New Statistical Bound for Inference Problems on Riemannian Manifolds

João Xavier and Victor Barroso

Instituto Sistemas e Robótica (ISR), Instituto Superior Técnico (IST)
Lisbon, Portugal

Outline

- ☐ **Motivation:** Riemannian manifolds in parametric estimation problems

- ☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

- ☐ **Framework**: Parametric statistical models over Riemannian manifolds

- ☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

- ☐ **Applications of IVLB**:
  - ▷ Parametric estimation with constraints
  - ▷ Parametric estimation over quotient spaces

- ☐ **Open problems**

Outline

☐ **Motivation:** Riemannian manifolds in parametric estimation problems

☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

☐ **Framework**: Parametric statistical models over Riemannian manifolds

☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

☐ **Applications of IVLB**:

- ▷ Parametric estimation with constraints
- ▷ Parametric estimation over quotient spaces

☐ **Open problems**

Motivation

☐ **Parameter estimation problem:** Given

▷ parametric family $\mathcal{F} = \{f_\theta \,:\, \theta \in \Theta\}$

▷ observed data vector $y \sim f_{\theta_0}$

find the unknown θ_0 in parameter space Θ

☐ **Key-idea**: Riemannian manifold theory unifies treatment of:

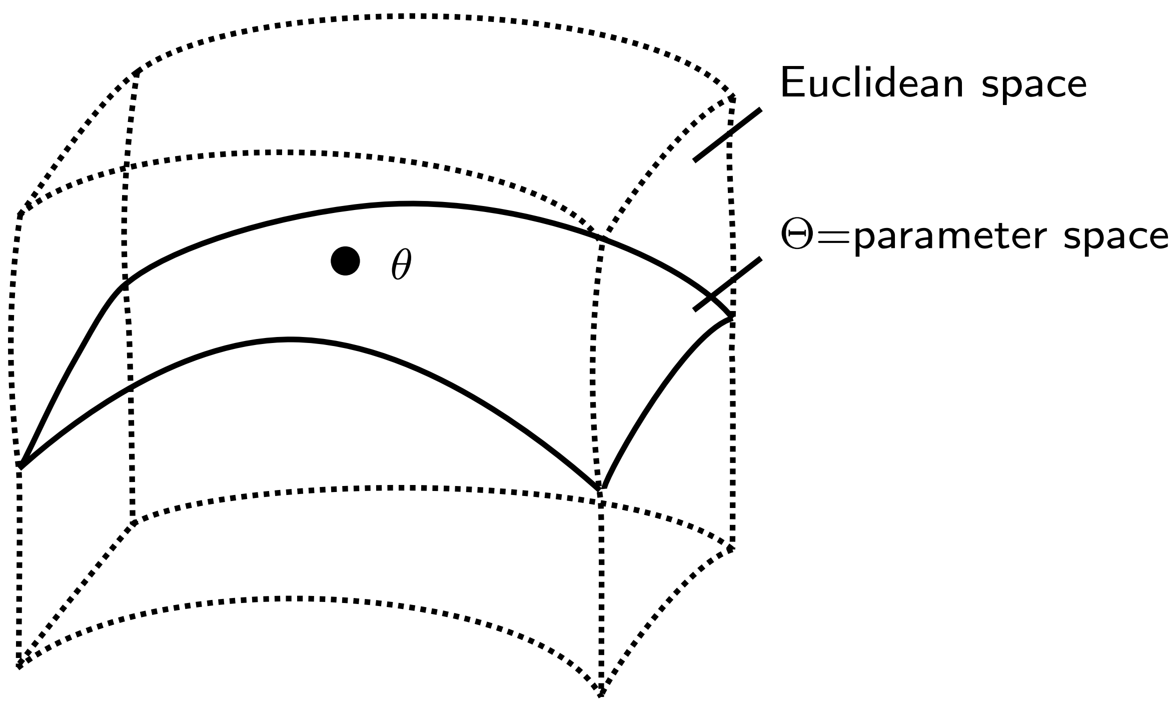
▷ Parametric estimation with constraints

▷ Parametric estimation over quotient spaces

Motivation: Parametric Estimation with Constraints

☐ There are deterministic constraints on parameter θ (e.g. θ is a rotation matrix)

☐ Parameter space Θ becomes a submanifold of an Euclidean space

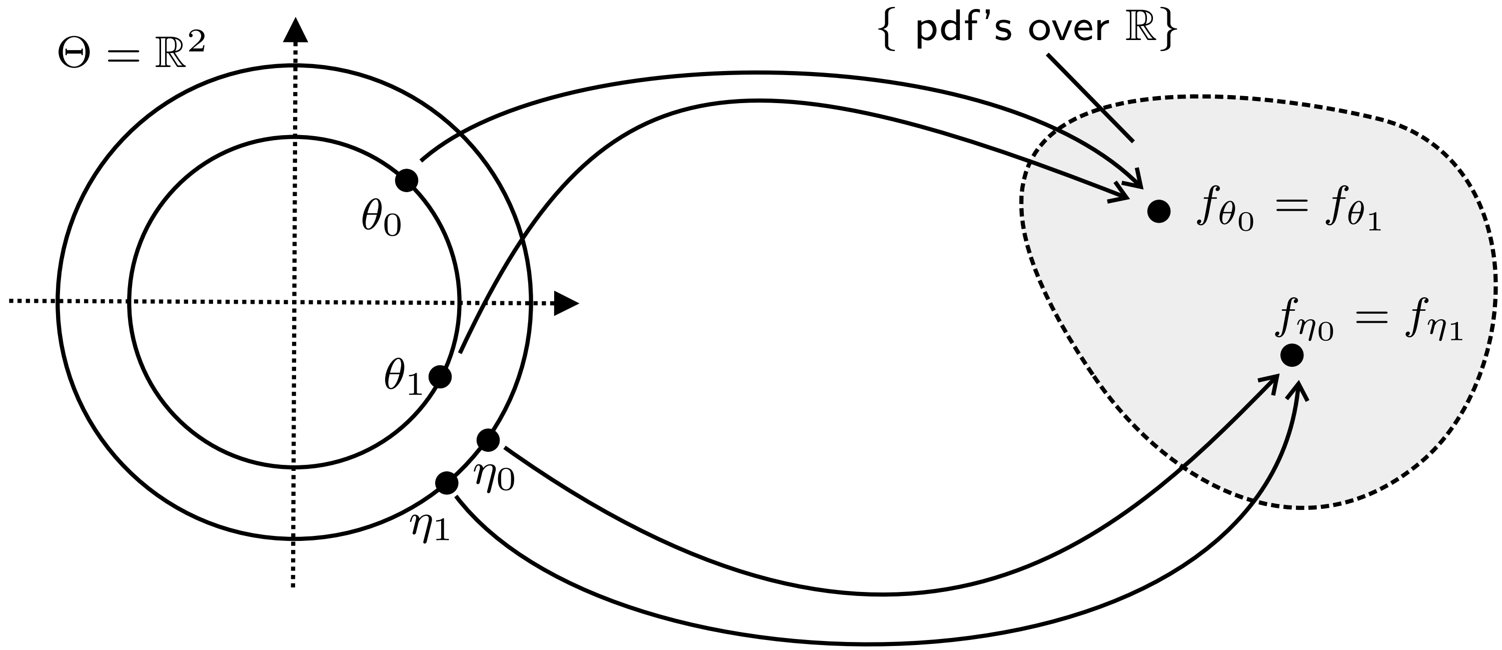


Motivation: Parametric estimation over quotient spaces

☐ Occurs in over-parameterized statistical models

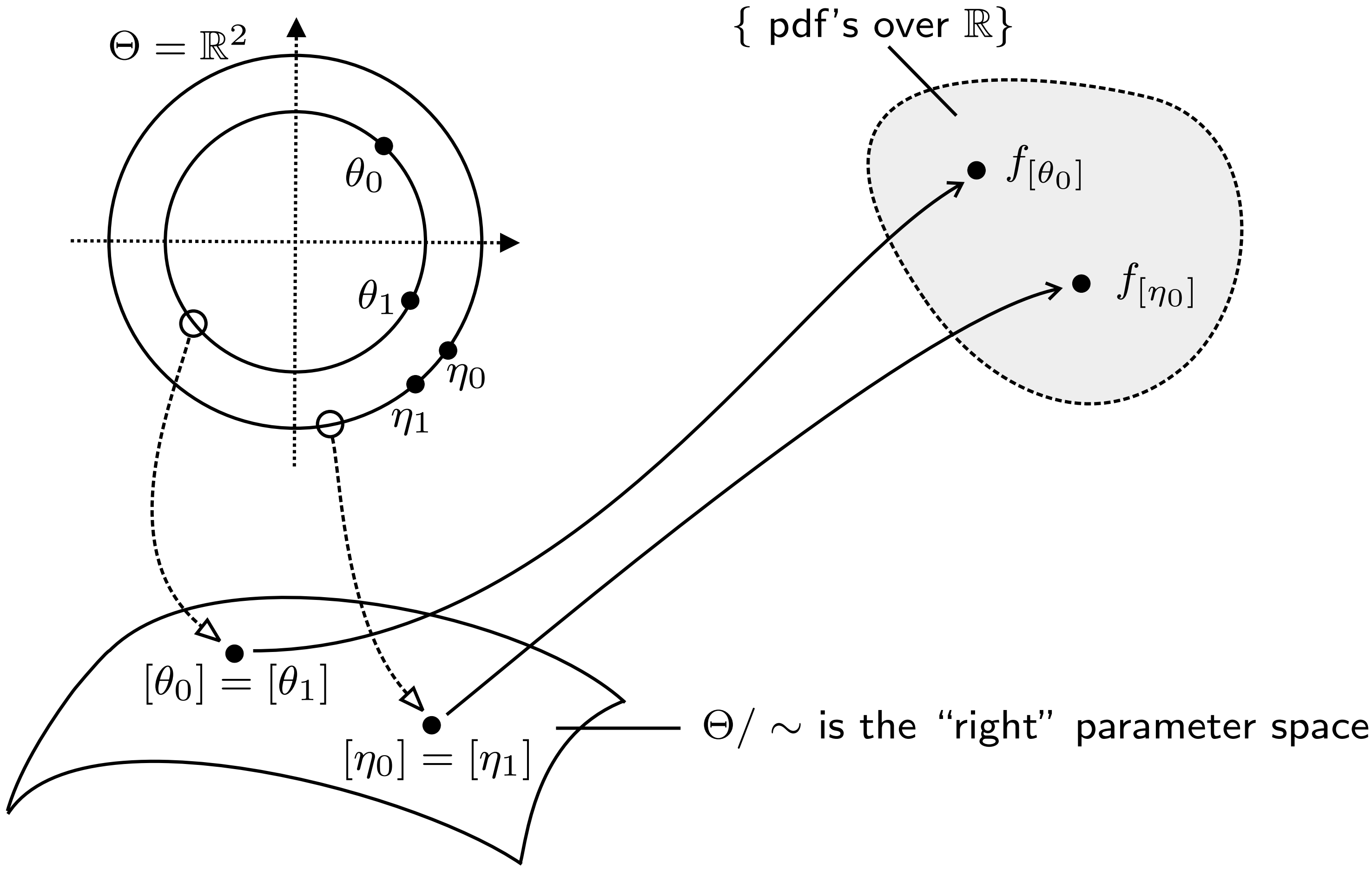
☐ **Simple example:** $\Theta = \mathbb{R}^2$

▷ Observation model: $y = \|\theta\| + \text{AWGN}$



Motivation: Parametric estimation over quotient spaces

☐ Introduce equivalence relation on Θ: $\theta_1 \sim \theta_2 \;\Leftrightarrow\; \|\theta_1\| = \|\theta_2\|$



Outline

☐ **Motivation:** Riemannian manifolds in parametric estimation problems

☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

☐ **Framework**: Parametric statistical models over Riemannian manifolds

☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

☐ **Applications of IVLB**:

▷ Parametric estimation with constraints

▷ Parametric estimation over quotient spaces

☐ **Open problems**

Cramér-Rao Bound (CRB)

☐ Classical setup:

- ▷ $\Omega = \mathbb{R}^n$ is the observation space and $y \in \Omega$ is the observed data point
- ▷ $\mathcal{F} = \{f_\theta \,:\, \theta \in \Theta\}$ is a given parametric family of positive pdf's
- ▷ $\widehat{\theta} \,:\, \Omega \to \Theta$ is an unbiased estimator of θ, i.e, $\mathsf{E}_\theta \left\{ \widehat{\theta}(Y) \right\} = \theta, \ \forall_{\theta \in \Theta}$
- ▷ Θ denotes an **open** subset of the Euclidean space $\mathbb{R}^p$

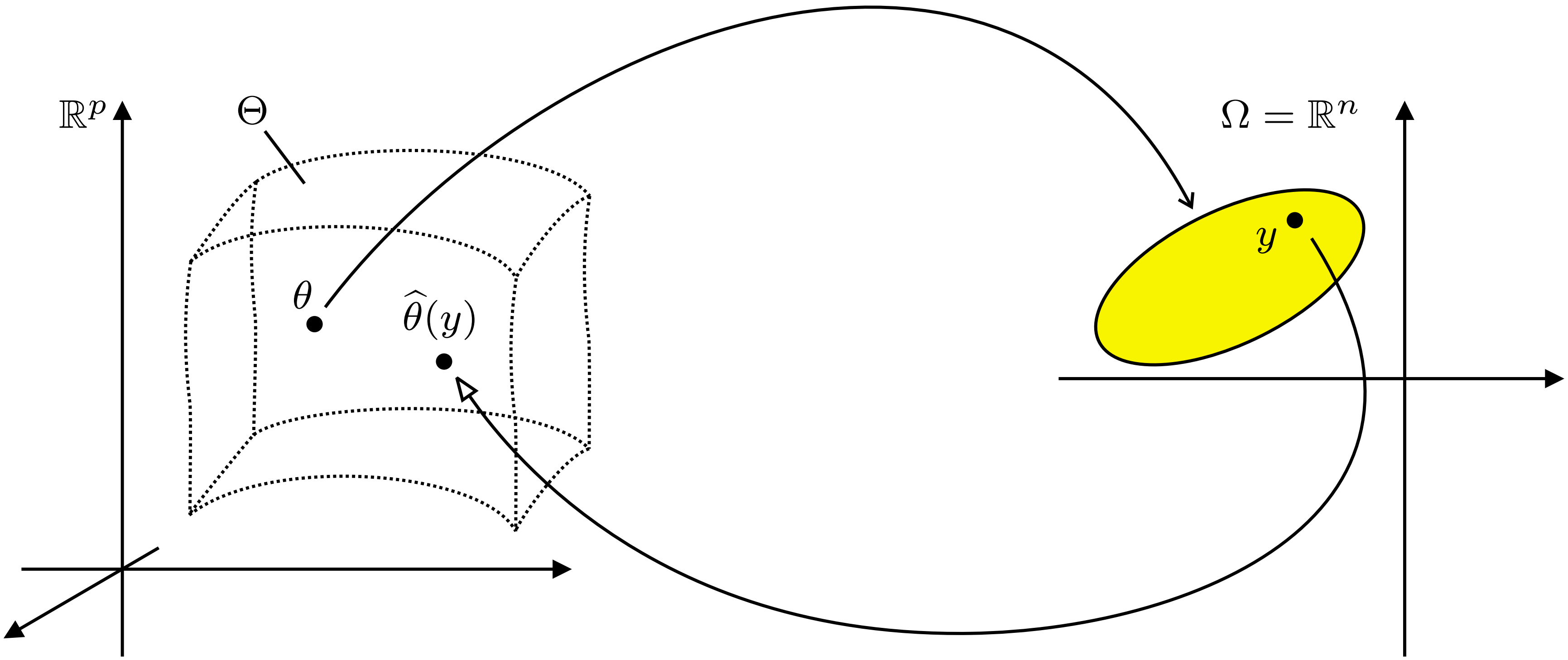
☐ CRB inequality:

$$\mathsf{Cov}_\theta \left(\widehat{\theta} \right) \succeq I(\theta)^{-1}$$

- ▷ $\mathsf{Cov}_\theta \left(\widehat{\theta} \right) = \mathsf{E}_\theta \left\{ \left(\widehat{\theta}(Y) - \theta \right) \left(\widehat{\theta}(Y) - \theta \right)^T \right\}$ is the covariance matrix of $\widehat{\theta}$
- ▷ $I(\theta) = \mathsf{E}_\theta \left\{ \nabla_\theta \ln f(Y;\theta) \, \nabla_\theta \ln f(Y;\theta)^T \right\}$ is the Fisher Information Matrix (FIM)

