Positive Curvature and Hamiltonian Monte Carlo

Christof Seiler, Simon Rubinstein-Salzedo, and Susan Holmes
Department of Statistics, Stanford University

Goal

Running time estimates for sampling from high dimensional probability distributions using Hamiltonian Monte Carlo (HMC).

Hamiltonian Monte Carlo

Bayesian posterior distribution:

$$\pi(q \mid \text{data}) = \frac{1}{Z}\mathcal{L}(\text{data} \mid q)\,\pi(q)$$

parameter $q \in \mathbb{R}^d$, $d \gg$ observations, and unknown normalizing constant Z.

Expectation of $f : \mathbb{R}^d \to \mathbb{R}$

$$I := \mathbb{E}_\pi(f) = \int f(q)\,\pi(q \mid \text{data})\,dq$$

Empirical mean:

$$\hat{I} := \frac{1}{T}\sum_{t=1}^{T} Q_t, \qquad Q_1, \dots, Q_T \sim \pi(q \mid \text{data})$$

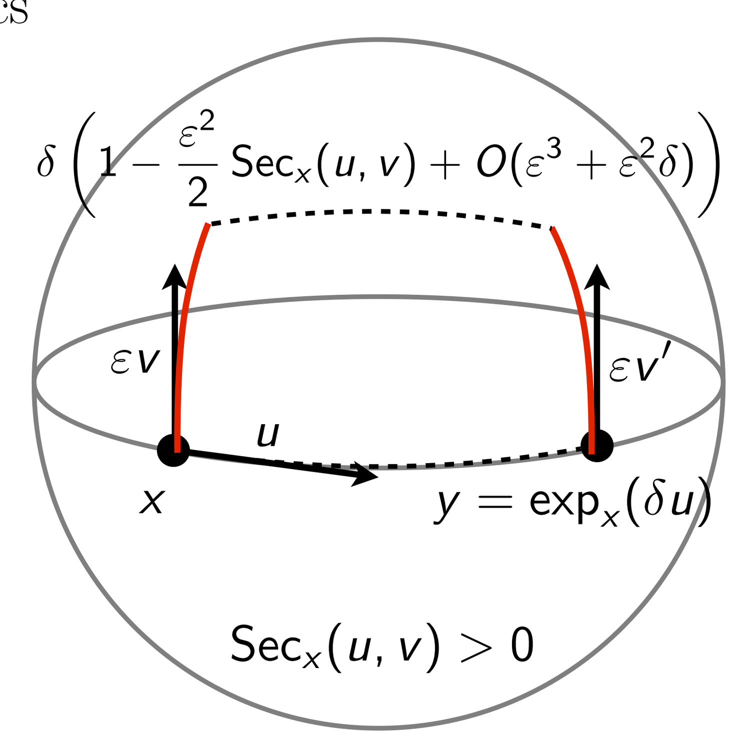
simulated using HMC for running time T

HMC algorithm: Metropolis-Hastings with proposal step following Hamiltonian trajectories

Error and running time: Bound the error $|\hat{I} - I|$ as a function of the running time T

Sectional Curvature

Sectional curvature **Sec** describes the nature of geodesics



Positive curvature: geodesics meet

Flat and negative curvature: geodesics deviate

Probability Meets Geometry

Coarse Ricci curvature $\boldsymbol{\kappa}$ describes the deviation from independence of subsequent draws from Markov chain Monte Carlo

- The larger κ the more independent
- The larger κ the smaller T for the same error
- We calculate the κ of HMC

Concentration inequality (JO10)

$$\mathbb{P}\left(|\hat{I} - I| \geq r\|f\|_{\text{Lip}}\right) \leq 2e^{-r^2/(16V^2(\kappa,T))}$$

Jacobi Metric

Jacobi metric $\boldsymbol{g_h}$ links Hamiltonians to Riemannian geometry

Hamiltonian function:

$$h = V(q) + K(p)$$

Potential energy:

$$V(q) = -\log \pi(q \mid \text{data}) + C$$

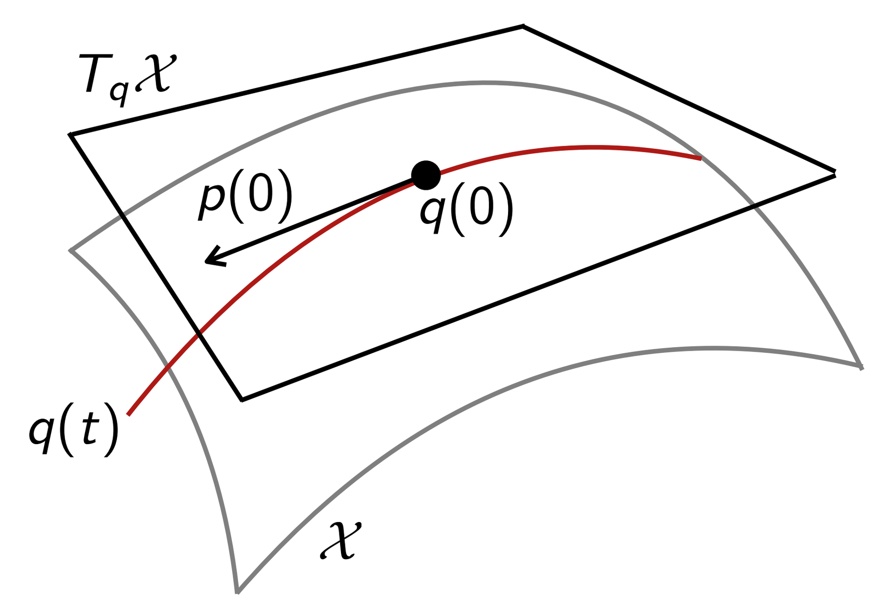
Kinetic energy:

$$K(p) = \frac{1}{2}\langle p, p\rangle$$

Hamiltonian equations:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \quad \text{and} \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q}$$

Jacobi-Maupertuis Principle:



Hamiltonian trajectories $q(t)$ with fixed h are geodesics on Riemannian manifolds $\mathcal{X}$ with the Jacobi metric:

$$g_h(p,p) = 2\,(h - V)\,\langle p, p\rangle.$$

HMC Meets Geometry

Reference metric $\langle \cdot, \cdot \rangle$

Conformal scaling of a metric $2\,(h - V)$

Equivalence table:

HMC	Riemannian geometry
posterior distribution	scaling of reference metric
covariance of proposal	reference metric
pick from proposal	geodesics

Gaussian Example

- HMC sampling from a Gaussian in $\mathbb{R}^d$ with zero mean and covariance matrix Λ^{-1}
- Calculate sectional curvature of Jacobi metric
- Sec is random because p is random
- Calculate coarse Ricci curvature from Sec

Our Lemma on positive curvature of HMC:

$$\mathbb{P}(d^2\,\text{Sec} \geq K_1) \geq 1 - K_2 e^{-K_3\sqrt{d}}$$

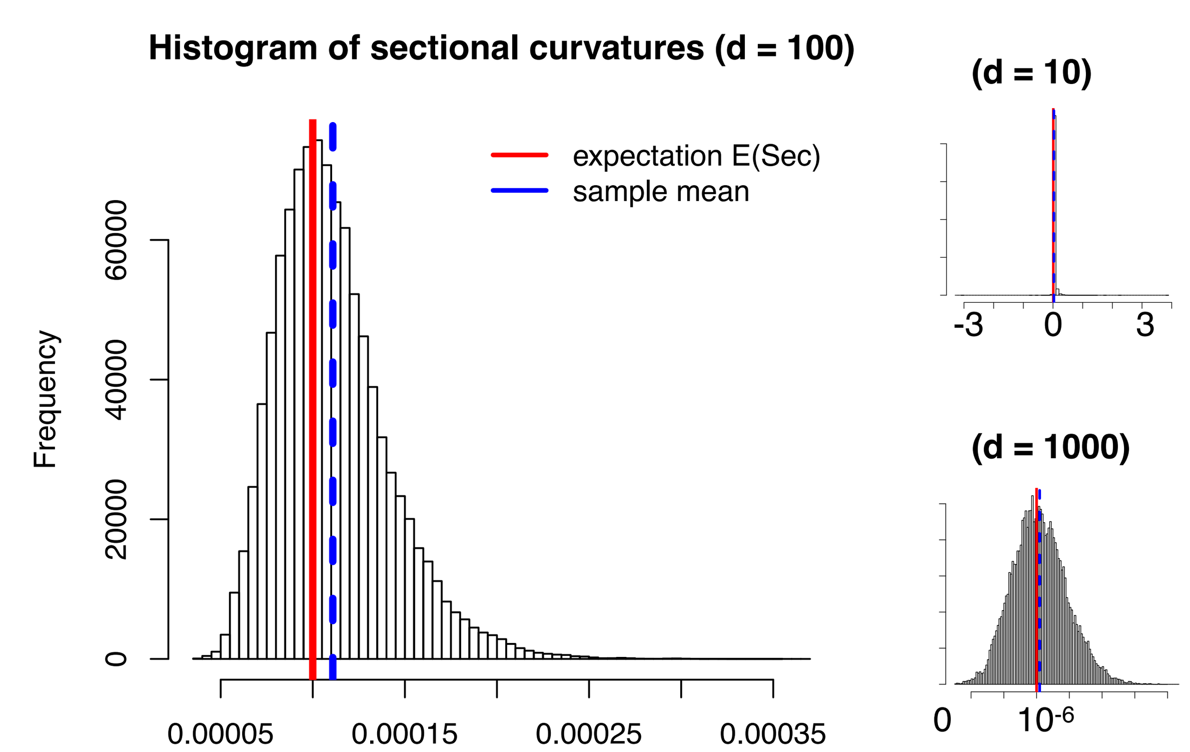
Our Corollary in high dimensions:

$$\text{Sec} \geq \frac{\text{Tr}(\Lambda)}{d^3} \quad \text{as} \quad d \to \infty$$

