New Results of Fully Bayesian

JIN XU

UCI

April 3, 2012

Background

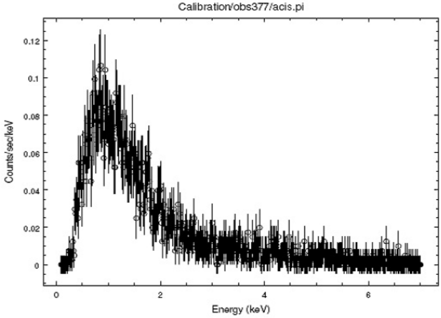
- High-Energy Astrophysics
- Spectral Analysis
- Calibration Products
- Scientific Goals

High-Energy Astrophysics

- Provide understanding into high-energy regions of the Universe.
- Chandra X-ray Observatory is designed to observe X-rays from high-energy regions of the Universe.
- X-ray detectors typically count a small number of photons in each of a large number of pixels.
- Spectral Analysis aims to explore the parameterized pattern between the photon counts and energy.

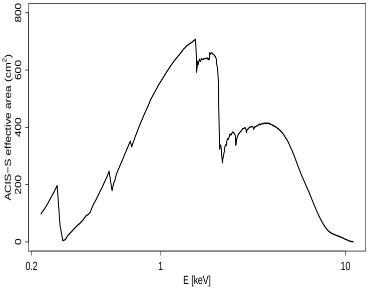
An Example of One Dataset

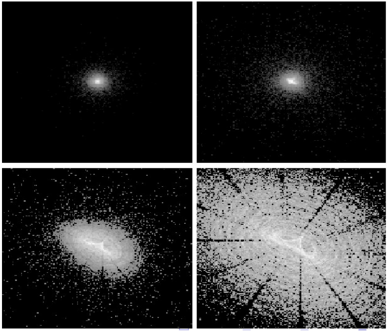
TITLE = EXTENDED EMISSION AROUND A GIGAHERTZ PEAKED RADIO SOURCE
DATE = 2006-12-29 T 16:10:48



Calibration Uncertainty

- Effective area records sensitivity as a function of energy.
- Energy redistribution matrix can vary with energy/location.
- Point Spread Functions can vary with energy and location.





Incorporate Calibration Uncertainty

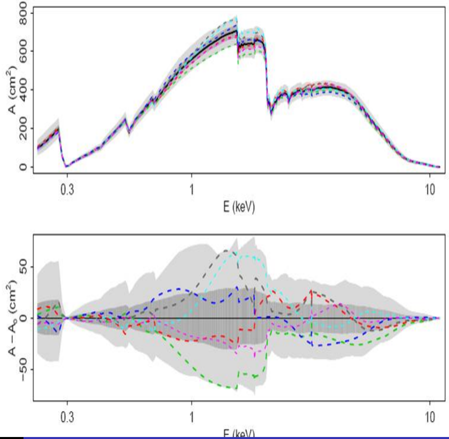
- Calibration Uncertainty in astronomical analysis have been generally ignored.
- No robust principled method is available.
- Our goal is to incorporate the uncertainty by Bayesian Methods.
- In this talk, we focus on uncertainty in the effective area.

Two Main Problems

- The true effective area curve can't be observed, when we try to incorporate calibration uncertainty in estimating source parameters.
- We don't have parameterized form for effective area curve. It makes sampling hard to approach.

Generating Calibration Samples

- Drake et al. (2006), suggests to generate calibration samples of effective area curves to represent the uncertainty.
- Calibration Samples: $\{A_1, A_2, A_3, ..., A_L\}$



Three Main Steps

- Use Principle Component Analysis to parameterize effective area curve.
- Model Building, that it combining source model with calibration uncertainty.
- Three source parameter sampling schemes.

