Controller Synthesis through Riemannian Optimization

Spencer Kraisler

Advisor: Mehran Mesbahi

Committee: Mehran Mesbahi · Behçet Açıkmeşe · Amir Taghvaei · Martin Nisser · Jing Wu

Robotics, Aerospace, and Information Networks (RAIN) Lab
William E. Boeing Department of Aeronautics & Astronautics
University of Washington

January 2026

Table of Contents

1 Introduction

2 Background

3 The Space of Controllers is a Riemannian Manifold and How to Optimize Over It

4 Trajectory Optimization over Smooth Manifolds: Exploiting Geometry

5 *Actual* Average Consensus over Lie Groups

6 Conclusion

Control Problems with Intrinsic Geometry

Many spaces arising in control theory are fundamentally *Riemannian*

1. Set of stable transfer functions
   - $T(s) = \frac{p(s)}{q(s)}$ vs.
     $\dot{x} = Ax + Bu, y = Cx + Du$ vs.
     $\mathcal{H} = (CB, CAB, CA^2B, \dots)$
2. Error-state EKF for IMUs – exploits $\mathrm{SE}(3)$ structure
3. Funnel synthesis, differential LMIs
4. Constrained optimization / Barrier methods
5. Geometric control theory

Figure: Geometry of feasible controllers

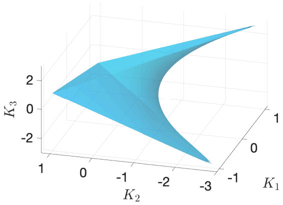
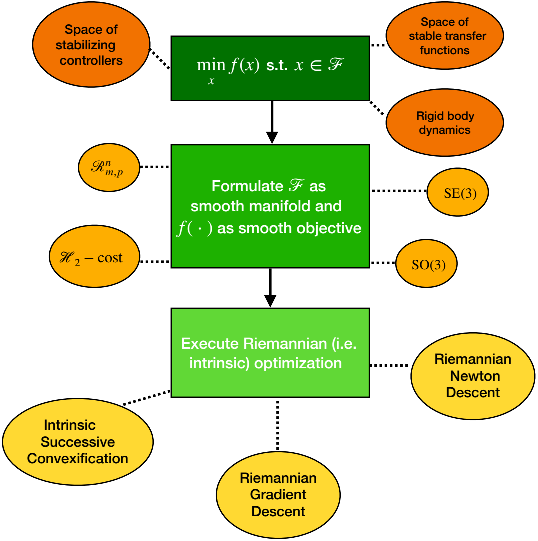


Figure: Static controller manifold

Unifying Viewpoint



Intrinsic Modeling

- Policies and trajectories as points on a manifold rather than vectors in a linear parameter space
- Dynamics, constraints, perturbations, defined intrinsically

Intrinsic Optimization

- Gradients and Hessians from the Riemannian metric
- Descent directions parameterization-invariant
- Updates performed via retractions

Motivation and Overview of Contributions

Motivation:

- Often $\mathcal{F}$ will have smaller dimension as a manifold than as a subset of the linear parameter space
- Uncover insights hidden by the reliance on parameters
- Resulting algorithm is invariant of any choosen coordinate system

This dissertation develops geometry-aware methods for both open-loop and closed-loop control. The main contributions are:

- **Linear Controller Policy Optimization**
- **Intrinsic Successive Convexification (iSCvx)**
- **Average Consensus on Lie Groups**

Table of Contents

1. Introduction
2. Background
3. The Space of Controllers is a Riemannian Manifold and How to Optimize Over It
4. Trajectory Optimization over Smooth Manifolds: Exploiting Geometry
5. *Actual* Average Consensus over Lie Groups
6. Conclusion

Smooth Manifold

Definition (Smooth Manifold)

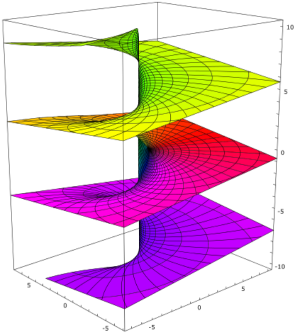
A *smooth manifold* is a space of points $\mathcal{M}$ and an open cover $\mathcal{A}$ of *local coordinate systems*,

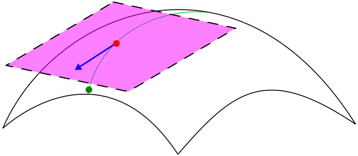
$$x : U \subset \mathcal{M} \to x(U) \subset \mathbb{R}^n$$

which are *smoothly compatible*: for all $(U, x), (V, y) \in \mathcal{A}$,

$$y \circ x^{-1} : x(U \cap V) \subset \mathbb{R}^n \to y(U \cap V) \subset \mathbb{R}^n$$

is classically smooth.





- A *thing* is smooth on $\mathcal{M}$ its its *coordinatization* is classically smooth
- Tangent space $T_x\mathcal{M} \cong \mathbb{R}^n$
    - $T_q\mathcal{Q} = \{v \in \mathbb{R}^4 : q^\top v = 0\} \cong \mathbb{R}^3$
- Calculus can be done on smooth manifolds

Table of Contents

1. Introduction

2. Background

3. The Space of Controllers is a Riemannian Manifold and How to Optimize Over It

4. Trajectory Optimization over Smooth Manifolds: Exploiting Geometry

5. *Actual* Average Consensus over Lie Groups

6. Conclusion

What is Direct Policy Optimization (PO)?

