`pinta`*: A uGMRT pulsar pipeline

Abhimanyu Susobhanan, TIFR, Mumbai

IPTA Catch-up meeting 2020

* **Pi**peline for **In**dian Pulsar **T**imining **A**rray

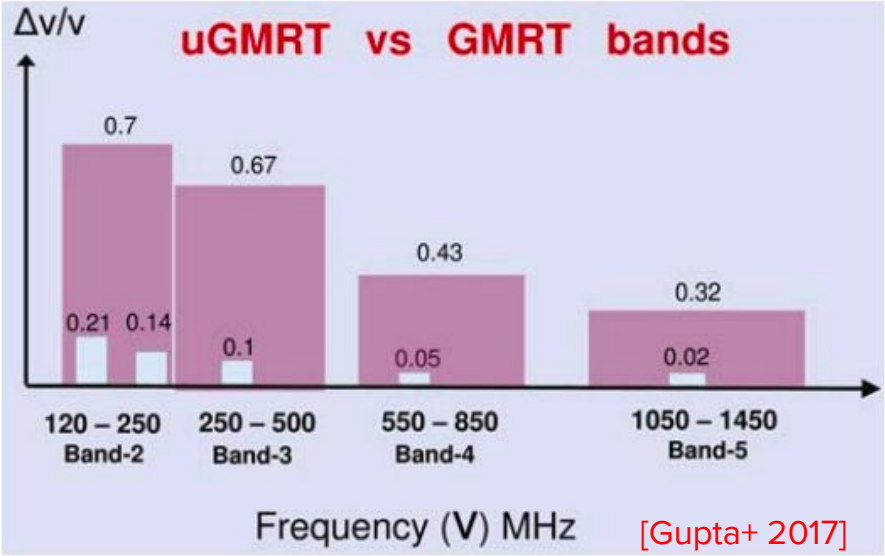
InPTA
Indian Pulsar Timing Array

The Indian Pulsar Timing Array

- We observe a subset of PTA pulsars with bi-weekly cadence using GMRT and ORT.
- Operational since 2015
- Aim: To complement international PTA efforts with unique strengths of GMRT
    - GMRT is an ideal instrument for studying ISM variability, which is an important source of noise in PTAs.

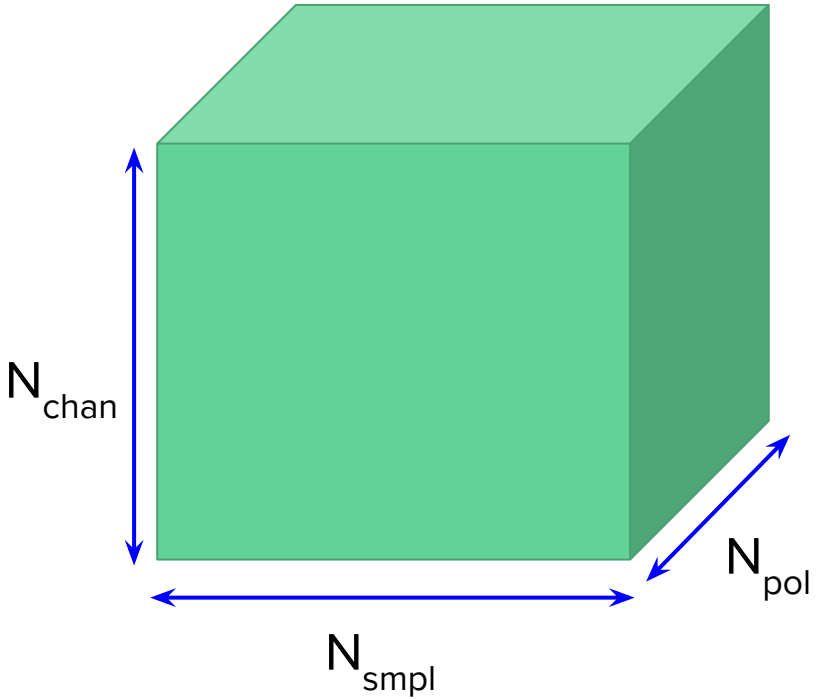
The upgraded GMRT

- 30-element interferometer
- New wide-band feeds provide seamless frequency coverage between 150-1460 MHz
- 4 phased array beams – simultaneous multi-frequency observations.
- A new GPS-synchronized hydrogen maser for precision timing – 1-10 ns precision timing.



