Diffeomorphic Logarithm of Special Orthogonal Matrices

Speaker: Wen Huang

Xiamen University

April 11, 2026

Joint work with Zhifeng Deng, P.-A. Absil, and Kyle A. Gallivan

ICOTA2026

Outline

Special Orthogonal Group

$$\begin{cases} \mathbb{SO}_n := \{Q \in \mathbb{R}^{n\times n} : Q^{\mathsf{T}}Q = \mathrm{I}_n, \det(Q) = 1\} & \text{Special orthogonal group,} \\ \mathbf{Skew}_n := \{A \in \mathbb{R}^{n\times n} : A + A^{\mathsf{T}} = \mathbf{0}\} & \text{Skew-symmetric matrices,} \\ T_Q\mathbb{SO}_n := \{QX : X \in \mathbf{Skew}_n\} & \text{Tangents of } \mathbb{SO}_n. \end{cases}$$

Special Orthogonal Group

$$\begin{cases} \mathbb{SO}_n := \{Q \in \mathbb{R}^{n\times n} : Q^\top Q = \mathrm{I}_n, \det(Q) = 1\} & \text{Special orthogonal group,} \\ \mathbf{Skew}_n := \{A \in \mathbb{R}^{n\times n} : A + A^\top = \mathbf{0}\} & \text{Skew-symmetric matrices,} \\ T_Q\mathbb{SO}_n := \{QX : X \in \mathbf{Skew}_n\} & \text{Tangents of } \mathbb{SO}_n. \end{cases}$$

- Quantum theory [Wil67], computer vision [TVSC11], Deep learning [ZBL+19], and etc.

Special Orthogonal Group

$$\begin{cases} \mathbb{SO}_n := \{Q \in \mathbb{R}^{n\times n} : Q^\mathsf{T} Q = \mathrm{I}_n, \det(Q) = 1\} & \text{Special orthogonal group,} \\ \mathbf{Skew}_n := \{A \in \mathbb{R}^{n\times n} : A + A^\mathsf{T} = \mathbf{0}\} & \text{Skew-symmetric matrices,} \\ T_Q\mathbb{SO}_n := \{QX : X \in \mathbf{Skew}_n\} & \text{Tangents of } \mathbb{SO}_n. \end{cases}$$

- Quantum theory [Wil67], computer vision [TVSC11], Deep learning [ZBL+19], and etc.
- The matrix exponential: $\exp : \mathbf{Skew}_n \to \mathbb{SO}_n$.

Special Orthogonal Group

$$\begin{cases} \mathbb{SO}_n := \{Q \in \mathbb{R}^{n\times n} : Q^\top Q = \mathrm{I}_n, \det(Q) = 1\} & \text{Special orthogonal group,} \\ \mathbf{Skew}_n := \{A \in \mathbb{R}^{n\times n} : A + A^\top = \mathbf{0}\} & \text{Skew-symmetric matrices,} \\ T_Q\mathbb{SO}_n := \{QX : X \in \mathbf{Skew}_n\} & \text{Tangents of } \mathbb{SO}_n. \end{cases}$$

- Quantum theory [Wil67], computer vision [TVSC11], Deep learning [ZBL+19], and etc.
- The matrix exponential: $\exp : \mathbf{Skew}_n \to \mathbb{SO}_n$.
- The inverse $\exp^{-1}$: multivalued.

Special Orthogonal Group

$$\begin{cases} \mathbb{SO}_n := \{Q \in \mathbb{R}^{n\times n} : Q^\mathsf{T} Q = \mathrm{I}_n, \det(Q) = 1\} & \text{Special orthogonal group,} \\ \mathbf{Skew}_n := \{A \in \mathbb{R}^{n\times n} : A + A^\mathsf{T} = \mathbf{0}\} & \text{Skew-symmetric matrices,} \\ T_Q\mathbb{SO}_n := \{QX : X \in \mathbf{Skew}_n\} & \text{Tangents of } \mathbb{SO}_n. \end{cases}$$

- Quantum theory [Wil67], computer vision [TVSC11], Deep learning [ZBL+19], and etc.
- The matrix exponential: $\exp : \mathbf{Skew}_n \to \mathbb{SO}_n$.
- The inverse $\exp^{-1}$: multivalued.
- The principal matrix logarithm: $\log(Q)$ of $Q \in \mathbb{SO}_n$ where any eigenvalue $\pm\imath\theta_i$ of $\log(Q)$ is constrained by $|\theta_i| < \pi$ for all i.

Special Orthogonal Group

$$\begin{cases} \mathbb{SO}_n := \{Q \in \mathbb{R}^{n\times n} : Q^{\mathsf{T}}Q = \mathrm{I}_n, \det(Q) = 1\} & \text{Special orthogonal group,} \\ \mathbf{Skew}_n := \{A \in \mathbb{R}^{n\times n} : A + A^{\mathsf{T}} = \mathbf{0}\} & \text{Skew-symmetric matrices,} \\ T_Q\mathbb{SO}_n := \{QX : X \in \mathbf{Skew}_n\} & \text{Tangents of } \mathbb{SO}_n. \end{cases}$$

- Quantum theory [Wil67], computer vision [TVSC11], Deep learning [ZBL+19], and etc.
- The matrix exponential: $\exp : \mathbf{Skew}_n \to \mathbb{SO}_n$.
- The inverse $\exp^{-1}$: multivalued.
- The principal matrix logarithm: $\log(Q)$ of $Q \in \mathbb{SO}_n$ where any eigenvalue $\pm\imath\theta_i$ of $\log(Q)$ is constrained by $|\theta_i| < \pi$ for all i.