- **Goal:** bridge control theory and RL with stabilization guarantees + constraints
- Train linear controllers (via policy gradient methods)
  - **Novel focus:** optimizer performance $\implies$ study geometry of policy space + performance measure
- **Idea:** if training in real time, we need to know when policy will be safe

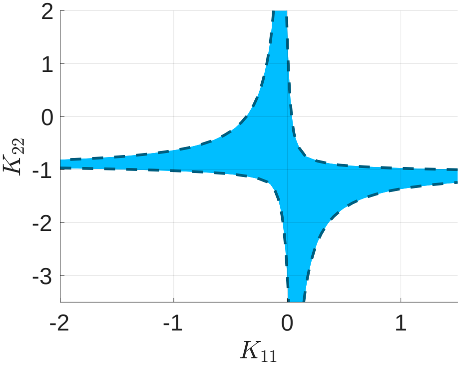


Figure: The set of stabilizing 2×2 **diagonal** control matrices K

LQR

- **Domain:** stabilizing static feedback matrices
- $\mathcal{H}_2$-cost is analytic, non-convex, gradient dominant
- global convergence under gradient descent (GD) with linear rate

LQG

- **Domain:** stabilizing dynamic linear controllers
- $\mathcal{H}_2$-cost is analytic, non-convex, non-strict saddle points, degenerate stationary points
- sublinear convergence rate under GD
- no local convergence guarantee

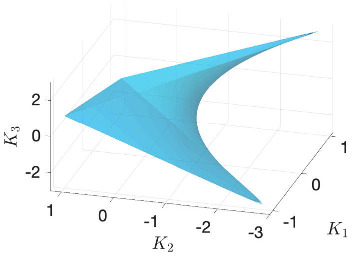


Figure: Set of stabilizing static controllers

Questions: Why sub-linear convergence rate? Why no convergence guarantee?

Stochastic LTI Systems

Consider the *minimal* (controllable + observable) linear system $\mathsf{P} = (A, B, C)$:

$$\dot{x} = Ax + Bu + w,$$
$$y = Cx + v$$

in feedback with a minimal linear controller $\mathsf{K} = (A_c, B_c, C_c)$:

$$\dot{x}_c = A_c x_c + B_c y,$$
$$u = C_c x_c$$

Set of stabilizing q-order minimal controllers:

$$\widetilde{\mathcal{C}} \subset \mathbb{R}^{q\times q} \times \mathbb{R}^{q\times p} \times \mathbb{R}^{m\times q}$$

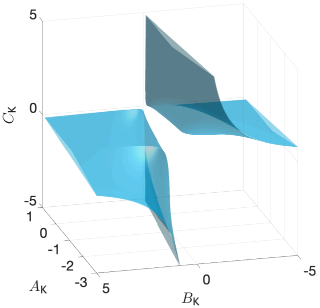


Figure: Illustration of the set of dynamic stabilizing policies $\widetilde{\mathcal{C}}_1$ for an LTI system with $A = 1.1$ and $B = C = 1$, resulting in two path-connected components

Stochastic LTI System (Cont.)

Coordinate-transformation: $\mathcal{T}_S(\mathsf{K}) = (SA_cS^{-1},\ SB_c,\ C_cS^{-1})$

$$\begin{cases} \dot{x}_c = A_c x_c + B_c y, \\ u = C_c x_c \end{cases} \implies \begin{cases} \dot{z}_c = SA_cS^{-1} z_c + SB_c y, \\ u = C_c S^{-1} z_c \end{cases}$$

$\mathcal{H}_2$ **Cost:**

$$\widetilde{J}(\mathsf{K}) := \lim_{T\to\infty} \mathbb{E}_w \frac{1}{T} \int_0^T (x^\top Q x + u^\top R u)\, dt$$

Theorem (Zheng, Tang, & Li, 2021)

$\widetilde{J} : \widetilde{\mathcal{C}} \to \mathbb{R}$ is analytic, non-convex, all minima are global, admits saddle points. Also, coordinate-invariance: $\widetilde{J}(\mathcal{T}_S(\mathsf{K})) = \widetilde{J}(\mathsf{K})$.

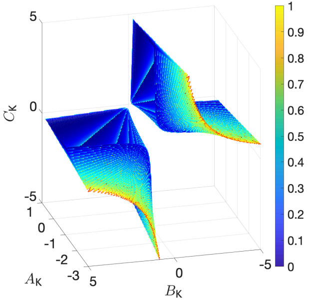


Figure: A colored plot of $\widetilde{J}(\cdot)$ over $\widetilde{\mathcal{C}}_1$.

Direct PO Re-visited

Coordinate-invariance $\implies$ n^2 dimensions of redundancy

An orbit: $\pi(\mathsf{K}) := \{\mathcal{T}_S(\mathsf{K}) : S \in \mathrm{GL}(n)\}$

Problem

Minimize $\widetilde{J}$ over $\widetilde{\mathcal{C}}$ fast (at least linear local rate) while exploiting coordinate-invariance.

Solution: Reformulate $\widetilde{\mathcal{C}}$ as a Riemannian manifold such that "$\nabla \widetilde{J}_{\mathsf{K}} \perp \pi(\mathsf{K})$".

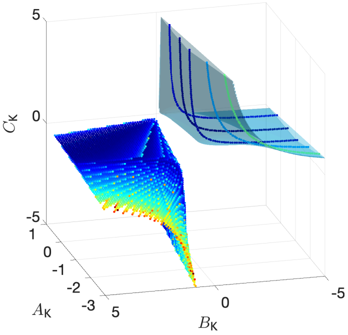


Figure: $\widetilde{J}(\cdot)$ over $\widetilde{\mathcal{C}}_1$ with a few orbits highlighted.

Example:

$$f(x) := x^4, \quad g_x(v, w) := v \cdot x^2 \cdot w$$
$$\overline{\nabla} f_x = 4x^3, \quad \nabla f_x = 4x$$

Riemannian metric:

- Inner product $g_x(\cdot, \cdot)$ on each tangent space $T_x\mathcal{M}$
- induces length, angle, gradient, Hessian

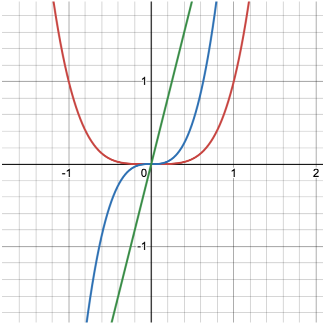


Figure: Graphs of $f(x)$, $\overline{\nabla} f(x)$, and ∇f_x.

Riemannian gradient:

$$\nabla f_x = G(x)^{-1}\, \overline{\nabla} f_x$$

where $\overline{\nabla} f$ is the Euclidean gradient.

Intuition: Riemannian gradient is preconditioning (think barrier-like scaling).

