Solar Spectral Analyses with Uncertainties in Atomic Physical Models

Xixi Yu

Imperial College London

xixi.yu16@imperial.ac.uk

3 April 2018

Acknowledgments

Joint work with the International Space Science Institute (ISSI) team

"Improving the Analysis of Solar and Stellar Observations"

The Solar Corona

- The solar corona is a complex and dynamic system
- Measuring physical properties in any solar region is important for understanding the processes that lead to these events

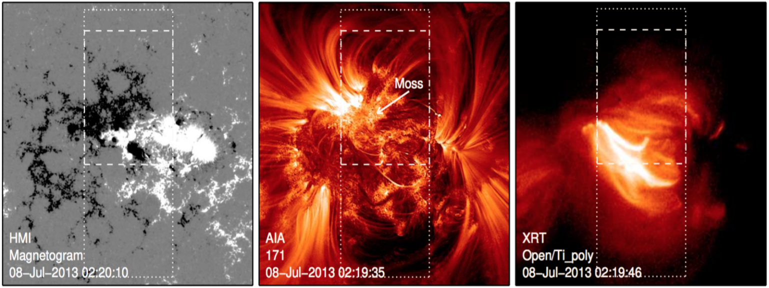


Figure: The photospheric magnetic field measured with HMI, million degree emission observed with the AIA Fe IX 171,Å channel, and high temperature loops observed with XRT

Aim

- We want to infer physical quantities of the solar atmosphere (density, temperature, path length, etc.), but we only observe intensity
- Inferences rely on models for the underlying atomic physics
- How to address **uncertainty** in the atomic physics models?

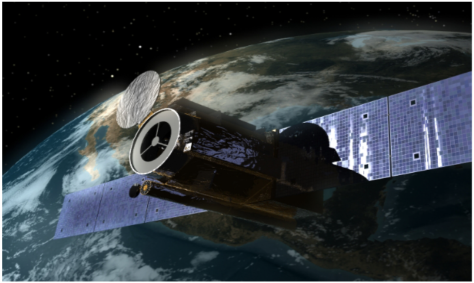


Figure: Hinode spacecraft.

Physical Parameters

- n_k [1]: number of free electrons per unit volume in plasma
- T_k: electron temperature
- d_k: path length through the solar atmosphere
- $\theta_k = (\log \mathrm{n}_k, \log \mathrm{d}_k)$
- m: index of the emissivity curve
- Expected intensity of line with wavelength λ:

$$\epsilon_\lambda^{(m)}(\mathrm{n}_k, \mathrm{T}_k)\mathrm{n}_k^2\mathrm{d}_k$$

- $\epsilon_\lambda^{(m)}(\mathrm{n}_k, \mathrm{T}_k)$ is the plasma emissivity for the line with wavelength λ in pixel k

[1] Subscript k is the pixel index

Data: Observed Intensity

- Data from the Extreme-Ultraviolet Imaging Spectrometer (EIS) on *Hinode* spacecraft.

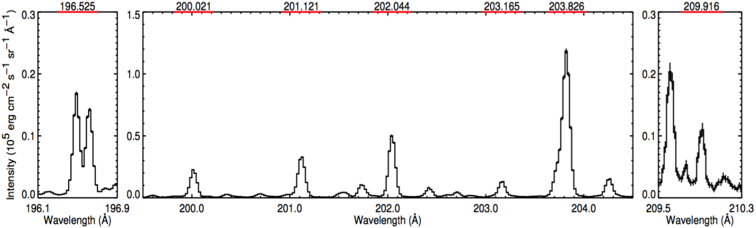


Figure: Example EIS spectrum of seven Fe XIII lines

- Spectral lines with wavelengths $\Lambda = \{\lambda_1, \ldots, \lambda_J\}$
- Observed intensities for K pixels and J wavelengths:

$$\hat{D} = \{D_k = (I_{k\lambda_1}, \ldots, I_{k\lambda_J}), k = 1, \ldots, K\}$$

- Standard deviation $\sigma_{k\lambda_j}$ are also known

Uncertainty: Emissivity

- Emissivity: how strongly energy is radiated at a given wavelength
- Simulated from a model accounting for uncertainty in the atomic data
- Suppose a collection of M emissivity curves are known

$$\mathcal{M} = \{\epsilon_{\lambda}^{(m)}(\mathrm{n}_k, \mathrm{T}_k), \lambda \in \Lambda, m = 1, \ldots, M\}$$

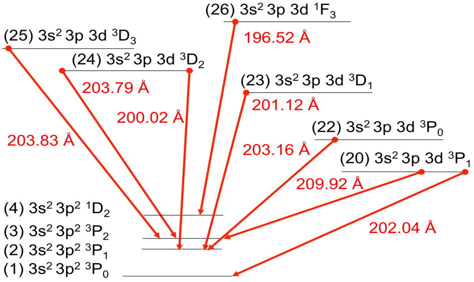


Figure: A simplified level diagram for the transitions relevant to the 7 lines considered here.