Use PCA to represent effective area curve

$$A = A_0 + \bar{\delta} + \sum_{j=1}^{m} e_j r_j v_j$$

A_0 : default effective area,

$\bar{\delta}$: mean deviation from A_0,

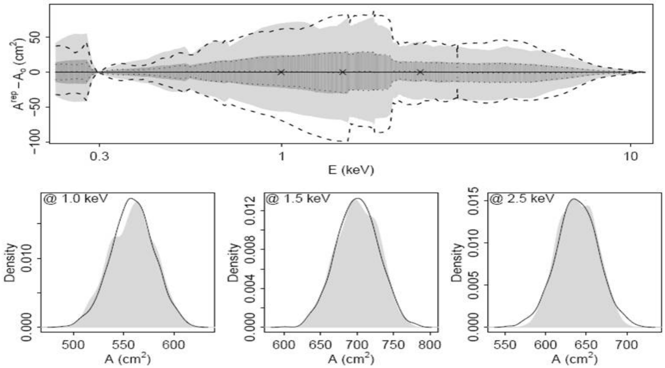
r_j and v_j : first m principle component eigenvalues & vectors,

e_j : independent standard normal deviations.

Capture 95% of uncertainty with m = 6 - 9.

Use PCA to represent effective area curve

PCA method has nicely parameterized effective area curve.



A simplified model of telescope response, only concerning effective area uncertainty

$$M(E;\theta) = S(E;\theta) * A(E)$$

$M(E;\theta)$: Observed Photon Distribution,

$S(E;\theta)$: True Source Model, we set it as poisson distribution with expectation equal to $exp(-n_H * sigma(E)) * Amp * E^{(-gamma)} + bkg$

$A(E)$: Effective Area Curve.

θ: source parameter, $\theta = \{n_H, Amp, gamma, bkg\}$

Scheme One: Fixed Effective Area Curved

- We assume $A = A_0$, where A_0 is the default affective area curve, and may not be the true one,
- This scheme doesn't incorporate any calibration uncertainty,
- The estimation may be biased and error bars may be underestimated.
- Only one sampling step involved: $p(\theta|M, A_0) \propto L(M|\theta, A_0)p(A_0)$

Scheme Two: Pragmatic Bayesian, Lee et al(2011, Apj)

- Main purpose is to reduce complexity of sampling.
- This scheme "completely" incorporates the calibration uncertainty,
- Step One: sample A from $p(A)$
- Step Two: sample θ from $p(\theta|M, A) \propto L(M|\theta, A)p(\theta)$

Scheme Three: Fully Bayesian

- Use correct Bayesian Approach,
- This scheme concerns about letting the current data influence calibration products,
- Step One: sample A from $p(A|M,\theta) \propto L(M|\theta,A)p(A)$
- Step Two: sample θ from $p(\theta|M,A) \propto L(M|\theta,A)p(\theta)$
- Most difficult approach to sample.

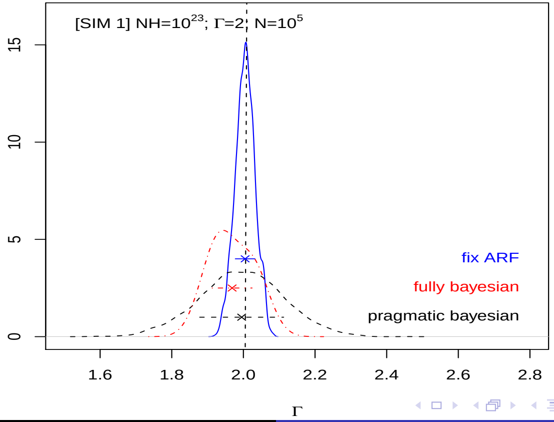
Eight simulated data sets

The first four data sets were all simulated without background contamination using the XSPEC model wabs*powerlaw, nominal default effective area A_0 from the calibration sample of Drake et al. (2006), and a default RMF for ACIS-S.

