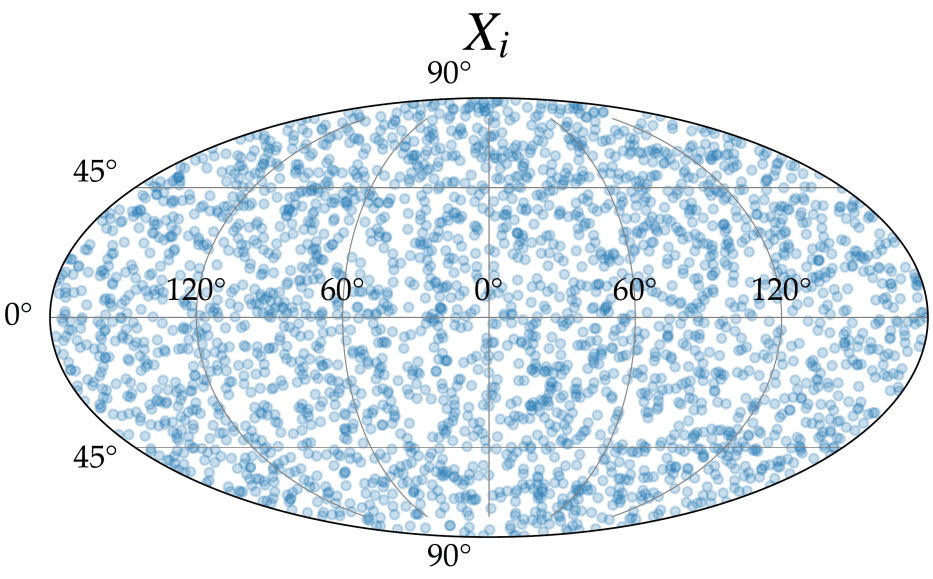
Nonparametric Riemannian empirical Bayes, and denoising measurements on manifolds

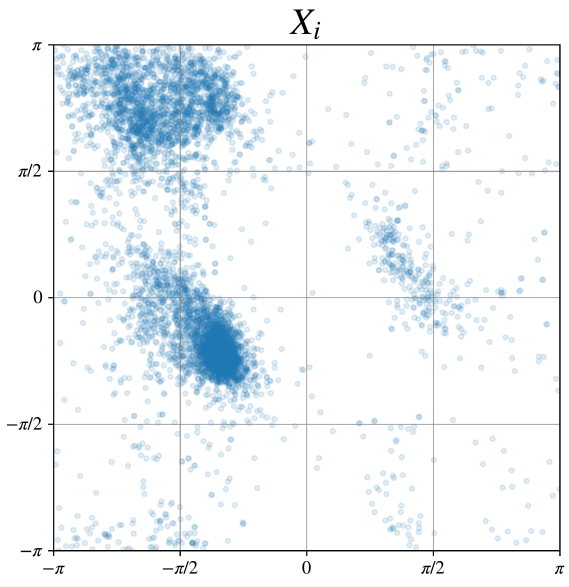
Adam Quinn Jaffe

With Leonardo Santoro and Bodhisattva Sen

Denoising manifold-valued data?



gamma ray burst localization



Ramachandran plot

Fix a known compact Riemannian manifold $\mathcal{M}$, geodesic distance d.

Model for data on $\mathcal{M}$:

$$\Theta \sim G \qquad \text{and} \qquad (X \,|\, \Theta = \theta) \sim p_{\theta,\sigma^2}(\cdot)$$

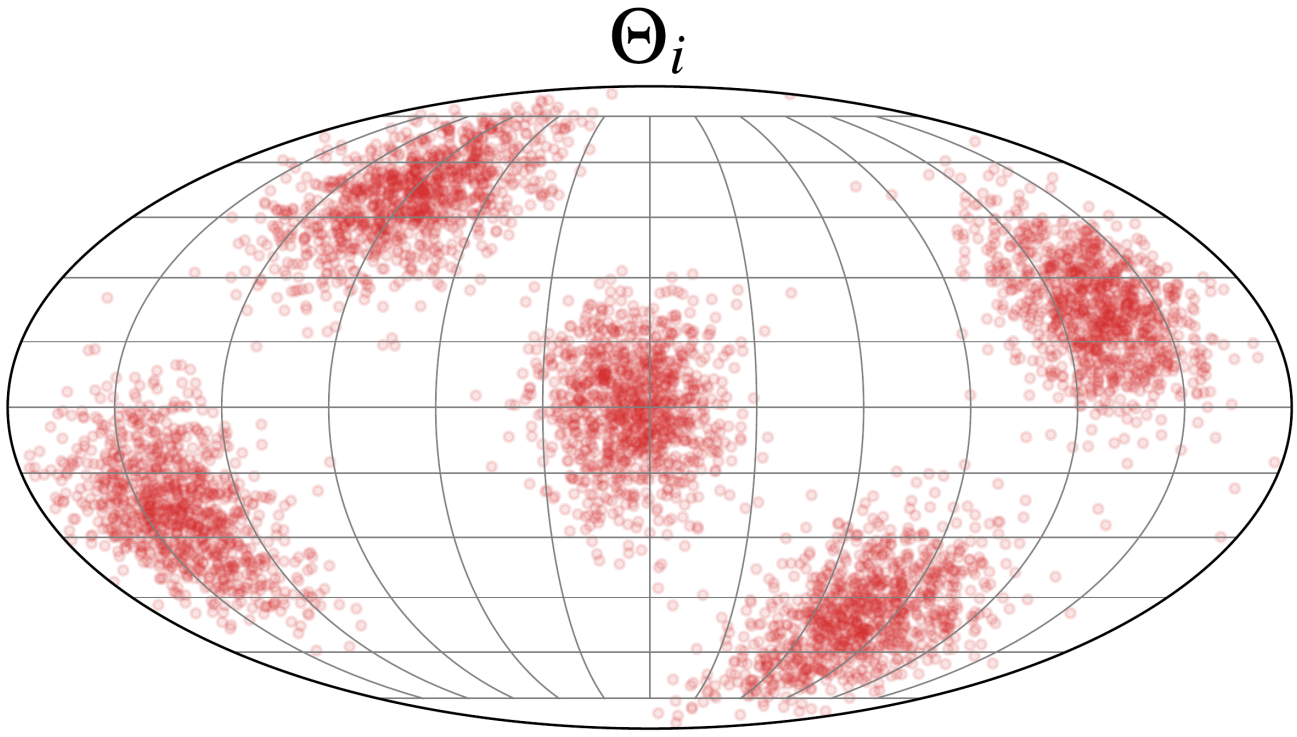
where G is unknown distribution on $\mathcal{M}$, and $\sigma^2 > 0$ is known.

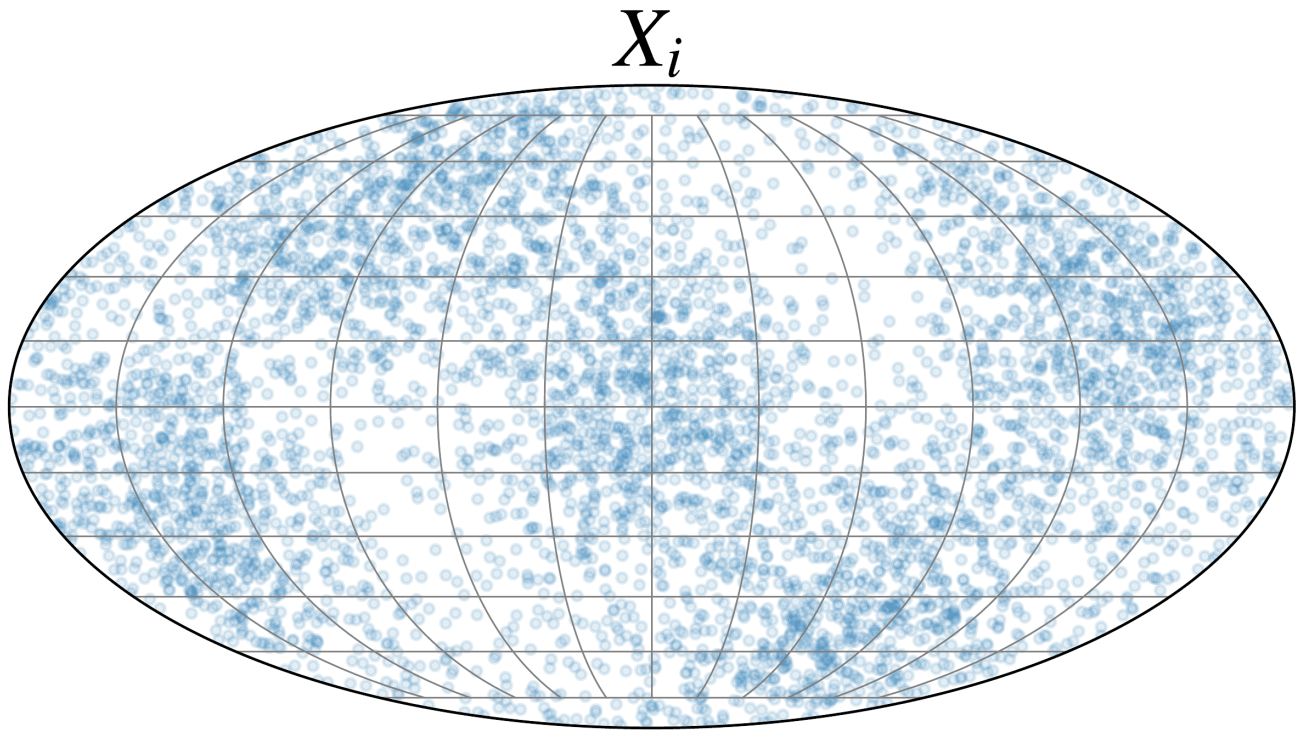
Riemannian Gaussian distribution (Pennec 2006) has density

$$p_{\theta,\sigma^2}(x) \propto \exp\left(-\frac{d^2(x,\theta)}{2\sigma^2}\right)$$

with respect to Riemannian volume form.

If $(\Theta_1, X_1), \ldots (\Theta_n, X_n)$ are i.i.d. samples from above, the *denoising* problem is to estimate the latent variables $\Theta_1, \ldots, \Theta_n$ from the measurements $X_1, \ldots, X_n$


Θ_i


X_i

Mathematically, goal is to minimize risk:

$$\begin{cases} \text{minimize} & \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} d^2(\delta(X_i), \Theta_i)\right] \\ \text{over} & \delta : \mathcal{M} \to \mathcal{M} \end{cases}$$

Solution is the posterior Fréchet mean,

$$\delta_{\mathcal{B}}(x) \in \operatorname*{arg\,min}_{\vartheta \in \mathcal{M}} \mathbb{E}\left[d^2(\vartheta, \Theta) \,|\, X = x\right],$$

called the *oracle Bayes denoiser* since it depends on the unknown G.

We can still try to estimate $\delta_{\mathcal{B}}$ directly from the data, leading to *empirical Bayes denoisers* (Robbins 1956, Efron 2011).

Baseline is the *naive denoiser* $\delta_{\mathcal{N}}(x) := x$, which coincides with the Bayes denoiser if $\mathcal{M}$ is homogeneous and G is uniform

Let G be distribution of Θ, let $f = f_G$ be marginal density of X.

In Euclidean case $\mathcal{M} = \mathbb{R}^m$, we have the *Tweedie-Eddington formula*

$$\delta_{\mathcal{B}}(x) = \mathbb{E}[\Theta \,|\, X = x] = x + \sigma^2 \nabla \log f(x)$$

which shows that $\delta_{\mathcal{B}}(x)$ is the naïve denoiser plus a correction.

Oracle Bayes denoiser depends on G only through f! So, estimate:

$$\hat{\delta}_{\mathcal{B}}(x) = x + \sigma^2 \nabla \log \hat{f}_n(x).$$

Parametric rate of convergence of $\hat{\delta}_{\mathcal{B}}$ to $\delta_{\mathcal{B}}$ (Zhang 1997).

How to do this in the Riemannian case, where oracle Bayes denoiser depends on G implicitly through Riemannian optimization?

Idea: consider an intermediate oracle denoiser $\delta_{\mathcal{T}}$ which depends on G only through f hence can be estimated by plug-in as $\hat{\delta}_{\mathcal{T}}$.

Then we will show:

Small-noise result, stating $\delta_{\mathcal{T}} \approx \delta_{\mathcal{B}}$ as $\sigma^2 \to 0$

Large-sample result, stating $\hat{\delta}_{\mathcal{T}} \approx \delta_{\mathcal{T}}$ as $n \to \infty$

This shows that $\hat{\delta}_{\mathcal{T}}$ has nearly optimal risk if σ^2 is small and n is large:

$$\hat{\delta}_{\mathcal{T}} \overset{n\to\infty}{\approx} \delta_{\mathcal{T}} \overset{\sigma^2\to 0}{\approx} \delta_{\mathcal{B}}$$

Related work:

- Nonparametric empirical Bayes in Euclidean setting (Zhang 1997, Jiang-Zhang 2009, Saha-Guntuboyina 2020, Soloff et al. 2025, ...)
- Shrinkage in nonpositive curvature (McCormack-Hoff 2022)
- Parametric empirical Bayes for matrices (Yang-Doss-Vemuri 2024)

Closely related nonparametric statistics on manifolds:

- Density estimation (Hendriks 1990, Pelletier 2005)
- Deconvolution (Healy-Kim 1996, Kim 1998, Huckemann 2010)
- Score matching (De Bortoli et al. 2022)

I. **Problem Statement**

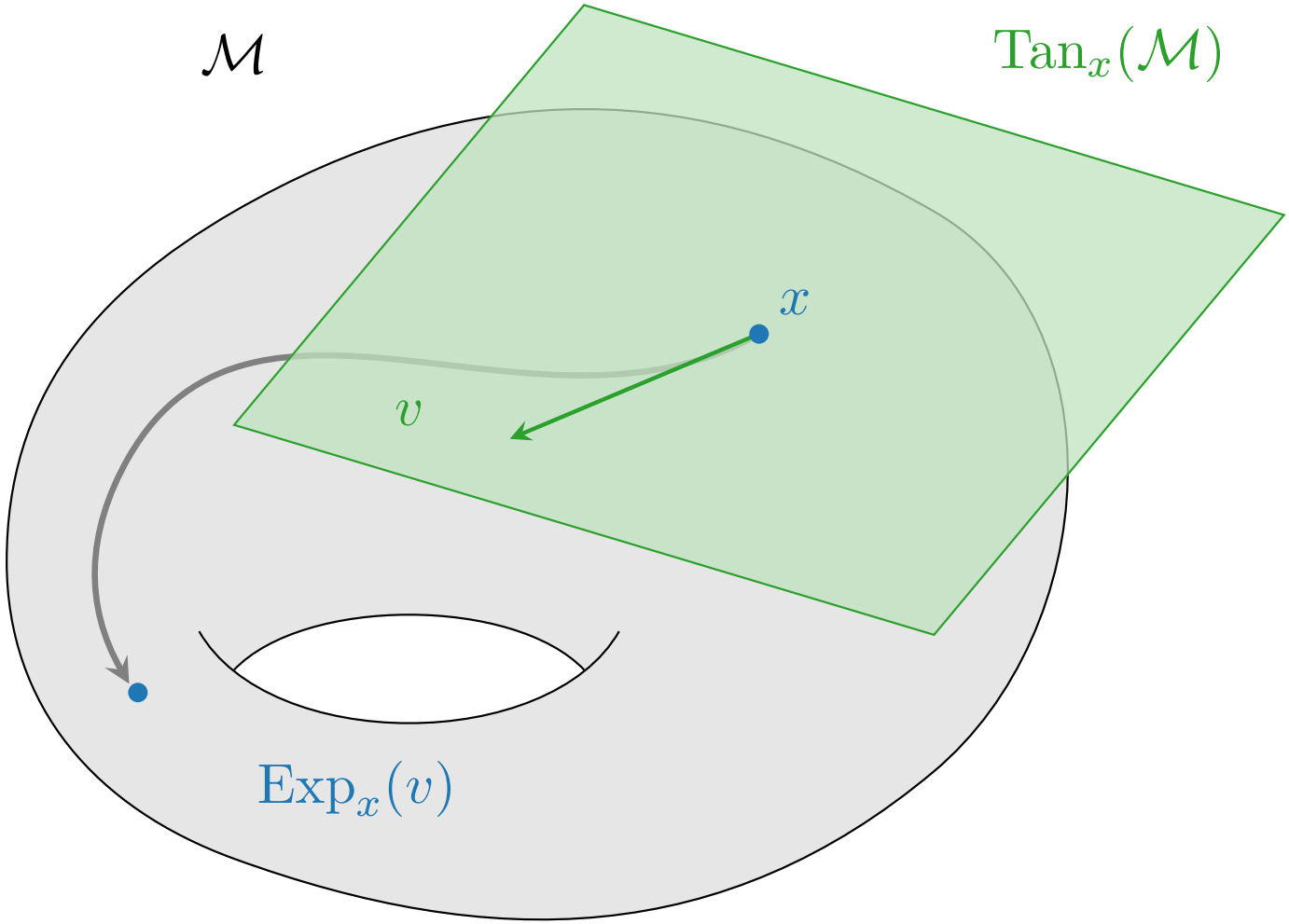
II. **Small Noise Regime, $\delta_{\mathcal{T}} \approx \delta_{\mathcal{B}}$**

