भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

TIMING AND ANALYSIS OF MILLISECOND PULSARS
B.Tech. Project Fall 2021

Divyansh Kharbanda
EP19BTECH11002
Engineering Physics
Indian Institute of Technology, Hyderabad

Supervisor : Dr. Shantanu Desai

November 20, 2022

Contents

ABSTRACT

The pulsar timing array experiment (PTA) employs an ensemble of millisecond pulsar clocks to detect gravitational waves. PTAs are expected to open the nano-Hertz window of the gravitational wave spectrum. Although the millisecond pulsars are known to be very stable rotators, on rare occasions there could be irregularities such as glitches, extreme scattering events, profile changes, etc seen in these pulsars. Monitoring such events and their recovery is crucial for modelling the noise processes in pulsar data sets used in PTAs. (**taylor2021nanohertz**)

This report summarizes the work I have done, while contributing to the Indian Pulsar Timing Array (InPTA) group, under the guidance of Prof. Shantanu Desai. InPTA is an Indo-Japanese consortia consisting of members from NCRA-TIFR, TIFR, IMSC, IIT Hyderabad, IIT Roorkee, BITS Hyderabad, Kumamoto University and led by Prof. B.C. Joshi (NCRA) and has been observing pulsars since 2014. The objective of InPTA is to detect, analyse and study nanoHertz Gravitational Waves. While having monitored a few observations remotely, the bulk of my contribution lies in the analysis of the observed pulsar data.

1   INTRODUCTION

1.1   PULSARS

Pulsars are a special class of neutron star, which themselves correspond to the collapsed M_o cores of massive stars that have undergone supernovae, leaving only small (10-15 km) carcasses with masses of $1 - 2M_\odot$ that are supported against total collapse by neutron degeneracy pressure. They emit beams of radiation from their magnetic poles, which is observed mainly in the radio, X-ray and gamma spectrum.
Pulsars have shed light on strong-field gravity, the equation of state of nuclear matter, evolutionary scenarios for massive binary systems, the structure of the ionized interstellar medium, the existence of exoplanets, and much more.

The "lighthouse model" provides our basic framework for understanding and modeling pulsars as rapidly rotating, highly magnetized neutron stars resulting from stellar collapse. Due to conservation of angular momentum and magnetic flux, these pulsars are far more rapidly spinning and magnetized than their progenitor stars. Their magnetic field (whose
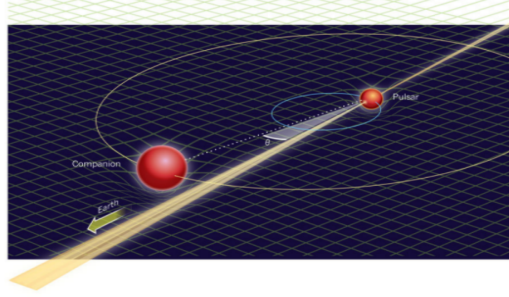

Figure 1: Diagrammatic Representation Shapiro Delay (Source : Manjari Bagchi, InPTA Student Week 2021)

axis may not necessarily align with its rotational axis) is such that the star acts as a rotating magnetic dipole, generating a local electric field along which charged particles within the co-rotating magnetospheric plasma are accelerated. It is expected that these particles excite beams of radio emission high in the pulsar atmosphere that we observe whenever the rotating beam intersects our line-of-sight. The pulse period is then a measure of the rotation period of the pulsar itself.
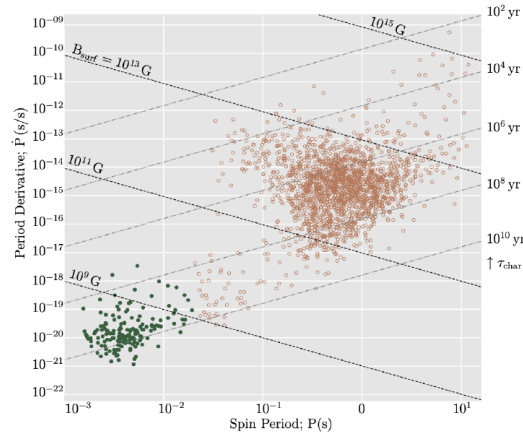

Figure 2: $P - \dot{P}$ diagram of known pulsars. Millisecond pulsars are filled green circles, while canonical pulsars are open circles. The dashed black lines show estimated surface magnetic fields strengths (B_{surf}), while the dot-dashed grey lines show the lines of characteristic age. Data were taken from the ATNF pulsar catalogue [11] version 1.56 (**taylor2021nanohertz**)

1.2 PULSAR TIMING

We observe the pulses of radio emission separated by the observational period of the pulsar. However, the shape of each pulse from one rotation to the next varies randomly, possibly associated with stochasticity in the emission region through which our line of sight traverses. But the pulse shape averaged over the rotations is remarkably stable and reproducible on timescale from minutes to decades. It is this stability at a given radio frequency that permits precision timing; the pulse shape is unique to each pulsar and can be relied upon to mark the passage of rotations when receiving a train of radio pulses. (**taylor2021nanohertz**)
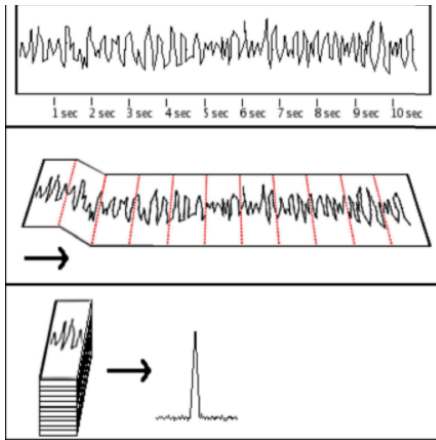

