Enhanced Gibbs sampling for an application to X-ray imaging

John Bardsley[†], Kevin Joyce[†*], Aaron Luttman[*]

[†]The University of Montana
[*]National Security Technologies LLC

April 25, 2015

Outline

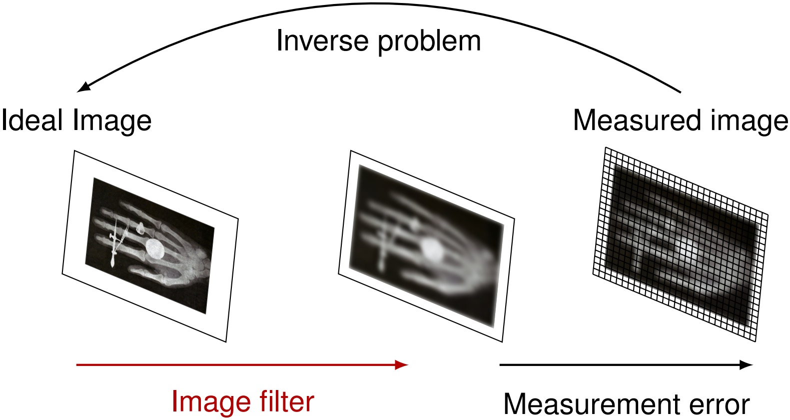
Modeling Imaging Systems

- Convolution with a point spread function
- Estimating the PSF with calibration images

Markov Chain Monte Carlo methods for posterior inference

- Introduction to Monte Carlo estimation
- The posterior density
- Gibbs Sampling
- Collapsing the Gibbs Sampler
- Results

Imaging Model



Imaging Model Assumptions

- A general model blur in imaging is as a linear filter $b = \mathcal{A}f$, where $f : \mathbb{R}^2 \to \mathbb{R}, b : \mathbb{R}^2 \to \mathbb{R}, a : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ and

$$b(\boldsymbol{x}) = \iint a(\boldsymbol{s}; \boldsymbol{x}) f(\boldsymbol{s}) d\boldsymbol{s}.$$

- When the blur is translation invariant, the model reduces to integral convolution when there exists a $k : \mathbb{R}^2 \to \mathbb{R}$ so that

$$b(\boldsymbol{x}) = \iint k(\boldsymbol{x} - \boldsymbol{s}) f(\boldsymbol{s}) d\boldsymbol{s}$$

- When blur is direction invariant or isotropic, this results in one more reduction with $p : [0, \infty) \to \mathbb{R}$ and

$$b(\boldsymbol{x}) = \iint p(|\boldsymbol{x} - \boldsymbol{s}|) f(\boldsymbol{s}) d\boldsymbol{s}$$

- Measurement error is modeled as Gaussian white noise.

Imaging Model Assumptions

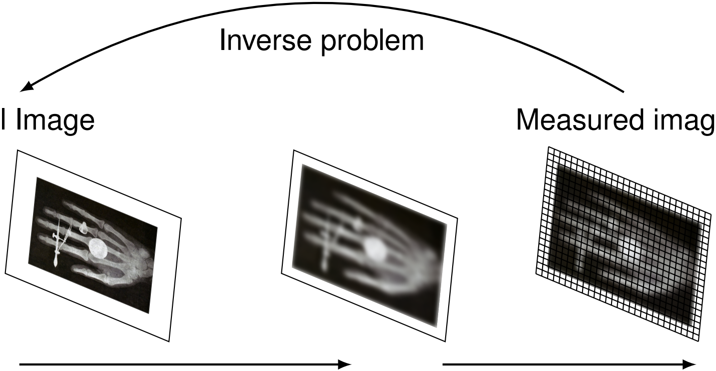


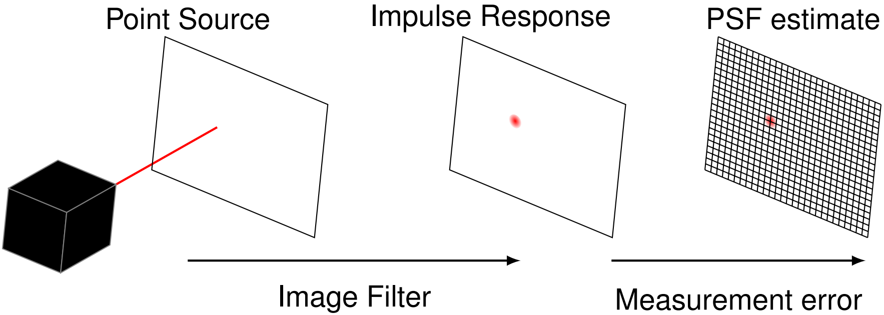
Image filter

$$b(x) = \iint p(|x - s|) \cdot f(s)\, ds$$

Measurement error

$$+\varepsilon \sim \text{white noise}$$

Point Spread Function Estimation

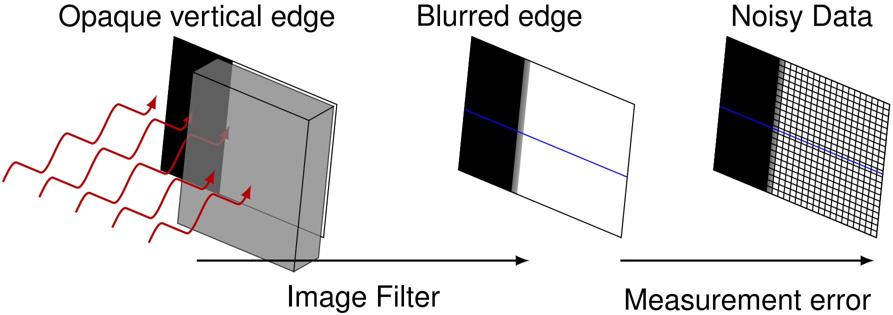


$$\iint p(|x - s|)\delta(s)\, ds$$

$$\iint p(|s|)\delta(x - s)\, ds = p(|x|)$$

$+\varepsilon \sim$ white noise

X-ray Radiography



$$\iint p(|x-s|)E(s)\,ds$$

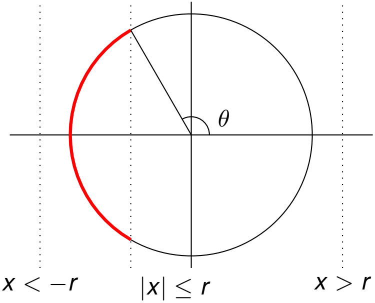
$$\iint p(|s|)E(x-s)\,ds = b(x)$$

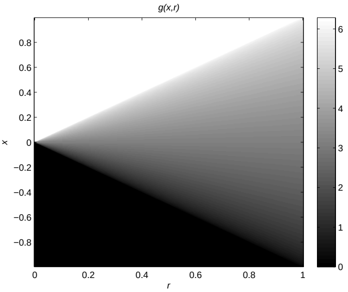
$+\varepsilon \sim$ white noise

Radially Symmetric PSF

We distinguish the radial profile from the kernel by $k(s,t) = p\left(\sqrt{s^2+t^2}\right)$

$$
\begin{aligned}
b(x,y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(s,t)E(x-s)dtds + \varepsilon_{x,y} \\
&= \int_0^{\infty} p(r)\left(\int_0^{2\pi} E(x - r\cos\theta)d\theta\right) rdr + \varepsilon_{x,y} \\
&= \int_0^{\infty} p(r)\cdot g(x,r)\, rdr + \varepsilon_{x,y}.
\end{aligned}
$$





