Computing geodesics on the Stiefel manifold

Marco Sutti

Postdoctoral fellow at NCTS

Sun Yat-sen University, Shenzhen Campus

中山大学深圳校区

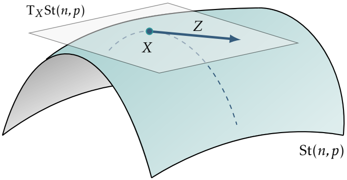
June 18, 2025

Shooting method for computing geodesics on Stiefel/1

Paper: A single shooting method with approximate Fréchet derivative for computing geodesics on the Stiefel manifold, M. Sutti, Electron. Trans. Numer. Anal. (ETNA), Vol. 60, 501–519, September 2024.

- Many applications in diverse fields deal with data belonging to the Stiefel manifold

  $$\mathrm{St}(n,p) = \{X \in \mathbb{R}^{n\times p} : \; X^\top X = I_p\}.$$



- Evaluation of the distance between two points on $\mathrm{St}(n,p)$.
- No closed-form solution is known for $\mathrm{St}(n,p)$!

Shooting method for computing geodesics on Stiefel/2

This talk:

I. Motivating examples: shape analysis and interpolation on manifolds.

II. Geometry of the Stiefel manifold.

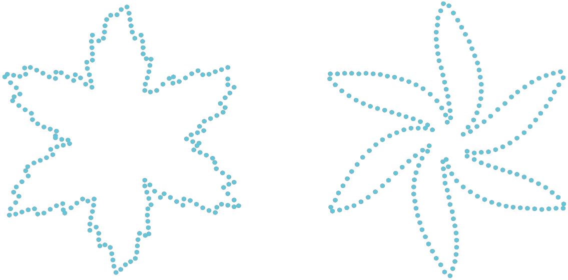
III. Computational framework based on the classical shooting method for BVPs, with an approximate formula for the Fréchet derivative of the geodesic involved.

IV. Numerical experiments show that the algorithm is competitive with other state-of-the-art methods.

I. Motivation

A motivating example: imaging/1

- Need to deal with transformations that are more complicated than similarity transformations (translation/rotation/scaling).
- E.g., distortion, or imaging the same scene from different viewing angles.
- Example: two shapes from the MPEG-7 dataset, with a certain degree of similarity.

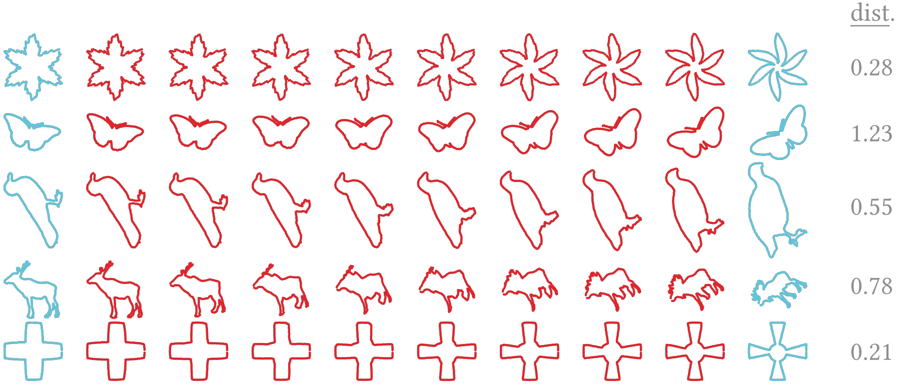


$\rightsquigarrow$ How "far" are they from each other?

MPEG-7: [Bober 2001], affine-standardized shapes: [Bryner 2017]

A motivating example: imaging/2

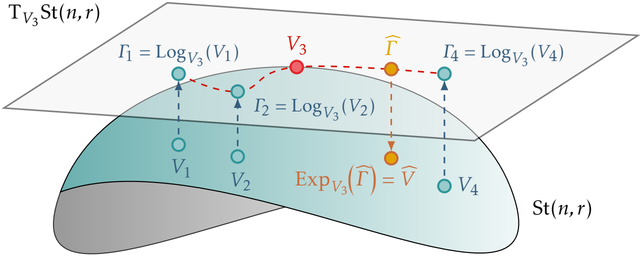
- One usually goes beyond the similarity group to define shape equivalences.
- Geodesics on $\mathrm{St}(n, 2)$, with shapes from the MPEG-7 dataset.



MPEG-7: [Bober 2001], affine-standardized shapes: [Bryner 2017]

A motivating example: interpolation on manifolds

- Model order reduction (MOR) for dynamical systems parametrized according to $p = [p_1, \dots, p_d]^\top$.
- Suppose we have a set of local orthonormal basis matrices $\{V_1, V_2, \dots, V_K\}$.
- Given a new parameter value $\hat{p}$, a basis $\widehat{V}$ can be obtained by **interpolating** the local basis matrices on a tangent space to $\mathrm{St}(n, r)$.
- For interpolation on $\mathrm{T}_{V_3}\mathrm{St}(n, r)$, the distance is needed.