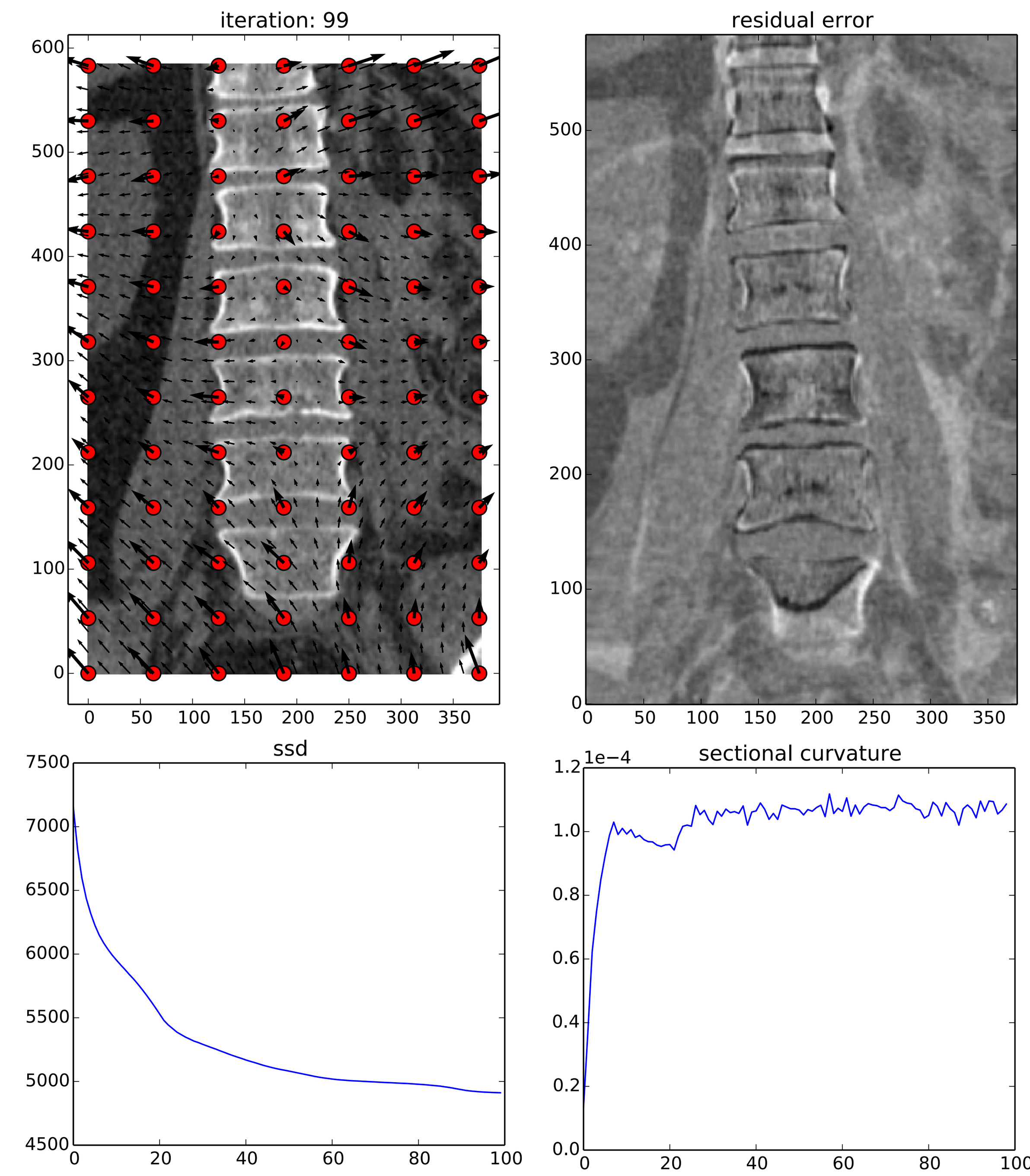
Our Proposal on coarse Ricci curvature of HMC:

$$\kappa \geq \frac{\text{Tr}(\Lambda)}{6d^2} + O\left(\frac{1}{d^2}\right) \quad \text{as} \quad d \to \infty$$

Numerical simulations: Sample Sec's uniformly at each step, $q(0), \dots, q(T-1)$



Computational Anatomy



Conclusion

Linking HMC to geodesics of the Jacobi metric enables us to analyze HMC through the rich mathematical framework of Riemannian geometry.

References

(JO10) A. Joulin and Y. Ollivier. Curvature, Concentration and Error Estimates for Markov Chain Monte Carlo. Ann. Probab., 2010.

(Pin75) O. C. Pin. Curvature and Mechanics. Advances in Math., 1975.

Acknowledgments

CS is supported by a postdoctoral fellowship from the Swiss National Science Foundation and a travel grant from the France-Stanford Center. SRS and SH are supported by NIH grant R01-GM086884.