[Gupta+ 2017]

uGMRT Pulsar Data Format



Data Volume = N_{chan} * N_{smpl} * N_{pol} * 2 bytes

- The binary raw data file stores **Npol** polarization products in **Nchan** channels for every time sample as **16-bit integers**.
  - Npol = 1 : Total intensity
  - Npol = 4 : Stokes I,Q,U,V
- The timestamp at the start of observation is stored as an ASCII file.

```
#Start time and date
IST Time: 19:59:57.633098240
Date: 25:08:2018
#Start ACQ SEQ NO = 17
```

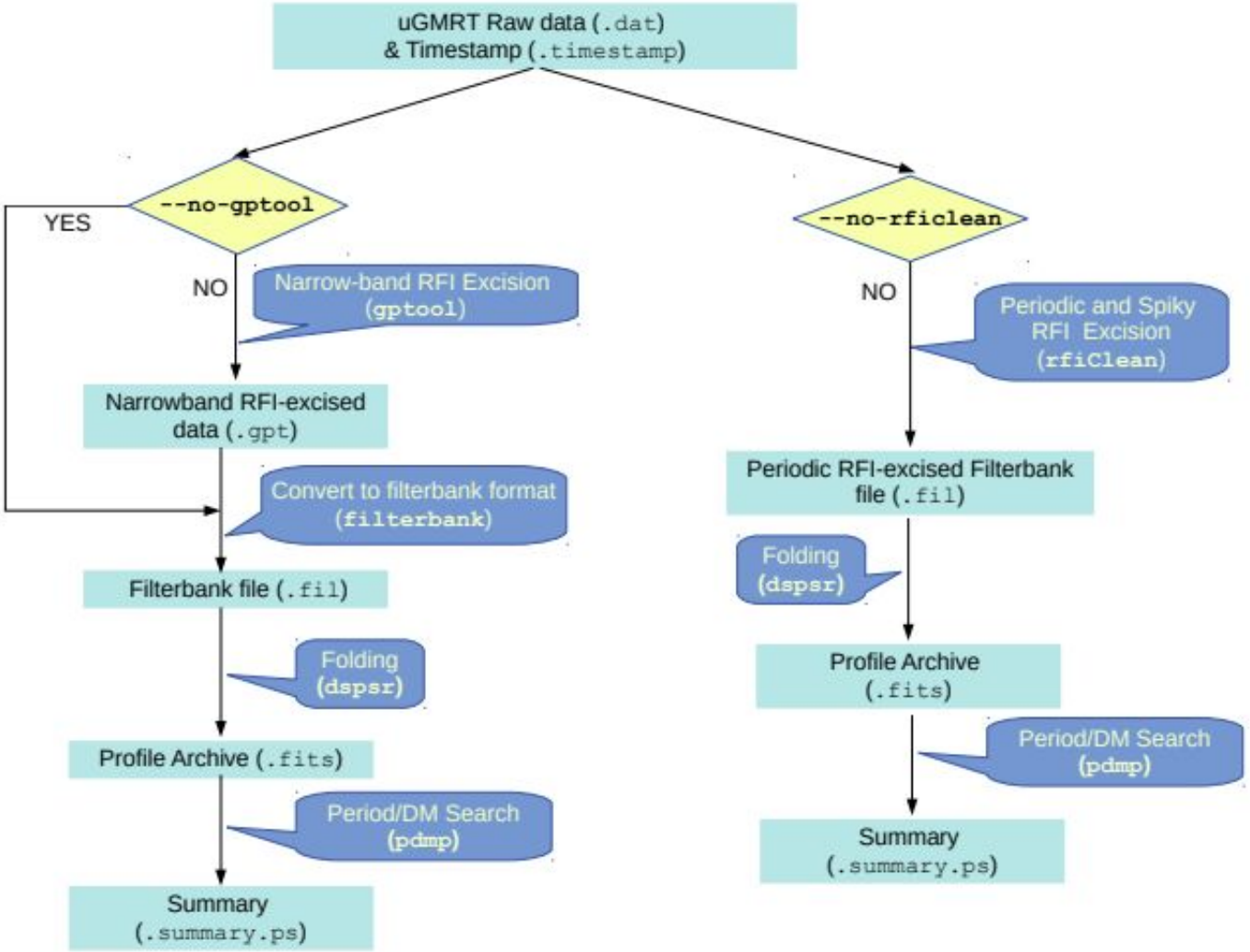
Motivation for developing `pinta`

- uGMRT raw data files do not contain metadata required for analysis.
  - Eg: Source name, Observation frequency, Bandwidth
- All InPTA datasets must be analyzed using a uniform method to avoid inconsistencies and systematic errors.
- Data analysis involves the usage of multiple software packages.
  - Hard to keep track of all the settings and command line options when multiple people are analyzing the data.
- `pinta` is installed in a GMRT server to minimize data transfer and it will also be installed at the new InPTA storage server at TIFR

`pinta` Overview

- pinta is a python script which calls various pulsar data processing codes to reduce the GMRT pulsar raw data to a folded Timer archive.
  - Timer format is compatible with packages used for downstream processing.
- Performs RFI mitigation using two different packages.
  - **`rfiClean`** : Removes periodic RFI (Eg: 50 Hz power line) and time-localized RFI [Maan et al., in prep]
  - **`gptool`** : Removes RFI localized in time or frequency [Chowdhary & Gupta, In prep]
- Metadata required for processing is provided as an ASCII input file.

Workflow


uGMRT Raw data (.dat)
& Timestamp (.timestamp)
--no-gptool
YES
NO
Narrow-band RFI Excision
(gptool)
Narrowband RFI-excised
data (.gpt)
Convert to filterbank format
(filterbank)
Filterbank file (.fil)
Folding
(dspsr)
Profile Archive (.fits)
Period/DM Search
(pdmp)
Summary
(.summary.ps)
--no-rficlean
NO
Periodic and Spiky
RFI Excision
(rfiClean)
Periodic RFI-excised Filterbank