Why is the inverse of the matrix exponential map important?

Karcher Mean in $\mathbb{SO}_n$

Karcher Mean in $\mathbb{SO}_n$, [Moa02] and [KHM07].

The Karcher mean of $\{Q_i\}_{i=1}^K \in \mathbb{SO}_n$ minimizes the sum of squared distances (under the Riemannian metric) to each data point, i.e., it solves the minimization problem:

$$\min_{X \in \mathbb{SO}_n} \mathcal{K}(X) := \frac{1}{2K} \sum_{i=1}^K d(X, Q_i)^2, \tag{1}$$

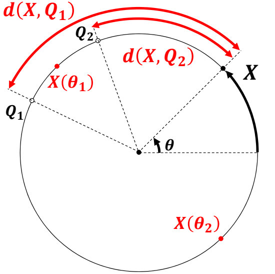
where the distance is given by the Frobenius norm of the principal logarithm: $d(X, Y)^2 = \frac{1}{2} \| \log(X^\mathsf{T} Y) \|_\mathrm{F}^2$.

[Moa02] Moakher, Maher. Means and averaging in the group of rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1-16, 2002.
[KHM07] Krakowski, Krzysztof and et al., On the computation of the Karcher mean on spheres and special orthogonal groups. In *Conference Paper, Robomat.* 2007

Subtlety of Mean Computation in $\mathbb{SO}_n$

Remark: Neither smooth nor convex.

The objective $\mathcal{K} : \mathbb{SO}_n \to \mathbb{R}$ is globally nonsmooth and nonconvex.



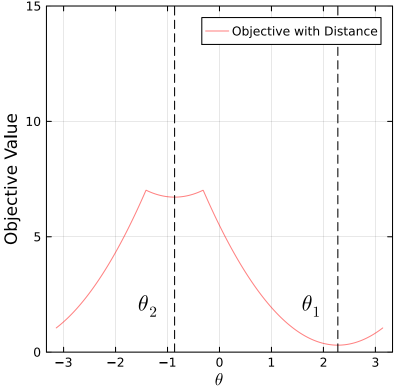


Figure: Karcher Mean of two points in $\mathbb{SO}_2$

Subtlety of Mean Computation in $\mathbb{SO}_n$

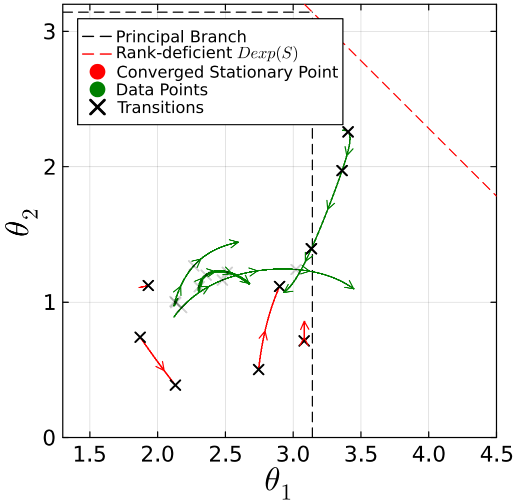
Remark: Neither smooth nor convex.

The objective $\mathcal{K} : \mathbb{SO}_n \to \mathbb{R}$ is globally nonsmooth and nonconvex.

- Multiple local minimizers arise.
- Classical smooth optimization theory does not apply.
- Global optimality is guaranteed within a geodesic ball of radius $< \pi/2$, [KHM07].
- These challenges are well recognized in both the theoretical literature [ATV13] and applications, e.g., [DRW$^+$20].

[KHM07] Krakowski, Krzysztof and et al., On the Computation of the Karcher Mean on Spheres and Special Orthogonal Groups. In *Conference Paper, Robomat.* 2007
[ATV13] Afsari, Bijan and et al., On the Convergence of Gradient Descent for Finding the Riemannian Center of Mass. *SIAM Journal on Control and Optimization*, 51(3):2230–2260, 2013
[DRW+20] Dellaert, Frank and et al., Shonan Rotation Averaging. *ECCV*, 2020

Subtlety of Mean Computation in $\mathbb{SO}_n$

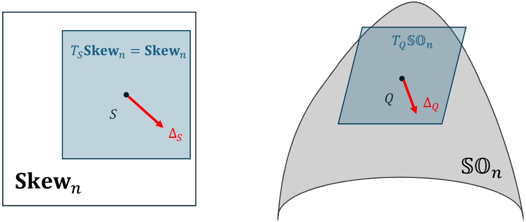


Converged stationary point $M(t)$ of $\mathcal{K}(X)$ of four moving points $Q_i(t)$ in $\mathbb{SO}_4$

Pullback of the Special Orthogonal Group

$$\exp\colon \mathbf{Skew}_n \to \mathbb{SO}_n, \qquad S \mapsto Q$$

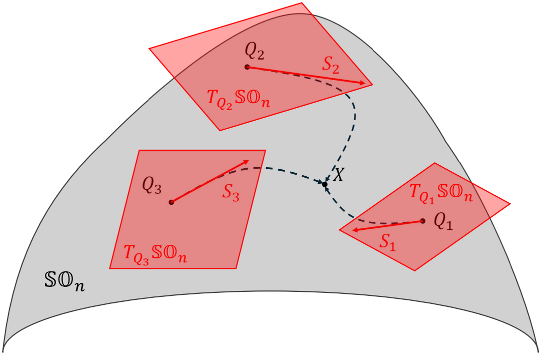
$$\mathrm{dexp}_S\colon T_S\mathbf{Skew}_n \to T_Q\mathbb{SO}_n, \qquad \Delta_S \mapsto \Delta_Q$$



