TEMPONEST tutorial - MeerKAT Pulsar Timing Workshop

Aditya Parthasarathy and Ryan Shannon

September 17, 2021

Abstract

The exceptional rotational stability of millisecond pulsars combined with sophisticated and powerful analysis tools has led to the pursuit of a variety of science goals, from the detection of planets around pulsars to some of the most precise tests of general relativity in the strong field regime. In this tutorial you will be briefly learning how to use TEMPONEST for modelling timing noise in pulsars. TEMPONEST is a Bayesian analysis software that performs a **simultaneous analysis** of the linear and non-linear deterministic pulsar timing model and additional stochastic parameters. It uses the Bayesian inference tool *MultiNest* to explore this joint parameter space, whilst using Tempo2 as an established means of evaluating the timing model at each point in that space. Please refer to **Lentati et al. (2014)** and *https://github.com/aparthas3112/*TEMPONEST for more detailed information.

1 General Introduction

Let's start with a brief recap of pulsar timing. The arrival times of pulses (ToAs) from a pulsar are recorded by an observatory in a series of discrete observations over a certain period time. To account for the motion of the Earth, these ToAs must be transformed into a common frame of reference, the solar system barycenter. Once that is done, a model of the pulsar is fit to the ToAs. This model generally describes the position, spin down, orbital parameters (if it's a binary) etc. Any discrepancies between the observed ToAs and those predicted by the pulsar model are manifested in the timing residuals. The timing residuals are extensively analyzed to understand and model the observed discrepancies. For a more detailed review of pulsar timing, start by referring to **Lorimer & Kramer (2012)**.

TEMPO2 is the most commonly used software for the above analysis. While fitting the pulsar model to the measured ToAs, it uses an initial guess to generate pre-fit timing residuals and then a linear least squares fit is used to improve the timing residuals. Typically multiple iterations of the fit are performed until convergence is reached.

2 What's the problem then?

The assumption made in Tempo2 is that the ToAs can be solely represented as a sum of the deterministic signal due to the timing model and it uses a white noise model that is dictated only by the uncertainties of the ToAs. The issue with this assumption is that, it is **not true** in most realistic data sets.

Power from stochastic signals like timing noise, which is thought to be due to the rotational irregularities of the neutron star or the correlation induced due to the stochastic gravitational wave backgound (GWB) signal, is absorbed by the timing model, hence affecting the process of parameter estimation and also breaking down the assumptions made previously.

Over the years, there have been several techniques developed to account for this issue. The Cholesky decomposition method as described in **Coles et al. (2011)** uses the Cholesky decomposition of the covariance matrix that is estimated from the power spectral density of the timing residuals to whiten the residuals. However as stated later by **van Haasteren & Levin (2013)**, the post fit timing residuals aren't time stationary making the estimation of the power spectral density poorly defined.

So, perhaps the preferred approach is to perform a joint analysis of the timing parameters simultaneously along with the stochastic parameters.

The above description of the problems and techniques is by no means a comprehensive one. Timing noise in pulsars, young and old, have been studied over many years (as you would have seen from Ryan's talk) and providing a descriptive review is beyond the scope of this tutorial.

3 What does TempoNest do?

TEMPONEST provides a means of performing a simultaneous analysis of linear (and) non-linear timing model parameters with stochastic parameters using a Bayesian inference tool called **MultiNest** to efficiently sample this joint parameter space. It also uses Tempo2 to evaluate the solution at each point in this space and finally provides a Bayesian evidence value to compare and select different models used for the modelling the data.

Now, if you didn't follow through the last paragraph, these are the key words that you might need to focus on, for the purposes of this tutorial:

- Bayesian inference
- Bayesian evidence
- Model selection

3.1 What's Bayesian inference?

If we have a set of model parameters defined as Θ in a model or hypothesis, H given the data, D. Bayes' theorem states that:

$$\Pr(\Theta \mid D, H) = \frac{\Pr(D \mid \Theta, H)\Pr(\Theta \mid H)}{\Pr(D \mid H)}, \tag{1}$$

where,

- $\Pr(\Theta \mid D, H) \equiv \Pr(\Theta)$ is the posterior probability distribution of the parameters,
- $\Pr(D \mid \Theta, H) \equiv L(\Theta)$ is the likelihood,
- $\Pr(\Theta \mid H) \equiv \pi(\Theta)$ is the prior probability distribution,
- and $\Pr(D \mid H) \equiv Z$ is the Bayesian Evidence.

3.2 What's Bayesian evidence?

From the above equation, we can estimate the evidence by the following equation:

$$Z = \int L(\Theta)\pi(\Theta)\mathrm{d}^n\Theta, \tag{2}$$

where n is the dimensionality of the parameter space.

3.3 What's model selection?

The equation for evidence $\mathbf{Z}$ states that it is the average of the likelyhood over the prior.

That states that:

- The evidence is larger for a model if more of its parameter space is likely
- and smalled for a model if its parameter space has low likelyhood values

Thus the evidence is a very useful parameter that helps us compare and select the best model that is significantly better at explaining the data.

4 How does TempoNest work?

Briefly, TEMPONEST uses the Bayesian inference package called MultiNest, which is a nested sampling framework which provides an efficient means of sampling from posteriors and also calculates the evidence. Please refer to the **Feroz et al. (2011)** for a more detailed understanding of the algorithm.

5 How does all of this apply to pulsar timing?

For any given pulsar, we can describe the ToAs for the pulses as a sum of a deterministic and a stochastic component.

$$\mathbf{t}_{\mathrm{total}} = \mathbf{t}_{\mathrm{deterministic}} + \mathbf{t}_{\mathrm{stochastic}}, \tag{3}$$

Here is an excerpt from the **temponest paper** that is quite important and perhaps summarizes the model used for the analysis of data.