III. **Large Sample Regime, $\hat{\delta}_{\mathcal{T}} \approx \delta_{\mathcal{T}}$**

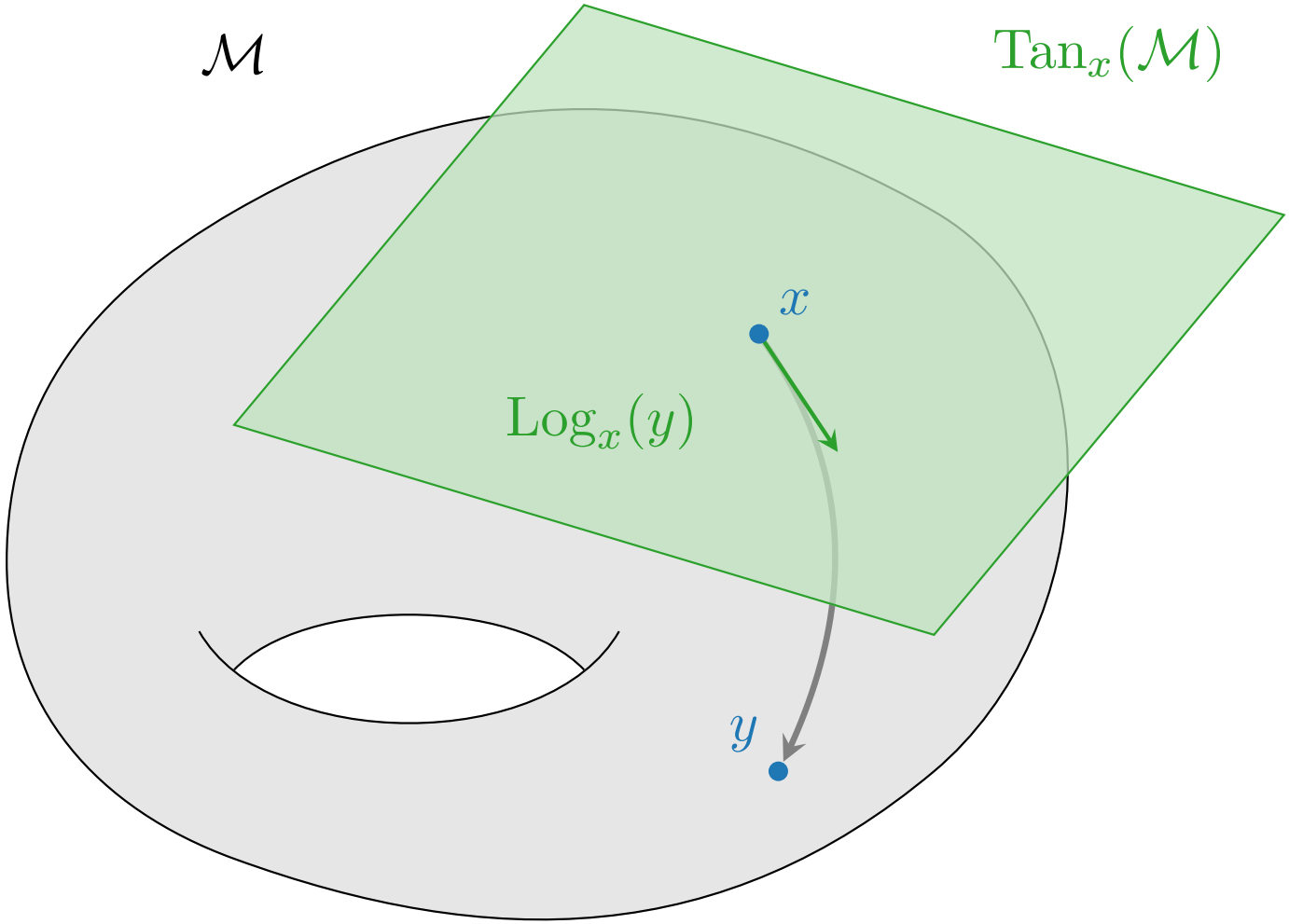
IV. **Applications**

II. Small Noise Regime, $\delta_{\mathcal{T}} \approx \delta_{\mathcal{B}}$

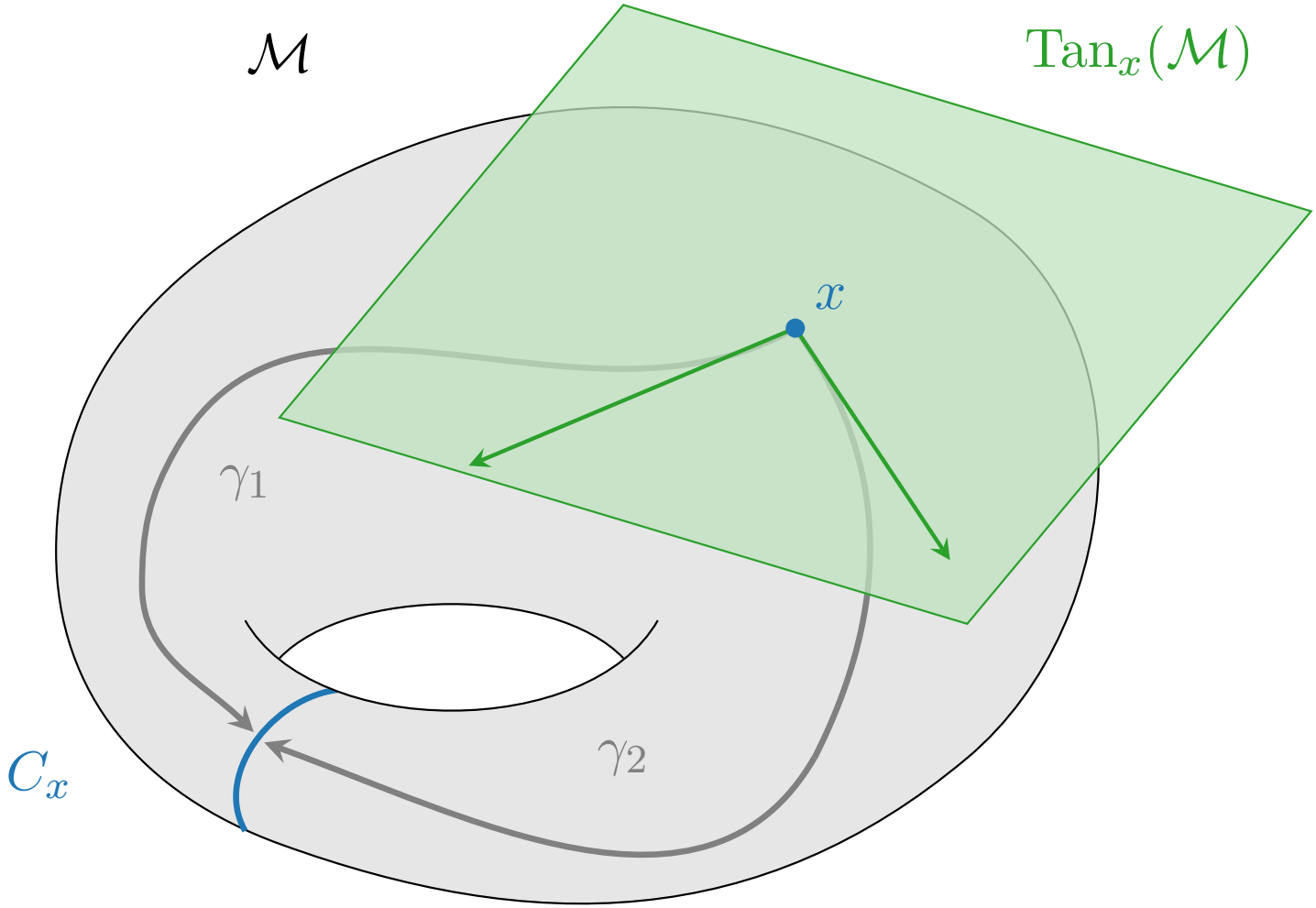
Geometric background:



Geometric background:



Geometric background:



Lemma (AQJ-Santoro-Sen, 2026)

If G has a density with respect to Riemannian volume form, then f is differentiable and

$$\mathbb{E}\left[\operatorname{Log}_x \Theta \,|\, X = x\right] = \sigma^2 \nabla_x \log f(x).$$

for all $x \in \mathcal{M}$.

Similar to Tweedie-Eddington but not quite

Proof sketch. Differentiate $f(x) = \int_{\mathcal{M}} p_{\theta,\sigma^2}(x)\, \mathrm{d}G(\theta)$ under the integral, and use the well-known fact $\nabla_x \frac{1}{2} d^2(x,\theta) = -\operatorname{Log}_x \theta$ for $x \notin C_\theta$

Define the *oracle tangential Bayes denoiser*:

$$\delta_{\mathcal{T}}(x) := \operatorname{Exp}_x \left(\mathbb{E} \left[\operatorname{Log}_x \Theta \,|\, X = x \right] \right) = \operatorname{Exp}_x \left(\sigma^2 \nabla_x \log f(x) \right).$$

Like Euclidean case, $\delta_{\mathcal{T}}$ depends on G only through f.

If $\mathcal{M} = \mathbb{R}^m$, then $\delta_{\mathcal{T}} = \delta_{\mathcal{B}}$ is Tweedie-Eddington formula.

If $\mathcal{M}$ is homogeneous and G is uniform, then $\delta_{\mathcal{N}} = \delta_{\mathcal{B}} = \delta_{\mathcal{T}}$.

In general, $\delta_{\mathcal{T}} \neq \delta_{\mathcal{B}}$.

Why "tangential"? Define conditional Fréchet functional:

$$\Phi_x(\vartheta) := \mathbb{E}\left[d^2(\vartheta, \Theta) \,|\, X = x\right],$$

so that

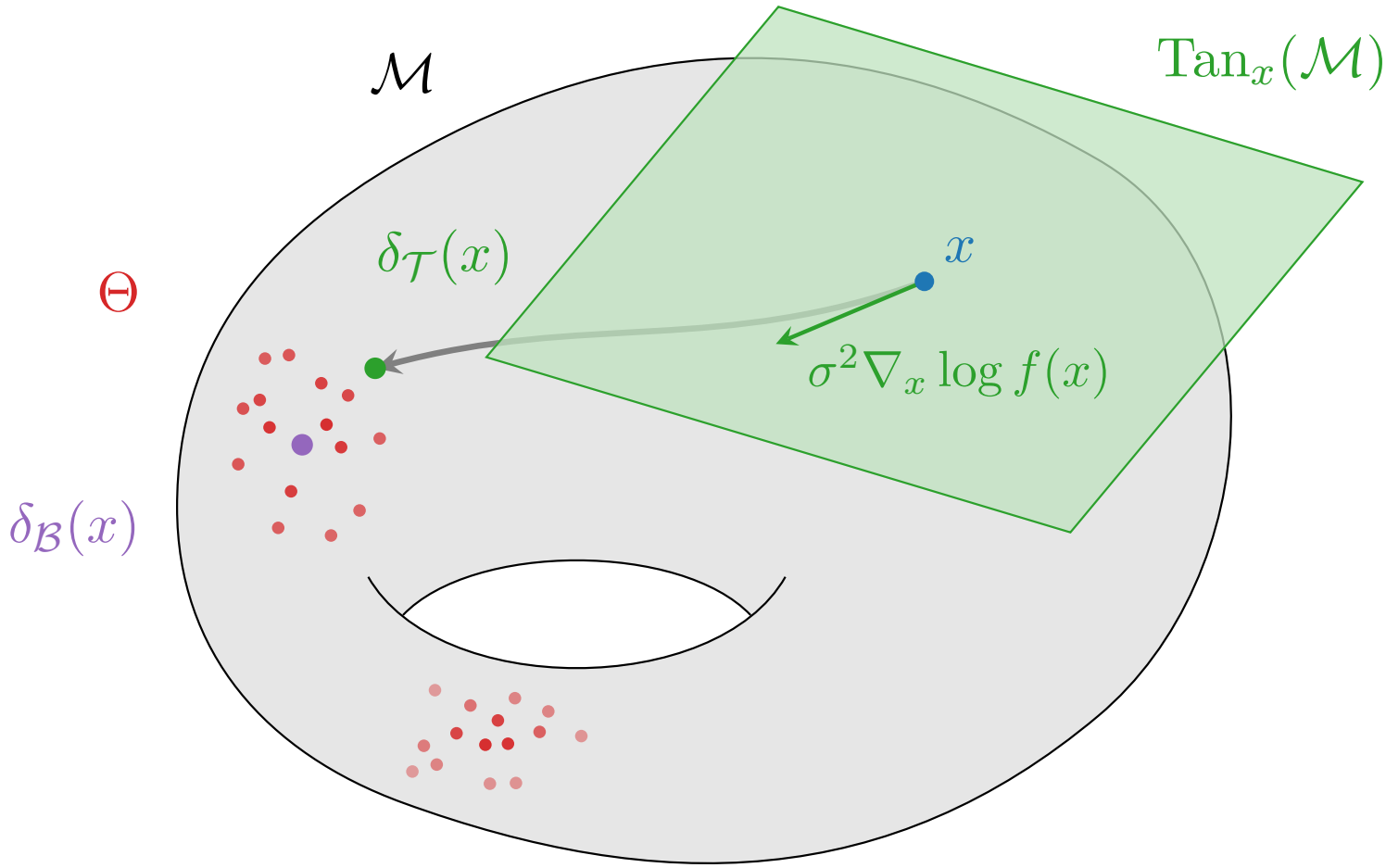
$$\nabla_\vartheta \Phi_x(\vartheta) = -2\,\mathbb{E}\left[\mathrm{Log}_\vartheta\, \Theta \,|\, X = x\right],$$

hence

$$\delta_{\mathcal{T}}(x) = \mathrm{Exp}_x\left(-\tfrac{1}{2}\nabla_\vartheta \Phi_x(x)\right).$$

This says $\delta_{\mathcal{T}}(x)$ is the result of one step of Riemannian gradient descent for Φ_x, with stepsize $\frac{1}{2}$ and initialization at $\delta_{\mathcal{N}}(x) = x$ (Boumal 2023).

So $\delta_{\mathcal{T}}(x)$ is a one-step improvement of the naïve baseline $\delta_{\mathcal{N}}(x)$


$\mathcal{M}$
$\mathrm{Tan}_x(\mathcal{M})$
$\delta_{\mathcal{T}}(x)$
x
Θ
$\sigma^2 \nabla_x \log f(x)$
$\delta_{\mathcal{B}}(x)$

Define the risk $R(\delta, \sigma^2) := \mathbb{E}\left[d^2(\delta(X), \Theta)\right]$

Theorem (AQJ-Santoro-Sen, 2026)

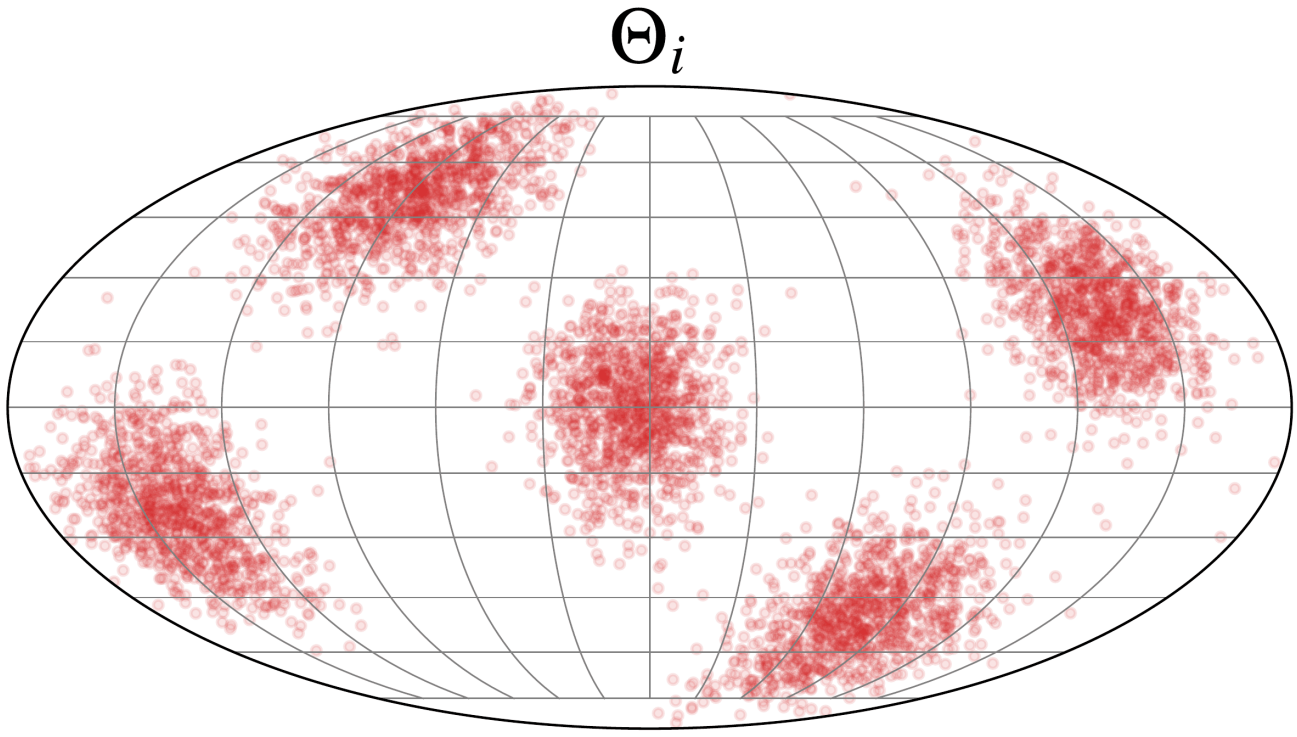
If G has a strictly positive, continuously differentiable density with respect to the Riemannian volume form, then

