MANCHESTER
1824
The University of Manchester

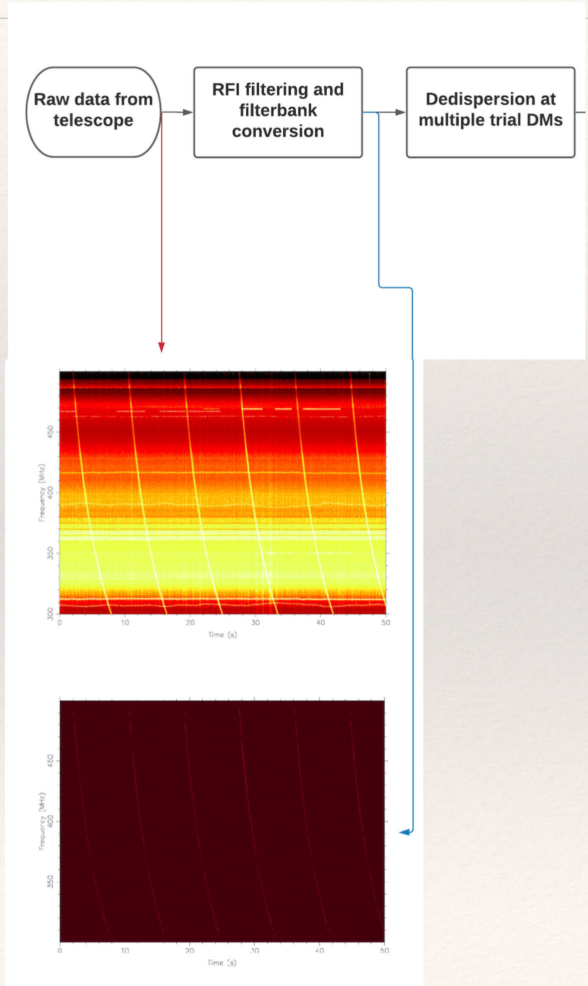
Ben Stappers, University of Manchester

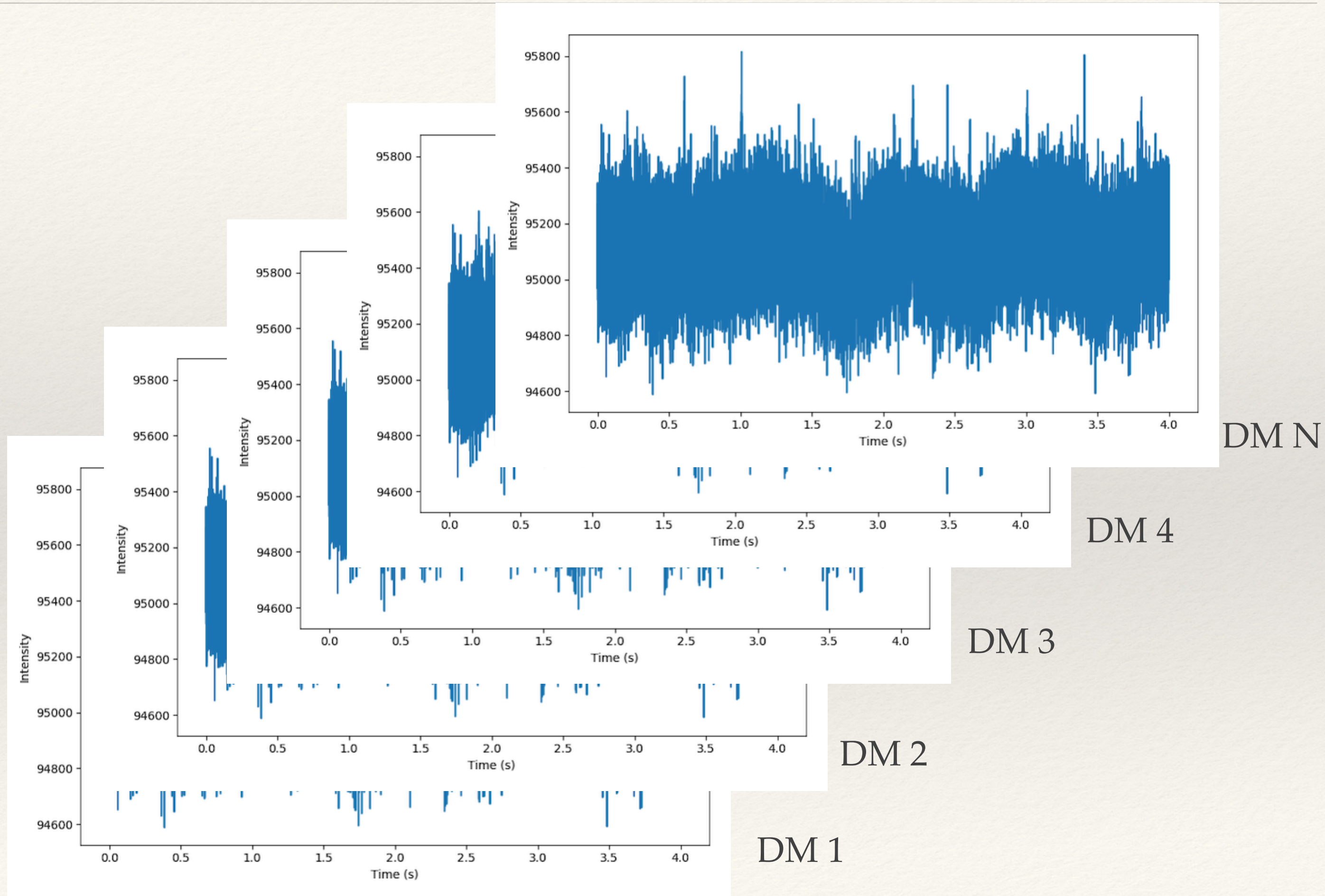
Pulsars and Fast Transients:

Find them, study them, understand them (and use them)!

DARA Lectures

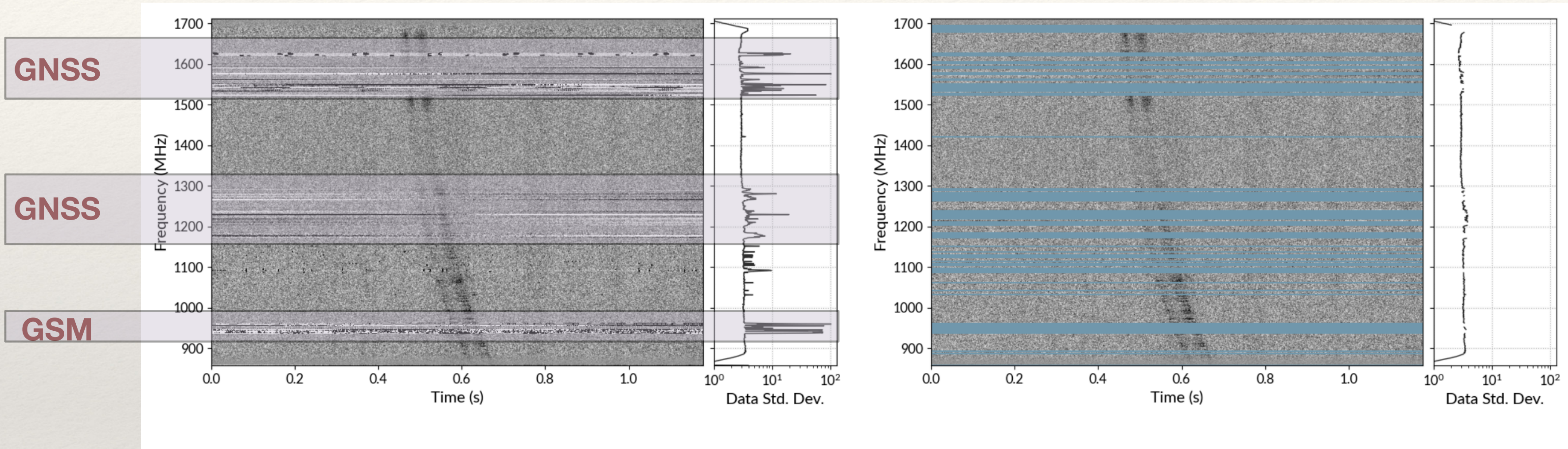
How do we search for single pulses?





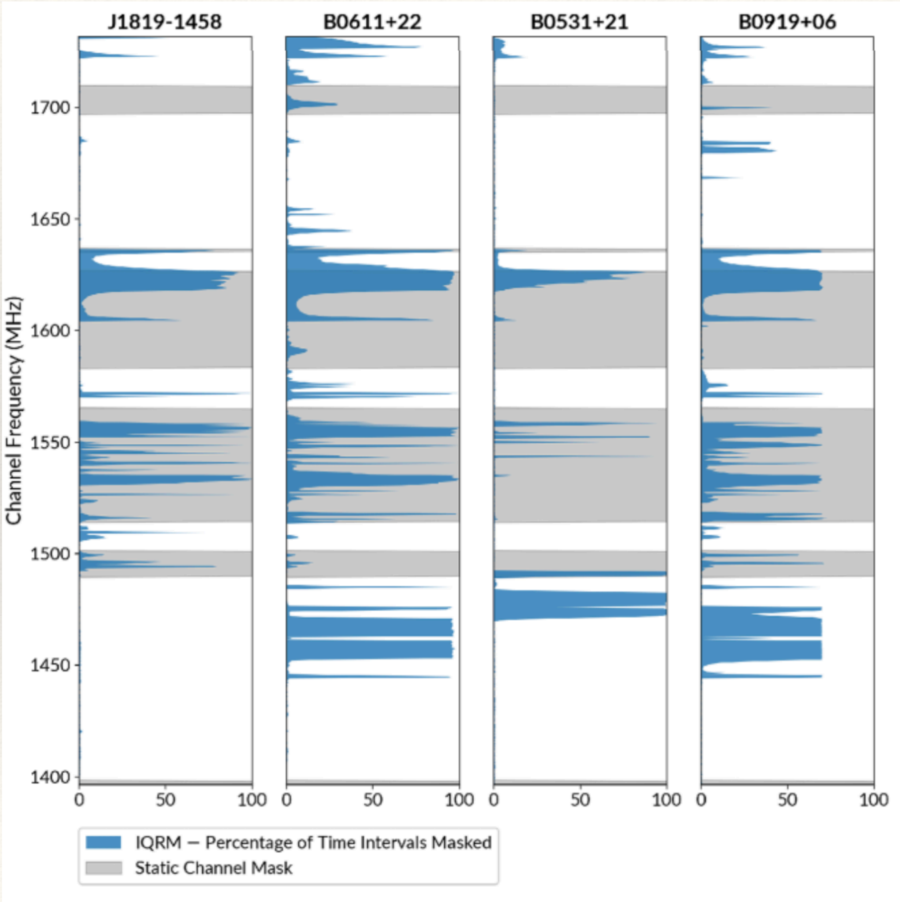
Radio Frequency Interference (RFI)

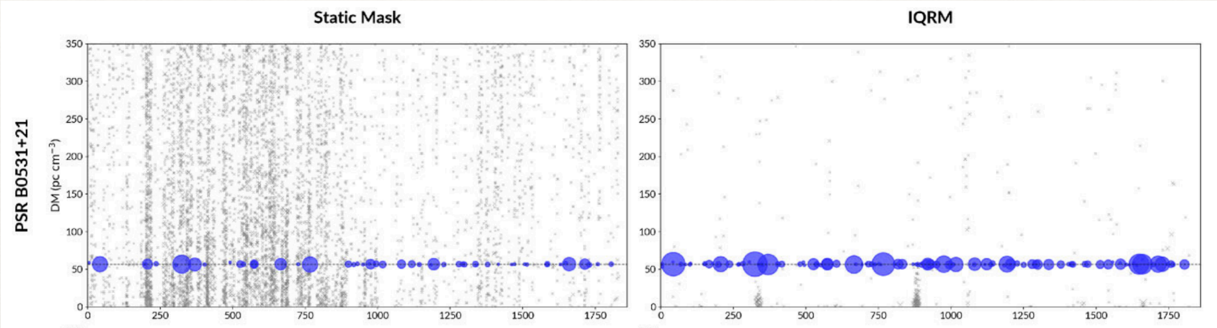
RRAT J1226–3223



Morello et al. 2022

Radio Frequency Interference (RFI)

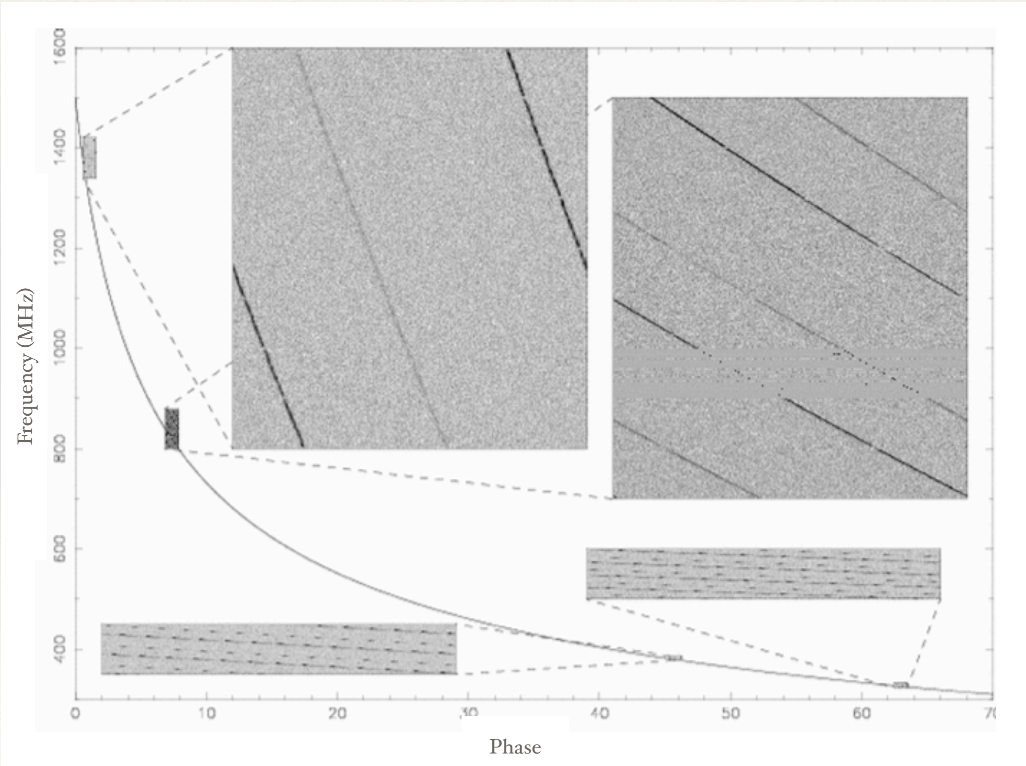




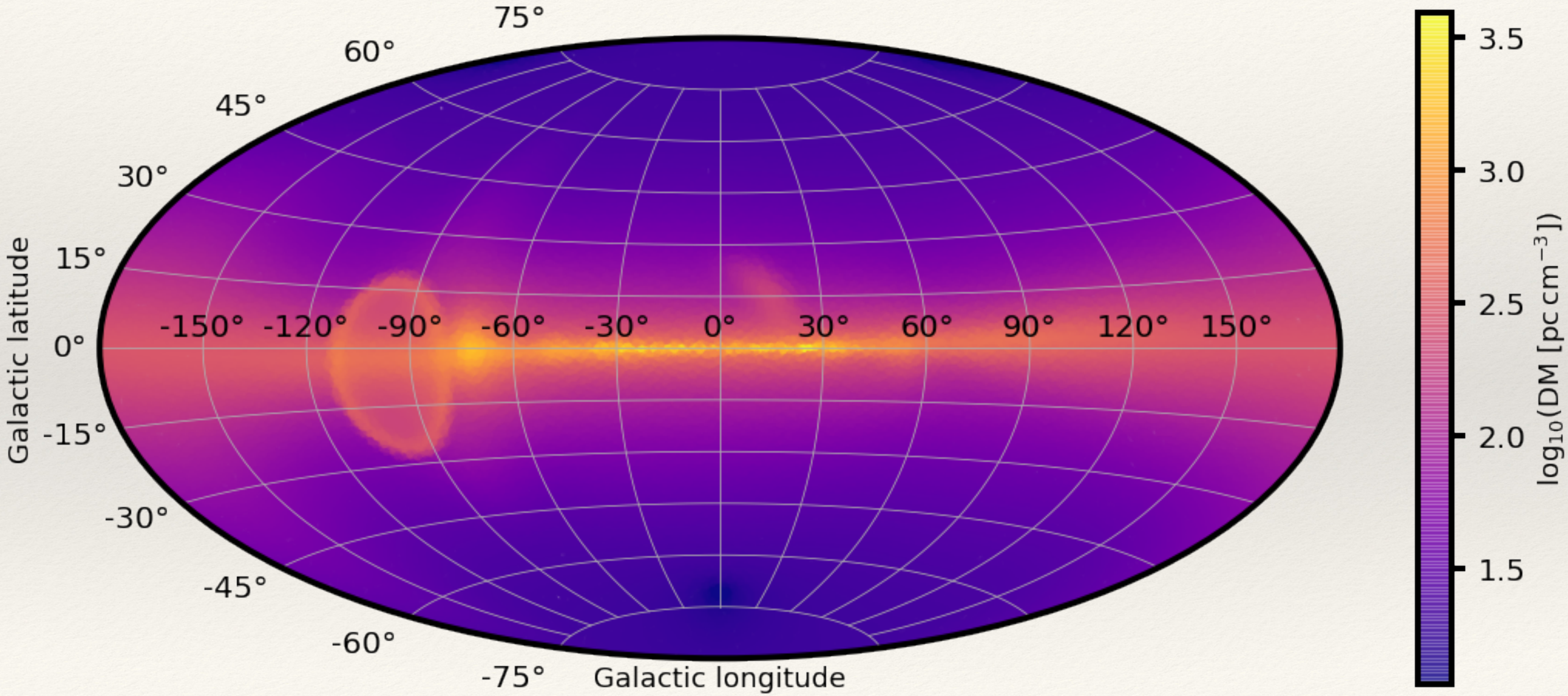
Morello et al. 2022

Dedispersion

$$\Delta t = 4.15 \times 10^3 \times \mathrm{DM} \times \left[\frac{1}{\nu_{\mathrm{low}}^2} - \frac{1}{\nu_{\mathrm{high}}^2}\right]$$