Pullback Cost:
$\hat{f}\colon \mathbf{Skew}_n \to \mathbb{R}$,
$S \mapsto (f \circ \exp)\ (S)$

Cost function:
$f\colon \mathbb{SO}_n \to \mathbb{R}$

Pullback of the Karcher Mean Problem



Pullback in $(\mathbf{Skew}_n)^K$	**Objective in $\mathbb{SO}_n$**
$\mathcal{J}(S_1,\dots,S_K)=\frac{1}{4K}\sum_{i=1}^{K}\Vert S_i\Vert_F^2$	$\mathcal{K}(X)=\frac{1}{2K}\sum_{i=1}^{K}d(X,Q_i)^2$

Pullback of the Karcher mean problem.

Pullback of the Karcher Mean Problem

Definition: Relaxed Karcher Mean of the Pullback Optimization

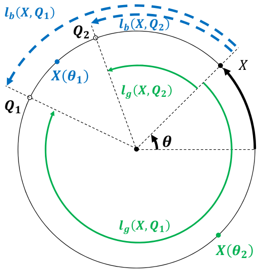
Consider a set of Riemannian geodesics $Q_i \exp(tS_i), S_i \in \mathbf{Skew}_n$ that arrive at the same endpoint $Q_i \exp(S_i) = Q_j \exp(S_j)$ for $i, j \leq K$. Denote

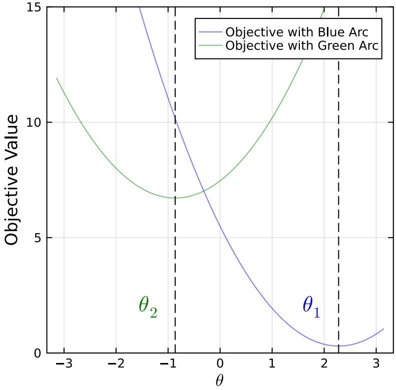
$$\mathcal{S}(Q_1, \ldots, Q_K) := \left\{(S_1, \ldots, S_K) \in (\mathbf{Skew}_n)^K : Q_i \exp(S_i) = Q_j \exp(S_j)\right\}.$$

A relaxed Karcher mean $\overline{Q}$ is the shared endpoint $Q_i \exp(S_i), \forall i \leq K$ of a local minimizer in

$$\underset{(S_1,\ldots,S_K)\in\mathcal{S}(Q_1,\ldots,Q_K)}{\arg\min} \mathcal{J}(S_1, \ldots, S_K) := \tfrac{1}{4K} \textstyle\sum_{i=1}^{K} \|S_i\|_{\mathrm{F}}^2. \tag{2}$$

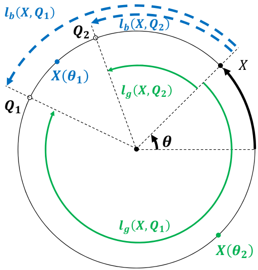
Relaxed Karcher Mean in $\mathbb{SO}_2$

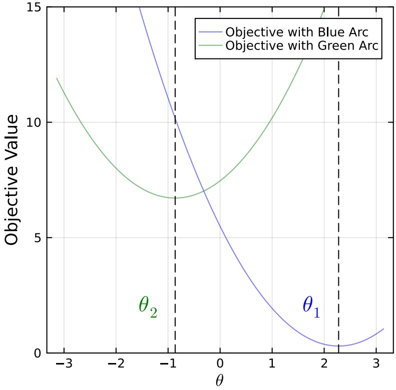




Relaxed Karcher Mean of two points in $\mathbb{SO}_2$

Relaxed Karcher Mean in $\mathbb{SO}_2$





Relaxed Karcher Mean of two points in $\mathbb{SO}_2$

Not necessarily use the length of the shortest geodesic

Ingredients for Pullback Optimization in $\mathbb{SO}_n, n \geq 3$

Pullback: $\hat{f} = f \circ \exp(S)$ versus original: $f(Q)$

S can be viewed as a preimage of $\exp$ at Q.

How to rigorously define such a preimage?

Ingredients for Pullback Optimization in $\mathbb{SO}_n, n \geq 3$

Pullback: $\hat{f} = f \circ \exp(S)$ versus original: $f(Q)$

S can be viewed as a preimage of $\exp$ at Q.

How to rigorously define such a preimage?

- A reference point $S \in \mathbf{Skew}_n$;
- A matrix $Q \in \mathbb{SO}_n$

Find a preimage of $\exp$ at Q that is closest to S in certain sense.

Ingredients for Pullback Optimization in $\mathbb{SO}_n$, $n \geq 3$

Pullback: $\hat{f} = f \circ \exp(S)$ versus original: $f(Q)$
S can be viewed as a preimage of $\exp$ at Q.
How to rigorously define such a preimage?

- A reference point $S \in \mathbf{Skew}_n$;
- A matrix $Q \in \mathbb{SO}_n$

Find a preimage of $\exp$ at Q that is closest to S in certain sense.