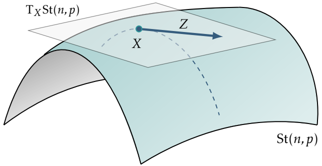
MOR, POD: [Benner/Gugercin/Willcox 2015]
Interpolation on manifolds: [Hüper/Silva Leite 2007, Amsallem 2010, Amsallem/Farhat 2011]

II. The Stiefel manifold

The Stiefel manifold and its tangent space

- Set of matrices with orthonormal columns:

$$\mathrm{St}(n,p) = \{X \in \mathbb{R}^{n\times p} :\ X^\top X = I_p\}.$$



- Tangent space to $\mathcal{M}$ at x: set of all tangent vectors to $\mathcal{M}$ at x, denoted $\mathrm{T}_x\mathcal{M}$. For $\mathrm{St}(n,p)$,

$$\mathrm{T}_X\mathrm{St}(n,p) = \{\xi \in \mathbb{R}^{n\times p} :\ X^\top \xi + \xi^\top X = 0\}.$$

- Alternative characterization of $\mathrm{T}_X\mathrm{St}(n,p)$:

$$\mathrm{T}_X\mathrm{St}(n,p) = \{X\Omega + X_\perp K :\ \Omega = -\Omega^\top,\ K \in \mathbb{R}^{(n-p)\times p}\},$$

where $\mathrm{span}(X_\perp) = \big(\mathrm{span}(X)\big)^\perp$. An orthogonal completion of $X \in \mathbb{R}^{n\times p}$ is $X_\perp \in \mathbb{R}^{n\times(n-p)}$ s.t. $[X\ X_\perp] \in \mathrm{O}(n)$.

- The projection onto the tangent space $\mathrm{T}_X\mathrm{St}(n,p)$ is

$$\mathrm{P}_X\,\xi = X\mathrm{skew}(X^\top \xi) + (I - XX^\top)\,\xi.$$

Stiefel manifold: [Stiefel, 1935]

Riemannian manifold

A manifold $\mathcal{M}$ endowed with a smoothly-varying inner product (called Riemannian metric g) is called Riemannian manifold.

$\rightsquigarrow$ A couple $(\mathcal{M}, g)$, i.e., a manifold with a Riemannian metric on it.

$\rightsquigarrow$ For the Stiefel manifold:

- Embedded metric inherited by $\mathrm{T}_X\mathrm{St}(n,p)$ from the embedding space $\mathbb{R}^{n\times p}$

$$\langle \xi, \eta \rangle = \mathrm{Tr}(\xi^\top \eta), \qquad \xi, \eta \in \mathrm{T}_X\mathrm{St}(n,p).$$

- Canonical metric by seeing $\mathrm{St}(n,p)$ as a quotient of the orthogonal group $\mathrm{O}(n)$: $\mathrm{St}(n,p) = \mathrm{O}(n)/\mathrm{O}(n-p)$

$$\langle \xi, \eta \rangle_\mathrm{c} = \mathrm{Tr}(\xi^\top (I - \tfrac{1}{2} X X^\top) \eta), \qquad \xi, \eta \in \mathrm{T}_X\mathrm{St}(n,p).$$

Metrics and geodesics on $\mathrm{St}(n,p)$

Embedded metric:

$$\langle \xi, \eta \rangle = \mathrm{Tr}(\xi^\top \eta).$$

Canonical metric:

$$\langle \xi, \eta \rangle_{\mathrm{c}} = \mathrm{Tr}(\xi^\top (I - \tfrac{1}{2} X X^\top) \eta).$$

Length of a tangent vector $\xi = X\Omega + X_\perp K$:

$$\|\xi\|_{\mathrm{F}} = \sqrt{\langle \xi, \xi \rangle} = \sqrt{\|\Omega\|_{\mathrm{F}}^2 + \|K\|_{\mathrm{F}}^2}.$$

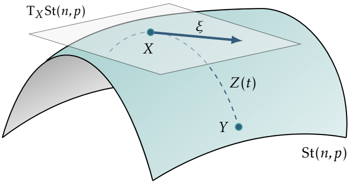
$$\|\xi\|_{\mathrm{c}} = \sqrt{\langle \xi, \xi \rangle_{\mathrm{c}}} = \sqrt{\tfrac{1}{2}\|\Omega\|_{\mathrm{F}}^2 + \|K\|_{\mathrm{F}}^2}.$$

- Closed-form solution (with the canonical metric) for a geodesic $Z(t)$ that realizes ξ with base point X:

$$Z(t) = Q \exp_{\mathrm{m}}(A(\xi)t)\, I_{n,p},$$

where $Q := [X \;\; X_\perp]$, $A(\xi) := \begin{bmatrix} X^\top \xi & -(X_\perp^\top \xi)^\top \\ X_\perp^\top \xi & O \end{bmatrix}$,

and $I_{n,p} := [I_p \;\; O_{p\times(n-p)}]^\top$.



Riemannian exponential and logarithm

- Given $x \in \mathcal{M}$ and $\xi \in T_x\mathcal{M}$, the exponential mapping $\mathrm{Exp}_x : T_x\mathcal{M} \to \mathcal{M}$ s.t. $\mathrm{Exp}_x(\xi) := \gamma(1)$, with γ being the geodesic with $\gamma(0) = x$, $\dot{\gamma}(0) = \xi$.
- Corollary: $\mathrm{Exp}_x(t\xi) := \gamma(t)$, for $t \in [0, 1]$.
- $\forall x,\ y \in \mathcal{M}$, the mapping $\mathrm{Exp}_x^{-1}(y) \in T_x\mathcal{M}$ is called logarithm mapping.