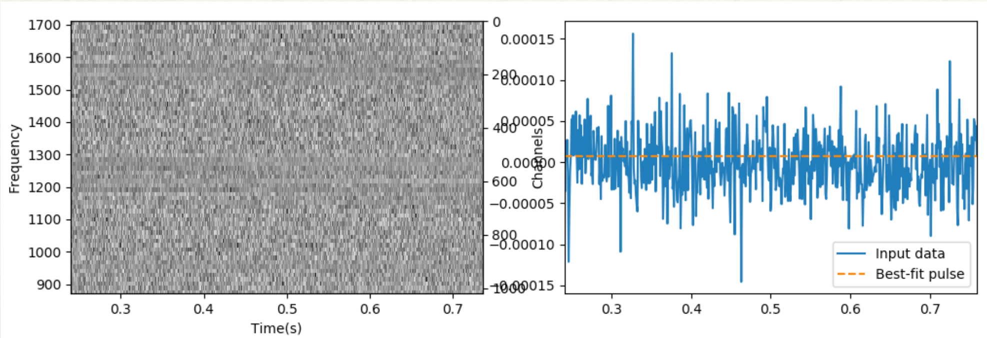
Galactic Electron Distribution



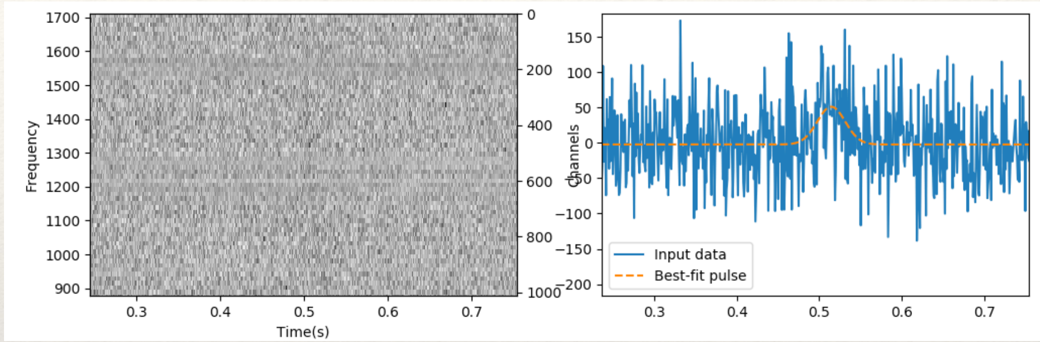
YMW16 & Driessen & MeerTRAP

Dispersion Measure Trials

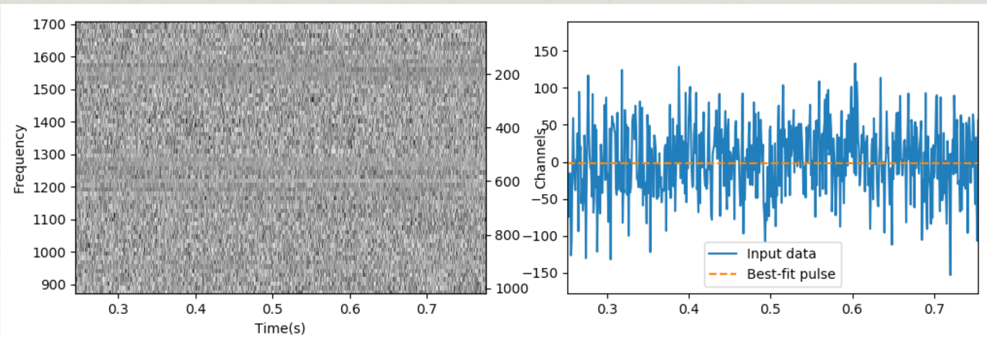
DM = 0.0



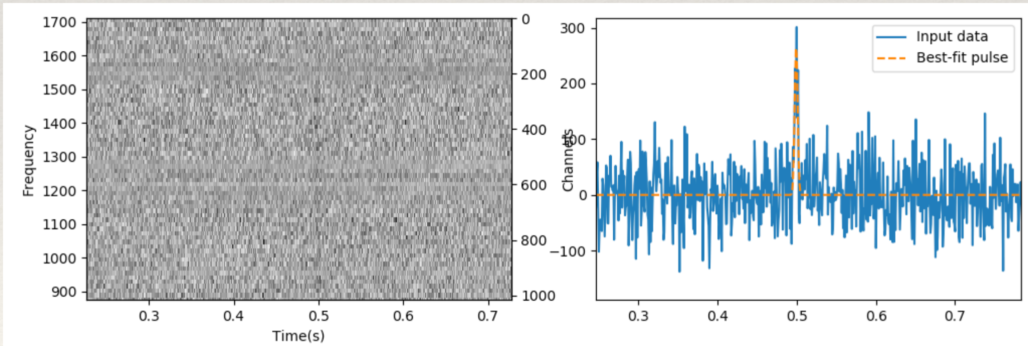
DM = 190.0



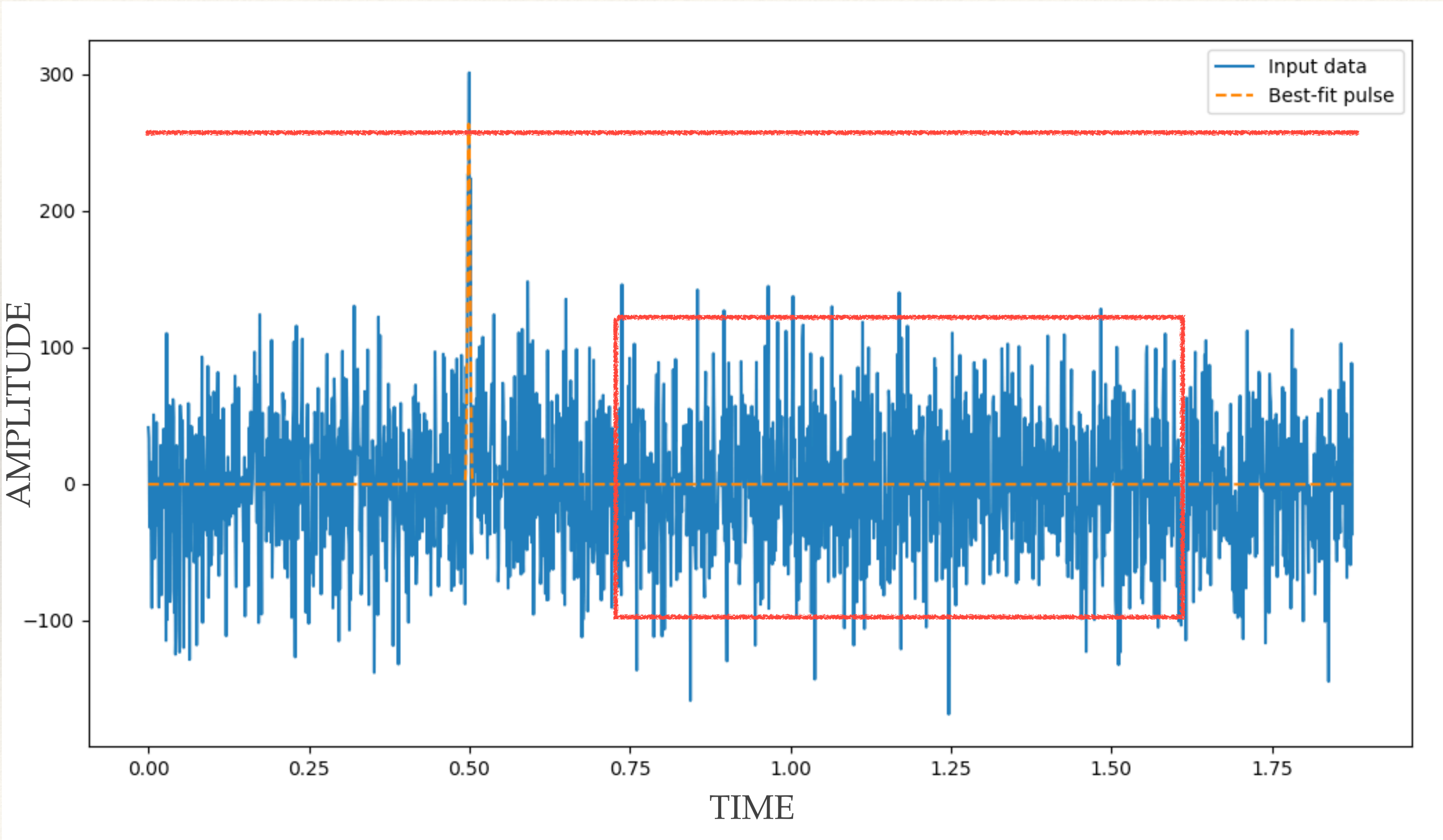
DM = 100.0



DM = 206.6



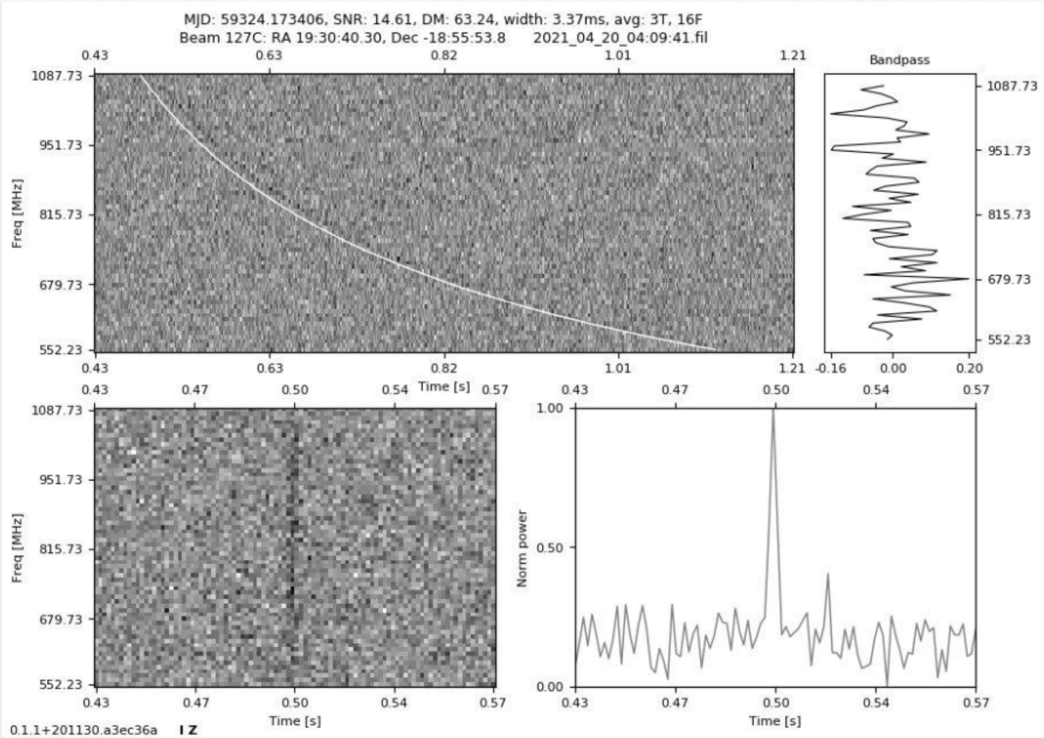
Single Pulse Detection



Set threshold at appropriate level given number of trials.