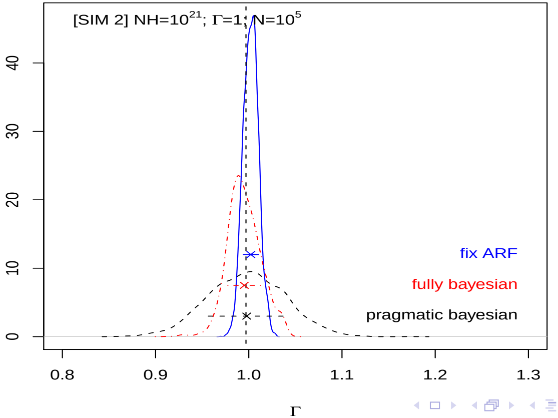
- Simulation 1: $\Gamma = 2, N_H = 2^{23} cm^{-2}$, and 10^5 counts;
- Simulation 2: $\Gamma = 1, N_H = 2^{21} cm^{-2}$, and 10^5 counts;
- Simulation 3: $\Gamma = 2, N_H = 2^{23} cm^{-2}$, and 10^4 counts;
- Simulation 4: $\Gamma = 1, N_H = 2^{21} cm^{-2}$, and 10^4 counts;

The other four data sets (Simulation 5-8) were generated using an extreme instance of an effective area.