Preceding Work

- Fully Bayesian Model I:
  - Use **Bayesian Methods** to incorporate information in the data for narrowing the uncertainty in the atomic physics calculation
  - There are only $M = 1000$ **equally likely** emissivity curves as a priori
- Solution:
  - Obtain a sample of $\mathcal{M}$ that accounts for the uncertainty in the atomic data, m
  - m is treated as an unknown parameter
  - Obtain sample from $p(m, \theta \mid D)$ via **two-step Monte Carlo (MC) samplers** or **Hamiltonian Monte Carlo** (HMC)
- Conclusions:
  - We are able to incorporate uncertainties in atomic physics calculations into analyses of solar spectroscopic data

Preceding Work: Fully Bayesian Model I

Independent prior distributions

$$p(m, \theta_k) = p(m)\ p(\log \mathrm{n}_k)\ p(\log \mathrm{d}_k) \tag{1}$$

$$m \sim \text{DiscreteUniform}(\{1, \ldots, M\}) \tag{2}$$

$$\log_{10} \mathrm{n}_k \sim \text{Uniform}(\text{min} = 7, \text{max} = 12) \tag{3}$$

$$\log_{10} \mathrm{d}_k \sim \text{Cauchy}(\text{center} = 9, \text{scale} = 5) \tag{4}$$

Likeihood $L(m, \theta_k \mid D_k)$

$$I_{k\lambda} \mid m, \mathrm{n}_k, \mathrm{d}_k \overset{\text{indep}}{\sim} \text{Normal}\left(\epsilon_\lambda^{(m)}(\mathrm{n}_k, \mathrm{T}_k)\mathrm{n}_k^2\mathrm{d}_k,\ \sigma_{k\lambda}^2\right), \quad \text{for } \lambda \in \Lambda \tag{5}$$

Joint posterior distribution

$$p(m, \theta_k \mid D_k) \propto L(m, \theta_k \mid D_k)\ p(m, \theta_k), \tag{6}$$

Preceding Work: Pragmatic VS Fully Bayesian Models

Pragmatic Bayesian posterior distribution

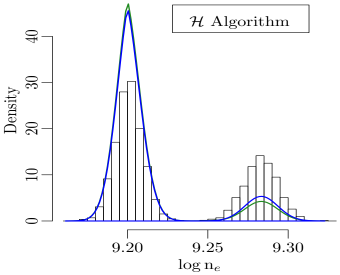
$$p(m, \theta_k \mid D_k) = p(\theta_k \mid D_k, m) \; p(m). \tag{7}$$

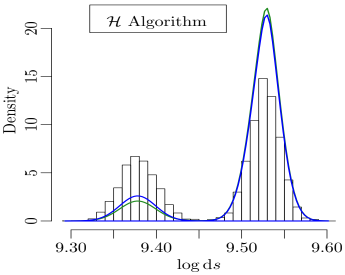
Fully Bayesian posterior distribution

$$p(m, \theta_k \mid D_k) = p(\theta_k \mid D_k, m) \; p(m \mid D_k). \tag{8}$$

Preceding Work: Multimodal Posterior Distributions

- **Bimodal** posterior distributions occur
  - **Two modes** correspond to **two emissivity curves**
  - **Inaccurate relative size** of two modes in HMC
- Reason: **Not enough** emissivity curves
- Challenge: **Sparse** selection of emissivity curves