Cramér-Rao Bound (CRB)

☐ **Classical Euclidean setup:**



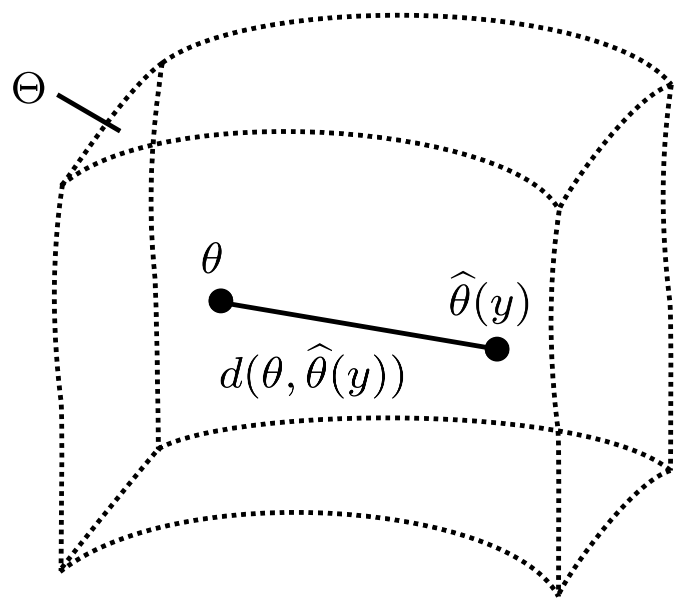
Cramér-Rao Bound (CRB)

☐ Distance lower bound:

$$\mathsf{Cov}_\theta\left(\widehat{\theta}\right) \succeq I(\theta)^{-1} \quad \Rightarrow \quad \mathsf{var}_\theta\left(\widehat{\theta}\right) \geq \mathsf{tr}\left(I(\theta)^{-1}\right)$$

▷ $\mathsf{var}_\theta\left(\widehat{\theta}\right) = \mathsf{E}_\theta\left\{d\left(\theta, \widehat{\theta}(Y)\right)^2\right\}$ is the variance of the estimator $\widehat{\theta}$

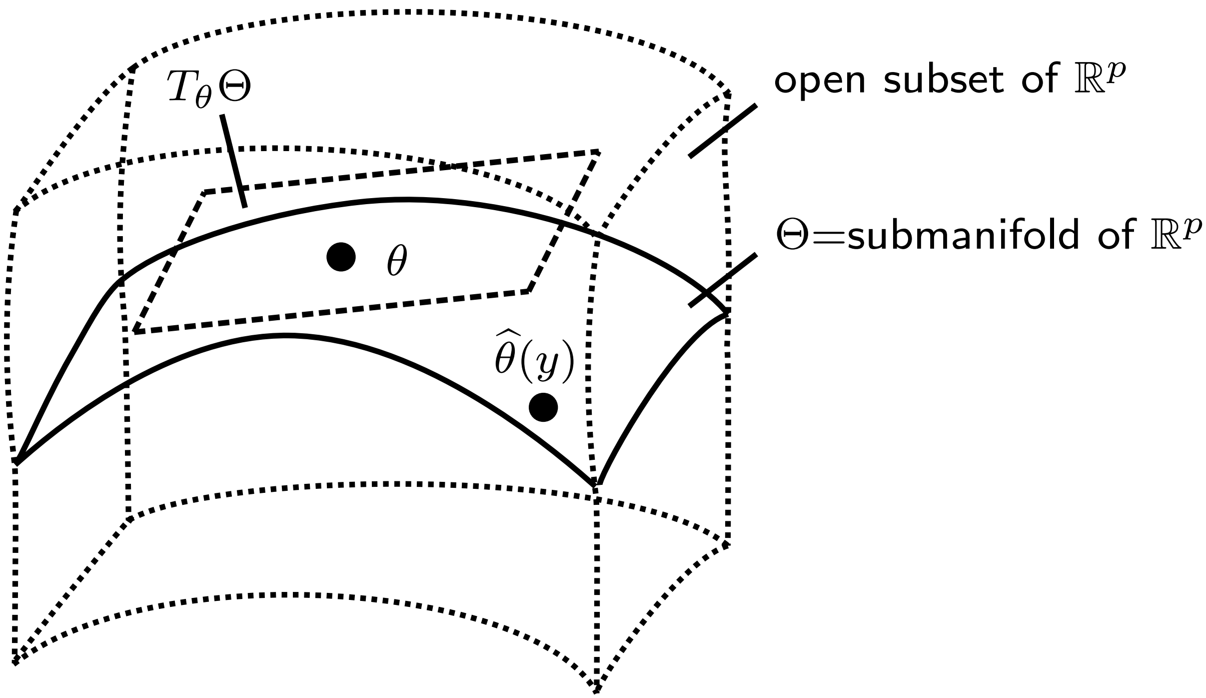
▷ $d\left(\theta, \widehat{\theta}(y)\right) = \left\|\theta - \widehat{\theta}(y)\right\|$ is the Euclidean distance between θ and $\widehat{\theta}(y)$



CRB Extension: Parametric Estimation with Constraints

☐ Θ contracts to a submanifold of $\mathbb{R}^p$:

[Gorman & Hero'90, Marzetta'93, Stoica & Ng'98, etc]



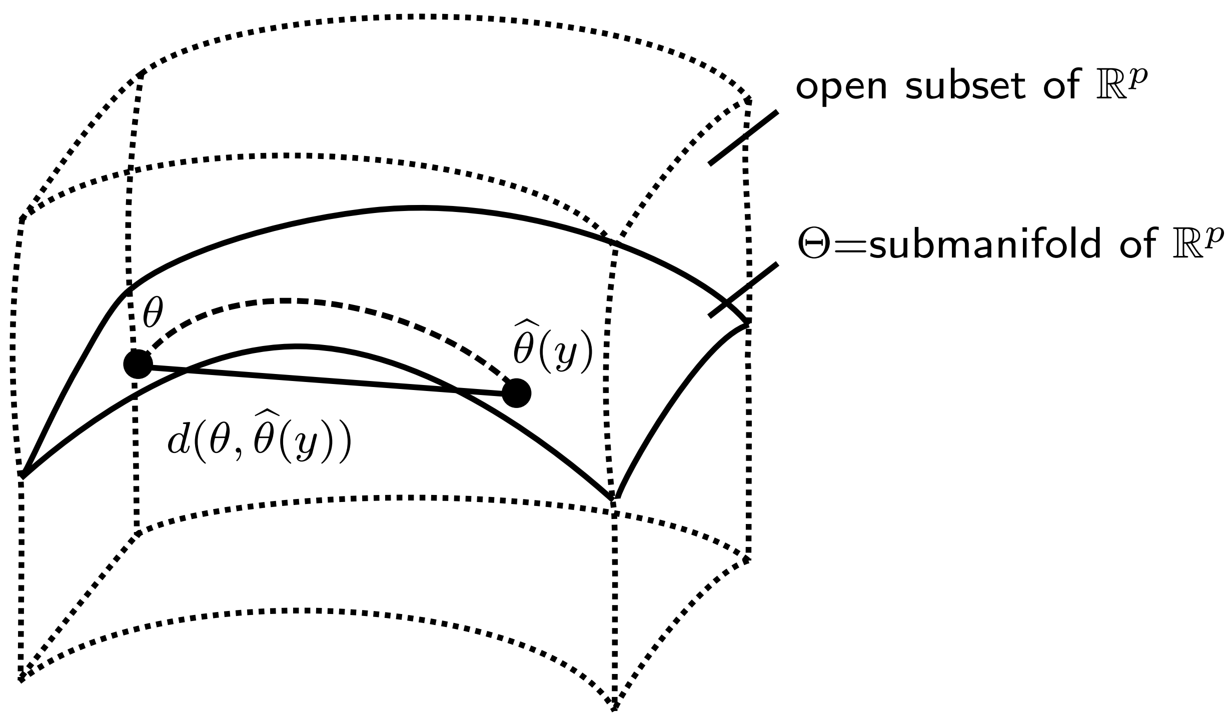
☐ Extended CRB inequality: $\text{Cov}_\theta\left(\widehat{\theta}\right) \succeq U_\theta \left(U_\theta^T I(\theta) U_\theta\right)^{-1} U_\theta^T$

▷ U_θ is an orthonormal basis for $T_\theta\Theta$, the tangent space to Θ at θ

CRB Extension: Parametric Estimation with Constraints

☐ Extrinsic distance lower bound: even with $\mathsf{E}_\theta\left\{\widehat{\theta}(Y)\right\} \simeq \theta$,

$$\mathsf{var}_\theta\left(\widehat{\theta}\right) = \mathsf{E}_\theta\left\{d(\theta, \widehat{\theta}(Y))^2\right\} \geq \mathsf{tr}\left(\left(U_\theta^T I(\theta) U_\theta\right)^{-1}\right)$$



CRB Extension: Parametric Estimation with Singular FIM

☐ $\widehat{b} : \Omega = \mathbb{R}^n \to \mathbb{R}^m$, $y \mapsto \widehat{b}(y)$ some estimator with mean value $b : \Theta \to \mathbb{R}^m$,

$$b(\theta) = \mathsf{E}_\theta \left\{ \widehat{b}(Y) \right\}$$

☐ Extended CRB Inequality:

$$\mathrm{Cov}_\theta \left(\widehat{b} \right) \succeq \nabla b(\theta) \, I(\theta)^+ \, \nabla b(\theta)^T$$

▷ $\nabla b(\theta)$ derivative of map b at θ

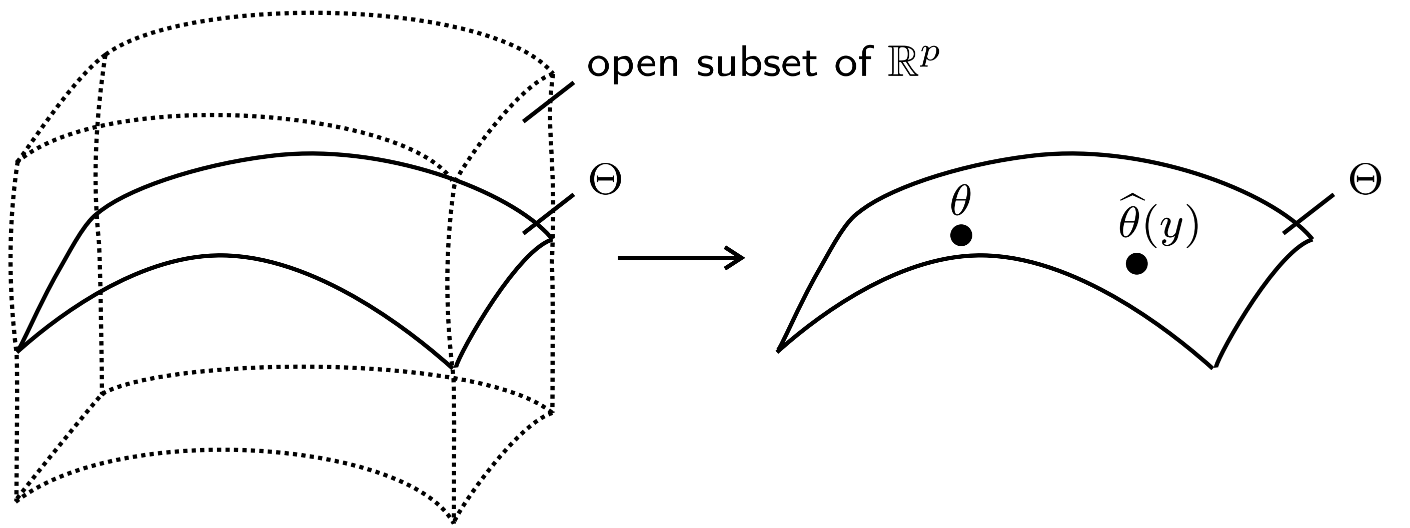
▷ $I(\theta)^+$ = Moore-Penrose pseudoinverse of $I(\theta)$

[Stoica & Marzetta'01, etc]

☐ **Note:** Pure algebraic result (lacks geometrical interpretation)

CRB Extension: Smooth Manifolds

☐ Θ is an abstract smooth manifold (topological space with smooth atlas)



☐ Bounds formulated in terms of intrinsic-only objects:

▷ **Tensor Information Inequality**: $\mathrm{Cov}_\theta\left(\widehat{\theta}\right) \succeq I_\theta^{-1}$

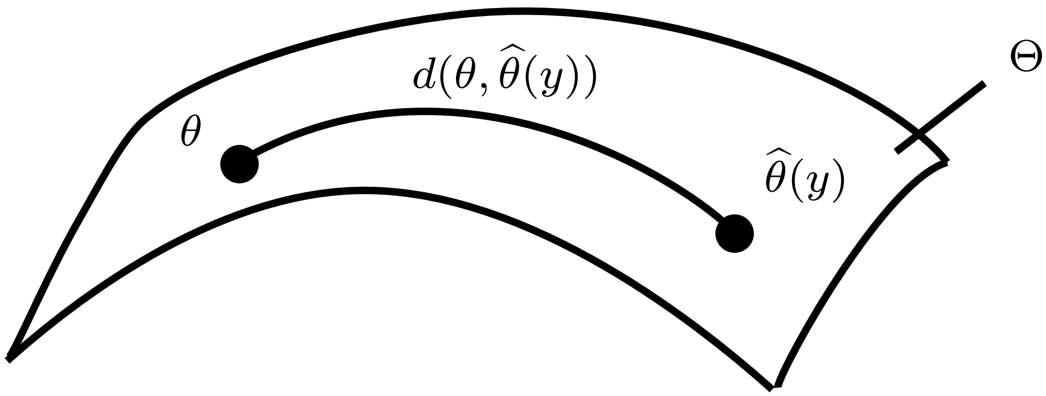
▷ $\mathrm{Cov}_\theta(\widehat{\theta})$ and I_θ are contravariant/covariant tensors

[Hendriks'91, *Journal of Multivariate Analysis*]