Consider $S \in \mathbf{Skew}_n$, $Q \in \mathbb{SO}_n$, and a neighborhood $\mathcal{M}$ of S where $Q \in \mathcal{Z} := \{\exp(X) : X \in \mathcal{M}\}$

- Diffeomorphism in $\exp : \mathcal{M} \to \mathcal{Z}$ (the inverse function theorem):
  - Invertible $\mathrm{D}\exp(X) : T_X\mathbf{Skew}_n \to T_{\exp(X)}\mathbb{SO}_n, \forall X \in \mathcal{M}$, [DAGH25];
  - Bijective nature of $X \leftrightarrow \exp(X), \forall X \in \mathcal{M}$.
- Reliable and efficient algorithms:
  - Derivative and its inverse in $T_X\mathbf{Skew}_n \leftrightarrow T_{\exp(X)}\mathbb{SO}_n$, [DAGH25];
  - Inversion $\exp^{-1} : \mathcal{Z} \to \mathcal{M}$.

[DAGH25] Deng, Zhifeng, Absil, P.-A., Gallivan, Kyle and Huang, Wen, The Exponential of Skew-Symmetric Matrices: A Nearby Inverse and Efficient Computation of Derivatives.

Ingredients for Pullback Optimization in $\mathbb{SO}_n$, $n \geq 3$

The main contributions of the work described in this talk include:

- Fully addressed diffeomorphism structure in $\exp : \mathbf{Skew}_n \to \mathbb{SO}_n$.
- Algorithm for the inversion termed *diffeomorphic logarithm* when a reference point $S \in \mathbf{Skew}_n$ is given.
- Application in the relaxed Karcher mean problem.

Outline

D-Notations

- The inverse of $\exp : \textbf{Skew}_n \to \mathbb{SO}_n$ is multivalued;
- **Ideally**: Divide $\textbf{Skew}_n$ into multiple components such that $\exp^{-1}$ is single-valued in one component;

D-Notations.

D-notations refers to those points $X \in \textbf{Skew}_n$ where $\mathrm{D}\exp(X) : T_X\textbf{Skew}_n \to T_{\exp(X)}\mathbb{SO}_n$ is invertible:

- D-point X: **invertible** $\mathrm{D}\exp(X)$;
- D-component $\mathcal{M}$ of X: a maximal connected subset of D-points containing X;
- D-connected X and Y: two D-points in a same D-component.

Moreover, the set of X with **rank-deficient** $\mathrm{D}\exp(X)$ is denoted by

$$\mathfrak{S} := \{X \in \textbf{Skew}_n : \mathrm{D}\exp(X) \text{ is rank-deficient}\}.$$

$\mathfrak{S}$ divides **Skew** into multiple components!

Rank Deficient Derivative and D-components

Rank-deficiency condition [DAGH25, Theorem 3.1].

$X \in \mathbf{Skew}_n$ is not a D-point iff X has a pair of angles θ_i and θ_j satisfying

$$\exists i \neq j \leq k \text{ where } n = 2k, \exists \ell \in \mathbb{Z} \setminus \{0\}, \quad \theta_i \pm \theta_j = 2\pi\ell. \tag{3}$$

An angle θ is obtained from a pair of conjugate eigenvalues $\pm\imath\theta$ of X. The set of points with rank deficient derivative is denoted by

$$\mathfrak{S} = \{X \in \mathbf{Skew}_n : \exists i \neq j \leq k, \exists \ell \in \mathbb{Z} \setminus \{0\}, \text{ s.t. } \theta_i \pm \theta_j = 2\ell\pi\}.$$

We only consider even n in this talk.
The odd n can be treated similarly, see details in [DAGH25].

Rank Deficient Derivative and D-components

Rank-deficiency condition [DAGH25, Theorem 3.1].

$X \in \mathbf{Skew}_n$ is not a D-point iff X has a pair of angles θ_i and θ_j satisfying

$$\exists i \neq j \leq k \text{ where } n = 2k, \exists \ell \in \mathbb{Z} \setminus \{0\}, \quad \theta_i \pm \theta_j = 2\pi\ell. \tag{3}$$

An angle θ is obtained from a pair of conjugate eigenvalues $\pm\imath\theta$ of X. The set of points with rank deficient derivative is denoted by

$$\mathfrak{S} = \{X \in \mathbf{Skew}_n : \exists i \neq j \leq k, \exists \ell \in \mathbb{Z} \setminus \{0\}, \text{ s.t. } \theta_i \pm \theta_j = 2\ell\pi\}.$$

D-Components in the complement [DAGH25, Corollary 4.4].

$\mathbf{Skew}_n \setminus \mathfrak{S}$ consists of countably many open D-components, denoted by

$$\mathcal{M}_e, e \in \mathcal{E}, \quad \text{where } \mathcal{E} \text{ is a countable set of indices.}$$

Rank Deficient Derivative and D-components

Rank-deficiency condition [DAGH25, Theorem 3.1].