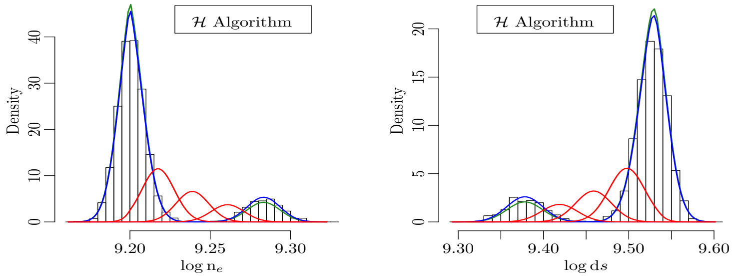


Preceding Work: Multimodal Posterior Distributions

- A computational issue: **Inaccurate relative size** of two modes
- A computational solution:
  Adding a few **synthetic** replicate emissivity curves with augmented set $\mathcal{M}^{\text{aug}} \supset \mathcal{M}$:

$$\mathcal{M}^{\text{aug}}/\mathcal{M} = \{w_1 * \text{Emis}_{471} + w_2 * \text{Emis}_{368}\},$$

where $(w_1, w_2) = (0.75, 0.25), (0.50, 0.50), \&(0.25, 0.75)$



Aim

Come up with a way to efficiently represent

the high dimensional joint distribution

of the uncertainty of the emissivity curves.

Comparison of Current and Preceding Models

Model I (Done)

- Joint posterior distribution:
$$p(m, \theta_k \mid D_k) \propto L(m, \theta_k \mid D_k)\ p(m)\ p(\theta_k)$$
- $p(m) = \frac{1}{M}$
- A computational trick: adding a few **synthetic** replicate emissivity curves

Model II (On going ...)

- Joint posterior distribution:
$$p(C(r_k), \theta_k \mid D_k) \propto L(C(r_k), \theta_k \mid D_k)\ p(r_k, \theta_k)$$
- r_k is the PCA transformation of emissivity curve, C
- $p(r)$ is a high dimensional **distribution**
- An algorithm: summarizing the distribution with multivariate standard Normal distribution via **principal component analysis (PCA)**

Principal Component Analysis

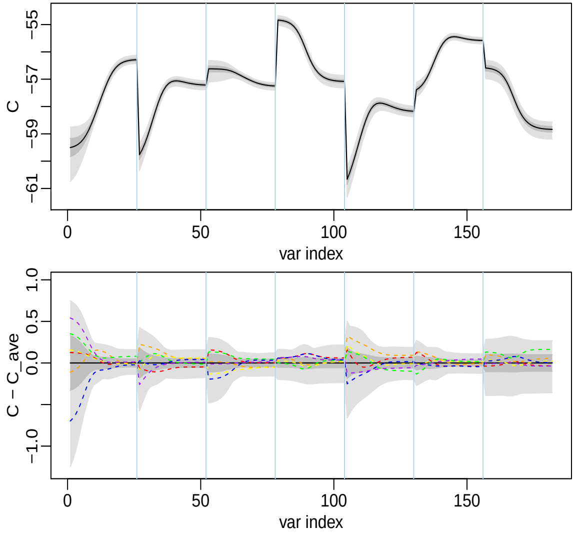
- In log space
- $J = 16$ PCs capture 99% of total variation
- **PCA generated** emissivity curve replicate based on **the first J PCs**

$$C^{\text{rep}} = \bar{C} + \sum_{j=i}^{J} r_j \beta_j v_j,$$

where

- $\bar{C}$: average of all 1000 emissivity curves
- r_j: random variate generated from the standard Normal distribution
- β_j^2, v_j: eigenvalue and eigenvector of component j in the PCA representation

Plot of Original Emissivity Curves



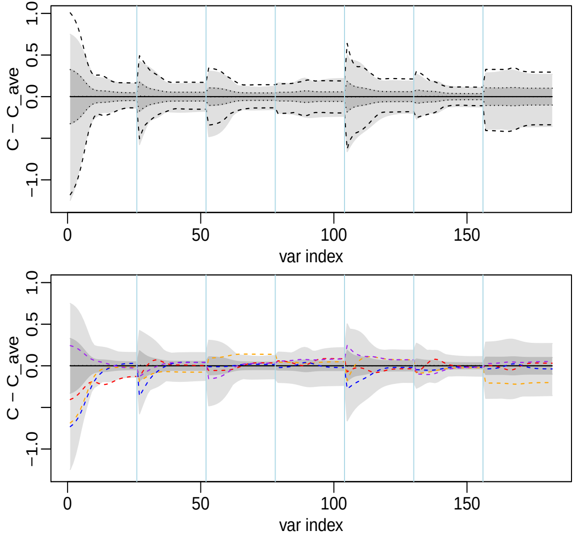
Top panel:

- Each part corresponds one of the seven lines
- Light gray area: all 1000 emissivities
- Dark gray area: middle 68% of emissivities
- Solid black curve: $\bar{C}$

Bottom panel:

- Same as above, but using $C - \bar{C}$
- Colored dashed curves: six randomly selected curves

Plot of PCA Generated Emissivity Curves



Light & dark gray areas are same as above

Top panel:

- Dashed lines:
  all 1500 PCA generated emissivities
- Dotted lines:
  the middle 68% of PCA emissivities

Bottom panel:

- Colored dashed curves:
  selection of PCA curves

Plot of Eigenvectors

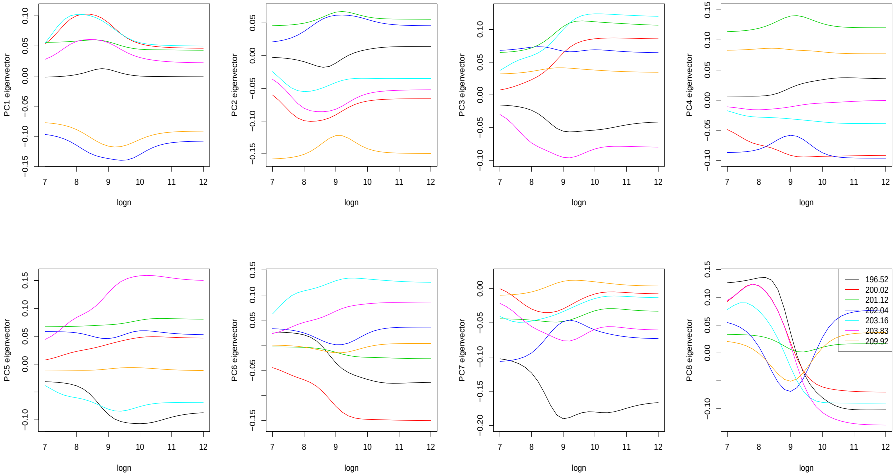


Figure: Plot of eigenvectors of 7 lines along $\log \mathrm{n}$ grid for the first 16 PCs.

Principal Component Analysis

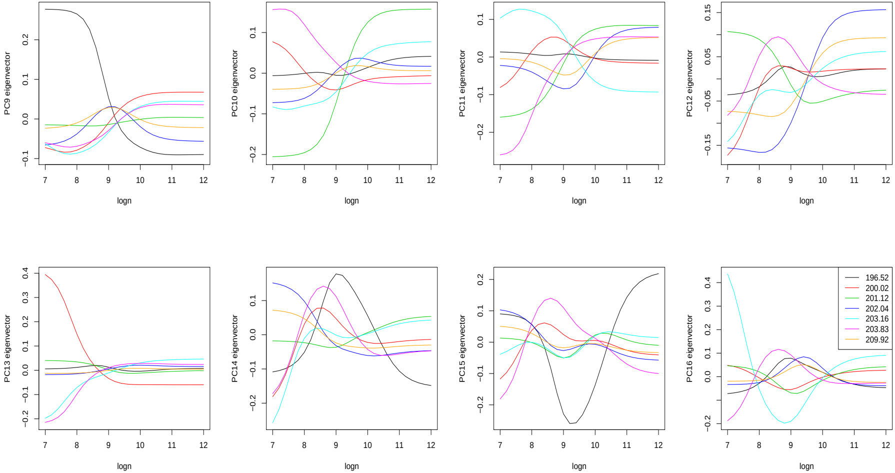


Figure: Plot of eigenvectors of 7 lines along $\log n$ grid for the first 16 PCs.

Alg I: Hamiltonian Monte Carlo via Stan

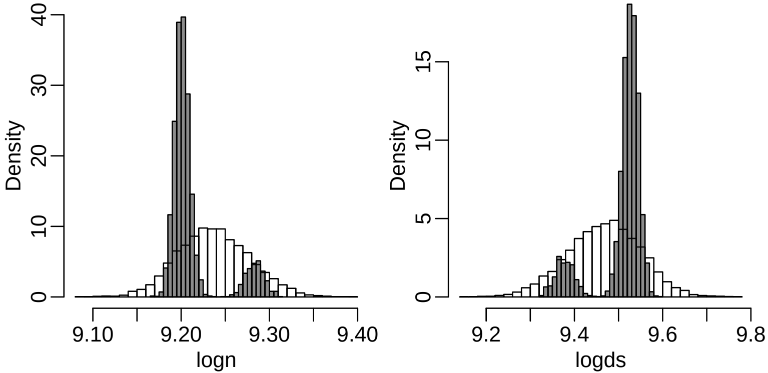
Use **Stan** (`mc-stan.org`) to sample

$$(r^{(l)}, \theta^{(l)}) \sim p(r, \theta \mid D), \text{ for } l = 1, \ldots, L.$$