Our CRB Extension: Riemannian Manifolds

☐ Θ is a (connected) Riemannian manifold

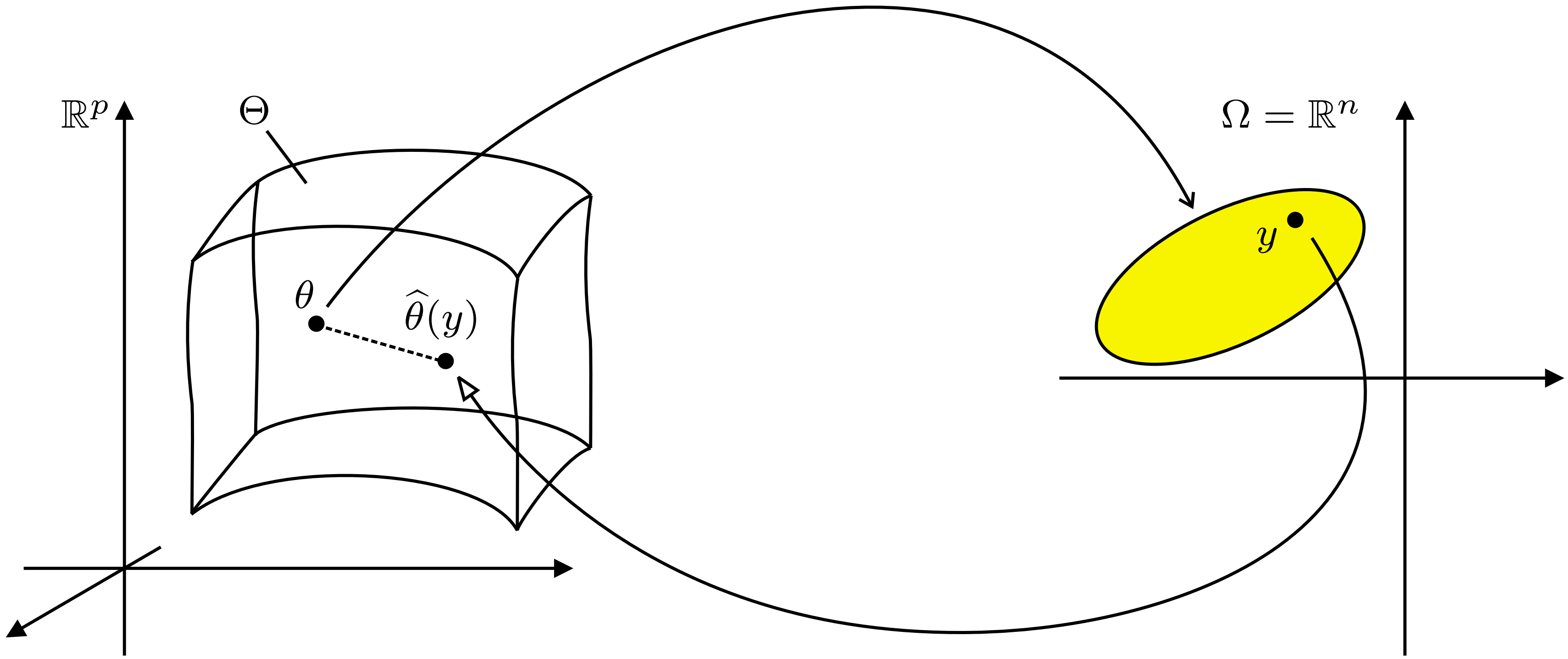
▷ **Key-point:** A (geodesic) distance d is available on Θ



☐ The performance of estimators is assessed with respect to d

CRB Extension: Riemannian Manifolds

☐ **Classical Euclidean setup:**

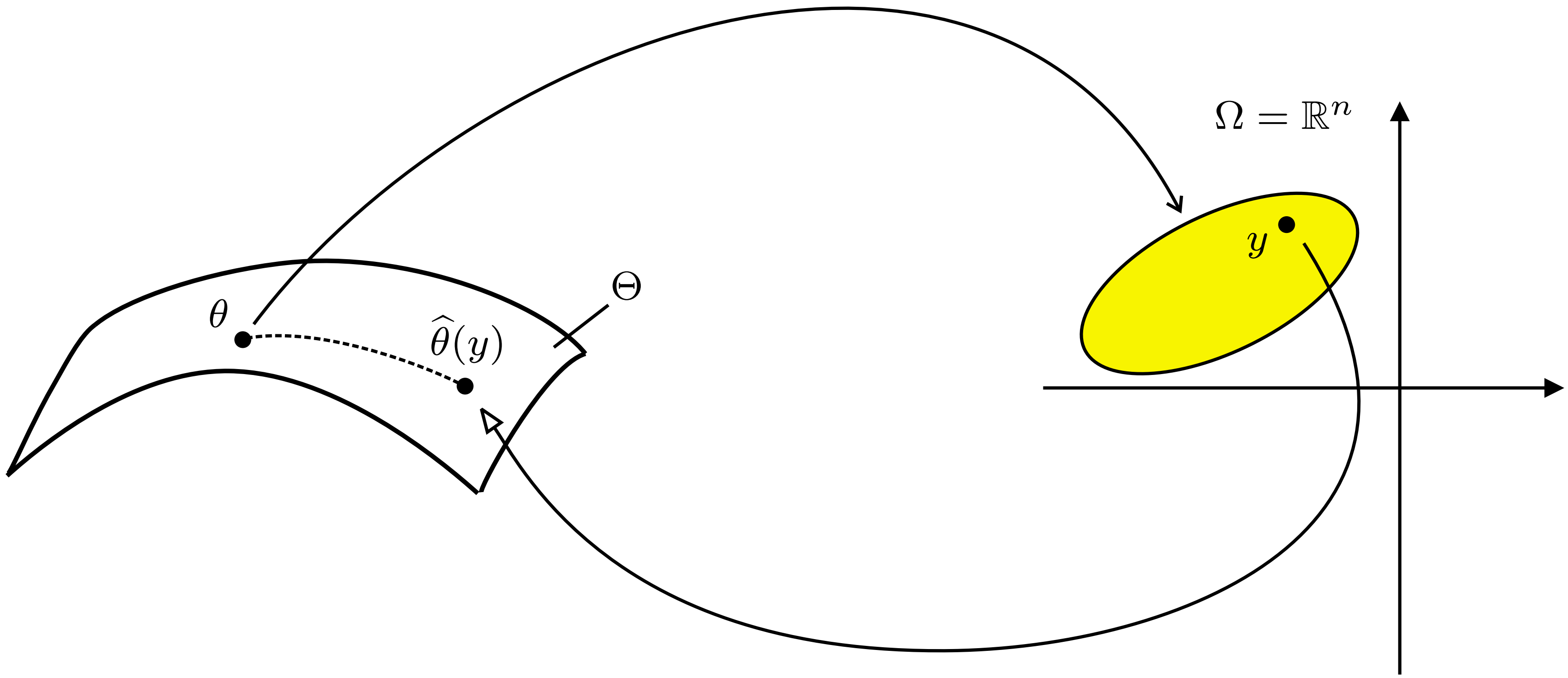


☐ **Cramér-Rao Bound (CRB):**

$$\mathsf{var}_\theta\left(\widehat{\theta}\right) = \mathsf{E}_\theta\left\{d\left(\theta, \widehat{\theta}(Y)\right)^2\right\} \geq \mathsf{tr}\left(I(\theta)^{-1}\right)$$

CRB Extension: Riemannian Manifolds

☐ **Riemannian setup:**



☐ **Intrinsic Variance Lower Bound (IVLB):**

$$\mathsf{var}_\theta\left(\widehat{\theta}\right) = \mathsf{E}_\theta\left\{d\left(\theta, \widehat{\theta}(Y)\right)^2\right\} \geq ?$$

Outline

☐ **Motivation:** Riemannian manifolds in parametric estimation problems

☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

☐ **Framework**: Parametric statistical models over Riemannian manifolds

☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

☐ **Applications of IVLB**:

- ▷ Parametric estimation with constraints
- ▷ Parametric estimation over quotient spaces

☐ **Open problems**

Problem Statement

☐ Setup:

▷ $\Omega = \mathbb{R}^n$ is the observation space and $y \in \Omega$ is the observed data point

▷ $\mathcal{F} = \{f_\theta \, : \theta \in \Theta\}$ is a given parametric family of positive pdf's

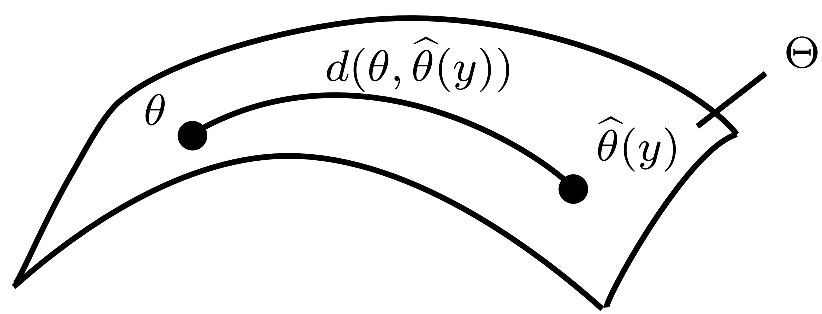
▷ $\widehat{\theta} \, : \, \Omega \longrightarrow \Theta$ is an unbiased estimator of θ:

$$\mathrm{iE}_\theta \left\{ \widehat{\theta}(Y) \right\} = \theta, \ \forall_{\theta \in \Theta}$$

▷ Θ denotes a connected **Riemannian manifold**

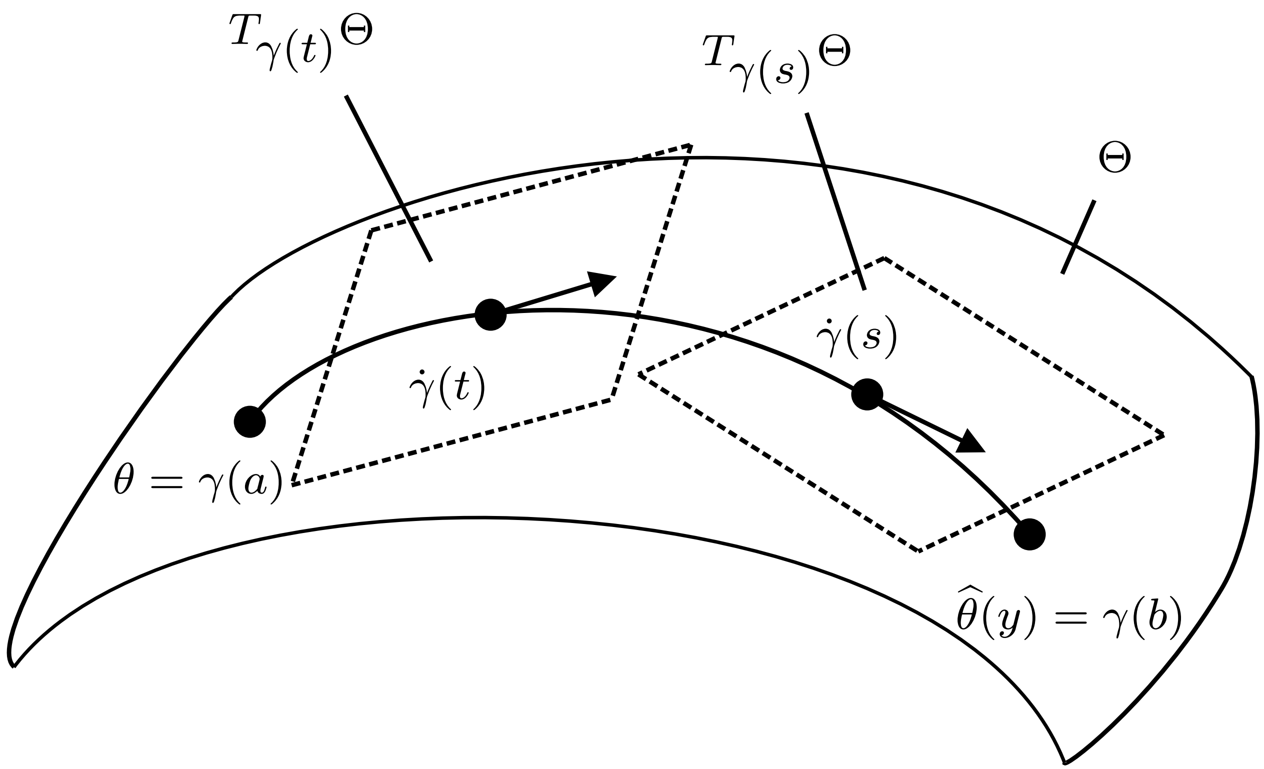
☐ **Goal**: to lower bound the intrinsic variance of $\widehat{\theta}$

$$\mathrm{var}_\theta \left(\widehat{\theta} \right) = \mathrm{E}_\theta \left\{ d\left(\theta, \widehat{\theta}(Y) \right)^2 \right\} \geq ?$$



Differential-Geometric Framework

☐ Length of a curve segment $\gamma : [a, b] \to \Theta$ is $L(\gamma) = \int_a^b |\dot{\gamma}(t)|\, dt$

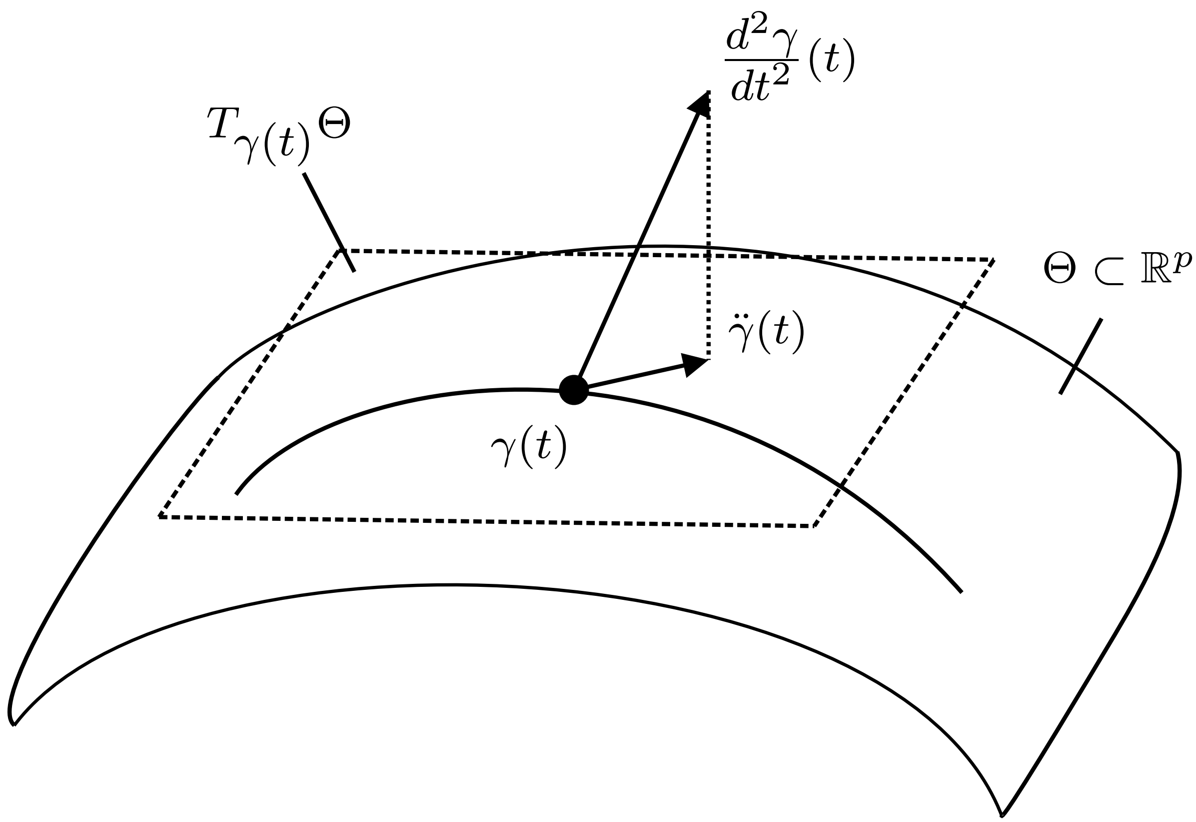


☐ Geodesic distance between θ and $\widehat{\theta}(y)$ in Θ:

$$d\left(\theta, \widehat{\theta}(y)\right) = \inf \left\{ L(\gamma) \,:\, \gamma(a) = \theta,\, \gamma(b) = \widehat{\theta}(y) \right\}$$