file (.fil)
Folding
(dspsr)
Profile Archive
(.fits)
Period/DM Search
(pdmp)
Summary
(.summary.ps)

Usage

```
$ pinta [--help] [--test] [--no-gptool] [--no-rficlean]
[--nodel] [--retain-aux] [--log-to-file] [--gptdir <...>]
[--pardir <...>] [--rficconf <...>] <input_dir>
<working_dir>
```

- `<input_dir>` contains input raw data files.
- Output will be written to `<working_dir>`
- Metadata is communicated to the pipeline through `<working_dir>/pipeline.in`

Metadata

#JName	RawData	Timestamp	Freq	Nbin	NChan	BandWidth	TSmpl	SB	NPol	TSubint	Cohded
J1939+2134	J1939+2134.25032019.B3.cdp.dat	J1939+2134.25032019.B3.cdp.timestamp	500	128	1024	100	0.00008192	LSB	1	10.0	1
J1939+2134	J1939+2134.25032019.B4.pa.raw	J1939+2134.25032019.B4.pa.hdr	750	128	1024	100	0.00008192	LSB	1	10.0	0
J1939+2134	J1939+2134.25032019.B5.cdp.dat	J1939+2134.25032019.B5.cdp.timestamp	1460	128	1024	100	0.00008192	LSB	1	10.0	1

Column	Parameter	Description	Data Type	Unit
1	JName	The name of the pulsar in J2000 epoch.	String	
2	RawDataFile	Raw data file name. Only the file name is required and not the full path.	String	
3	TimestampFile	Timestamp file name. Only the file name is required and not the full path.	String	
4	Frequency ($F_{\rm LO}$)	Local oscillator frequency of the observing band.	Float	MHz
5	NBins ($N_{\rm bin}$)	Number of phase bins for the folded profile.	Integer	
6	NChans ($N_{\rm chan}$)	Number of frequency channels.	Integer	
7	BandWidth (ΔF)	Bandwidth of the observing band.	Float	MHz
8	TSample ($T_{\rm smpl}$)	The sampling time used for observation.	Float	s
9	SideBand	The side-band. This should be either LSB (lower side-band) or USB (upper side-band).	String	
10	NPol ($N_{\rm pol}$)	Number of polarizations (1:=(I), 4:=(I,Q,U,V))	Integer	
11	TSubInt ($T_{\rm subint}$)	The duration of individual sub-integrations within which the data will be folded over the pulsar period.	Float	s
12	Cohded	Whether the data has been coherently dedispersed (De & Gupta, 2016). 1 represents Yes and 0 represents No.	Boolean	