where

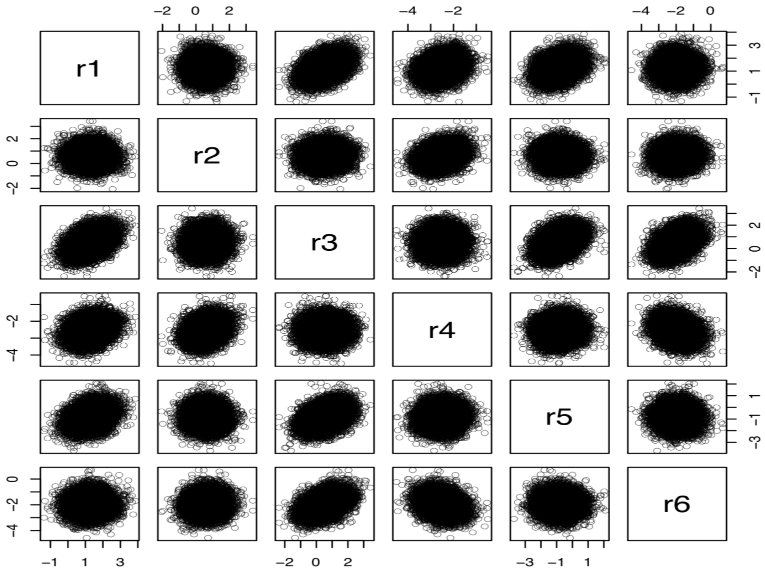
- $p(r, \theta \mid D) \propto p(D \mid r, \theta)\ p(r)\ p(\theta)$
- $p(r) \sim \text{Normal}(0, I)$
- $\log_{10} \mathrm{n} \sim \text{Uniform}(\text{min} = 7, \text{max} = 12)$
- $\log_{10} \mathrm{d} \sim \text{Cauchy}(\text{center} = 9, \text{scale} = 5)$

Alg I: Compare Stan Results From Model I & Model II

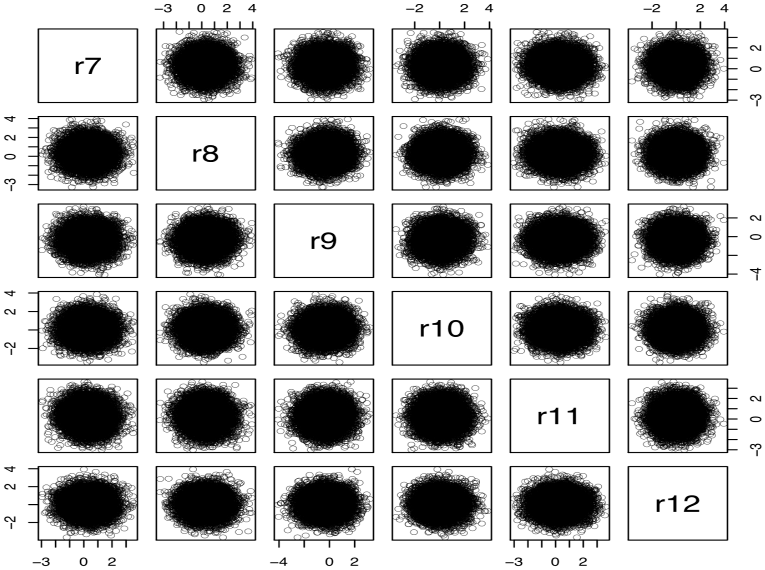


- The results from Model II (white hist) can **mitigate** the multimodal from Model I (gray hist)
- Problem: Stan is **time-consuming**