$X \in \mathbf{Skew}_n$ is not a D-point iff X has a pair of angles θ_i and θ_j satisfying

$$\exists i \neq j \leq k \text{ where } n = 2k, \exists \ell \in \mathbb{Z} \setminus \{0\}, \quad \theta_i \pm \theta_j = 2\pi\ell. \tag{3}$$

An angle θ is obtained from a pair of conjugate eigenvalues $\pm\imath\theta$ of X. The set of points with rank deficient derivative is denoted by

$$\mathfrak{S} = \{X \in \mathbf{Skew}_n : \exists i \neq j \leq k, \exists \ell \in \mathbb{Z} \setminus \{0\}, \text{ s.t. } \theta_i \pm \theta_j = 2\ell\pi\}.$$

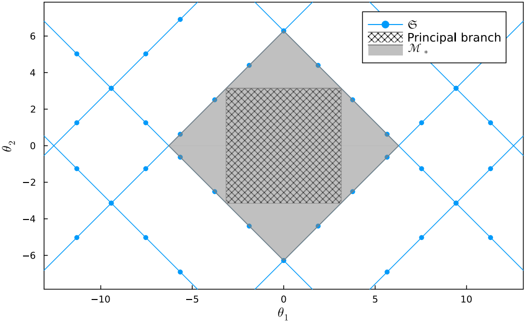
D-Components in the complement [DAGH25, Corollary 4.4].

$\mathbf{Skew}_n \setminus \mathfrak{S}$ consists of countably many open D-components, denoted by

$$\mathcal{M}_e, e \in \mathcal{E}, \quad \text{where } \mathcal{E} \text{ is a countable set of indices.}$$

How to characterize $\mathcal{M}_e$?

D-Components



Skew$_4$ plot by the two angles.

Remark: The special D-component.

The principal matrix logarithms of $\mathbb{SO}_n$ form a proper subset of

$$\mathcal{M}_* := \{X \in \mathbf{Skew}_n : |\theta_i \pm \theta_j| < 2\pi, \forall i \neq j\}.$$

Characterization of D-Component

Proposition: Canonical Schur Decomposition.

All $Q \in \mathbb{SO}_n$ admit a Schur decomposition, termed a *canonical Schur decomposition*, satisfying (i) $R \in \mathbb{SO}_n$ and (ii)

$$\pi \geq \theta_1 \geq \cdots \geq \theta_{k-1} \geq |\theta_k| \geq 0 \tag{4}$$

KEY RESULT: Characterization of D-components.

For $\xi = (x_1, \ldots, x_k) \in \mathbb{Z}^k$, denote

$$\mathcal{M}_\xi := \left\{ RAR^\mathsf{T} : \begin{array}{l} R \in \mathbb{SO}_n, A = \operatorname{diag}\left(\begin{bmatrix} 0 & -\alpha_1 \\ \alpha_1 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & -\alpha_k \\ \alpha_k & 0 \end{bmatrix} \right); \\ \alpha_i = \theta_i + 2\pi x_i,\ \theta \text{ satisfies (4) and } |\theta_2| < \pi; \\ \text{if } x_i \neq \pm x_{i+1}, \text{ then } \theta_i \neq \pm\theta_{i+1} \text{ for } i \leq k-1 \end{array} \right\}.$$

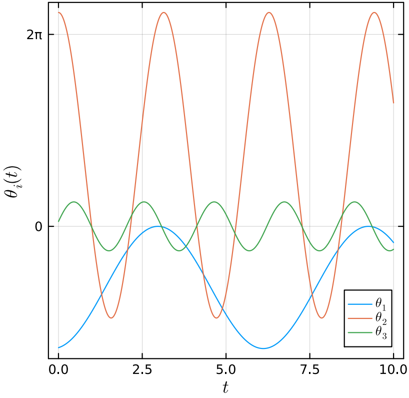
Then, $\mathcal{M}_* = \mathcal{M}_{(0,\ldots,0)} \cup \mathcal{M}_{(-1,0,\ldots,0)}$, and M_ξ is a D-component for $\xi \neq (0, \ldots, 0), (-1, 0, \ldots, 0)$.

Characterization of D-Component

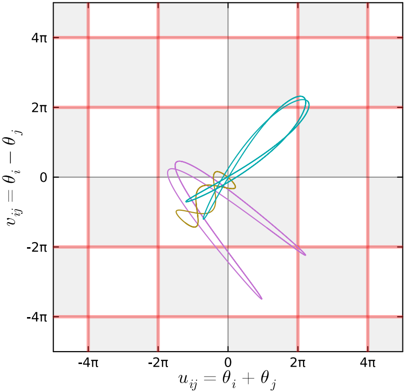
- Fully characterize $\mathfrak{S}$;
- Fully characterize all D-component $\mathcal{M}_e$, $\forall e \in \mathcal{E}$;
- $\mathfrak{S}$ is a zero-measure set of $\mathbf{Skew}_n$;
- $\mathbf{Skew}_n / \mathfrak{S} = \cup_{e \in \mathcal{E}} \mathcal{M}_e$;
- Given a matrix $Q \in \mathbb{SO}_n$, does $\mathcal{M}_e$ only contain one preimage of Q for all $e \in \mathcal{E}$?
  If yes, then one could choose a preimage of Q based on the selection of the index e!

Consider the sum and difference of a pair of angles (θ_i, θ_j) denoted by

$$(u_{ij} = \theta_i + \theta_j, v_{ij} = \theta_i - \theta_j) \quad \text{for } i < j.$$



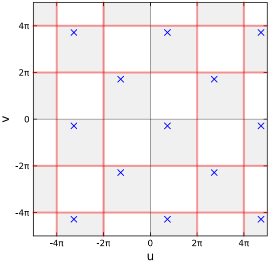
(a) Example of $\theta_1, \theta_2, \theta_3 : [0, 1] \to \mathbb{R}$.