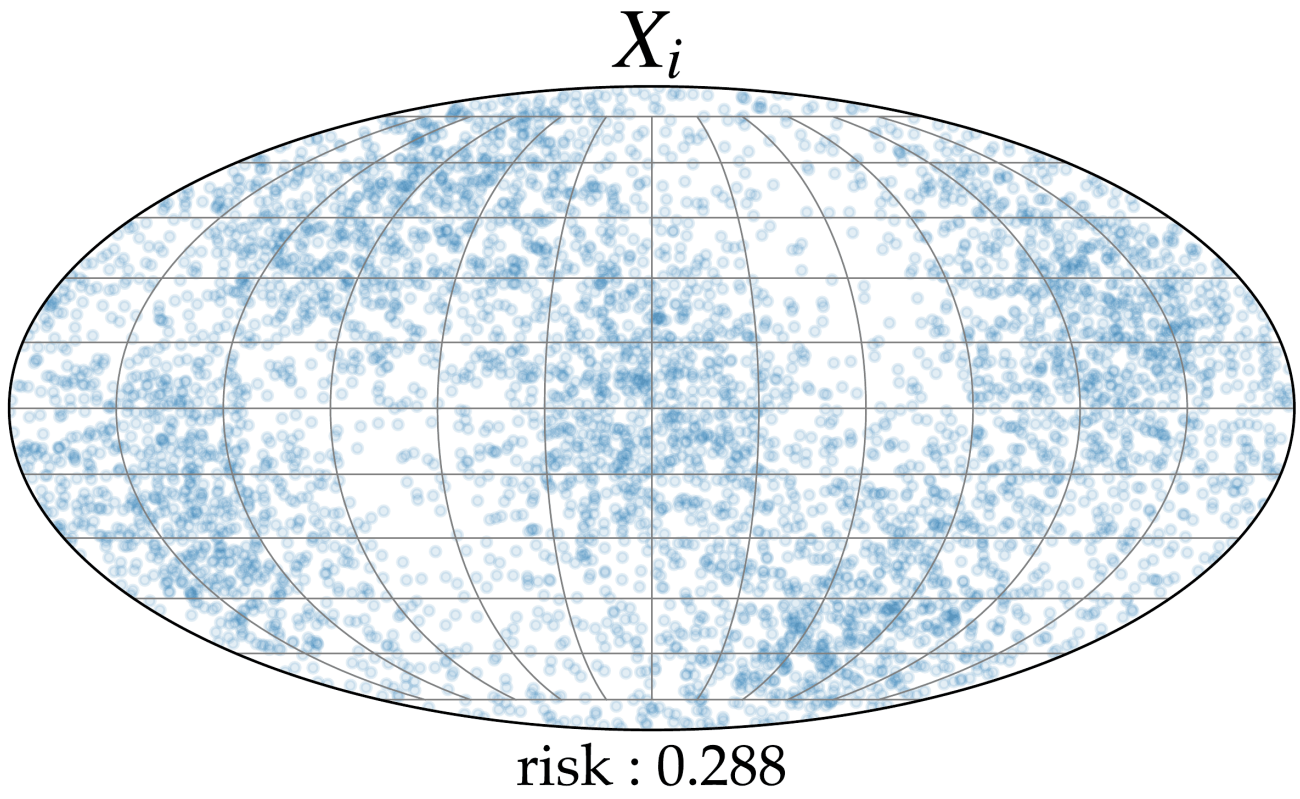
$$0 \leq R(\delta_{\mathcal{T}}, \sigma^2) - R(\delta_{\mathcal{B}}, \sigma^2) = O(\sigma^5)$$

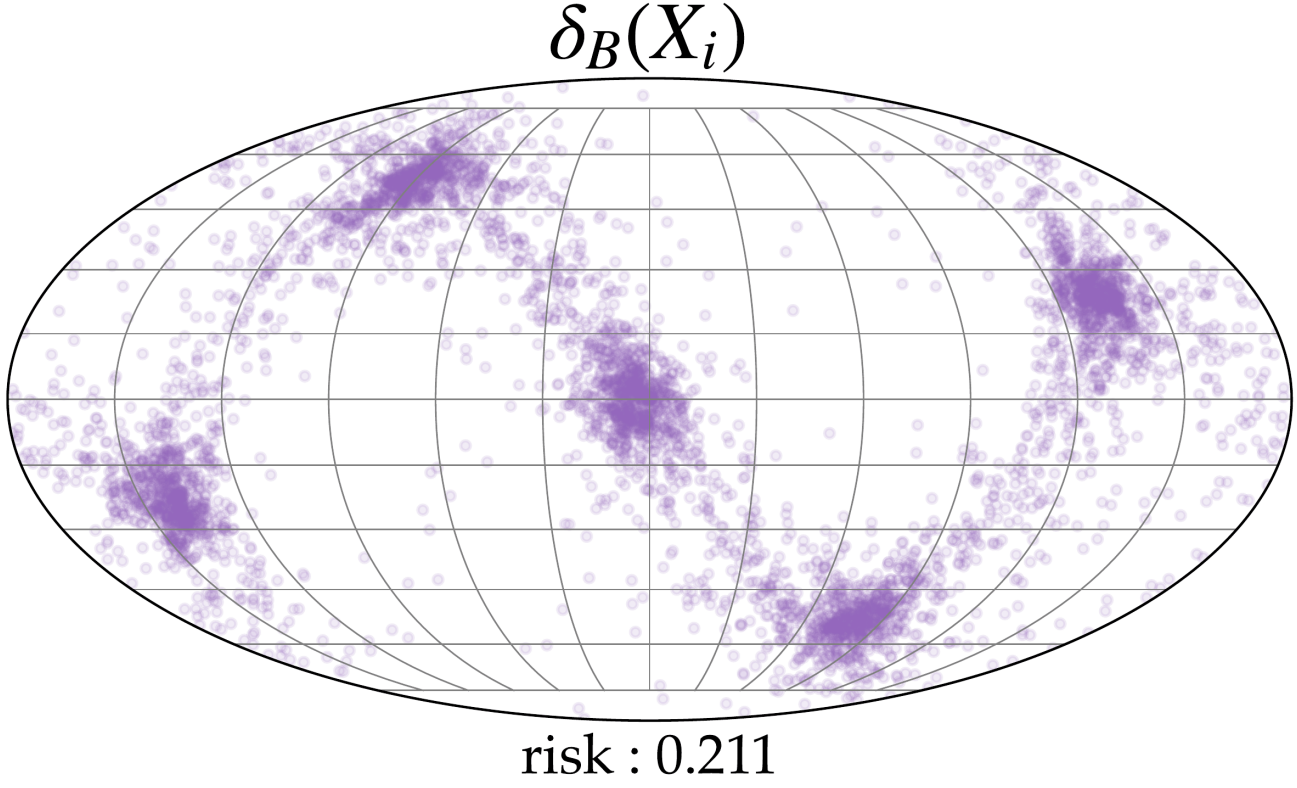
as $\sigma^2 \to 0$.

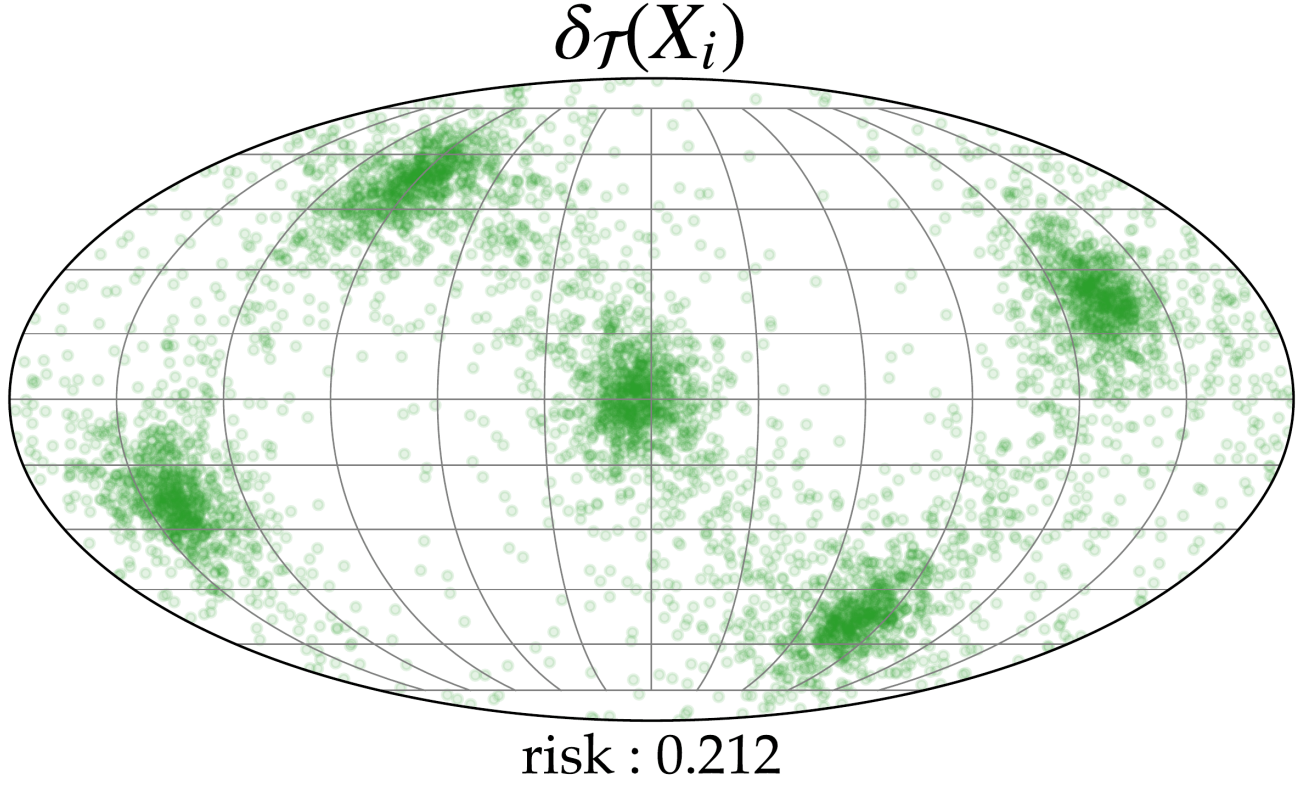
For comparison, $R(\delta_{\mathcal{N}}, \sigma^2) \sim \dim(\mathcal{M})\sigma^2$ as $\sigma^2 \to 0$.

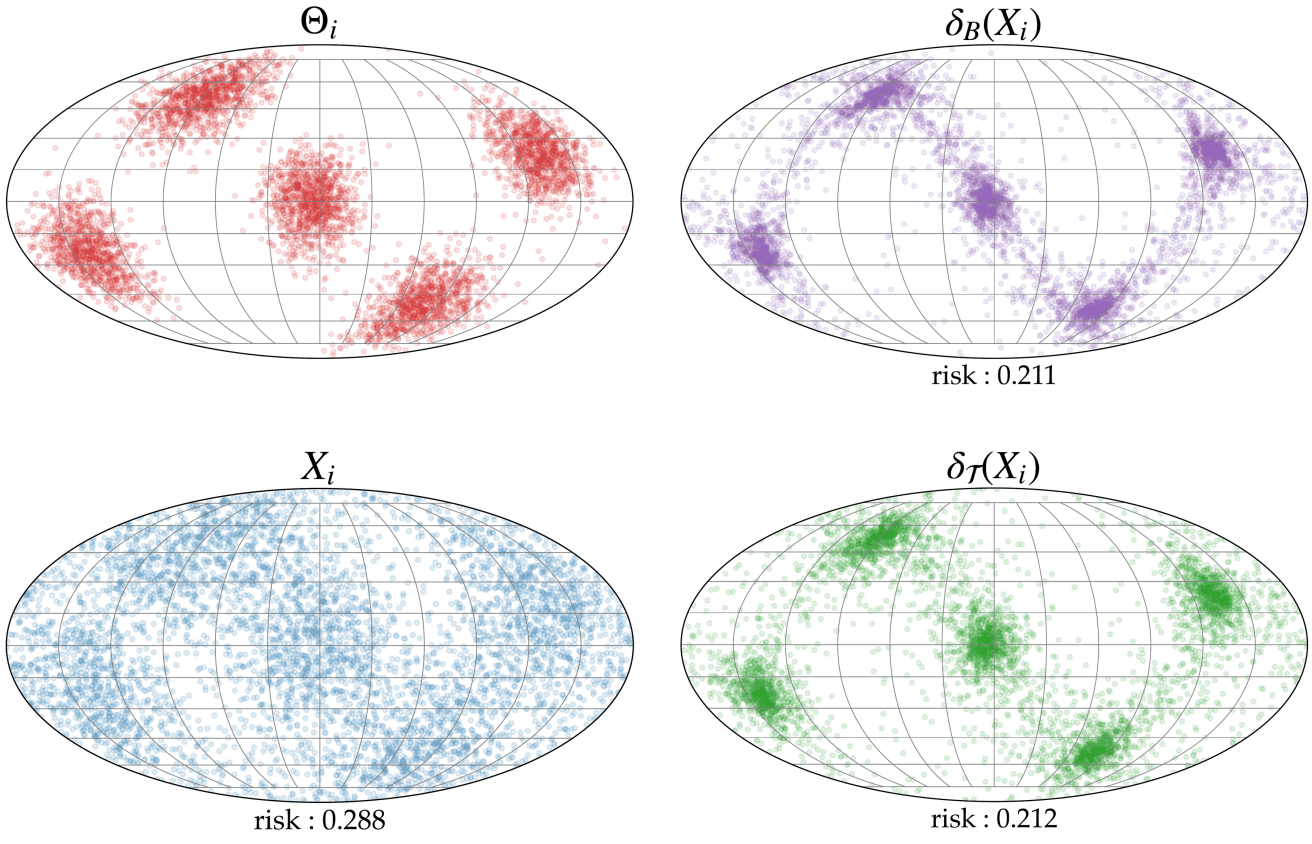
So risks of $\delta_{\mathcal{T}}$ and $\delta_{\mathcal{B}}$ are very close!