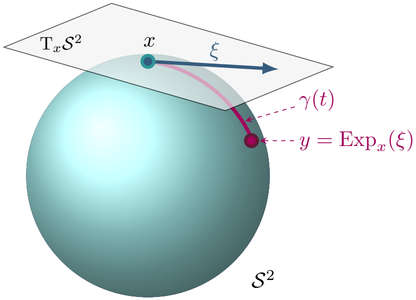
Example. Let $\mathcal{M} = S^{n-1}$, then the exponential mapping at $x \in S^{n-1}$ is

$$y = \mathrm{Exp}_x(\xi) = x\cos(\|\xi\|) + \frac{\xi}{\|\xi\|}\sin(\|\xi\|),$$

and the Riemannian logarithm is

$$\mathrm{Log}_x(y) = \xi = \arccos(x^\top y)\frac{\mathrm{P}_x y}{\|\mathrm{P}_x y\|},$$

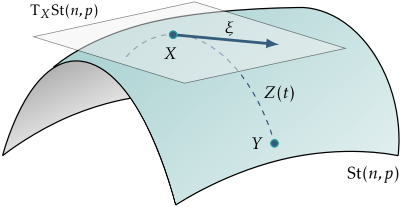
where $y \equiv \gamma(1)$ and P_x is the projector onto $\left(\mathrm{span}(x)\right)^\perp$, i.e., $\mathrm{P}_x = I - xx^\top$.



Riemannian distance on $\mathrm{St}(n,p)$

- Property: Given $X, Y \in \mathrm{St}(n,p)$, s.t. $\mathrm{Exp}_X(\xi) = Y$, the Riemannian distance $d(X,Y)$ equals the length of $\xi \equiv \dot{Z}(0) \in \mathrm{T}_X\mathrm{St}(n,p)$:

$$d(X,Y) = \|\xi\|_{\mathrm{c}} = \sqrt{\langle \xi, \xi \rangle_{\mathrm{c}}}.$$



Equivalent to: Compute the length of the Riemannian logarithm of Y with base point X, i.e.,

$$\mathrm{Log}_X(Y) = \xi.$$

- No closed-form solution is known for $\mathrm{St}(n,p)$!

$\rightsquigarrow$ How do we compute $d(X,Y)$ in practice / numerically?

III. The shooting method

Single shooting for BVPs

- Boundary value problem (BVP): Find $w(x)\colon [a,b] \to \mathbb{R}$ that satisfies

$$w'' = f(x, w, w'), \quad \text{with BCs} \quad \begin{cases} w(a) = \alpha, \\ w(b) = \beta. \end{cases}$$

- Recast it as an initial value problem (IVP): Find $w(x)$ that satisfies

$$w'' = f(x, w, w'), \quad \text{with ICs} \quad \begin{cases} w(a) = \alpha, \\ w'(a) = s. \end{cases}$$

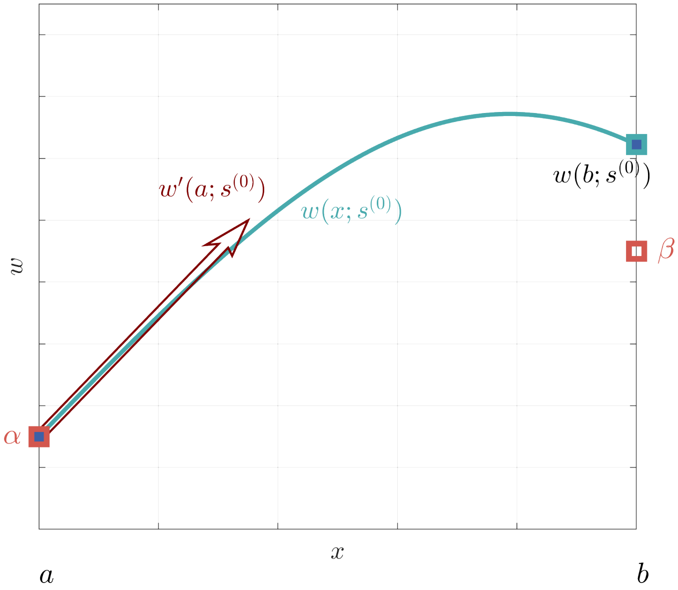
$\rightsquigarrow$ In general, this has a unique solution $w(x) \equiv w(x; s)$ which depends on s (Picard–Lindelöf theorem). Analytical or numerical solution (e.g., Runge–Kutta).

$\rightsquigarrow$ Single shooting method for BVPs:

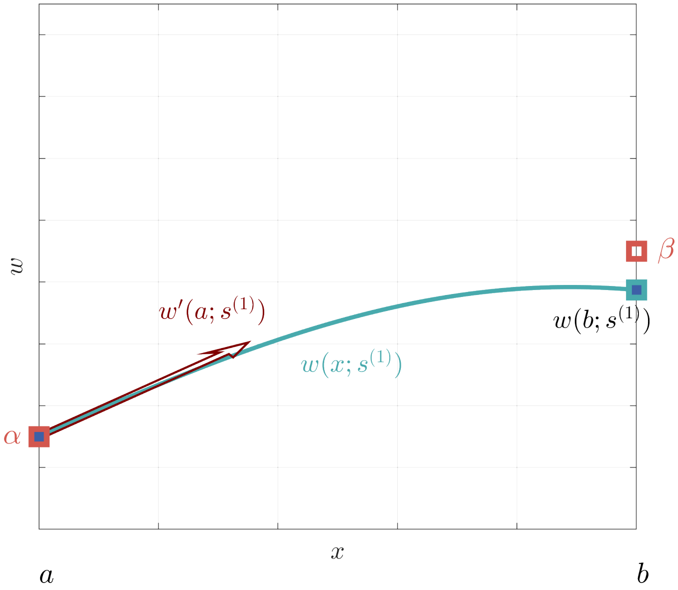
- Define $F(s) = w(b; s) - \beta$.
- Find $\bar{s}$ s.t. $F(\bar{s}) = 0$. Usually, with Newton's method.