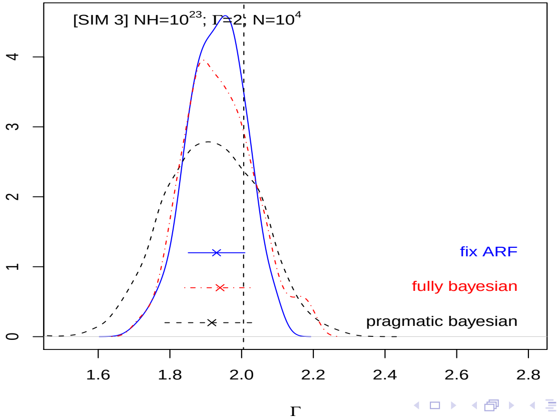
Results for Simulation 1



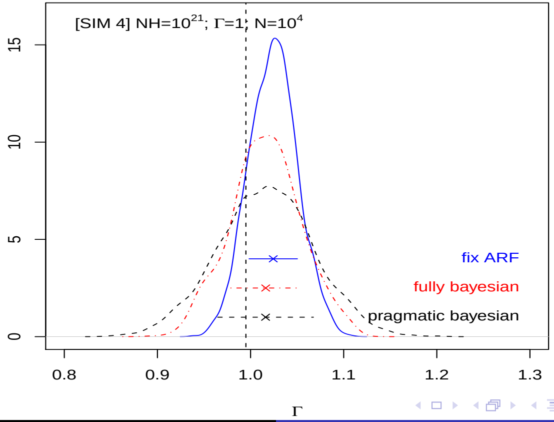
Results for Simulation 2



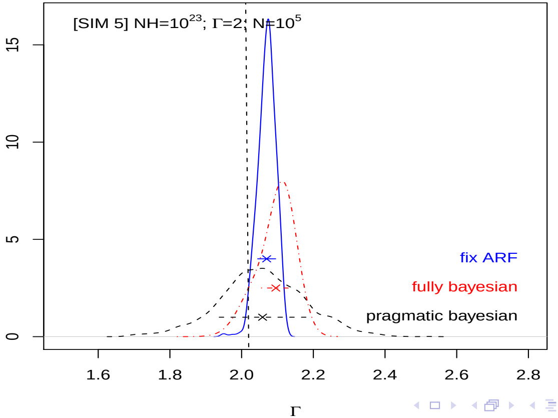
Results for Simulation 3



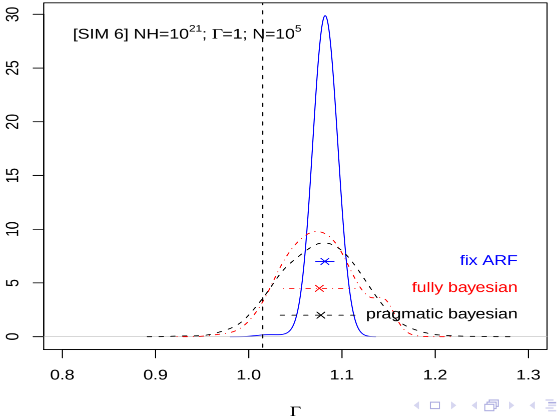
Results for Simulation 4



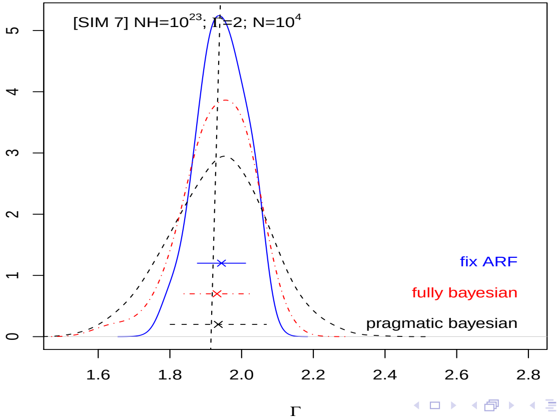
Results for Simulation 5



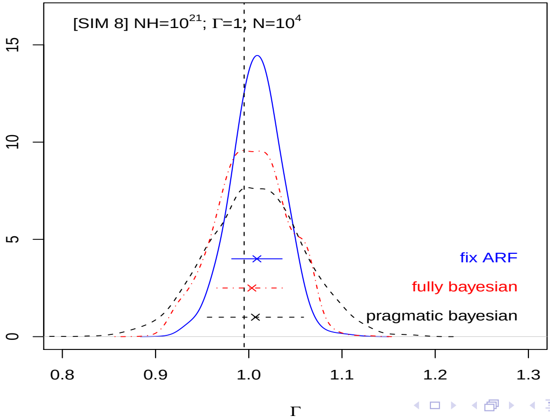
Results for Simulation 6



Results for Simulation 7



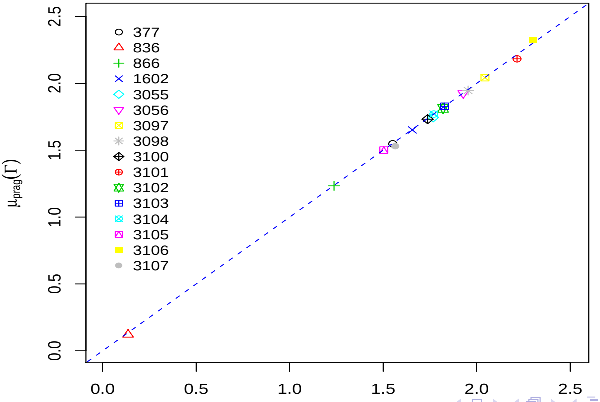
Results for Simulation 8



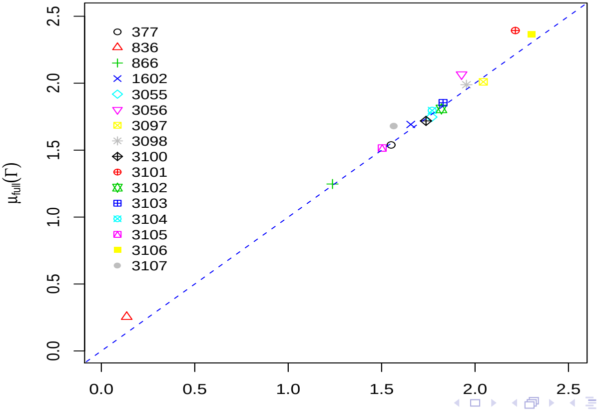
Quasar results

- 16 Quasar data sets were fit by these three models: 377, 836, 866, 1602, 3055, 3056, 3097, 3098, 3100, 3101, 3102, 3103, 3104, 3105, 3106, 3107.
- Most interesting founding for fully bayesian model is shift of parameter fitting, besides the change of standard errors.
- Both comparisons of mean and standard errors among three models are shown below.