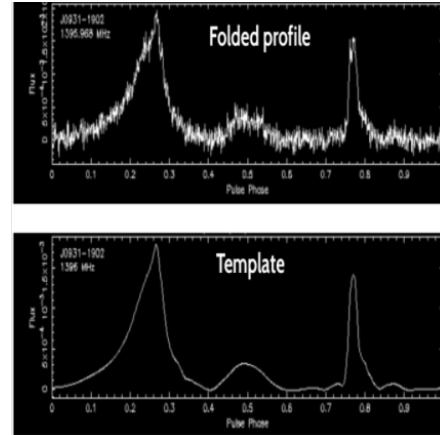Figure 3: Ryan Lynch: Pulsar searching tutorial for PSC

Figure 4: James McKee: Pulsar timing tutorial, IPTA 2018 student week

Pulsar Timing involves tracking the pulsar rotation by measuring pulse arrival times (TOAs). Through these time of arrivals, we calculate the time of emissions at the pulsar. The complications involved in this calculation include:

1. Pulsars slow down, hence the emission times are not constant. We fit our models to obtain the true fit for time of emissions.

2. The space between the pulsar and the Earth is not vacuum, but is the interstellar medium (ISM). This causes the pulse to travel at sub-light-speed, dependent on the frequency of the spectrum. The pulse is composed of a band of frequencies, each of which experiences this phenomena. To obtain the correct results, the whole band is divided into N channels, data recorded at these channels, the added. Signals at different channels arrive at different times - addition will not give a good signal. We need to correct for this dispersive delay to get a good pulse.

3. Romer delay due to the motion of the Earth around the sun.

4. Effects arising out of special and general relativity :

   (a) Curvature of space time due to masses of other planets in our solar system, any companion to the pulsar and the pulsar itself.

   (b) Motion of the pulsar relative to our planet.

The Romer, Einstein, and Shapiro delays all contribute to the difficulties in the analysis. To ensure uniform and complete analysis of the data, a data processing pipeline was developed by the InPTA for use.

1.3 uGMRT : Upgraded Giant Metrewave Radio Telescope

The pulsars are observed by the Upgraded Giant Metrewave Radio Telescope (uGMRT), located near Pune at Khodad and is operated by the National Centre for Radio Astrophysics, TIFR.

The antennas form an interferometric array, which is composed of a total of 30 antennas: 14 in the central square and 16 distributed in a Y-shaped arm. Each antenna is a parabolic reflector having a diameter of 45 metres. The uGMRT provides a frequency coverage from 30 MHz to 1500 MHz. Simultaneous multi-band observations across four bands with high sensitivity with an instantaneous bandwidth of 400 MHz. (**10.2307/26293915**)
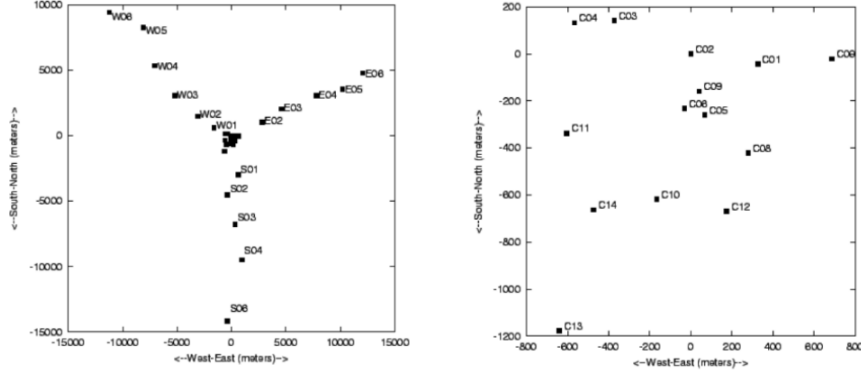

Figure 5: Layout of Antennas at the uGMRT (Credit : NCRA-TIFR)

The uGMRT has four bands which allow for simultaneous wideband observations in multiple bands:

1. BAND 5: 1000 to 1500 MHz
2. BAND 4: 550 to 900 MHz
3. BAND 3: 250 to 500 MHz
4. BAND 2: 120 - 250 MHz

InPTA has observed it's pulsars in bands 3, 4 and 5. Recent cycles have had observations in bands 3 and 5. The 30 antennas of GMRT are split into 2 (or 3) sub-arrays, where each sub-array observes sources in a different band. Observations occur with a bandwidth of 100 or 200 MHz.
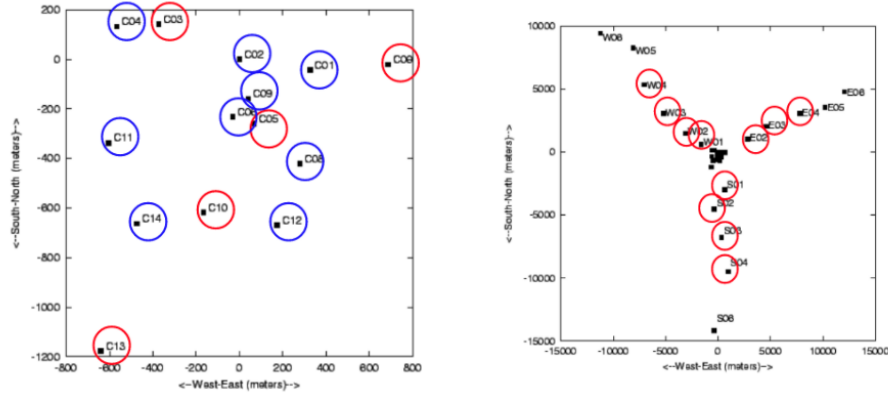