Retraction: $R_x : T_x\mathcal{M} \to \mathcal{M}$

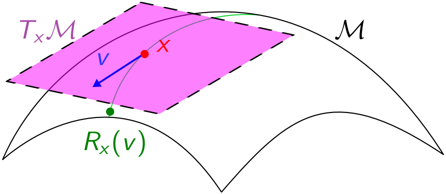


Figure: Retraction visualization

Riemannian Gradient Descent (RGD):

$$x_{k+1} = R_{x_k}(-\alpha\nabla f_{x_k})$$

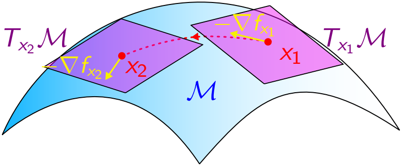


Figure: RGD visualization; $x_2 = R_{x_1}(-\alpha\nabla f_{x_1})$.

Important: the right metric can speed up RGD from sublinear to linear (locally).

The Krishnaprasad-Martin Metric

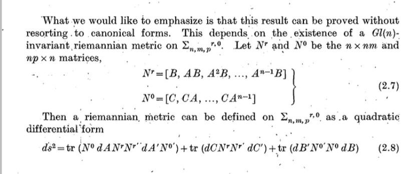
Intuition for the *right* Riemannian metric $g_{\mathrm{K}}^{\mathrm{KM}} : T_{\mathrm{K}}\widetilde{\mathcal{C}} \times T_{\mathrm{K}}\widetilde{\mathcal{C}} \to \mathbb{R}$:

1. Coordinate-invariant: $g_{\mathrm{K}}^{\mathrm{KM}}(\mathcal{T}_S(\mathbf{V}), \mathcal{T}_S(\mathbf{W})) = g_{\mathrm{K}}^{\mathrm{KM}}(\mathbf{V}, \mathbf{W})$.
2. Explodes as K becomes less stabilizing (barrier-like behavior).

Definition (Krishnaprasad-Martin metric)

$$
\begin{aligned}
g_{\mathrm{K}}^{\mathrm{KM}}(\mathbf{V}, \mathbf{W}) :=& \\
& c_1 \operatorname{tr}\left(W_o(\mathrm{K}) E(\mathbf{V}) W_c(\mathrm{K}) E(\mathbf{W})^\top\right) \\
& + c_2 \operatorname{tr}\left(F(\mathbf{V})^\top W_o(\mathrm{K}) F(\mathbf{W})\right) \\
& + c_3 \operatorname{tr}\left(G(\mathbf{V}) W_c(\mathrm{K}) G(\mathbf{W})^\top\right),
\end{aligned}
$$

where $W_c(\cdot)$, $W_o(\cdot)$ are closed-loop controllability/observability Gramians.



What we would like to emphasize is that this result can be proved without resorting to canonical forms. This depends on the existence of a $Gl(n)$-invariant riemannian metric on $\Sigma_{n,m,p}{}^{r,0}$. Let N^r and N^0 be the $n \times nm$ and $np \times n$ matrices,

$$
\left.
\begin{aligned}
N^r &= [B, AB, A^2B, \ldots, A^{n-1}B] \\
N^0 &= [C, CA, \ldots, CA^{n-1}]
\end{aligned}
\right\} \qquad (2.7)
$$

Then a riemannian metric can be defined on $\Sigma_{n,m,p}{}^{r,0}$ as a quadratic differential form

$$ds^2 = \operatorname{tr}(N^0\, dA N^r N^{r\prime}\, dA' N^{0\prime}) + \operatorname{tr}(dC N^r N^{r\prime}\, dC') + \operatorname{tr}(dB' N^{0\prime} N^0\, dB) \qquad (2.8)$$

Figure: Developed in 1983 to study manifolds of stable LTI systems.

The Algorithm and Convergence Analysis

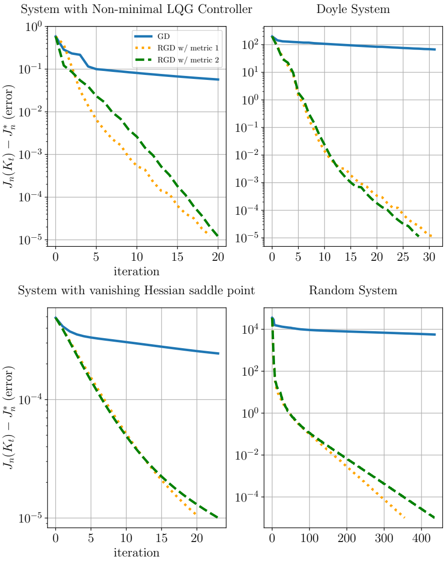
RGD of $(\widetilde{\mathcal{C}}, g^{\mathrm{KM}})$ over $\widetilde{J}(\cdot)$ with fixed step size:

$$\mathsf{K}_{t+1} = \mathsf{K}_t - \alpha \nabla \widetilde{J}_{\mathsf{K}_t} \qquad (3)$$

Theorem (Kraisler & Mesbahi, 2024)

Suppose the LQG controller is controllable + observable, and $\ker(\mathrm{Hess}\widetilde{J}_{\mathsf{K}^*}) = T_{\mathsf{K}^*}[\mathsf{K}^*]$. *Then there exists* $\alpha > 0$ *and a neighborhood* U *of* K^* *such that the sequence defined by* (3) *with* $\mathsf{K}_0 \in U$ *exists and converges to* $[\mathsf{K}^*]$ *with at least linear rate.*

Summary: local convergence + linear rate guarantee.



An Interpretation: *Smooth Orbit Manifolds*

- Orbit: $\pi(\mathsf{K}) = \{\mathcal{T}_S(\mathsf{K}) : S \in \mathrm{GL}(n)\}$
- Smooth orbit manifold: $\mathcal{C} := \{\pi(\mathsf{K}) : \mathsf{K} \in \widetilde{\mathcal{C}}\}$
- $\dim(\widetilde{\mathcal{C}}) = n^2 + nm + np$ and $\dim(\mathcal{C}) = nm + np$
- **Key point:** RGD over $\widetilde{\mathcal{C}} \iff$ RGD over $\mathcal{C}$