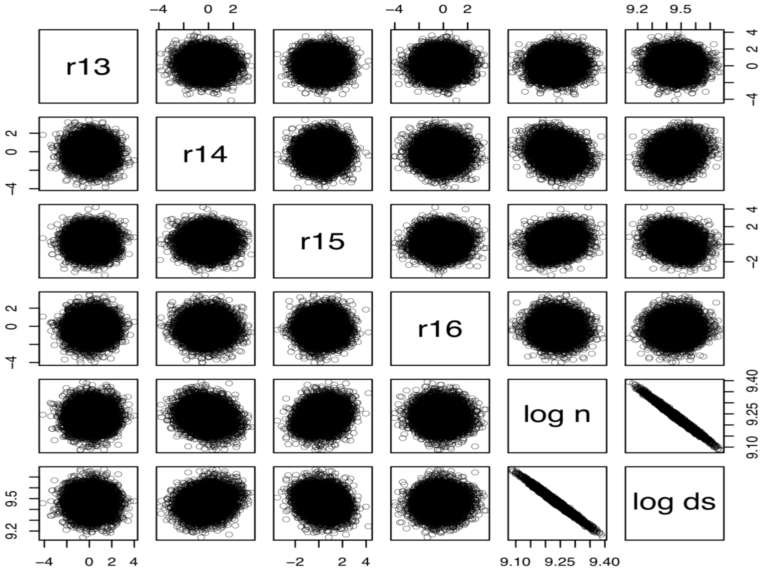
Alg I: Scatterplot matrices from Stan Results



Alg I: Scatterplot matrices from Stan Results



Alg I: Scatterplot matrices from Stan Results



Metropolis Hastings (MH) within Fully Bayesian Gibbs Sampler

Step 1: For $j = 1, \ldots, J$, sample $r_j^{[prop]} \sim \mathcal{N}(\mu = 0, \text{sd} = 1)$.

Step 2: Set $r^{[prop]} = (r_1^{[prop]}, \ldots, r_J^{[prop]})$ and
$C^{[\text{prop}]} = \bar{C} + \sum_{j=1}^{J} r_j^{[t+1]} \beta_j \, v_j$.

Step 3: Sample $\theta^{[\text{prop}]} \sim Q_{\mathcal{MVN}}(\theta | r^{[\text{prop}]})$.

Step 4: Compute

$$\rho = \frac{p(C^{[\text{prop}]}, \theta^{[\text{prop}]} | D) \; Q_{\mathcal{MVN}}(\theta^{(t)}, r^{(t)})}{p(C^{(t)}, \theta^{(t)} | D) \; Q_{\mathcal{MVN}}(\theta^{[\text{prop}]}, r^{[\text{prop}]})} \tag{9}$$

Alg II: Independence Sampler

Use a multivariate normal distribution $Q_{\mathcal{MVN}}(\theta, r)$ to fit a sample obtained from the **pragmatic** Bayesian posterior.

- Suppose we have the pragmatic Bayesian samples, $\{(r_{pB}^{(1)}, \theta_{pB}^{(1)}), \ldots, (r_{pB}^{(T)}, \theta_{pB}^{(T)})\}$
- Set the multivariate normal distribution $Q_{\mathcal{MVN}}(\theta, r)$ to be,

$$\begin{pmatrix} \theta \\ r \end{pmatrix} \sim \text{Normal}\left(\begin{pmatrix} \hat{\mu}_\theta \\ \mu_r \end{pmatrix}, \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \Sigma_{22} \end{pmatrix}\right)$$

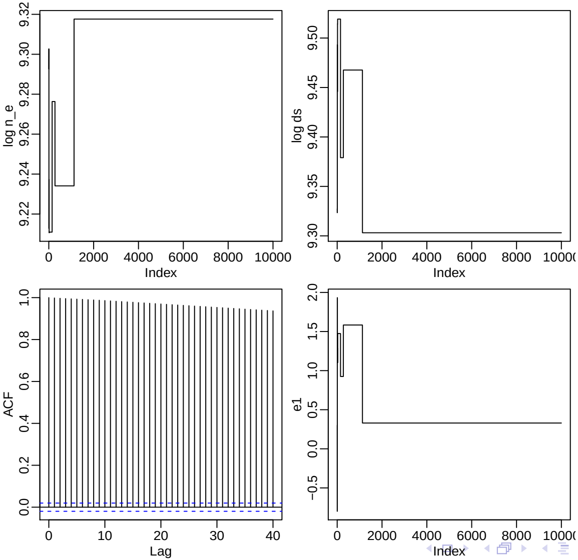
- $r \sim \text{Normal}(\mu_r = 0, \Sigma_{22} = I)$
- $\hat{\mu}_\theta$, $\hat{\Sigma}_{11}$, $\hat{\Sigma}_{12}$: sample mean and sample variance of θ obtained from the pragmatic Bayesian sample, sample covariance between θ and r
- The conditional distribution of θ given r is,
$$Q_{\mathcal{MVN}}(\theta \mid r) \sim \text{Normal}(\hat{\mu}_\theta + \hat{\Sigma}_{12} r, \hat{\Sigma}_{11} - \hat{\Sigma}_{12}\hat{\Sigma}_{21})$$
achieved via **multivariate normal linear regression**.