Figure 6: The above image shows an setup with blue antennas observing in BAND3, with red antennas observing in BAND 5. (Credit : NCRA-TIFR)

6

1.4 PINTA

PINTA, which stands for the "Pipeline for the Indian Pulsar Timing Array" is the software developed for the InPTA experiment to address these concerns as well as to improve the efficiency, reliability, and user friendliness of the data reduction process and to ensure faster turnaround time from observations to PTA analysis. This allows for the homogeneity in data reduction practices to avoid non-uniformity in the data products used for PTA analysis, which can introduce systematic errors. (**Susobhanan+2021**)
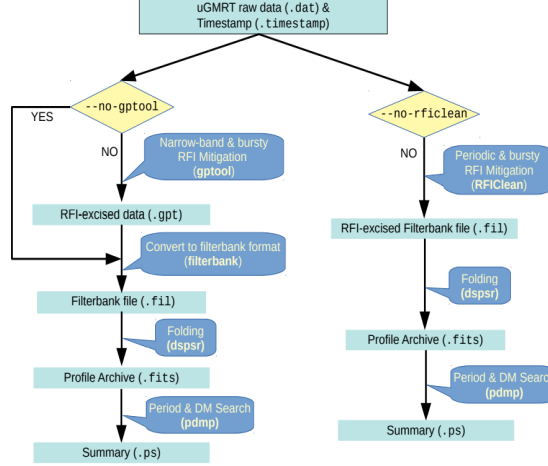

*Figure 7: The workflow of pinta. The pinta pipeline uses uses two separate packages for the RFI mitigation, namely gptool and RFIClean. A typical data reduction workflow can optionally engage these RFI mitigation choices. Note that the profile archives generated by pinta are in the Timer format although their extension is '.fits'. (***Hotan+2004***)*

gptool : It is both an RFI mitigation and a data reduction tool for the beamformer data from GMRT. It mitigates both narrow-band spectral line RFI and broadband bursty time-domain RFI.
RFIClean : It excises periodic RFI in the Fourier domain, and then mitigates narrow-band spectral line RFI and broadband bursty time-domain RFI using robust statistics of conventional RFI mitigation techniques.
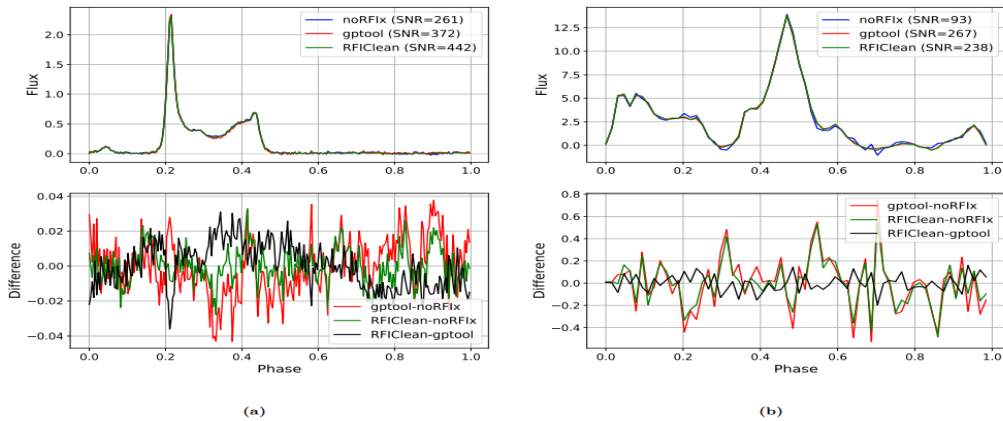

*Figure 8: Comparison of frequency collapsed profiles obtained using gptool, RFIClean, and without any RFI mitigation (noRFIx).Both epochs show significant improvement in SNR while using RFI mitigation. (a) PSR J2145-0750 observed on 16 June 2020 in Band 5 (1260-1460 MHz) with 40.96 μs sampling time and no coherent dedispersion. The total integration time is 55 min. (b) PSR J2124-3358 observed on 25 August 2018 in Band 3 (400-500 MHz) with 81.92 μs sampling time with coherent dedispersion. The total integration time is 24 min (***Susobhanan+2021***)*

1.5 PROCESSING INDIVIDUAL EPOCHS

The processing pipeline takes in .raw/.dat files and timestamps, as obtained from an observation as its input. It also takes in a pipeline.in file containing the metadata parameters with which to process the raw data files. The resulting output is in the form of archive files, a pulsar data format widely used among other PTAs. (**Susobhanan+2021**)
The pipeline is made up of multiple softwares, which perform the functions of de-dispersion, folding, Single Pulse Analysis, RFI cleaning, Template making, ToA generation, Flux calibration and Pulsar Calibration; to obtain the accurate results.