Calculate statistics

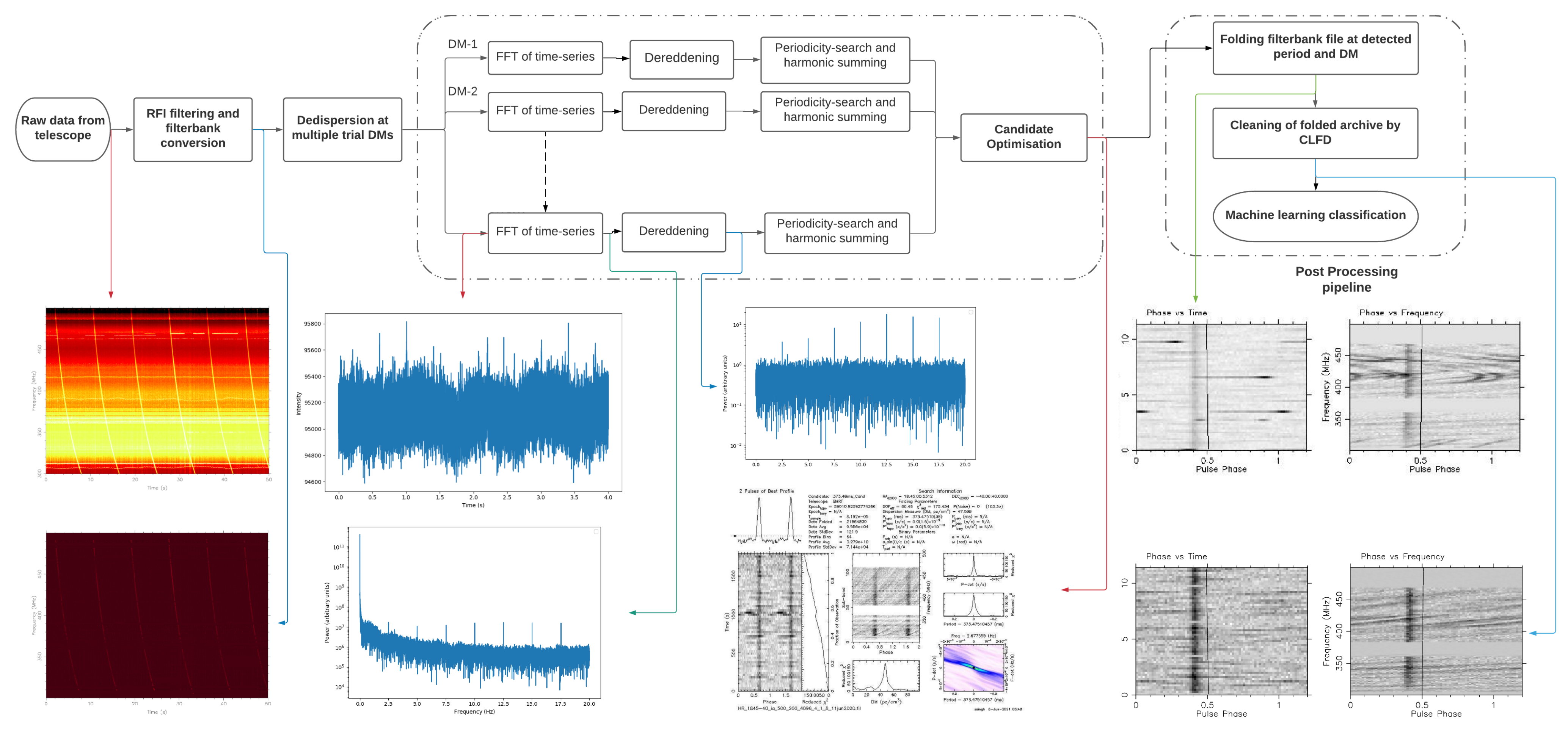
Candidate Plots



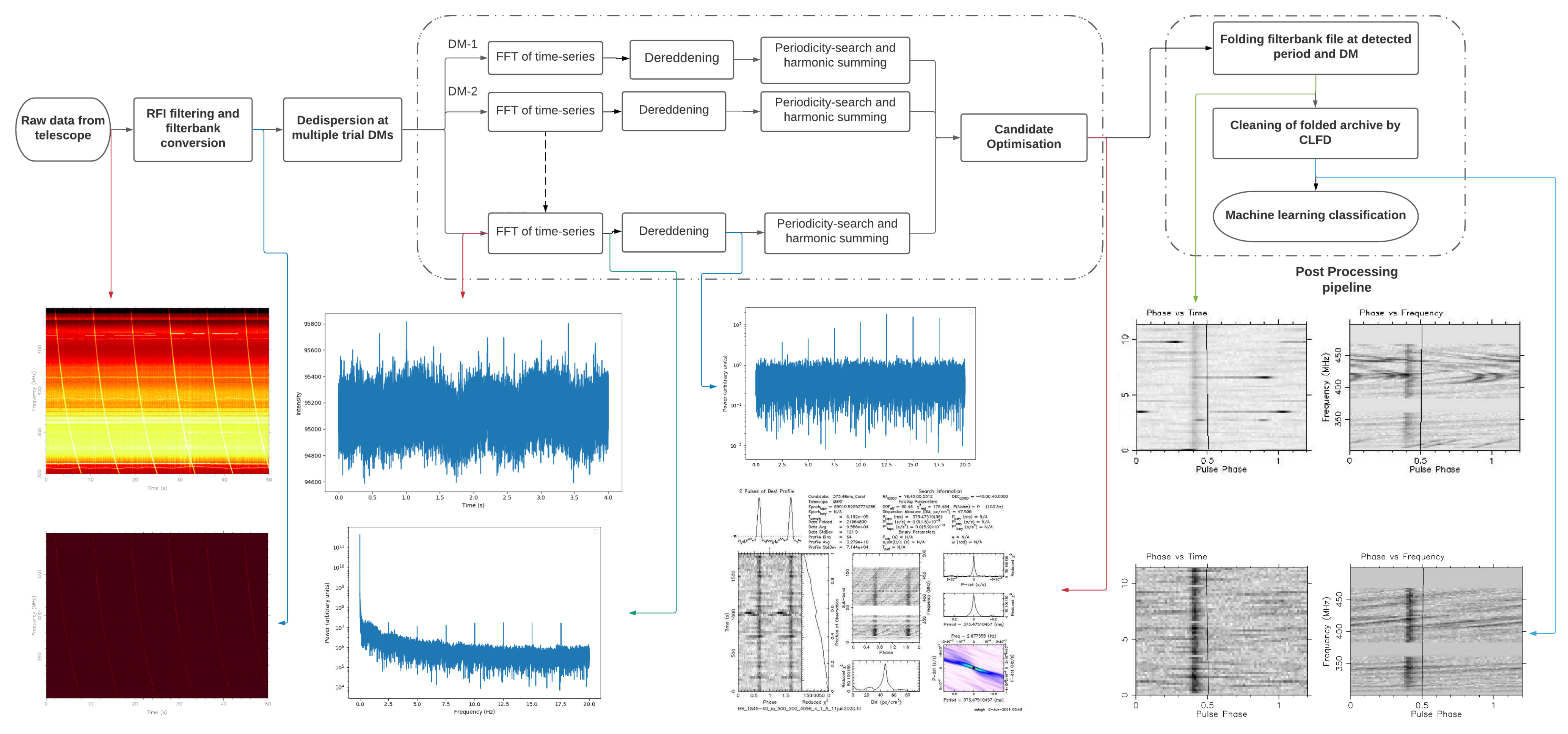
RFI: 57 **Candidates: 5**

45 out of 62: 59324.1734060347_DM_63.24_beam_127C.jpg

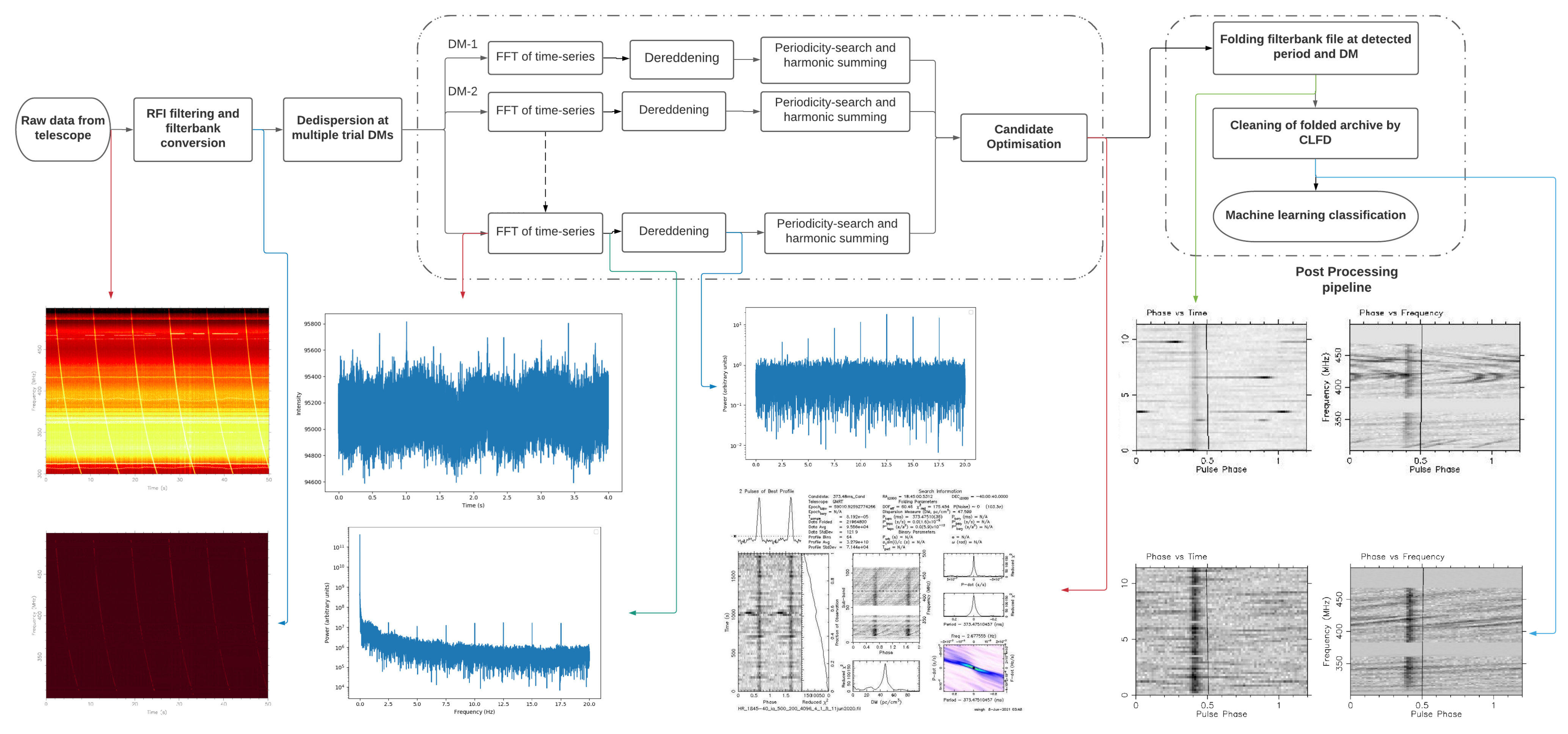
Searching using the Fast Fourier Transform



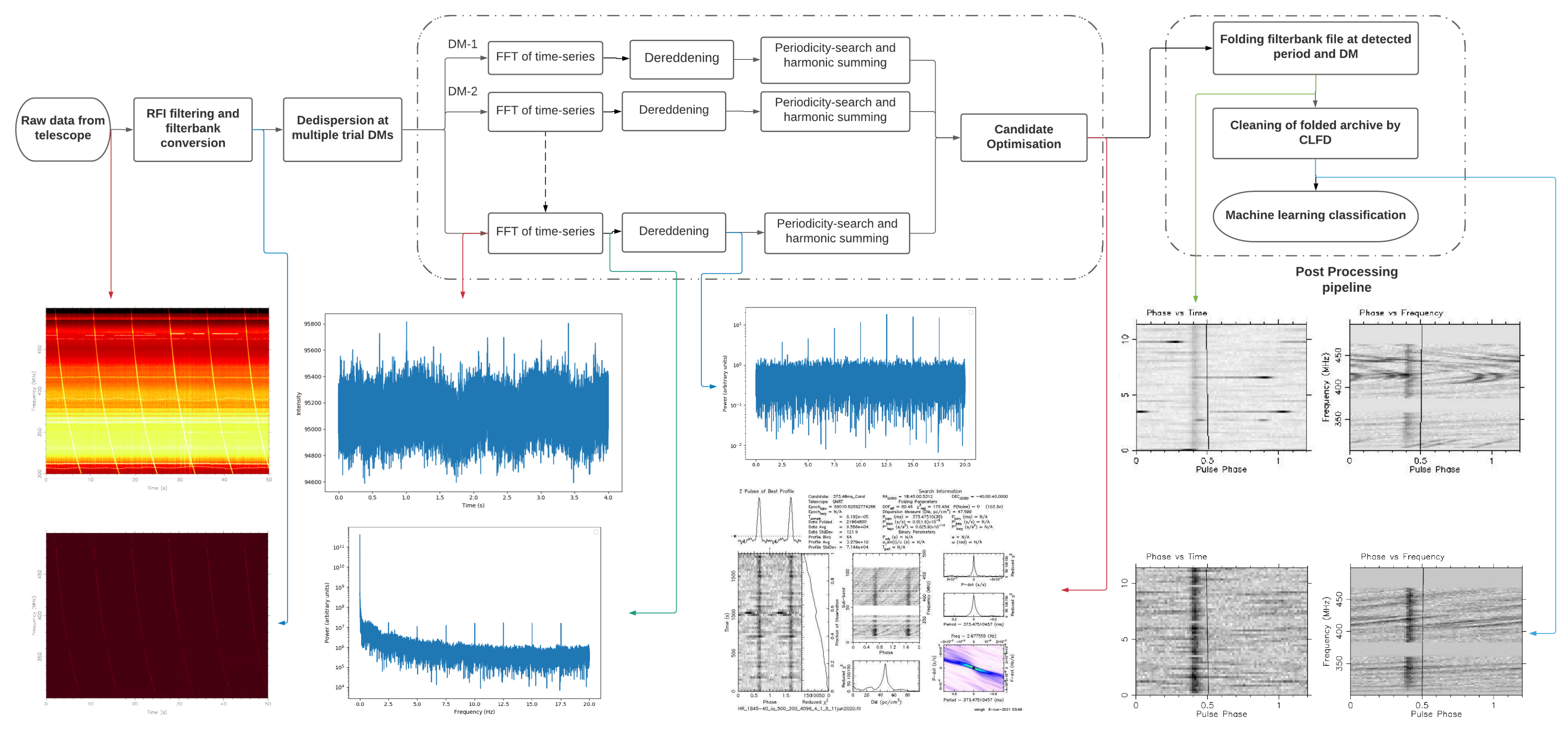
Searching using the Fast Fourier Transform



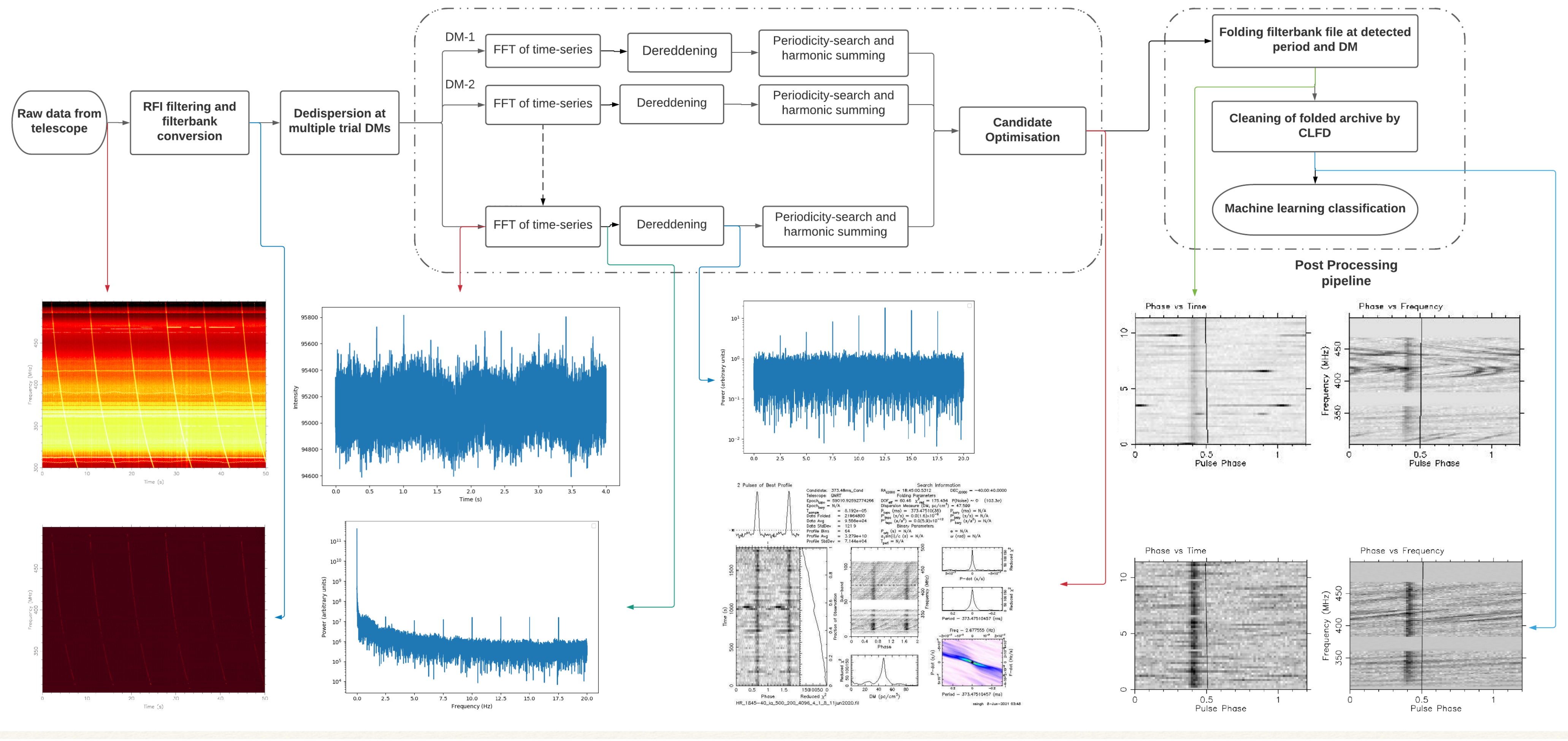
Searching using the Fast Fourier Transform



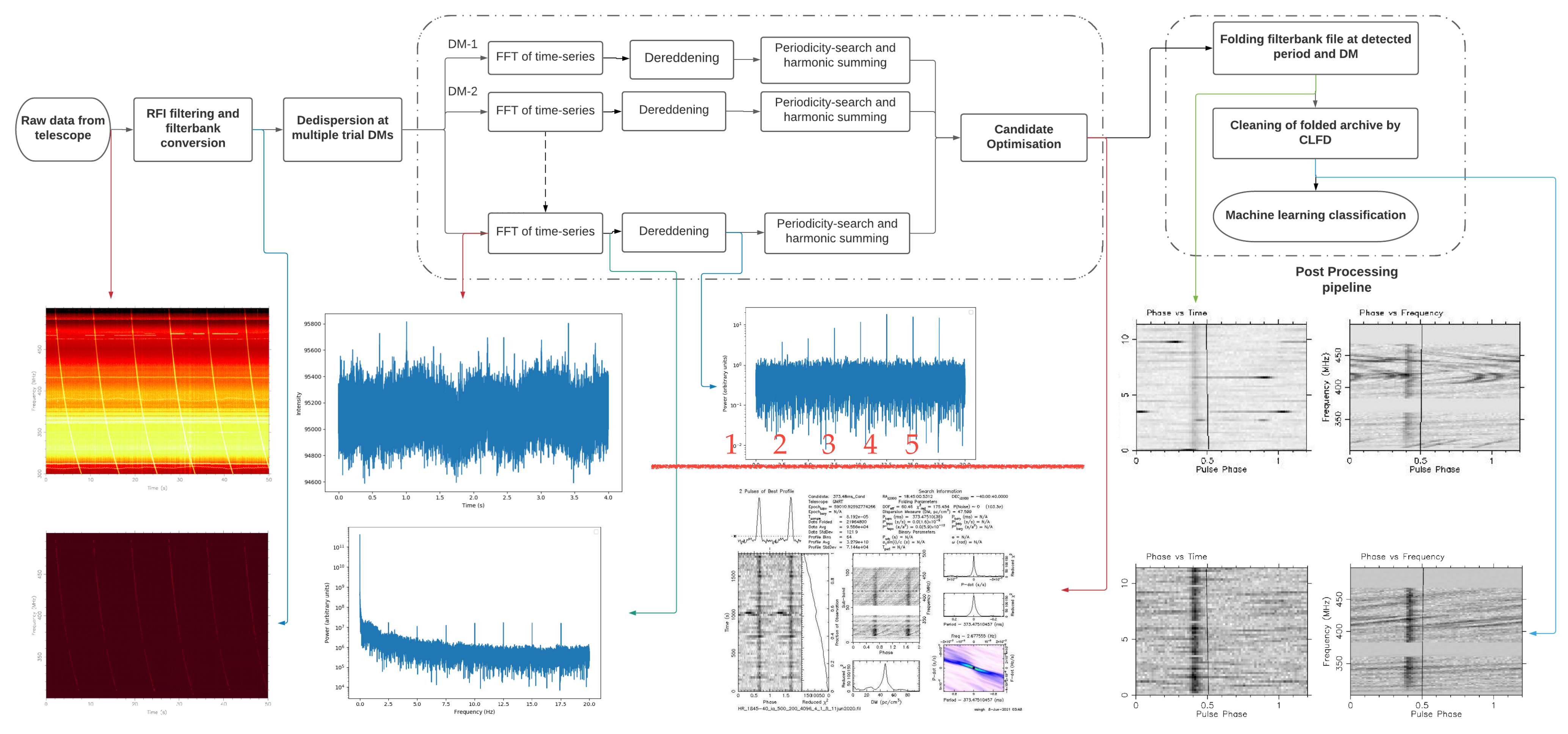
Searching using the Fast Fourier Transform



