NANOGrav Data Analysis Tutorial

Jerry Sun
6/24/24

Goals for today

1. Understand the fundamental concepts behind using pulsar timeseries to look for gravitational waves (GWs)

2. Understand how Bayes' theorem can be applied to compute probability distributions for GW signal parameters

3. Analyze a pulsar timing array (PTA) data set for a simple GW signal using the state-of-the-art analysis software (PTMCMCSampler, ENTERPRISE)

The Idea

- Millisecond pulsars are *extremely stable* rotating objects

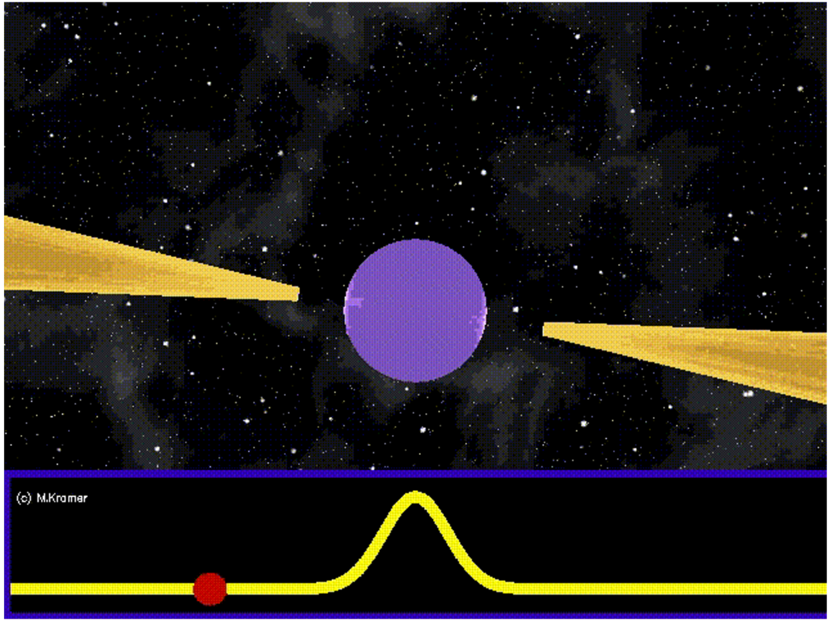


Image credit: Michael Kramer

The Idea

- Millisecond pulsars are *extremely stable* rotating objects
- GWs passing through the pulsar-Earth line of sight change the distance between the pulsar and Earth

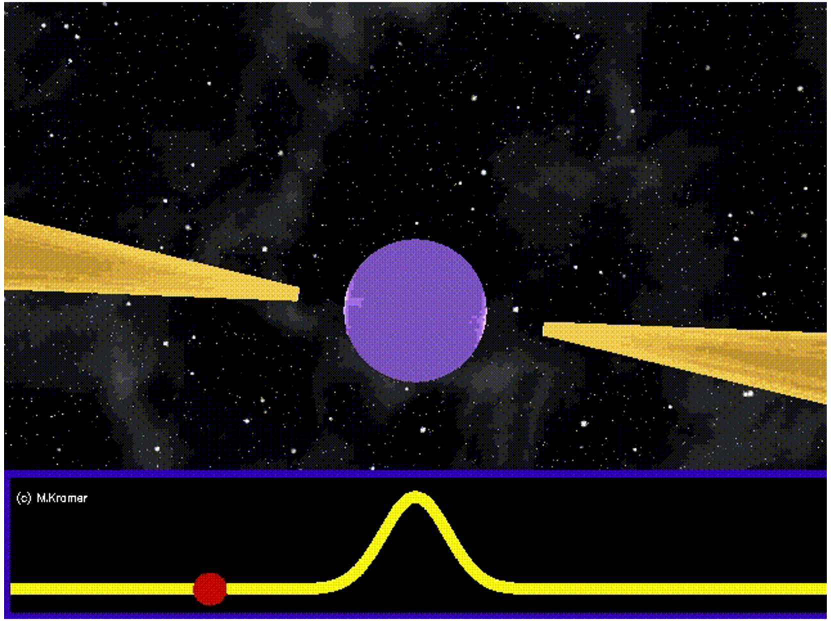


Image credit: Michael Kramer

The Idea

- Millisecond pulsars are *extremely stable* rotating objects
- GWs passing through the pulsar-Earth line of sight change the distance between the pulsar and Earth

This causes the photon arrive a little too soon or a little too late (the **timing residual**)!

Pulsars are stable enough that we can detect this.

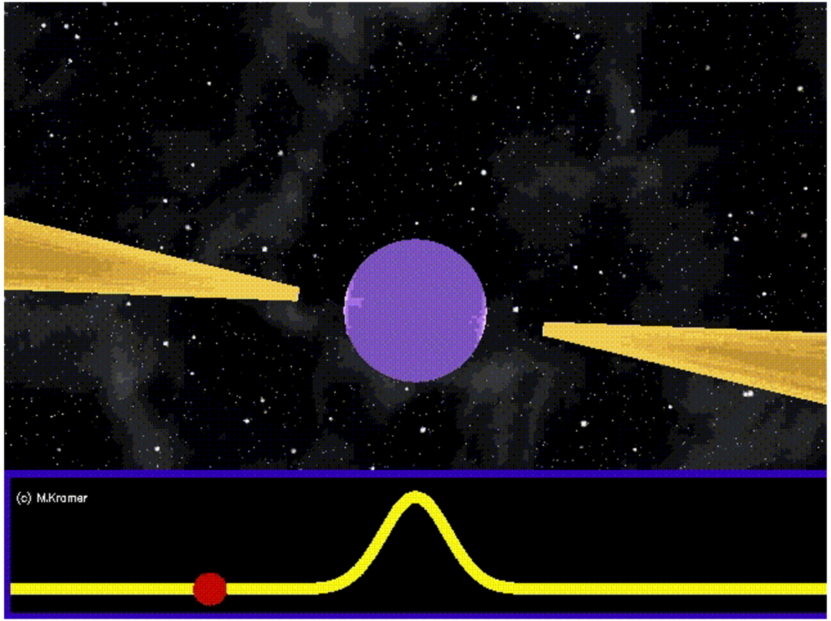
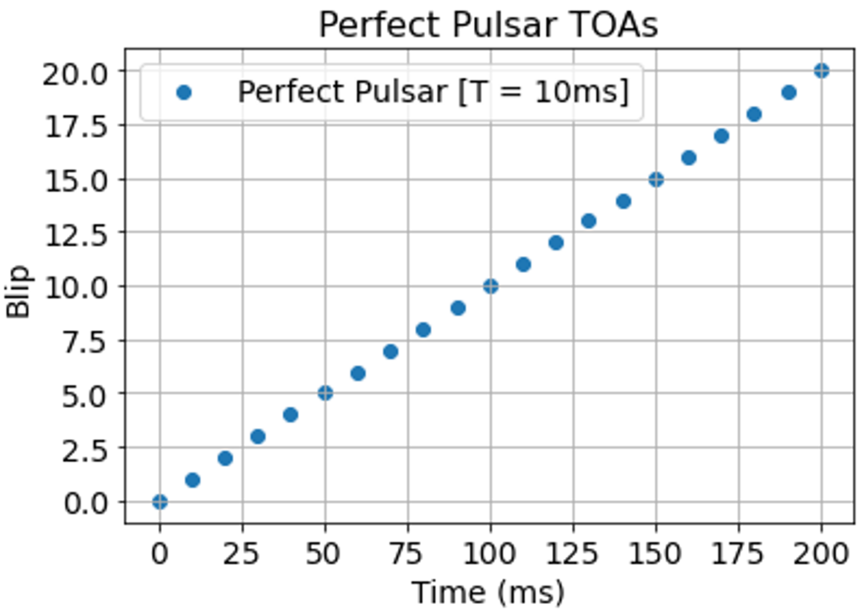