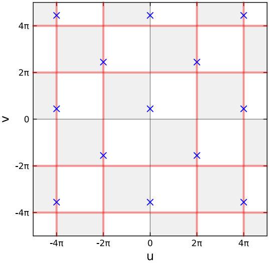
(b) $(u_{ij}(t), v_{ij}(t))$ for $i < j \leq 3$.

Preimage Distribution on the (u, v)-Chessboard

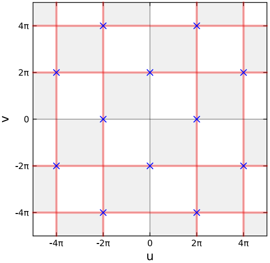
Let (θ_i, θ_j) generate the scatters with $(\theta_i + 2\pi\ell_i, \theta_j + 2\pi\ell_j)$, $\ell_i, \ell_j \in \mathbb{Z}$:



Pattern (a).



Pattern (b).



Pattern (c).

Remark: D-Connected Preimages

If X and Y are distinct D-connected preimages of Q, at least one of trajectories connects distinct blue crosses within a red box, which only occurs in pattern (a).

Distribution of Preimages among D-Components

KEY RESULT: Distribution of preimages.

Consider the preimages $\mathsf{exp}^{-1}(Q) := \{X \in \mathbf{Skew}_n : \mathsf{exp}(X) = Q\}$ of any $Q \in \mathbb{SO}_n$ and a D-component $\mathcal{M}_e$:

- if $\mathcal{M}_e = \mathcal{M}_*$, then $\mathcal{M}_e \cap \mathsf{exp}^{-1}(Q)$ has at most **two** preimages;
- if $\mathcal{M}_e \neq \mathcal{M}_*$, then $\mathcal{M}_e \cap \mathsf{exp}^{-1}(Q)$ has at most **one** preimage.

Characterization of D-Connected Preimages

Characterization of D-Connected preimages.

Consider a Schur decomposition of $Q \in \mathbb{SO}_n$, where $n = 2k$,

$$Q = \sum_{i=1}^{k} \begin{bmatrix} R_{2i-1} & R_{2i} \end{bmatrix} \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \begin{bmatrix} R_{2i-1} & R_{2i} \end{bmatrix}^{\mathsf{T}},$$

where $\pi \geq |\theta_1| \geq |\theta_2| \geq \ldots \geq |\theta_k|$. If Q has two distinct D-connected preimages X and Y, then $|\theta_1| > |\theta_2|$. Moreover, X and Y are in $\mathcal{M}_*$ in the forms of

$$\begin{cases} X = \sum_{i=1}^{k} \begin{bmatrix} R_{2i-1} & R_{2i} \end{bmatrix} \begin{bmatrix} 0 & -\theta_i \\ \theta_i & 0 \end{bmatrix} \begin{bmatrix} R_{2i-1} & R_{2i} \end{bmatrix}^{\mathsf{T}} \\ Y = X - \begin{bmatrix} R_1 & R_2 \end{bmatrix} \begin{bmatrix} 0 & -2\pi \\ 2\pi & 0 \end{bmatrix} \begin{bmatrix} R_1 & R_2 \end{bmatrix}^{\mathsf{T}}. \end{cases}$$

Note that X is in the closure of the principal branch.

Characterization of D-Connected Preimages

- Fully characterize $\mathfrak{S}$;
- Fully characterize all D-component $\mathcal{M}_e$, $\forall e \in \mathcal{E}$;
- $\mathfrak{S}$ is a zero-measure set of $\mathbf{Skew}_n$;
- $\mathbf{Skew}_n / \mathfrak{S} = \cup_{e \in \mathcal{E}} \mathcal{M}_e$;
- Given a matrix $Q \in \mathbb{SO}_n$, $\mathcal{M}_e$ may contain zero/one/two preimages of Q;

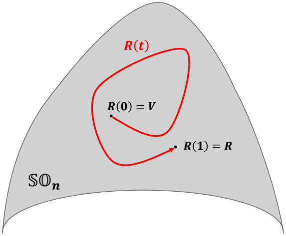
How dense is $\exp(\mathcal{M}_e)$ in $\mathbb{SO}_n$?

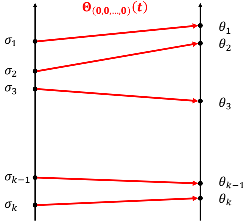
Path-Connected Image in $\mathbb{SO}_n$

For what kind of $P, Q \in \mathbb{SO}_n$,
we always can construct a D-curve connecting them?

Schur bases $R, V \in \mathbb{SO}_n$ of P and Q respectively; angles σ, θ of P and Q respectively satisfies $\pi \geq \theta_1(\sigma_1) \geq \cdots \geq \theta_{k-1}(\sigma_{k-1}) \geq |\theta_k|(|\sigma_k|) \geq 0$:

- construct $R(t)$ connects V and R in $\mathbb{SO}_n$;
- construct $\Theta_\xi(t) := \operatorname{diag}\left(\begin{bmatrix} 0 & -\theta_1(t) \\ \theta_1(t) & 0 \end{bmatrix}, \ldots\right)$ where $\xi = (x_1, \ldots, x_k) \in \mathbb{Z}^k$ and $\theta_i(t) = 2\pi x_i + (1-t)\sigma_i + t\theta_i$ for $i \leq m$.





Path-Connected Image in $\mathbb{SO}_n$

For what kind of $P, Q \in \mathbb{SO}_n$,
we always can construct a D-curve connecting them?