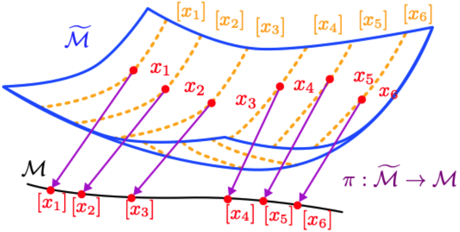


Figure: Visualization of a orbit manifold

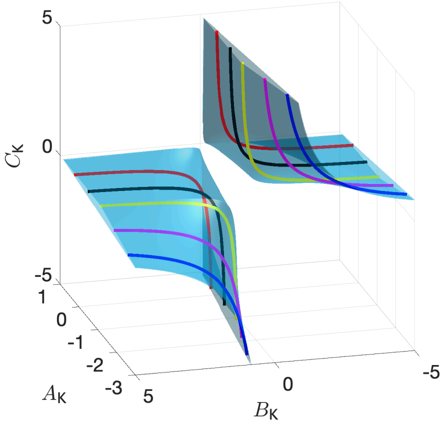


Figure: The orbits of $\widetilde{\mathcal{C}}_1$ colored

Key objects

- $\widetilde{\mathcal{C}}_q^*$: all order-$q$ controllers
- $\widetilde{\mathcal{C}}_q \subset \widetilde{\mathcal{C}}_q^*$: minimal controllers
- $\mathcal{C}_q := \widetilde{\mathcal{C}}_q / \mathrm{GL}(q)$: **controller orbit manifold**
- $\widetilde{\mathcal{X}} := \widetilde{\mathcal{C}}_q^* \setminus \widetilde{\mathcal{C}}_q$: non-minimal controllers

Why removing non-minimal controllers doesn't break connectivity

- $\mathrm{codim}(\widetilde{\mathcal{X}}) = \min(m, p) \;\geq 2$
- $\implies$ $\widetilde{\mathcal{X}}$ is "too thin" to separate $\widetilde{\mathcal{C}}_q^*$

Full-order punchline ($q = n$)

- $\widetilde{\mathcal{C}}_n$ has at most two path components and $\mathcal{C}_n$ is path-connected.

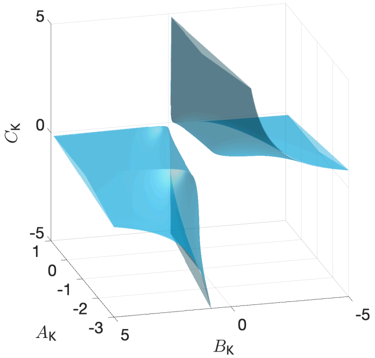


Figure: $\widetilde{\mathcal{C}}_n$ has at most 2 path-connected components, and the orbit space $\mathcal{C}_n$ has 1.

Orbit Riemannian Newton Descent (Orbit-RND)

Goal: achieve *quadratic* local convergence by doing Newton on $\mathcal{C}$:

Orbit-RND step: Let $\mathsf{k}_t := \pi(\mathsf{K}_t) \in \mathcal{C}$. Solve $\nabla^2 J_{\mathsf{k}_t}(\mathbf{h}_t) = \nabla J_{\mathsf{k}_t}$ and iterate

$$\mathsf{k}_{t+1} = R_{\mathsf{k}_t}(-\mathbf{h}_t).$$

Theorem (Local quadratic rate on the quotient)

If $\nabla^2 J_{\pi(\mathsf{K}^\star)}$ is invertible, then there exists a neighborhood U such that for $\mathsf{K}_0 \in U$, Orbit-RND is well-defined and $\pi(\mathsf{K}_t) \to \pi(\mathsf{K}^\star)$ with at least quadratic convergence.

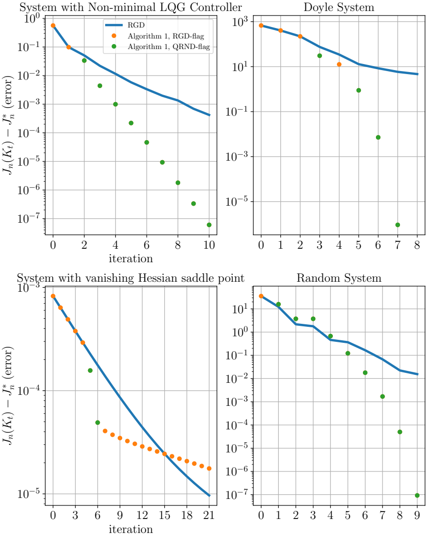


Figure: QRND vs. RGD: QRND exhibits

Conclusion

Summary:

- RGD is order(s) of magnitude faster than GD
- Local convergence guarantee
- Interpretation: RGD over the much smaller orbit controller manifold $\mathcal{C}$
- Local convergence guarantee Riemann Newton Descent with quadratic rate
- $\mathcal{C}_n$ is path-connected, $\widetilde{\mathcal{C}}_n$ has at most 2 path-connected components

Future directions:

- $\mathcal{H}_\infty$ nonsmooth optimization over $\mathcal{C}$
- Data-driven formulations

Table of Contents

1. Introduction

2. Background

3. The Space of Controllers is a Riemannian Manifold and How to Optimize Over It

4. Trajectory Optimization over Smooth Manifolds: Exploiting Geometry

5. *Actual* Average Consensus over Lie Groups

6. Conclusion

Trajectory Optimization over Smooth Manifolds

$$\begin{cases} \min\limits_{\mathbf{x},\mathbf{u}} & C(\mathbf{x},\mathbf{u}) := \sum_{k=0}^{N-1} \phi(x_k,u_k) + \phi_f(x_N) \\ \text{s.t.} & x_{k+1} = f(x_k,u_k), \\ & g(x_k,u_k) \le 0,\ x_0 \text{ given} \end{cases}$$

What if state x and input u are points on *smooth manifolds* $\mathcal{M} \subset \mathbb{R}^N$, $\mathcal{U} \subset \mathbb{R}^M$? Can we exploit their geometry and lower dimensions?