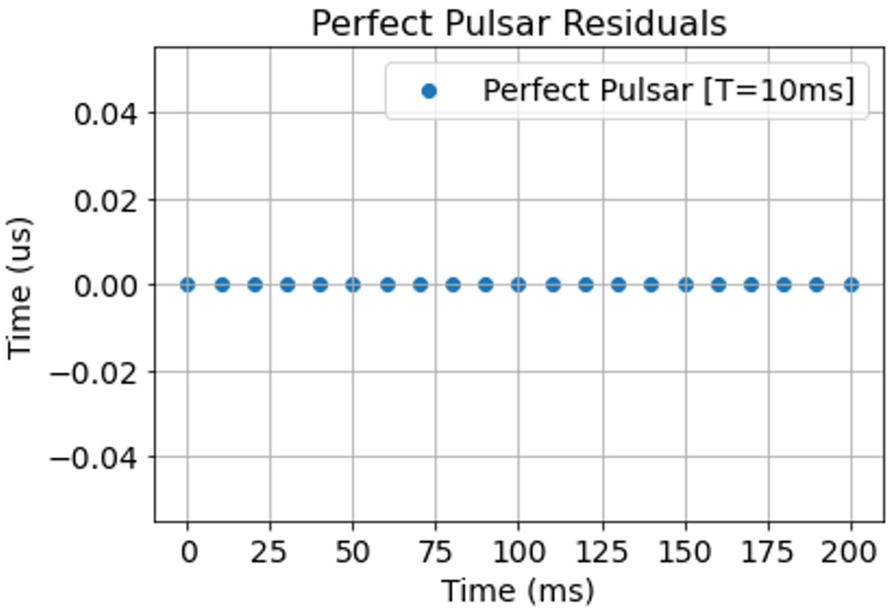
Image credit: Michael Kramer

Pulsar Timing Data

If pulsars were perfect, the pulses would arrive exactly on time:



Then, if we subtract the expected arrival time from the actual:

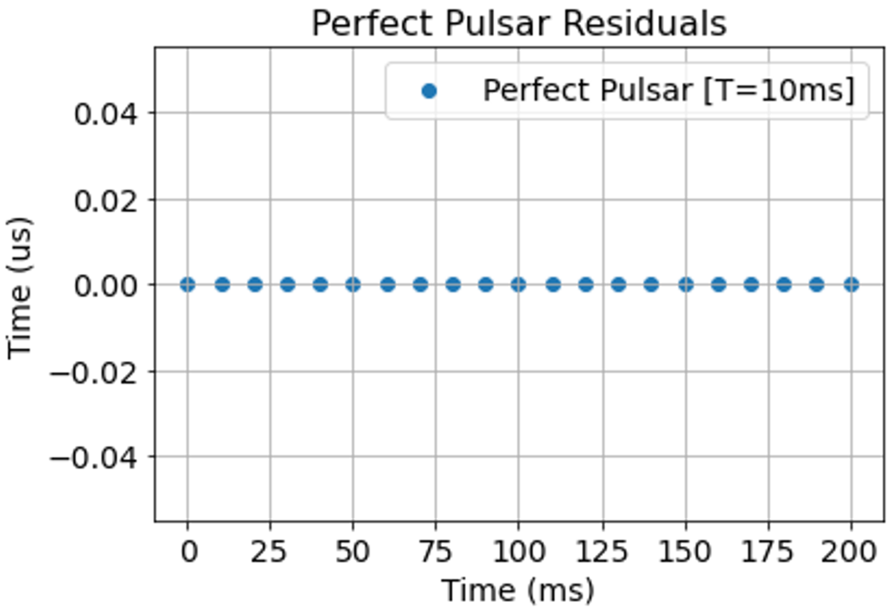


Pulsar Timing Data

A perfect pulsar has no residuals!
We get blips exactly on time.

$$\vec{r} = \vec{0}$$

Then, if we subtract the expected arrival time from the actual:



Pulsar Timing Data

Effect of errors in timing model appear as a linear (mismodeled frequency) and a quadratic (mismodeled frequency-derivative)

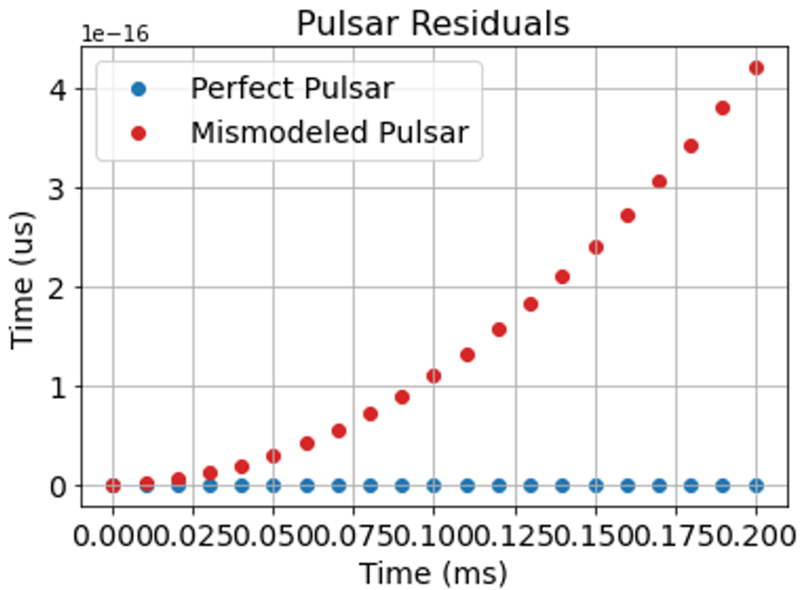
Don't worry! We can fit for a quadratic and subtract it out to deterministically remove this effect.

$$\vec{r} = \mathbf{M}\vec{\epsilon}$$

design matrix → $\mathbf{M}$

timing model errors → $\vec{\epsilon}$

However, we may imperfectly estimate the frequency or frequency-derivative of the pulsar.