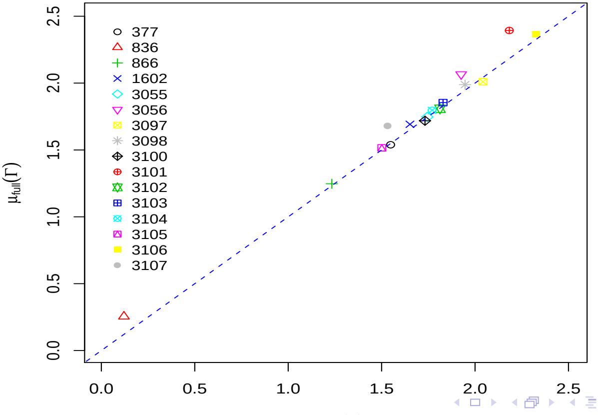
mean: fix-prag



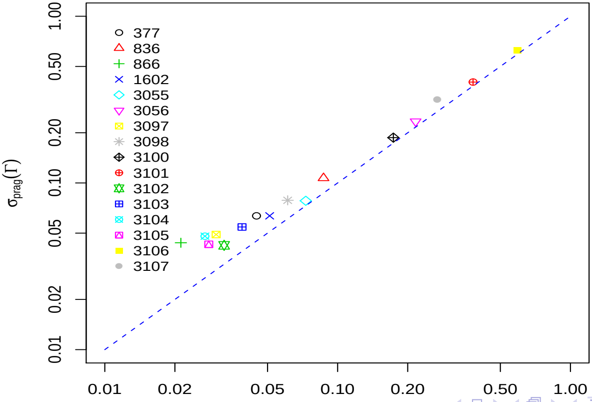
mean: fix-full



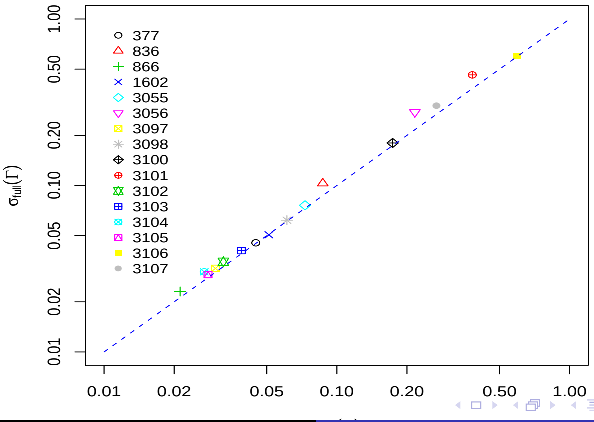
mean: prag-full



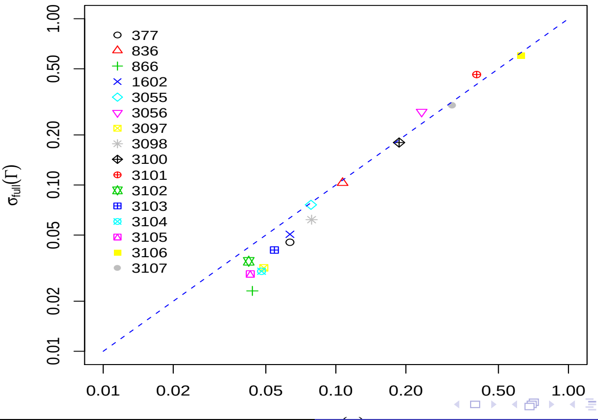
sd: fix-prag


377
836
866
1602
3055
3056
3097
3098
3100
3101
3102
3103
3104
3105
3106
3107
$\sigma_{prag}(\Gamma)$
0.01
0.02
0.05
0.10
0.20
0.50
1.00