Remark: Invariant integer shift along a D-curve.

Any $P, Q \in \mathbb{SO}_n$ can be connected by a curve in the form of $Q(t) = R(t)\exp(\Theta_\xi(t))R(t)^\mathsf{T}$. Moreover, if $x_i = \pm x_{i+1}$ holds for $\theta_i(t) = \pm\theta_{i+1}(t)$, then $S_\xi(t) = R(t)\Theta_\xi(t)R(t)^\mathsf{T}$ is a D-curve.

Proposition: Closure of D-components (guaranteed fallback).

For $\xi = (x_1, \ldots, x_k) \in \mathbb{Z}^k$, the closure of $\mathcal{M}_\xi$ is

$$\mathrm{closure}(\mathcal{M}_\xi) = \left\{ RAR^\mathsf{T} : \begin{array}{l} R \in \mathbb{SO}_n, A = \mathrm{diag}\left(\begin{bmatrix} 0 & -\alpha_1 \\ \alpha_1 & 0 \end{bmatrix}, \ldots \right); \\ \alpha_i = \theta_i + 2\pi x_i,\ \theta \text{ satisfies } (4.6) \end{array} \right\},$$

whose image is $\mathbb{SO}_n$, i.e., $\{\exp(X) : X \in \mathrm{closure}(\mathcal{M}_\xi)\} = \mathbb{SO}_n$.

Characterization of the density of $\exp(\mathcal{M}_e)$

- Fully characterize $\mathfrak{S}$;
- Fully characterize all D-component $\mathcal{M}_e$, $\forall e \in \mathcal{E}$;
- $\mathfrak{S}$ is a zero-measure set of $\mathbf{Skew}_n$;
- $\mathbf{Skew}_n / \mathfrak{S} = \cup_{e \in \mathcal{E}} \mathcal{M}_e$;
- $\exp(\mathcal{M}_e)$ is dense in $\mathbb{SO}_n$;
- A preimage of Q always can be found in the closure of $\mathcal{M}_e$ for any $e \in \mathcal{E}$;

Outline

Diffeomorphic Logarithm

Definition: Diffeomorphic Logarithm.

For $S \in \mathcal{M}_e$ and $Q \in \mathbb{SO}_n$, the diffeomorphic logarithm of Q with S, denoted by $\operatorname{dlog}_S(Q)$ is given by

- the only preimage in $\mathcal{M}_e$, when $S \notin \mathcal{M}_*$;
- the nearest preimage in Frobenius norm in $\mathcal{M}_*$, when $S \in \mathcal{M}_*$.

Moreover, the nearest candidates $X \in \mathcal{M}_{(0,\ldots,0)}$ and $Y \in \mathcal{M}_{(-1,\ldots,0)}$ is determined by the largest θ_1 and the Schur vectors R_1 and R_2:

$$\|X - S\|_{\mathrm{F}} \leq \|Y - S\|_{\mathrm{F}} \Leftrightarrow |R_2^{\top} S R_1 - \theta_1| \leq |R_2^{\top} S R_1 - \theta_1 + 2\pi|.$$

- $S \in \mathcal{M}_e$: a reference point;
- $Q \in \mathbb{SO}_n$: a preimage;
- the index e: specified by S;
- a preimage of Q: in closure$(\mathcal{M}_e)$ and closest to S in F-norm;

Diffeomorphic Logarithm

Algorithm 1 Diffeomorphic Logarithm

Input: $Q \in \mathbb{SO}_n$, $S \in \textbf{Skew}_n \setminus \mathfrak{S}$ and (optional) $\xi = (x_1, \ldots, x_k) \in \mathbb{Z}^k$ of $\mathcal{M}_\xi \ni S$.

Output: The diffeomorphic logarithm X of Q with S, or its guaranteed fallback.

$(R, \theta \in (-\pi, \pi]^k) \xleftarrow[R \in \mathbb{SO}_n,\ \theta \text{ subjected to (4)}]{\text{Canonical Schur Decomposition}} Q$

if ξ is not provided **then**

    $(\sim, \sigma + 2\pi\xi \in \mathbb{R}^k) \xleftarrow[\sigma \in (-\pi, \pi]^k \text{ subjected to (4)}]{\text{Canonical Schur Decomposition}} S$ `// Identify` $\mathcal{M}_\xi \ni S$

end if

if $\xi = (0, 0, \ldots, 0)$ or $\xi = (-1, 0, \ldots, 0)$, i.e., $S \in \mathcal{M}_*$ **then**

    if $|R_2^\mathsf{T} S R_1 - \theta_1| \leq |R_2^\mathsf{T} S R_1 - \theta_1 + 2\pi|$ **then**

        $\xi \leftarrow (0, 0, \ldots, 0)$

    else

        $\xi \leftarrow (-1, 0, \ldots, 0)$

    end if

end if

return $X \leftarrow \sum_{i=1}^m R_{[i]} \begin{bmatrix} 0 & -\theta_i - 2\pi x_i \\ \theta_i + 2\pi x_i & 0 \end{bmatrix} R_{[i]}^\mathsf{T}$ `// Logarithm with shifted angles.`

Review: Pullback of the Karcher Mean Problem

Review: Relaxed Karcher Mean of the Pullback Optimization.