Pulsar Timing Data

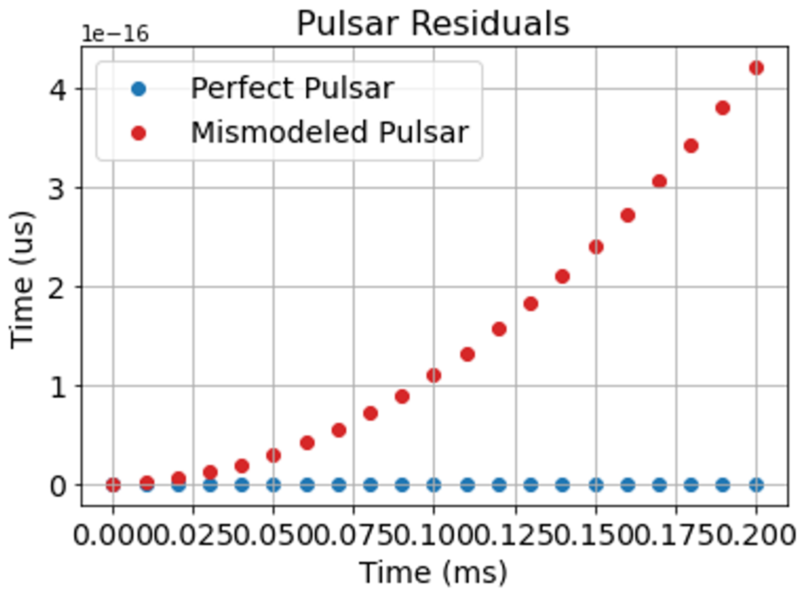
Effect of errors in timing model appear as a linear (mismodeled frequency) and a quadratic (mismodeled frequency-derivative)

Don't worry! We can fit for a quadratic and subtract it out to deterministically remove this effect.

$$\vec{r} = \mathbf{M}\vec{\epsilon}$$

$$\vec{\epsilon} \equiv \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \epsilon_2 \end{bmatrix} \qquad M \equiv \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ & \vdots & \\ 1 & t_{N_{\text{toa}}} & t_{N_{\text{toa}}}^2 \end{bmatrix}$$

However, we may imperfectly estimate the frequency or frequency-derivative of the pulsar.



Pulsar Timing Data

In reality, it's more complicated that just this. There are lots of physical parameters that affect our residuals.

(a) good timing model fit!
(b) bad frequency-derivative
(c) wrong sky position (yearly variation)
(d) wrong proper motion of pulsar (drifting in the sky)