Differential-Geometric Framework

☐ The acceleration of a curve segment $\gamma : [a,b] \longrightarrow \Theta$ is $\ddot{\gamma}(t) = \dfrac{D}{dt}\dot{\gamma}(t)$



☐ A curve segment γ is a geodesic iff $\ddot{\gamma} \equiv 0$ (its acceleration vanishes identically)

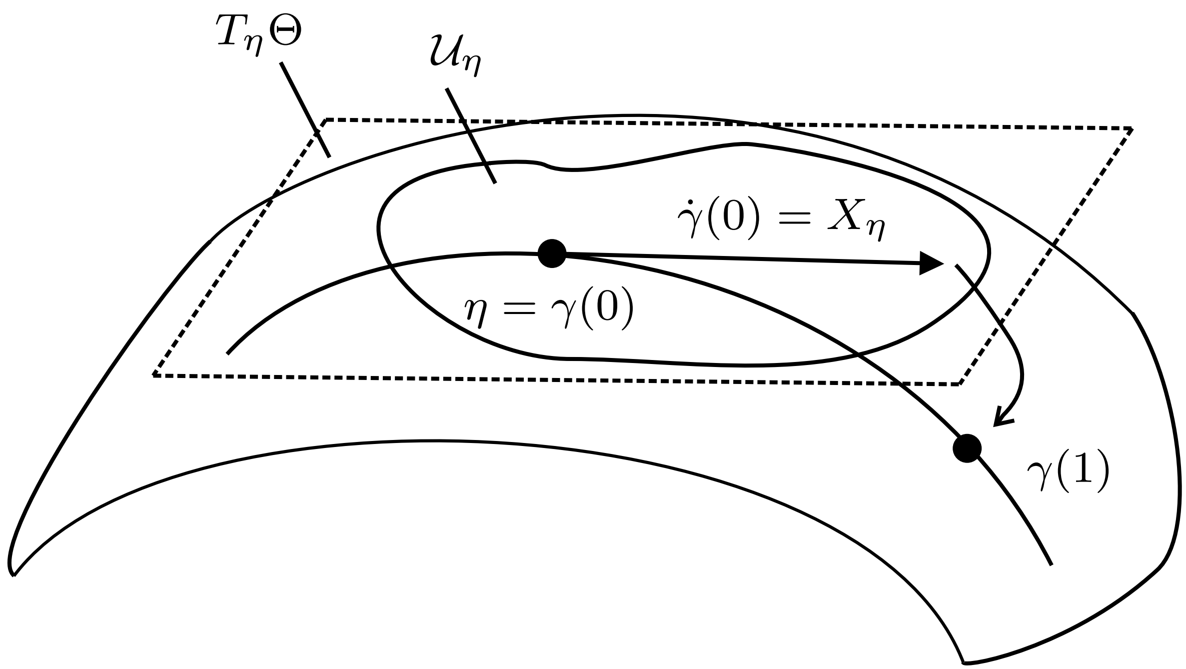
Differential-Geometric Framework

☐ The exponential mapping

▷ $\exp_\eta$: $\mathcal{U}_\eta$ (open) $\subset T_\eta\Theta \to \Theta$

▷ Let γ denote a geodesic such that $\gamma(0) = \eta$ and $\dot{\gamma}(0) = X_\eta \in T_\eta\Theta$

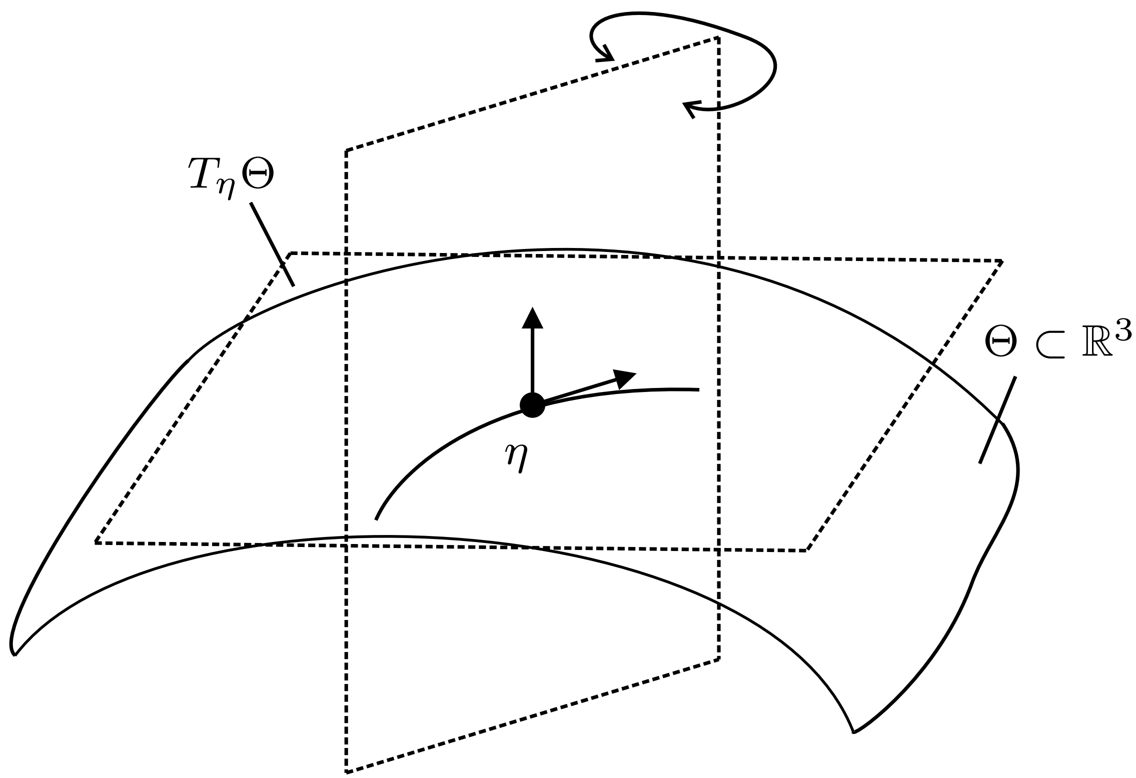
▷ $\exp_\eta(X_\eta) = \gamma(1)$



☐ Geodesic ball: $\mathrm{B}(\eta;\epsilon) = \exp_\eta(\{X_\eta : |X_\eta| < \epsilon\})$ (diffeomorphic image)

Differential-Geometric Framework

☐ If γ is a unit-speed curve, then $\kappa = |\ddot{\gamma}(0)|$ is its curvature at $t = 0$



☐ Gaussian curvature (GC) of a surface $\Theta \subset \mathbb{R}^3$ at η: $\kappa_\eta = \pm\kappa_{\max} \cdot \kappa_{\min}$

☐ Sectional curvature of a 2D-plane $\Pi \subset T_\eta\Theta$: $C_\eta(\Pi) =$ GC of $\exp_\eta(\Pi)$

Differential-Geometric Framework

☐ The Fisher-Information Form I is a section of the bundle $T^2_0(\Theta)$: I_θ is a $(2,0)$ tensor on $T_\theta\Theta$

$$(X_\theta, Y_\theta) \in T_\theta\Theta \times T_\theta\Theta \quad \mapsto \quad I_\theta\,(X_\theta, Y_\theta) = \mathsf{E}_\theta\,\{X_\theta \ln f(Y;\theta)\, Y_\theta \ln f(Y;\theta)\}$$

▷ **Euclidean**: $T_\theta\Theta \simeq \mathbb{R}^p$ and $I_\theta(X_\theta, Y_\theta) = \mathsf{E}_\theta\left\{X_\theta^T \nabla_\theta \ln f(Y;\theta) Y_\theta^T \nabla_\theta \ln f(Y;\theta)\right\}$

☐ Θ is a measurable space with σ-algebra generated by its topology

☐ A random point $\widehat{\theta}$ in Θ is a measurable mapping $\widehat{\theta} : \Omega \to \Theta$

☐ $\mathrm{i}\mathsf{E}_\theta\left\{\widehat{\theta}(Y)\right\} = \theta$ iff $x = \theta$ globally minimizes $\mathsf{E}_\theta\left\{d\left(x, \widehat{\theta}(Y)\right)^2\right\}$, $\forall_{x} \in \Theta$

[Fréchet mean, Riemannian center of mass]

☐ The variance of $\widehat{\theta}$ is $\mathrm{var}_\theta\left(\widehat{\theta}\right) = \mathsf{E}_\theta\left\{d\left(\theta, \widehat{\theta}(Y)\right)^2\right\}$

Outline

- ☐ **Motivation:** Riemannian manifolds in parametric estimation problems

- ☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

- ☐ **Framework**: Parametric statistical models over Riemannian manifolds

- ☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

- ☐ **Applications of IVLB**:
  - ▷ Parametric estimation with constraints
  - ▷ Parametric estimation over quotient spaces

- ☐ **Open problems**

Intrinsic Variance Lower Bound (IVLB)

☐ **Theorem (IVLB).** Suppose:

▷ The sectional curvature of Θ is upper bounded by $C \geq 0$

▷ $\mathsf{iE}_\theta \left\{ \widehat{\theta}(Y) \right\} = \theta, \quad \forall_{\theta \in \Theta}$

▷ (simplified) $\forall_{\theta \in \Theta}$, $\mathsf{Prob}_\theta \left\{ \widehat{\theta}(Y) \in \mathsf{B}(\theta; \epsilon) \right\} = 1$ where $\epsilon < \frac{\text{const.}}{\sqrt{C}}$

Then,

$$\text{var}_\theta \left(\widehat{\theta} \right) \geq \begin{cases} \lambda_\theta & , \quad \text{if } C = 0 \\ \dfrac{\lambda_\theta C + 1 - \sqrt{2\lambda_\theta C + 1}}{C^2 \lambda_\theta / 2} & , \quad \text{if } C > 0 \end{cases}$$

where:

▷ $\lambda_\theta = \mathsf{tr}(I_\theta^{-1})$

▷ $I_\theta^{-1} : T_\theta\Theta \times T_\theta^*\Theta \to \mathbb{R}$ is a $(1, 1)$-tensor obtained from the Fisher tensor I_θ

Intrinsic Variance Lower Bound (IVLB)

☐ Remarks:

▷ If $\Theta = \mathbb{R}^p$ (Euclidean case), then $C = 0$ and

IVLB = Cramér-Rao Bound

▷ The upper-bound C on the sectional curvatures of Θ should be tight:

$$\phi(C) = \frac{\lambda_\theta C + 1 - \sqrt{2\lambda_\theta C + 1}}{C^2 \lambda_\theta / 2} \text{ decreases with } C$$

Outline

☐ **Motivation:** Riemannian manifolds in parametric estimation problems

☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

☐ **Framework**: Parametric statistical models over Riemannian manifolds

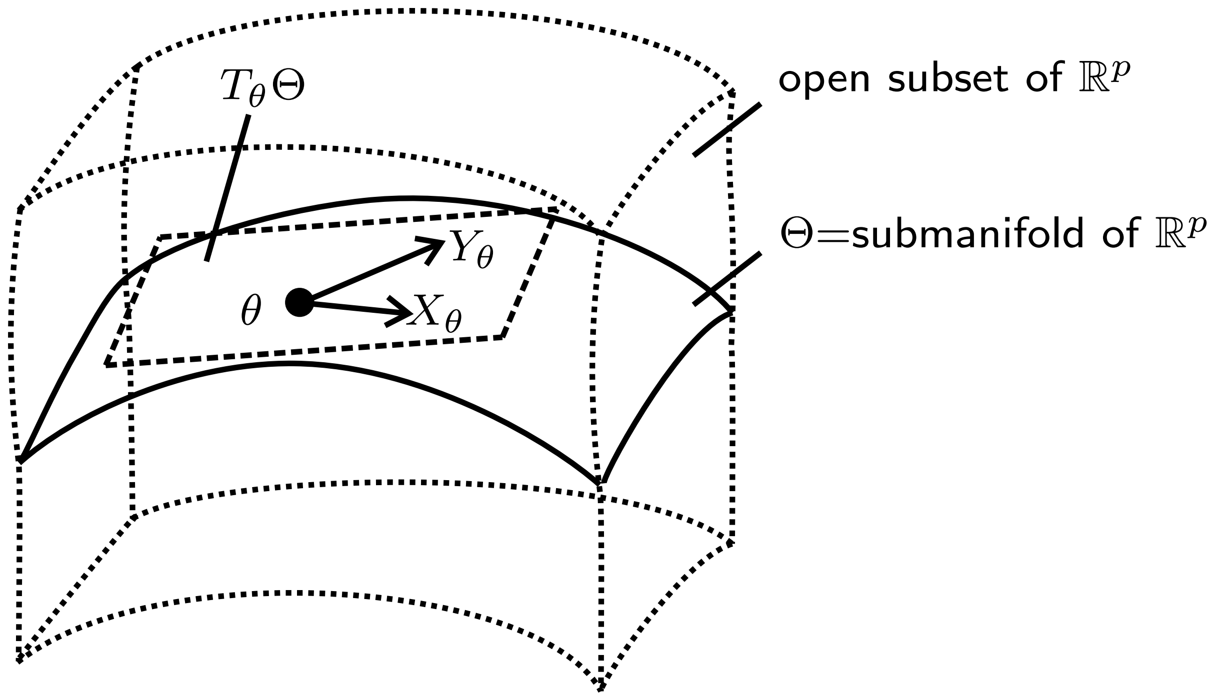
☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

☐ **Applications of IVLB**:

- ▷ Parametric estimation with constraints
- ▷ Parametric estimation over quotient spaces

☐ **Open problems**

IVLB: Parametric estimation with constraints



- ☐ Geometry of Θ is inherited from the ambient Euclidean space: $\langle X_\theta, Y_\theta \rangle = Y_\theta^T X_\theta$

- ☐ $\lambda_\theta = \text{tr}\left((U_\theta^T I(\theta) U_\theta)^{-1}\right)$ where:
  - ▷ $I(\theta) =$ Fisher information matrix
  - ▷ U_θ is an orthonormal basis for $T_\theta\Theta$

IVLB: Parametric estimation with constraints

☐ Upper-bound on sectional curvatures:

▷

$$C = \max_{\theta \in \Theta} K_\theta$$

▷

$$K_\theta = \max_{X_\theta, Y_\theta \text{ (orthonormal) } \in T_\theta \Theta} \mathsf{Rm}(X_\theta, Y_\theta, Y_\theta, X_\theta)$$

▷ Rm : $T_\theta\Theta \times T_\theta\Theta \times T_\theta\Theta \times T_\theta\Theta \to \mathbb{R}$ is the Riemannian tensor

☐ Suppose $\Theta = \{x \in \mathbb{R}^p \, : \, f_1(x) = 0, \ldots, f_M(x) = 0\}$

☐ Assume

$$\text{rank}\left[\, \nabla f_1(\theta)\, \nabla f_2(\theta) \, \cdots \, \nabla f_M(\theta) \,\right] = M, \quad \forall_{\theta \in \Theta}$$