BVPs and shooting methods: see, e.g., [Stoer/Bulirsch 1991]

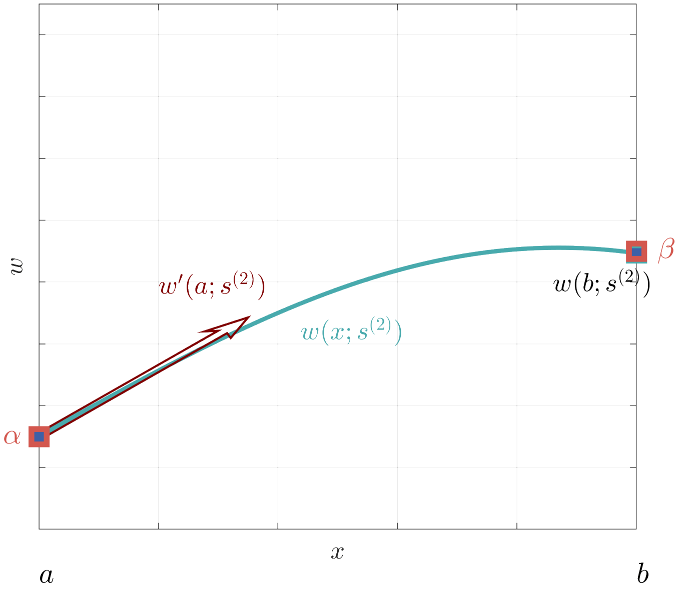
Single shooting for BVPs: example



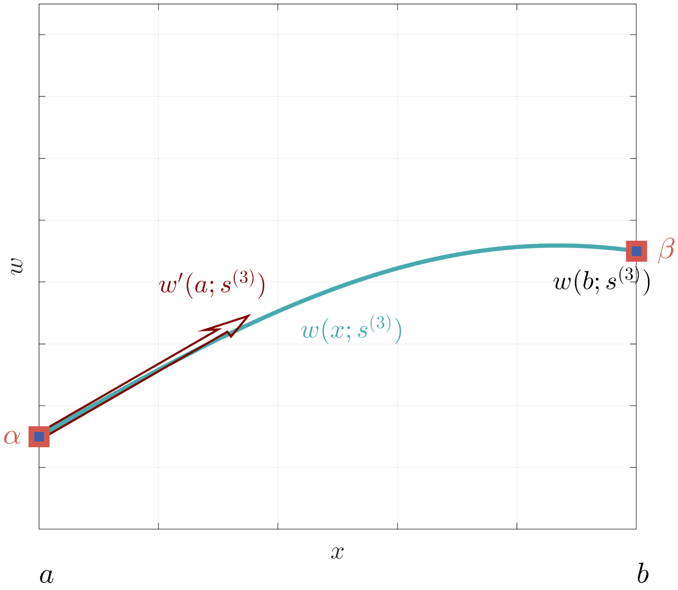
Single shooting for BVPs: example



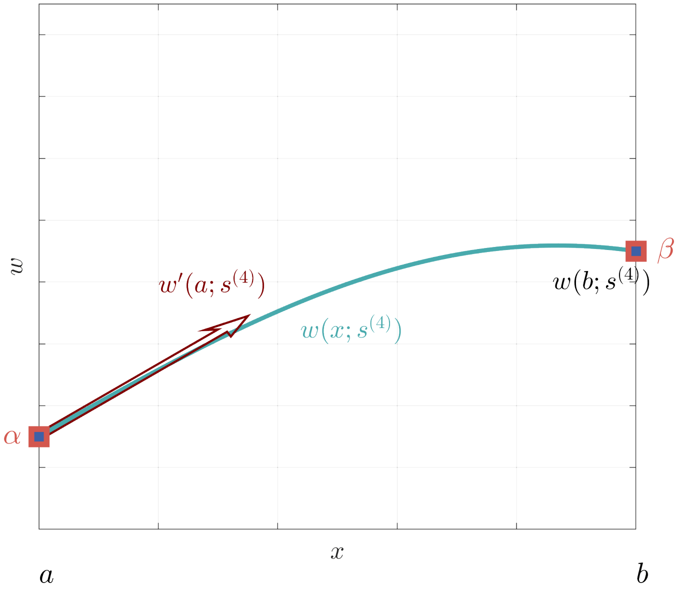
Single shooting for BVPs: example



Single shooting for BVPs: example



Single shooting for BVPs: example

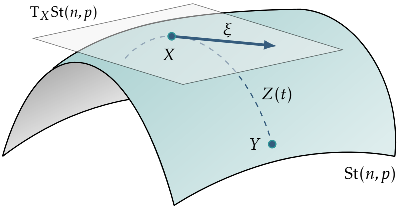


Stiefel geodesics via single shooting

- Problem statement:
  Find $\xi \equiv \dot{Z}(0) \in \mathrm{T}_X\mathrm{St}(n,p)$
  that satisfies the BVP

