

Gravitational Wave Data Analysis for PTAs

Paul T. Baker

Center for Gravitational Waves and Cosmology
West Virginia University

29 June 2017



Our Likelihood is an N_{toa} D gaussian

$$\mathcal{L} \propto \exp \left(-\frac{1}{2} \vec{r} \cdot \mathbb{C}^{-1} \cdot \vec{r} \right)$$

- \vec{r} : residuals (data - model)
- \mathbb{C} : covariance matrix (noise)

likelihood is probability that our data is observed,
if our model is correct

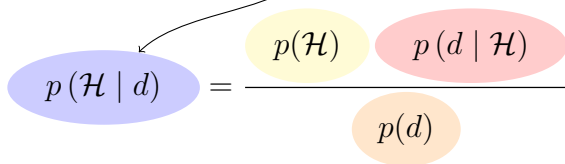
- $\max[p(d | m(\lambda))] \rightarrow$
model most likely to have produced our data

What is the probability that this (or any) model is correct?

Bayes Theorem

Probability hypothesis \mathcal{H} is true given you observed data d :

- **Posterior probability, \mathcal{P}**



The diagram illustrates Bayes' Theorem with the following components:

- A blue oval containing the posterior probability $p(\mathcal{H} | d)$.
- An equals sign.
- A fraction where the numerator consists of a yellow oval $p(\mathcal{H})$ and a pink oval $p(d | \mathcal{H})$.
- A denominator consisting of an orange oval $p(d)$.
- A curved arrow pointing from the text "Posterior probability, \mathcal{P} " to the blue oval.

$$p(\mathcal{H} | d) = \frac{p(\mathcal{H}) p(d | \mathcal{H})}{p(d)}$$

Bayes Theorem

Probability hypothesis \mathcal{H} is true given you observed data d :

- Posterior probability, \mathcal{P}

The diagram illustrates Bayes' Theorem. On the left, the posterior probability $p(\mathcal{H} | d)$ is enclosed in a blue oval. An arrow points from this oval to the text 'Posterior probability, \mathcal{P} ' above. To the right, the equation $p(\mathcal{H} | d) = \frac{p(\mathcal{H}) p(d | \mathcal{H})}{p(d)}$ is shown. The numerator terms $p(\mathcal{H})$ and $p(d | \mathcal{H})$ are in yellow and pink ovals, respectively. The denominator term $p(d)$ is in an orange oval. An arrow points from the text 'Prior probability p ' below to the $p(\mathcal{H})$ term in the numerator.

$$p(\mathcal{H} | d) = \frac{p(\mathcal{H}) p(d | \mathcal{H})}{p(d)}$$

- **Prior probability** p

Bayes Theorem

Probability hypothesis \mathcal{H} is true given you observed data d :

- Posterior probability, \mathcal{P}

The diagram illustrates Bayes' Theorem with the following components:

- A blue oval on the left contains the expression $p(\mathcal{H} | d)$.
- An equals sign follows.
- A horizontal line is drawn across the middle.
- Above the line, a yellow oval contains $p(\mathcal{H})$ and a red oval contains $p(d | \mathcal{H})$.
- Below the line, an orange oval contains $p(d)$.
- Arrows point from the text labels below to their respective terms in the diagram.

$$p(\mathcal{H} | d) = \frac{p(\mathcal{H}) p(d | \mathcal{H})}{p(d)}$$

- Prior probability p
- **Likelihood**, \mathcal{L}

Bayes Theorem

Probability hypothesis \mathcal{H} is true given you observed data d :

- Posterior probability, \mathcal{P}

The diagram illustrates Bayes' Theorem with the following components:

- A blue oval on the left contains the posterior probability $p(\mathcal{H} | d)$. An arrow points from the text "Posterior probability, \mathcal{P} " to this oval.
- An equals sign follows the posterior probability.
- A horizontal line represents the denominator, with an orange oval below it containing the evidence $p(d)$. An arrow points from the text "Evidence (marginal likelihood), \mathcal{E} " to this oval.
- Two ovals are positioned above the horizontal line: a yellow one on the left containing the prior probability $p(\mathcal{H})$, and a red one on the right containing the likelihood $p(d | \mathcal{H})$.
- An arrow points from the text "Prior probability p " to the yellow oval.
- An arrow points from the text "Likelihood, \mathcal{L} " to the red oval.

$$p(\mathcal{H} | d) = \frac{p(\mathcal{H}) p(d | \mathcal{H})}{p(d)}$$

- Prior probability p
- Likelihood, \mathcal{L}
- **Evidence (marginal likelihood), \mathcal{E}**

Need a way to compute Posterior distribution without computing Evidence

Sample from unnormalized distribution:

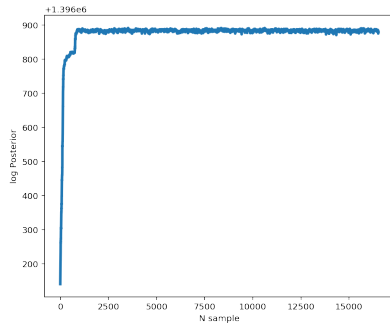
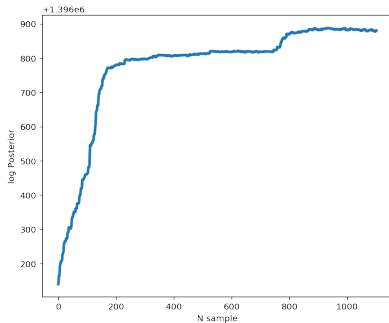
- Markov chain Monte Carlo
(e.g. emcee, kombine, PTMCMCSampler)

- Nested Sampling (e.g. MultiNest)

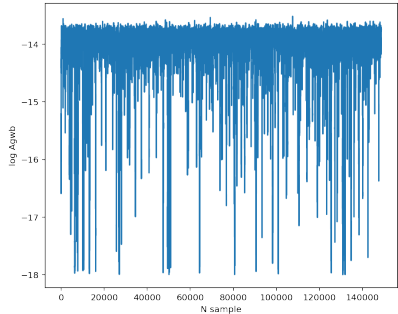
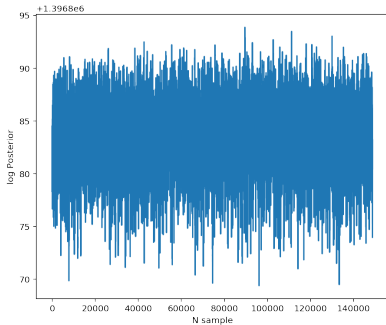
histogram of N samples \rightarrow Posterior as $N \rightarrow \infty$

inference with samples

in practice first samples are biased – **burn in**



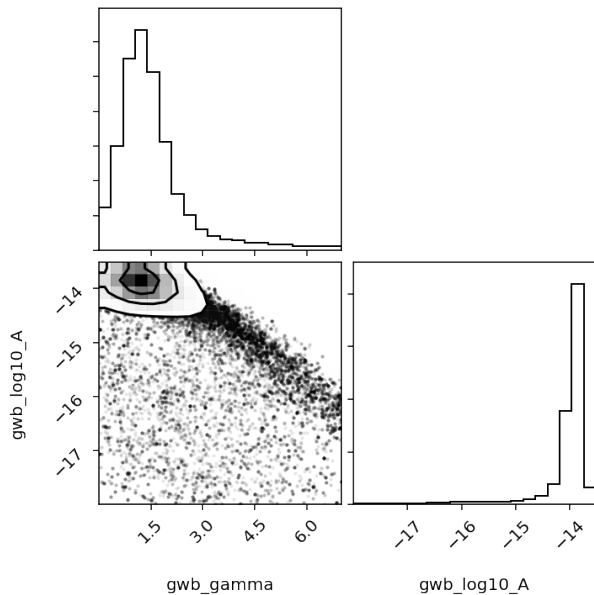
a properly converged chain of samples



Marginalization – posterior for subset of parameters

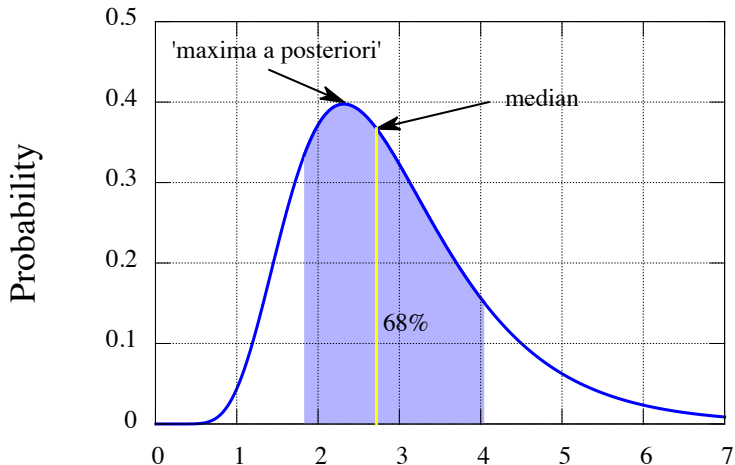
$$\mathcal{P}(x) = \int \mathcal{P}(x, y, z) dy dz$$