IVLB: Parametric estimation with constraints

☐ Then:

$$K_\theta = \max_{v_1, v_2 \text{ (orthonormal)} : \nabla f_m(\theta)^T v_i = 0} \langle II(v_1, v_1), II(v_2, v_2) \rangle - \langle II(v_1, v_2), II(v_1, v_2) \rangle$$

☐ $II : T_\theta\Theta \times T_\theta\Theta \to T_\theta\Theta^\perp$ is the second fundamental form of Θ:

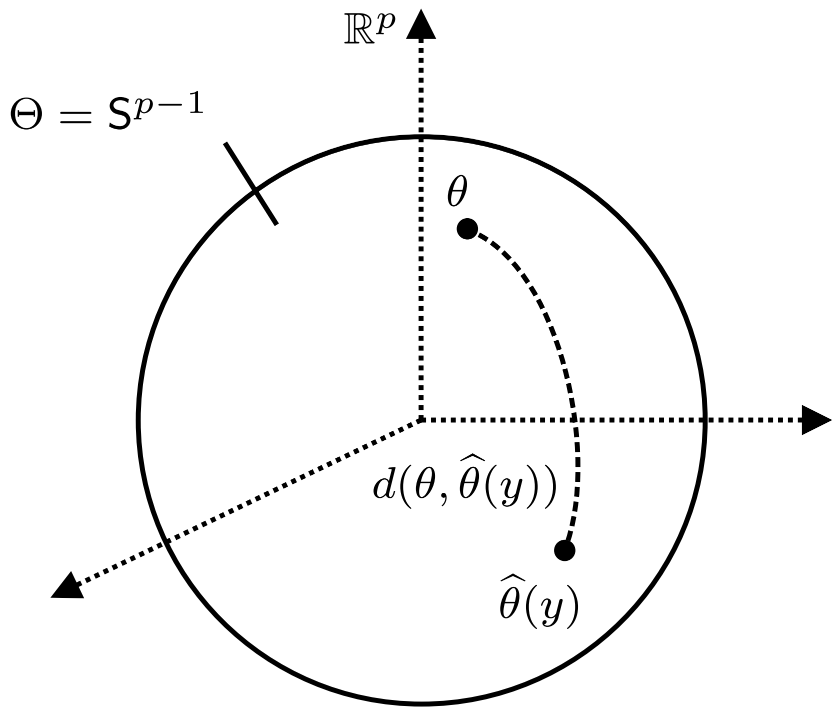
$$II(a,b) = -\nabla f(\theta) \left(\nabla f(\theta)^T \nabla f(\theta) \right)^{-1} \begin{bmatrix} a^T \nabla^2 f_1(\theta) b \\ a^T \nabla^2 f_2(\theta) b \\ \vdots \\ a^T \nabla^2 f_M(\theta) b \end{bmatrix}$$

where:

- ▷ $\nabla f(\theta) = [\, \nabla f_1(\theta) \, \nabla f_2(\theta) \, \cdots \, \nabla f_M(\theta) \,]$
- ▷ $\nabla^2 f_m(\theta)$ is the Hessian of f_m at θ

Example: inference on S^{p-1}

☐ $\mathsf{S}^{p-1} = \{x \in \mathbb{R}^p \ : \ \|x\| = 1\}$ is the unit-sphere in $\mathbb{R}^p$



☐ **Geometry of** Θ: $d(\theta, \widehat{\theta}(y)) = \mathrm{acos}(\theta^T \widehat{\theta}(y))$ and $C = 1$

Example: inference on S^{p-1}

- **Observation**: $y = \theta + w \in \mathbb{R}^p \ (p = 10)$
  - $\triangleright\ \theta \in \Theta = \mathsf{S}^{p-1}$
  - $\triangleright\ w \sim \mathcal{N}(0, \sigma^2 I_p)$

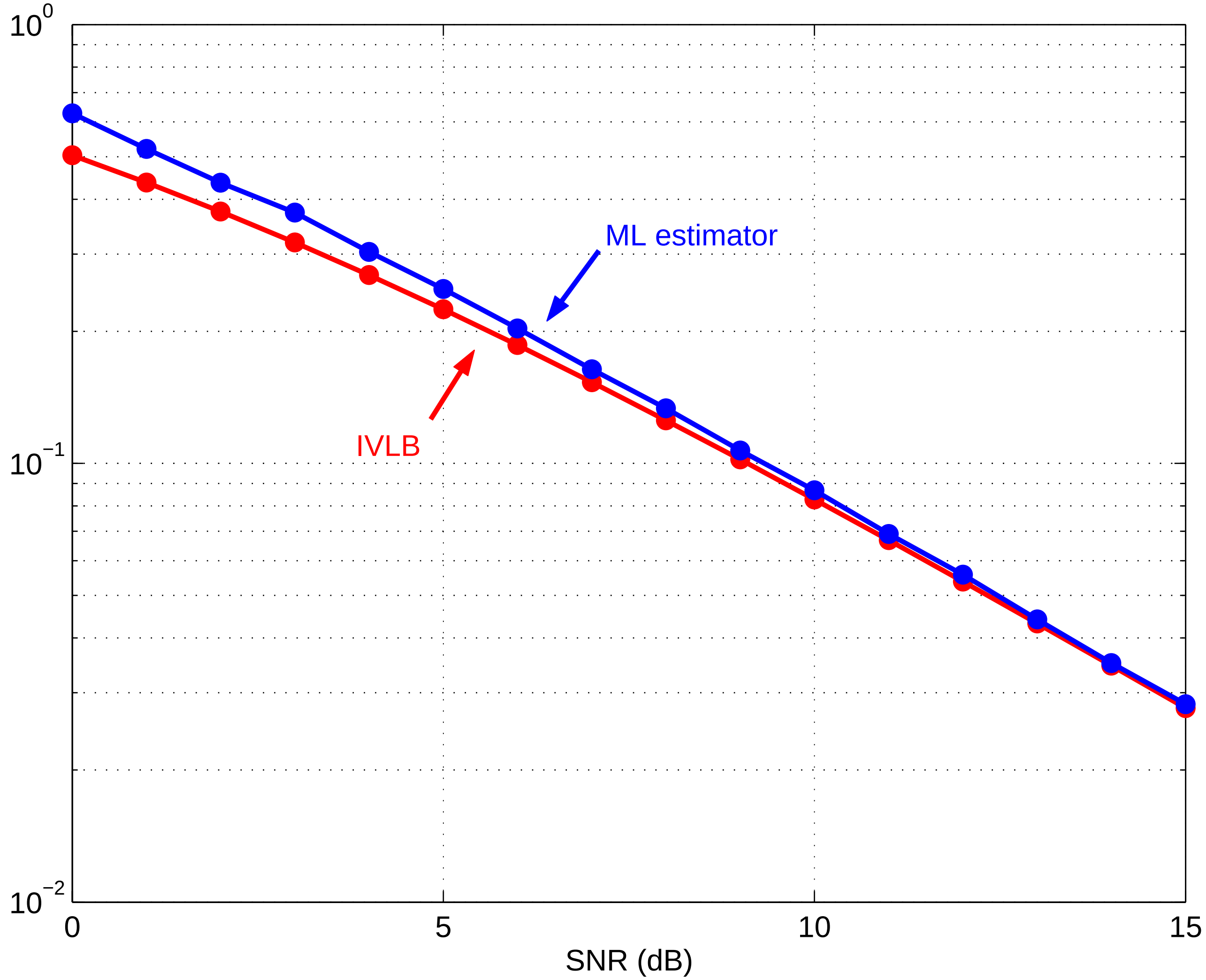
- Maximum-likelihood estimator:

$$\widehat{\theta}(y) = \frac{y}{\|y\|}$$

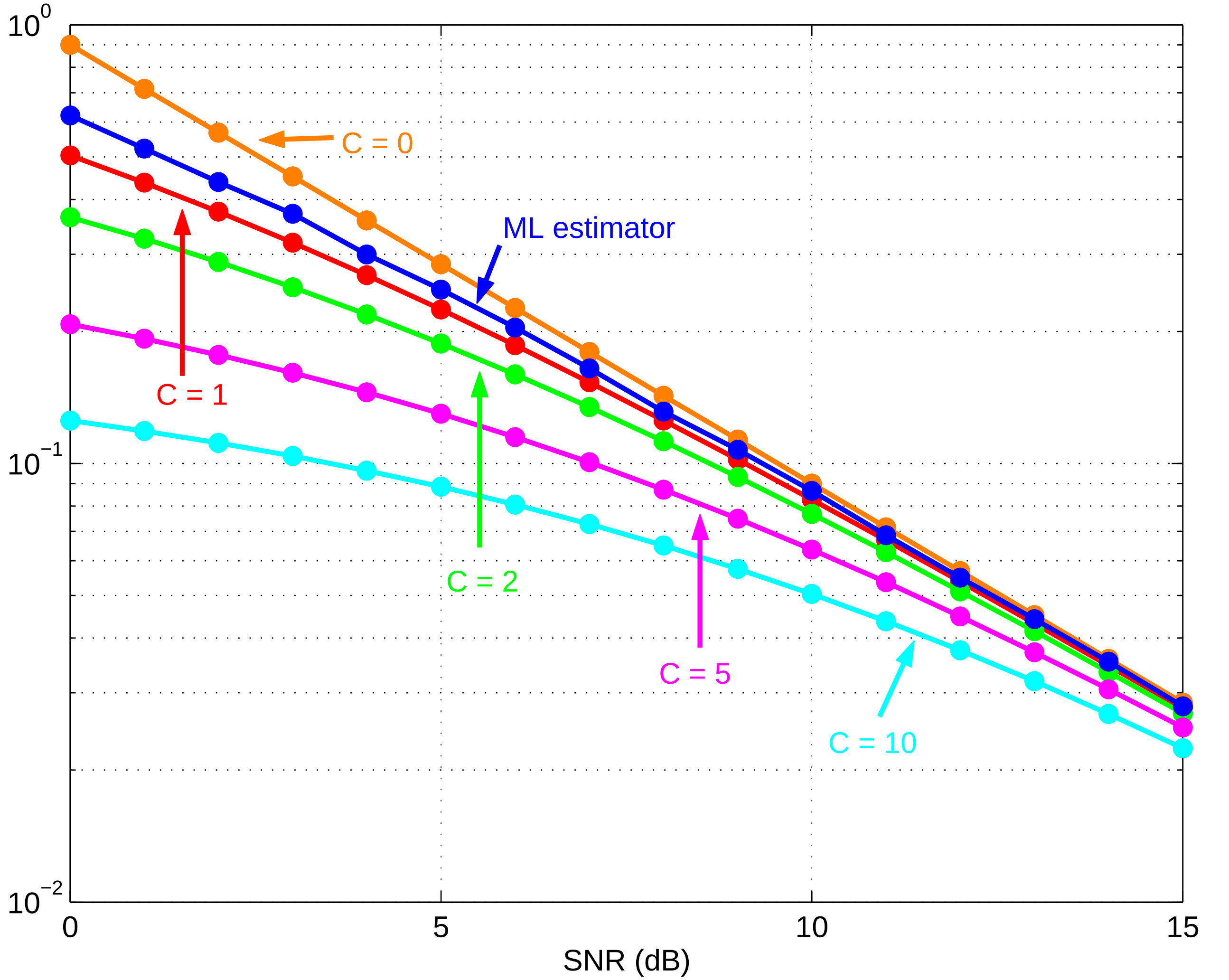
- Signal-to-noise ratio:

$$\mathsf{SNR} = \frac{\mathsf{E}\left\{\|\theta\|^2\right\}}{\mathsf{E}\left\{\|w\|^2\right\}} = \frac{1}{p\,\sigma^2}$$

Example: inference on $\mathbb{S}^{p-1}$



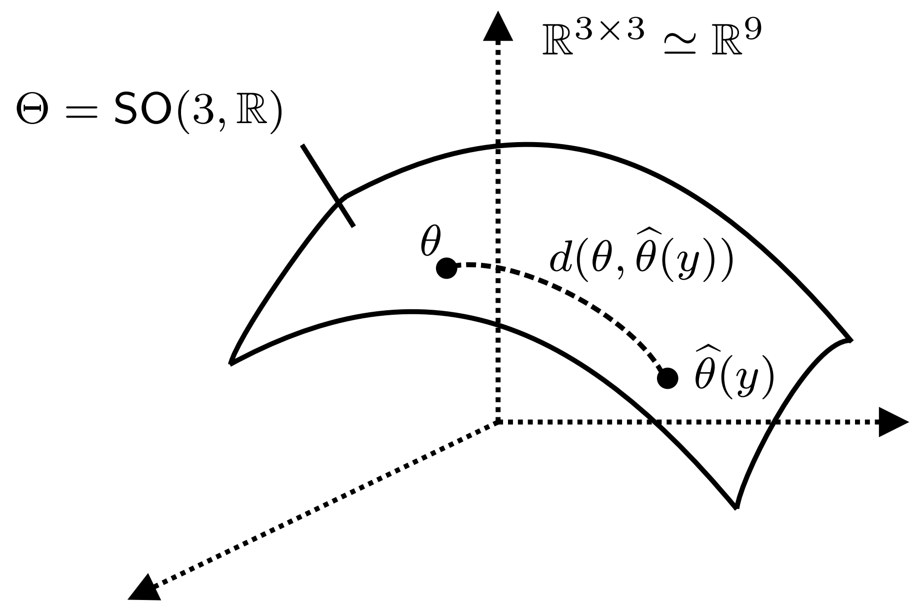
Example: inference on $\mathbb{S}^{p-1}$



Example: inference on $\mathsf{SO}(3,\mathbb{R})$

☐ $\mathsf{SO}(3,\mathbb{R})$ is the special orthogonal group:

$$\mathsf{SO}(3,\mathbb{R}) = \left\{ Q \in \mathbb{R}^{3\times 3} \,:\, Q^T Q = I_3, \det(Q) = 1 \right\}$$



☐ **Geometry of** Θ: $d(\theta,\widehat{\theta}(y)) = \sqrt{2}\,\mathsf{acos}(0.5[\mathsf{tr}(\theta^T\widehat{\theta}(y)) - 1])$ and $C = 1/8$

Example: inference on $\mathsf{SO}(3,\mathbb{R})$

☐ **Observation**: $Y = \theta X + W \in \mathbb{R}^{3\times k}$ $(k = 10)$

▷ $\theta \in \Theta = \mathsf{SO}(3,\mathbb{R})$: unknown rotation matrix [Procrustean analysis]

▷ $X = [\, x_1\, x_2\, \cdots\, x_k\,]$: constellation of known k landmarks in $\mathbb{R}^3$ $(XX^T = I_3)$

▷ $W = [\, w_1\, w_2\, \cdots\, w_k\,]$, $w_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2 I_3)$: additive observation noise