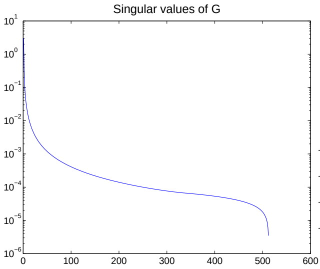
Observe that g is symmetric about $x = 0$.

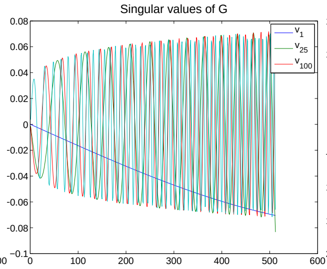
Radial Profile Inverse Problem

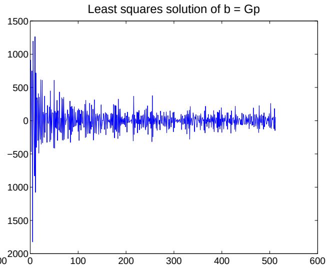
For the inverse problem

$$b = \mathcal{G}p + \varepsilon,$$

- Hence, the discretized problem $\boldsymbol{Gp} = \boldsymbol{b}$ results in an ill-conditioned a matrix
- The SVD of a matrix: $\boldsymbol{G} = U\Sigma V^*$, so the left-inverse is $\boldsymbol{G}^\dagger = V\Sigma^\dagger U^*$.







The discrete problem

In order to carry out estimation on a computer, we discretize the integral operator using mid-point quadrature

$$b = \mathcal{G}p + \epsilon \quad \Longrightarrow \quad \boldsymbol{b} = \boldsymbol{G}\boldsymbol{p} + \boldsymbol{\epsilon}$$

Further, we discretize the regularization operator $\mathcal{L}$ using finite differencing

$$\|\mathcal{L}p\|^2_{rad} = \int \left[\frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right)\right]^2 p(r) r dr \Longrightarrow \boldsymbol{L}\boldsymbol{p} = \boldsymbol{r}^{-1/2} \odot \boldsymbol{D}\left(\boldsymbol{r} \odot \boldsymbol{D}\boldsymbol{p}\right)$$

and midpoint quadrature for the inner products

$$\frac{\lambda}{2}\left\|b - \mathcal{G}p\right\|^2_{L^2} \Longrightarrow \underbrace{\frac{\lambda}{2m}}_{\lambda}\left\|\boldsymbol{b} - \boldsymbol{G}\boldsymbol{p}\right\|^2_{\mathbb{R}^m}$$

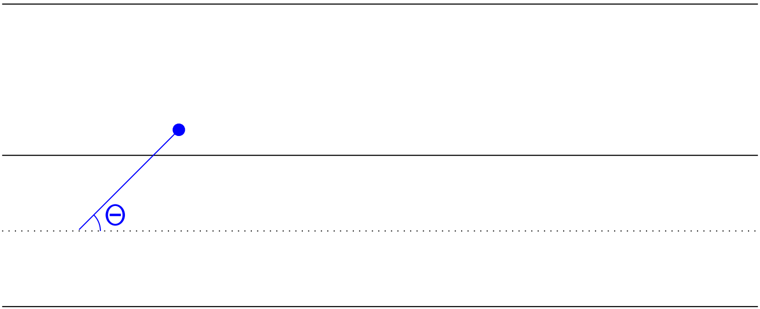
and

$$\frac{\delta}{2}\left\|\nabla^2 k\|_{L^2} \Longrightarrow \frac{\delta}{2}\left\|\mathcal{L}p\right\|^2_{rad} \Longrightarrow \underbrace{\frac{\delta}{2n}}_{\delta}\left\|\boldsymbol{L}\boldsymbol{p}\right\|^2_{\mathbb{R}^n}$$

Buffon’s Needle

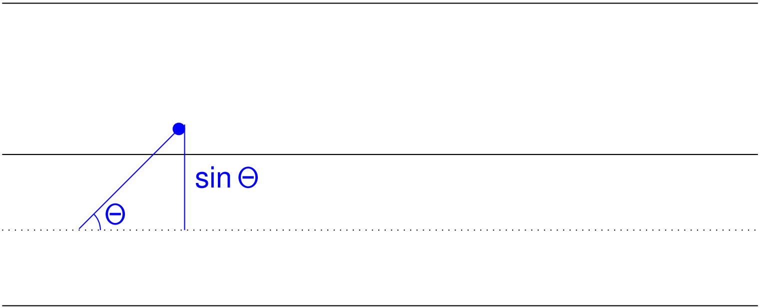
- Suppose a needle of unit length is dropped on a panelled floor with unit width panels.
- What is the probability that the needle crosses a panel line?

Buffon's Needle



- $C = 1$ if the needle crosses and 0 otherwise,
- Let Θ be the angle the bottom point of the needle makes with the horizontal,

Buffon's Needle



- $\mathbb{P}(C = 1|\Theta = \theta) = \sin\theta$ and $\mathbb{P}(C = 0|\Theta = \theta) = 1 - \sin\theta$.
- $\mathbb{P}(\Theta = \theta) = \frac{1}{\pi} I_{[0,\pi]}(\theta)$ where I_S is the indicator function for S.

Bayes' "Theorem"

- Bayes:
  (i) $\mathbb{P}(A = a | B = b)\mathbb{P}(B = b) = \mathbb{P}(A = a, B = b)$
  (ii) $\int \mathbb{P}(A = a, B = b)db = \mathbb{P}(A = a)$.