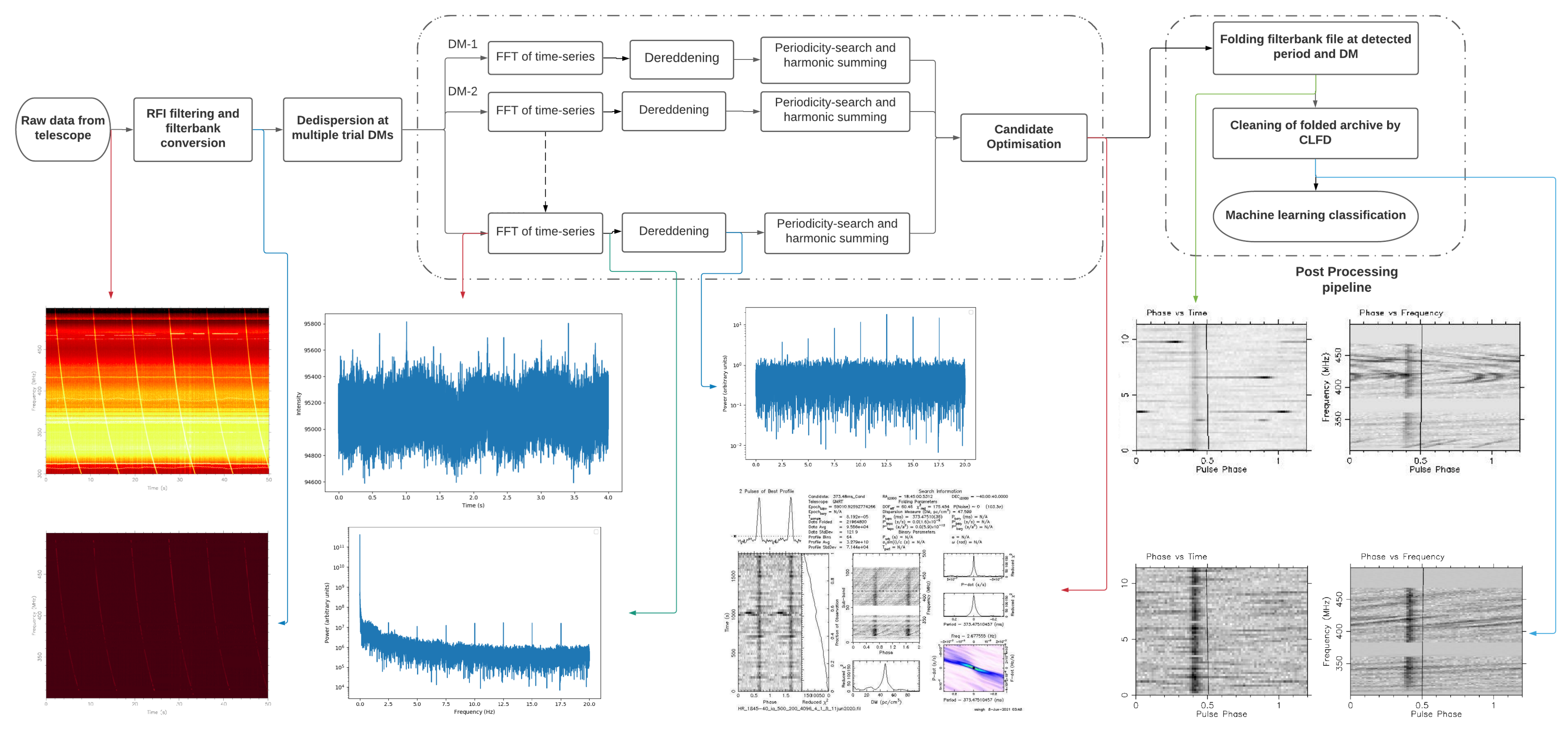
Searching using the Fast Fourier Transform



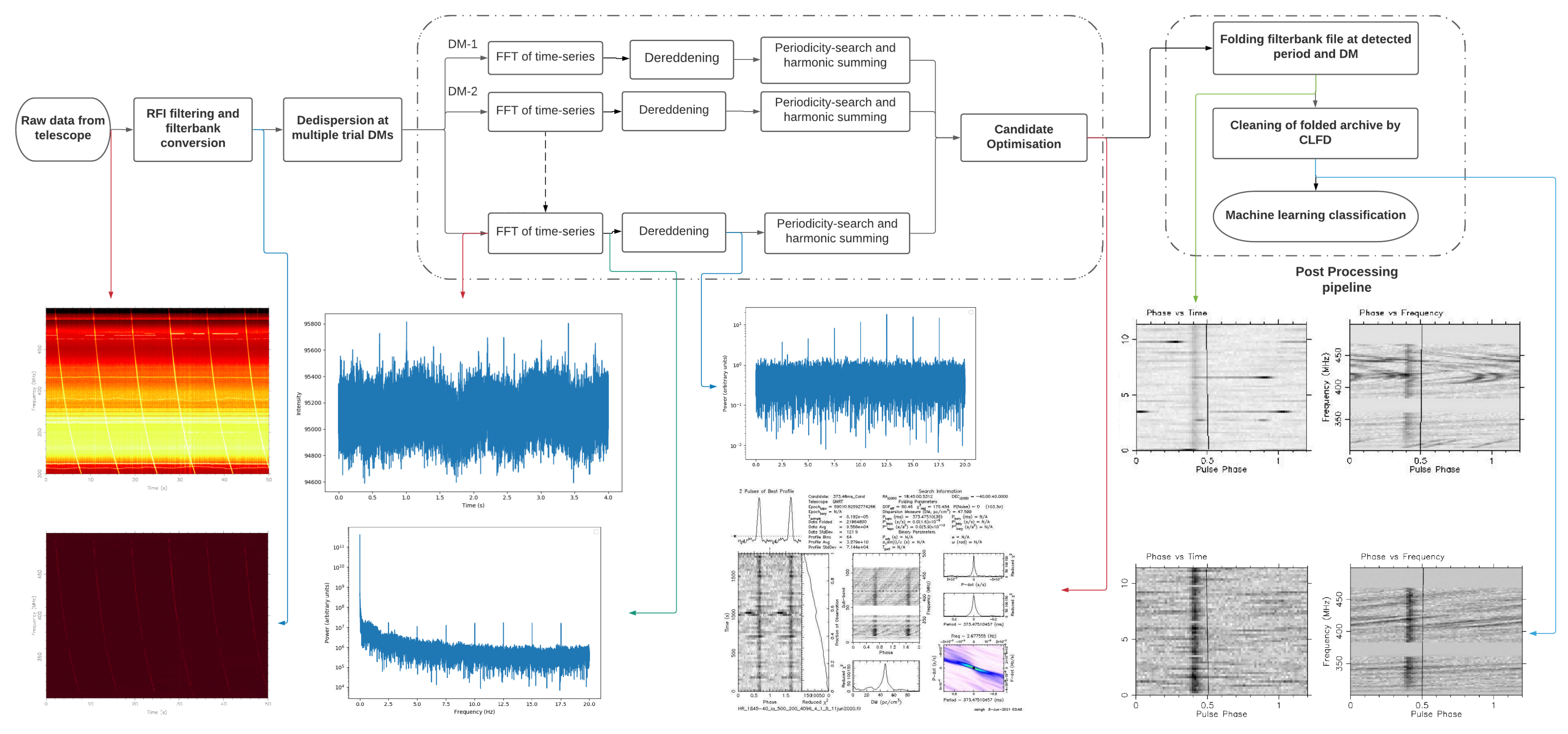
Searching using the Fast Fourier Transform



Searching using the Fast Fourier Transform

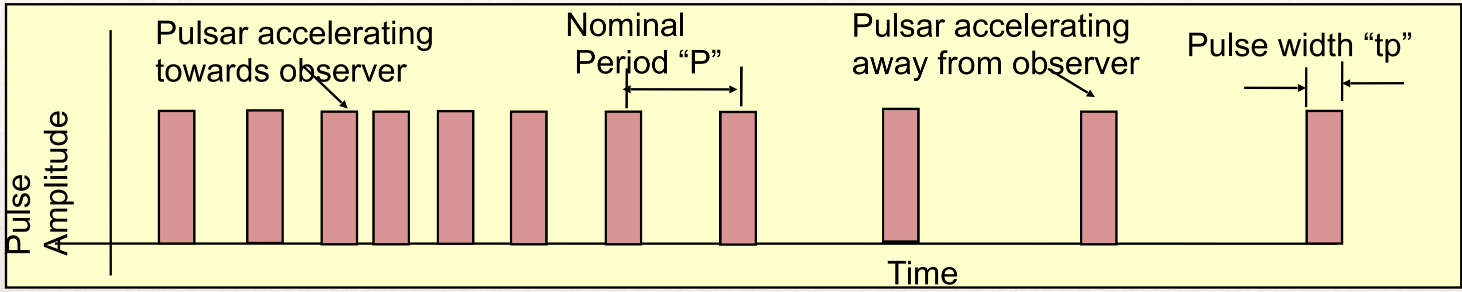


Searching using the Fast Fourier Transform

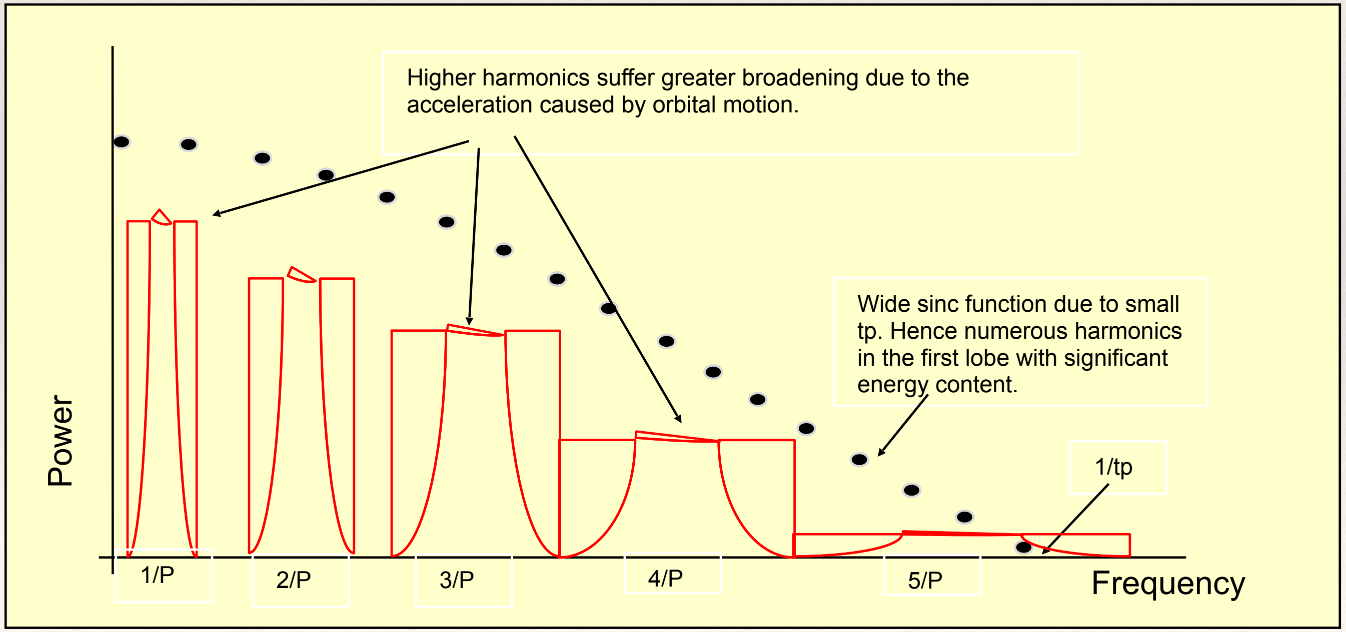


FINDING BINARY PULSARS

Time Domain



Frequency Domain



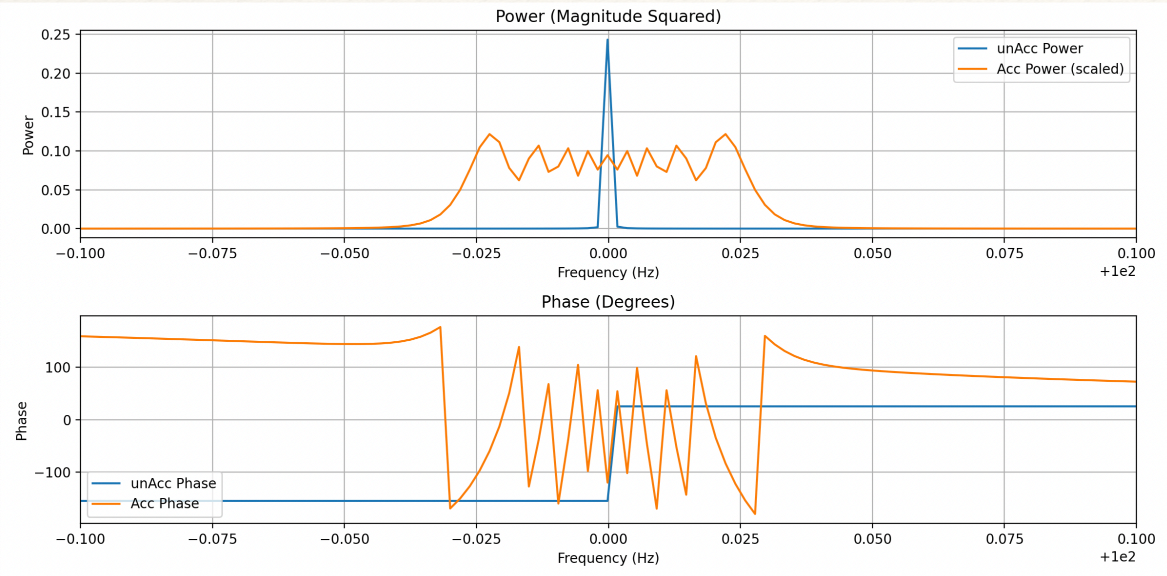
Pulsar in a binary system may accelerate towards and/or away from an observer.

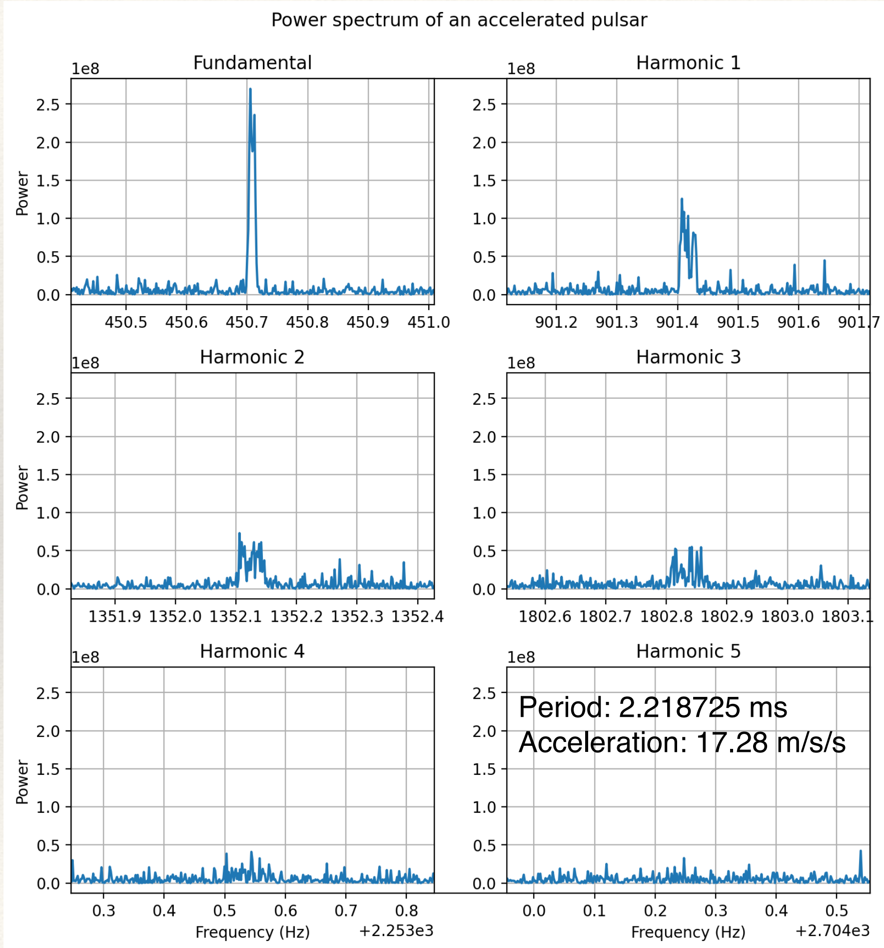
Due to the Doppler effect this results in pulses bunching up or spreading apart.

- Pulsar acceleration results in a broadening of the harmonics.
- Drifting magnitude proportional to the acceleration and the harmonic number (i.e. the frequency).
- Narrow pulse width leads to more harmonics with significant power.

FINDING BINARY PULSARS

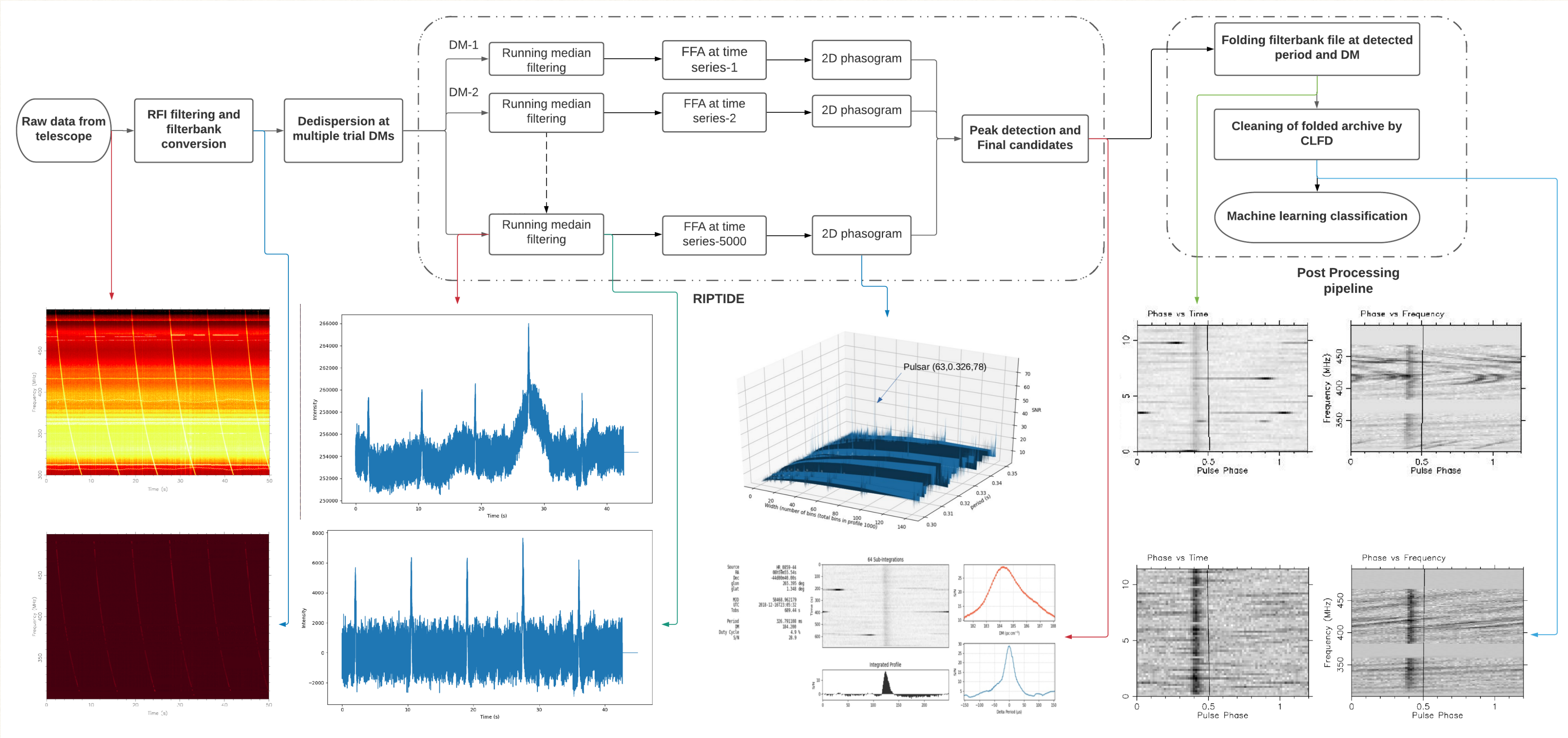
Effect of Linear Acceleration/Templates — 100 Hz / 350 m/s/s





Need dozens of filters to get sensitivity to different accelerations
Need to consider each of the harmonics.
Need many trial dispersion measures

Searching using the Fast Folding Algorithm



The University of Manchester

SOOO…MANY CANDIDATES!