Alg II: Several Possible Proposals of r

- Normal$(0, I)$
- Normal$(r^{(471)}, I)$ or Normal$(r^{(368)}, I)$
- Mixture distribution:
  $$p \text{ Normal}(r^{(471)}, \sigma^2 I) + (1 - p) \text{ Normal}(r^{(368)}, \sigma^2 I)$$
  where p and σ are tuning parameters
- Normal$(\mu_{Stan}, \Sigma_{Stan})$ or Normal$(\mu_{Stan}, \Sigma_{Stan} * 2)$
- Student t_4

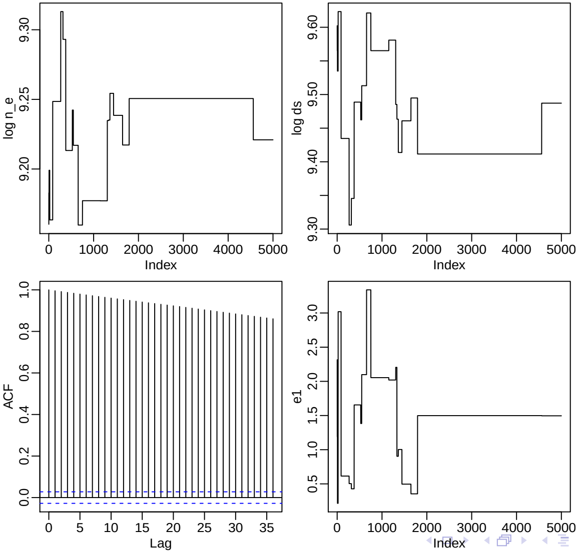
Alg II: Proposals of r

$r \sim \text{Normal}(0, I)$



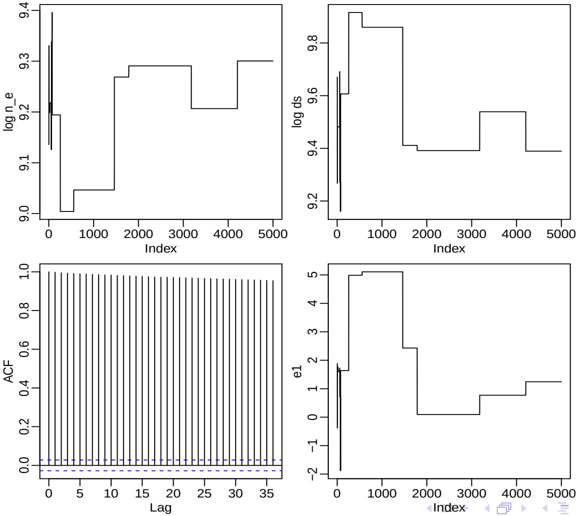
Alg II: Proposals of r

$r \sim \text{Normal}(r^{(471)}, I)$ or $\text{Normal}(r^{(368)}, I)$



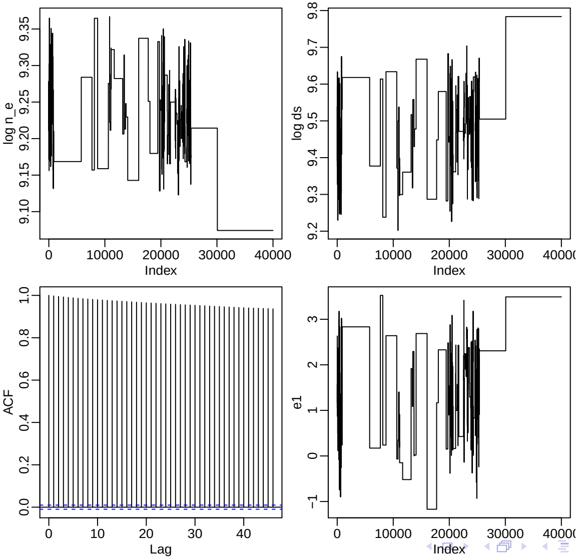
Alg II: Proposals of r

$r \sim p \text{ Normal}(r^{(471)}, \sigma^2 I) + (1-p) \text{ Normal}(r^{(368)}, \sigma^2 I)$
where $p = 0.5$ and $\sigma = 2$ are tuning parameters