Sample output

```
bash-4.2$ cat pinta.20200713T182709.log
[INFO] pinta invoked by asusobhanan at 2020-07-13T18:27:09 on tapti.ncra.tifr.res.in
[                                    /asusobhanan/Work/V5/pinta for permissions... OK...
[                                    INPTA/Pipeline_Tests for permissions... OK...
[                                    INPTA/Pipeline_Tests/pinta_mod_tests for permissions... OK...
[CHECK] Checking directory /Data/bcj/INPTA/Pipeline_Tests/pinta_mod_tests/wdir for permissions... OK...
[LOCK] Creating lock file...
[CMD] touch /Data/bcj/INPTA/Pipeline_Tests/pinta_mod_tests/wdir/pinta.lock
[CONFIG] Reading config from /misc/home/asusobhanan/Work/V5/pinta/pinta.yaml
[CONFIG] Will run gptool.
[CONFIG] Will run rficlean.
[CHECK] Checking directory /Data/bcj/INPTA/30june2018/gwbh7/parfilesinpta for permissions... OK...
[CHECK] Checking directory /misc/home/asusobhanan/Work/gptool_files for permissions... OK...
[CHECK] Checking file /misc/home/ymaan/inpta_pipeline/inpta_rficlean.flags for permissions... OK...
[CHECK] Checking for current group... OK... Current group is ugmrtpsr.
[CONFIG] Will remove intermediate products.
[CHECK] Checking for dspsr... OK...
[CHECK] Checking for filterbank... OK...
[CHECK] Checking for tempo2... OK...
[CHECK] Checking for pdmp... OK...
[CHECK] Checking for gptool... OK...
[CHECK] Checking for crp_rficlean_gm.sh... OK...
[CHECK] Checking file /misc/home/asusobhanan/Work/gptool_files/gptool.in.499 for permissions... OK...
[CHECK] Checking file /misc/home/asusobhanan/Work/gptool_files/gptool.in.749 for permissions... OK...
[CHECK] Checking file /misc/home/asusobhanan/Work/gptool_files/gptool.in.1459 for permissions... OK...
[CHECK] Checking file /Data/bcj/INPTA/Pipeline_Tests/pinta_mod_tests/wdir/pipeline.in for permissions... OK...
[INPUT] Reading /Data/bcj/INPTA/Pipeline_Tests/pinta_mod_tests/wdir/pipeline.in... Done. 1 item(s) to be processed.
[CONFIG] Will remove auxiliary files.
[CHECK] Checking file /Data/bcj/INPTA/Pipeline_Tests/pinta_mod_tests/J2124-3358.25august2018.band3cdp.dat.dat.1G for permissions
... OK...
[CHECK] Checking file /Data/bcj/INPTA/Pipeline_Tests/pinta_mod_tests/J2124-3358.25august2018.band3cdp.dat.timestamp for permissi
ons... OK...
[INPUT] The timestamp is MJD 58355.670067517540740
[CHECK] Checking file /Data/bcj/INPTA/30june2018/gwbh7/parfilesinpta/J2124-3358.par for permissions... OK...
[INPUT] Pulsar spin-frequency found :  202.793894
[INFO] Entering working directory.
```

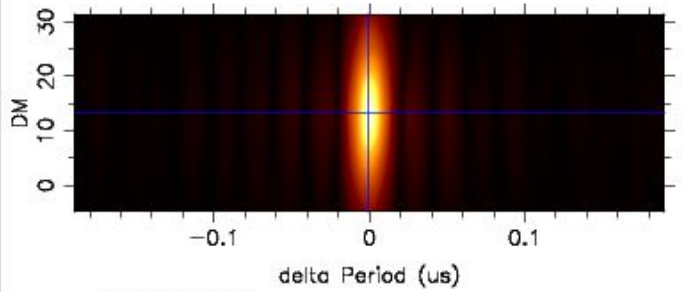
J1857+0943: J1857+0943.58355.60413927197.1459.90234375.32768M.rflClean.fits