Examples:

- Rigid body dyanmics: $\dot{q} = \frac{1}{2} q \otimes \omega$, $q \in \mathcal{Q} \subset \mathbb{H} \cong \mathbb{R}^4$
- Optimal Covariance Steering: $P_{k+1} = A_k P_k A_k^\top + Q_k,\ P_k \in \mathrm{SPD}(n)$
- D-LMIs: $\mathcal{L}(K(t), P(t)) \preceq \dot{P}(t)$
- Constrained mechanical systems $\ddot{x} = f(x,\dot{x},u)$, $(x,\dot{x}) \in T\mathcal{M}$

Figure: Rigid body $SE(3)$

Outline of Successive Convexification (SCvx)

- 1st-order solver for constrained optimal control problems
- Allows infeasible initial trajectories
- Scalable
- Industry standard

$$\begin{cases} \min\limits_{\mathbf{x},\mathbf{u}} & C(\mathbf{x},\mathbf{u}) \\ \text{s.t.} & x_{k+1} = f(x_k, u_k) \\ & g(x_k, u_k) \le 0 \end{cases}$$

$\downarrow$ convexify about $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$

$$\begin{cases} \min\limits_{\delta\mathbf{x},\, \delta\mathbf{u},\, \mathbf{v},\, \mathbf{s}} & C(\bar{\mathbf{x}} + \delta\mathbf{x},\ \bar{\mathbf{u}} + \delta\mathbf{u}) + \lambda P(\mathbf{v}, \mathbf{s}) \\ \text{s.t.} & f(\bar{x}_k, \bar{u}_k) - \bar{x}_{k+1} + A_k \delta x_k + B_k \delta u_k \le s_k \\ & g(\bar{x}_k, \bar{u}_k) + S_k \delta x_k + T_k \delta u_k = v_k \\ & s \ge 0, \quad \|\delta x\| \le r \end{cases}$$

$\downarrow$

$$(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \leftarrow (\bar{\mathbf{x}}, \bar{\mathbf{u}}) + (\delta\mathbf{x}^*, \delta\mathbf{u}^*)$$

From Shortcomings to Goals

Shortcomings:

1. Different parameterizations $\implies$ Different performance
2. Redundant dimensions
   - Ex. $\dot{q} = \frac{1}{2} q \otimes \omega$, $A_k \in \mathbb{R}^{4\times4}$ instead of $\in \mathbb{R}^{3\times3}$
   - $\delta q_k \notin T_{q_k}\mathcal{Q}$
3. Next trajectory off manifold: $\bar{q} \leftarrow \bar{q} + \delta\bar{q}$

Goals:

1. Make SCvx *intrinsic* to the manifold, not to the parameterization
2. Intelligent perturbations: must be *tangent* to the manifold
3. Retraction will *add* perturbations to trajectory points, intrinsically

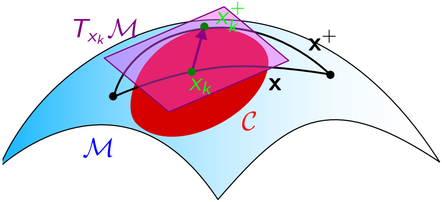


Figure: Visualization of perturbing a trajectory along a tangent space.

SCvx vs. Intrinsic SCvx: Key Differences

SCvx ($x \in \mathbb{R}^N, u \in \mathbb{R}^U$)	**iSCvx** ($x \in \mathcal{M}^n, u \in \mathcal{U}^m$)
Uses Jacobians $$A := D_x f(\bar{x}, \bar{u}) \in \mathbb{R}^{N \times N}, \quad B := D_u f(\bar{x}, \bar{u}) \in \mathbb{R}^{N \times M}$$ in the ambient space ($N \geq n$, $M \geq m$).	Uses intrinsic differentials $$d_x f_{(\bar{x},\bar{u})} : T_{\bar{x}}\mathcal{M} \to T_{f(\bar{x},\bar{u})}\mathcal{M}, \quad d_u f_{(\bar{x},\bar{u})} : T_{\bar{u}}\mathcal{U} \to T_{f(\bar{x},\bar{u})}\mathcal{M},$$ $$\mathbf{A} := [D_x f(\bar{x}, \bar{u})\|_{T_{\bar{x}}\mathcal{M}}] \in \mathbb{R}^{n \times n}, \quad \mathbf{B} := [D_u f(\bar{x}, \bar{u})\|_{T_{\bar{u}}\mathcal{U}}] \in \mathbb{R}^{n \times m}$$

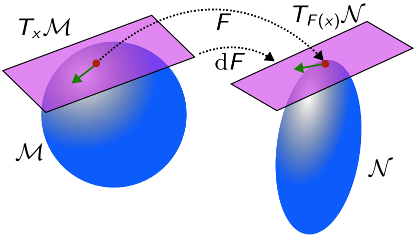


Figure: The differential $\mathrm{d}F_x : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$ of the function $F : \mathcal{M} \to \mathcal{N}$ at a point $x \in \mathcal{M}$

SCvx vs. Intrinsic SCvx: Key Differences (Cont.)

SCvx ($x \in \mathbb{R}^N, u \in \mathbb{R}^M$)	**iSCvx** ($x \in \mathcal{M}^n, u \in \mathcal{U}^m$)
Representation-dependent	Representation-invariant
Linearizes dynamics as $$\delta x_{k+1} \approx f(\bar{x}_k, \bar{u}_k) - \bar{x}_{k+1} + A_k\, \delta x_k + B_k\, \delta u_k$$ (LTV, redundant dimensions, does not account for projection)	Intrinsic linearization: $$\delta x_{k+1} \approx R^{-1}_{\bar{x}_{k+1}}(f(\bar{x}_k, \bar{u}_k)) + \mathbf{D}_k \circ (\mathbf{A}_k[\delta x_k] + \mathbf{B}_k[\delta u_k])$$ (LTV, minimal dimensions, accounts for projection)
Perturbations live in ambient $\mathbb{R}^N$ and $\mathbb{R}^M$ (redundant)	Perturbations live in $T_{\bar{x}}\mathcal{M} \cong \mathbb{R}^n$ and $T_{\bar{u}}\mathcal{U} \cong \mathbb{R}^m$ (minimal dimension)
Local problem exploits cost convexity: $C(\bar{\mathbf{x}} + \delta\mathbf{x}, \bar{\mathbf{u}} + \delta\mathbf{u})$	Local problem exploits cost *geodesic* convexity: $\hat{C}^{(2)}_{(\bar{\mathbf{x}},\bar{\mathbf{u}})}(\delta\mathbf{x}, \delta\mathbf{u}) := C(\bar{\mathbf{x}}, \bar{\mathbf{u}}) + \mathrm{d}C_{(\bar{\mathbf{x}},\bar{\mathbf{u}})}[\delta\mathbf{x}, \delta\mathbf{u}] + \frac{1}{2}\nabla^2 C_{(\bar{\mathbf{x}},\bar{\mathbf{u}})}[\delta\mathbf{x}, \delta\mathbf{u}]$
State update: $\bar{\mathbf{x}} \leftarrow \bar{\mathbf{x}} + \delta\mathbf{x}$, $\bar{\mathbf{u}} \leftarrow \bar{\mathbf{u}} + \delta\mathbf{u}$ (not manifold-preserving)	State update: $\bar{\mathbf{x}} \leftarrow R_{\bar{\mathbf{x}}}(\delta\mathbf{x})$, $\bar{\mathbf{u}} \leftarrow R_{\bar{\mathbf{u}}}(\delta\mathbf{u})$ (manifold-preserving)