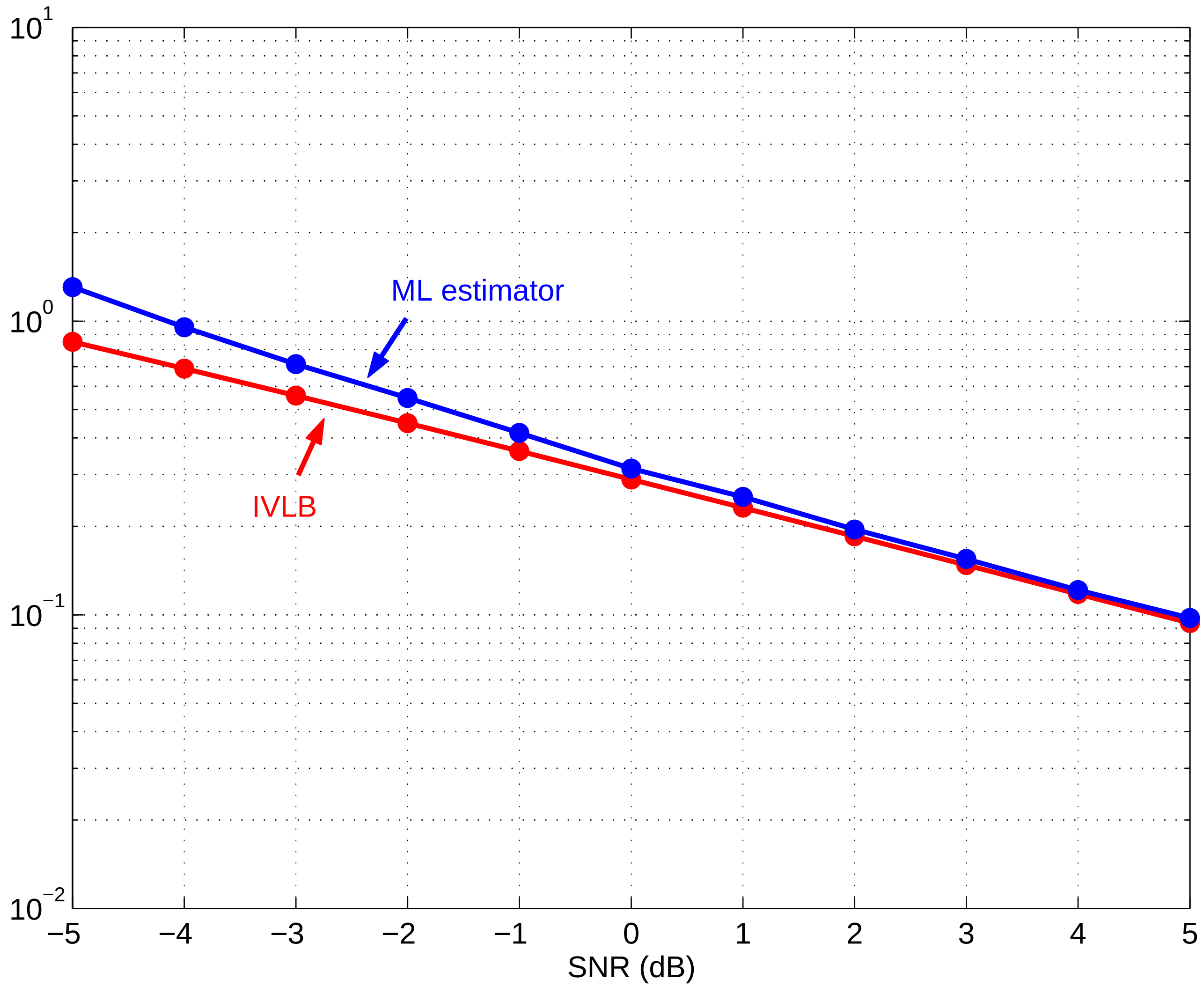
☐ Maximum-likelihood estimator:

$$\widehat{\theta}(Y) = \cdots \text{ (closed} - \text{form)}$$

☐ Signal-to-noise ratio:

$$\mathsf{SNR} = \frac{\mathsf{E}\left\{\|\theta X\|^2\right\}}{\mathsf{E}\left\{\|W\|^2\right\}} = \frac{1}{k\,\sigma^2}$$

Example: inference on $\mathrm{SO}(3, \mathbb{R})$



Outline

☐ **Motivation:** Riemannian manifolds in parametric estimation problems

☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

☐ **Framework**: Parametric statistical models over Riemannian manifolds

☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

☐ **Applications of IVLB**:

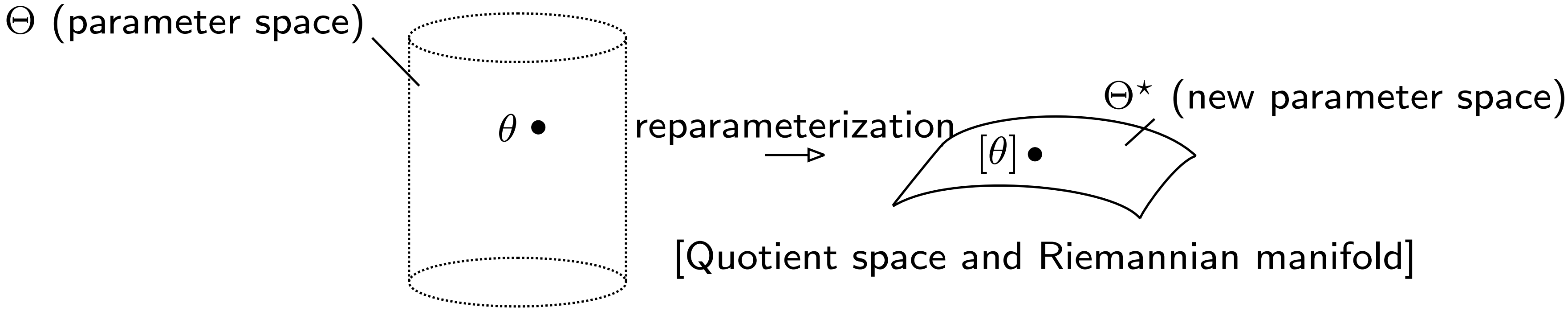
▷ Parametric estimation with constraints

▷ Parametric estimation over quotient spaces

☐ **Open problems**

IVLB: Parametric estimation over quotient spaces

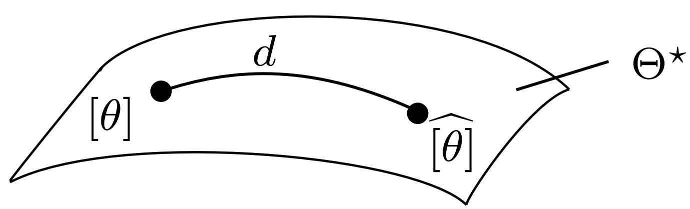
☐ New parameter space is a quotient space



[Quotient space and Riemannian manifold]

[Euclidean space]

☐ Lower bound on the variance of unbiased estimators for equivalence classes

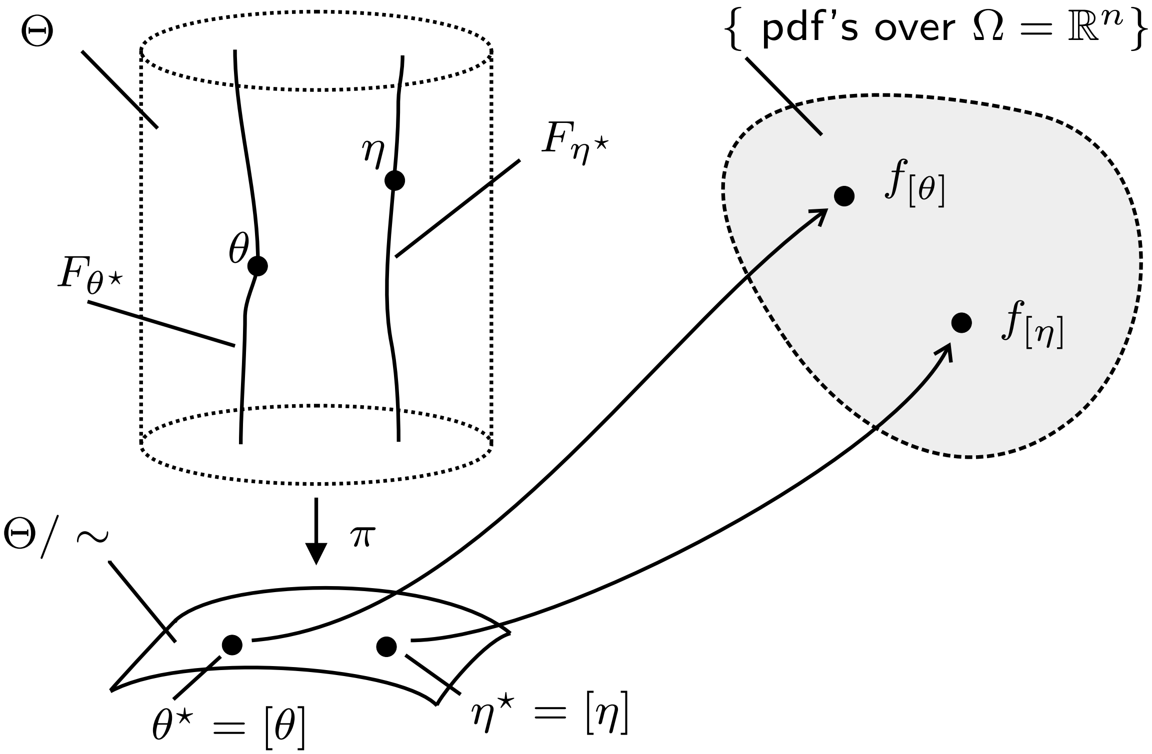


$d = d\left([\theta], \widehat{[\theta]}\right)$ is the geodesic distance

$$\mathsf{var}_{[\theta]}(\widehat{[\theta]}) = \mathsf{E}_{[\theta]}\left\{d(\widehat{[\theta]}, [\theta])^2\right\} \geq \text{IVLB}$$

Model Reparameterization

- ☐ New parametric family $\mathcal{F}^\star = \{f_{\theta^*} \, : \, \theta^* \in \Theta^*\}$
  - ▷ Introduce equivalence relation on Θ: $\theta_1 \sim \theta_2$ iff $f_{\theta_1} = f_{\theta_2}$
  - ▷ $\pi \, : \, \Theta \to \Theta/\sim$ projects θ to its equivalence class $\pi(\theta) = [\theta]$



  - ▷ $F_{\theta^\star} = \pi^{-1}(\theta^\star) = \{\theta \, : \, \pi(\theta) = \theta^\star\}$ is the fiber over $\theta^\star$

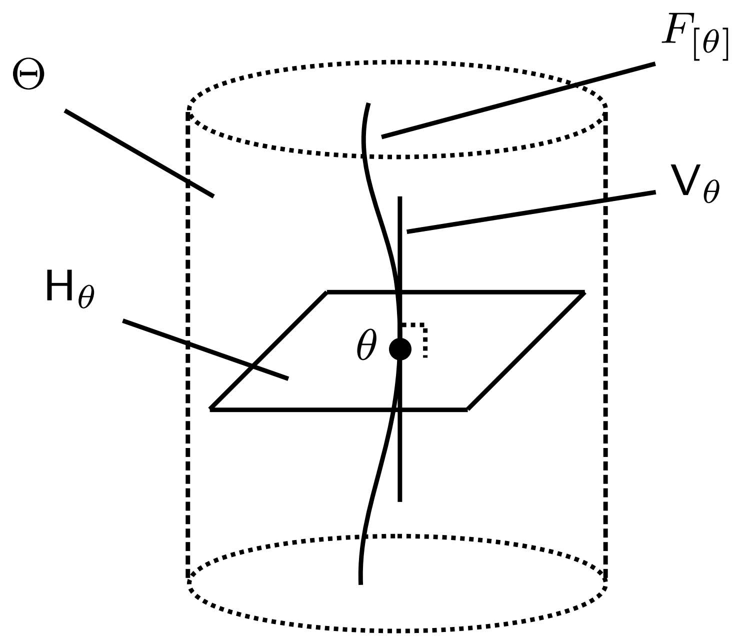
Model Reparameterization

☐ Assumptions:

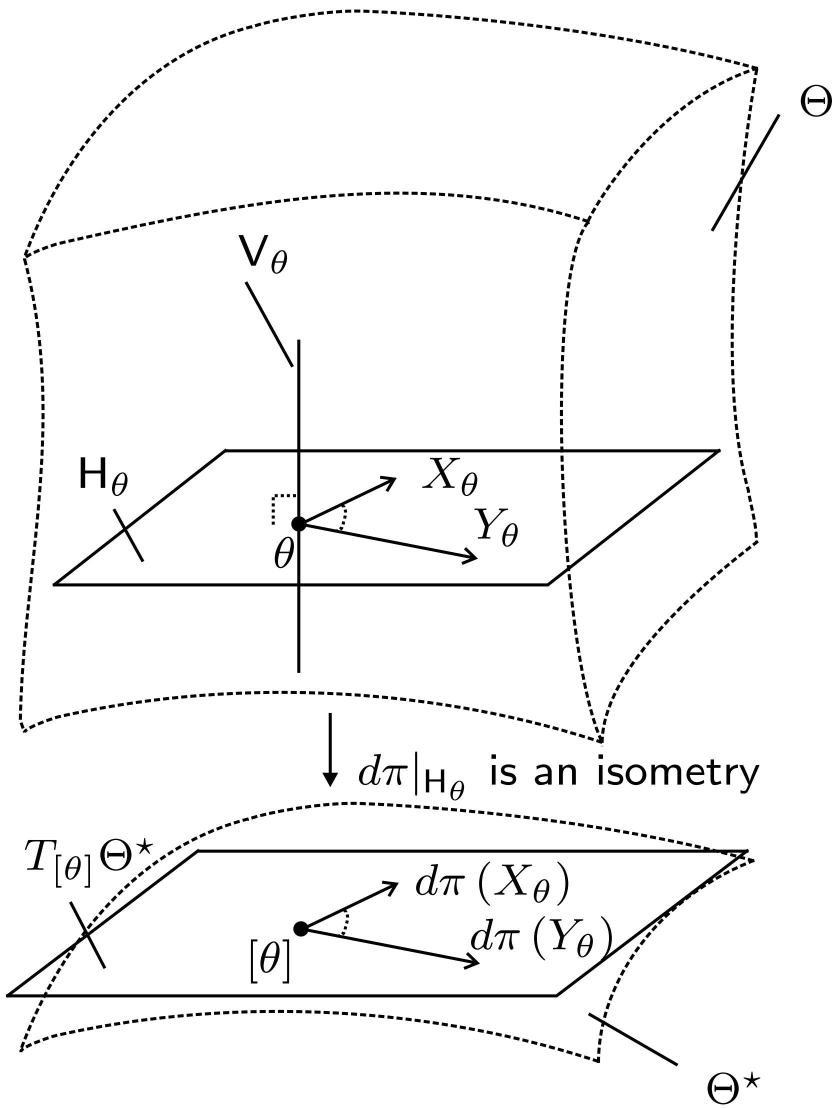
▷ $\Theta^\star$ can be given the structure of a Riemannian manifold

▷ $\pi : \Theta \rightarrow \Theta^\star$ is a Riemannian submersion

- **Intuition**: geometries of Θ and $\Theta^\star$ interface nicely through π
- Vertical subspace V_θ : tangent to the fiber $F_{[\theta]}$
- Horizontal subspace H_θ: orthogonal complement of V_θ