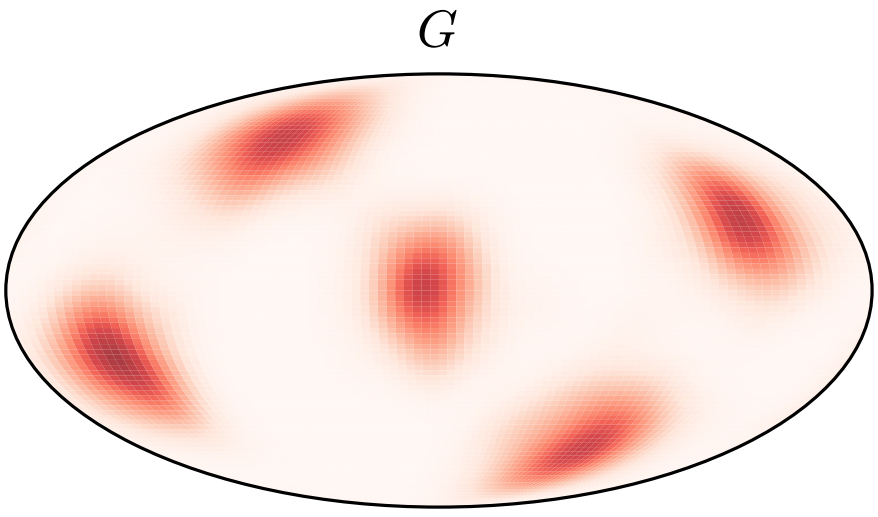

Θ_i

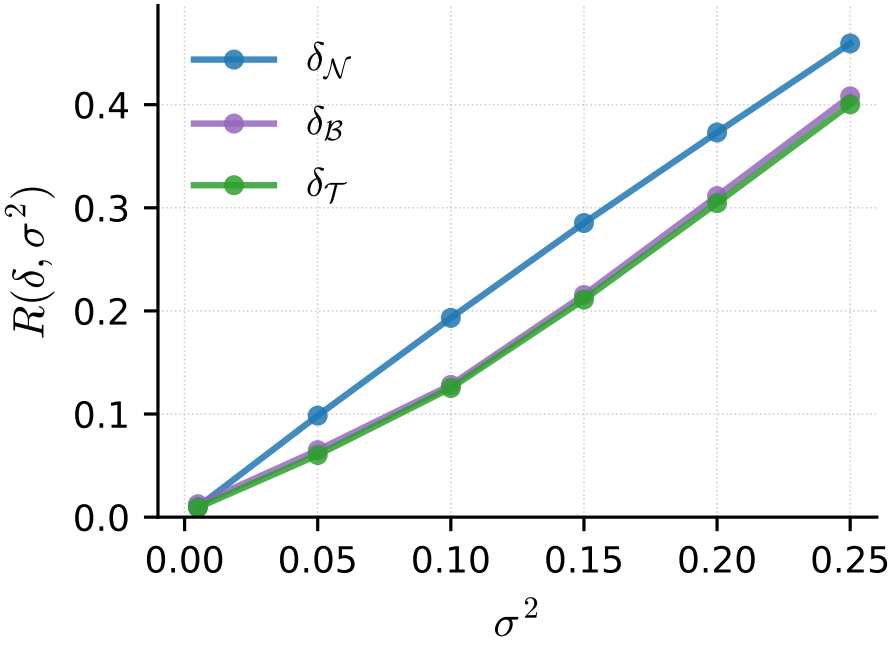

X_i
risk : 0.288

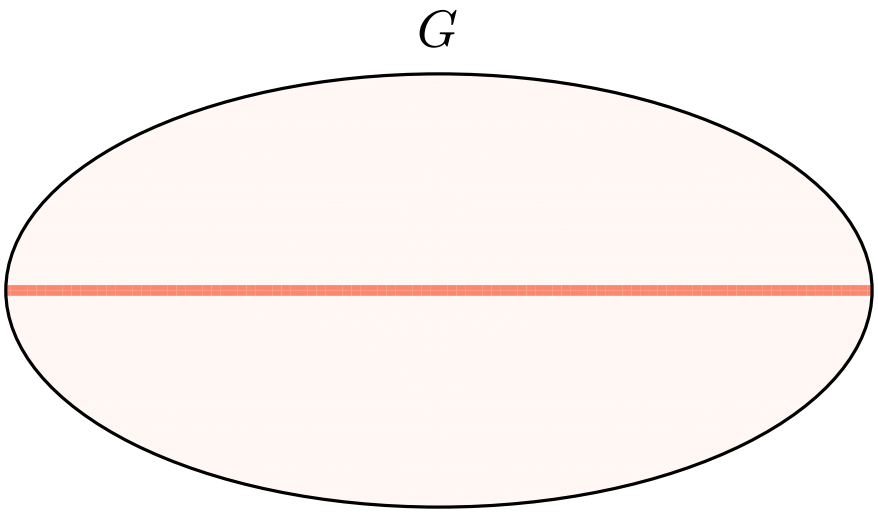

$\delta_B(X_i)$
risk : 0.211

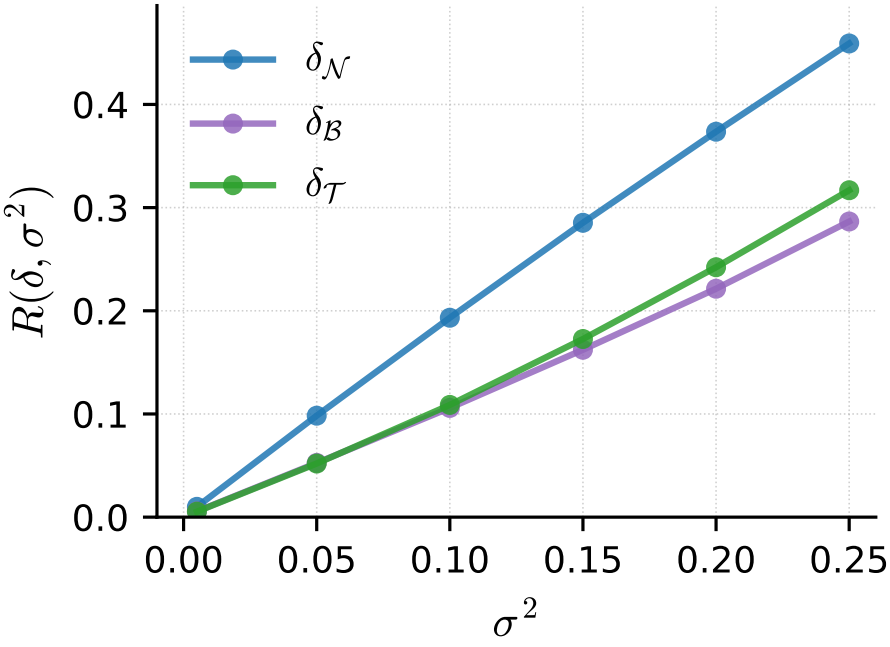

$\delta_{\mathcal{T}}(X_i)$
risk : 0.212

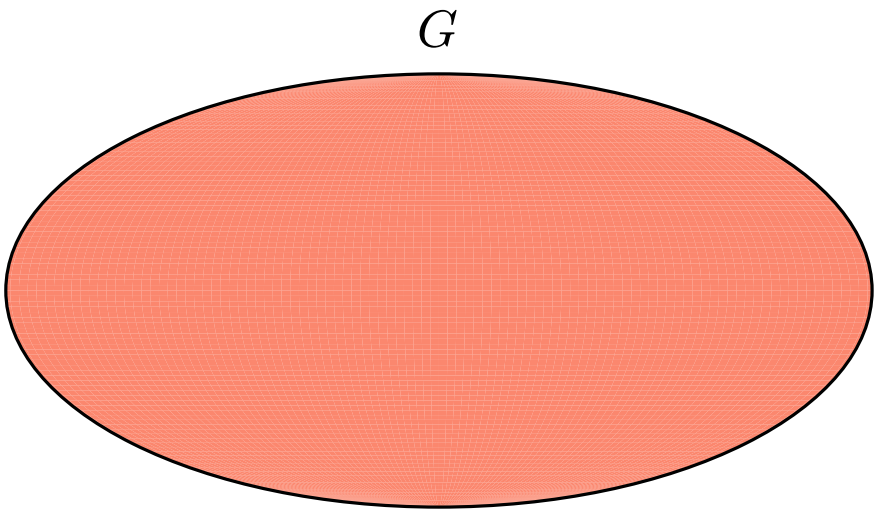

Θ_i
$\delta_B(X_i)$
risk : 0.211
X_i
risk : 0.288
$\delta_\pi(X_i)$
risk : 0.212

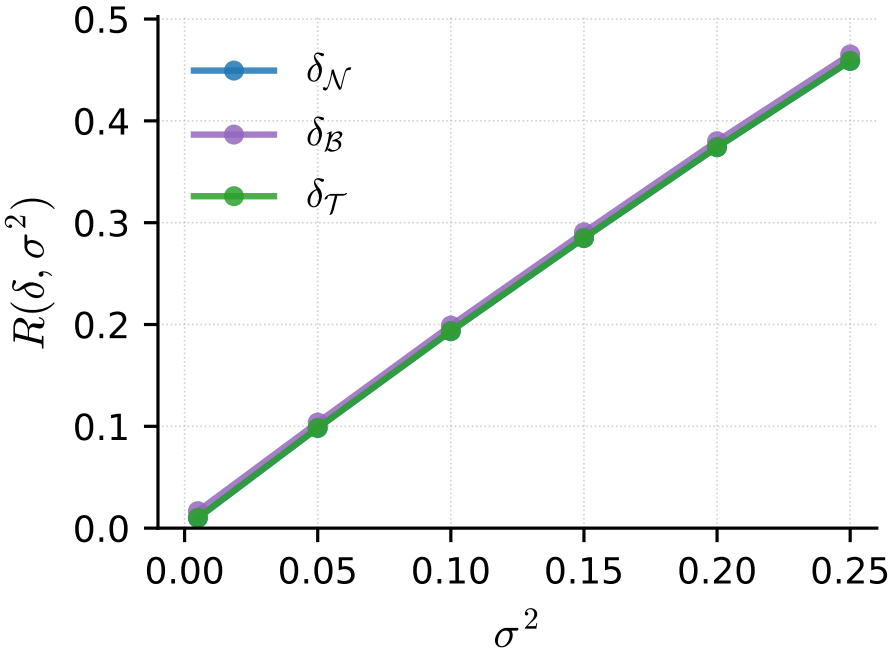

G


$\delta_{\mathcal{N}}$
$\delta_{\mathcal{B}}$
$\delta_{\mathcal{T}}$
$R(\delta, \sigma^2)$
0.0
0.1
0.2
0.3
0.4
0.00
0.05
0.10
0.15
0.20
0.25
σ^2


G


$\delta_{\mathcal{N}}$
$\delta_{\mathcal{B}}$
$\delta_{\mathcal{T}}$
$R(\delta, \sigma^2)$
0.0
0.1
0.2
0.3
0.4
0.00
0.05
0.10
0.15
0.20
0.25
σ^2


G


0.5
0.4
0.3
0.2
0.1
0.0
$R(\delta, \sigma^2)$
$\delta_{\mathcal{N}}$
$\delta_{\mathcal{B}}$
$\delta_{\mathcal{T}}$
0.00
0.05
0.10
0.15
0.20
0.25
σ^2

III. Large Sample Regime: $\hat{\delta}_T \approx \delta_T$

Oracle tangential Bayes denoiser:

$$\delta_{\mathcal{T}}(x) = \mathrm{Exp}_x\left(\sigma^2 \frac{\nabla_x f(x)}{f(x)}\right)$$

Plug-in approach? If $\hat{f}_n$ is a density estimate of f, then set:

$$\hat{\delta}_{\mathcal{T}}(x) := \mathrm{Exp}_x\left(\sigma^2 \frac{\nabla_x \hat{f}_n(x)}{\hat{f}_n(x)}\right)$$

Only need to estimate f and ∇f, no need to estimate G.

Efficiently computable (e.g., in `geomstats`), since it only relies on local geometry of $\mathcal{M}$ near x, not on global optimization over $\mathcal{M}$.

Density estimation via eigenfunctions $\{\phi_j\}_{j\geq 0}$ of Laplace-Beltrami operator Δ on $\mathcal{M}$ (Hendriks 1990):

$$\hat{f}_n(x) := \frac{1}{n}\sum_{i=1}^{n} K(x, X_i, \ell_n) \quad \text{where} \quad K(x, y, \ell) := \sum_{j=0}^{\ell} \phi_j(x)\overline{\phi_j(y)}.$$

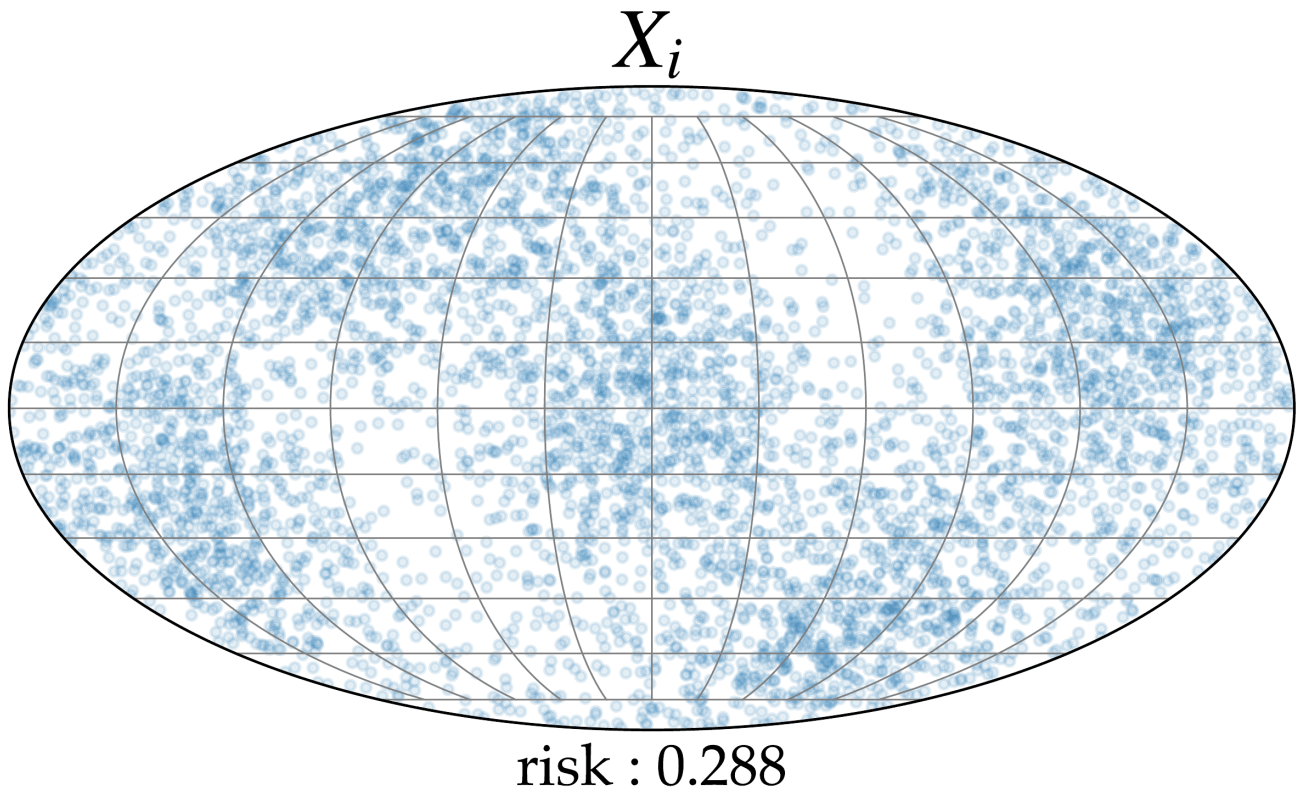
Example. For $\mathcal{M} = \mathbb{S}^1$, Fourier series

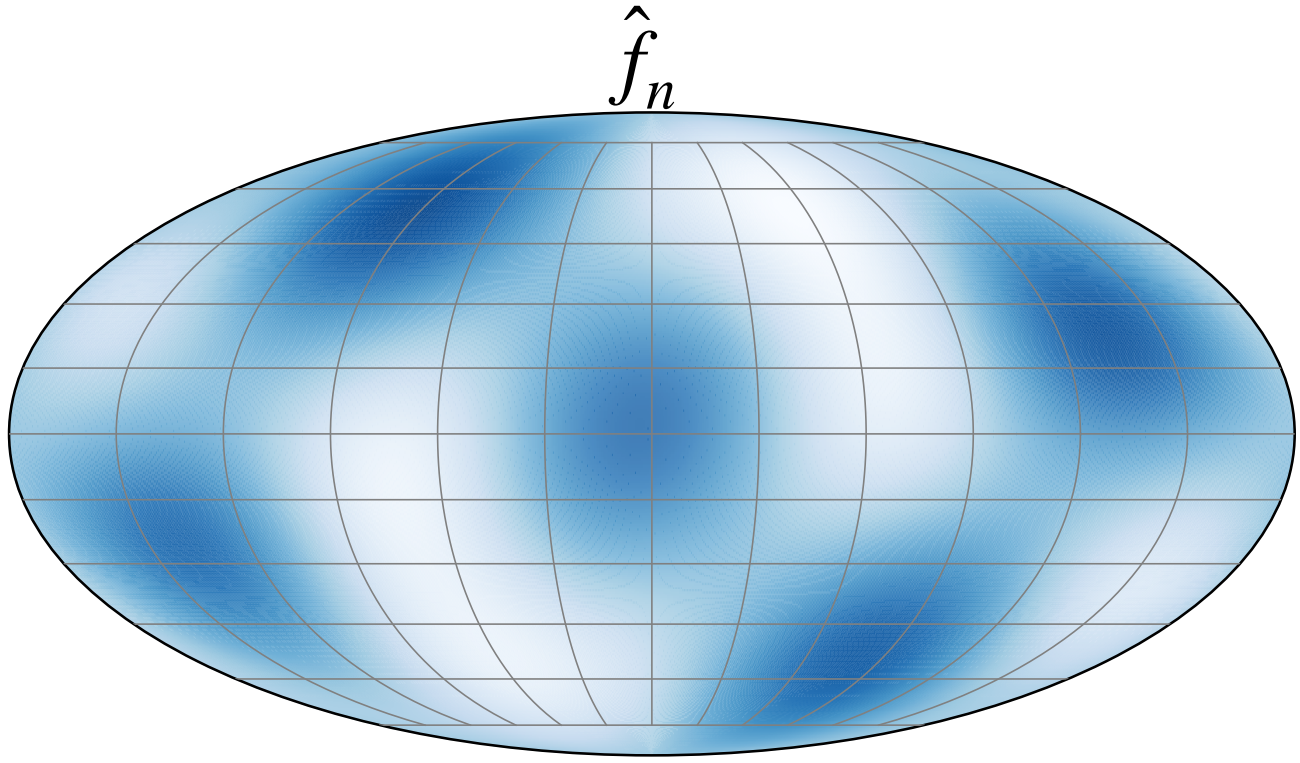
Example. For $\mathcal{M} = \mathbb{S}^2$, spherical harmonics

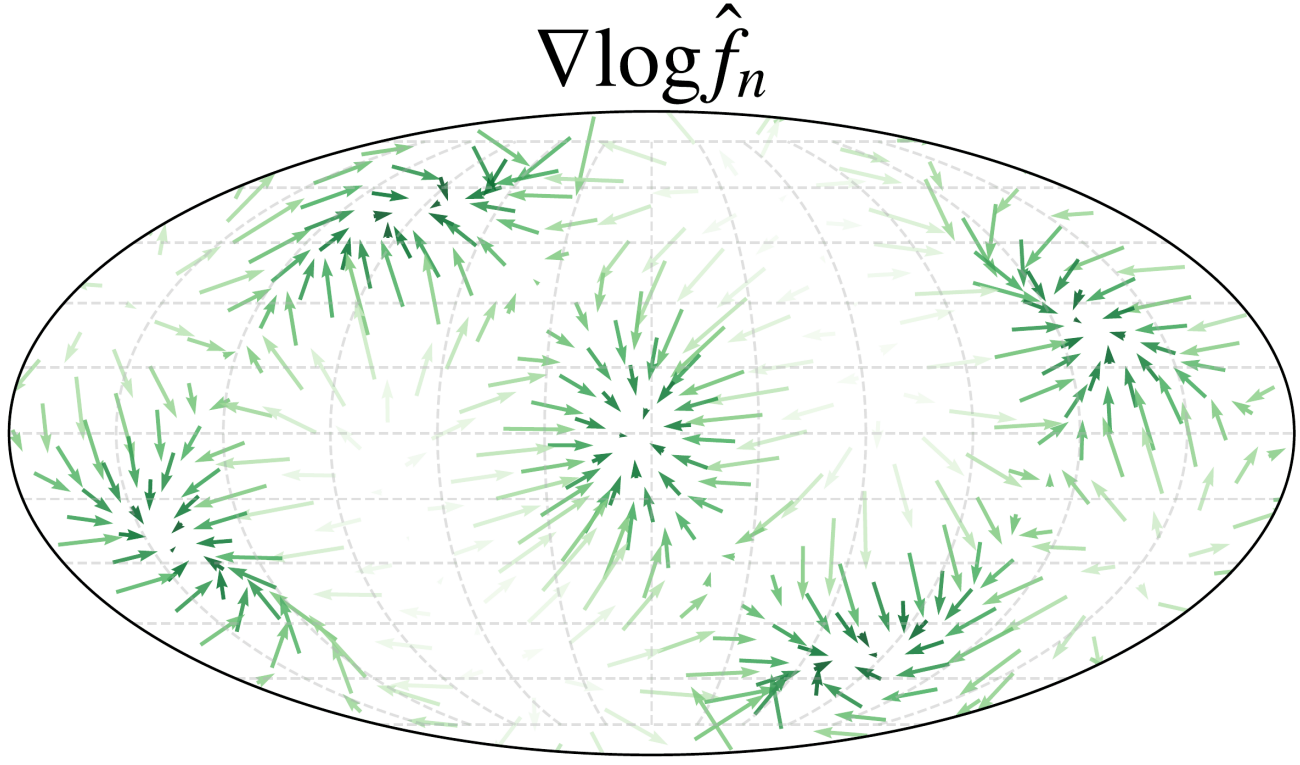
Empirical tangential Bayes denoiser:

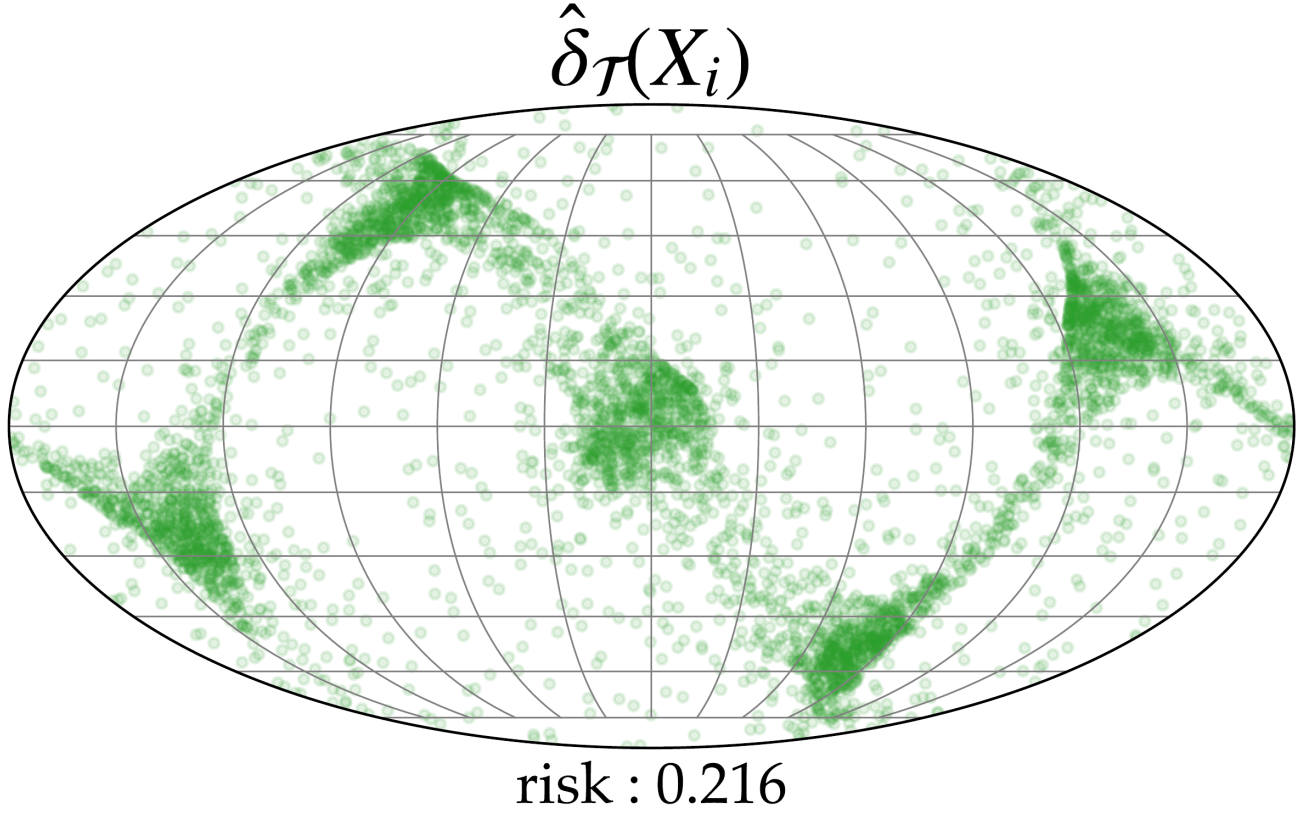
$$\hat{\delta}_{\mathcal{T}}(x) := \operatorname{Exp}_x\left(\sigma^2 \frac{\nabla_x \hat{f}_n(x)}{\max\{\hat{f}_n(x), \rho\}}\right)$$

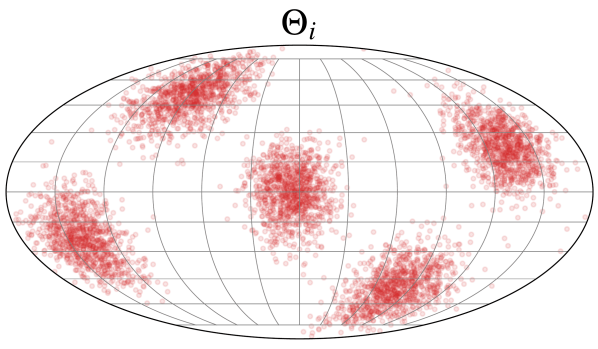
with tuning parameters $\ell_n \geq 0$ and $\rho > 0$.

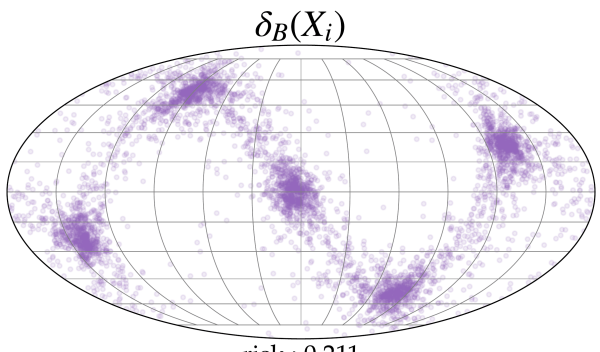

X_i
risk : 0.288


$\hat{f}_n$

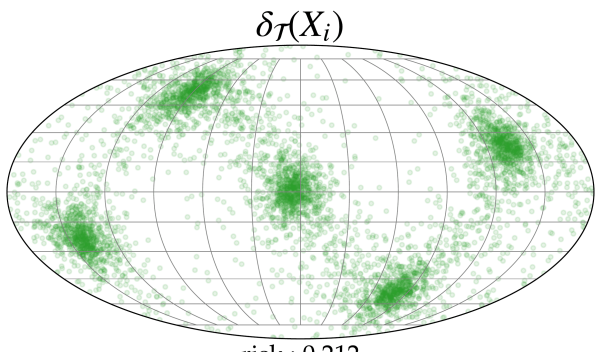

$\nabla \log \hat{f}_n$


$\hat{\delta}_{\mathcal{T}}(X_i)$
risk : 0.216

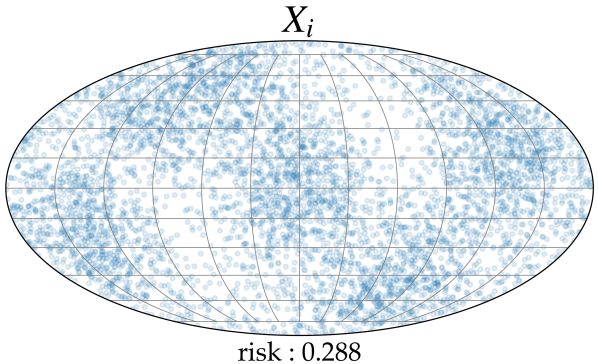




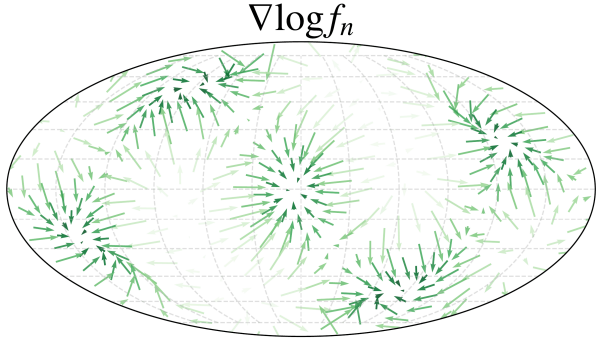
risk : 0.211

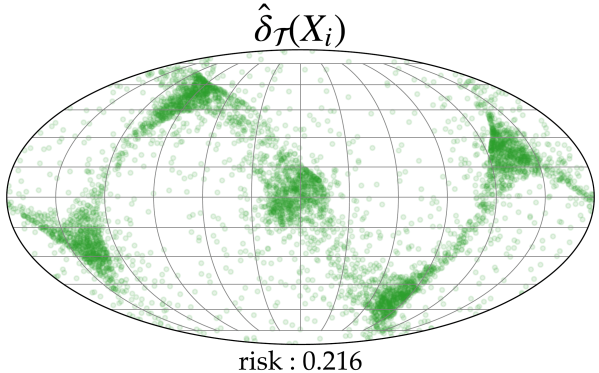


risk : 0.212



risk : 0.288





risk : 0.216

Recall that G is distribution of Θ and $f = f_G$ is density of X.

Theorem (AQJ-Santoro-Sen, 2026)

Suppose G satisfies $\|f\|_{\mathcal{H}^r(\mathcal{M})} < \infty$ for some $r > 1$, where $\mathcal{H}^r(\mathcal{M})$ is Sobolev space of order r. Then, for suitable ℓ_n and ρ, we have

$$\mathbb{E}\left[d^2\left(\hat{\delta}_{\mathcal{T}}(X), \delta_{\mathcal{T}}(X)\right)\right] \lesssim n^{-\frac{2r-2}{2r+m}}.$$

As $r \to \infty$ this approaches the parametric rate n^{-1}.

But, how does the smoothness G impact the smoothness of $f = f_G$?

In Euclidean setting, f is infinitely differentiable for every G, leading to the parametric rate of convergence of $\frac{1}{n}$ up to logarithmic factors.

In Riemannian setting, smoothness of f depends delicately on G. This all because $x \mapsto d^2(x, \theta)$ is not smooth on C_θ.

Sobolev regularity of Riemannian Gaussian mixture models:

- On $\mathbb{S}^1$: $\|f\|_{\mathcal{H}^{s+2}(\mathbb{S}^1)} \asymp \|g\|_{\mathcal{H}^s(\mathbb{S}^1)}$
- On $\mathbb{S}^2$: $\|f\|_{\mathcal{H}^2(\mathbb{S}^2)} \lesssim \|g\|_{L^\infty(\mathbb{S}^2)}$
- On Lie group $\mathcal{M}$: $\|f\|_{\mathcal{H}^r(\mathcal{M})} \lesssim \|g\|_{\mathcal{H}^r(\mathcal{M})}$
- In general: $\|f\|_{\mathcal{H}^1(\mathcal{M})} \lesssim \|g\|_{L^\infty(\mathcal{M})}$

Lower bound for $\mathbb{S}^1$ with $g \in \mathcal{H}^s(\mathbb{S}^1)$ gives minimax rate $n^{-\frac{2s+2}{2s+5}}$

Qualitatively similar behavior to "smeariness" (Hotz-Huckemann 2015)

Choosing tuning parameters in practice? Riemannian version of score matching (Hyvärinen 2005, De Bortoli et al. 2022).

For smooth positive density q on $\mathcal{M}$, integration by parts on $\mathcal{M}$ gives

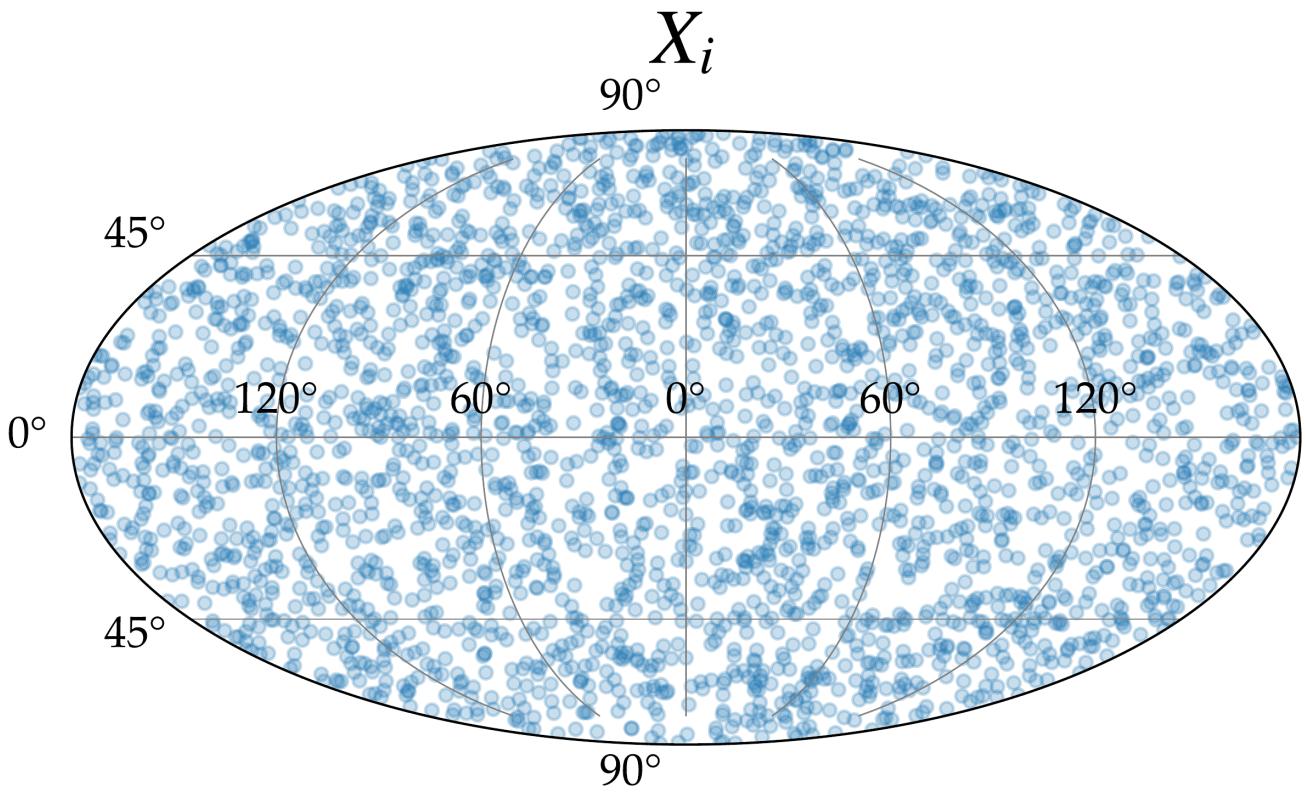
$$\mathbb{E}\left[\tfrac{1}{2}\left\|\nabla \log q(X) - \nabla \log f(X)\right\|_X^2\right]$$

$$= \mathbb{E}\left[\tfrac{1}{2}\left\|\nabla \log q(X)\right\|_X^2 + \Delta \log q(X)\right] + C_f.$$

Right side can be estimated from $X_1, \ldots, X_n$ by cross-validation.