Outline of Intrinsic SCvx

$$\begin{cases} \min\limits_{\mathbf{x}\subset\mathcal{M},\mathbf{u}\subset\mathcal{U}} & C(\mathbf{x},\mathbf{u}) \\ \text{s.t.} & x_{k+1} = f(x_k, u_k) \\ & g(x_k, u_k) \leq 0 \end{cases}$$

$\downarrow$ convexify about $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$

$$\begin{cases} \min\limits_{\delta\mathbf{x},\,\delta\mathbf{u},\,\mathbf{v},\,\mathbf{s}} & \hat{C}^{(2)}_{(\bar{\mathbf{x}},\bar{\mathbf{u}})}(\delta\mathbf{x}, \delta\mathbf{u}) + \lambda P(\mathbf{v}, \mathbf{s}) \\ \text{s.t.} & R^{-1}_{\bar{x}_{k+1}}(f(\bar{x}_k, \bar{u}_k)) + \mathbf{D}_k \circ (\mathbf{A}_k[\delta x_k] + \mathbf{B}_k[\delta u_k]) \leq s_k \\ & g(\bar{x}_k, \bar{u}_k) + \mathbf{S}_k[\delta x_k] + \mathbf{T}_k[\delta u_k] = v_k \\ & s \geq 0, \quad \|\delta x\| \leq r \end{cases}$$

$\downarrow$

$$(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \leftarrow (R_{\bar{\mathbf{x}}}(\delta\mathbf{x}^*), R_{\bar{\mathbf{u}}}(\delta\mathbf{u}^*))$$

Advantages

1. The convex sub-problem has dimension $\dim \mathcal{M}$, independent of the ambient representation.
    - E.g., quaternions (4 vars) or rotation matrices (9 vars) both yield an $n = 3$ LTV sub-problem.
2. The formulation is parameterization-invariant: any coordinate choice for $\mathcal{M}$ gives the same convex sub-problem.
3. Each updated trajectory remains on the manifold.
4. Convex sub-problem is feasible, as with SCvx.

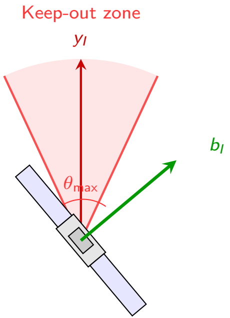
Example Problem: Constrained Attitude Control

Dynamics:

$$q_{k+1} = f(q_k, \omega_k) := q_k \otimes \exp(\Delta t \cdot \omega_k)$$

Retraction: $R_q(v) := q \otimes \exp\left(q^{-1} \otimes v\right)$
Input: Angular velocity $\omega_k \in \mathbb{R}^3$
State: Attitude $q_k \in \mathcal{Q}$
Constraints:

1. Keep-out zone:
   $g(q) := b_I \cdot y_I - \cos\theta_{\max}$,
   - y_I = inertial-fixed keep-out direction
   - b_B = body-fixed boresight direction
   - $\theta_{\max}$ = keep-out angular radius
2. Initial attitude: $q_0 \in \mathcal{Q}$
3. Target attitude: $q_{des} \in \mathcal{Q}$



- SCvx stage cost:

$$\phi(q_k, \omega_k) = \frac{1}{2}\|q_k - q_{des}\|_2^2 + \frac{1}{2}\|\omega_k\|_2^2$$

- iSCvx stage cost:

$$\phi(q_k, \omega_k) = \frac{1}{2} d_{\mathcal{Q}}(q_k, q_{des})^2 + \frac{1}{2}\|\omega_k\|_2^2$$

Experiments and Results

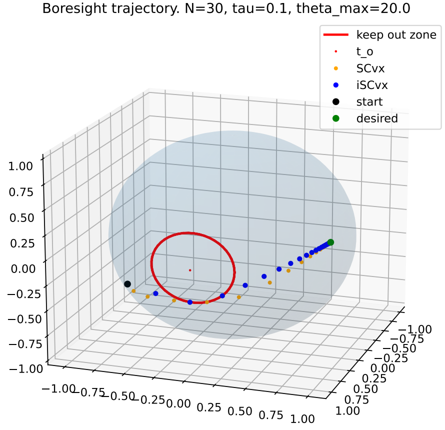


Figure: Trajectory of boresight vector, red circle is keep-out zone

TABLE I

SCvx and iSCvx Comparison: Case 1

	$\theta_{\max} = 10°$		$\theta_{\max} = 30°$	
	SCvx	**iSCvx**	**SCvx**	**iSCvx**
Avg. Iter.	40.21	24.89	45.8	26.8
Std. Iter.	9.23	2.14	18.28	1.88
Time (s)	6.26	4.40	7.10	4.70
Avg. Geo. Cost	4.98	4.73	6.50	5.67
Avg. Eucl. Cost	7.21	6.91	9.65	8.17

Results compare SCvx and iSCvx for spacecraft attitude control under different angle constraints with $N = 30$ time steps and discretization rate $\Delta t = 0.1$ sec. Values are averages over multiple runs.

TABLE II

SCvx and iSCvx Comparison: Case 2

	$\theta_{\max} = 10°$		$\theta_{\max} = 30°$	
	SCvx	**iSCvx**	**SCvx**	**iSCvx**
Avg. Iter.	67.9	24.75	65.72	25.65
Std. Iter.	34.86	2.22	17.02	2.45
Time (s)	22.18	9.09	21.43	9.37
Avg. Geo. Cost	9.04	8.96	10.59	10.50
Avg. Eucl. Cost	11.94	13.34	13.98	15.42

$N = 60$ time steps and discretization rate $\Delta t = 0.05$ sec.