Model Reparameterization

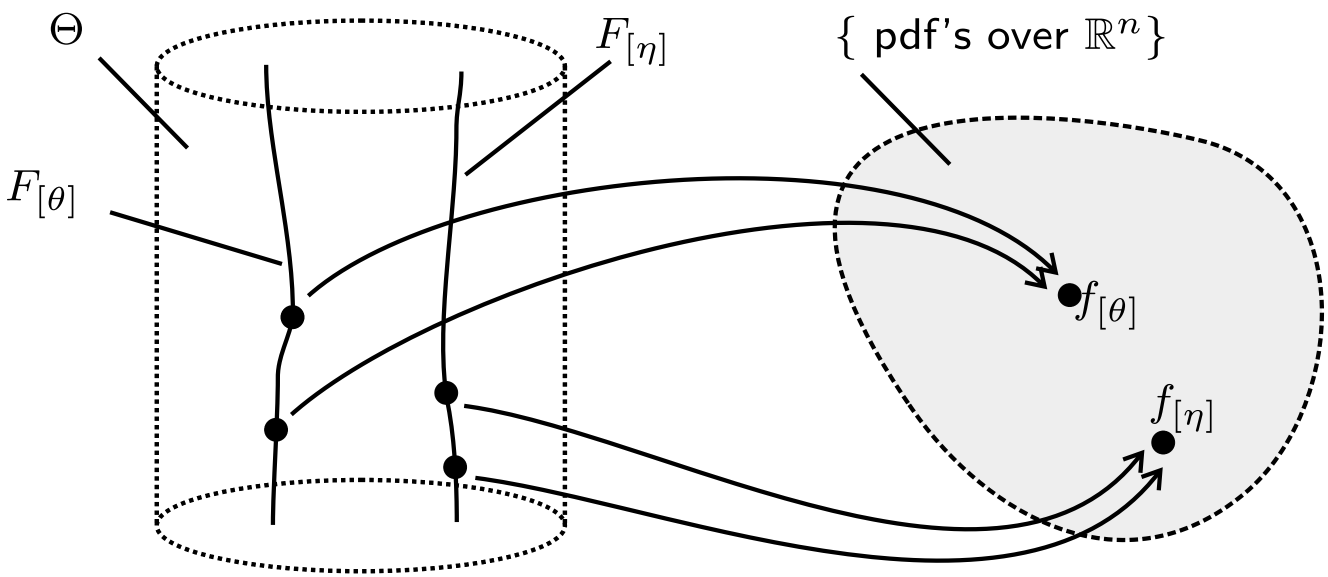


Model Reparameterization

☐ Assumptions (cont):

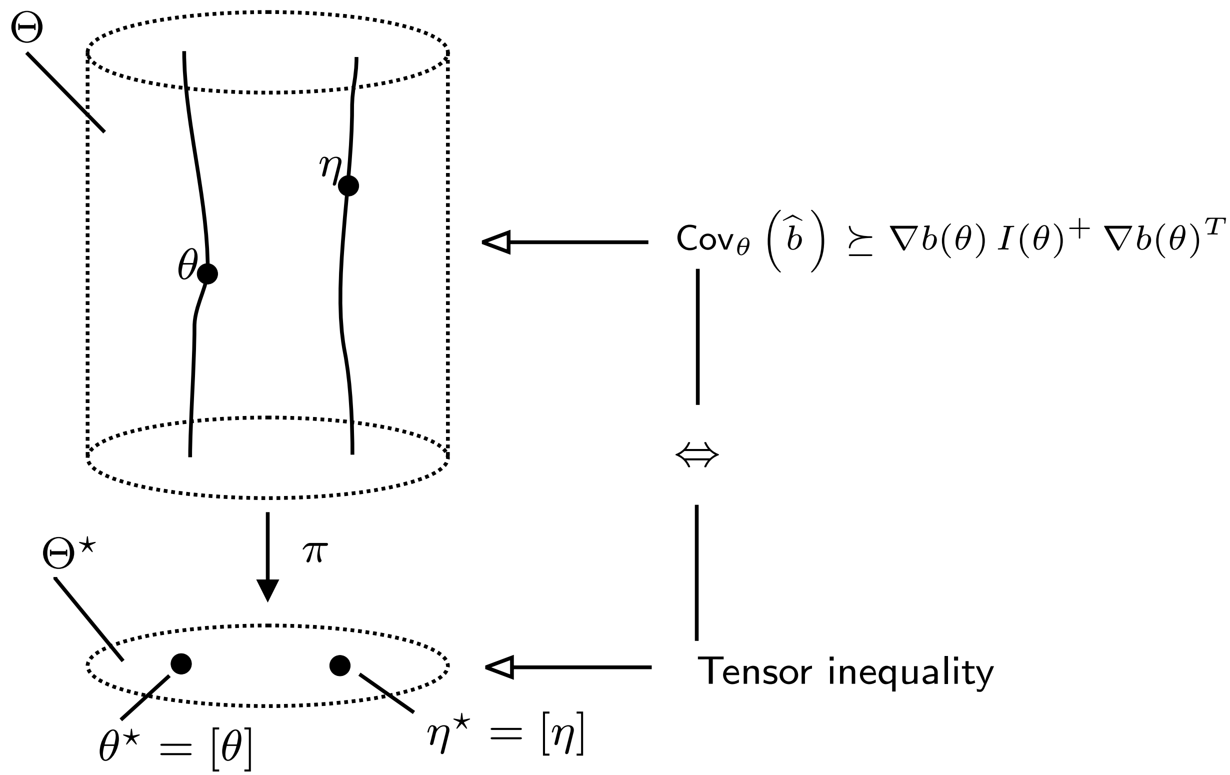
▷ Kernel $I(\theta) = \mathsf{V}_\theta$

∘ **Intuition**: motions along fibers are ambiguous, motions across fibers are not



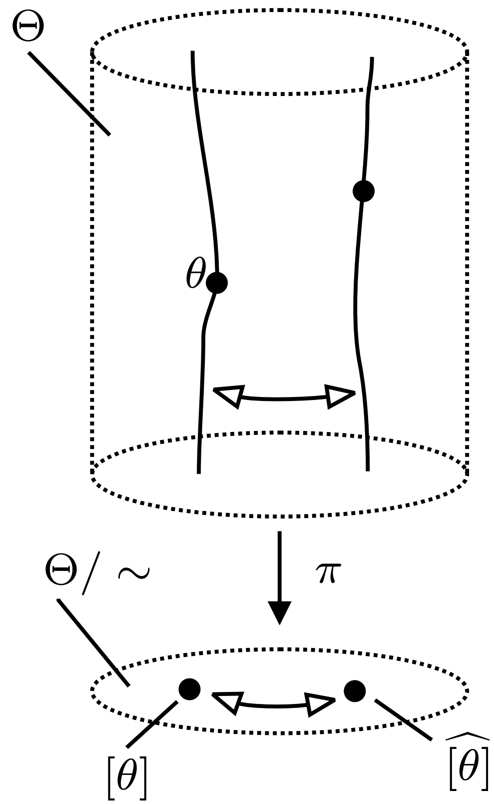
Geometrical Interpretation of Extended CRB

☐ **Result:** it can be shown that ...



☐ **Key-point:** The inequality $\text{Cov}_\theta\left(\widehat{b}\right) \succeq \nabla b(\theta)\, I(\theta)^+\, \nabla b(\theta)^T$ is nothing but the "classical" CRB read in the quotient space $\Theta^\star$

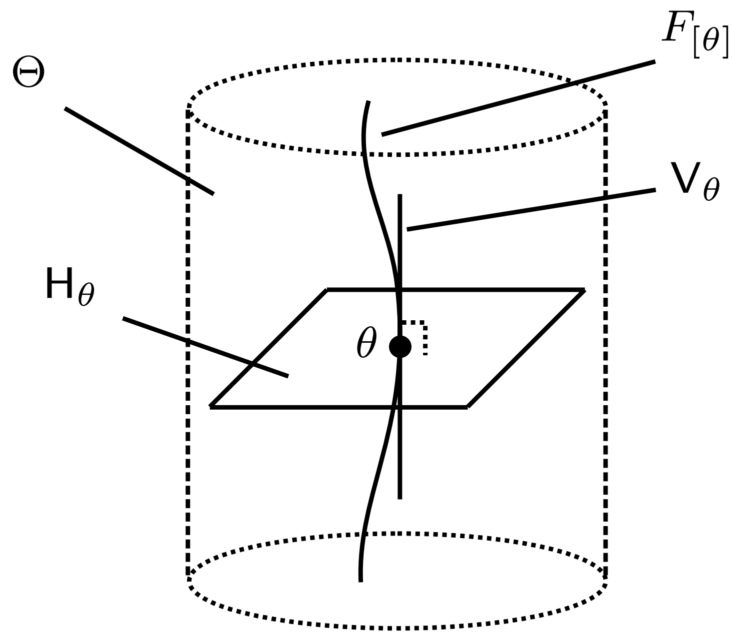
IVLB: Geometrical Interpretation



$$\mathsf{var}_{[\theta]}(\widehat{[\theta]}) = \mathsf{E}_{[\theta]}\left\{ d(\widehat{[\theta]}, [\theta])^2 \right\} \geq \text{IVLB}$$

☐ IVLB lower bounds the accuracy of estimators in discriminating adjacent fibers in Θ

IVLB: Parametric estimation over quotient spaces



- $\lambda_{[\theta]} = \mathrm{tr}\left((U_\theta^T I(\theta) U_\theta)^{-1}\right)$ where:
  - ▷ $I(\theta)$ = Fisher information matrix
  - ▷ U_θ is an orthonormal basis for H_θ

IVLB: Parametric estimation over quotient spaces

☐ Upper-bound on sectional curvatures:

$$C = \max_{[\theta] \in \Theta^\star} K_{[\theta]}(X_{[\theta]}, Y_{[\theta]})$$

where $X_{[\theta]}$ and $Y_{[\theta]}$ are orthonormal vectors in $T_{[\theta]}\Theta^\star$

☐ Since π is a Riemannian submersion, by O'Neill formula:

$$K_{[\theta]}(X_{[\theta]}, Y_{[\theta]}) = \frac{3}{4} |[\widetilde{X}, \widetilde{Y}]^{\mathsf{ver}}|^2$$

where $\widetilde{X}, \widetilde{Y}$ (vector fields) denote horizontal pre-images of $X_{[\theta]}, Y_{[\theta]}$ and $[\cdot, \cdot]$=Lie bracket

☐ Simple formulas for naturally reductive homogeneous spaces (e.g. Grassmann manifold)

Example: inference on Grassmann $\mathrm{G}(4,2)$

□ Array snapshot: $y(t) = Us(t) + w(t) \in \mathbb{R}^4$

- ▷ $U \in \mathbb{R}^{4\times 2}$: **unknown** orthonormal frame ($U^T U = I_2$)
- ▷ $s(t) \in \mathbb{R}^2$: vector of i.i.d., zero-mean, unit-power, Gaussian sources
- ▷ $w(t) \in \mathbb{R}^4$: zero-mean, white spatio-temporal Gaussian noise with power σ^2
- ▷ **Observation:** $y = \mathrm{vec}([\, y(1)\, y(2) \; \cdots \; y(T)\,]) \in \mathbb{R}^{4T}$

□ **Parameter space:** $\Theta = \left\{ U \in \mathbb{R}^{4\times 2} \; : \; U^T U = I_2 \right\}$ [Stiefel manifold]

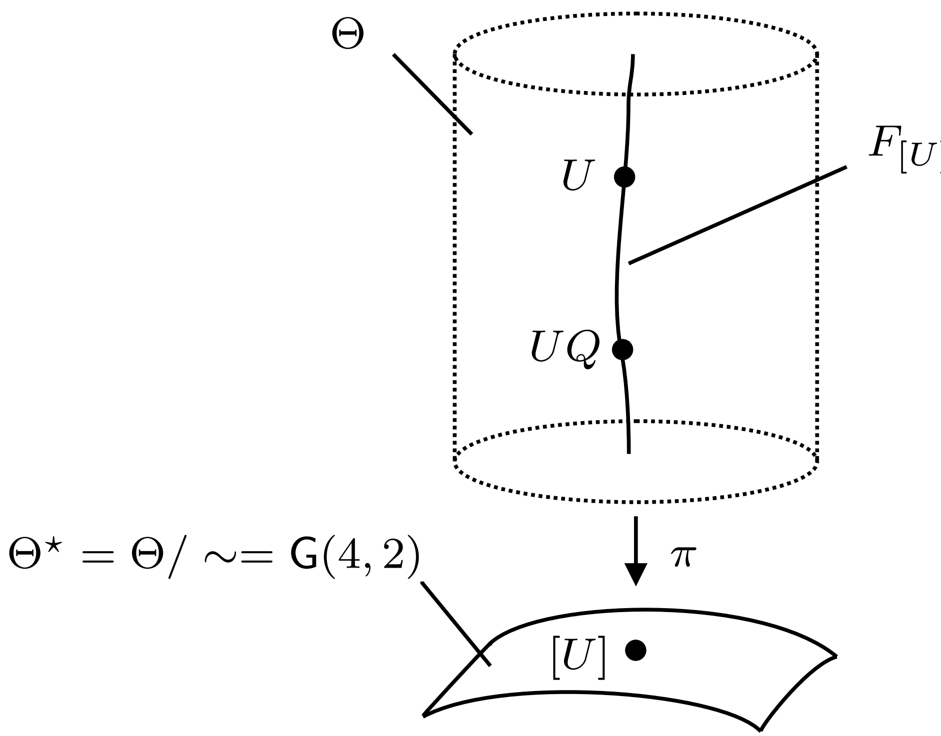
Example: inference on Grassmann $\mathrm{G}(4,2)$

☐ **Ambiguous parameterization:** y is distributed as $\mathcal{N}(0, C(U))$ where

$$C(U) = I_T \otimes (UU^T + \sigma^2 I_4)$$

$C(U) = C(UQ)$ for $QQ^T = I_2 \Rightarrow$ only the 2D-subspace spanned by U is identifiable

☐ **New parameter space:** $\Theta^\star = \Theta/\sim$ where $U \sim V$ iff $U = VQ$ with $QQ^T = I_2$



Example: inference on Grassmann $\mathrm{G}(4,2)$

☐ $\Theta^\star$ can be given the structure of a Riemannian manifold with π a Riemannian submersion