$$\ddot{Z} = -\dot{Z}\dot{Z}^\top Z - Z((Z^\top \dot{Z})^2 + \dot{Z}^\top \dot{Z}),$$

$$\text{with BCs} \quad \begin{cases} Z(0) = X, \\ Z(1) = Y. \end{cases}$$



- Recall: we have the explicit solution: $Z(t) = Q\,\exp_{\mathrm{m}}(A(\xi)t)\,I_{n,p}$, where
  $Q := [\,X\ \ X_\perp],\ A(\xi) := \begin{bmatrix} X^\top\xi & -(X_\perp^\top\xi)^\top \\ X_\perp^\top\xi & O \end{bmatrix}$, and $I_{n,p} := [\,I_p\ \ O_{p\times(n-p)}]^\top$.

$\rightsquigarrow$ Single shooting for Stiefel geodesics:

- Define $F(\xi) = Z_{(t=1,\xi)} - Y$. $\leftsquigarrow$ main difference w.r.t. the previous example!
- Find $\bar{\xi}$ s.t. $F(\bar{\xi}) = 0$ with Newton's method.

Linearization and approximation/1

- Starting from the nonlinear matrix equation

$$F(\xi) = Z_{(t=1,\xi)} - Y,$$

- we linearize it using matrix perturbation theory

$$F(\xi + \delta\xi) = Z(\xi + \delta\xi) - Y = 0,$$

where

$$Z(\xi + \delta\xi) = Z(\xi) + \mathrm{D}Z\,[\delta\xi] + o(\|\delta\xi\|).$$

- Neglecting the higher-order terms in $\delta\xi$, we get

$$Z(\xi) + \mathrm{D}Z\,[\delta\xi] - Y = 0,$$

$$Z(\xi) + Q\,\mathrm{D}\exp_{\mathrm{m}}(A(\xi))\Big[\mathrm{D}A(\xi)[\delta\xi]\Big] \cdot I_{n,p} - Y = 0.$$

This is equation (4.5) in in the paper [S. 2024].

Linearization and approximation/1

- Starting from the nonlinear matrix equation

$$F(\xi) = Z_{(t=1,\xi)} - Y,$$

- we linearize it using matrix perturbation theory

$$F(\xi + \delta\xi) = Z(\xi + \delta\xi) - Y = 0,$$

where

$$Z(\xi + \delta\xi) = Z(\xi) + \mathrm{D}Z[\delta\xi] + o(\|\delta\xi\|).$$

- Neglecting the higher-order terms in $\delta\xi$, we get

$$Z(\xi) + \mathrm{D}Z[\delta\xi] - Y = 0,$$

$$Z(\xi) + Q\,\boxed{\mathrm{D}\exp_{\mathrm{m}}(A(\xi))\Big[\mathrm{D}A(\xi)[\delta\xi]\Big]} \cdot I_{n,p} - Y = 0.$$

This is equation (4.5) in in the paper [S. 2024].

Linearization and approximation/2

- The perturbation of the matrix exponential by a matrix $E \in \mathbb{R}^{n\times n}$ is

$$\exp_{\mathrm{m}}(A+E) = \exp_{\mathrm{m}}(A) + \mathrm{D}\exp_{\mathrm{m}}(A)[E] + o(\|E\|).$$

- Representation for the Fréchet derivative of the matrix exponential

$$\mathrm{D}\exp_{\mathrm{m}}(A)[E] = E + \frac{AE+EA}{2} + \frac{A^2E+AEA+EA^2}{3!} + \cdots.$$

- Approximation: keep only the first two terms in the expansion, i.e.,

$$\mathrm{D}\exp_{\mathrm{m}}(A)[E] \approx E + \frac{AE+EA}{2}.$$

Function of matrices, Fréchet derivative of the matrix exponential: [Higham 2008]

Linearization and approximation/3

- Identifying E with $\mathrm{D}A(\xi)[\delta\xi]$, the formula $\mathrm{D}\exp_{\mathrm{m}}(A)[E] \approx E + \frac{1}{2}(AE + EA)$ can be used to approximate $\mathrm{D}\exp_{\mathrm{m}}(A(\xi))\Big[\mathrm{D}A(\xi)[\delta\xi]\Big]$ in

$$Z(\xi) + Q\,\mathrm{D}\exp_{\mathrm{m}}(A(\xi))\Big[\mathrm{D}A(\xi)[\delta\xi]\Big] \cdot I_{n,p} - Y = 0,$$

  obtaining

$$Q \cdot \Big(\mathrm{D}A(\xi)[\delta\xi] + \frac{1}{2}(A \cdot \mathrm{D}A(\xi)[\delta\xi] + \mathrm{D}A(\xi)[\delta\xi] \cdot A)\Big) \cdot I_{n,p} = Y - Z.$$

- This is now a linear matrix equation to be solved for $\delta\xi$. In practice, we work with the factors $\delta\Omega$ and δK of $\delta\xi$; see Appendix B in [S. 2024].
- We obtain a (small-sized) Sylvester equation which can be efficiently solved with MATLAB's command `lyap` to obtain the update $\delta\xi$.
- Then the tangent vector is updated as

$$\xi^{(k+1)} = \xi^{(k)} + Q \cdot \begin{bmatrix} \delta\Omega^{(k)} \\ \delta K^{(k)} \end{bmatrix}.$$

Algorithm pseudocode

Algorithm 1: A single shooting method on the Stiefel manifold with an approximation of the Fréchet derivative (SSAF method).