Summary and Future Work

Intrinsic SCvx is:

- an SCvx-like non-linear optimal control solver
- Formulated intrinsically on system manifolds
- Solves sequence of smaller-dimensional convex sub-problems
- Invariant of parameterization

Future work:

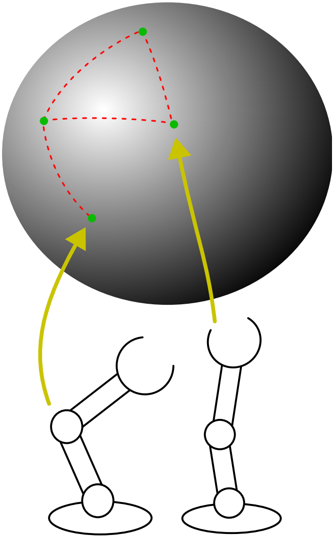
- Extend existing convergence guarantees for SCvx to the iSCvx
- Develop an in-depth tutorial of iSCvx with simple worked examples to make the method more accessible to practitioners
- Continue extending and refining the open-source iSCvx library

Table of Contents

1. Introduction
2. Background
3. The Space of Controllers is a Riemannian Manifold and How to Optimize Over It
4. Trajectory Optimization over Smooth Manifolds: Exploiting Geometry
5. *Actual* Average Consensus over Lie Groups
6. Conclusion

Motivation

- Examples:
  - 3D localization via network of cameras
  - Coordinated motion of robot arms
  - **Non-Euclidean** state space (Dome camera, covariance matrix, $SO(3)$, robot arm)
- Consensus is the **foundation** of distributed computation
  - Synchronize states of network of processors $\iff$ steer agents to a **single point**
- Consensus point needs geometric+statistical **significance**



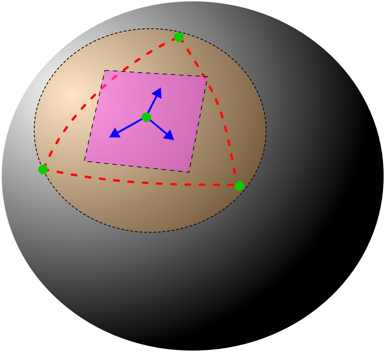
Average on a manifold: Frechet mean

Let $Z_1, \ldots, Z_N \in \mathcal{M}$

Fréchet function

$$f(R) := \frac{1}{2} \sum_{i=1}^{N} d_g(R, Z_i)^2, \qquad R \in \mathcal{M}.$$

Any global minimizer is a **Fréchet mean**



- In $\mathbb{R}^n$: $d_g(p, z) = \|p - z\|$ minimizer is the usual average.
- On manifolds: mean depends on geometry and can be non-unique and non-existant

Matrix Lie Groups and Bi-Invariant Metric

- A matrix Lie group $\mathcal{G}$ is a closed matrix subgroup of $\mathrm{GL}(n)$
- Tangent space at the identity $\mathfrak{g} := T_I\mathcal{G}$ encodes local group structure
- Left translation $L_S(R) := SR$ induces an isomorphism

$$\mathrm{d}L_{R^{-1}} : T_R\mathcal{G} \to \mathfrak{g}$$

  allowing tangent vectors to be identified in $\mathfrak{g}$
- A bi-invariant metric g satisfies

$$g_R(V, W) = g_{AR}(AV, AW) = g_{RA}(VA, WA), \quad \forall A \in \mathrm{GL}(n)$$

Why Lie groups for consensus?

- Algebraic structure enables intrinsic averaging and cancellation.
- Relationship between geodesic distance and matrix log: $d_g(R, S) = \|\log\left(R^{-1}S\right)\|_2$
- Antisymmetry: $\log\left(R^{-1}S\right) = -\log\left(S^{-1}R\right)$.

Problem statement: distributed average consensus on a Lie Group

Let $\mathcal{G}$ be a Lie group with a bi-invariant Riemannian metric, and let $Z_1, \ldots, Z_N \in \mathcal{G}$ admit a unique Fréchet mean $\bar{Z}$.

Agents and communication

Agents have states $R_i(t) \in \mathcal{G}$ on a fixed, connected, undirected graph $\mathbf{G} = ([N], E)$. Each agent uses only neighbor info.

Goal

Design distributed dynamics so that

$$\lim_{t \to \infty} R_i(t) = \bar{Z}, \qquad \forall i \in [N],$$

using only intrinsic operations.

Reformulation: mean = consensus-constrained optimization

Lift to the product manifold $\mathcal{G}^N$ with $\mathbf{R} = (R_1, \ldots, R_N)$:

$$F(\mathbf{R}) := \frac{1}{2} \sum_{i=1}^{N} d_g(R_i, Z_i)^2.$$

Diagonal (consensus) set

$$\mathcal{D} := \{(R_1, \ldots, R_N) \in \mathcal{G}^N : R_1 = \cdots = R_N\}.$$

We have

$$F(R, \ldots, R) = f(R).$$

Equivalent problem

$$\min_{\mathbf{R} \in \mathcal{G}^N} \; F(\mathbf{R}) \quad \text{s.t.} \quad \mathbf{R} \in \mathcal{D}.$$

This makes Riemannian average consensus a distributed optimization problem.

(Euclidean Case) Consensus optimization

- Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be strongly convex with Lipschitz gradient
- Consider $f(x) = \sum_i f_i(x)$ and the lifted objective $F(\mathbf{x}) = \sum_{i=1}^N f_i(x_i)$.
- Our goal is to solve

$$\left\{ \min_x f(x) = \sum_{i=1}^N f_i(x) \right\} \sim \left\{ \min_{\mathbf{x}} F(\mathbf{x}) = \sum_{i=1}^N f_i(x_i) \text{ s.t. } x_1 = \ldots = x_N \right\}$$

in a distributed way

(Euclidean Case) Gradient Tracking

If each agent had the global average gradient, the problem is solved:

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} (x_j - x_i) - \frac{1}{N} \sum_{j=1}^{N} \nabla f_j(x_j).$$

But not distributed ...

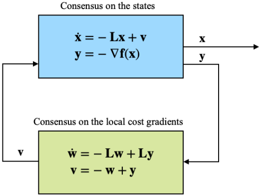
Gradient tracking (GT) idea

Each agent runs a dynamic consensus algorithm to asymptotically reconstruct the network average gradient using only neighbor communication.