- $\mathbb{P}(C = 1 | \Theta = \theta)\mathbb{P}(\Theta = \theta) = \mathbb{P}(C = 1, \Theta = \theta)$

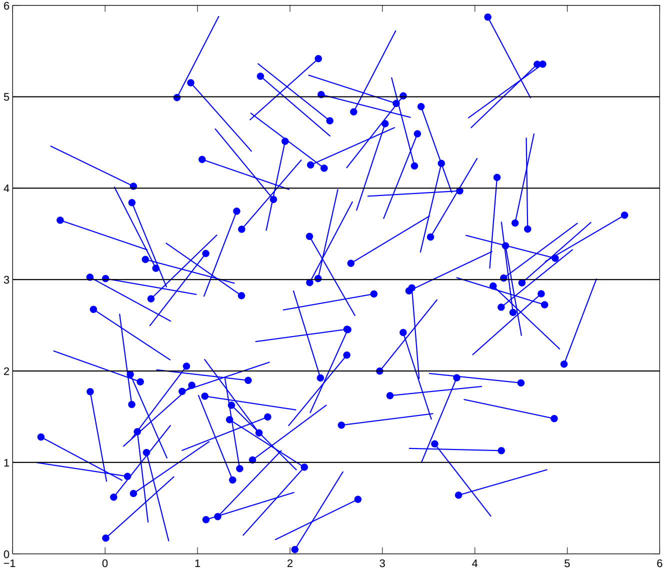
So,

$$\mathbb{P}(C = 1, \Theta = \theta) = \frac{1}{\pi} I_{[0,\pi]}(\theta) \cdot \sin\theta$$

and

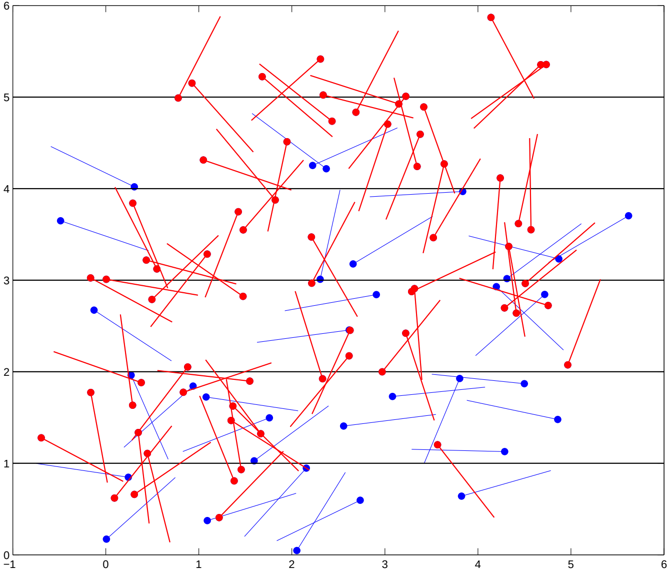
$$\mathbb{P}(C = 1) = \int \mathbb{P}(C = 1, \theta)d\theta = \int_0^{\pi} \frac{\sin\theta}{\pi} d\theta = \frac{2}{\pi}$$

Buffon's Monte Carlo Needles



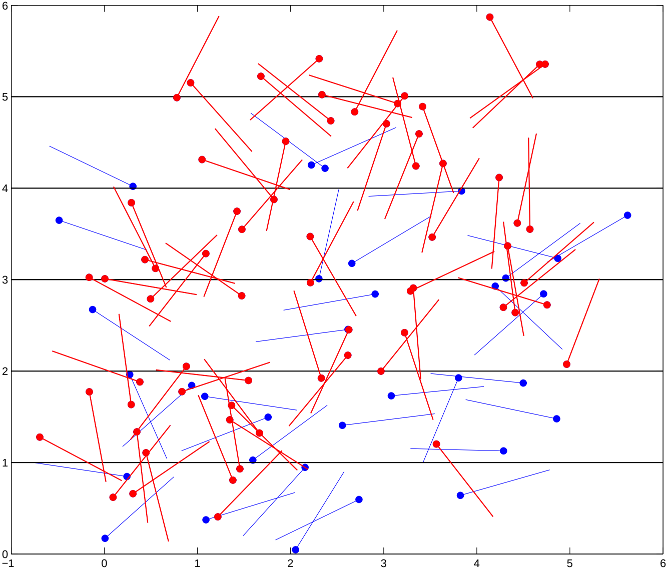
- Drop *N* needles.

Buffon's Monte Carlo Needles



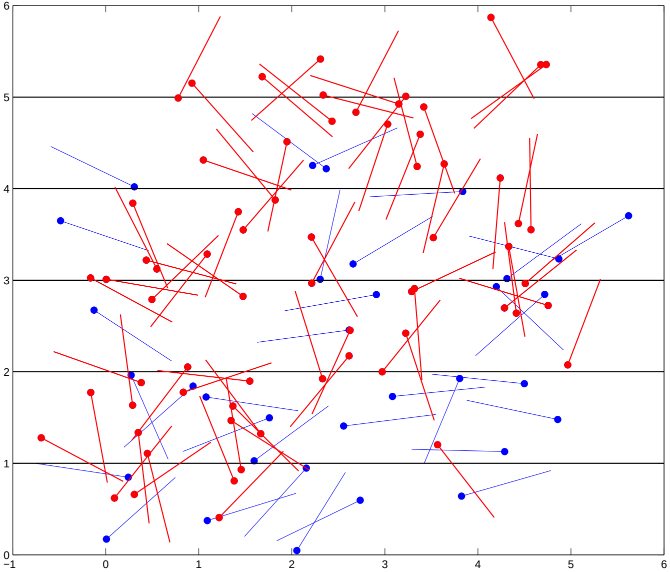
- Let $\{C_1, \ldots C_N\}$ be the simulated needle crossings.

Buffon's Monte Carlo Needles



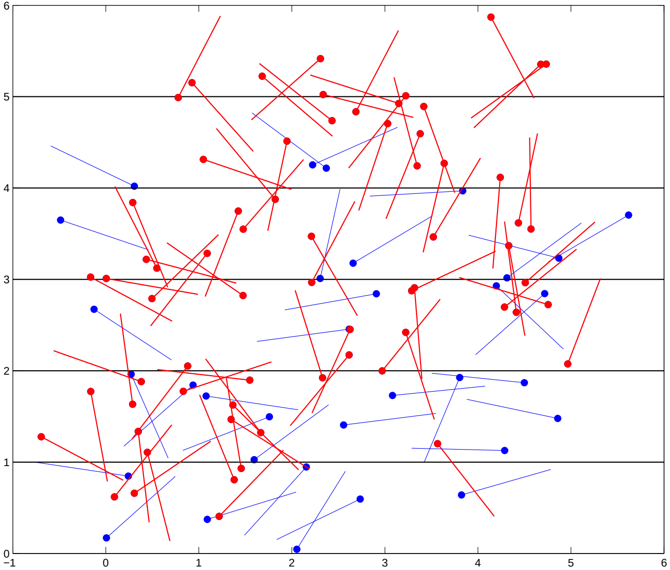
- The Law of Large Numbers says $2N / \sum C_i \to \pi$.