IV. Applications

Astronomy

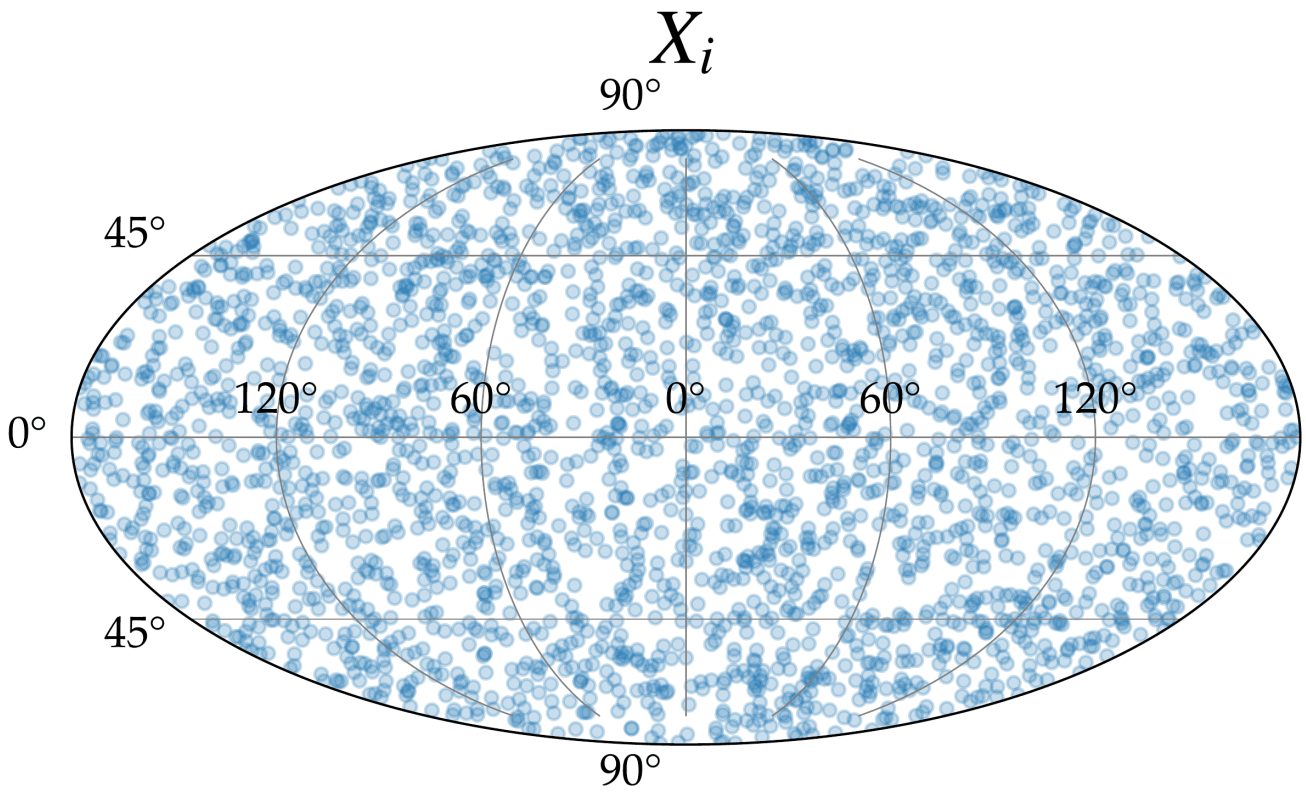

X_i
90°
45°
120°
60°
0°
60°
120°
0°
45°
90°

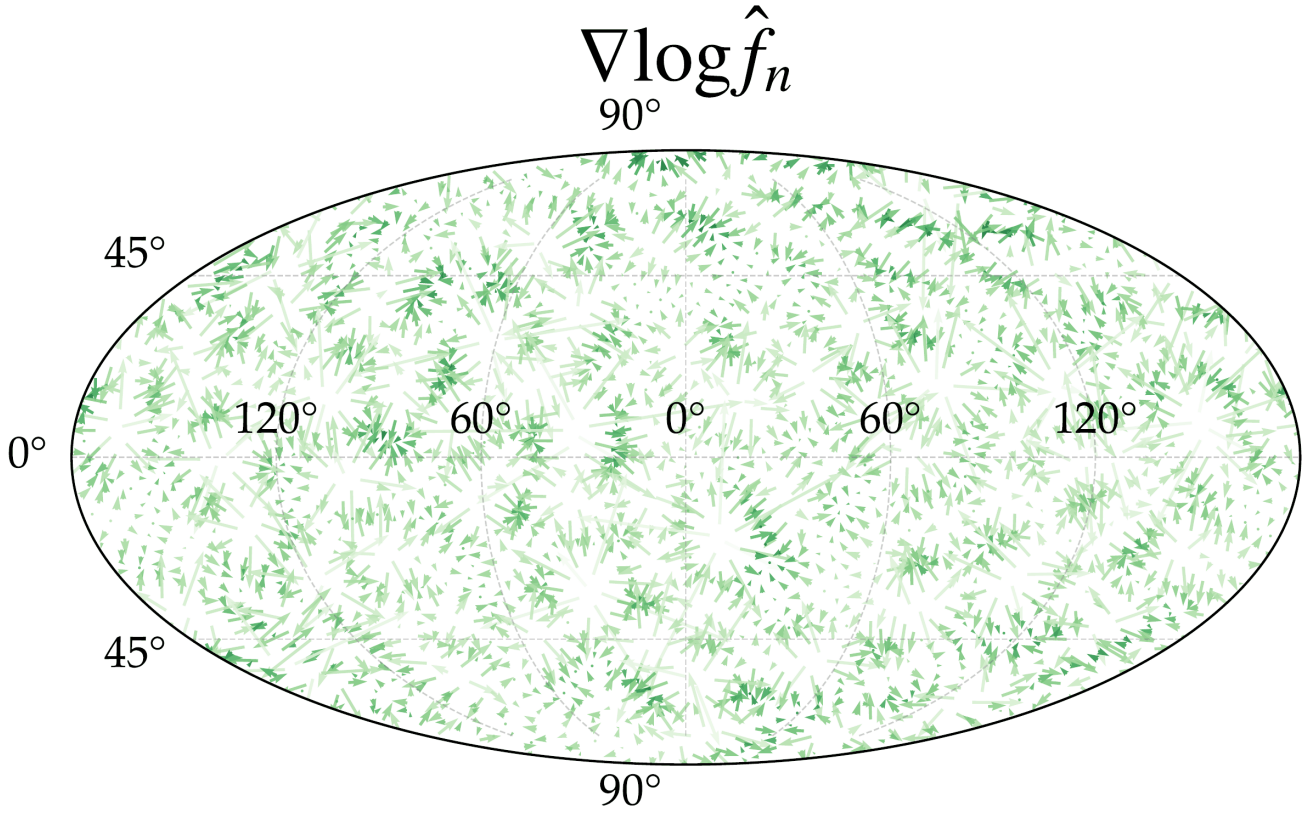
Interested in the directions of $n = 2702$ gamma ray bursts on the celestial sphere $\mathbb{S}^2$, from the BATSE 4B catalog (Paciesas et al. 1999).

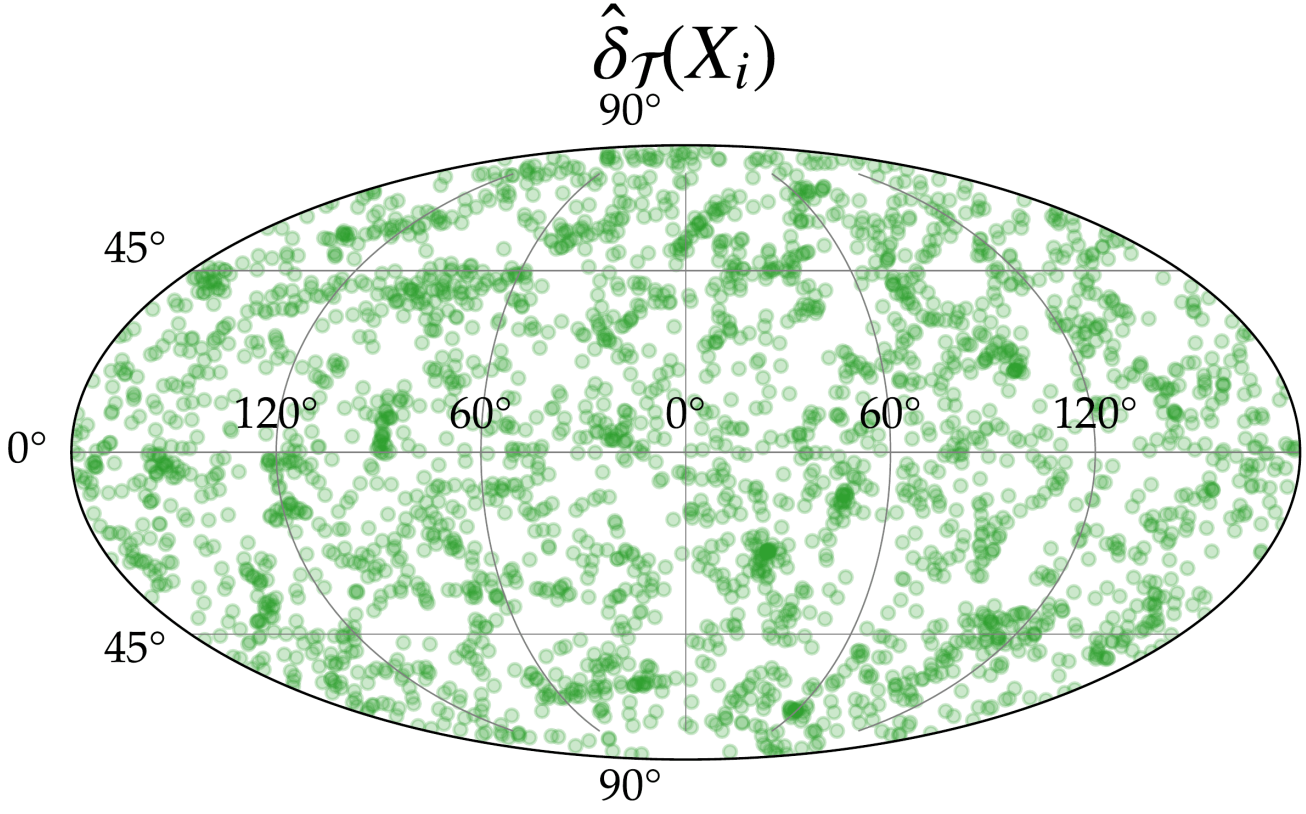
Latent directions $\Theta_1, \ldots, \Theta_n \in \mathbb{S}^2$ are i.i.d. from unknown G.

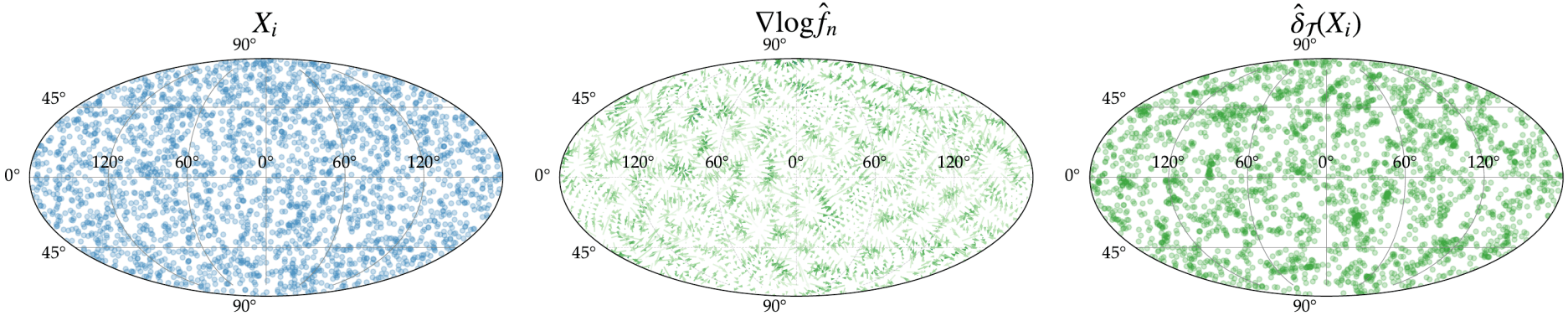
Measurements $X_1, \ldots, X_n \in \mathbb{S}^2$ have error, we take $\sigma^2 = 3.05 \times 10^{-2}$ from reported calibration error.

Whether or not G is uniform is related to the origins of gamma ray bursts, which is unknown (Meegan et al. 1992, Piran 2004).

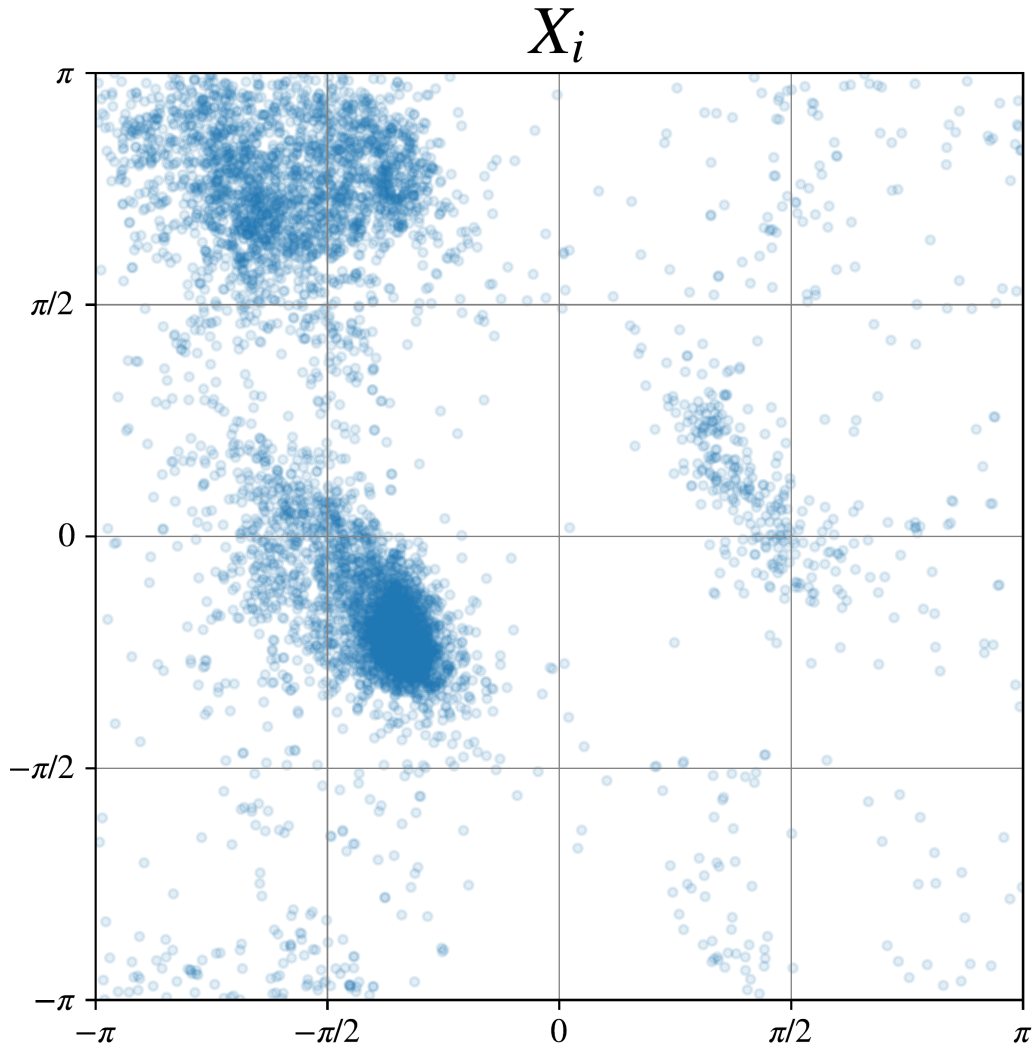

X_i
90°
45°
120°
60°
0°
60°
120°
0°
45°
90°


$\nabla \log \hat{f}_n$
90°
45°
0°
120°
60°
0°
60°
120°
45°
90°


$\hat{\delta}_{\mathcal{T}}(X_i)$
90°
45°
0°
45°
90°
120°
60°
0°
60°
120°


X_i
$\nabla \log \hat{f}_n$
$\hat{\delta}_{\mathcal{T}}(X_i)$
90°
45°
0°
120°
60°
0°
60°
120°