The primary processing of pulsar data results in an output that looks like the following. It can easily be observed that a pulsar was detected in the first observation, while it was not in the second one (note the pulse shape in the bottom plot).
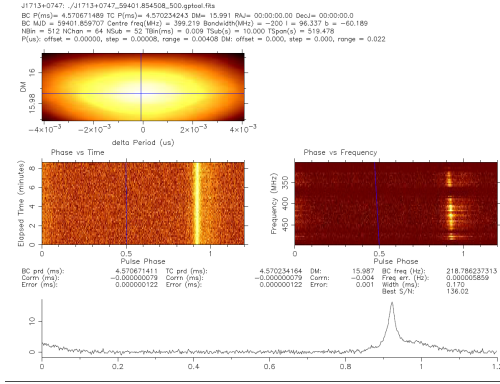

Figure 9: Output from the gptool processing of an observa-tion data on Jul 07 (monitored on by me)
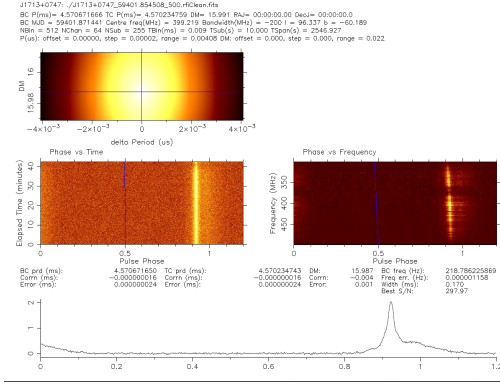

Figure 10: Output from rfiClean processing of an observa-tion data on Jul 07 (monitored on by me)
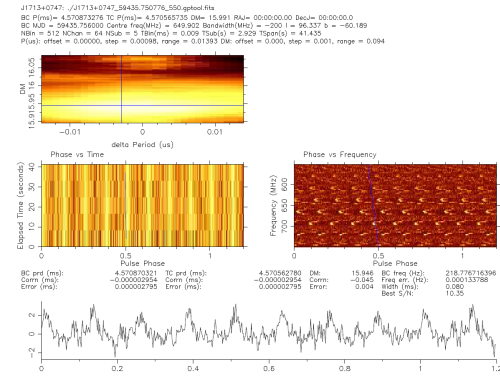

Figure 11: Output from the gptool processing of an obser-vation data on Aug 09 2021 (monitored on by me)
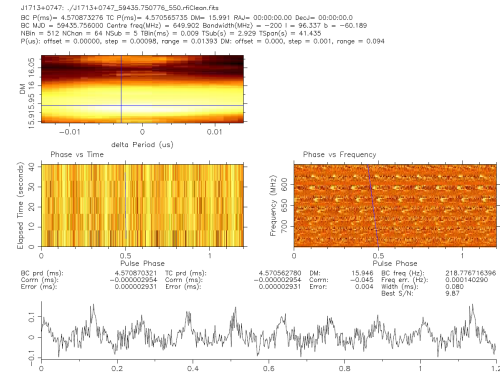

Figure 12: Output from the rfiClean processing of an obser-vation data on Aug 09 2021 (monitored on by me)

The process of observation itself requires careful consideration of the rise and set time of pulsars, optimal choice of phasing sources to calibrate the antennas, awareness of sources of radio frequency interference like satellites, which can pass nearby and affect the observations itself. Such careful measures are necessary in the art of pulsar analysis, and are the reason why most of the observations do not yield unusable results.

1.6 OVERVIEW OF PROCESSING

The figure below summarizes the entire procedure followed from observation to folding, pre-processing, processing to analysis.
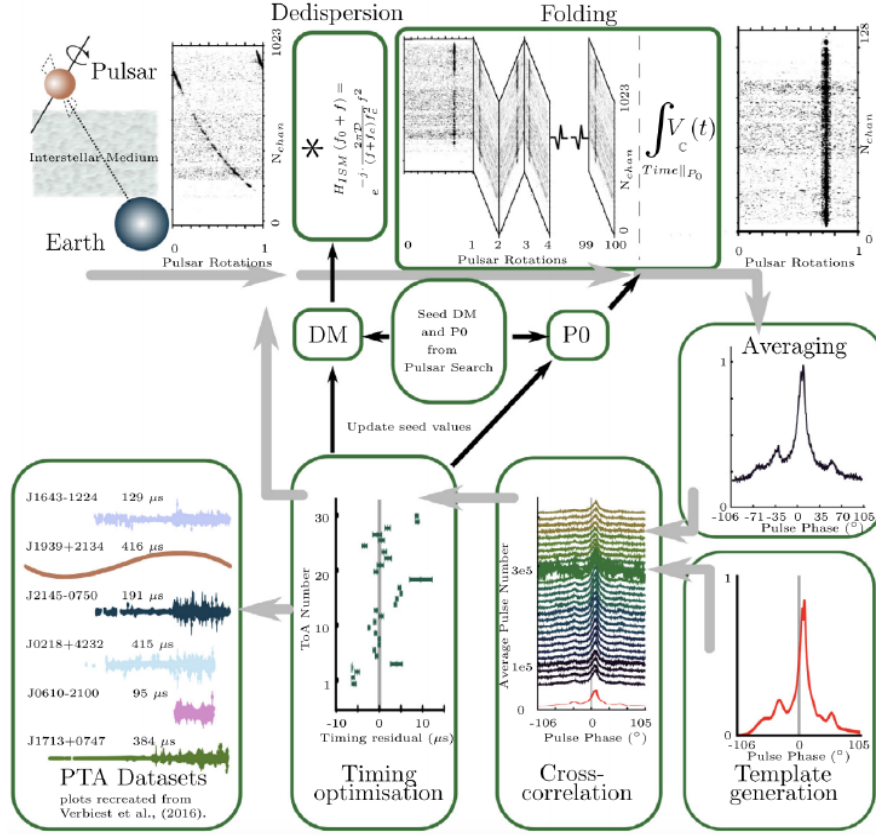