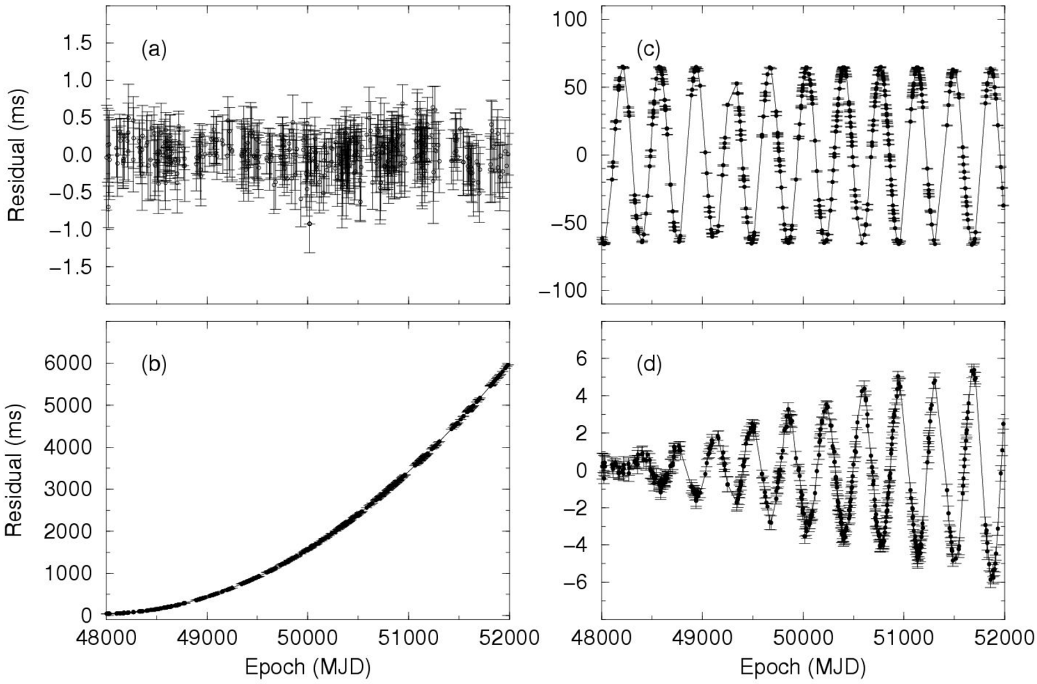


Image credit: Handbook of Pulsar Astronomy, Lorimer and Kramer

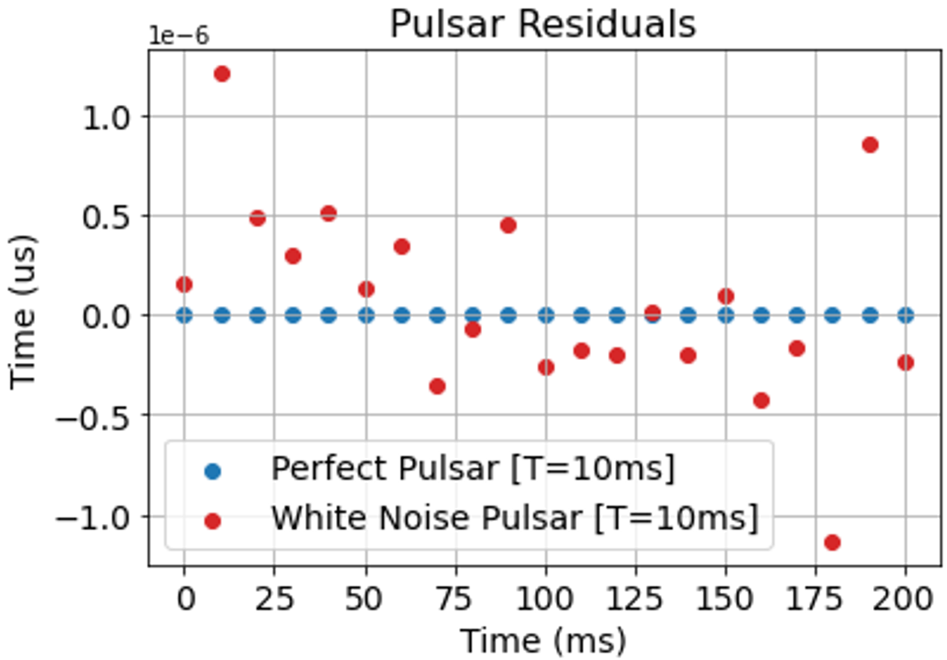
Pulsar Timing Data

Our observations also have some gaussian white noise

$$\vec{r} = \mathbf{M}\vec{\epsilon} + \vec{n}$$

$$\vec{n} \sim N(0, \sigma_{\mathrm{wn}})$$

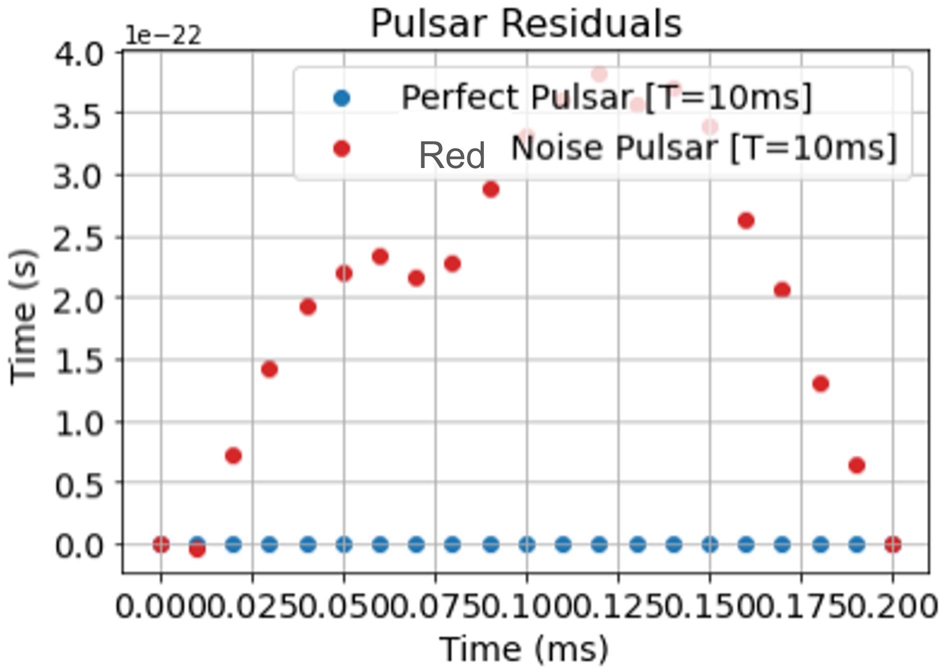
$$p(n_i|\sigma_{\mathrm{wn}}) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{wn}}^2}} \exp\left(\frac{-n_i^2}{2\sigma_{\mathrm{wn}}^2}\right)$$



Pulsar Timing Data

Pulsars can also exhibit “random walk” noise.

Over a long timescale, their frequencies wander around a little bit. We model this with a Fourier series



$$\vec{r} = \mathbf{M}\vec{\epsilon} + \vec{n} + \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}}$$

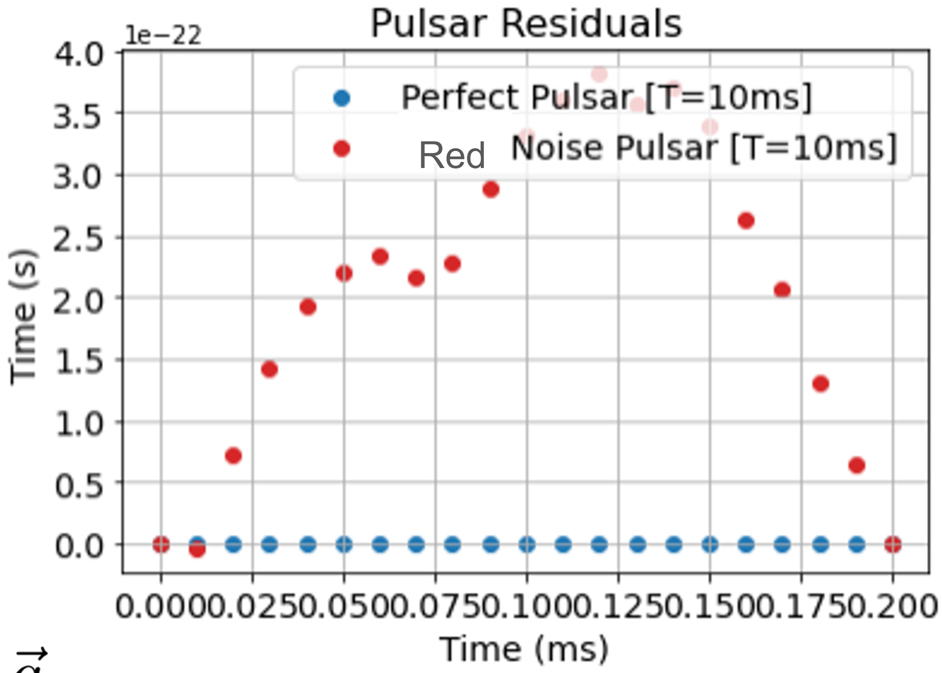
$$\mathbf{F} = \begin{pmatrix} \sin(2\pi t_1/T) & \cos(2\pi t_1/T) & \dots & \sin(2\pi N_f t_1/T) & \cos(2\pi N_f t_1/T) \\ \sin(2\pi t_2/T) & \cos(2\pi t_2/T) & \dots & \sin(2\pi N_f t_2/T) & \cos(2\pi N_f t_2/T) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sin(2\pi t_N/T) & \cos(2\pi t_N/T) & \dots & \sin(2\pi N_f t_N/T) & \cos(2\pi N_f t_N/T) \end{pmatrix}$$

$$\vec{a} \equiv \begin{bmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \vdots \\ a_{N_{\mathrm{freqs}}} \\ b_{N_{\mathrm{freqs}}} \end{bmatrix}$$

Pulsar Timing Data

Finally, the gravitational wave background also looks like a “random walk” noise!