MACHINE LEARNING TECHNIQUES

Pulsars – increasing candidate volumes

- 1,500 search beams for just SKA Mid. Each producing 1,000 candidates per scan.
- 1.5 million candidates per scan.
- A scan can be as short as 600 seconds.
- That's 72 million candidates per day.
- 7.2 billion candidates for a full sky survey!
- SKA Mid expected to find around 10,000 pulsars
- <u>There are >100,000 non-target examples for each interesting candidate</u>
- <u>Remember that the survey is done in real time, i.e. no second chances!</u>

BUT — candidate plots and metadata stored for future improved methods.

Transients – increasing candidate volumes

- The numbers are less well defined in the literature right now.
- Heavily dependent on the RFI situation and highly variable
- The MeerTRAP survey at MeerKAT can be 10,000s per day
- MeerTRAP (and others like CHIME and ASKAP) are already running real time ML implementations (see later) to limit these candidate numbers
- Remember that the SKA survey is done in real time, i.e. no second chances!

In this case real time is more important as we need to be able to trigger a transient buffer for:

- **Accurate localisation**
- **Polarisation studies**
- **High time resolution pulse studies.**

BUT — candidate plots and metadata stored for future improved methods.

Candidate selection methods up to 2014

Method	Advantages	Disadvantages	Examples
Manual selection, Section 3.5.1	Accurate human decisions	Slow, error prone, non-reproducible	All surveys
Summary selection, Section 3.5.2	Accurate human decision making, slightly faster analysis	Same as above, success relies on plot used, assumes data visually separable	RUNVIEW (Burgay et al., 2006), REAPER (Faulkner et al., 2004), JREAPER (Keith et al., 2009)
Ranking, Section 3.5.3	Prioritises good candidates for attention, fast, reproducible	Needs human input, prone to error due to human input, ranking used determines success	PEACE Lee et al. (2013) CLUSTERRANK (Deneva et al., 2016)
Automated selection, Section 3.5.4	Fast, reproducible, no human interaction	Difficult to train well, requires access to training data	SPINN (Morello et al., 2014), PICS (Zhu et al., 2014)

Table 3.2: Summary of candidate selection methods.

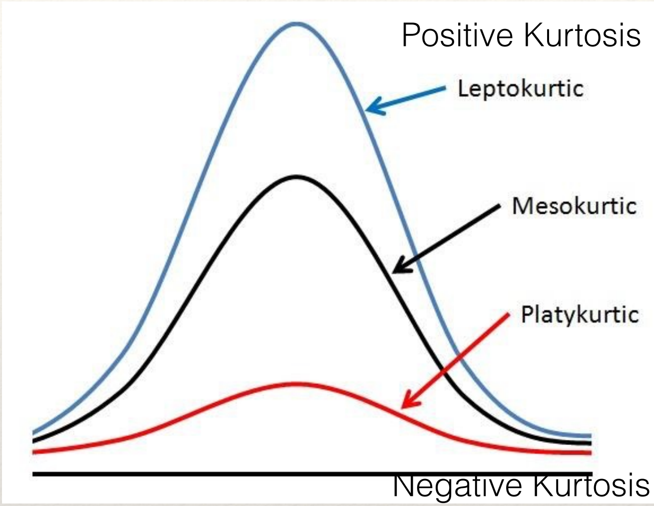
Classical ML, CNNs, transformers

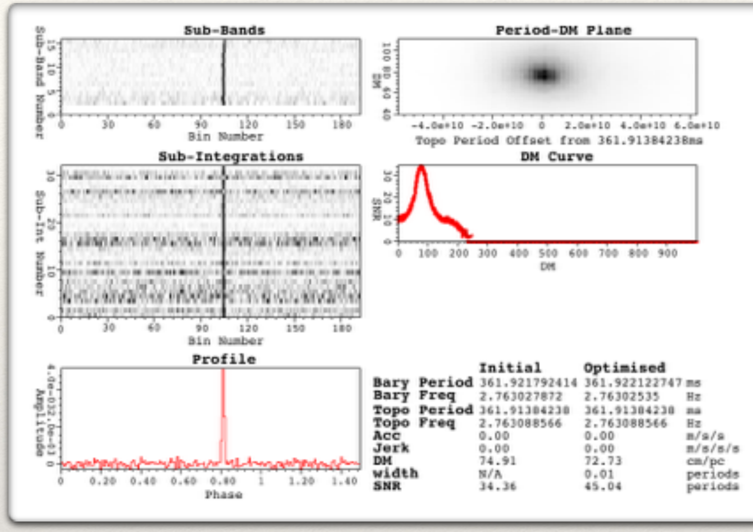
Modified and updated versions being used — including for TRAPUM

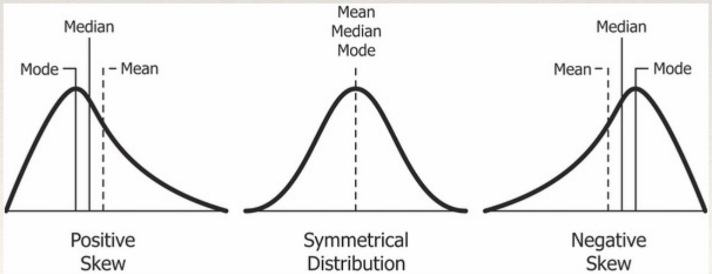
Lyon et al. Features

Feature	Description	Definition
$Prof_{\mu}$	Mean of the integrated profile P.	$\frac{1}{n}\sum_{i=1}^{n} p_i$
$Prof_{\sigma}$	Standard deviation of the integrated profile P.	$\sqrt{\frac{\sum_{i=1}^{n}(p_i - \bar{P})^2}{n-1}}$
$Prof_{k}$	Excess kurtosis of the integrated profile P.	$\frac{\frac{1}{n}(\sum_{i=1}^{n}(p_i - \bar{P})^4)}{(\frac{1}{n}(\sum_{i=1}^{n}(p_i - \bar{P})^2))^2} - 3$
$Prof_{s}$	Skewness of the integrated profile P.	$\frac{\frac{1}{n}\sum_{i=1}^{n}(p_i - \bar{P})^3}{(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - \bar{P})^2})^3}$
DM_{μ}	Mean of the DM-SNR curve D.	$\frac{1}{n}\sum_{i=1}^{n} d_i$
DM_{σ}	Standard deviation of the DM-SNR curve D.	$\sqrt{\frac{\sum_{i=1}^{n}(d_i - \bar{D})^2}{n-1}}$
DM_{k}	Excess kurtosis of the DM-SNR curve D.	$\frac{\frac{1}{n}(\sum_{i=1}^{n}(d_i - \bar{D})^4)}{(\frac{1}{n}(\sum_{i=1}^{n}(d_i - \bar{D})^2))^2} - 3$
DM_{s}	Skewness of the DM-SNR curve D.	$\frac{\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{D})^3}{(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{D})^2})^3}$

Table 4. The eight features derived from the integrated pulse profile $P = \{p_1, \ldots, p_n\}$, and the DM-SNR curve $D = \{d_1, \ldots, d_n\}$. For both P and D, all p_i and $d_i \in \mathbb{N}$ for $i = 1, \ldots, n$.







Summary of Results using a VFDT

Dataset	Algorithm	G-Mean	F-Score	Recall	Precision	Specificity	FPR	Accuracy
HTRU 1	C4.5	0.962*	0.839*	0.961	0.748	0.962	0.038	0.962
	MLP	**0.976**	0.891	**0.976**	0.820	0.975	0.025*	0.975
	NB	0.925	0.837*	0.877	0.801	0.975	0.025*	0.965
	SVM	0.967	0.922	0.947	0.898	0.988	0.012	0.984
	GH-VFDT	0.961*	**0.941**	0.928	**0.955**	**0.995**	**0.005**	**0.988**
HTRU 2	C4.5	0.926	0.740	**0.904**	0.635*	0.949*	0.051*	0.946*
	MLP	**0.931**	0.752	0.913	0.650*	0.950*	0.050*	0.947*
	NB	0.902	0.692	0.863	0.579	0.943	0.057	0.937
	SVM	0.919	0.789	0.871	0.723	0.969	0.031	0.961
	GH-VFDT	0.907	**0.862**	0.829	**0.899**	**0.992**	**0.008**	**0.978**
LOTAAS 1	C4.5	0.969	0.623	0.948	0.494	0.991	0.009	0.990
	MLP	**0.988**	0.846*	**0.979**	0.753	0.998	0.002	0.997*
	NB	0.977	0.782	0.959	0.673	0.996	0.004	0.996
	SVM	0.949	**0.932**	0.901	**0.966**	**0.999***	**0.001***	**0.999**
	GH-VFDT	0.888	0.830*	0.789	0.875	**0.999***	**0.001***	0.998*

Table 11. Results obtained on the three test data sets. Bold type indicates the best performance observed. Results with an asterisk indicate no statistically significant difference at the $\alpha = 0.01$ level.

The University of Manchester

Improvements

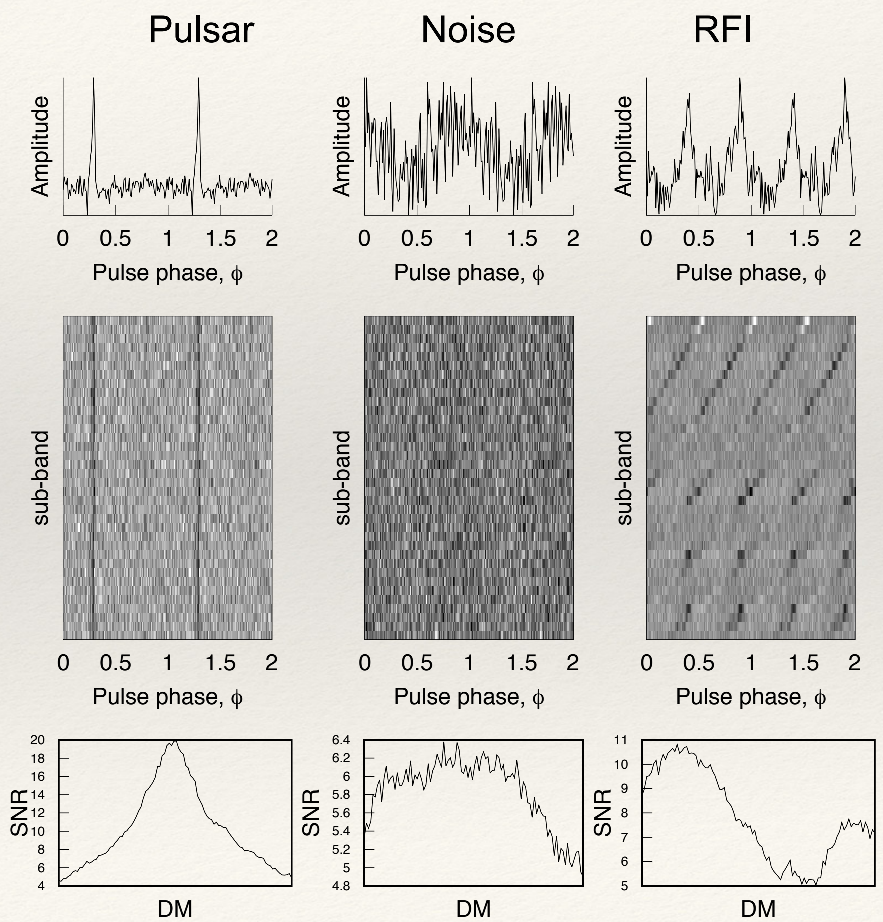
Tan et al. 2018.

Adding an RFI class

- Separate known RFI instances from other non-pulsars
- 3 class classifier
- Greatly improved the performance.

Ensemble Classifier

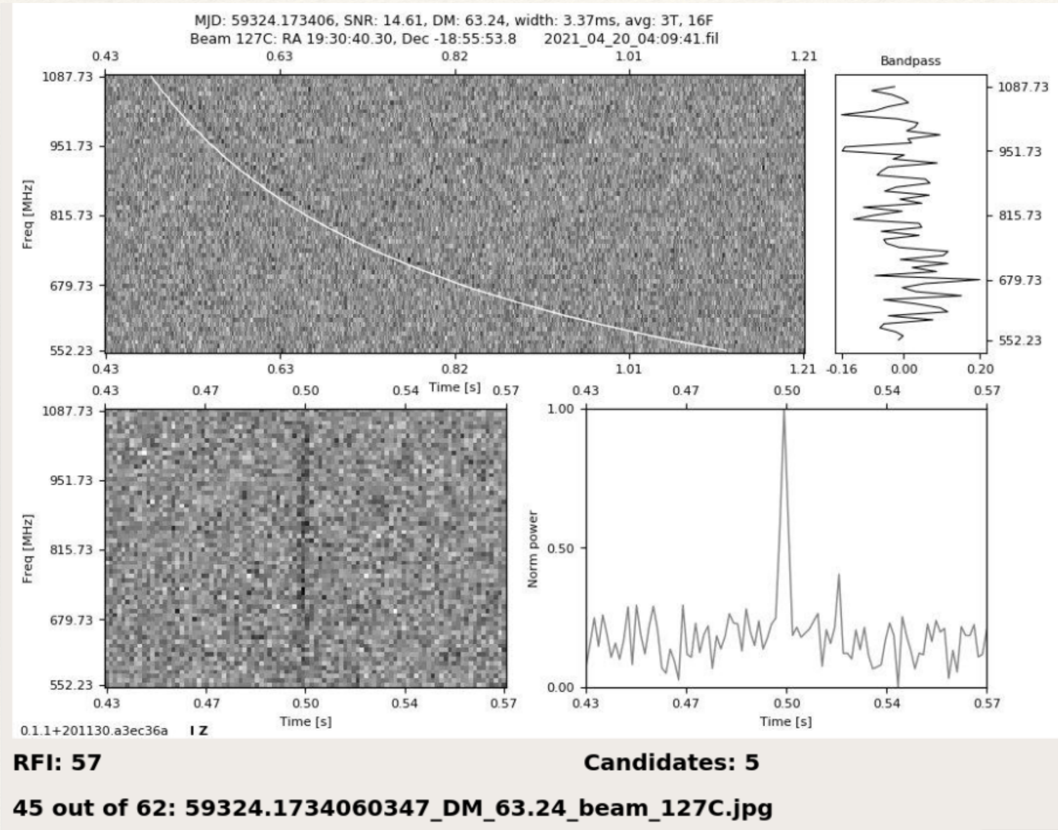
- 5 VFDT classifiers with 5 different training datasets
- Noise and RFI class
- Candidate is pulsar if >=3 VFDT classifiers said so.
- Pulsar recall rate from 96.2% to 98.7%
- False positive rate from 2.5% to 1% — 60% reduction in candidates