Buffon’s Monte Carlo Needles



- The Strong Law of Large Numbers says σ for $\frac{1}{N}\sum C_i$ is $O(1/\sqrt{N})$.

Buffon's Monte Carlo Needles



- So for $N = 100$, $\pi \approx 200/67 \approx 2.9851$ with $\sigma = O(10^{-1})$.

Hierarchical Model for PSF estimation

Let $\pi(\boldsymbol{x}) = \mathbb{P}(X = x)$ denote the probability density. For λ, δ and $\boldsymbol{p}$ and

$$\boldsymbol{b} = \boldsymbol{G}\boldsymbol{p} + \epsilon$$

assume

- The likelihood $\pi(\boldsymbol{b}|\boldsymbol{p}, \lambda, \delta) = \pi(\boldsymbol{b}|\boldsymbol{p}, \lambda) \propto \lambda^{M/2} \exp\left(-\frac{\lambda}{2}\|\boldsymbol{b} - \boldsymbol{G}\boldsymbol{p}\|^2\right)$ since ϵ is independent Gaussian noise.
- The prior $\pi(\boldsymbol{p}|\delta, \lambda) = \pi(\boldsymbol{p}|\delta) \propto \delta^{N/2} \exp\left(-\frac{\delta}{2}\|\boldsymbol{L}\boldsymbol{p}\|^2\right)$ since $k \sim \mathcal{N}(0, \nabla^{-2}) \implies p \sim N(0, \mathcal{L}^{-2})$
- The hyperpriors $\pi(\lambda) \propto \exp\left(-10^{-4}\lambda\right)$ and $\pi(\delta) \propto \exp\left(-10^{-4}\delta\right)$ are independent "unobjective" Gamma distributions.

Bayesian Posterior

With $\pi(\boldsymbol{b}|\boldsymbol{p},\lambda,\delta), \pi(\boldsymbol{p}|\delta,\lambda)$, and $\pi(\lambda,\delta)$, use Bayes' "Theorem" to obtain

$$\pi_{\mathbf{b}}(\boldsymbol{p},\lambda,\delta) \stackrel{\text{def}}{=} \pi(\boldsymbol{p},\lambda,\delta|\boldsymbol{b}) = \pi(\boldsymbol{b},\boldsymbol{p},\lambda,\delta)/\pi(\boldsymbol{b})$$
$$\propto \lambda^{M/2}\delta^{N/2}\exp\left(-\frac{\lambda}{2}\left\|\boldsymbol{b}-\boldsymbol{G}\boldsymbol{p}\right\|_{\mathbb{R}^m} - \frac{\delta}{2}\left\|\boldsymbol{L}\boldsymbol{p}\right\|_{\mathbb{R}^n} - 10^{-4}(\lambda+\delta)\right)$$

- This is not a "common" probability density, hence simulations from a computer are not readily available.
- Leveraging conditional density ideas will give us a method to simulate a Monte Carlo method like Buffon's needle problem.
- Bayes' "Theorem" will allow simulations from the full conditionals $\pi_{\mathbf{b}}(\lambda|\delta,\boldsymbol{p}), \pi_{\mathbf{b}}(\delta|\lambda,\boldsymbol{p})$ and $\pi_{\mathbf{b}}(\boldsymbol{p}|\lambda,\delta)$.

Markov Chains

- A transition operator is a linear map on probability densities π, so that

$$\mathcal{K}[\pi](\boldsymbol{x}') = \int K(\boldsymbol{x}, \boldsymbol{x}')\pi(\boldsymbol{x})d\boldsymbol{x}.$$

  where $K(\boldsymbol{x}, \cdot)$, the transition kernel.

- A Markov chain is a stochastic process $\{\boldsymbol{X}^0, \boldsymbol{X}^1, \boldsymbol{X}^2, \dots, \}$ with $\boldsymbol{X}^k : \Omega \to \mathbb{R}^M$ such that

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{X}^{k+1} = x^{k+1} | \boldsymbol{X}^k = \boldsymbol{x}^k, \dots, \boldsymbol{X}^0 = \boldsymbol{x}^0\right) &= \mathbb{P}\left(\boldsymbol{X}^{k+1} = \boldsymbol{x}^{k+1} | \boldsymbol{X}^k = \boldsymbol{x}^k\right) \\ &= K(\boldsymbol{x}^k, \boldsymbol{x}^{k+1}).\end{aligned}$$

  So, given an initial density $\pi_0(\boldsymbol{x}^0)$

$$\mathbb{P}(\boldsymbol{X}^N = \boldsymbol{x}^N) = \mathcal{K}^N[\pi_0](\boldsymbol{x}^N)$$

- Markov chains are not independent and not identically distributed.

An Ergodic Theorem

Theorem[Tierney1994]: Suppose $\mathcal{K}$ defines a stationary Markov chain with invariant density π. If the chain is π-irreducible and Harris recurrent, then π is unique and for any initial density π_0 and all $\boldsymbol{x}$ but a subset whose measure under π is zero. Moreover,

(i) Almost surely with respect to π, for any integrable h

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} h(\boldsymbol{X}^n) = \int h(\boldsymbol{x})\pi(\boldsymbol{x})d\boldsymbol{x}. \tag{1}$$

(ii) If in addition, the chain is aperiodic, then

$$\lim_{N\to\infty} \|\mathcal{K}^N \pi_0 - \pi\|_{TV} = 0,$$

where $\|\pi\|_{TV}$ denotes the supremum of $\int_A \pi(x)dx$ over all Borel sets A.

Ergodic Theorem Hypotheses

- A Markov chain is stationary and invariant with respect to π if $\mathcal{K}\pi = \pi$.

- MCMC idea: Design a Markov chain that has $\pi_{\mathbf{b}}(\boldsymbol{p}, \lambda, \delta)$ as its invariant density, i.e. $\mathcal{K}\pi_{\mathbf{b}} = \pi_{\mathbf{b}}$.