Figure 13: A representation of the main stages involved in the precision timing of pulsars. (**taylor2021nanohertz**)

1.7 CONTRIBUTIONS TO InPTA

My work on INPTA is summarized below:

1. My contributions over the recent months have been the processing of Pulsar Data using the Pinta pipeline, which was used under the team effort to collect evidence for the profile changes in pulsar J1713+0747. I was also involved in a couple of pulsar observations under the same effort. This work has been published in MNRAS **10.1093/mnrasl/slab098**.

2. When it was recognised that there might be an issue with the configuration processes of the Pinta pipeline itself, I was part of the subgroup that was set-up to validate the effects of various configurations of the data processing software.

3. During the same time, it was noticed that there was a discrepancy in the DMs of the processed data. I was given the task to find out the exact difference and report the same (which was then incorporated into the observation process itself). The results of this study have been used by the GMRT team to issue a technical note regarding the frequency shift.

4. I am continuing to participate in observations, data reduction and backup during ongoing Cycle 41 observations.

2   EVIDENCE FOR PROFILE CHANGES IN PSR J1713+0747 USING THE UGMRT

2.1   INTRODUCTION

PSR J1713+0747 is a 4.6 ms pulsar with a low-mass white dwarf companion (**FosterWC+1993**), and has the second lowest measured timing noise behaviour in the PTA ensemble (**pdd+19**), making one of the most precisely timed pulsars in the international pulsar timing array experiment. This pulsar showed an abrupt profile shape change between 2021 April 16, (MJD 59320) and 2021 April 17 (MJD 59321). My contribution in the group effort to report the results involved processing multiple batches of pulsar data.
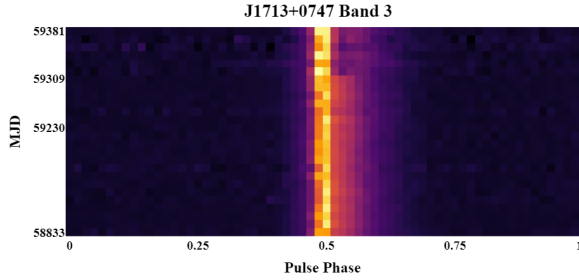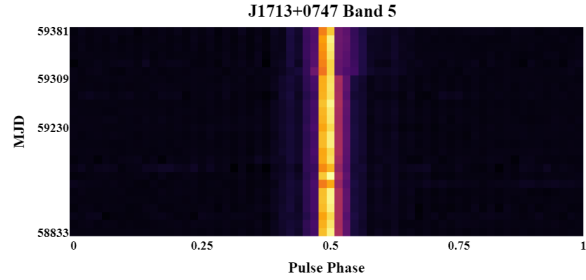

Figure 14: Band 3 Phase Change Plot

Figure 15: Band 5 Phase Change Plot

2.2   OBSERVATIONS

PSR J1713+0747 was observed with the uGMRT (**Gupta+2017**) as part of the InPTA (**JoshiAB+18**) observing program before and after the profile change event reported here, as well as via a Director's Discretionary Time (DDT) project after the event. The InPTA observations were carried out using two sub-arrays of 10 and 15 antennae at 300–500 MHz (Band 3) and 1260–1460 MHz (Band 5), respectively.

All observations used 200 MHz bandwidth. This bandwidth was split into 256 and 128 sub-bands (before and after MJD 59348, respectively) for Band 3 and 1024 sub-bands for Band 5 observations, respectively.

The raw time-series were further reduced to Timer format (**vanStratenBailes2011**) using the pinta pipeline (**Susobhanan+2021**). Further analysis was carried out using the widely-used pulsar software PSRCHIVE (**Hotan+2004**). The time-series were folded using pulsar ephemeris derived from the IPTA Data Release 2 (**Perera+2019**). Out of the two radio-frequency mitigation methods available in pinta, we exclusively used RFIClean (**Maan+2020**)in this work.
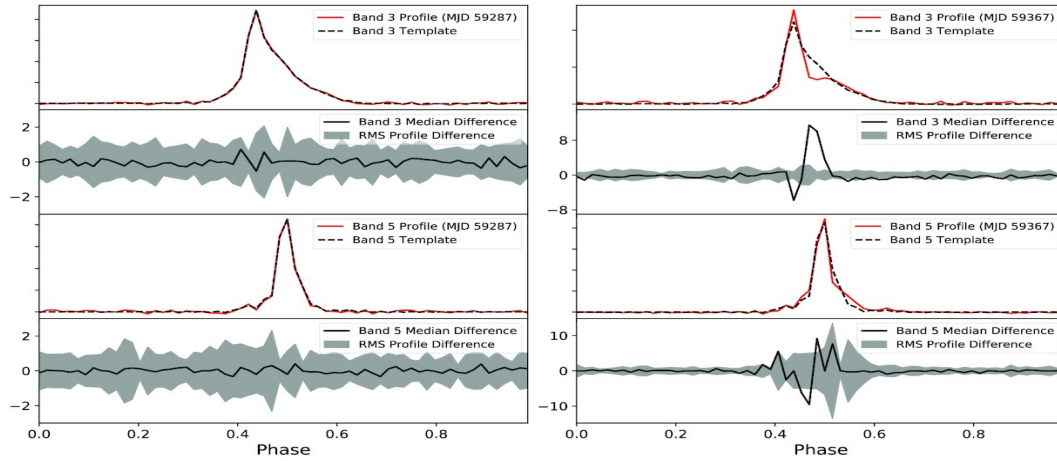
2.3 THE PROFILE CHANGE

