Pulsar Timing Array Data Analysis for Periodic and Near-periodic Gravitational Wave Sources

Xihao Deng and Lee Samuel Finn

The Pennsylvania State University
NANOGrav Collaboration

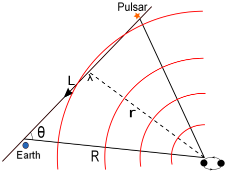
Outline

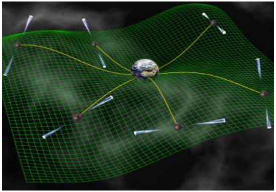
1. An overview on pulsar timing response to a periodic gravitational wave
   - Plane-wave response—"Earth term" and "Pulsar term"
   - Conventional point of view on "Pulsar term"
2. Bayesian Data Analysis
   - Parameter Inference
   - Bayes Factor
3. Importance of the "Pulsar term"
   - Enhance the detection probability of gravitational waves
   - Improve the accuracy and precision of gravitational wave parameter estimation
4. Conclusion and prospect

Pulsar Timing Response to A Monochromatic Gravitational Wave

In a pulsar timing array, the timing response of the jth pulsar to a monochromatic gravitational wave (GW) is

$$\tau_j = A_j \left(\sin\left(2\pi f\, t + \phi\right) - \sin\left(2\pi f\, t + \phi - 2\pi f\, L_j(1 - \cos\theta_j)\right) \right)$$





A_j is the jth pulsar timing amplitude, depending on GW amplitude & frequency (h, f), polarization & inclination angle (ψ, ι) and source sky location (α, λ).

"Earth term" and "Pulsar term"

The pulsar timing response function is the sum of two terms:

- "Earth term": $\tau_{jE} = A_j \sin\left(2\pi f\, t + \phi\right)$
  - Only depends on GW parameters—they are the *same* for different pulsars
  - Coherent for different pulsars in a pulsar timing array
- "Pulsar term": $\tau_{jP} = A_j \sin\left(2\pi f\, t + \phi - 2\pi f\, L_j(1 - \cos\theta_j)\right)$
  - Depends on both GW parameters and pulsar distances L_j, which are *different* for different pulsars
  - Incoherent for different pulsars in a pulsar timing array

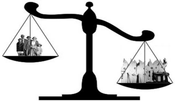
Therefore, for a pulsar timing array composed of N pulsars, the parameter set $\boldsymbol{\theta} =$ GW parameters $+N$ incoherent phases $2\pi f\, L_j(1 - \cos\theta_j)$, totally $7 + N$ dimensions.

Conventional Point of View on "Pulsar term"

— "Pulsar term" can be thrown away in data analysis!

"Pulsar term" is conventionally thought be negligible, because

- "Pulsar term" vanishes after incoherent average
  (e.g. Sesana *et al*, 2009)



- Computationally impossible with "Pulsar term"
  (Ellis *et al*, 2012)

"Pulsar term" may be included in the data analysis *only if*

- GW chirp is important
  (Corbin & Cornish, 2010)
  But: high f sources are rare

- Parsec pulsar distance measurement
  (Lee *et al*, 2011)
  But: hard to achieve without SKA

Is "Pulsar term" Really Negligible ?

In most likely cases, GW chirp is not important and we are not able to measure pulsar distances within parsec. Does this mean we can throw away "Pulsar term"?

It is the *likelihood function* that determines the detection.

How do we know the "Pulsar term" can factor out of the *likelihood function*?

In a more careful Bayesian analysis, we have found that "Pulsar term" can NOT factor out of the likelihood function and they will NOT vanish after average over all the pulsars.

Likelihood Function

— Data Dependent Contribution to Detection Probability

The Likelihood function of the jth pulsar is

$$\Lambda_j(\mathrm{d}_j|\boldsymbol{\theta}) = \left(\begin{array}{l} \text{probability of observing } \mathrm{d}_j \text{ assuming the gravitational} \\ \text{wave signal characterized by the parameter } \boldsymbol{\theta} \text{ is present} \end{array} \right)$$

For Gaussian noises, the likelihood function of a PTA can be written as:

$$\Lambda(\mathbf{d}|\boldsymbol{\theta}) = \prod_{j=1}^{n} \Lambda_j(\mathrm{d}_j|\boldsymbol{\theta}) = \prod_{j=1}^{n} N(\mathrm{d}_j - \tau_j(\boldsymbol{\theta})|C_j)$$

where d_j and τ_j respectively represent the data and the timing response for the jth pulsar, and C_j is the noise power spectral density of the j pulsar.

Remember the parameter $\boldsymbol{\theta}$ is composed of 7 GW parameters plus N incoherent phases, totally $7 + N$ dimensions.

Likelihood Function (Continued)

The Likelihood function of a PTA is a normal distribution function,

$$\Lambda(\mathbf{d}|\boldsymbol{\theta}) \propto \exp\left[\sum_{j=1}^{N} -\frac{1}{2}(\mathrm{d}_j - \tau_j(\boldsymbol{\theta}))C_j^{-1}(\mathrm{d}_j - \tau_j(\boldsymbol{\theta}))\right]$$

$$\propto \exp\left[-\frac{1}{2}\Big(\sum_{j=1}^{N} \mathrm{d}_j C_j^{-1}\mathrm{d}_j + 2\mathrm{d}_j C_j^{-1}\tau_{jE}(\boldsymbol{\theta}) + \mathrm{d}_j C_j^{-1}\tau_{jP}(\boldsymbol{\theta}) + \ldots + \tau_{jP}(\boldsymbol{\theta})C_j^{-1}\tau_{jP}(\boldsymbol{\theta})\Big)\right] \to \text{quadratic term}$$