Consider a set of Riemannian geodesics $Q_i \exp(tS_i), S_i \in \mathbf{Skew}_n$ that arrive at the same endpoint $Q_i \exp(S_i) = Q_j \exp(S_j)$ for $i, j \le K$. Denote

$$\mathcal{S}(Q_1, \ldots, Q_K) := \left\{ (S_1, \ldots, S_K) \in (\mathbf{Skew}_n)^K : Q_i \exp(S_i) = Q_j \exp(S_j) \right\}.$$

A relaxed Karcher mean $\overline{Q}$ is the shared endpoint $Q_i \exp(S_i), \forall i \le K$ of a local minimizer in

$$\operatorname*{arg\,min}_{(S_1,\ldots,S_K)\in\mathcal{S}(Q_1,\ldots,Q_K)} \mathcal{J}(S_1, \ldots, S_K) := \frac{1}{4K} \sum_{i=1}^{K} \|S_i\|_{\mathrm{F}}^2.$$

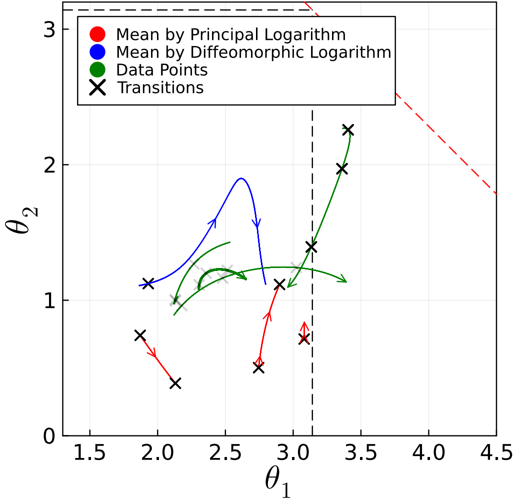
Algorithm 2 Relaxed Karcher mean

Input: Data $\{Q_i\}_{i=1}^K \subset \mathbb{SO}_n$, initial $X, \{S_i\}_{i=1}^K \in \mathbf{Skew}_n$ where $Q_i \exp(S_i) = \exp(X)$, step size $\alpha > 0$, tolerance $\mathrm{AbsTol} > 0$.

Output: A Relaxed Mean $M = \exp(X) \in \mathbb{SO}_n$.

$M \leftarrow \exp(X)$
for $i = 1, \ldots, K$ **do**
  $S_i \leftarrow \log(Q_i^\top M)$ or $S_i \leftarrow \mathrm{dlog}_{S_i}(Q_i^\top M)$ `// log: principal log; dlog: diffeomorphic log.`
end for
$\Delta \leftarrow \frac{1}{2K} \sum_{i=1}^K S_i$
while $\|\Delta\|_{\mathrm{F}} > \mathrm{AbsTol}$ **do**
  $M \leftarrow M \exp(\alpha \Delta)$
  $X \leftarrow \log(M)$ or $X \leftarrow \mathrm{dlog}_X(M)$
  for $i = 1, \ldots, K$ **do**
    $S_i \leftarrow \log(Q_i^\top M)$ or $S_i \leftarrow \mathrm{dlog}_{S_i}(Q_i^\top M)$
  end for
  $\Delta \leftarrow \frac{1}{2K} \sum_{i=1}^K S_i$
end while
return M and X

The Relaxed Karcher Mean



The relaxed Karcher means computed by principal or diffeomorphic logarithm.

Summary

This work exploits a local diffeomorphism structure on $\mathbb{SO}_n$, attained by the *diffeomorphic logarithm*, which provides a *continuous inverse* of exp *beyond the principal branch*.

It enables skew-symmetric formulations of special orthogonal constrained problems in a less restrictive domain, including

- the Karcher mean problem in $\mathbb{SO}_n$;
- the interpolation problem in $\mathbb{SO}_n$;
- the Stiefel geodesic problem, and etc.

Thank You!

References I

[ATV13] Bijan Afsari, Roberto Tron, and René Vidal. On the Convergence of Gradient Descent for Finding the Riemannian Center of Mass. *SIAM Journal on Control and Optimization*, 51(3):2230–2260, 2013.

[DAGH25] Zhifeng Deng, P. A. Absil, Kyle A. Gallivan, and Wen Huang. The Exponential of Skew-Symmetric Matrices: A Nearby Inverse and Efficient Computation of Derivatives, 2025.

[DRW+20] Frank Dellaert, David M. Rosen, Jing Wu, Robert Mahony, and Luca Carlone. Shonan Rotation Averaging: Global Optimality by Surfing SO(p)n. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI*, page 292–308, Berlin, Heidelberg, 2020. Springer-Verlag.

[KHM07] Krzysztof Krakowski, Knut Hüper, and J Manton. On the Computation of the Karcher Mean on Spheres and Special Orthogonal Groups. In *Conference Paper, Robomat*. Citeseer, Coimbra, Portugal, 2007.

[Moa02] Maher Moakher. Means and Averaging in the Group of Rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1–16, 2002.

References II

[TVSC11] Pavan Turaga, Ashok Veeraraghavan, Anuj Srivastava, and Rama Chellappa. Statistical Computations on Grassmann and Stiefel Manifolds for Image and Video-Based Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11):2273–2286, 2011.

[Wil67] Ralph M Wilcox. Exponential Operators and Parameter Differentiation in Quantum Physics. *Journal of Mathematical Physics*, 8(4):962–982, 1967.

[ZBL+19] Yi Zhou, Connelly Barnes, Jingwan Lu, Jimei Yang, and Hao Li. On the Continuity of Rotation Representations in Neural Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5745–5753, 2019.