We used 20 observing epochs before the event with a cadence of around 14 days, and 6 epochs after the event with cadence ranging from 3–10 days. After de-dispersion and collapsing all the sub-integrations and sub-bands, the folded profiles were normalized with the area under the profile and then stacked together to produce a single multi-epoch file for each of Band 3 and 5. These stacked profiles are shown as color-map plots.
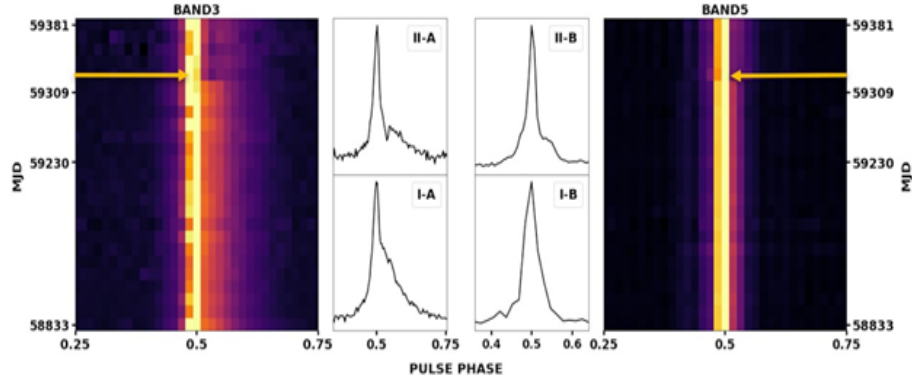
A distinct change is seen in both Band 3 and Band 5 before and after MJD 59321, reported to be the epoch of the event (**Meyers+2021**). The reduced data were collapsed in frequency and time for typical profiles before and after the event for both the bands and these are shown above.

The degree of change in the profile was quantified through the following process. First, a high signal-to-noise ratio (S/N) template profile was formed by averaging all the profiles before the event. Then, the profiles at individual epochs, both before and after the event, were scaled and shifted to align with this template using a frequency-domain cross-correlation method (**Taylor1992**). The scaled and shifted profiles were then subtracted from the template to form profile differences, and were examined for significant deviations. The results of this process can be viewed in the following plot.

2.4 RESULTS

The profile change is seen in both Band 3 and Band 5 data after MJD 59321 with a peak deviation of 12 and 8 times off-pulse RMS, and an enhanced variability in Band 5 profiles after the event. This variability appears to be linked with a jump that is followed by an exponential recovery seen in the timing residuals with a magnitude of 48 μs and a recovery time-scale of 159 days.

We infer that the event is unlikely to be due to a change in the ISM alone as

1. the event appears to be more pronounced in Band 5 than in Band 3 with a chromatic index of -1.34
2. a profile change is seen in both high and low frequency bands

Profile mode changes are believed to arise from a re-organization of pulsar beams, probably due to global changes in the pulsar magnetosphere (**Timokhin2010**), and can exhibit quasi-periodicity (**Lyne+2010**) accompanied by changes in the spin-down rate. This leads to higher timing noise (**arun2017**) in the pulsar, which has the potential to degrade the precision achievable for PTA experiments relying on the high stability of MSPs.

3   PINTA VERSION 6 VALIDATION

3.1   INTRODUCTION

The Data processing pipeline, pinta, receives regular updates from its maintainers. The latest version of the same is version 6, with which a majority of the pulsar data for the year was processed. This processing pipeline runs across multiple production grade servers to help provide maximum work efficiency. However, on upgrading it from version 5 to version 6, a discrepancy was created due to different configuration files being sourced on different servers. This could possibly have affected the processed data.

To check if the above mentioned scenario had really created an issue, a plan was formulated to process data from pulsars with high SNR, using configuration files with different parameters (in local server directories), to check the impact of having different/ no configuration file on the processed data.

This task was treated with the highest priority for the following reason. In the event that this minuscule change actually affected the processed data, the processed data from their entire batch would have become unfit for use. The raw data for an observation is deleted after its processing to make space for new data, hence multiple observations would have been lost due to this issue.