Given X, Y;
Result: ξ^* such that $\mathrm{Exp}_X(\xi^*) = Y$.
Compute the initial guess $\xi^{(0)}$;
Set $k = 0$;
while *a stopping criterion is met* **do**
    Compute $F^{(k)} = Z(1, \xi^{(k)}) - Y$;
    Solve $F^{(k)} + \mathrm{D}Z\,[\delta\xi^{(k)}] = 0$ for $\delta\xi^{(k)}$;
    Update $\xi^{(k+1)} \leftarrow \xi^{(k)} + \delta\xi^{(k)}$;
    Project $\xi^{(k+1)}$ onto $\mathrm{T}_X\mathrm{St}(n, p)$: $\xi^{(k+1)} \leftarrow \mathrm{P}_X\left(\xi^{(k+1)}\right)$;
    $k = k + 1$;
end while

- The code was implemented in MATLAB and is freely available on the repository https://github.com/MarcoSutti/SSAF_2024_repo

IV. Numerical experiments

Comparisons with other methods/1

Comparisons with existing algorithms:

- Bryner's method ("shooting", based on discretized parallel transport).
- Zimmermann's method (matrix-algebraic approach, based on the matrix logarithm).

Table: Comparisons on $\mathrm{St}(1500, p)$ with large values of p, for a prescribed $d(X, Y) = 0.5\,\pi$. Results are averaged over 10 pairs of randomly generated endpoints on $\mathrm{St}(1500, p)$.

p	Avg. comput. time (s)			Avg. no. of iterations		
	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**
500	12.09	3.30	**1.89**	3.00	2.00	**4.00**
700	31.27	8.21	**4.56**	3.00	2.00	**4.00**
1000	77.37	20.39	**8.82**	3.00	2.00	**4.00**

Existing algorithms: [Bryner 2017], [Zimmermann 2017]
This is Table 5.2 in the paper [S. 2024].

Comparisons with other methods/2

Table: Comparisons on $\mathrm{St}(n, 2)$, for doubling values of n, for a prescribed $d(X, Y) = 0.5\,\pi$. $T = 20$, tolerance $\tau = 10^{-3}$. Results are averaged over 100 experiments.

n	Avg. comput. time (s)			Avg. no. of iterations		
	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**
10	0.00400	0.00091	**0.00080**	4.08	3.73	**7.77**
20	0.00367	0.00093	**0.00091**	3.85	3.87	**7.35**
40	0.00337	0.00095	**0.00075**	3.49	3.61	**6.96**
80	0.00312	0.00101	**0.00081**	3.30	3.61	**6.90**
160	0.00310	0.00105	**0.00086**	3.15	3.42	**6.86**
320	0.00328	0.00107	**0.00096**	3.02	3.08	**6.86**
640	0.00371	0.00105	**0.00091**	3.00	3.02	**6.89**
1 280	0.00543	0.00104	**0.00100**	3.00	2.72	**6.87**
2 560	0.00856	0.00135	**0.00121**	3.00	2.47	**6.87**
5 120	0.01056	0.00131	**0.00132**	3.00	2.34	**6.93**
10 240	0.01596	0.00144	**0.00141**	3.00	2.12	**6.97**

- As the ratio $p/n \to 0$, solving the endpoint geodesic problem requires fewer iterations. Similar observation in [Nguyen 2022] and [Zimmermann 2017].

This is Table 5.4 in the paper [S. 2024].

Comparisons with other methods/3

Table: Comparisons on $\mathrm{St}(500, p)$, for doubling values of p, for a prescribed $d(X, Y) = 0.5\,\pi$. $T = 20$, tolerance $\tau = 10^{-3}$. Results are averaged over 100 experiments.

p	Avg. comput. time (s)			Avg. no. of iterations		
	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**
2	0.00353	0.00103	**0.00086**	3.01	2.95	**6.78**
4	0.00533	0.00156	**0.00128**	3.00	2.81	**5.28**
8	0.00711	0.00182	**0.00115**	3.00	2.00	**4.08**
16	0.01173	0.00369	**0.00173**	3.00	2.00	**4.00**
32	0.02912	0.01354	**0.00453**	3.00	2.00	**4.00**
64	0.08762	0.03582	**0.01150**	3.00	2.00	**3.00**
128	0.40437	0.10052	**0.05657**	3.00	1.00	**3.00**
256	1.94025	0.47720	**0.25847**	3.00	1.00	**3.00**

This is Table 5.5 in the paper [S. 2024].

Comparisons with other methods/4

Plots of the data from Tables 5.4 and 5.5.