Structural biology

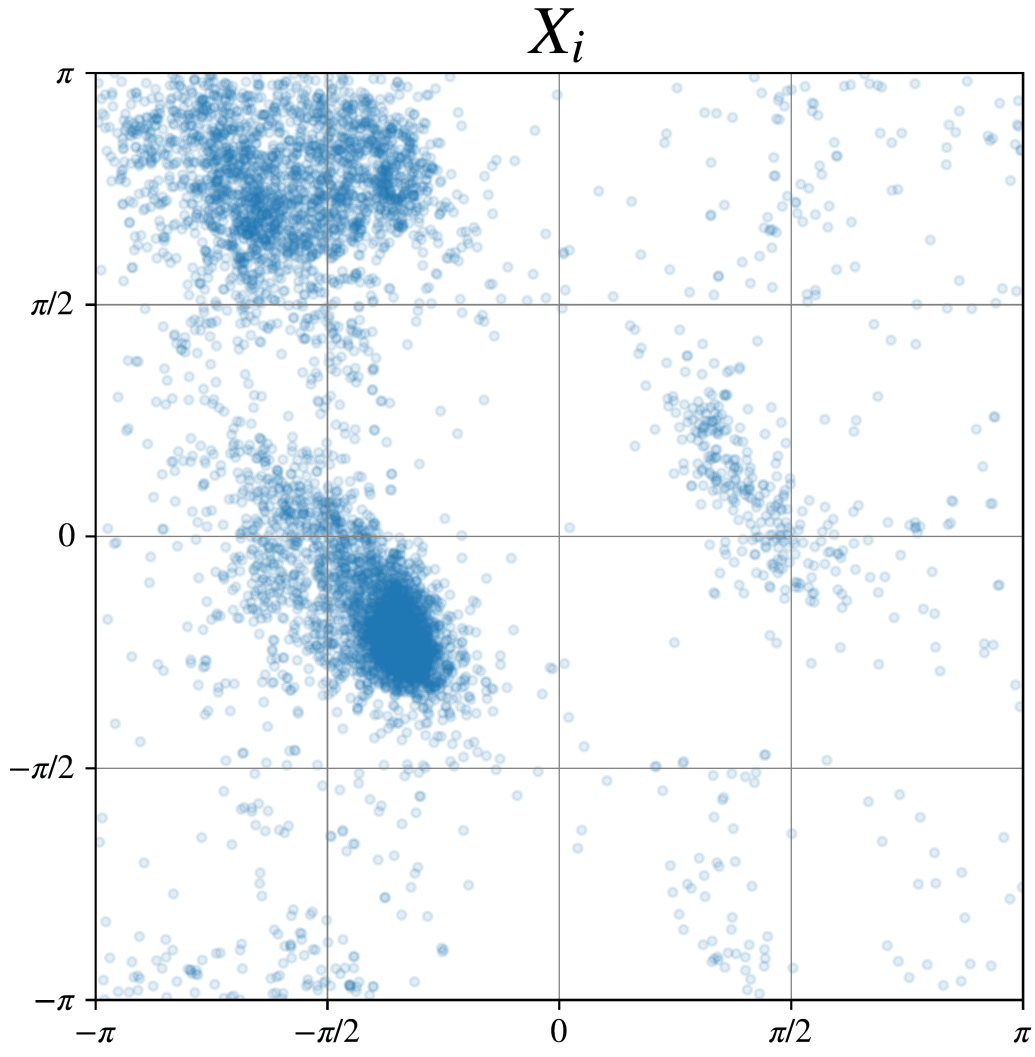

X_i
π
$\pi/2$
0
$-\pi/2$
$-\pi$
$-\pi$
$-\pi/2$
0
$\pi/2$
π

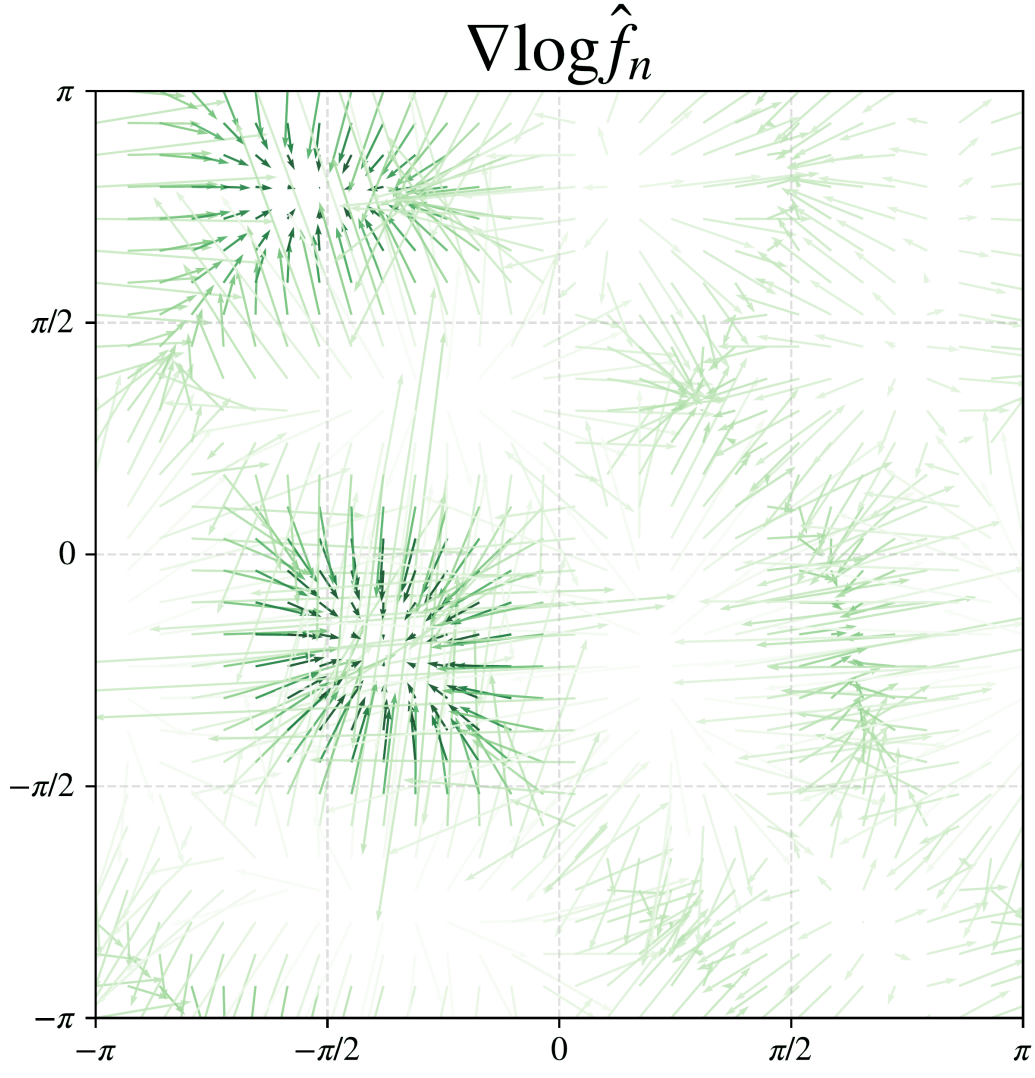
Interested in shape of proteins: for each pair adjacent amino acids, compute the pair of torsion angles relative to the backbone $(\phi, \psi) \in \mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ (Ramachandran et al. 1963).

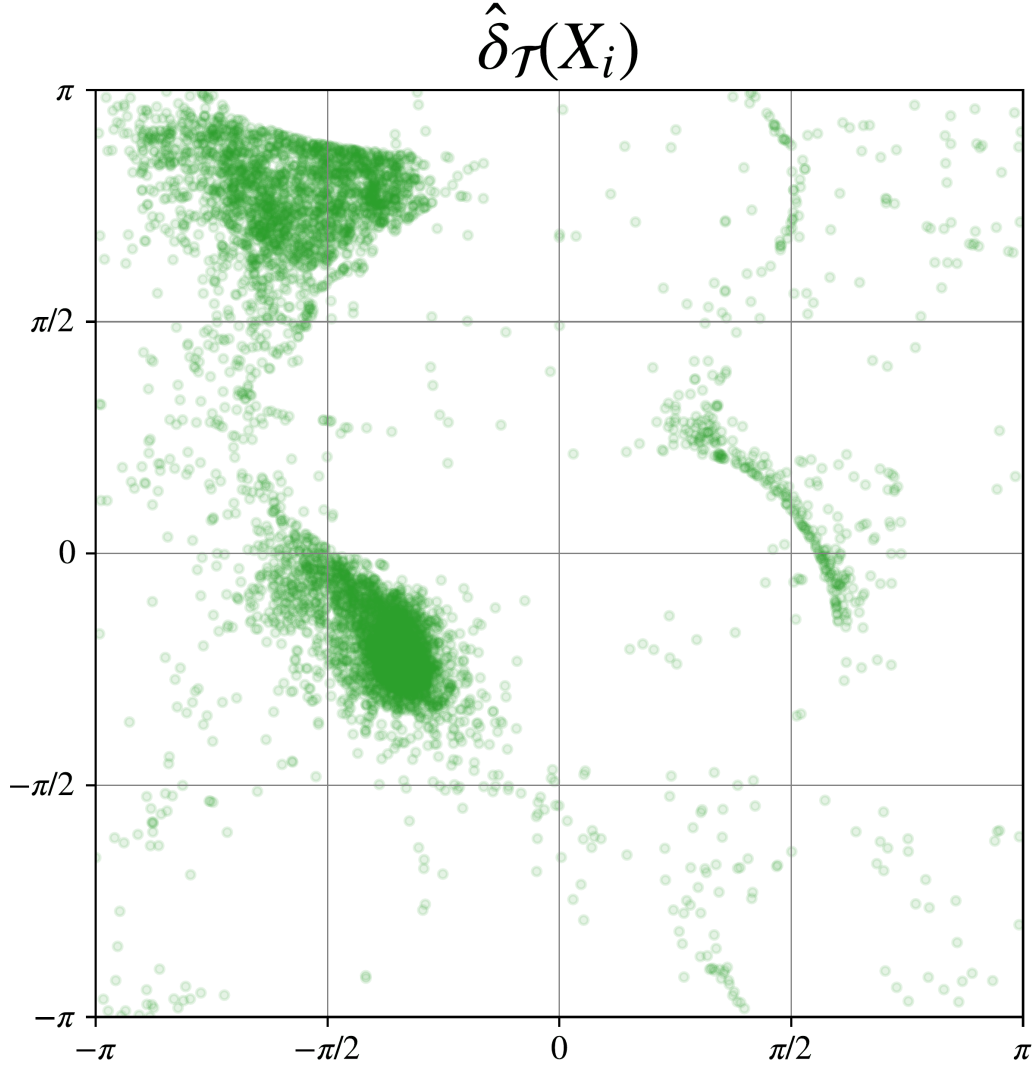
Set of 10 common glycolytic enzymes in the Protein Data Bank (PDB) lead to $n = 7181$ angle pairs.

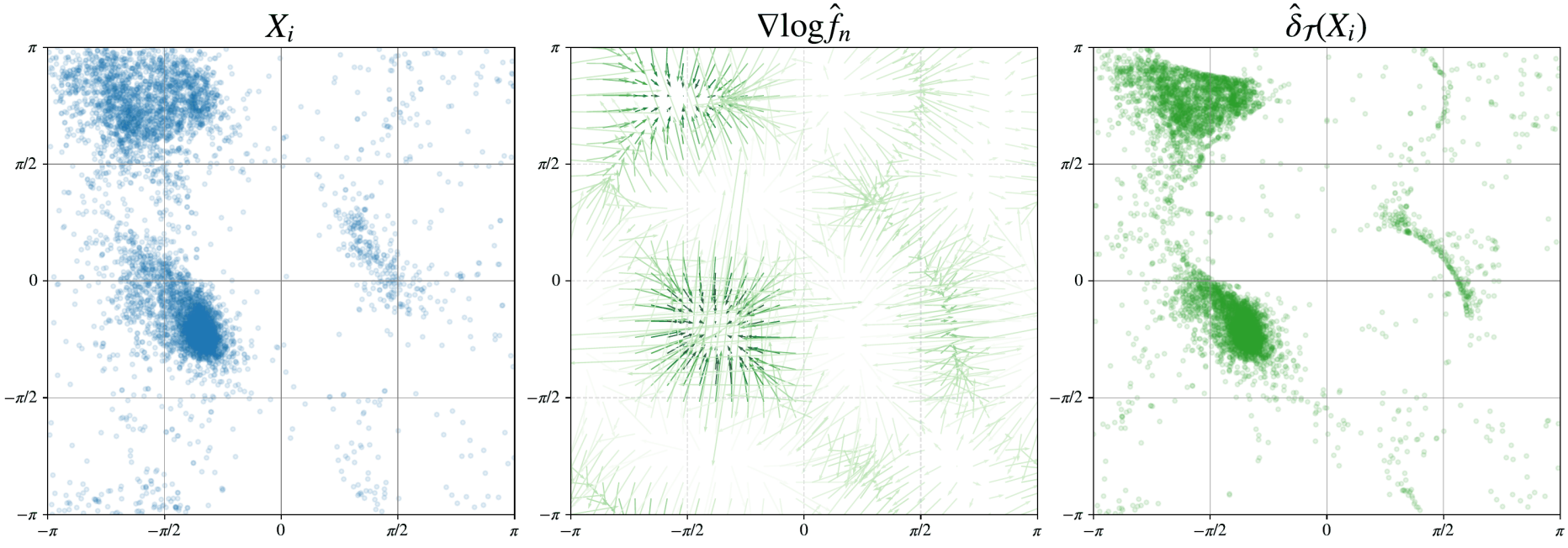
Latent angle pairs $\Theta_1, \ldots, \Theta_n \in \mathbb{T}^2$ are i.i.d. from unknown G.

Measurements $X_1, \ldots, X_n \in \mathbb{T}^2$ have known noise level $\sigma^2 \approx 0.135$


X_i
π
$\pi/2$
0
$-\pi/2$
$-\pi$
$-\pi$
$-\pi/2$
0
$\pi/2$
π


$\nabla \log \hat{f}_n$
π
$\pi/2$
0
$-\pi/2$
$-\pi$
$-\pi$
$-\pi/2$
0
$\pi/2$
π


$\hat{\delta}_{\mathcal{T}}(X_i)$
π
$\pi/2$
0
$-\pi/2$
$-\pi$
$-\pi$
$-\pi/2$
0
$\pi/2$
π


X_i
$\nabla \log \hat{f}_n$
$\hat{\delta}_{\mathcal{T}}(X_i)$
π
$\pi/2$
0
$-\pi/2$
$-\pi$

Cryo-electron microscopy

Preliminary work in single-particle cryo-EM (Singer-Sigworth 2020):

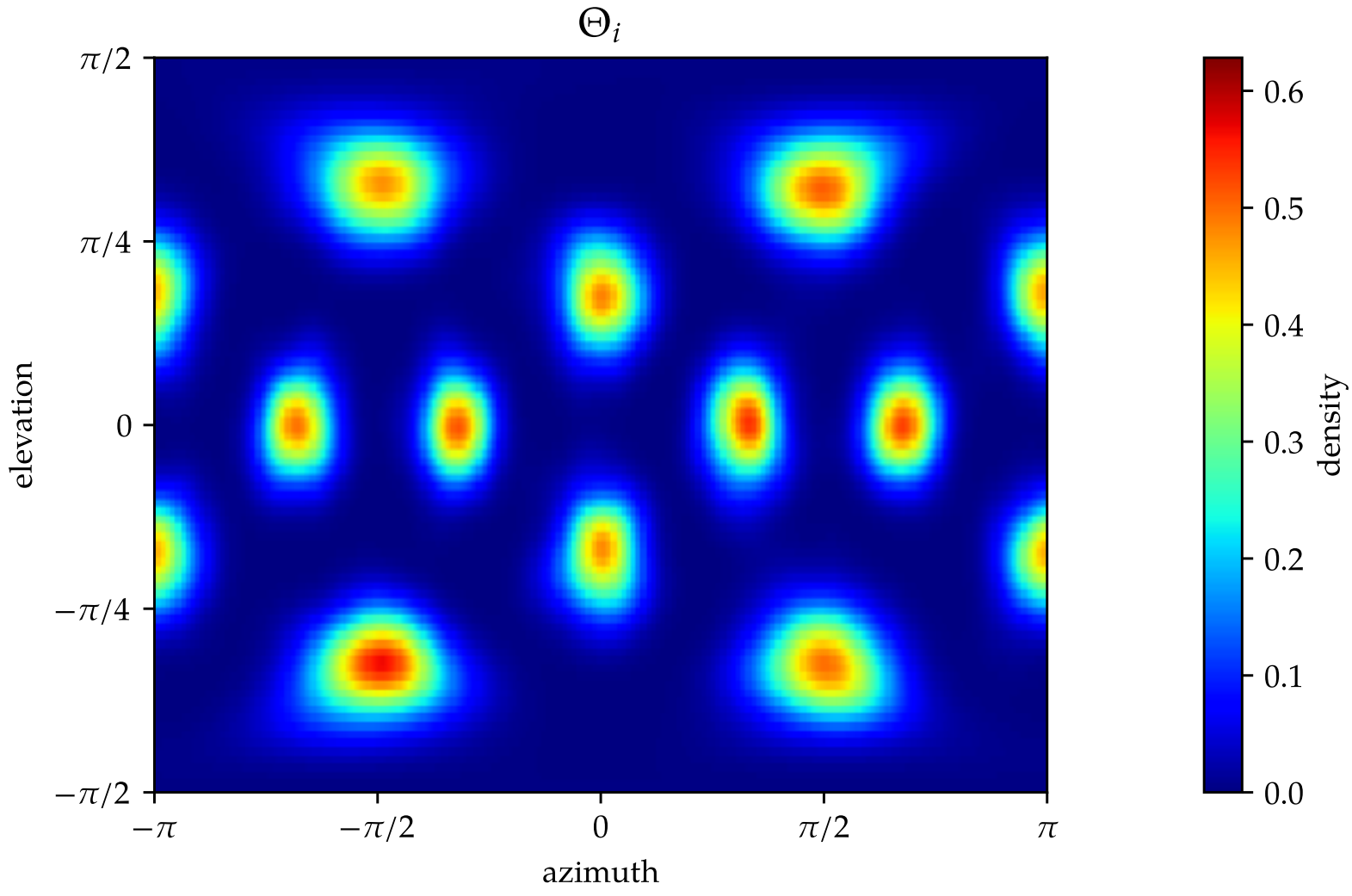
Imaging process for batch of similar molecules, obtaining one image for each: rotate, freeze, image (hence destroying the molecule)

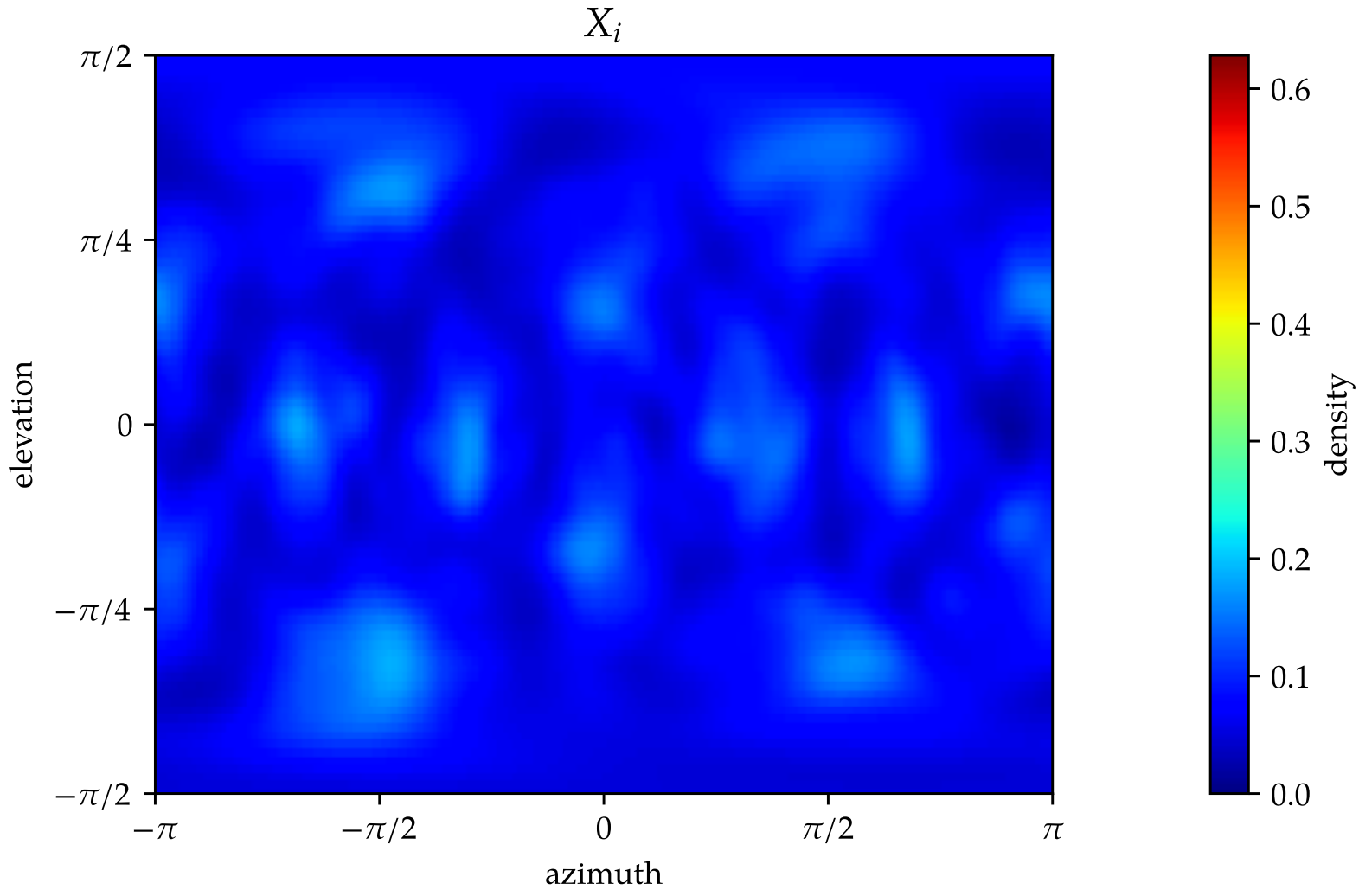
Latent rotations $\Theta_1, \ldots, \Theta_n \in \mathrm{SO}(3)$ are i.i.d. from unknown G which reflects molecules' physical symmetries (Harpaz-Shkolnisky 2025).