3.2   ACTION TAKEN

The work was divided to completely reprocess specific epochs of two pulsars J1713+0747 and J1939+2134, as they both had high SNRs. The configuration files were to be identical, with a change in the two parameters :

1. Barycentric Correction (Value could be 0 or 1).
2. Rounding Correction (Value could be 0 or 1).

A subset of the configuration file with the relevant parameters can be seen below :

```
#########################################################################
#
# Append :: must_match – Ensure that Archive parameters match [boolean]
#
# If true , then TimeAppend and FrequencyAppend (used by psradd) will fail
# if certain observational parameters do not match (see ArchiveMatch)
#
# Append :: must_match = 1
ArrivalTime :: default_format = Tempo2
#########################################################################
#
# Dispersion :: barycentric_correction – Dedisperse using barycentric freqs [boolean]
#
# If true , interchannel dispersion delays are calculated and corrected
# using barycentric radio frequency values . The historical psrchive
# behavior was to use topocentric frequencies .
#
# Dispersion :: barycentric_correction = 1
#########################################################################
#
# WeightedFrequency :: round_to_kHz – Round mean frequency to nearest kHz [boolean]
#
# If true , then WeightedFrequency (used by tscrunch and fscrunch)
# will round the weighted mean frequency to the nearest value in kHz
#
WeightedFrequency :: round_to_kHz = 0
```

One person was to process without the configuration file. This created 5 possible ways of processing a pulsar. I was to process the J1939+2134 pulsar with Bar 1 and Round 0 configuration. The process involved running pdmp and pat on V5 and V6 Pinta processed J1909 Pulsar, and pdmp, pat and profile comparison on V6 Pinta processed J939 to find any discrepancies. The results obtained were summarized in the following plots.
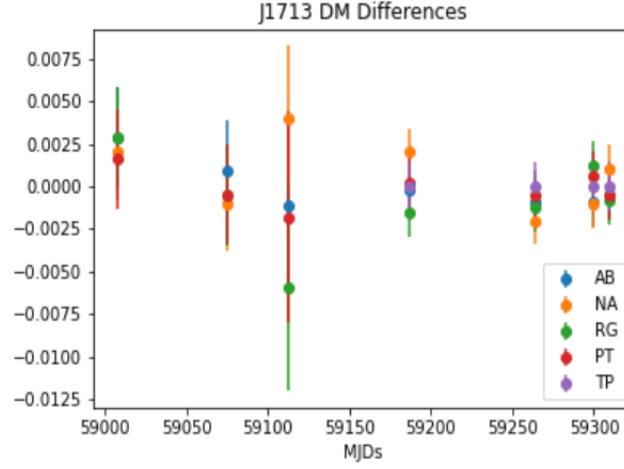
Figure 16: Final Result of DM Differences for Different Configuration files. The legend denotes a separate configuration file for each team member's initials.
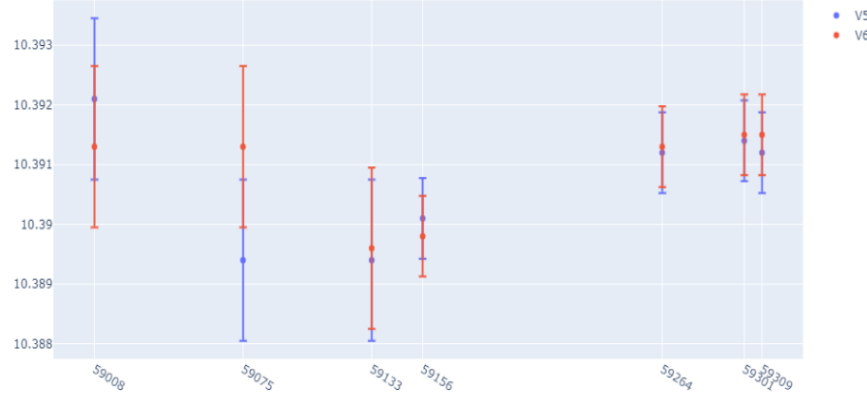
PDMP on J1909 , Bar = 1, Round = 0 (DK)

Figure 17: Result of using the `pdmp` *Analysis on J1909 Pulsar, with configuration as Bar = 1, Round = 0, done by me*
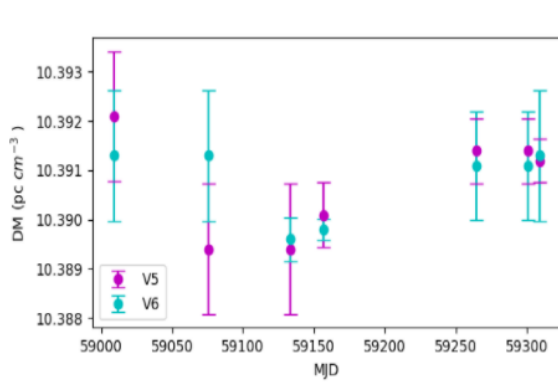

Figure 18: Result of using the `pdmp` *Analysis on J1909 Pulsar, without any configuration*
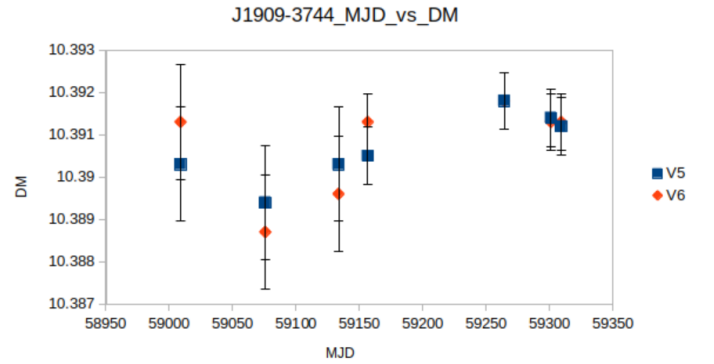
Figure 19: Result of using the `pdmp` *Analysis on J1909 Pulsar, with configuration as Bar = 1, Round = 1*

14

The initial task was succeeded by a complete reduction and processing, including profile comparison of J1713+0747 and J1939+2134 Pulsars. These test were much more explicit, in that they tested the complete processing to rule out any discrepancies outside of the Pinta processing.