marginalization



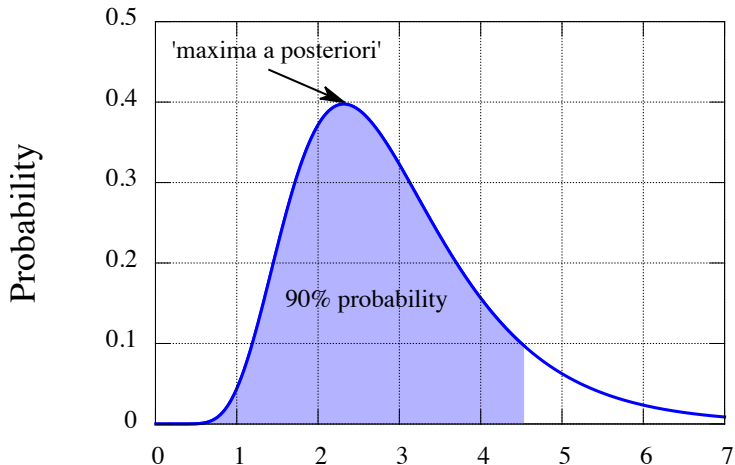
parameter estimation

posterior density function

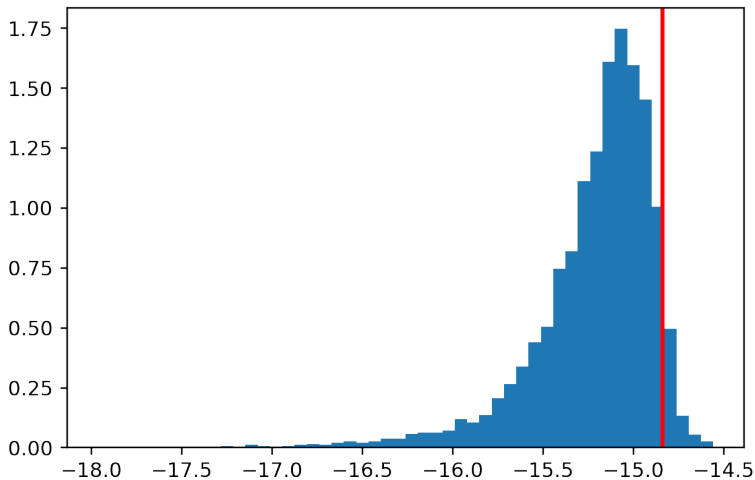


Upper Limits

posterior density function



PTA upper limit on A_{gw} – first marginalize over all other params, then integrate to 95%



$$\mathcal{E}_1 = \int \mathcal{L}_1(x, y, z) \cdot p(x, y, z) dx dy dz$$

$$\mathcal{E}_2 = \int \mathcal{L}_2(x, y) \cdot p(x, y) dx dy$$

$$\mathcal{B}_{1,2} = \frac{\mathcal{E}_1}{\mathcal{E}_2}$$

Bayes factor is betting odds.

model selection – Bayes factor

$$\mathcal{B}_{1,2} = \frac{\mathcal{E}_1}{\mathcal{E}_2}$$

\mathcal{B}	“sigma”	
2	1σ	no evidence
20	2σ	weak evidence
2×10^6	5σ	very strong evidence

(sometimes reported as $\log \mathcal{B}$)

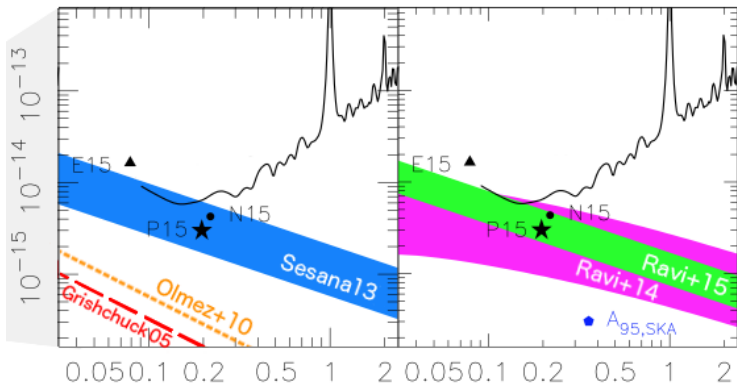
Currently no GW detection from PTAs

95 % upper limit for stochastic GW background:

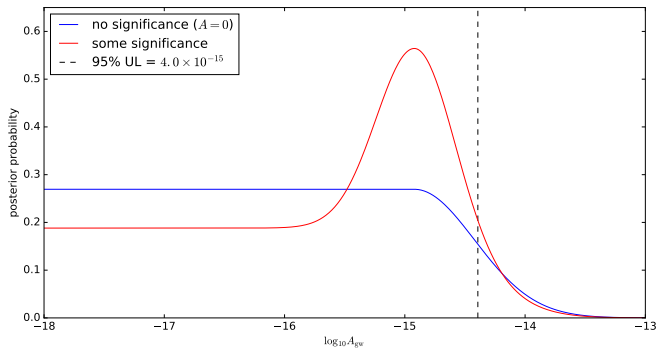
EPTA	3.0×10^{-15}	Lentati + (2015)
PPTA	1.0×10^{-15}	Shannon + (2015)
NANOGrav	1.5×10^{-15}	Arzoumanian + (2015)
IPTA	1.7×10^{-15}	Verbiest + (2016)

current ULs are ephemeris dependent

some current PTA results



Detection of GWB will be a gradual process



Detection claim must compare multiple models

How are you sure it's GWB
and not some other source of red noise?

- common red process
- correlated red process
 - quadrupole (Hellings-Downs)
 - dipole (ephemeris)
 - monopole (clocks)
- different spectral indices