.

[15] D. Gabay, "Minimizing a differentiable function over a differential manifold," *Journal of Optimization Theory and Applications*, vol. 37, no. 2, pp. 177–219, 1982.

[16] S. T. Smith, "Optimization techniques on Riemannian manifolds," *Fields Institute Communications*, vol. 3, no. 3, pp. 113–135, 1994.

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

[18] J. Bu, A. Mesbahi, and M. Mesbahi, "On topological and metrical properties of stabilizing feedback gains: the MIMO case," *arXiv preprint arXiv:1904.02737*, 2019.

[19] J. M. Lee, *Introduction to Smooth Manifolds*. Springer, 2nd ed., 2013.

[20] J. M. Lee, *Introduction to Riemannian Manifolds*. Cham, Switzerland: Springer Nature, 2nd ed., 2018.

[21] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.