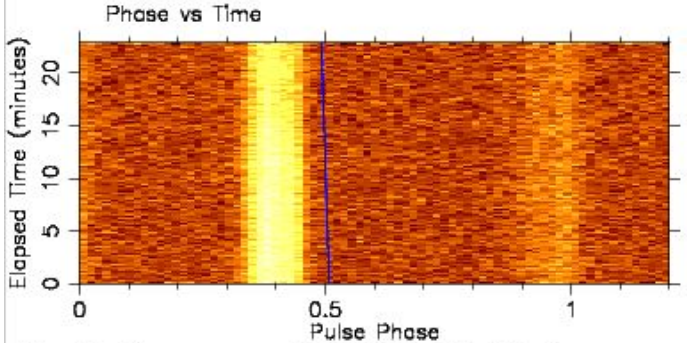
BC P(ms)= 5.362952698 TC P(ms)= 5.362704506 DM= 13.299 RAJ= 00:00:00.00 DecJ= 00:00:00.0

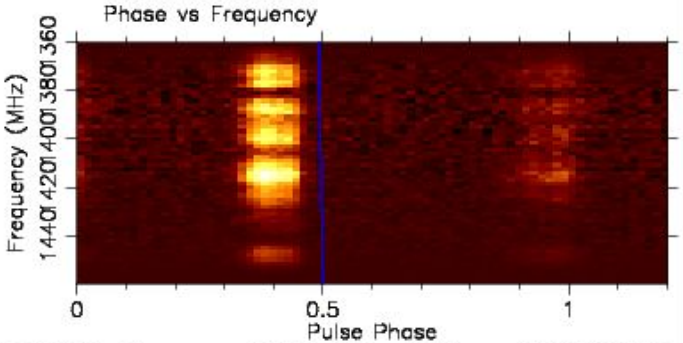
BC MJD = 58355.617996 Centre freq(MHz) = 1409.951 Bandwidth(MHz) = −100 l = 96.337 b = −60.189

NBin = 64 NChan = 64 NSub = 139 TBin(ms) = 0.084 TSub(s) = 2.367 TSpan(s) = 1374.388

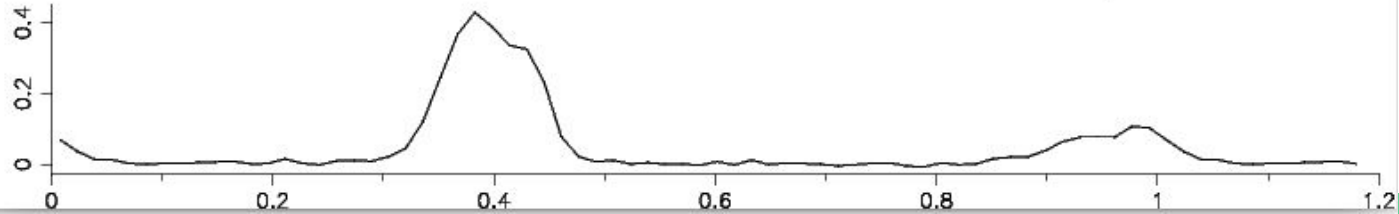
P(us): offset = 0.00000, step = 0.00033, range = 0.18985 DM: offset = 0.000, step = 0.140, range = 18.110