DM Stands for Dispersion Measure, BC denotes Barycentric period, TC denotes Topocentric period and SNR denotes Signal to Noise Ratio. The processing results in a data sample output that looks as follows : (Note that this output is for a different pulsar epoch, mentioned as an example)

Best S/N = 204.66
BC MJD = 59106.664774
BC Period (ms) = 4.570535503 TC Period (ms) = 4.570474288 DM = 15.9910
Best BC Period (ms) = 4.570535454 Correction (ms) = -4.94526339e-08 Error (ms) = 7.424341696e-08
Best TC Period (ms) = 4.570474239 Correction (ms) = -4.945197141e-08 Error (ms) = 7.424341696e-08
Best DM = 15.9876 Correction = -0.00341 Error = 0.002095
Best BC Frequency (Hz) = 218.7927454 Error (Hz) = 3.554052078e-06
Pulse width (bins) = 13
Best Pdot = 0 s/s
Best Accn = 0 m/s

3.3 DETERMINATION OF FREQUENCY SHIFT

This was taken up, upon coming across a realization that the DM of all Pulsars showed a jump near January 2021 at the same time, which pointed to an offset in either the processing or the observation semantics. A +10 KHz offset in the local oscillator settings was determined using this methodology. The following steps were followed to check for disparity in frequencies in the process.

1. Find mean DM of J1713 pulsar before the jump event in Band 3.

2. Substitute in the formula for LSB the parameters F_{lo} as 0.5 (GHz), ΔF as 0.2 (GHz), N_{chan} = 256 and calculate F_{1X} which is actually ν_{hi}. (**Susobhanan+2021**)

$$F_{1X} = \begin{cases} F_{\text{LO}} - \frac{\Delta F}{N_{\text{chan}}} & \text{for LSB} \\ F_{\text{LO}} + \Delta F\left(1 - \frac{1}{N_{\text{chan}}}\right) & \text{for USB} \end{cases}$$

3. Use the formula $\nu_{lo} = \nu_{hi} - \Delta F$ to obtain ν_{lo}

4. Substitute in equation below, the values of DM1, ν_{lo} and ν_{hi} as calculated above and calculate Δt (**10.12942/lrr-2008-8**)

$$\Delta t = 4.15 \text{ ms} \times \left[\left(\frac{\nu_{\text{lo}}}{\text{GHz}}\right)^{-2} - \left(\frac{\nu_{\text{hi}}}{\text{GHz}}\right)^{-2}\right] \times \left(\frac{\text{DM}}{\text{cm}^{-3}\text{pc}}\right)$$

Now, to find if there is a change in frequency, we will reverse this process with the DMs after the jump and obtain the

1. Now find mean DM2 of J1713 pulsar after the DM jump.

2. Substitute Δ t calculated in the previous process, this DM2 and calculate the factor in square bracket equation mentioned above.

3. Obtain the ν_{lo} and ν_{hi} values. This was done using the `sympy` python library in python to solve the equation. The change in frequency is the difference in the values of F_{LO}.

4. Repeat the same for J1909 pulsar, the results should be same for two pulsars in case of no frequency change.

3.4 RESULTS

Table 1: Results obtained from calculation of frequency shift process

	Pre-jump DM	Δt(ms)	Post-jump DM	F_{LO}(GHz)	ΔF(GHz)
J1713 (V5)	15.998783	475.16733824824945	15.997628	0.499991180113022	8.819886977984304e-06
J1713 (V6)	15.988197	474.8529317435362	15.986748	0.499988927659255	11.07234074498864e-06
J1909 (V5)	10.391595	308.6326338887038	10.391000	0.499993004791562	6.995208437998723e-06
J1909 (V6)	10.391048	308.6163878696146	10.390474	0.499993251330391	6.74866960898024e-06

3.5 CONCLUSIONS

The main conclusion was that V6 Analysis was valid and can be used for science.
It was seen that pdmp DMs and pat frequency match for all 4 cases. This suggested that .psrchive.cfg settings do not matter for DSPSR, but these matter in post-processing by psrchive (**Hotan+2004**). For correct post-processing by psrchive, Bar 1 Round 0 config file is recommended.

The GMRT Operations team issued a technical note which summarised the situation and its resolution as follows :
The DM jump appears to be due to a frequency shift in LO of +10 KHz in Cycle 39, Cycle 40. For beam mode observations, The LO offset of 10 KHz introduces a systematic shift in frequencies interpreted by the analysis software. This can lead to an error in the DM estimates between 0.0005 to 0.001 pccm^-3 and a residual DM smear in the time-series due to dedispersion to incorrect frequencies. This can also introduce a small shift in the time-of-arrival of data. The magnitude of these errors depend on DM of the source being observed and will vary from one source to another. It also depends on observing frequency. The most significant effect is likely in Band 3 and 4, while Band 5 data is not affected except for very large DM sources. The header for all data taken with TGC system prior to 29 July 2021 needs to be corrected by adding +10 KHz to the frequency in the header before carrying out the data reduction.

4   ACKNOWLEDGMENTS