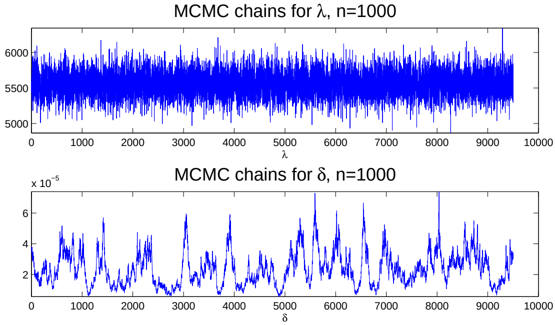
- Harris recurrence , irreducibility, and aperiodicity are necessary conditions so ensure that the evolution of the chain "explores" the invariant density π enough and avoids transient states.

- The first step to appealing to the ergodic theorem is to ensure invariance with the desired density.

MCMC Inference by the Ergodic Theorem

- The first conclusion of the Ergodic theorem says that statistics on the realizations of the chain converge as expected.

- The second conclusion is distributional, and implies that late stages in the chain are “approximately” identically distributed, but not independent.



Gibbs sampling

The Gibbs sampler [Geman and Geman 1984]:
Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda | \delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta | \lambda^k, \boldsymbol{p}^{k-1})$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p} | \lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to π:

$$
\begin{aligned}
[\mathcal{K}\pi_{\mathbf{b}}](\lambda', \delta', \boldsymbol{p}') &= \iiint \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \int \underbrace{\pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta, \mathbf{p}) / \pi_{\mathbf{b}}(\delta, \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \underbrace{\pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta', \mathbf{p}) / \pi_{\mathbf{b}}(\lambda', \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p}) d\delta d\boldsymbol{p} \\
&\cdots \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta'), \text{ so the Gibbs sampler is invariant.}
\end{aligned}
$$

Gibbs sampling

The Gibbs sampler [Geman and Geman 1984]:
Given $(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1})$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta|\lambda^k, \boldsymbol{p}^{k-1})$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to π:

$$[\mathcal{K}\pi_{\mathbf{b}}](\lambda', \delta', \boldsymbol{p}') = \iiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta'|\lambda', \boldsymbol{p})\pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p})\pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})d\lambda d\delta d\boldsymbol{p}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta') \int \pi_{\mathbf{b}}(\delta'|\lambda', \boldsymbol{p}) \int \underbrace{\pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta, \mathbf{p})/\pi_{\mathbf{b}}(\delta, \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})d\lambda d\delta d\boldsymbol{p}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta') \int \underbrace{\pi_{\mathbf{b}}(\delta'|\lambda', \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta', \mathbf{p})/\pi_{\mathbf{b}}(\lambda', \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p})d\delta d\boldsymbol{p}$$

$$\cdots$$

$= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta')$, so the Gibbs sampler is invariant.

Gibbs sampling

The Gibbs sampler [Geman and Geman 1984]:
Given $(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1})$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda | \delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta | \lambda^k, \boldsymbol{p}^{k-1})$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p} | \lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to π:

$$
\begin{aligned}
[\mathcal{K}\pi_{\mathbf{b}}](\lambda', \delta', \boldsymbol{p}') &= \iiint \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \int \underbrace{\pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta, \mathbf{p}) / \pi_{\mathbf{b}}(\delta, \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \underbrace{\pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta', \mathbf{p}) / \pi_{\mathbf{b}}(\lambda', \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p}) d\delta d\boldsymbol{p} \\
&\dots \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta'), \text{ so the Gibbs sampler is invariant.}
\end{aligned}
$$

Gibbs sampling

The Gibbs sampler [Geman and Geman 1984]:
Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda | \delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta | \lambda^k, \boldsymbol{p}^{k-1})$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p} | \lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to π:

$$
\begin{aligned}
[\mathcal{K}\pi_{\mathbf{b}}](\lambda', \delta', \boldsymbol{p}') &= \iiint \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \int \underbrace{\pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta, \mathbf{p}) / \pi_{\mathbf{b}}(\delta, \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \underbrace{\pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta', \mathbf{p}) / \pi_{\mathbf{b}}(\lambda', \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p}) d\delta d\boldsymbol{p} \\
&\cdots \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta'), \text{ so the Gibbs sampler is invariant.}
\end{aligned}
$$

Gibbs sampling

The Gibbs sampler [Geman and Geman 1984]:
Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta|\lambda^k, \boldsymbol{p}^{k-1})$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to π:

$$
\begin{aligned}
[\mathcal{K}\pi_{\mathbf{b}}](\lambda', \delta', \boldsymbol{p}') &= \iiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta'|\lambda', \boldsymbol{p})\pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p})\pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\,d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta') \int \pi_{\mathbf{b}}(\delta'|\lambda', \boldsymbol{p}) \int \underbrace{\pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta, \mathbf{p})/\pi_{\mathbf{b}}(\delta, \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\,d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta') \int \underbrace{\pi_{\mathbf{b}}(\delta'|\lambda', \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta', \mathbf{p})/\pi_{\mathbf{b}}(\lambda', \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p})\,d\delta d\boldsymbol{p} \\
&\dots \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta'), \text{ so the Gibbs sampler is invariant.}
\end{aligned}
$$

Gibbs sampling

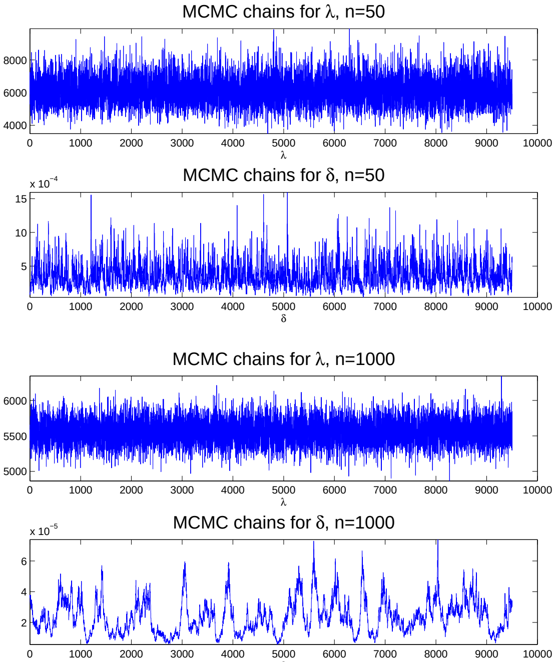
The Gibbs sampler [Geman and Geman 1984]:
Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda | \delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta | \lambda^k, \boldsymbol{p}^{k-1})$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p} | \lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to π:

$$
\begin{aligned}
[\mathcal{K}\pi_{\mathbf{b}}](\lambda', \delta', \boldsymbol{p}') &= \iiint \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p}) \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p}) \int \underbrace{\pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta, \mathbf{p}) / \pi_{\mathbf{b}}(\delta, \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) d\lambda d\delta d\boldsymbol{p} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \int \underbrace{\pi_{\mathbf{b}}(\delta' | \lambda', \boldsymbol{p})}_{\pi_{\mathbf{b}}(\lambda', \delta', \mathbf{p}) / \pi_{\mathbf{b}}(\lambda', \mathbf{p})} \int \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p}) d\delta d\boldsymbol{p} \\
& \dots \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta'), \text{ so the Gibbs sampler is invariant.}
\end{aligned}
$$