Thus, we need two terms to track both sources of noise

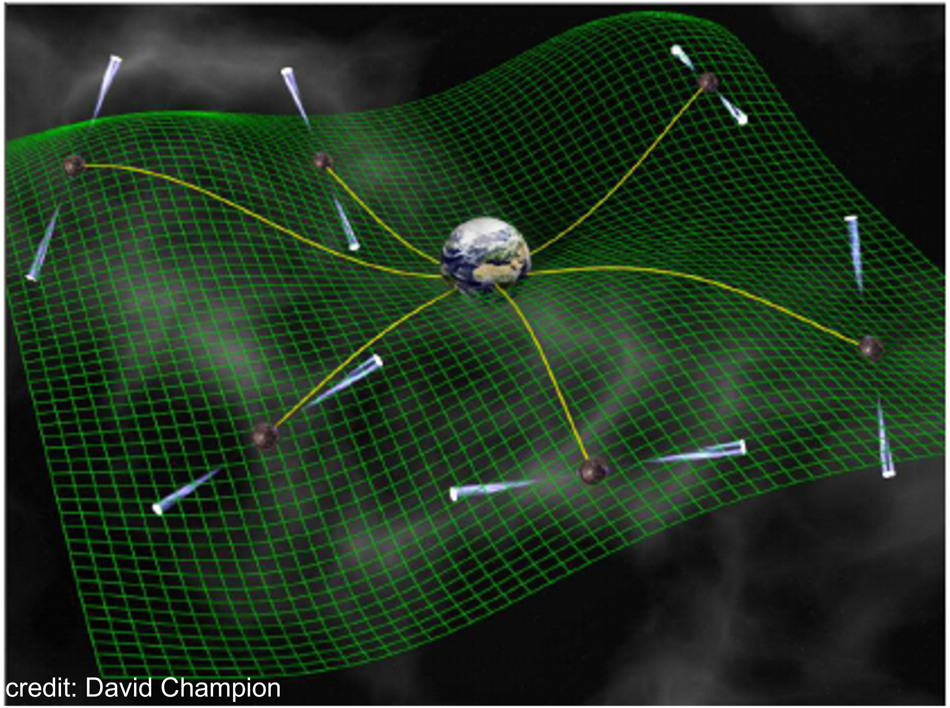


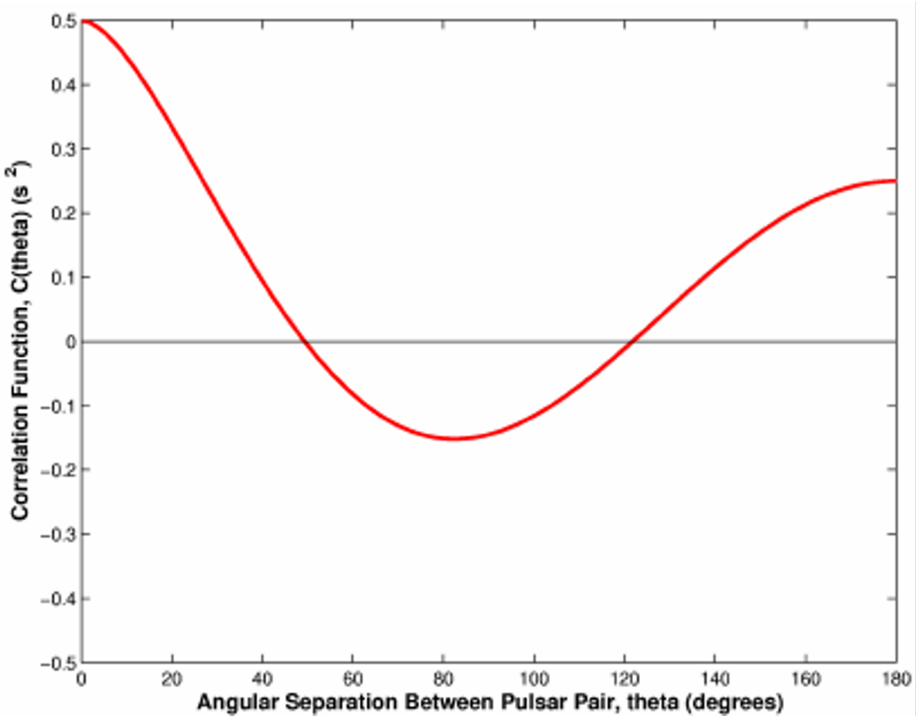
$$\vec{r} = \mathbf{M}\vec{\epsilon} + \vec{n} + \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}} + \mathbf{F}_{\mathrm{gw}}\vec{a}_{\mathrm{gw}}$$

$$\mathbf{F} = \begin{pmatrix} \sin(2\pi t_1/T) & \cos(2\pi t_1/T) & \dots & \sin(2\pi N_f t_1/T) & \cos(2\pi N_f t_1/T) \\ \sin(2\pi t_2/T) & \cos(2\pi t_2/T) & \dots & \sin(2\pi N_f t_2/T) & \cos(2\pi N_f t_2/T) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sin(2\pi t_N/T) & \cos(2\pi t_N/T) & \dots & \sin(2\pi N_f t_N/T) & \cos(2\pi N_f t_N/T) \end{pmatrix}$$

$$\vec{a} \equiv \begin{bmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \vdots \\ a_{N_{\mathrm{freqs}}} \\ b_{N_{\mathrm{freqs}}} \end{bmatrix}$$

Pulsar Timing Data



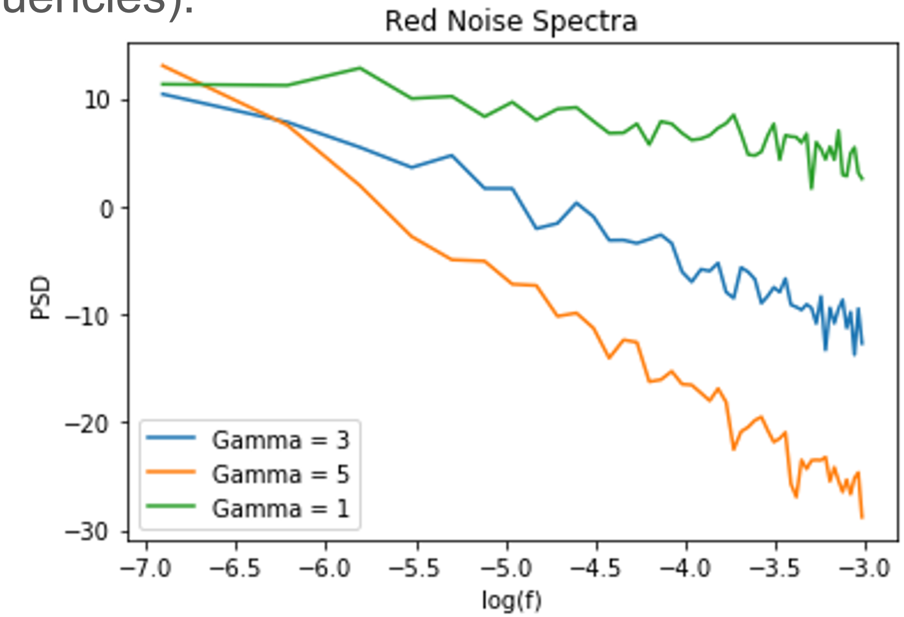


Because gravitational waves have a quadrupolar (+/×-polarization) pattern, we can reliably predict how the background will affect pulsars in different parts of the sky.