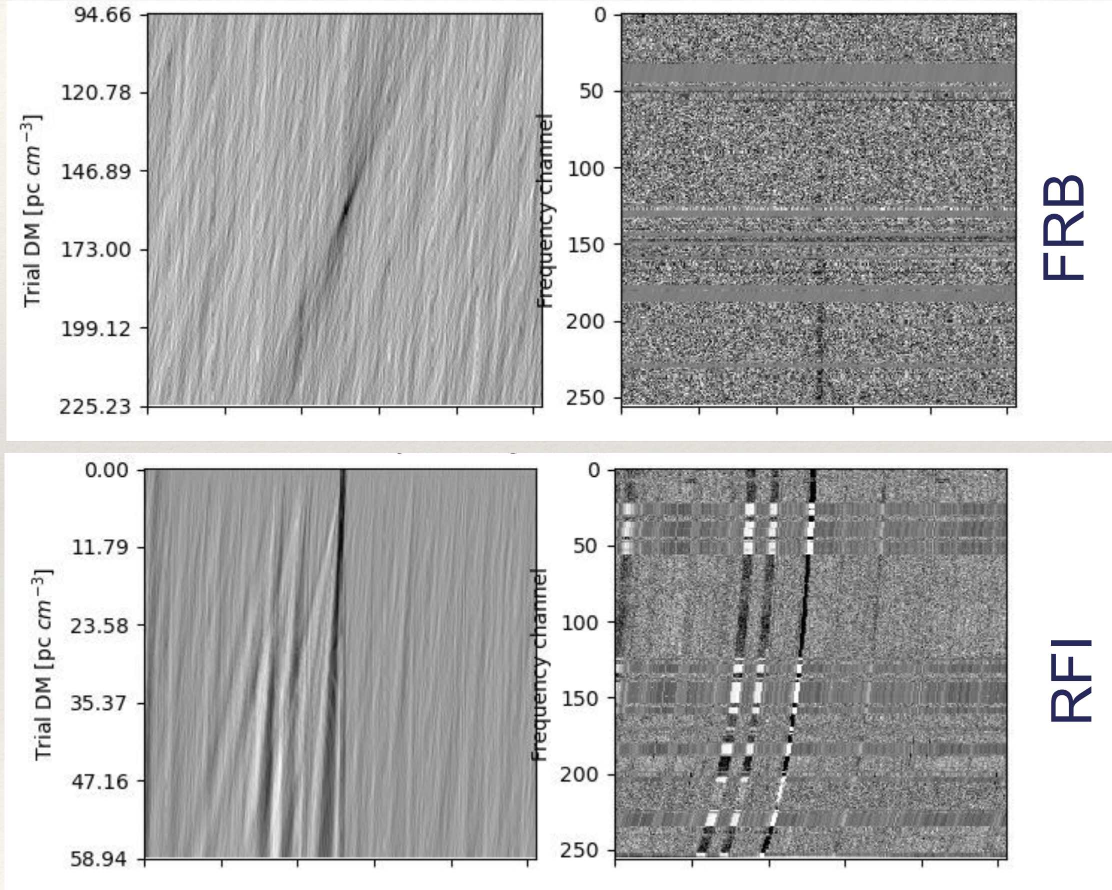


Transcient candidate data

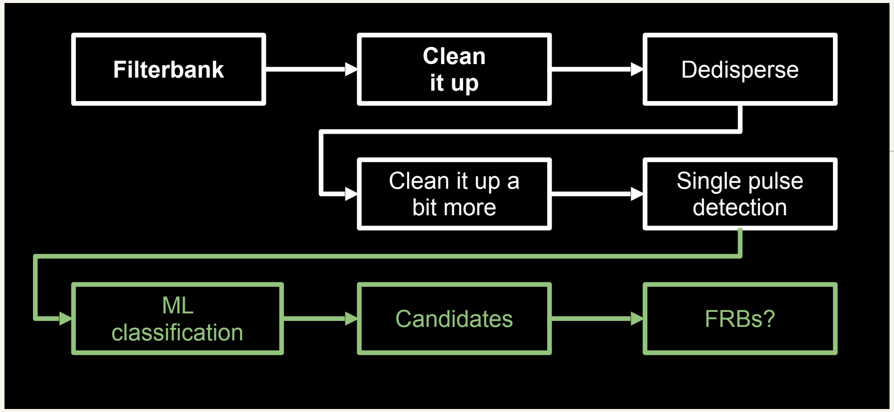
CNN binary classification

Human verification / data annotation





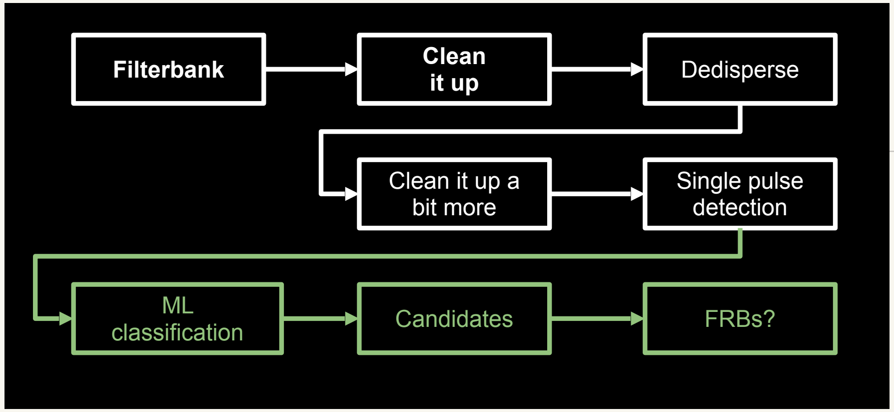
POSTDETECTION



- MeerTRAP Can have **100k+ candidates in one day**
- Use additional vetting: **multibeam clustering, known-source removal**
- Can still end up with 10,000+ candidates **per day**
- Takes **a lot** of time to look through all of them
- 250TB archive **gone in no time**

MANCHESTER 1824
The University of Manchester

POSTDETECTION

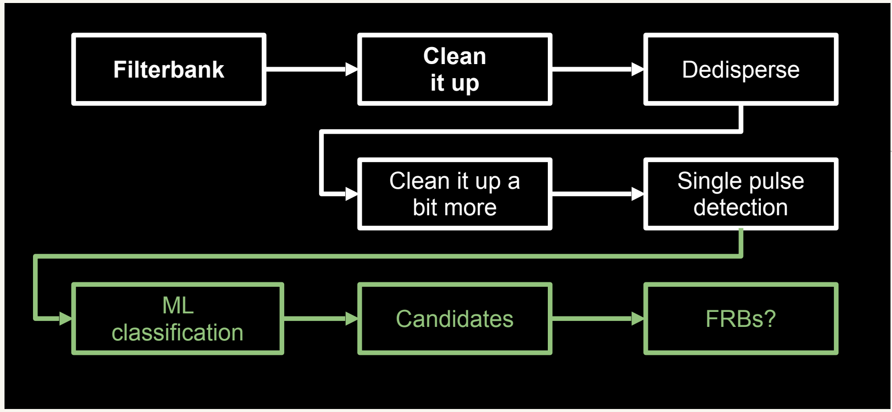


- MeerTRAP Can have **100k+ candidates in one day**
- Use additional vetting: **multibeam clustering, known-source remova**l
- Can still end up with 10,000+ candidates **per day**
- Takes **a lot** of time to look through all of them
- 250TB archive **gone in no time**

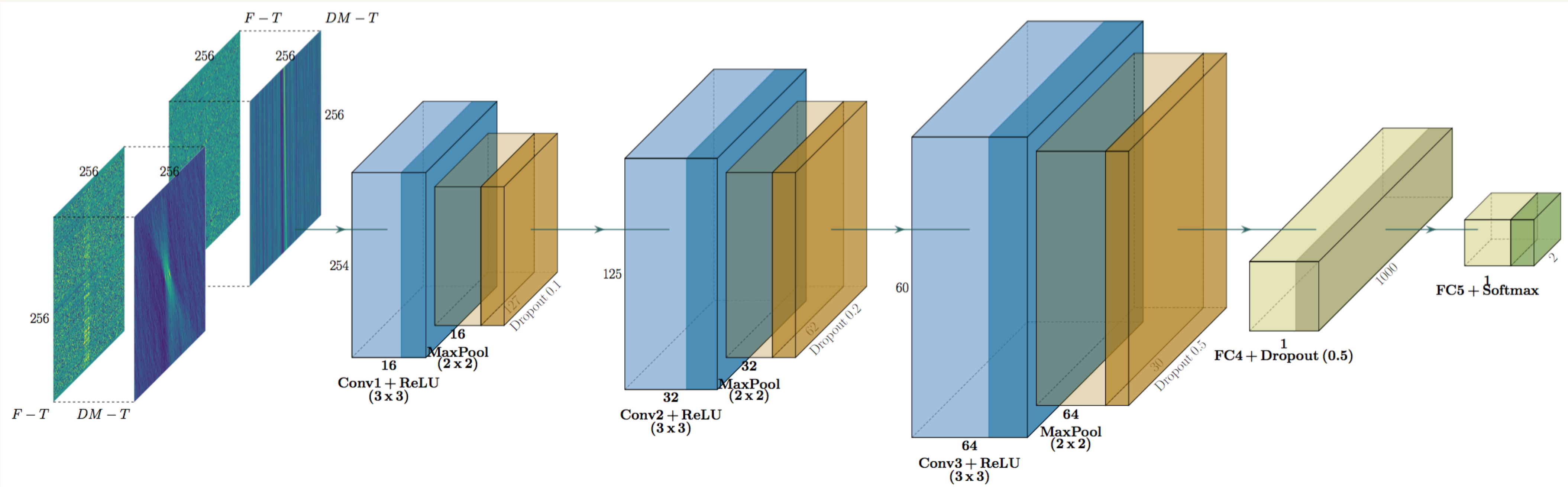
→ **Additional post-processing needed**

MANCHESTER 1824
The University of Manchester

POSTDETECTION - FRBID



- Inspired by **FETCH** (Agarwal et al. 2020)
- Deep CNN with DM-Time and Freq-Time input planes
- Stack the inputs and follow single path (**works very well**)



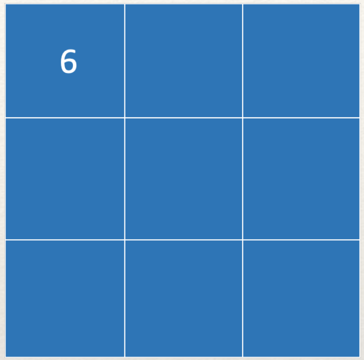
MANCHESTER 1824
The University of Manchester

Deep learning algorithms – Convolutional Neural Networks

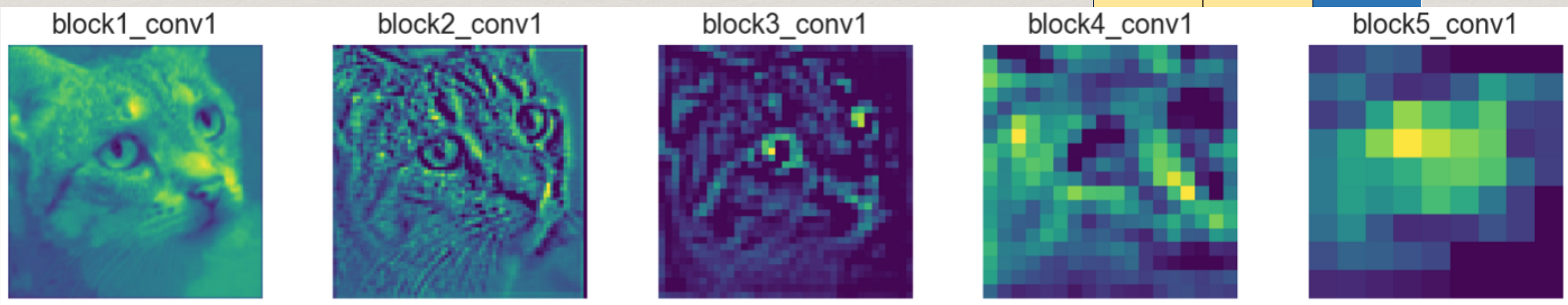
- Sparse interactions: reduce number of parameters and improve the efficiency.
- Parameter sharing.
- Equivariant representation: location of features is not important.

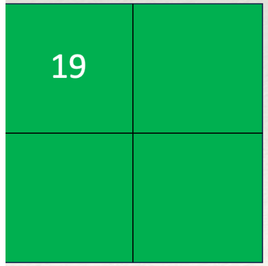
Convolution operation

1x1	0x0	1x2	0	0
3x0	1x3	0x1	3	1
0x2	1x1	0x1	3	2
2	3	1	2	0
0	0	2	0	2



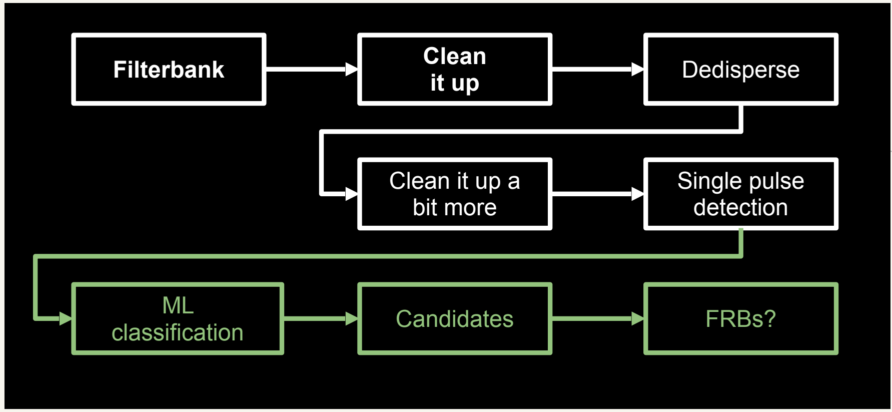
Maximum pooling

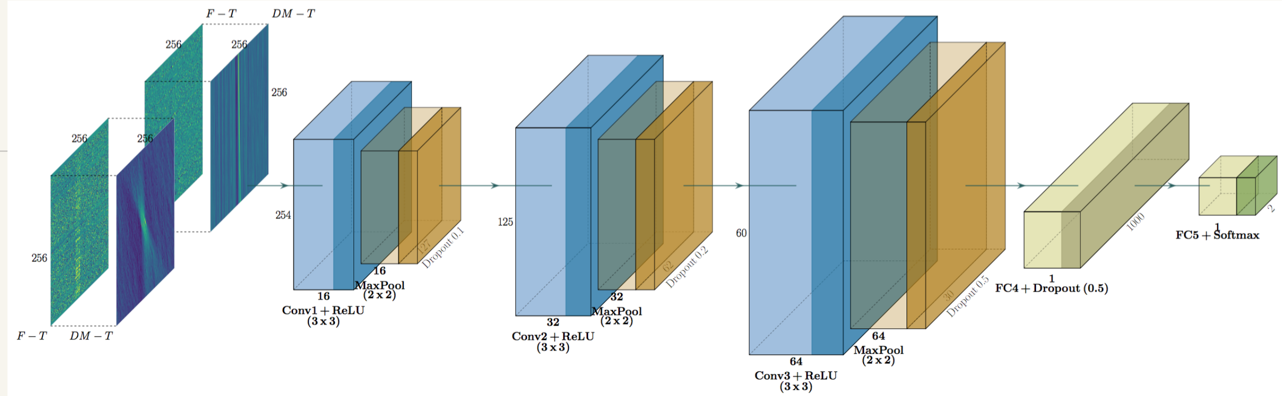




From towardsdatascience

POSTDETECTION - FRBID

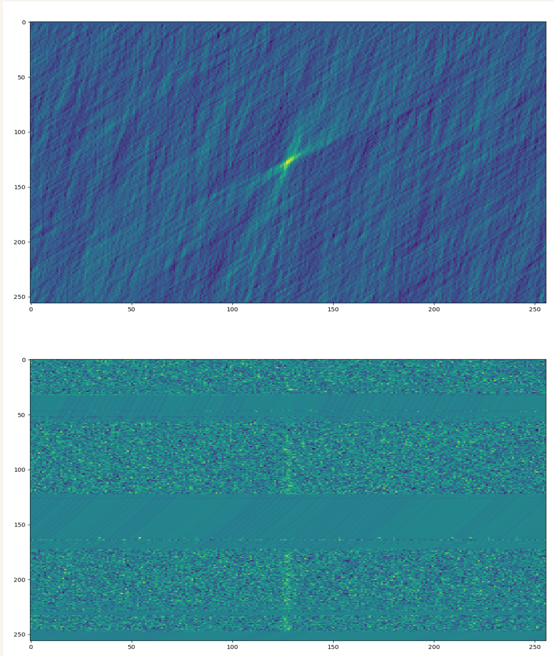


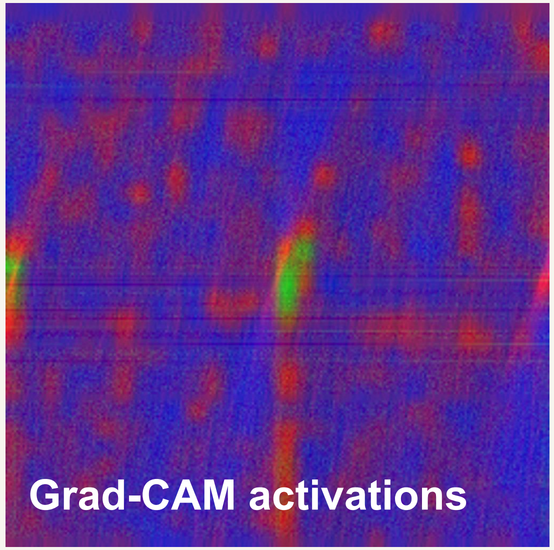


Z. Hosenie

- Inspired by **FETCH** (Agarwal et al. 2020)
- CNN with DM-Time and Freq-Time input planes
- Stack the inputs and follow single path (**works very well**)

- 20k candidates in the training set
- FP rate between 0.015% and 3%
- (<1% most of the time)