Correlated δ chains



Literature on the Issue

- [Agapiou,Bardsley,Stuart,Papaspiliopoulos, 2014] explained this phenomena theoretically for a general class of Laplacian based Hierarchical samplers for inverse problems.
- The issue arises when the discretization of $\boldsymbol{p}$ closer approximates the continuum, the correlation in the δ component of the Markov Chain becomes more correlated.
- [VanDyke, Park 2008] provide a general method for removing the dependence of problematic components in the Gibbs sampler, called partial collapse.
- The idea has been independently derived in many places, however, if done carelessly [VanDyke, Park 2008] showed that the resulting Markov chain is no longer invariant, although invariance was not proved there.

Marginalized Sampler

Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}, \widetilde{\boldsymbol{p}}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda | \delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $(\delta^k, \widetilde{\boldsymbol{p}}^k) \sim \pi_{\mathbf{b}}(\delta, \widetilde{\boldsymbol{p}} | \lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p} | \lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to $\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) \pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}} | \lambda, \delta)$:

$$
\begin{aligned}
[\mathcal{K}\widetilde{\pi_{\mathbf{b}}}] &= \iiiint \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}' | \lambda') \pi_{\mathbf{b}}(\lambda', \boldsymbol{p} | \delta) \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) d\lambda d\delta d\boldsymbol{p} d\widetilde{\boldsymbol{p}} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta') \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}' | \lambda') \iint \pi_{\mathbf{b}}(\lambda' | \delta, \boldsymbol{p}) \underbrace{\int\int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) d\widetilde{\boldsymbol{p}} d\lambda}_{\pi_{\mathbf{b}}(\delta, \mathbf{p})} d\delta d\boldsymbol{p} \\
&= \underbrace{\pi_{\mathbf{b}}(\boldsymbol{p}' | \lambda', \delta')}_{\pi_{\mathbf{b}}(\mathbf{p}', \delta', \lambda') / \pi_{\mathbf{b}}(\delta', \lambda')} \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}' | \lambda') \underbrace{\iint \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p}) d\delta d\boldsymbol{p}}_{\pi_{\mathbf{b}}(\lambda')} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta') \frac{\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}', \lambda', \delta')}{\pi_{\mathbf{b}}(\lambda', \delta')} \\
&= \widetilde{\pi_{\mathbf{b}}}(\lambda', \delta', \boldsymbol{p}', \widetilde{\boldsymbol{p}}')
\end{aligned}
$$

Marginalized Sampler

Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}, \widetilde{\boldsymbol{p}}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $(\delta^k, \widetilde{\boldsymbol{p}}^k) \sim \pi_{\mathbf{b}}(\delta, \widetilde{\boldsymbol{p}}|\lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to $\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}|\lambda, \delta)$:

$$[\mathcal{K}\widetilde{\pi_{\mathbf{b}}}] = \iiiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda')\pi_{\mathbf{b}}(\lambda', \boldsymbol{p}|\delta)\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})\,d\lambda\, d\delta\, d\boldsymbol{p}\, d\widetilde{\boldsymbol{p}}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \iint \pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p}) \underbrace{\int\int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})\,d\widetilde{\boldsymbol{p}}\,d\lambda}_{\pi_{\mathbf{b}}(\delta, \mathbf{p})}\, d\delta\, d\boldsymbol{p}$$

$$= \underbrace{\pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')}_{\pi_{\mathbf{b}}(\mathbf{p}', \delta', \lambda')/\pi_{\mathbf{b}}(\delta', \lambda')} \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \underbrace{\iint \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p})\,d\delta\, d\boldsymbol{p}}_{\pi_{\mathbf{b}}(\lambda')}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta') \frac{\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}', \lambda', \delta')}{\pi_{\mathbf{b}}(\lambda', \delta')}$$

$$= \widetilde{\pi_{\mathbf{b}}}(\lambda', \delta', \boldsymbol{p}', \widetilde{\boldsymbol{p}}')$$

Marginalized Sampler

Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}, \widetilde{\boldsymbol{p}}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $(\delta^k, \widetilde{\boldsymbol{p}}^k) \sim \pi_{\mathbf{b}}(\delta, \widetilde{\boldsymbol{p}}|\lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to $\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}|\lambda, \delta)$:

$$[\mathcal{K}\widetilde{\pi_{\mathbf{b}}}] = \iiiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda')\pi_{\mathbf{b}}(\lambda', \boldsymbol{p}|\delta)\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\lambda d\delta d\boldsymbol{p} d\widetilde{\boldsymbol{p}}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \iint \pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p}) \underbrace{\int \int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\widetilde{\boldsymbol{p}}d\lambda}_{\pi_{\mathbf{b}}(\delta, \mathbf{p})} d\delta d\boldsymbol{p}$$

$$= \underbrace{\pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')}_{\pi_{\mathbf{b}}(\mathbf{p}', \delta', \lambda')/\pi_{\mathbf{b}}(\delta', \lambda')} \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \underbrace{\iint \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p})d\delta d\boldsymbol{p}}_{\pi_{\mathbf{b}}(\lambda')}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta')\frac{\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}', \lambda', \delta')}{\pi_{\mathbf{b}}(\lambda', \delta')}$$

$$= \widetilde{\pi_{\mathbf{b}}}(\lambda', \delta', \boldsymbol{p}', \widetilde{\boldsymbol{p}}')$$