Alg II: Proposals of r

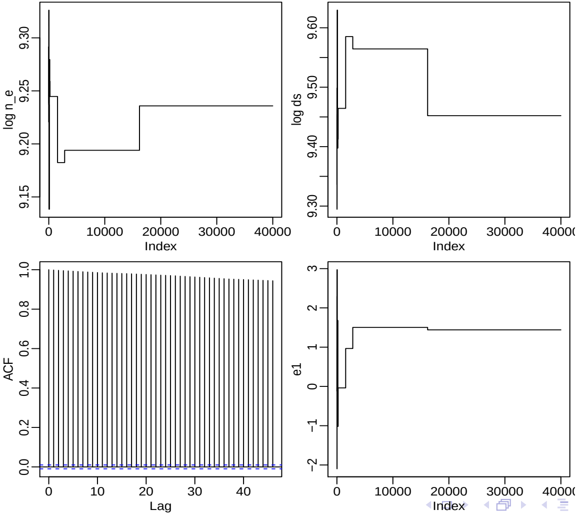
$r \sim \text{Normal}(\mu_{Stan}, \Sigma_{Stan})$ or $\text{Normal}(\mu_{Stan}, \Sigma_{Stan} * 2)$



Alg II: Proposals of r

$r \sim$ Student t_4

$\theta | r \sim$ multivariate t linear regression with $df = 4$



Alg III: Random Walk Sampler (TBD)

Use Random Walk

Step 1: For $j = 1, \ldots, J$, sample $r_j^{[prop]} \sim \mathcal{N}(\mu = r_j^{[t]}, \text{sd} = \sigma_r)$.

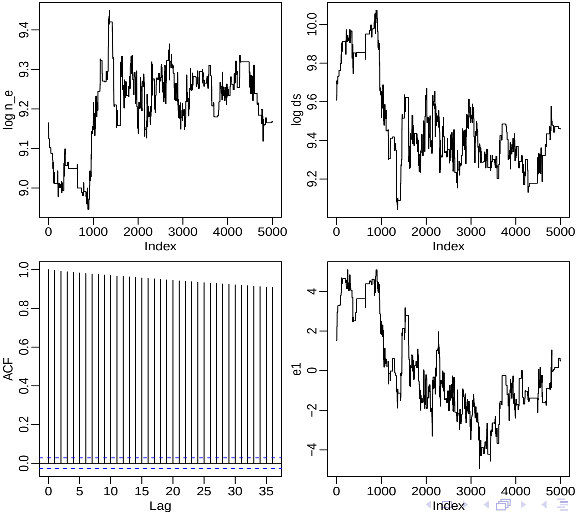
Step 2: Set $r^{[prop]} = (r_1^{[prop]}, \ldots, r_J^{[prop]})$ and
$C^{[\text{prop}]} = \bar{C} + \sum_{j=1}^{J} r_j^{[t+1]} \; \beta_j \; v_j$.

Step 3: Sample $\theta^{[\text{prop}]} \sim Q_{\mathcal{MVN}}(\theta | r^{[\text{prop}]})$.

Step 4: Compute

$$\rho = \frac{p(C^{[\text{prop}]}, \theta^{[\text{prop}]} | D) \; Q_{\mathcal{MVN}}(\theta^{(t)} \mid r^{(t)})}{p = (C^{(t)}, \theta^{(t)} | D) \; Q_{\mathcal{MVN}}(\theta^{[\text{prop}]} \mid r^{[\text{prop}]})} \tag{10}$$

Tuning parameter $\sigma = 0.6$
r is initialized from Normal$(r^{(471)}, I)$



Alg IV: Mixture of Independence Sampler and Random Walk Sampler (TBD)

Step 1: Sample $u_1 \sim \text{Uniform}(0, 1)$

Step 2: If $u_1 < p_m$, go to Alg III, Random Walk Sampler
else, go to Alg II, Independence Sampler

- Two tuning parameters p_m and σ_r