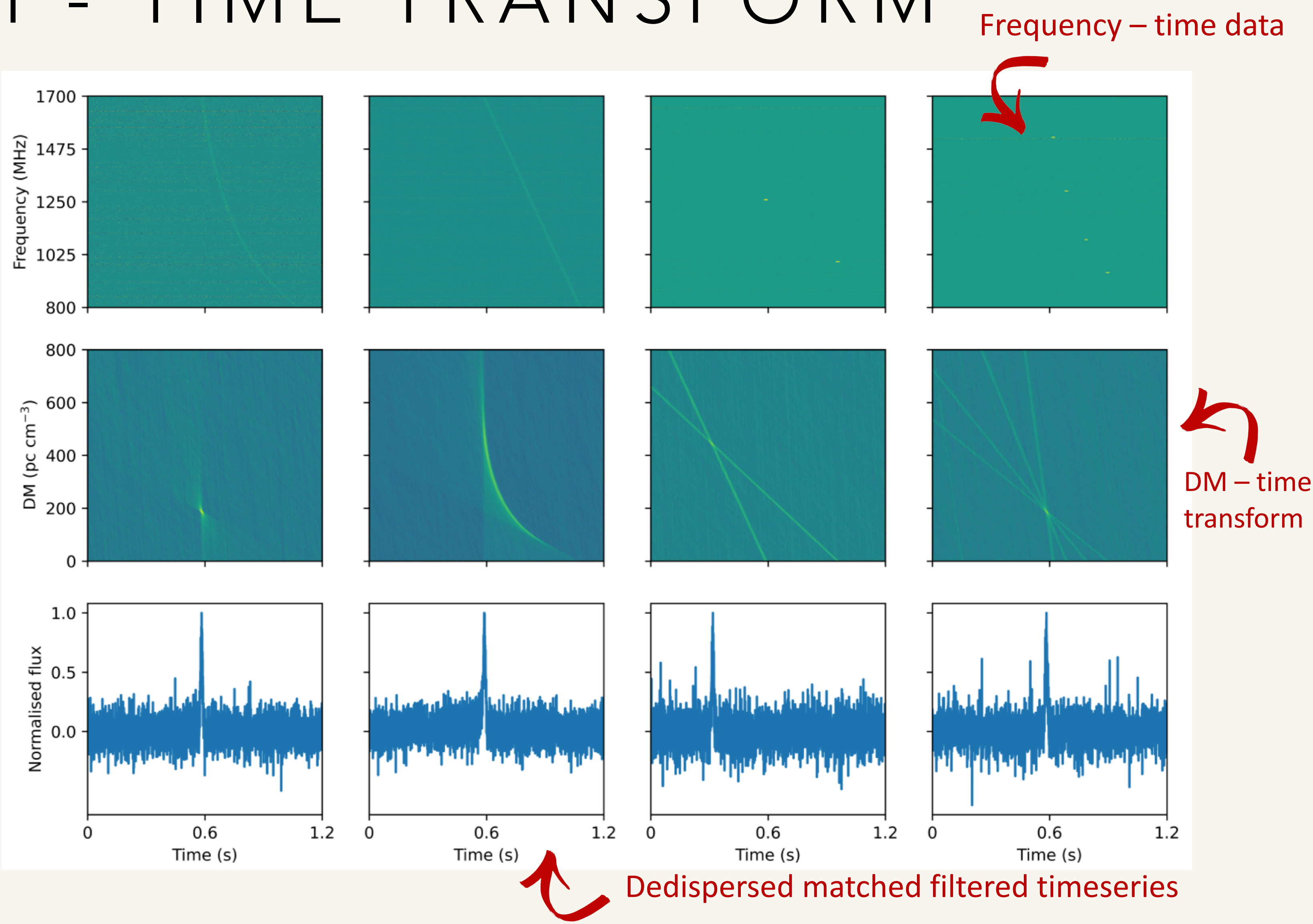




DM - TIME TRANSFORM

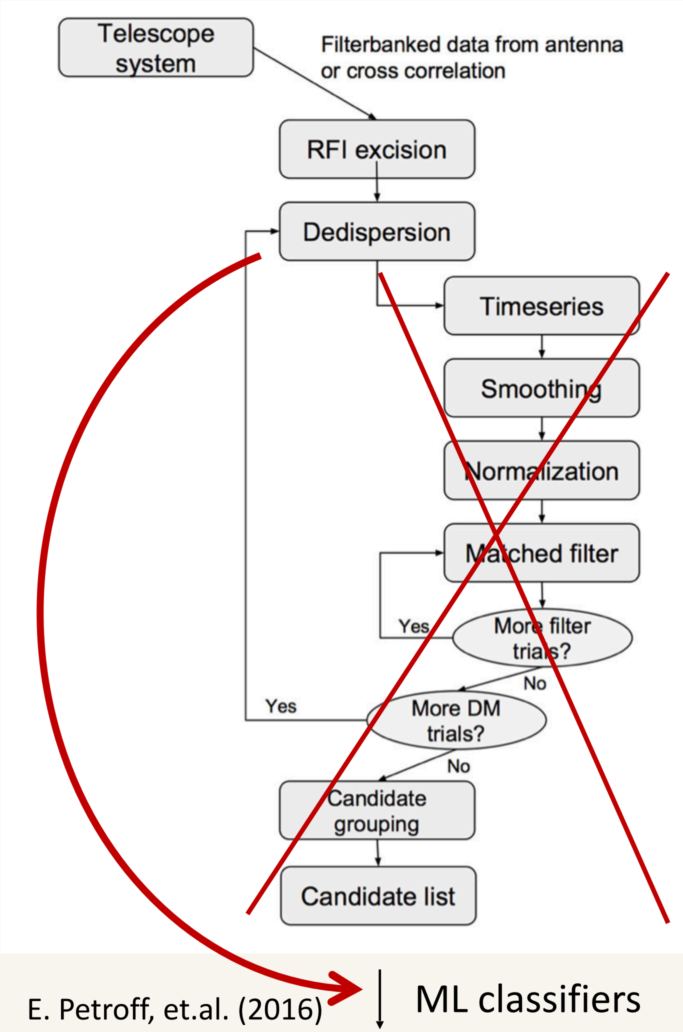
- The matched filtered timeseries provides information of the S/N values at a local DM value

- The DM – time transform provides information of the S/N degradation as a function of time and DM



Belmonte Diaz et al. 2025

OUR APPROACH – PERFORMING THE SEARCH IN DM - TIME USING ML



E. Petroff, et.al. (2016)

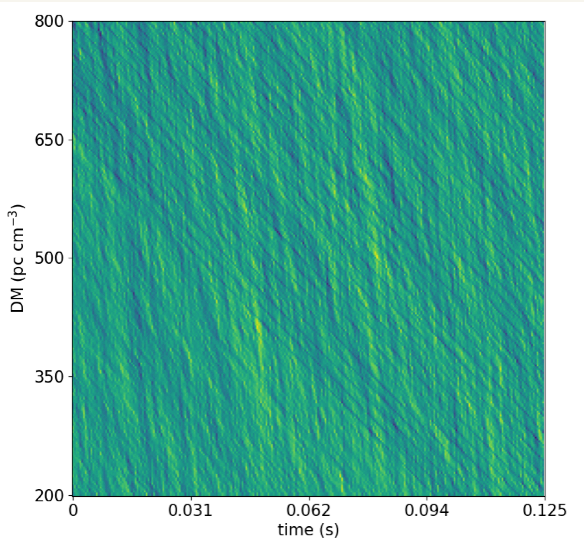
- Use the most available information as the pipeline input.
- Skip search on timeseries.
- Directly search for bursts in the DM – time plane using ML.

Belmonte Diaz et al. 2025

DEEP LEARNING ALGORITHMS – IMAGE SEGMENTATION

Image classification

Is there a burst in this image?

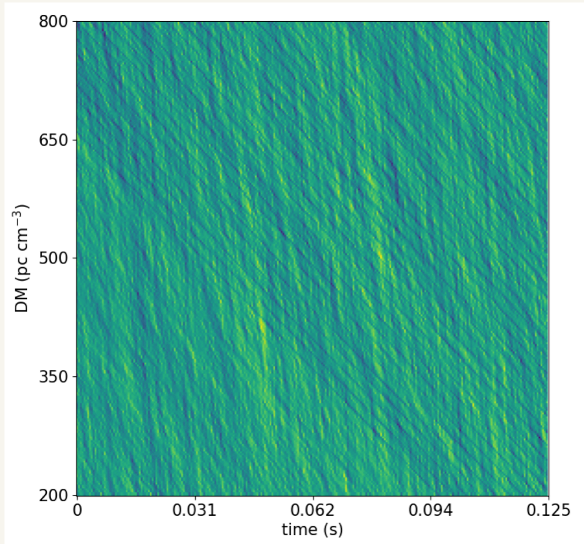


Yes

No

Image segmentation

Which pixels belong to a burst in this image?



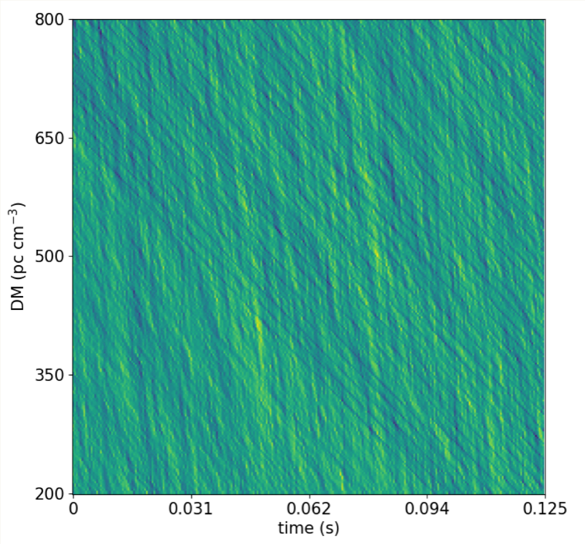
Localisation of pixels belonging to bursts

Belmonte Diaz et al. 2025

DEEP LEARNING ALGORITHMS – IMAGE SEGMENTATION

Image classification

Is there a burst in this image?

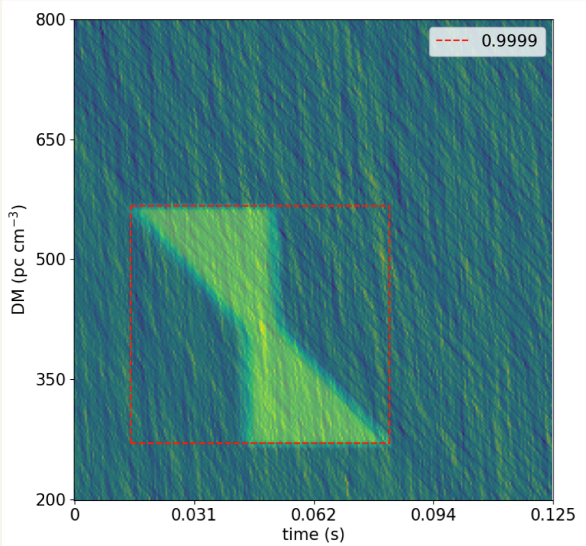


Yes

No

Image segmentation

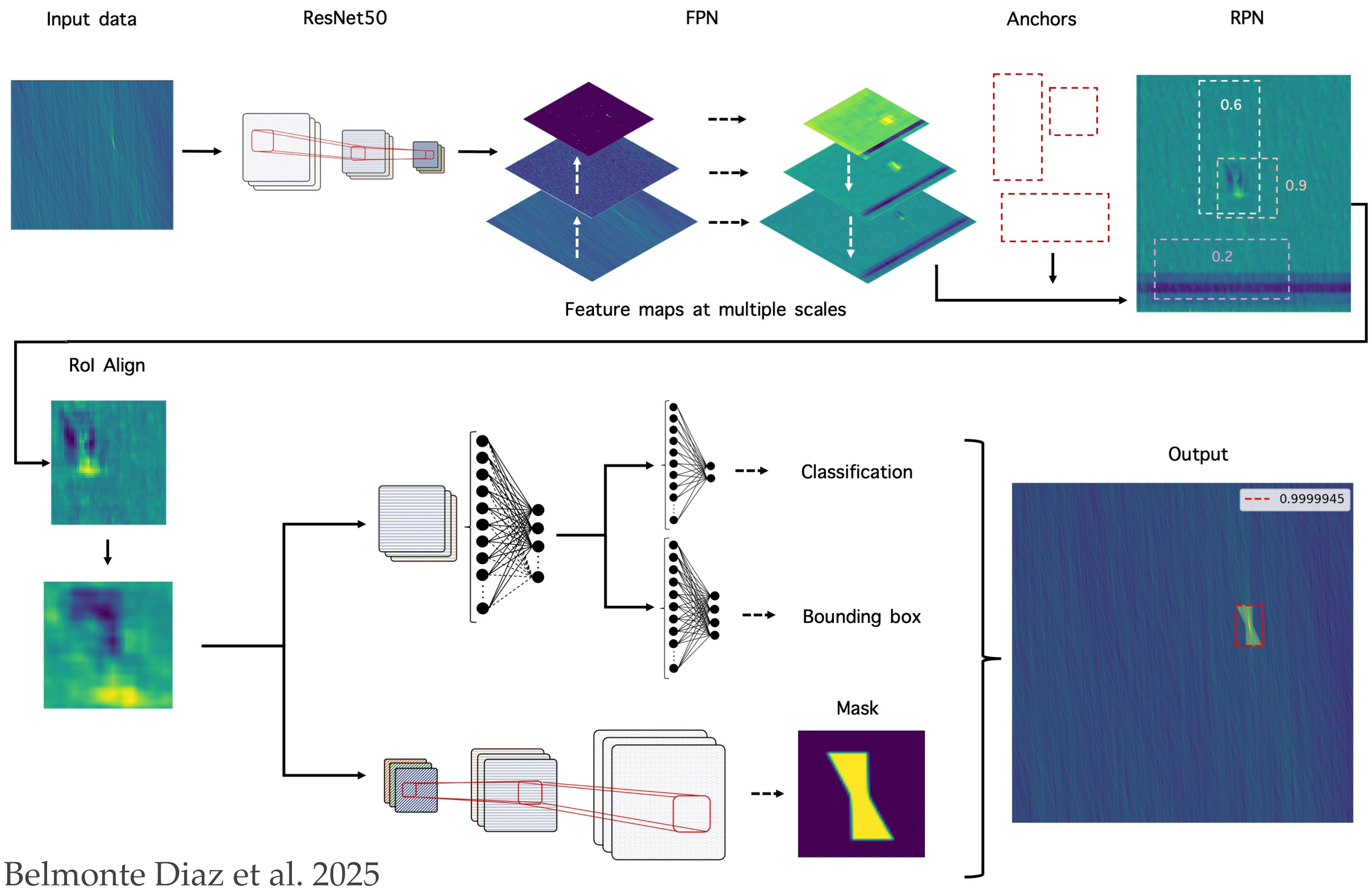
Which pixels belong to a burst in this image?



Localisation of pixels belonging to bursts

Belmonte Diaz et al. 2025

MASK RCNN ARCHITECTURE



Belmonte Diaz et al. 2025

Three main stages:

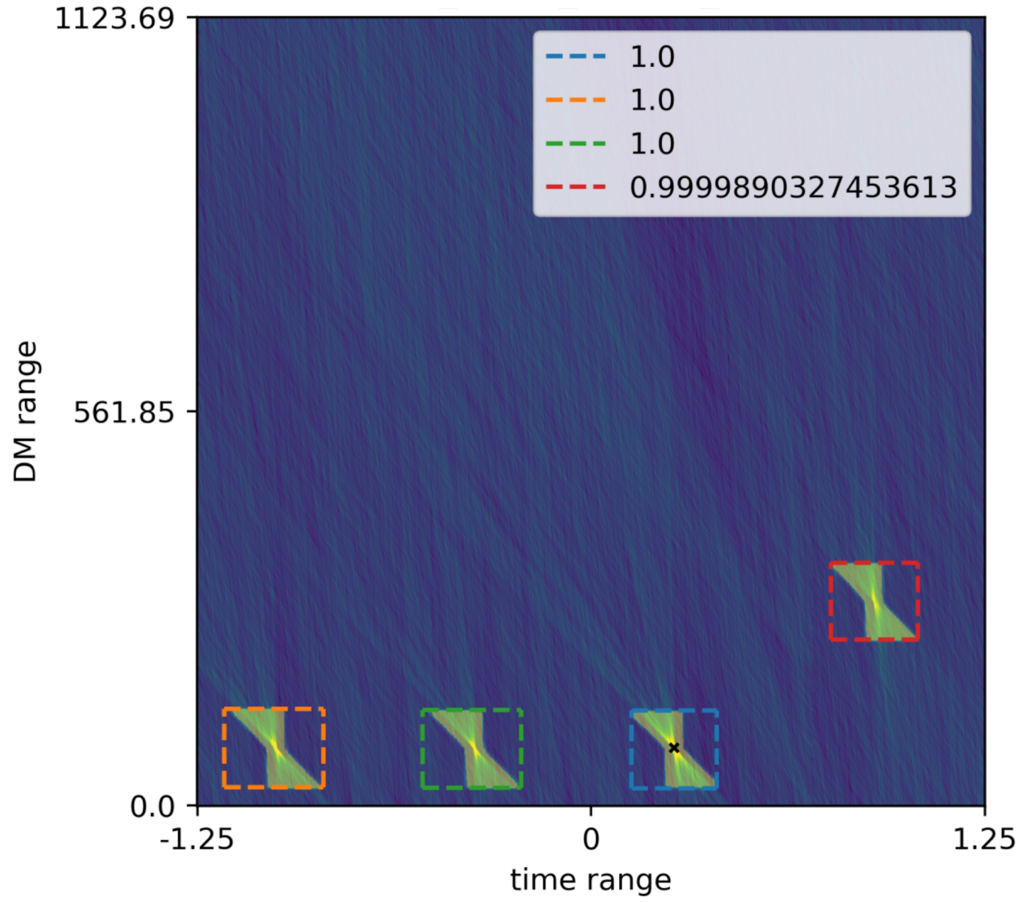
- Feature extraction with backbone network.
- Region of Interest regression and light weight classification.
- Deeper classification and mask regression.

TESTING THE MASK RCNN – REAL DATASET

- Tested the network against MeerKAT L-band, UHF and S-band observations containing bursts from nearby pulsars and FRBs.

- Detected all bursts in a wide range of DMs, widths and S/Ns.

- Multiple bursts can be detected in the same image.