The quadratic term $\tau_{jP}(\boldsymbol{\theta})C_j^{-1}\tau_{jP}(\boldsymbol{\theta})$ will never vanish after we sum over all pulsars, because the trigonometric identity

$$A^2 \sin^2\theta = \frac{A^2}{2}(1 - \cos 2\theta)$$

Estimating Gravitational Wave Parameters —Posterior Probability Density

Given the data set $\mathbf{d}$ of an array of pulsar, what is the probability density p that the gravitational wave is characterized by $\boldsymbol{\theta}$? It is mainly determined by the likelihood function $\Lambda(\mathbf{d}|\boldsymbol{\theta})$ and a priori knowledge on $\boldsymbol{\theta}$:

$$p(\boldsymbol{\theta}|\mathbf{d}) = \frac{1}{Z}\Lambda(\mathbf{d}|\boldsymbol{\theta})q(\boldsymbol{\theta})$$

$q(\boldsymbol{\theta})$ is the prior probability, a priori knowledge on $\boldsymbol{\theta}$ before we make observations. $p(\boldsymbol{\theta}|\mathbf{d})$ is the posterior probability and the data dependent likelihood function update the prior.

Therefore, **adding "Pulsar term" in likelihood function will result in different estimation of GW parameters**.

Gravitational Wave Detection — Bayes Factor

What odds should we give that a gravitational wave is present?

$$\mathrm{M}_1 = (\text{a gravitational wave signal present})$$
$$\mathrm{M}_0 = (\text{no gravitational wave signals present})$$

The relative probability of any two hypotheses is their odds-ratio $\mathcal{O}$

$$\mathcal{O} = \frac{p_{\mathrm{M}}(\mathrm{M}_1|\mathbf{d})}{p_{\mathrm{M}}(\mathrm{M}_0|\mathbf{d})} = \frac{\Lambda(\mathbf{d}|\mathrm{M}_1)}{\Lambda(\mathbf{d}|\mathrm{M}_0)} \frac{q(\mathrm{M}_1)}{q(\mathrm{M}_0)}$$

where

$$\Lambda(\mathbf{d}|\mathrm{M}_k) = (\text{probability of observation } \mathbf{d} \text{ if assuming } \mathrm{M}_k)$$
$$q(\mathrm{M}_k) = (\text{prior probability of hypothesis } \mathrm{M}_k)$$

Detection Probability — Bayes Factor

The data dependent contribution to the odds ratio $\mathcal{O}$ is Bayes Factor:

$$B(\mathrm{M}_1, \mathrm{M}_0|\mathbf{d}) = \frac{\Lambda(\mathbf{d}|\mathrm{M}_1)}{\Lambda(\mathbf{d}|\mathrm{M}_0)} = \frac{\int d^n\theta\, \Lambda(\mathbf{d}|\boldsymbol{\theta})q(\boldsymbol{\theta})}{\Lambda(\mathbf{d}|\mathrm{M}_0)}$$

The Bayes Factor reflects the evidence provided by the data **d** in favor of the hypothesis M_1 relative to M_0, which determines the detection probability.

It *marginalizes* over the parameter space because all of the parameters have some probability to describe the GW signal.

Therefore, **adding "Pulsar term" also results in a different Bayes Factor and different detection probability.**

Interpretation of Bayes Factor

Bayes Factor is the summary of the evidence provided by the data in favor of one hypothesis as opposed to the other.

The interpretation of Bayes Factor is (Kass & Raftery, 1995)

# of σ	Bayes Factor	Evidence against M_0
1 σ	3	A bare mention
2 σ	10	Positive
3 σ	120	Strong

“Pulsar term” should not be ignored !!!

It significantly enhances the detection probability

We apply Bayesian analysis to simulated observations by a pulsar timing array (PTA) to illustrate the importance of “Pulsar term”.

- Pulsar timing array: 20 pulsars with 100ns RMS timing noise
- Observation time: weekly timing for 10 year observation
- Hypothetical signal: a 6-year GW with 6.5ns peak timing residual
- Computational Method: Markov Chain Monte Carlo (MCMC)

Bayes Factor:

- Including “Pulsar term”: 2.69×10^5

  Decisive evidence on the presence of a gravitational wave !

- Ignoring “Pulsar term”: 4.95

  The presence of a gravitational wave is NOT obvious !

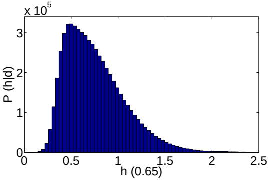
“Pulsar term” should not be ignored !!!

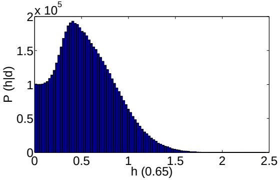
It significantly improves the quality of parameter inference

Marginalized posterior distribution of GW amplitude h and sky location α & λ:

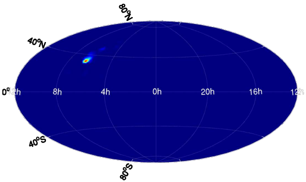
Parameter | Including “Pulsar term” | Ignoring “Pulsar term”

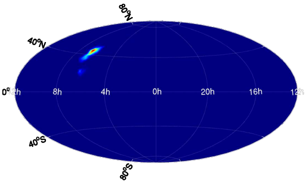
h (0.65)





sky location





“Pulsar term” improves the quality of parameter estimation.

Conclusion and Prospect
— New Detection Pipeline Including “Pulsar term”

We develop a new periodic gravitational wave detection pipeline. It applies Bayesian analysis and includes the “Pulsar term” contribution. It is computationally possible, and will significantly enhance the detection probability and improve the quality of parameter inference.

First claim of successful gravitational wave detection requires not just good hardwares such as sensitive radio telescopes, quiet pulsars and genius timing technique, but also an excellent sofware — an efficient gravitational wave data analysis method!