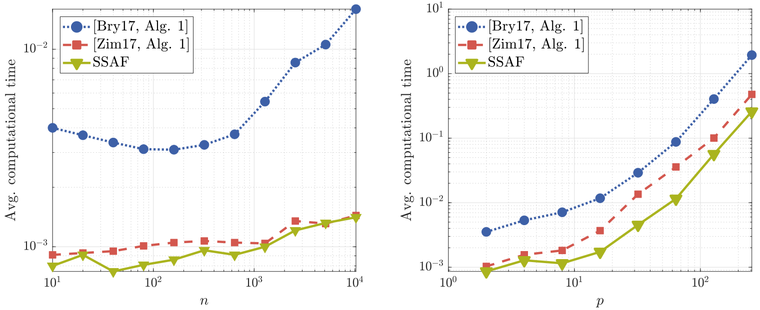


Figure: Average computation times for Bryner's shooting method, Zimmermann's matrix algebraic algorithm, and our SSAF method on $\mathrm{St}(n,p)$.

- Bryner did not use the smaller formulation on $\mathrm{St}(2p,p)$ when $p < n/2$, which makes its algorithm's complexity $\mathcal{O}(Tnp^2)$.
- All the other algorithms considered here use the smaller formulation on $\mathrm{St}(2p,p)$ when possible, and hence they are essentially $\mathcal{O}(p^3)$.

This is Figure 5.1 in the paper [S. 2024].

Comparisons with other methods/5

Experiment for larger value of n, i.e., $n = 1000$, and doubling values of p.

Table: Comparisons on $\mathrm{St}(1000, p)$, for doubling values of p, for a prescribed $d(X, Y) = 0.5\,\pi$. $T = 20$, tolerance $\tau = 10^{-5}$. Results are averaged over 100 experiments.

p	Avg. comput. time (s)			Avg. no. of iterations		
	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**	Bryner, 2017	Zimmermann, 2017	**SSAF, 2024**
20	0.03897	0.00641	**0.00391**	4.00	3.00	**5.02**
40	0.09512	0.02957	**0.01284**	3.00	3.00	**5.00**
80	0.25528	0.08044	**0.03969**	3.00	2.00	**4.00**
160	0.76246	0.24119	**0.13763**	3.00	2.00	**4.00**
320	3.99810	1.07286	**0.64483**	3.00	2.00	**4.00**
640	23.36386	5.62897	**2.80133**	3.00	2.00	**4.00**

- The numerical results demonstrate the competitiveness of our SSAF method in terms of both average computation time and number of iterations w.r.t. the existing algorithms considered here.

This is Table 5.6 in the paper [S. 2024].

Shooting for computing geodesics on Stiefel - summary

Main contributions:

- Computational framework with the classical shooting method for BVPs to compute the Riemannian distance on the Stiefel manifold.
- Computational trick: approximation of the Fréchet derivative of the geodesic.
- Competitiveness and efficiency w.r.t. other state-of-the-art algorithms.

Outlook and ongoing work:

- Explore the connection between shooting algorithms for computing geodesics and domain decomposition methods.
  - Nonlinear Schwarz methods to compute geodesics on manifolds. Joint work with Tommaso Vanzan.
- Target other manifolds that do not have an explicit formula for the distance.

V. Bonus material

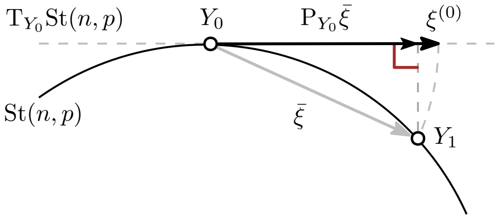
Initial guess

- In Newton's method, selecting a "good enough" initial guess is crucial.
- We use a first-order approximation of $\exp_{\mathrm{m}}(A) \approx I + A$ and solve for ξ.
- This yields a first-order approximation $\bar{\xi}$ to the solution ξ^* as $\bar{\xi} = Y_1 - Y_0$.
- We project $\bar{\xi}$ onto $\mathrm{T}_{Y_0}\mathrm{St}(n,p)$ to obtain a tangent vector:

$$\mathrm{P}_{Y_0}\bar{\xi} = Y_0\,\mathrm{skew}\Big(Y_0^\top(Y_1 - Y_0)\Big) + (I_n - Y_0Y_0^\top)(Y_1 - Y_0) = Y_1 - Y_0\,\mathrm{sym}(Y_0^\top Y_1).$$

- To get $\xi^{(0)}$, we rescale $\mathrm{P}_{Y_0}\bar{\xi}$ so that its norm is equal to the norm of $\bar{\xi}$, i.e.,

$$\xi^{(0)} = \frac{\|\bar{\xi}\|}{\|\mathrm{P}_{Y_0}\bar{\xi}\|}\,\mathrm{P}_{Y_0}\bar{\xi}.$$



Model order reduction/1

- Model order reduction (MOR) for dynamical systems parametrized according to $p = [p_1, \dots, p_d]^\top$.
- For each parameter p_i in a set $\{p_1, p_2, \dots, p_K\}$, use proper orthogonal decomposition (POD) to derive a reduced-order basis $V_i \in \mathrm{St}(n, r)$, $r \ll n$.

$$\begin{cases} \dot{x}(t;p) = A(p)\,x(t;p) + B(p)\,u(t), \\ y(t;p) = C(p)\,x(t;p), \end{cases} \quad \xrightarrow{\text{reduction}} \quad \begin{cases} \dot{x}_r(t;p) = A_r(p)\,x_r(t;p) + B_r(p)\,u(t), \\ y_r(t;p) = C_r(p)\,x_r(t;p), \end{cases}$$

$$x(t;p) \in \mathbb{R}^n,\ \ u(t) \in \mathbb{R}^m,\ y(t) \in \mathbb{R}^q, \qquad x_r = V^\top x,\ A_r = V^\top A V,\ B_r = V^\top B,$$