Belmonte Diaz et al. 2025

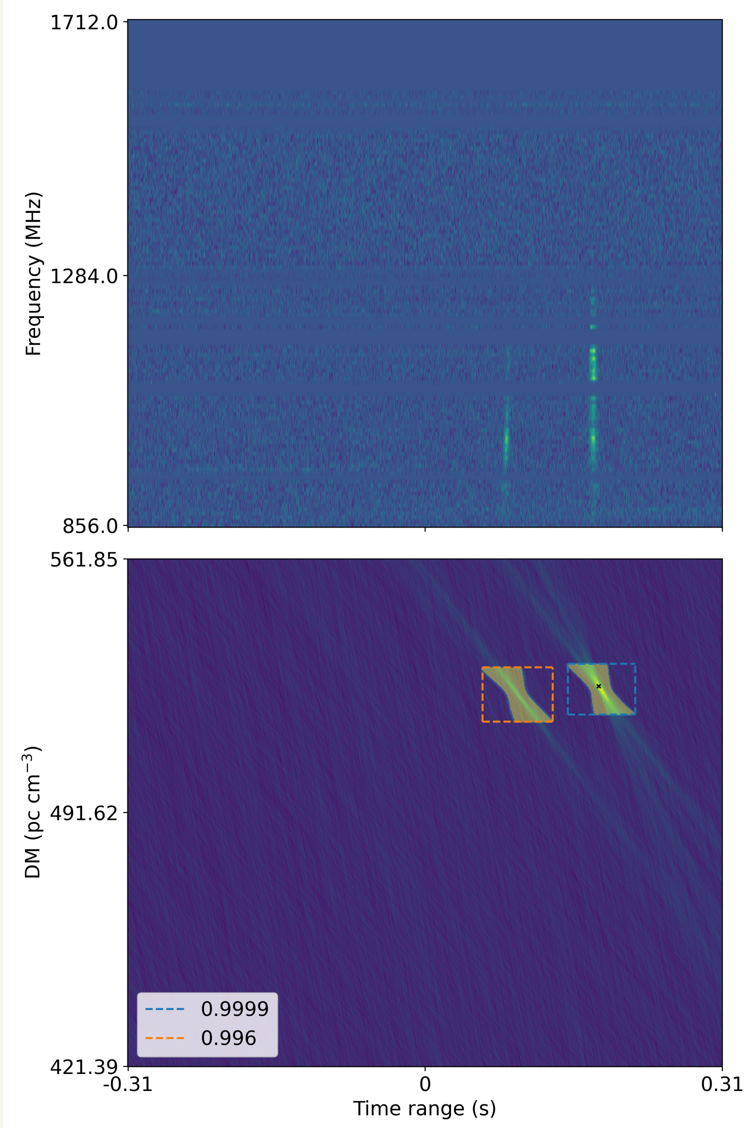
TESTING THE MASK RCNN – REAL DATASET

- Even though simulations were basic bursts, it can find pulses exhibiting an exotic behaviour.
- If separation between burst components is big enough, all components can be detected.

Dedispersed frequency – time data

DM – time transform with pipeline outputs

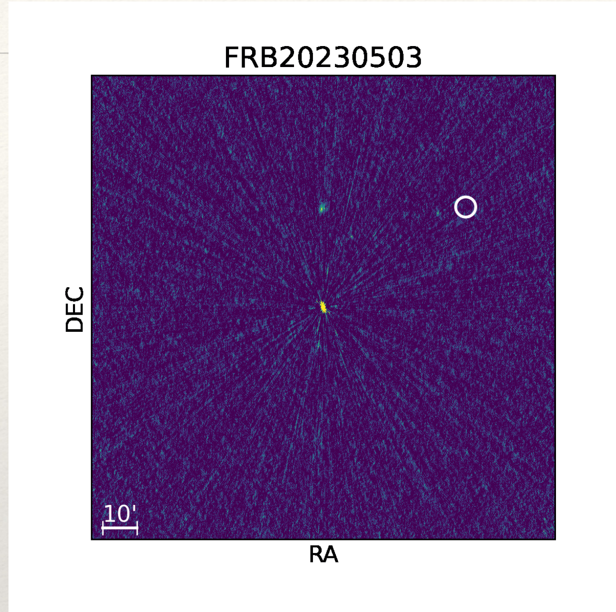
Bursts from FRB20240114A

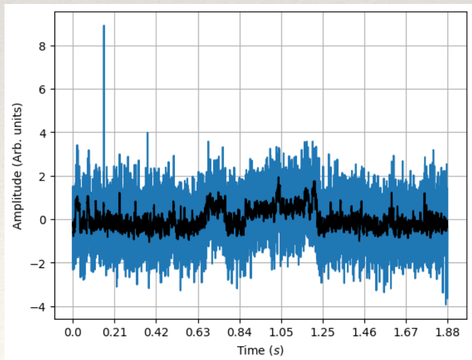


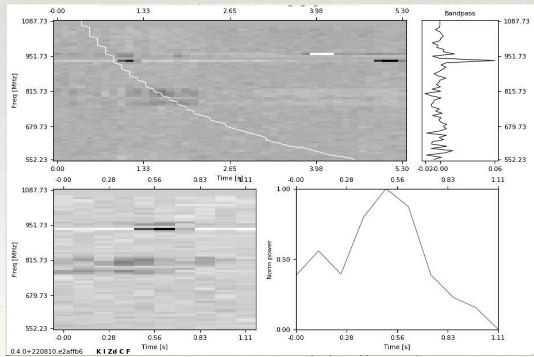
Belmonte Diaz et al. 2025

IMAGE PLANE SEARCHES

- Ideally we would image the entire sky down to whatever time resolution we wanted, to the minimum possible given the number of frequency channels.
- BUT the data rate is enormous, hence why we use beam forming (in interferometers instead — effectively reduces the number of pixels/resolution elements).
- However time domain searches start to become problematic when we get to timescales of ~1s or longer, i.e. for pulse durations of this length.
  - this is due to baseline variations/red noise in the data
  - and radio frequency interference/lightning mimicking real signals.
- There is a cross-over point, currently, also due to data rate limitations most interferometers are creating images on timescales of a few seconds (e.g. MeerKAT does 8 routinely, but under special cases can do 2).







EXAMPLE IMAGE PLANE SEARCH CODES

- One of the first was the TRAP (Swinbank et al. 2015) and extended by Antonia Rowlinson and collaborators (https://github.com/transientskp/tkp).
    - Originally developed for LOFAR, but used on wide range of telescopes including MeerKAT.
- TRON (Smirnov et al. 2025) a framework that considers both snapshot images and dynamic spectra — highly automated.
- VASTER (Yuanming et al. 2023) used for ASKAP/VAST searches
- LORDS (Rajwade et al.) includes dedispersion
- +++ others I have probably missed.

HEURISTICS

Modulation coefficient

$$V_\nu = \frac{s}{\overline{I_\nu}} = \frac{1}{\overline{I_\nu}}\sqrt{\frac{N}{N-1}\left(\overline{I_\nu^2} - \overline{I_\nu}^2\right)}$$

Swinbank et al. 2015

χ^2 coefficient

$$\eta_\nu = \frac{1}{N-1}\sum_{i=1}^{N}\frac{(I_{\nu,i} - \overline{I_\nu})}{\sigma_{\nu,i}^2} = \frac{N}{N-1}\left(\overline{wI^2} - \frac{\overline{wI}^2}{\overline{w}}\right)$$

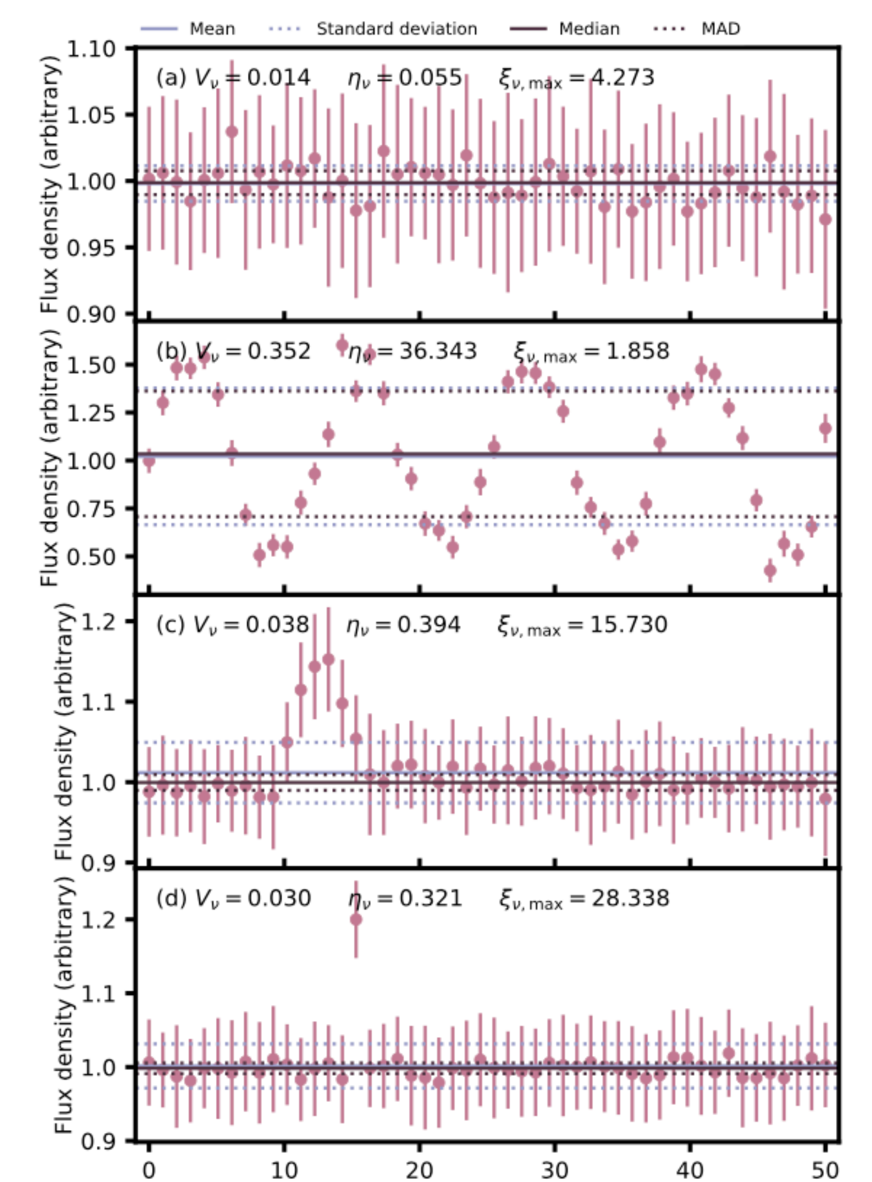
ascl:1412.011 2014

MADulation

actually use $\xi_{\nu,max}$ which is maximum

$$\xi_i = \frac{X_i - \widetilde{X}}{\text{median}\left(\left|X_i - \widetilde{X}\right|\right)}$$

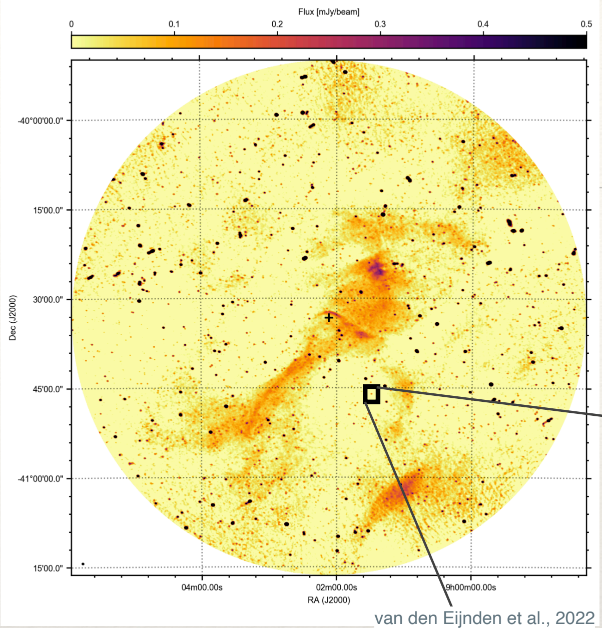
Driessen 2021



Driessen 2021

METHOD

Vela X-1 region



Generate Deep Image
& Source Find

PSR J0901-4046 — 76s pulsar

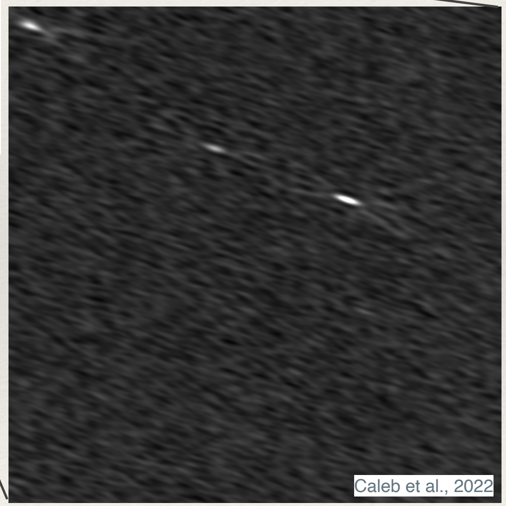


Figure Credit: Ian Heywood

Find source in snapshot images.
Difference imaging / UV subtraction being used here now…

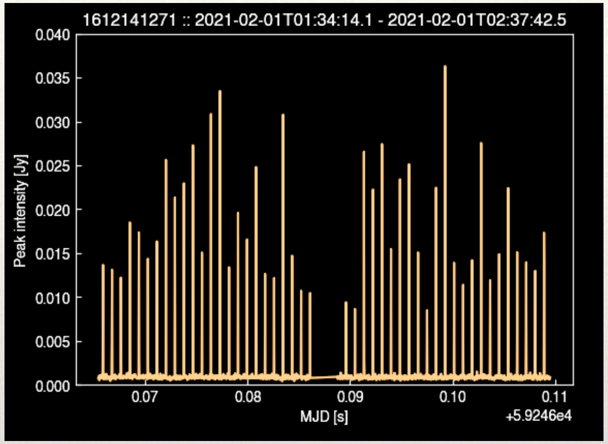
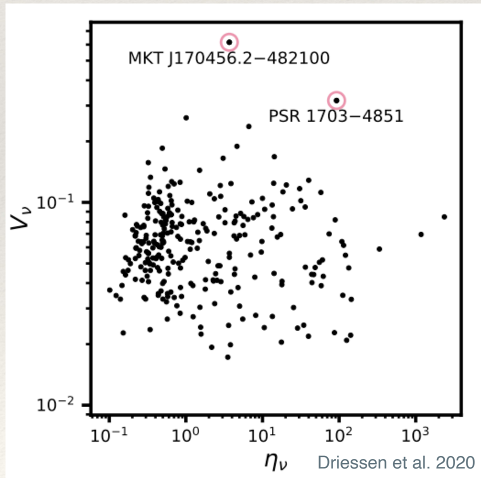


Figure Credit: Ian Heywood

Generate Light Curves & Heuristics



Look for outliers

EXTENSION EXAMPLES

LORDS - Rajwade et al.

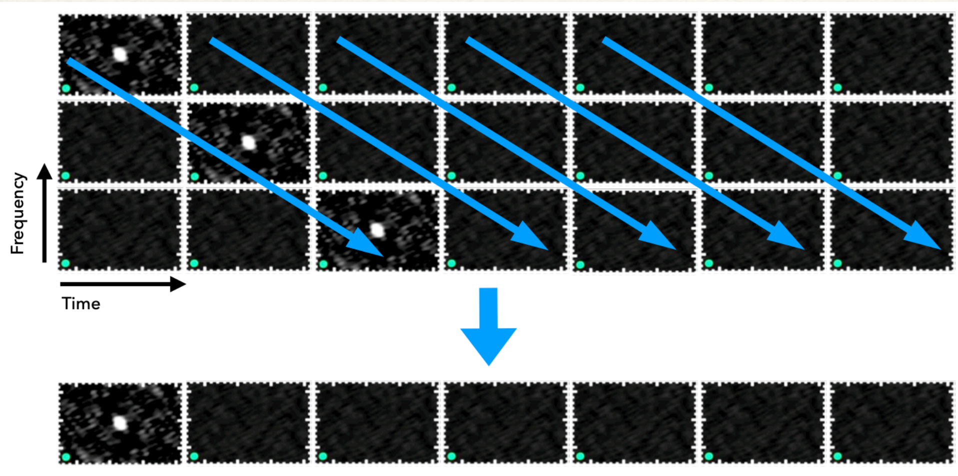
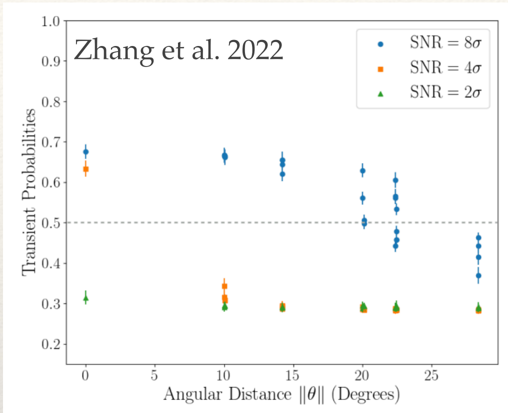
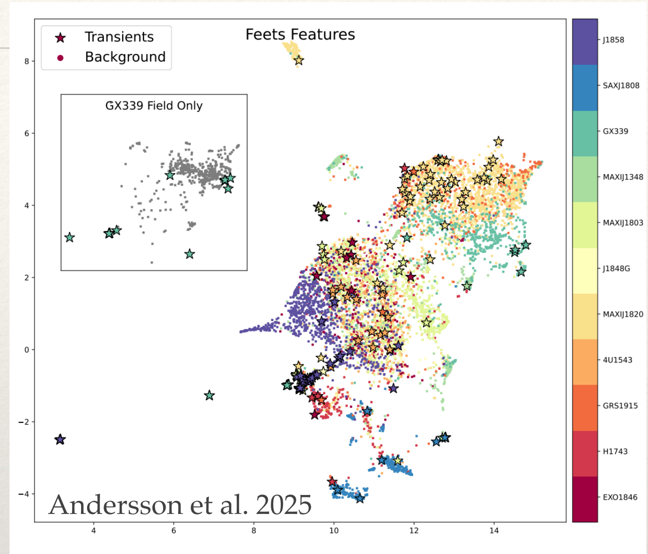


Image Plane Dedispersion (High DM / Low Freqs)





ML / Anomaly Detection / Citizen Science