Pulsar Timing Data

We actually don't care too much about the coefficients themselves. The gravitational wave background can be described statistically as having a red power spectrum (more power at low frequencies).

$$S_h(f) = A^2 \left(\frac{f}{f_{1\mathrm{yr}}} \right)^{-\gamma}$$



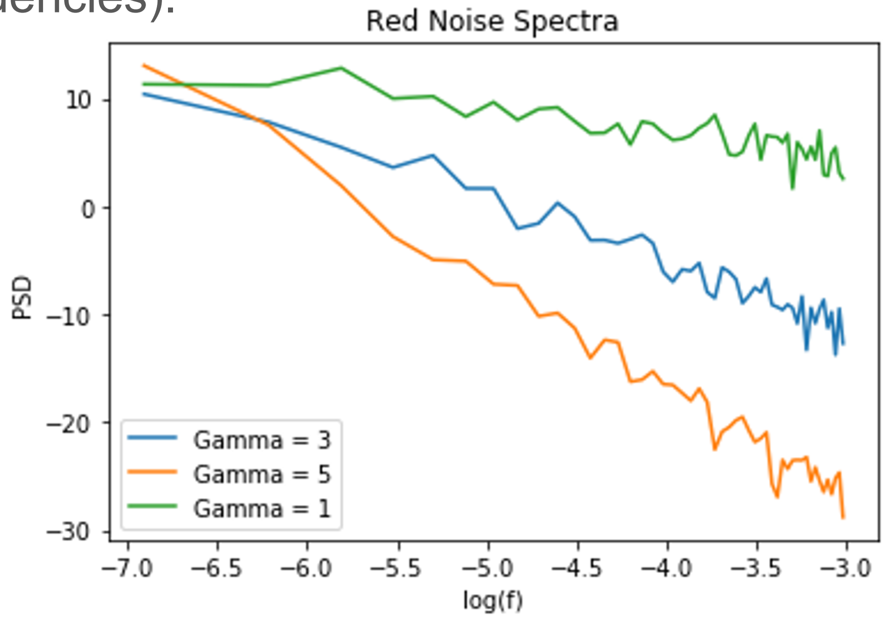
Pulsar Timing Data

We actually don't care too much about the coefficients themselves. The gravitational wave background can be described statistically as having a red power spectrum (more power at low frequencies).

$$S_h(f) = A^2 \left(\frac{f}{f_{1\mathrm{yr}}} \right)^{-\gamma}$$

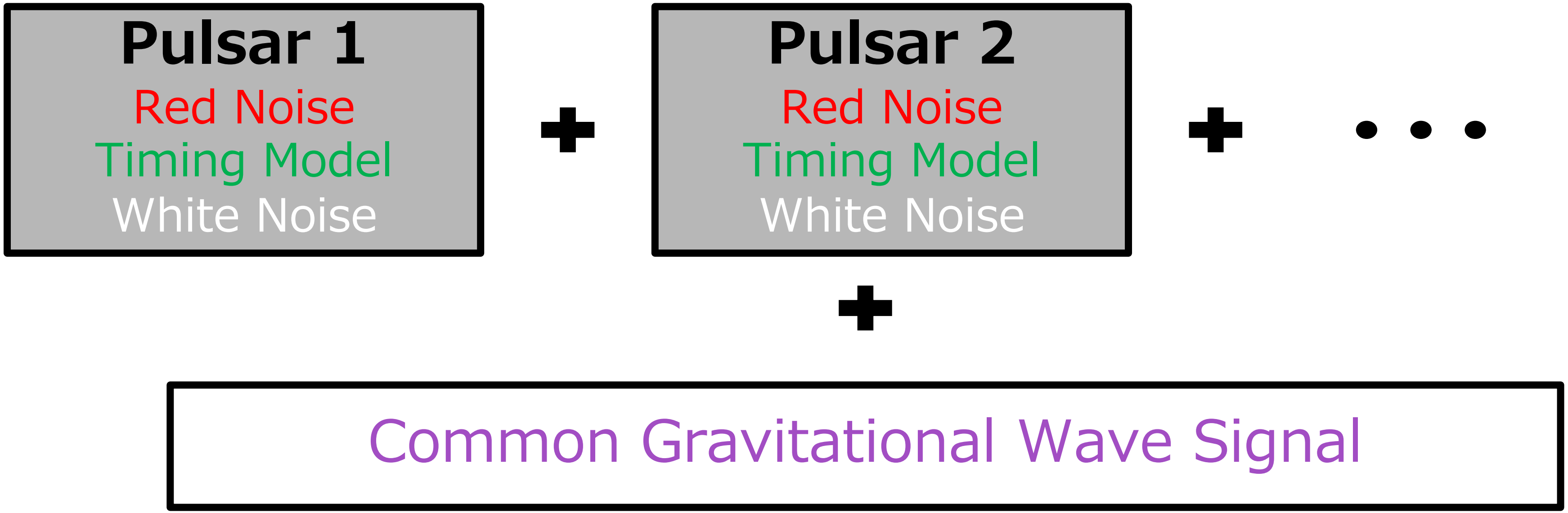
Amplitude

spectral index



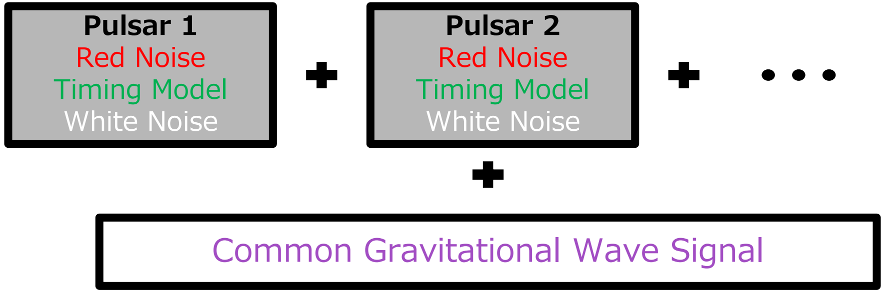
In reality, rather than estimating individual coefficients, we model the entire process with just these two hyperparameters for both GW and intrinsic pulsar noise

Pulsar Timing Data



$$\vec{r} = \mathbf{M}\vec{\epsilon} + \vec{n} + \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}} + \mathbf{F}_{\mathrm{gw}}\vec{a}_{\mathrm{gw}}$$

Pulsar Timing Data



These models are all built into ENTERPRISE

https://github.com/nanograv/enterprise

White Noise Parameters:

1) EFAC (Scaling factor)
2) EQUAD (additional noise in quadrature)
3) ECORR (correlated white noise)

Red Noise Parameters:

1) Amplitude
2) Spectral Index

$$S_h(f) = A^2 \left(\frac{f}{f_{1\mathrm{yr}}}\right)^{-\gamma}$$

Timing Model Parameters:

1) None; these can be analytically dealt with!

$$\vec{r} = \mathbf{M}\vec{\epsilon} + \vec{n} + \mathbf{F}_{\rm psr}\vec{a}_{\rm psr} + \mathbf{F}_{\rm gw}\vec{a}_{\rm gw}$$

Pulsar Timing Data

Pulsar 1		Pulsar 2		
Red Noise Timing Model White Noise	+	Red Noise Timing Model White Noise	+	• • •

+

Common Gravitational Wave Signal

These models are all built into ENTERPRISE

https://github.com/nanograv/enterprise

White Noise Parameters:

1) EFAC (Scaling factor)
2) EQUAD (additional noise in quadrature)
3) ECORR (correlated white noise)

Red Noise Parameters:

1) Amplitude
2) Spectral Index

$$S_h(f) = A^2 \left(\frac{f}{f_{1\rm yr}}\right)^{-\gamma}$$

Timing Model Parameters:

1) None; these can be analytically dealt with!

Questions?

How do we do analysis?

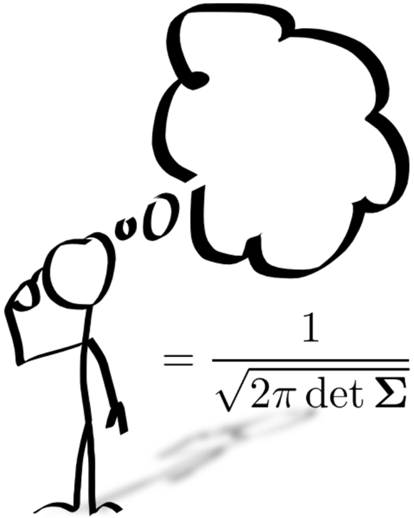
Key Idea: If I model everything out of my residuals correctly, I should be left with white noise. Then, the “correct” model parameters are the ones which maximize the probability that my final residuals are only white noise.

$$\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\text{psr}}\vec{a}_{\text{psr}} - \mathbf{F}_{\text{gw}}\vec{a}_{\text{gw}} = \vec{n}$$

$$p(\text{I}'\text{m right}) = p(\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\text{psr}}\vec{a}_{\text{psr}} - \mathbf{F}_{\text{gw}}\vec{a}_{\text{gw}} = \text{white noise})$$

How do we do analysis?

Key Idea: If I model everything out of my residuals correctly, I should be left with white noise. Then, the “correct” model parameters are the ones which maximize the probability that my final residuals are only white noise.



$$\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}} - \mathbf{F}_{\mathrm{gw}}\vec{a}_{\mathrm{gw}} = \vec{n}$$

$$p(\mathrm{I'm\ right}) = p(\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}} - \mathbf{F}_{\mathrm{gw}}\vec{a}_{\mathrm{gw}} = \text{white noise})$$

$$= \frac{1}{\sqrt{2\pi \det \mathbf{\Sigma}}} \exp\left((\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}} - \mathbf{F}_{\mathrm{gw}}\vec{a}_{\mathrm{gw}})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\mathrm{psr}}\vec{a}_{\mathrm{psr}} - \mathbf{F}_{\mathrm{gw}}\vec{a}_{\mathrm{gw}})\right)$$

How do we do analysis?

This is called our likelihood. This quantity tells us the probability that we would have observed the data we have ($\vec{r}$) if these were the underlying model parameters ($\vec{\theta}$)

$$L(\sigma_{\rm wn}, A_{\rm gw}, \gamma_{\rm gw}, A_{\rm psr}, \gamma_{\rm psr}, \vec{\epsilon}) \equiv L(\vec{\theta})$$

$$= \frac{1}{\sqrt{2\pi \det \mathbf{\Sigma}}} \exp\left((\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\rm psr}\vec{a}_{\rm psr} - \mathbf{F}_{\rm gw}\vec{a}_{\rm gw})^{\rm T} \mathbf{\Sigma}^{-1} (\vec{r} - \mathbf{M}\vec{\epsilon} - \mathbf{F}_{\rm psr}\vec{a}_{\rm psr} - \mathbf{F}_{\rm gw}\vec{a}_{\rm gw}) \right)$$

How do we do analysis?

Bayes' theorem allows us to turn a likelihood into a probability that those model parameters ($\vec{\theta}$) are correct

$$p(\vec{\theta}|\vec{r}) = \frac{L(\vec{\theta}|\vec{r})p(\vec{\theta})}{p(\vec{r})}$$

How do we do analysis?

Bayes' theorem allows us to turn a likelihood into a probability that those model parameters ($\vec{\theta}$) are correct

posterior

$$p(\vec{\theta}|\vec{r}) = \frac{L(\vec{\theta}|\vec{r})p(\vec{\theta})}{p(\vec{r})}$$

How do we do analysis?

Bayes' theorem allows us to turn a likelihood into a probability that those model parameters ($\vec{\theta}$) are correct

Likelihood

posterior

$$p(\vec{\theta}|\vec{r}) = \frac{L(\vec{\theta}|\vec{r})p(\vec{\theta})}{p(\vec{r})}$$

How do we do analysis?

Bayes' theorem allows us to turn a likelihood into a probability that those model parameters ($\vec{\theta}$) are correct

posterior

Likelihood

prior

$$p(\vec{\theta}|\vec{r}) = \frac{L(\vec{\theta}|\vec{r})p(\vec{\theta})}{p(\vec{r})}$$

How do we do analysis?

Recall the key idea: The “correct” model parameters are the ones which maximize the posterior probability that my final residuals are just white noise.