☐ **Geodesic distance on** $\Theta^\star$:

$$d([U],[V]) = \sqrt{2}\,\sqrt{(\mathsf{acos}(\sigma_1))^2 + (\mathsf{acos}(\sigma_2))^2}$$

where σ_1, σ_2 are the singular values of $U^T V$

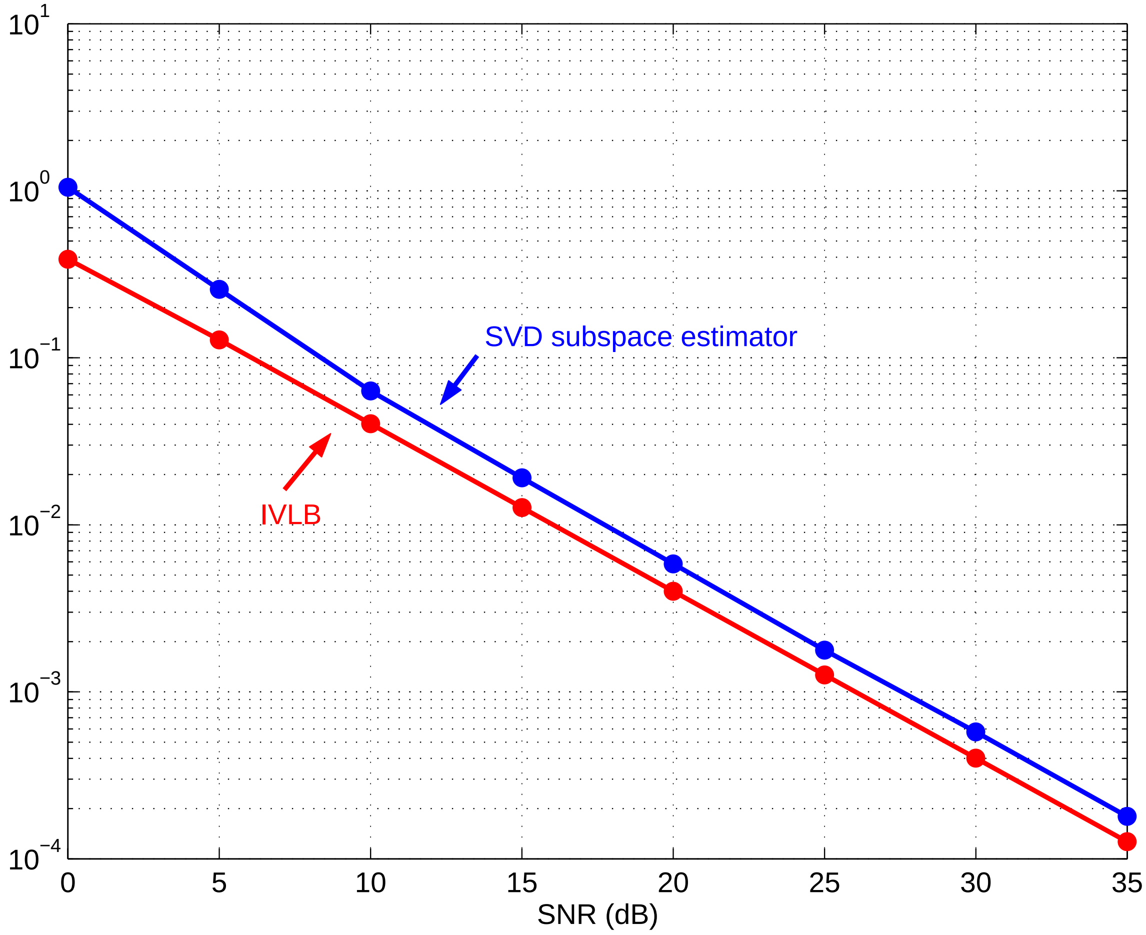
☐ Bound on sectional curvature:

$$C = 1$$

☐ $\widehat{[U]}$ is the dominant 2D-subspace from the SVD of $\widehat{R_y} = \frac{1}{T}\sum_{t=1}^{T} y(t)y(t)^T$

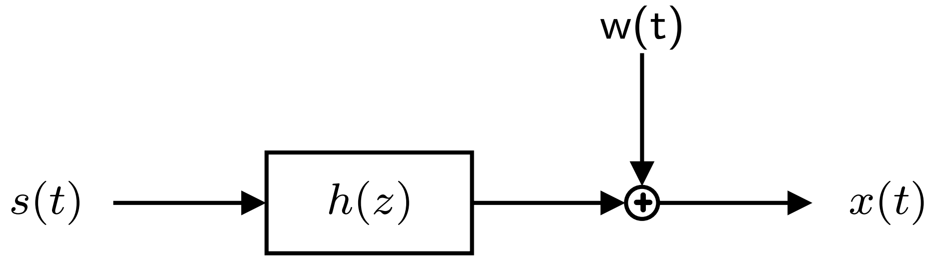
Example: inference on Grassmann $\mathrm{G}(4,2)$

☐ **Example:** $T = 10$ data samples



Example: Blind identification of SIMO channels

☐ Single-Input Multiple Output (SIMO) channel: $y(t) = h(z)s(t) + w(t) \in \mathbb{C}^L$



☐ **Input:** $s(t) \in \mathbb{C}$ is i.i.d., zero-mean, unit-power, complex circular Gaussian

☐ **Finite-impulse channel:** $h(z) = h(0) + h(1)z^{-1} + \cdots + h(D)z^{-D}$

☐ **Noise:** $w(t) \in \mathbb{C}^L$ is zero-mean, white spatio-temporal Gaussian with power σ^2

☐ **Observation:** $y = \text{vec}([\, y(1)\, y(2) \,\cdots\, y(T)\,]) \in \mathbb{C}^{TL}$

☐ **Parameter space:** $\Theta = \{h(z)\} \simeq \{[\, h(0)\, h(1) \,\cdots\, h(D)\,]\} = \mathbb{C}^{DL} - \{0\}$

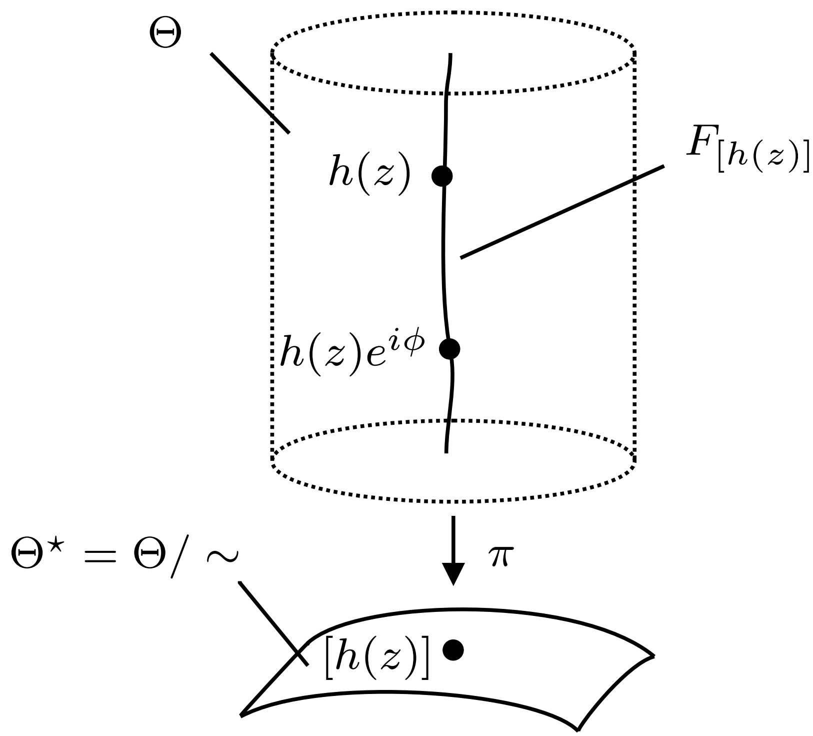
Example: Blind identification of SIMO channels

- **Ambiguous parameterization:** y is distributed as $\mathcal{N}(0, C(h(z))$ where

$$C(h(z)) = C(h(z)e^{i\phi})$$

Meaning: channel is identifiable only up to a phase ambiguity

- **New parameter space:** $\Theta^\star = \Theta/\sim$ where $h(z) \sim g(z)$ iff $h(z) = g(z)e^{i\phi}$



Example: Blind identification of SIMO channels

□ $\Theta^\star$ can be given the structure of a Riemannian manifold with π a Riemannian submersion

□ **Geodesic distance on** $\Theta^\star$:

$$d([h(z)],[g(z)]) = \sqrt{\|H\|^2 + \|G\|^2 - 2|\mathsf{tr}(G^H H)|}$$

where $H = [\, h(0)\, h(1)\, \cdots\, h(D)\,]$ and $G = [\, g(0)\, g(1)\, \cdots\, g(D)\,]$

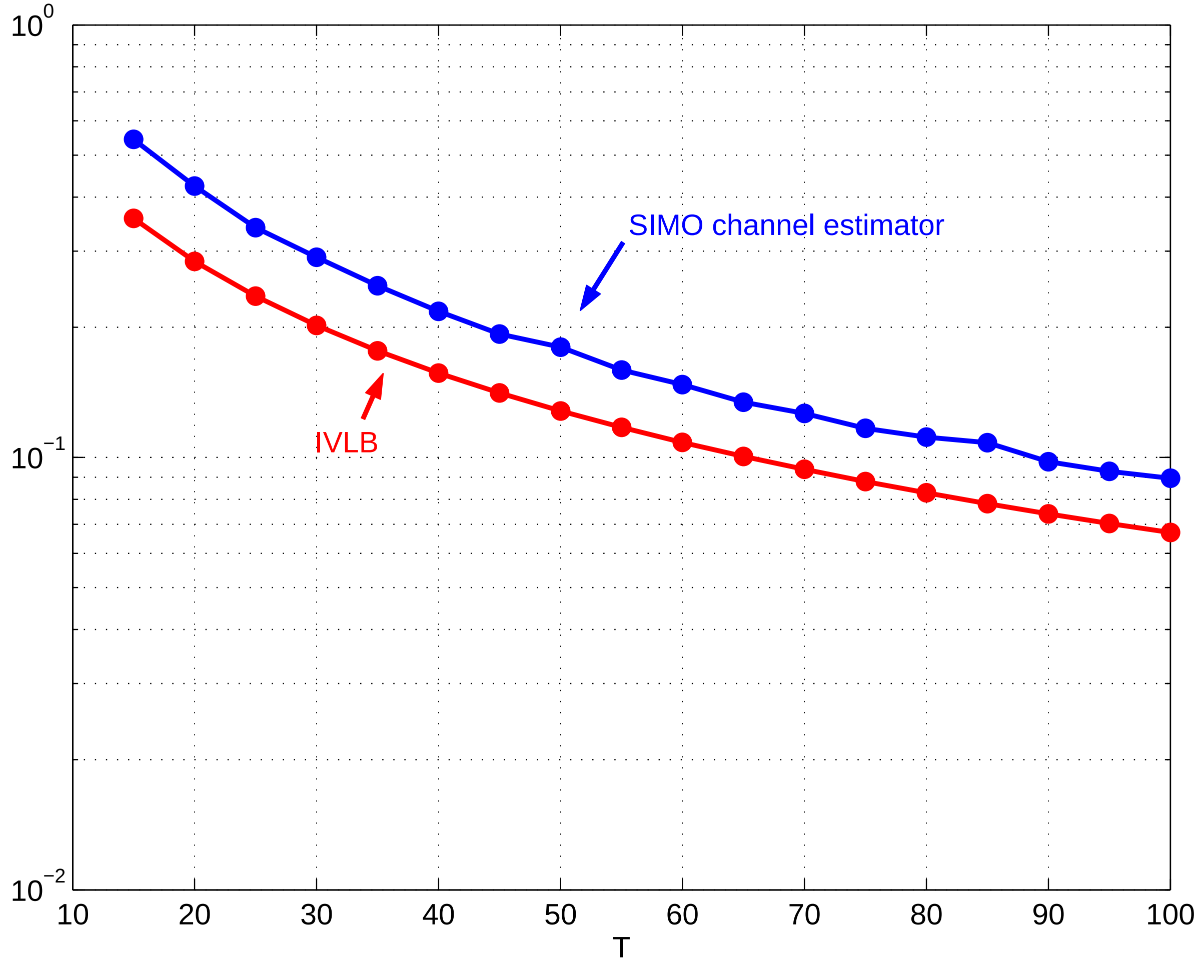
□ Bound on sectional curvature:

$$K_{[h(z)]} \leq \frac{3}{\|H\|^2}$$

□ $\widehat{[h(z)]}$ is the subspace channel estimator [Moulines et al'95, *IEEE SP*]

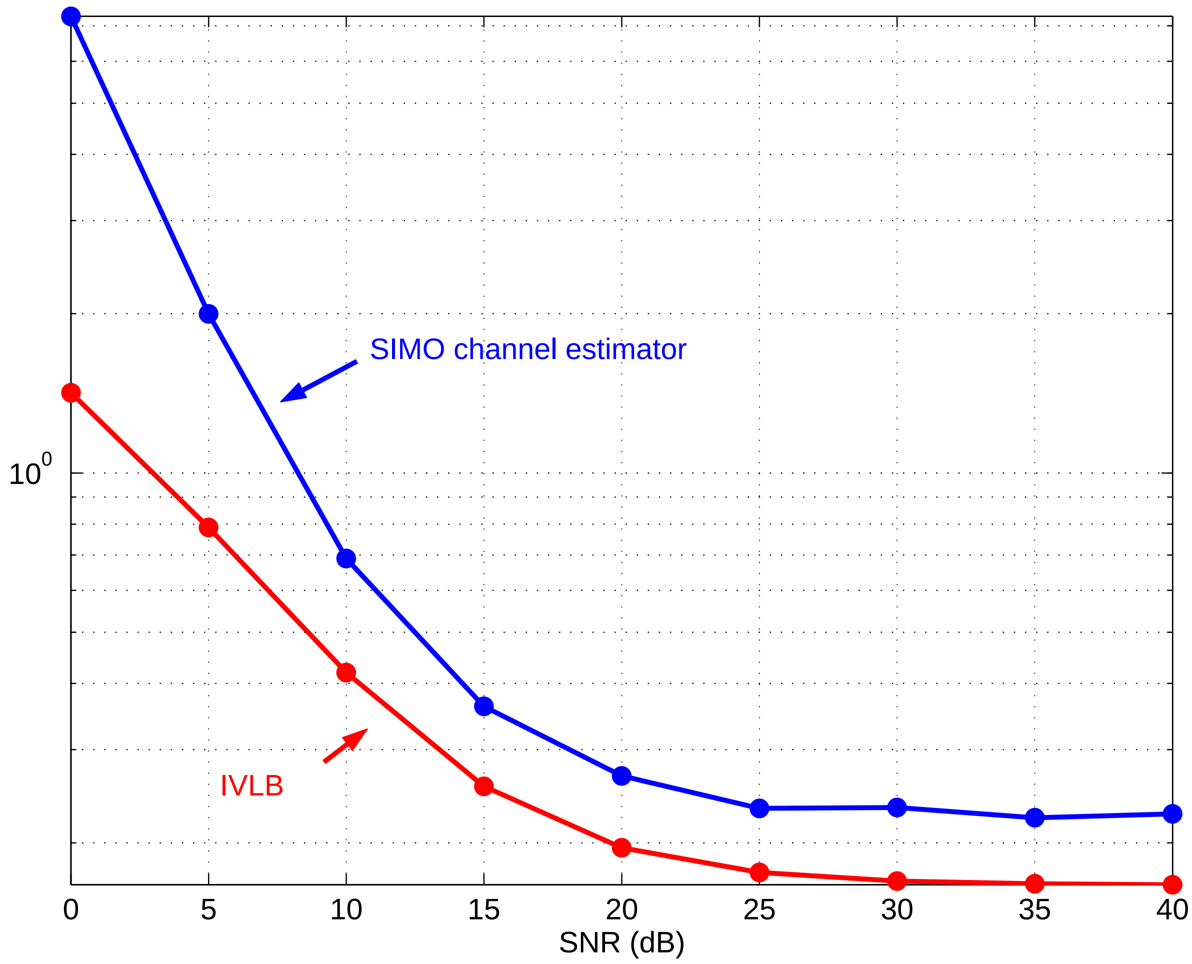
Example: Blind identification of SIMO channels

☐ **Example:** $L = 4$ antennas, $D = 2$ (filter memory), SNR$=20$dB



Example: Blind identification of SIMO channels

☐ **Example:** $L = 4$ antennas, $D = 2$ (filter memory), T=25 samples



Outline

☐ **Motivation:** Riemannian manifolds in parametric estimation problems

☐ **Performance bounds for parametric estimation:** Cramér-Rao bound & extensions

☐ **Framework**: Parametric statistical models over Riemannian manifolds

☐ **Contribution**: Intrinsic Variance Lower Bound (IVLB)

☐ **Applications of IVLB**:

▷ Parametric estimation with constraints

▷ Parametric estimation over quotient spaces

☐ **Open problems**

Open problems

- Interesting directions for research:
  - Explore other applications (blind source separation, image, etc)
  - Extend other Euclidean bounds: Ziv-Zakai, Barankin, etc
  - Bayesian (non-deterministic) setups