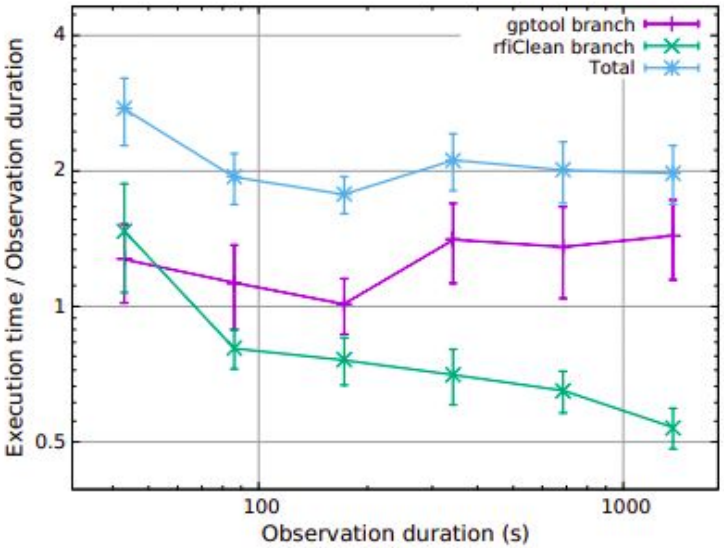


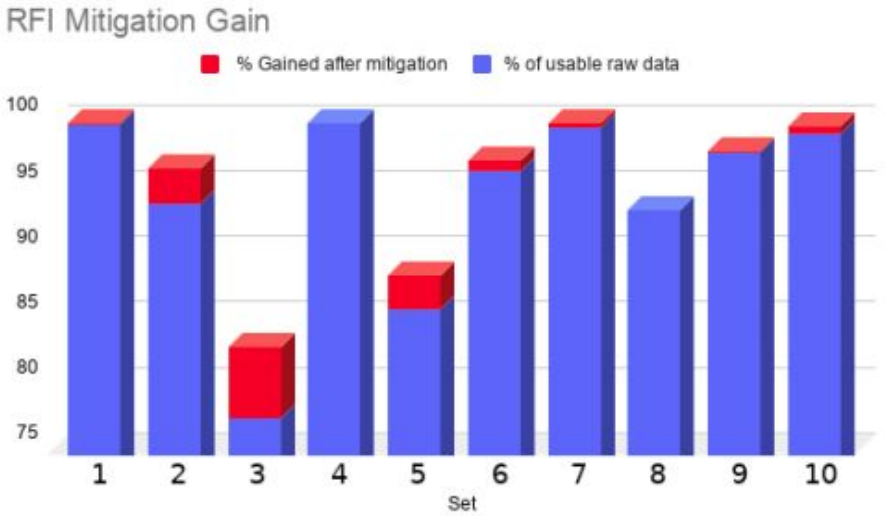


BC prd (ms):	5.362952371	TC prd (ms):	5.362704179	DM:	13.159	BC freq (Hz):	186.464456665
Corrn (ms):	−0.000000327	Corrn (ms):	−0.000000327	Corrn:	−0.140	Freq err. (Hz):	0.000017102
Error (ms):	0.000000492	Error (ms):	0.000000492	Error:	0.431	Width (ms):	0.503
						Best S/N:	203.23



Performance





pinta: The uGMRT Data Processing Pipeline for the Indian Pulsar Timing Array

Abhimanyu Susobhanan, Yogesh Maan, Bhal Chandra Joshi, T. Prabu, Shantanu Desai, Yashwant Gupta, A. Gopakumar, Neelam Dhanda Batra, Arpita Choudhary, Mayuresh P. Surnis, Lankeswar Dey, Jaikhomba Singha, K. Nobleson, Manjari Bagchi, Avishek Basu, Suryarao Bethapudi, Kishalay De, Raghav Girgaonkar, M. A. Krishnakumar, P. K. Manoharan, Arun Kumar Naidu, Dhruv Pathak, Sai Chaitanya Susarla

We introduce pinta, a pipeline for reducing the uGMRT raw pulsar timing data developed for the Indian Pulsar Timing Array experiment. We provide a detailed description of the workflow and usage of pinta, as well as its computational performance and RFI mitigation characteristics. Furthermore, the results of a calibration experiment carried out to determine the relative time offsets between different back-end modes and the correct interpretation of the observatory frequency settings at the uGMRT, which are crucial for performing precision pulsar timing, are also discussed.

Subjects: **Instrumentation and Methods for Astrophysics (astro-ph.IM)**; High Energy Astrophysical Phenomena (astro-ph.HE)
Cite as: arXiv:2007.02930 **[astro-ph.IM]**
(or arXiv:2007.02930v1 **[astro-ph.IM]** for this version)

To be submitted.

Code Repository: https://github.com/abhisrkckl/pinta

Summary

- Developed a pipeline to reduce uGMRT pulsar raw data to folded archives
- RFI mitigation using two different tools: `gptool` & `rfiClean`
- Standardizes InPTA data reduction process
- We plan to deploy `pinta` as a **GMRT observatory pipeline.**

Thank you