Marginalized Sampler

Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}, \widetilde{\boldsymbol{p}}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $(\delta^k, \widetilde{\boldsymbol{p}}^k) \sim \pi_{\mathbf{b}}(\delta, \widetilde{\boldsymbol{p}}|\lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to $\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}|\lambda, \delta)$:

$$
\begin{aligned}
[\mathcal{K}\widetilde{\pi_{\mathbf{b}}}] &= \iiiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda')\pi_{\mathbf{b}}(\lambda', \boldsymbol{p}|\delta)\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\lambda d\delta d\boldsymbol{p} d\widetilde{\boldsymbol{p}} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \iint \pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p}) \underbrace{\int\int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\widetilde{\boldsymbol{p}}d\lambda}_{\pi_{\mathbf{b}}(\delta, \mathbf{p})} d\delta d\boldsymbol{p} \\
&= \underbrace{\pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')}_{\pi_{\mathbf{b}}(\mathbf{p}', \delta', \lambda')/\pi_{\mathbf{b}}(\delta', \lambda')} \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \underbrace{\iint \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p})d\delta d\boldsymbol{p}}_{\pi_{\mathbf{b}}(\lambda')} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta')\frac{\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}', \lambda', \delta')}{\pi_{\mathbf{b}}(\lambda', \delta')} \\
&= \widetilde{\pi_{\mathbf{b}}}(\lambda', \delta', \boldsymbol{p}', \widetilde{\boldsymbol{p}}')
\end{aligned}
$$

Marginalized Sampler

Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}, \widetilde{\boldsymbol{p}}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $(\delta^k, \widetilde{\boldsymbol{p}}^k) \sim \pi_{\mathbf{b}}(\delta, \widetilde{\boldsymbol{p}}|\lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to $\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}|\lambda, \delta)$:

$$[\mathcal{K}\widetilde{\pi_{\mathbf{b}}}] = \iiiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda')\pi_{\mathbf{b}}(\lambda', \boldsymbol{p}|\delta)\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\lambda d\delta d\boldsymbol{p}d\widetilde{\boldsymbol{p}}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \iint \pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p}) \underbrace{\int\int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\widetilde{\boldsymbol{p}}d\lambda}_{\pi_{\mathbf{b}}(\delta, \mathbf{p})} d\delta d\boldsymbol{p}$$

$$= \underbrace{\pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')}_{\pi_{\mathbf{b}}(\mathbf{p}', \delta', \lambda')/\pi_{\mathbf{b}}(\delta', \lambda')} \quad \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \underbrace{\iint \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p})d\delta d\boldsymbol{p}}_{\pi_{\mathbf{b}}(\lambda')}$$

$$= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta')\frac{\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}', \lambda', \delta')}{\pi_{\mathbf{b}}(\lambda', \delta')}$$

$$= \widetilde{\pi_{\mathbf{b}}}(\lambda', \delta', \boldsymbol{p}', \widetilde{\boldsymbol{p}}')$$

Marginalized Sampler

Given $\left(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1}, \widetilde{\boldsymbol{p}}^{k-1}\right)$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda|\delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $(\delta^k, \widetilde{\boldsymbol{p}}^k) \sim \pi_{\mathbf{b}}(\delta, \widetilde{\boldsymbol{p}}|\lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p}|\lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to $\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}|\lambda, \delta)$:

$$
\begin{aligned}
[\mathcal{K}\widetilde{\pi_{\mathbf{b}}}] &= \iiiint \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda')\pi_{\mathbf{b}}(\lambda', \boldsymbol{p}|\delta)\widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}})d\lambda d\delta d\boldsymbol{p} d\widetilde{\boldsymbol{p}} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')\pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \iint \pi_{\mathbf{b}}(\lambda'|\delta, \boldsymbol{p}) \underbrace{\int \int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) d\widetilde{\boldsymbol{p}} d\lambda}_{\pi_{\mathbf{b}}(\delta, \mathbf{p})} d\delta d\boldsymbol{p} \\
&= \underbrace{\pi_{\mathbf{b}}(\boldsymbol{p}'|\lambda', \delta')}_{\pi_{\mathbf{b}}(\mathbf{p}', \delta', \lambda')/\pi_{\mathbf{b}}(\delta', \lambda')} \quad \pi_{\mathbf{b}}(\delta', \widetilde{\boldsymbol{p}}'|\lambda') \underbrace{\iint \pi_{\mathbf{b}}(\lambda', \delta, \boldsymbol{p}) d\delta d\boldsymbol{p}}_{\pi_{\mathbf{b}}(\lambda')} \\
&= \pi_{\mathbf{b}}(\boldsymbol{p}', \lambda', \delta') \frac{\pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}}', \lambda', \delta')}{\pi_{\mathbf{b}}(\lambda', \delta')} \\
&= \widetilde{\pi_{\mathbf{b}}}(\lambda', \delta', \boldsymbol{p}', \widetilde{\boldsymbol{p}}')
\end{aligned}
$$

Partially Collapsed Sampler

Given $(\lambda^{k-1}, \delta^{k-1}, \boldsymbol{p}^{k-1})$, simulate

1: $\lambda^k \sim \pi_{\mathbf{b}}(\lambda | \delta^{k-1}, \boldsymbol{p}^{k-1})$
2: $\delta^k \sim \pi_{\mathbf{b}}(\delta | \lambda^k)$
3: $\boldsymbol{p}^k \sim \pi_{\mathbf{b}}(\boldsymbol{p} | \lambda^k, \delta^k)$

The associated Markov Chain is invariant with respect to
$\int \widetilde{\pi_{\mathbf{b}}}(\lambda, \delta, \boldsymbol{p}, \widetilde{\boldsymbol{p}}) = \int \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p}) \pi_{\mathbf{b}}(\widetilde{\boldsymbol{p}} | \lambda, \delta) = \pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})$.

- The order of the chain matters in the previous arguments.
- Permuting steps 2 and 3 results in a chain that is no longer invariant with respect to $\pi_{\mathbf{b}}$.
- Cyclically permuting results in a different sampler as well, however, this does not practically effect the overall chain, only the first and last steps, e.g.
$(\mathcal{K}_1\mathcal{K}_3\mathcal{K}_2)^N = \mathcal{K}_1(\mathcal{K}_3\mathcal{K}_2\mathcal{K}_1)^{N-1}\mathcal{K}_2\mathcal{K}_3$

Marginalized Posterior Density

In order to sample $\pi_{\mathbf{b}}(\delta|\lambda)$, we complete the square of the quadratic form in $\pi_{\mathbf{b}}(\lambda, \delta, \boldsymbol{p})$ and integrate out λ, this results in

$$\pi_{\mathbf{b}}(\delta|\lambda) \propto \exp\left((n/2)\delta - \ln|\det \boldsymbol{J}_{\lambda,\delta}| - \frac{\lambda}{2}\langle \boldsymbol{b}, \boldsymbol{H}_{\lambda,\delta}\boldsymbol{b}\rangle - 10^{-4}\delta\right),$$

where $\boldsymbol{J}_{\lambda,\delta} = \lambda \boldsymbol{G}^T\boldsymbol{G} + \delta \boldsymbol{L}$ and $\boldsymbol{H}_{\lambda,\delta} = \boldsymbol{I} - \lambda \boldsymbol{G}\boldsymbol{J}_{\lambda,\delta}^{-1}\boldsymbol{G}^T$.