sd: fix-full

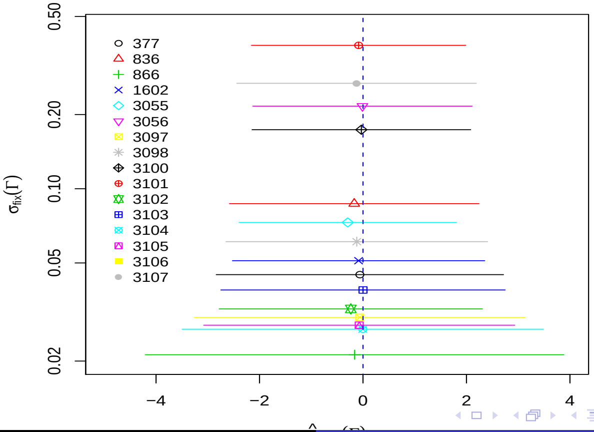

$\sigma_{full}(\Gamma)$
1.00
0.50
0.20
0.10
0.05
0.02
0.01
0.01
0.02
0.05
0.10
0.20
0.50
1.00
377
836
866
1602
3055
3056
3097
3098
3100
3101
3102
3103
3104
3105
3106
3107

sd: prag-full


377
836
866
1602
3055
3056
3097
3098
3100
3101
3102
3103
3104
3105
3106
3107
$\sigma_{full}(\Gamma)$
0.01
0.02
0.05
0.10
0.20
0.50
1.00

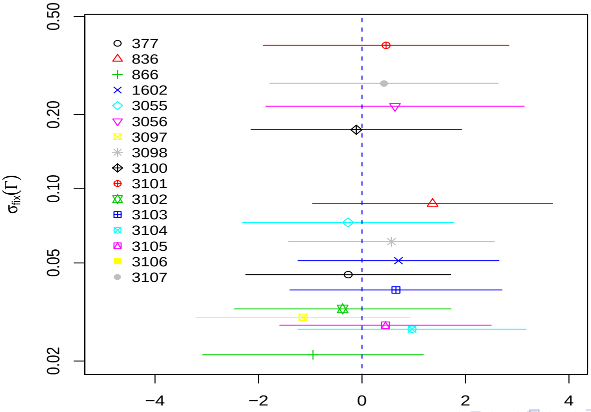
more plots

$\hat{\mu}_{prag}(\Gamma) = \frac{\mu_{prag}(\Gamma) - \mu_{fix}(\Gamma)}{\sigma_{fix}(\Gamma)}$, these lines cover 2 sd.



more plots

$\hat{\mu}_{full}(\Gamma) = \frac{\mu_{full}(\Gamma) - \mu_{fix}(\Gamma)}{\sigma_{fix}(\Gamma)}$, these lines cover 2 sd.



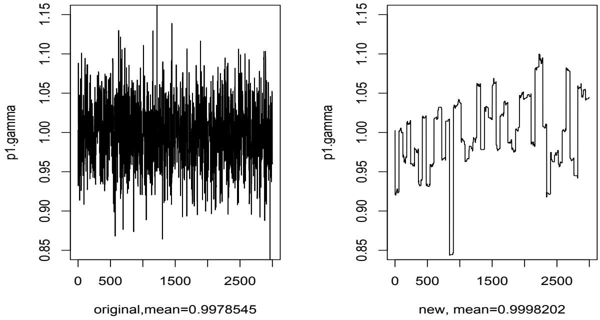
New approach for Pragmatic Bayesian

Repeat 50 times

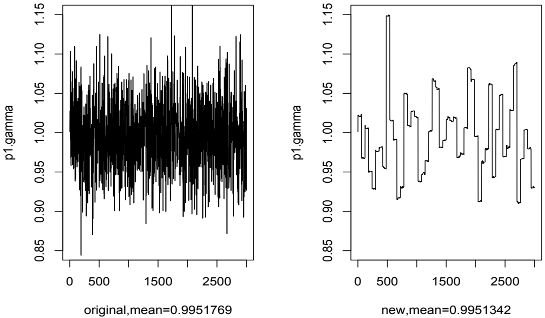
- Draw A
- run 10 inner iterations of PyBlocks, THROWING OUT the draws of theta
- run an additional L/50 iterations of PyBlocks, KEEPING the draws of theta.