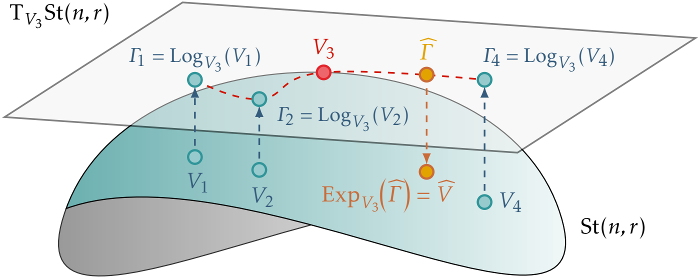
$$A(p) \in \mathbb{R}^{n\times n},\ B(p) \in \mathbb{R}^{n\times m},\ C(p) \in \mathbb{R}^{q\times n}. \qquad C_r = CV,\ \ V \equiv V(p) \in \mathrm{St}(n,r),\ r \ll n.$$

$\rightsquigarrow$ This gives a set of local basis matrices $\{V_1, V_2, \dots, V_K\}$.

MOR, POD: [Benner/Gugercin/Willcox 2015]

Model order reduction/2

- Given a new parameter value $\hat{p}$, a basis $\widehat{V}$ can be obtained by **interpolating** the local basis matrices on a tangent space to $\mathrm{St}(n, r)$.
- For interpolation on $\mathrm{T}_{V_3}\mathrm{St}(n, r)$, the distance is needed.

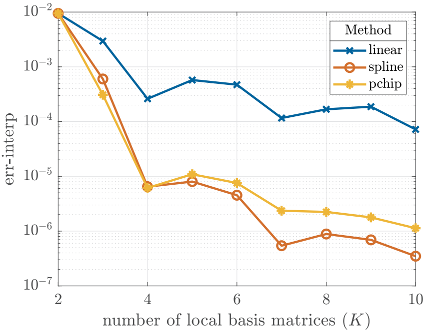


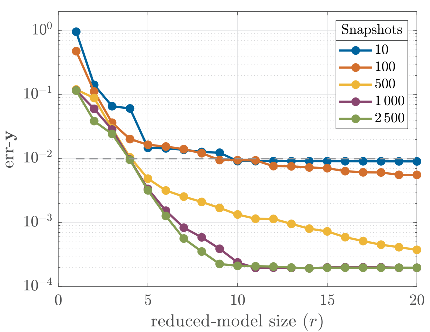
Interpolation in the tangent space to a manifold: [Hüper/Silva Leite 2007, Amsallem 2010, Amsallem/Farhat 2011]

Model order reduction/3

Transient heat equation on a square domain, with 4 disjoint discs.

- FEM discretization with $n = 1169$. Simulation for $t \in [0, 500]$, with $\Delta t = 0.1$.
- 500 snapshot POD over 5000 timeframes, with a reduced model of size $r = 4$.
- Relative error between $y(\cdot; \hat{p})$ and $y_r(\cdot; \hat{p})$ is about 1%.





Details for these experiments: [S. 2023]

Riemannian center of mass

- Notion of mean on a Riemannian manifold $\mathcal{M}$, defined by the optimization problem

$$\mu = \operatorname*{argmin}_{p \in \mathcal{M}} \frac{1}{2N} \sum_{i=1}^{N} d^2(p, q_i),$$

where $d(p, q_i)$ is the Riemannian distance on $\mathcal{M}$, and $q_i \in \mathcal{M}$, for $i = 1, \ldots, N$.

- For $\mathrm{St}(n, p)$, the distances $d(p, q_i)$ are computed with our algorithm.

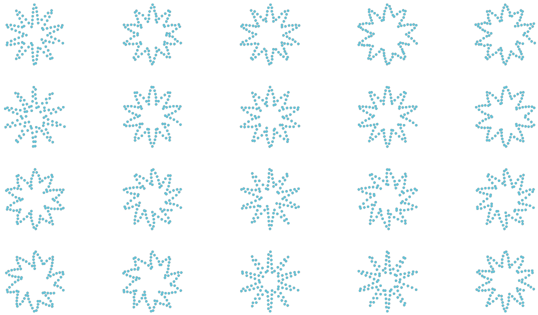
⚠ Caveat: On manifolds of positive curvature the Riemannian center of mass is general not unique. But if the data points are close enough, then uniqueness is guaranteed.

- $\mathrm{St}(n, p)$ has also positive curvature (an upper bound on its sectional curvature is given by 5/4).

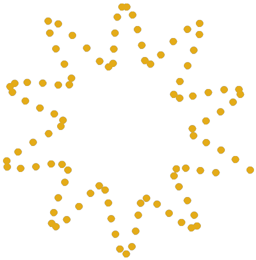
Riemannian center of mass: [Cartan 1920s, Calabi 1958, Grove/Karcher 1973]
Uniqueness of the Riemannian center of mass: [Afsari/Tron/Vidal 2013]
Upper bound on the sectional curvature of $\mathrm{St}(n, p)$: [Rentmeesters 2013]

Riemannian center of mass of a shape set

- "device7" shape set from the MPEG-7 dataset.



- Riemannian center of mass:



MPEG-7: [Bober 2001], affine-standardized shapes: [Bryner 2017]