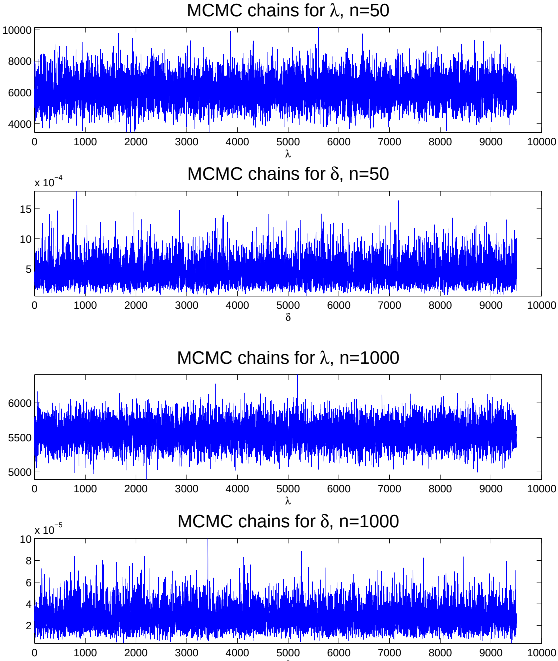
- To draw from this density, we embed a Metropolis-Hastings algorithm within the Gibbs sampler.
- It can be shown that the resulting MCMC algorithm remains invariant.
- Both constants can be carried out using a Cholesky factorization $\boldsymbol{R}_{\lambda,\delta}^T\boldsymbol{R}_{\lambda,\delta} = J_{\lambda,\delta}$ in $(O(n^3))$ flops, and will be required for each Metropolis step.

Metropolis-Hastings within Gibbs Sampling

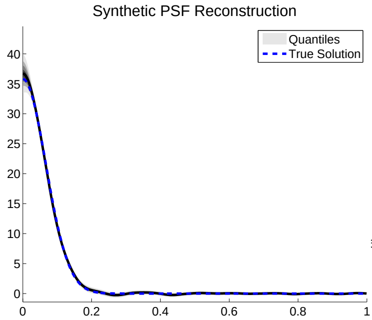
- We use n_{MH} Metropolis-Hastings steps using a Gaussian proposal with variance γ.
- Due to numerical overflow issues, all computations are carried out on the log scale.
- The full implementation is:

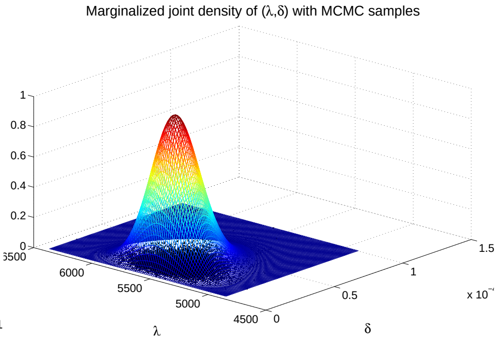
1: Let λ_0, δ_0,$\boldsymbol{p}$ and γ be given.
2: Draw λ_{k+1} from $\Gamma\left(n/2+\alpha,\ \frac{1}{2}\|\boldsymbol{Ax_k}-\boldsymbol{b}\|^2-\beta\right)$.
3: Set $j=1$, Compute $\boldsymbol{R}_0^T\boldsymbol{R}_0=\boldsymbol{J}_{\lambda^k,\delta^{k-1}}$, then $\pi_0=\log\pi_{\mathbf{b}}(\delta^{k-1}|\lambda^k)$.
4: **for** $1<j<n_{MH}$ **do**
5: Draw $\widetilde{\delta}$ from $N(\widetilde{\delta}_{j-1},\gamma)$.
6: Compute $\widetilde{\boldsymbol{R}}^T\widetilde{\boldsymbol{R}}=J_{\lambda^k,\widetilde{\delta}}$, then $\pi_j=\log p(\widetilde{\delta}_j|\boldsymbol{f}\lambda_k)$.
7: Set $\widetilde{\pi}=\pi_j$, $\boldsymbol{R}_j=\widetilde{\boldsymbol{R}}$ and $\delta_k=\widetilde{\delta}_j$ with probability $\widetilde{\pi}/\pi_j$
8: **end for**
9: Draw $\boldsymbol{p}_k$ from $N\left(\boldsymbol{R}_k^{-1}\lambda_k\boldsymbol{A}^T\boldsymbol{b},\boldsymbol{R}_k^{-1}\right)$.
10: Set $k=k+1$ and return to 2.

Marginalized δ chains



Results





Algorithm	$\widehat{\lambda}_{\text{MCMC}}$ $(\times 10^4)$	$\widehat{\delta}_{\text{MCMC}}$ $(\times 10^{-8})$	λ-p_{Geweke}	δ-p_{Geweke}	IACT	ESS	#Chol/ESS
Gibbs	1.102	6.132	0.998	0.850	36.2	138.0	72.4
PC Gibbs $n_{mh} = 1$	1.102	5.611	0.992	0.943	7.9	633.0	31.6
PC Gibbs $n_{mh} = 5$	1.102	5.515	0.999	0.985	1.3	3799.6	15.8

Summary and Future Work

- We introduced a novel Hierarchical Bayesian non-parametric model for estimating translation invariant and isotropic image blur with and edge.
- We developed the Partially Collapsed Gibbs sampler from the Gibbs sampler, and showed when partial collapse remained stationary.
- We then implemented the algorithm on a synthetic example using Metropolis with Partially Collapsed Gibbs, and showed that it improves the standard Gibbs sampler.
- Future: Develop the model and algorithm completely in infinite dimensions.
- Future: Adapt the strategies to other imaging models that incorporate radial geometry such as Abel and Radon transforms.

References

Sergios Agapiou, Johnathan M Bardsley, Omiros Papaspiliopoulos, and Andrew M Stuart.
Analysis of the Gibbs sampler for hierarchical inverse problems.
SIAM/ASA Journal on Uncertainty Quantification, 2(1):511–544, 2014.

Johnathan M Bardsley.
MCMC-based image reconstruction with uncertainty quantification.
SIAM Journal on Scientific Computing, 34(3):A1316–A1332, 2012.

Daniela Calvetti and Erkki Somersalo.
An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing, volume 2.
Springer Science & Business Media, 2007.

A. M. Stuart.
Inverse problems: A Bayesian perspective.
Acta Numerica, 19:451–559, 2010.