This approach speeds the sampling up a lot, since it only needs 50 sherpa fit(). eg, 3000 iterations, old approach needs 3 hours, new approach needs less than 15min

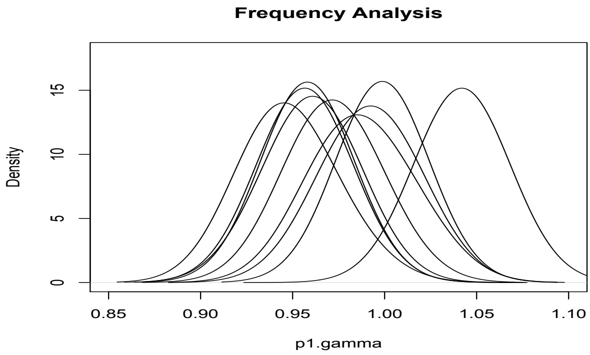
Dataset 0000-1-21-e4



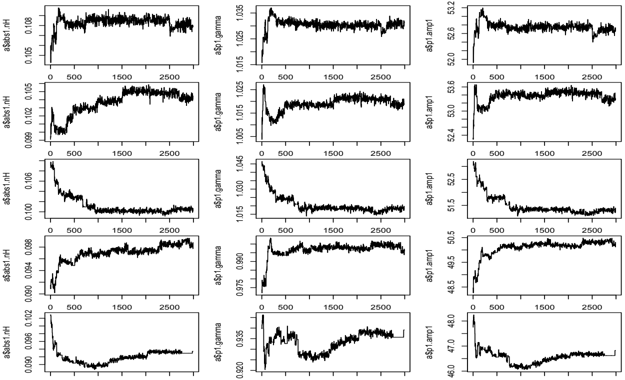
Dataset 0000-1-21-5e4



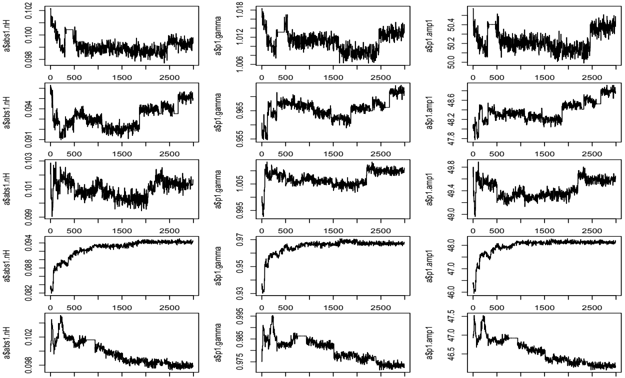
10 simulated Datasets 0000-1-21-e4



10 simulated Datasets 0000-1-21-5e4



10 simulated Datasets 0000-1-21-5e4



Scheme

- Introduce posterior distribution of pragmatic sampling as the biasing density
- Get the draws from Pragmatic Bayesian Method
- Calculate the ratio $r = P_{fully}(A, \theta | data)/P_{prag}(A, \theta | data)$

Now, I still have some problem to calculate R. Hope this method can solve all fully bayesian problems.

Use PCA to represent effective area curve

$$A = A_0 + \bar{\delta} + \sum_{j=1}^{m} e_j r_j v_j$$

A_0 : default effective area,

$\bar{\delta}$: mean deviation from A_0,

r_j and v_j : first m principle component eigenvalues & vectors,

e_j : independent standard normal deviations.

This approach not only reduces the dimension of ARF, but also provides a nice and convenient way to sample ARF.

SCA or Wavelets

- Usually, Spectral Clustering Analysis is only used for clustering. First step is nonlinear dimension reduction, Then use K-mean algorithm to do clustering.
- If we apply this method to samples of ARF, we can get K clusters of ARF. The main problem here is how to take the advantage of clustering to sample new ARF.
- Wavelet transformation is usually used to find out the hidden pattern inside the signal. If we use wavelets to analyze ARF, we can get a lot of parameters, making a summary of these parameters and then sampling back is still the main problem.