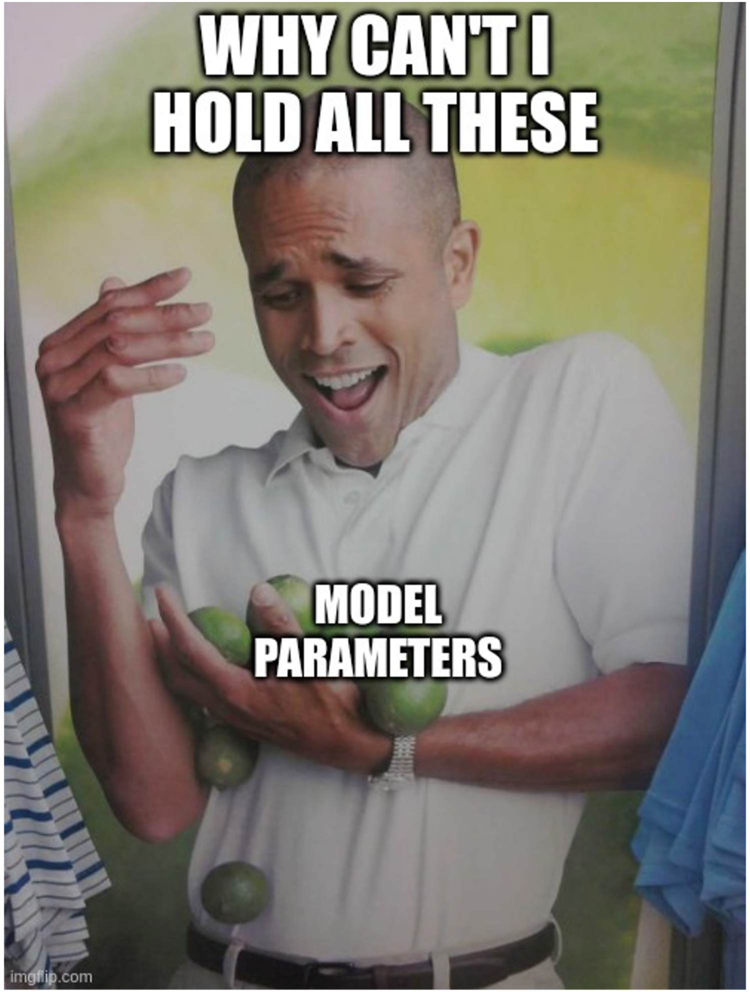
Question: How do I choose what model parameters to test? Do I test every single combination?

How do we do analysis?

Recall the key idea: The “correct” model parameters are the ones which maximize the posterior probability that my final residuals are just white noise.

Question: How do I choose what model parameters to test? Do I test every single combination?

NO



How do we do analysis?

Question: How do I choose what model parameters to test? Do I test every single combination?

The parameter space is too big. We need to use Markov Chain Monte Carlo sampling to explore the posterior.

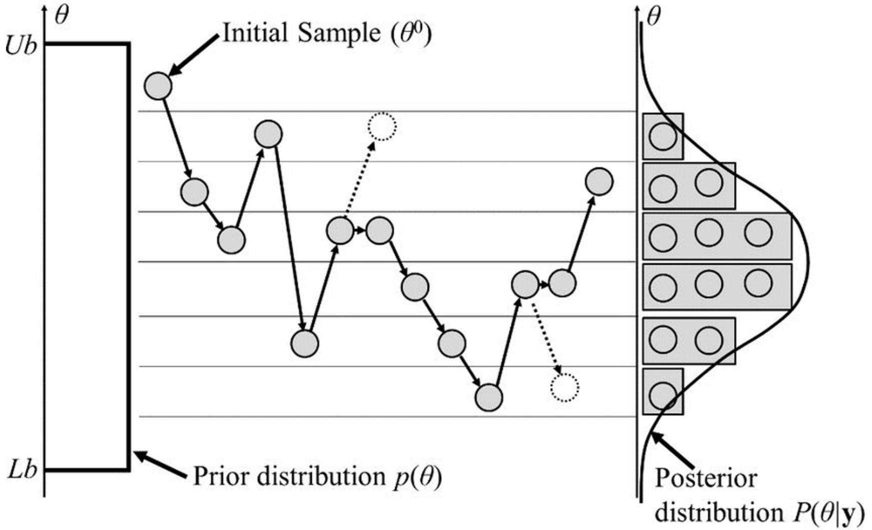
How do we do analysis?

Question: How do I choose what model parameters to test? Do I test every single combination?

The parameter space is too big. We need to use Markov Chain Monte Carlo sampling to explore the posterior.

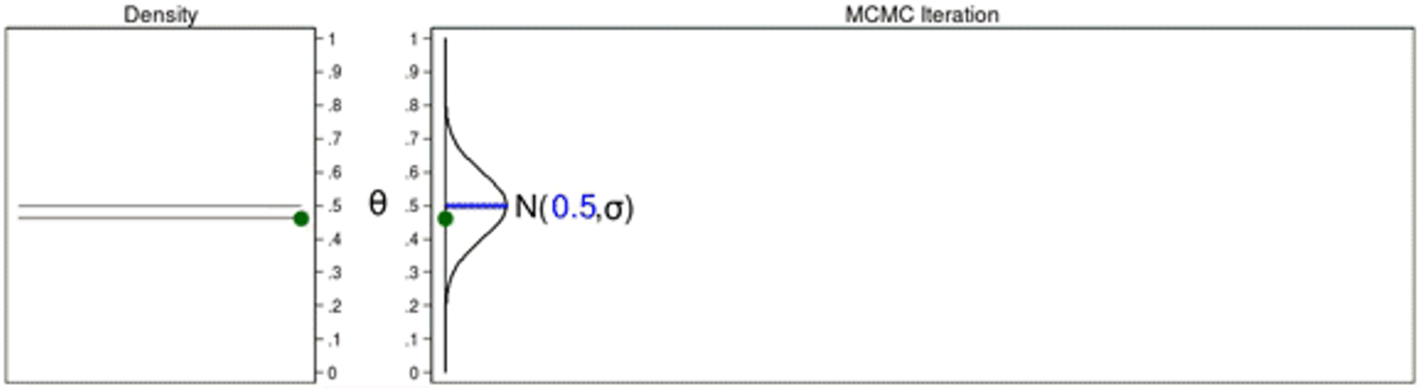
MCMC Sampling Explained:

1) Pick a random set of parameters and evaluate likelihood
2) Pick another set of random parameters and evaluate likelihood
3) Accept and write down the new set with probability determined by the likelihood-ratio
4) Repeat 2) and 3) until your advisor wants to see a plot



Jaewook, Woosuk, Joo-Ho. (2015). Energies. 8. 5538-5554. 10.3390/en8065538.

How do we do analysis?



https://github.com/nanograv/PTMCMCSampler

Draw $\theta_t \sim \text{Normal}(0.5,\sigma) = 0.460$

How do we do analysis

We now have numerically computed posterior probability distribution for our white noise and red noise parameters.

We can **use these to make inferences about the significance of our gravitational wave**!

Summary

Our data look like this:

$$\vec{r} = \mathbf{M}\vec{\epsilon} + \vec{n} + \mathbf{F}_{\rm psr}\vec{a}_{\rm psr} + \mathbf{F}_{\rm gw}\vec{a}_{\rm gw}$$

Pulsar 1
Red Noise
Timing Model
White Noise

+

Pulsar 2
Red Noise
Timing Model
White Noise

+ . . .

+

Common Gravitational Wave Signal

Use Bayes' theorem and MCMC sampling to compute our probability distributions for our model parameters (like the GWB)

$$p(\vec{\theta}|\vec{r}) = \frac{L(\vec{\theta}|\vec{r})p(\vec{\theta})}{p(\vec{r})}$$

tutorial notebooks and data here