Writing the deterministic signal due to the timing model as $\tau(\epsilon)$, and the uncertainty associated with a particular TOA i as σ_i we can write the likelihood that the data is described solely by the timing model as:

$$\Pr(\mathbf{t}|\epsilon) \propto \left(\prod_{i=1}^{n} \sigma_i^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{(t_i - \tau(\epsilon)_i)^2}{\sigma_i^2}\right). \tag{4}$$

This represents the simplest model choice possible in TEMPONEST*, including only those free parameters present in the TEMPO2 fit. From here we can now begin to* ***make our model for the stochastic contribution to the signal more realistic by introducing additional parameters to describe the white and red noise components****, in order to compare the evidence with this simpler model and determine the optimal set of parameters supported by the data.*

If you're interested in understanding how the white noise and red noise processes are modelled and how TEMPONEST uses all of this to marginalize over the timing model, please refer to the **Lentati et al. (2014)**

We will, however cover some of the noise modelling aspects in the following sections.

5.1 Using TempoNest: a brief introduction

For this exercise, you will use the files in the **TempoNest** directory. We shall be trying to model the timing noise in the Vela pulsar. You might be familiar at this stage with the contents of the .par and .tim files.

Use *tempo2* and plot the timing residuals. How do they look like? What can you find out about the data set?

You will notice that there seems to be a strong presence of red noise in the timing residuals. Young pulsars are known to exhibit significant timing noise, perhaps due to rotational irregularities of the neutron star. We will try using TEMPONEST to model and subtract the red noise from the timing residuals.

Since TEMPONEST is built as a plugin for TEMPO2, you can run TEMPONEST using:

```
tempo2 -gr temponest -f <parfile> <timfile> -cfile <.cfile>
```

But before we go ahead and run TEMPONEST, we have to understand a few things.

5.2 The TempoNest configuration file

Let's try to understand the configuration file for TEMPONEST. Open up the configuration file (*temponest.cfile*) in the **TempoNest** directory.

The parameters set in the configuration file dictates the settings used in that TEMPONEST run.

A detailed description of the configuration file is beyond the scope of this exercise. The parameters that are important for this exercise are listed above the rest in the configuration file.

5.2.1 Key parameters in the configuration file for this tutorial

The **root** parameter specifies the output directory for the TEMPONEST results and the prefix for the filenames. For example,

```
root = <directory>/<prefix>-
```

Please note that it is required for you to create the output directory manually and TEMPONEST does not do this for you.

To fit for the white noise parameter $EFAC$, we set **incEFAC=1**.

To fit for a power-law red noise model we set **incRed=3**.

The prior ranges for the white noise and red noise models (for both the spectral index and amplitude) are defined by the parameters **AlphaPrior,AmpPrior and EFACPrior**.

5.3 Running TempoNest

Try running TEMPONEST using the command mentioned above. Please remember to create the required output directory.

You will notice that a lot of text is printed in the terminal during this process and it stops once the fit converges after multiple iterations. This might take a few minutes.

5.4 TempoNest results

Let us now finally look at the results that TEMPONEST produces. You will notice that TEMPONEST has produced a series of files (within the results directory) with the prefix as assigned in the configuration file. The key files that we might have to look at for this tutorial are the following:

- The output par file (.par)
- The paramnames file (.paramnames)

TEMPONEST outputs a new .par file after each run. What difference do you notice between the new par file and the input par file?

There are a set of new parameters at the end of the new par file output by TEMPONEST. These parameters, **TNGlobalEF, TNRedAmp, TNRedGam and TNRedC** describe the white and red noise models. TNGlobalEF is the value of EFAC. TNRedAmp and TNRedGam are the amplitude and spectral index of the red noise model. TNRedC is the number of Fourier components used to sample the red noise.

So what did TEMPONEST do? To take a look at the result, use tempo2 to load the new par file with the old tim file.

```
tempo2 -gr plk -f <temponest_par> <timfile>
```

Look at the pre-fit and the post-fit residuals. What do you notice? Now press *'shift+k'*. Did that change something? Does the noise modelling work as expected?

Now, let's take a look at the posterior distributions. The *.paramnames* file lists the parameters that were fitted for in TEMPONEST. What parameters are listed?

Use the *posteriors.py* script to plot the posteriors distributions of these parameters (remember to use the *-h* flag). The posterior distributions are important in validating the results and are also important to study the correlation among the different fit parameters.

Were you able to plot the posteriors? What are the values of the red noise amplitude, spectral index and EFAC? Do you see any covariance between the three parameters?

`EFAC:`
`Red noise amplitude:`
`Red noise spectral index:`

References

Coles W., Hobbs G., Champion D. J., Manchester R. N., Verbiest J. P. W., 2011, http://dx.doi.org/10.1111/j.1365-2966.2011.19505.x , https://ui.adsabs.harvard.edu/abs/2011MNRAS.418..561C 418, 561

Feroz F., Hobson M. P., Bridges M., 2011, MultiNest: Efficient and Robust Bayesian Inference (ascl:1109.006)

Lentati L., Alexander P., Hobson M. P., Feroz F., van Haasteren R., Lee K. J., Shannon R. M., 2014, http://dx.doi.org/10.1093/mnras/stt2122 , https://ui.adsabs.harvard.edu/abs/2014MNRAS.437.3004L 437, 3004

Lorimer D. R., Kramer M., 2012, Handbook of Pulsar Astronomy

van Haasteren R., Levin Y., 2013, http://dx.doi.org/10.1093/mnras/sts097 , https://ui.adsabs.harvard.edu/abs/2013MNRAS.428.1147V 428, 1147