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} (x_j - x_i) - \omega_i,$$
$$\dot{\alpha}_i = \sum_{j \in \mathcal{N}_i} (\alpha_j - \alpha_i + \nabla f_i(x_i) - \nabla f_j(x_j)),$$
$$\omega_i = -\alpha_i + \nabla f_i(x_i).$$

Key idea

$$\omega_i(t) \approx:= \frac{1}{N} \sum_{j=1}^{N} \nabla f_j(x_j(t)).$$



Bi-invariant Lie groups Allow for GT

On a matrix Lie group $\mathcal{G}$ with a bi-invariant metric g:

- Local costs: $f_i(R) = \frac{1}{2} d_g(R, Z_i)^2$ with

$$\nabla f_i(R) = -R \log\left(R^{-1} Z_i\right)$$

- Consensus dynamics:

$$\dot{R}_i = R_i \sum_{j \in \mathcal{N}_i} \log\left(R_i^{-1} R_j\right)$$

Gradient tracking:

$$\dot{R}_i = R_i \left(\sum_{j \in \mathcal{N}_i} \log\left(R_i^{-1} R_j\right) - \omega_i \right)$$

$$\dot{\alpha}_i = \sum_{j \in \mathcal{N}_i} (\alpha_j - \alpha_i + R_i^{-1} \nabla f_i(R_i) - R_j^{-1} \nabla f_j(R_j))$$

$$\omega_i = -\alpha_i + R_i^{-1} \nabla f_i(R_i)$$

Stationary Points Characterization

If the agents converge to a stationary point, then

$$R_1^\star = \cdots = R_N^\star = \bar{Z}$$

(consensus at the unique Fréchet mean).

Simulations on SO(3): comparisons + convergence rate

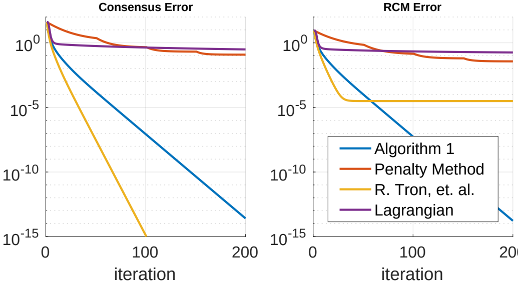


Figure: Consensus error and Frechet mean error vs iteration (proposed vs baselines).

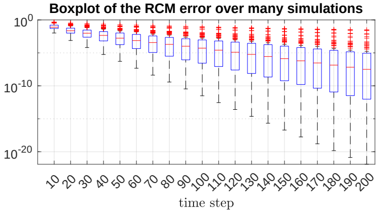


Figure: 100 random instances: Frechet mean error statistics show linear-rate decay.

Takeaways

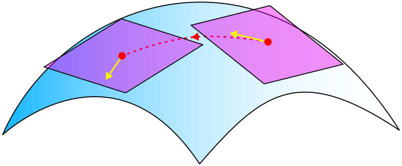
- Proposed method: clear linear-rate decay in both consensus + Frechet mean metrics.
- "Consensus-only" Riemannian schemes can agree but miss the Frechet mean.
- Next step: full convergence theory on Lie groups beyond Euclidean case.

Table of Contents

1. Introduction
2. Background
3. The Space of Controllers is a Riemannian Manifold and How to Optimize Over It
4. Trajectory Optimization over Smooth Manifolds: Exploiting Geometry
5. *Actual* Average Consensus over Lie Groups
6. Conclusion

Conclusion

- Many control problems reduce to optimization
- Treat the feasible set intrinsically as a **low-dimensional Riemannian manifold**
- Intrinsic geometry reveals structure hidden by coordinates
- Solve control problem using Riemannian optimization
- Applications in controller synthesis, trajectory optimization, and multi-agent control



Works

1. Bandyopadhyay, Deole, and Kraisler et al., *Vision-based Distributed Pose Estimation using a Spacecraft Constellation*, AIAA SciTech, 2023.
2. Deole, Talebi, and Kraisler et al., *Multi-Agent Passivity-based Control for Perception-based Guidance*, AIAA SciTech, 2023.
3. Kraisler and Talebi, *Distributed Consensus on Manifolds using the Riemannian Center of Mass*, IEEE CCTA, 2023.
4. Kraisler, Talebi, and Mesbahi, *Consensus on Lie Groups for the Riemannian Center of Mass*, IEEE CDC, 2023.
5. Kraisler and Mesbahi, *Output-feedback Synthesis Orbit Geometry: Quotient Manifolds and LQG Direct Policy Optimization*, IEEE Control Systems Letters, 2024.
6. Talebi, Zheng, and Kraisler et al., *Policy Optimization in Control: Geometry and Algorithmic Implications*, arXiv, 2024.
7. Kraisler, Mesbahi, and Açıkmeşe, *Intrinsic Successive Convexification: Trajectory Optimization on Smooth Manifolds*, IEEE Control Systems Letters, 2025.
8. Holder, Kraisler, and Mesbahi, *Centralized and Distributed Strategies for Handover-Aware Task Allocation in Satellite Constellations*, Journal of Guidance, Control, and Dynamics, 2025.
9. **Ongoing**: Kraisler, Talebi, and Mesbahi, *Direct Policy Optimization over the Orbit Space of Linear Dynamic Controllers*.

Acknowledgements

Thank you

- My wife, my parents, family and friends
- RAIN Lab: Mehran, Shahriar, Aditya, Newsha, Mengyuan, Dillon, Taylor, Jingjing, Matias, Unsik, Siavash, Dan, Thomas, Beniamino, Shiva, Jackson, Joseph, Justin, and the undergraduates
- UW Aero & Astro: Behcet, Amir, Jing, Martin, Lillian, ACL
- JPL Lab: Amir, Saptarshi, Vinod
- The controls community, CDC, CCTA, NASA JPL and JSC, Aerospace industry

RAIN

WILLIAM E. BOEING DEPARTMENT OF AERONAUTICS & ASTRONAUTICS

College of Engineering

JPL

Jet Propulsion Laboratory

California Institute of Technology

Conference on Decision and Control

CDC2025

Rio de Janeiro, Brazil. Dec. 09 - 12

NSF