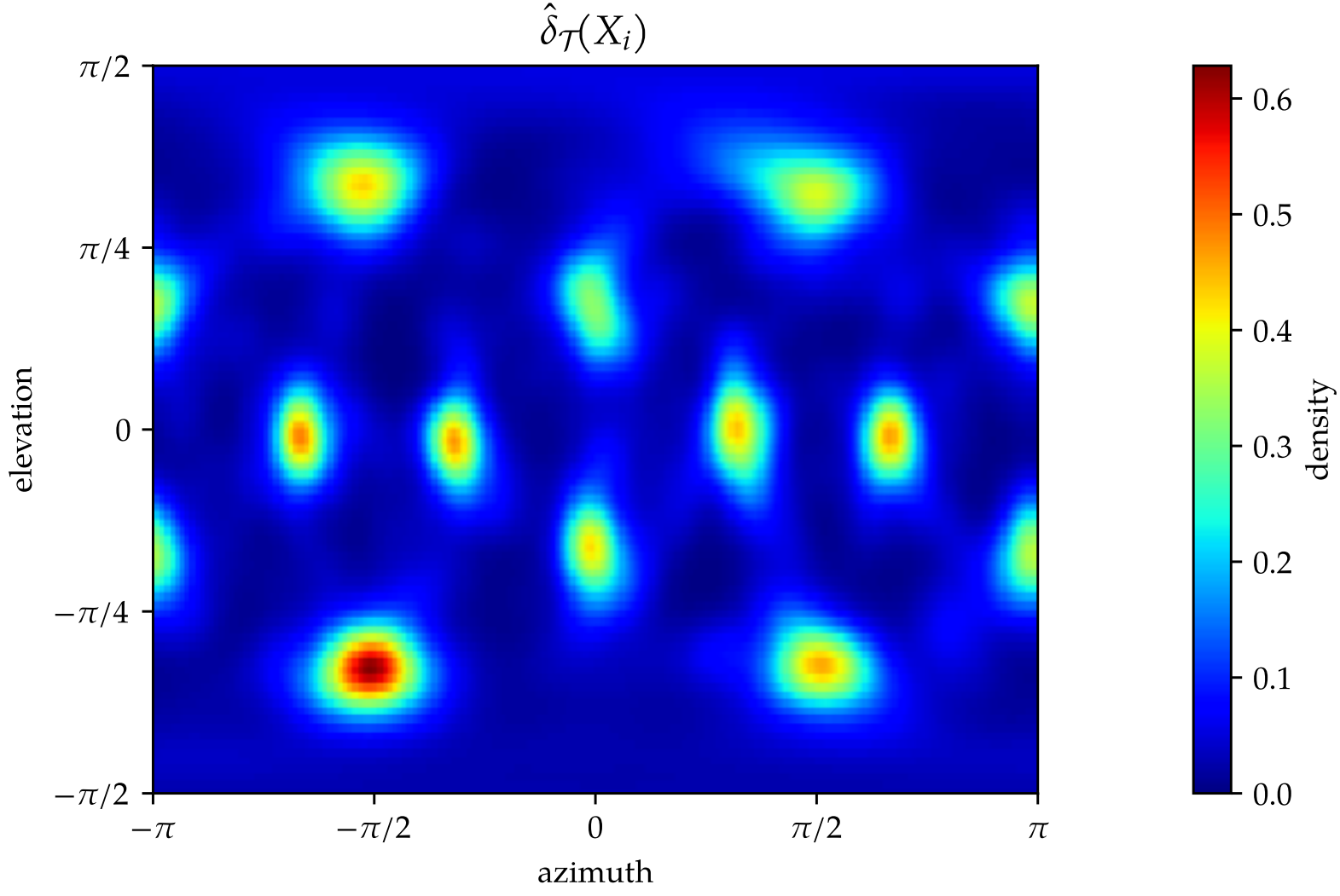
Alignment procedures produce estimates $X_1, \ldots, X_n \in \mathrm{SO}(3)$.

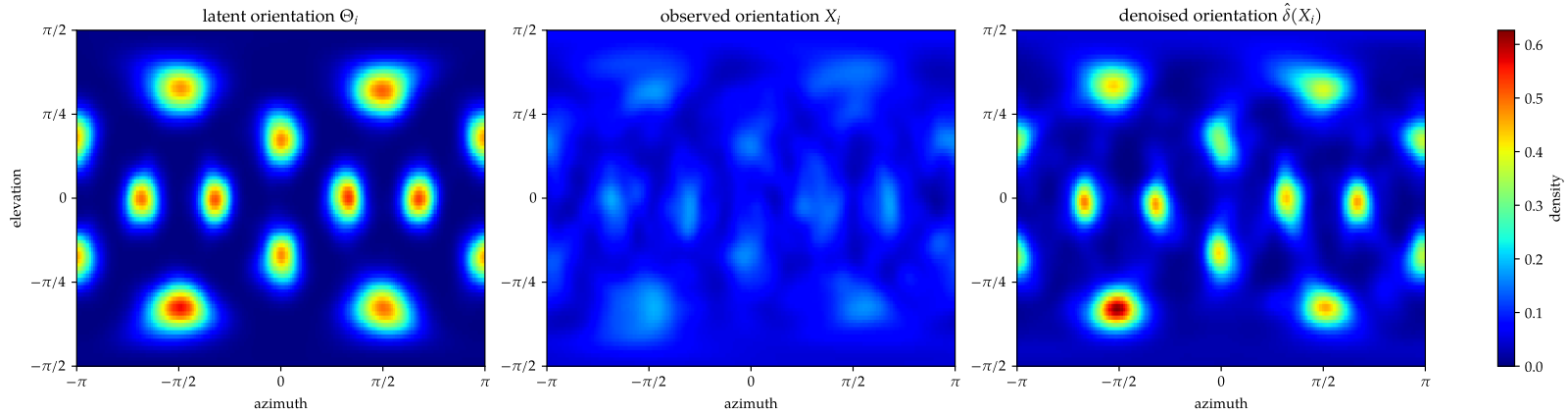
Recent interest in Bayesian approaches, which design priors to reflect expected symmetries of the unknown molecules (Xu et al. 2026).

With empirical Bayes, learn symmetries from the data!


Θ_i
$\pi/2$
$\pi/4$
0
$-\pi/4$
$-\pi/2$
elevation
$-\pi$
$-\pi/2$
0
$\pi/2$
π
azimuth
0.6
0.5
0.4
0.3
0.2
0.1
0.0
density


X_i
$\pi/2$
$\pi/4$
0
$-\pi/4$
$-\pi/2$
elevation
$-\pi$
$-\pi/2$
0
$\pi/2$
π
azimuth
0.6
0.5
0.4
0.3
0.2
0.1
0.0
density


$\hat{\delta}_{\mathcal{T}}(X_i)$
$\pi/2$
$\pi/4$
0
$-\pi/4$
$-\pi/2$
elevation
$-\pi$
$-\pi/2$
0
$\pi/2$
π
azimuth
0.6
0.5
0.4
0.3
0.2
0.1
0.0
density


latent orientation Θ_i
observed orientation X_i
denoised orientation $\hat{\delta}(X_i)$
elevation
azimuth
density
$\pi/2$
$\pi/4$
0
$-\pi/4$
$-\pi/2$
$-\pi$
π
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Challenges:

- Sobolev regularity of Riemannian Gaussian mixtures on SO(3)?
- Hyperparameter tuning for 3-dimensional density estimation
- In practice, measurement noise is very large!
- Is the likelihood approximately Riemannian Gaussian in practice?

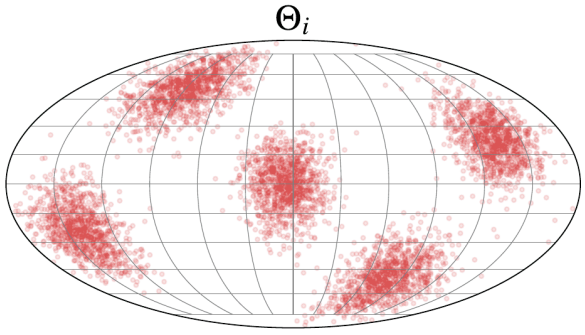
Conclusion

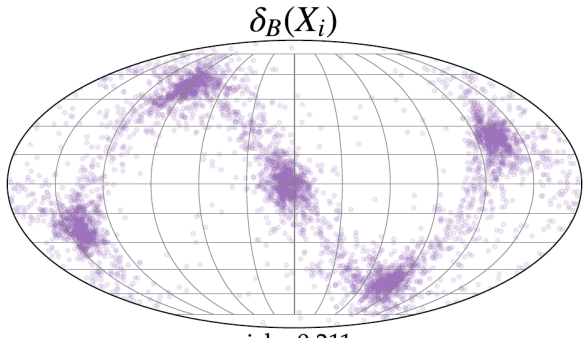
Summary:

- Many statistical problems involve denoising manifold-valued data
- Riemannian Tweedie-Eddington $\rightarrow$ intermediate oracle denoiser $\delta_{\mathcal{T}}$
- Theory shows $\hat{\delta}_{\mathcal{T}} \approx \delta_{\mathcal{T}} \approx \delta_{\mathcal{B}}$ as $\sigma^2 \rightarrow 0$ and $n \rightarrow \infty$
- Sobolev regularity results for Riemannian Gaussian mixtures
- Preliminary work towards application in cryo-EM

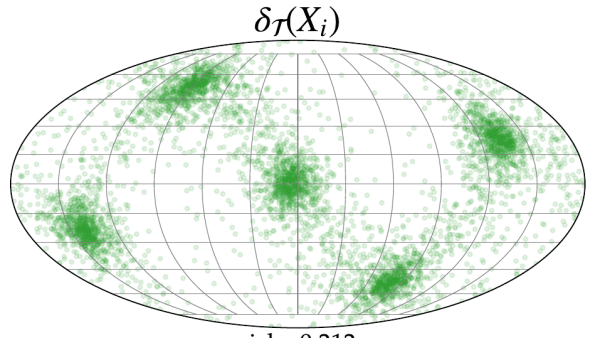
Future work:

- Other likelihoods (e.g., heat kernel, von Mises)?
- G-modeling (e.g., NPMLE)? (Jiang-Zhang 2009, Efron 2014)
- Sharp Sobolev estimates for Riemannian Gaussian mixtures?
- Minimax lower bounds beyond the circle?
- Unknown manifolds? (García Trillos-Li-Venkatraman 2026)

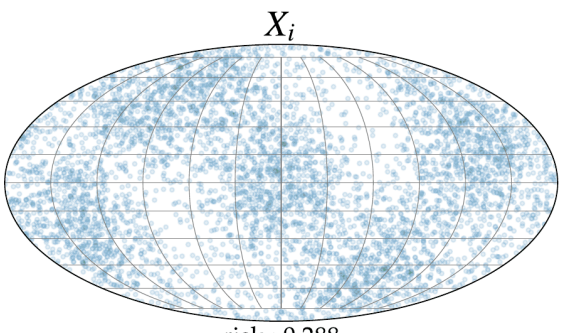




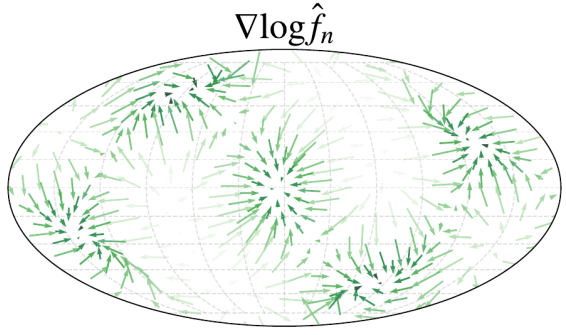
risk : 0.211

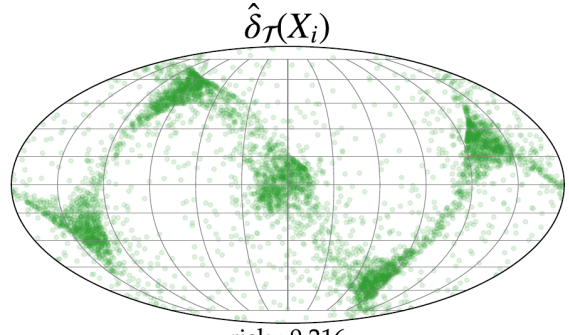


risk : 0.212



risk : 0.288





risk : 0.216

Thank you!

References

N. Boumal. *An Introduction to Optimization on Smooth Manifolds*. Cambridge University Press, 2023.

V. De Bortoli, E. Mathieu, M. Hutchinson, J. Thornton, Y. W. Teh, and A. Doucet. Riemannian score-based generative modelling. *NeurIPS*, 35, 2022.

F. W. Dyson. A method for correcting series of parallax observations. *Mon. Not. R. Astron. Soc.,* 86(9):686-706, 1926.

A. S. Eddington. The correction of statistics for accidental error. *Mon. Not. R. Astron. Soc.,* 100:354-361, 1940.

B. Efron. Tweedie's formula and selection bias. *J. Am. Stat. Assoc.,* 106(496):1602-1614, 2011.

B. Efron. Two modeling strategies for empirical Bayes estimation. *Statist. Sci.,* 29(2):285-301, 2014.

N. García Trillos, C. Li, and R. Venkatraman. Minimax rates for the estimation of eigenpairs of weighted Laplace-Beltrami operators on manifolds. *IEEE Trans. Inf. Theory*, 2026.

Y. Harpaz and Y. Shkolnisky. Three-dimensional alignment of density maps in cryo-electron microscopy. *Biol. Imaging*, 3:e8, 2025.

D. M. Healy Jr. and P. T. Kim. An empirical Bayes approach to directional data and efficient computation on the sphere. *Ann. Statist.*, 24(1):232-254, 1996.

H. Hendriks. Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *Ann. Statist.,* 18(2):832-849, 1990.

T. Hotz and S. Huckemann. Intrinsic means on the circle: uniqueness, locus and asymptotics. *Ann. Inst. Statist. Math.,* 67(1):177-193, 2015.

S. F. Huckemann, P. T. Kim, J.-Y. Koo, and A. Munk. Möbius deconvolution on the hyperbolic plane with application to impedance density estimation. *Ann. Statist.*, 38(4):2465-2498, 2010.

A. Hyvärinen. Estimation of non-normalized statistical models by score matching. *J. Mach. Learn. Res.,* 6:695-709, 2005.

W. Jiang and C.-H. Zhang. General maximum likelihood empirical Bayes estimation of normal means. *Ann. Statist.*, 37(4):1647-1684, 2009.

P. T. Kim. Deconvolution density estimation on SO(N). *Ann. Statist.*, 26(3):1083-1102, 1998.

References

A. McCormack and P. Hoff. The Stein effect for Fréchet means. *Ann. Statist.,* 50(6): 3647-3676, 2022.

C. A. Meegan, G. J. Fishman, R. B. Wilson, et al. Spatial distribution of γ-ray bursts observed by BATSE. *Nature,* 355:143-145, 1992.

W. S. Paciesas, C. A. Meegan, G. N. Pendleton, et al. The Fourth BATSE Gamma-Ray Burst Catalog (Revised). *Astrophys. J. Suppl. Ser.,* 122(2):465-495, 1999.

B. Pelletier. Kernel density estimation on Riemannian manifolds. *Statist. Probab. Lett.*, 73(3):297-304, 2005.

X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Math. Imaging Vis.,* 25(1):127-154, 2006.

T. Piran. The physics of gamma-ray bursts. *Rev. Mod. Phys.,* 76(4):1143-1210, 2004.

G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.,* 7(1):95-99, 1963.

H. Robbins. An empirical Bayes approach to statistics. *Berkeley Symp. on Math. Statist. and Prob.,* 3(1):157-163, 1956.

S. Saha and A. Guntuboyina. On the nonparametric maximum likelihood estimator for Gaussian location mixture densities with application to Gaussian denoising. *Ann. Statist.*, 48(2):738-762, 2020.

A. Singer and F. J. Sigworth. Computational methods for single-particle electron cryomicroscopy. *Annu. Rev. Biomed. Data Sci.,* 3:163-190, 2020.

J. A. Soloff, A. Guntuboyina, and B. Sen. Multivariate, heteroscedastic empirical Bayes via nonparametric maximum likelihood. *J. R. Stat. Soc. Ser. B*, 87(1):1-32, 2025.

S. Xu, A. Balanov, A. Singer, and T. Bendory. Bayesian perspective for orientation determination in cryo-EM with application to structural heterogeneity analysis. *Acta Cryst. D*, 82(4):274-294, 2026.

C.-H. Yang, H. Doss, and B. C. Vemuri. An empirical Bayes approach to shrinkage estimation on the manifold of symmetric positive-definite matrices. *J. Amer. Statist. Assoc.*, 119(545):259-272, 2024.

C.-H. Zhang. Empirical Bayes and compound estimation of normal means. *Statist. Sinica,* 7(1):181-193, 1997.