The SMART pulsar survey: Past, present, & future

Bradley Meyers
(on behalf of the SMART team)

MWA Project Meeting 25-28 July 2023

How it all began

Motivation and the inception of a MWA pulsar survey.

An idea is born...

By 2017 we had moved to a stage where **high time resolution processing** was becoming **more tractable** with the Voltage Capture System (VCS)

- *Offline software coherent beamformer*
    - + MWA now included in common pulsar processing packages
- *Tools and utilities to make managing processing easier* (VCSTools)
- *Local group knowledge and understanding*
    - R. Bhat, S. Ord, S. Tremblay, F. Kirsten, G. Sleap, S. McSweeney, B. Meyers, M. Xue, N. Swainston...

MWA Phase II introduced "compact configuration" making the concept of tesselating the entire southern sky feasible - still *very* daunting!

An MWA pulsar survey

With the **wide field of view** from the tile design, **"fat" tied-array beams** from MWA Phase II Compact tile layouts, and a **multi-pixel beamformer** on the way, it all seemed possible!

And so, the **Southern-sky MWA Rapid Two-metre pulsar survey** was conceived

- Only need *<100 observations* total (still a large ask given VCS data rates)
- *Better sensitivity* than previous Southern pulsar surveys at similar frequencies
- If all VCS observations can be stored, constitutes a *voltage record of the sky*
- Provides a *reference for SKA-Low* predictions and verification

An MWA pulsar survey

With the **wide field of view** from the tile design, **“fat” tied-array beams** from MWA Phase II Compact tile layouts, and a **multi-pixel beamformer** on the way, it all seemed possible!

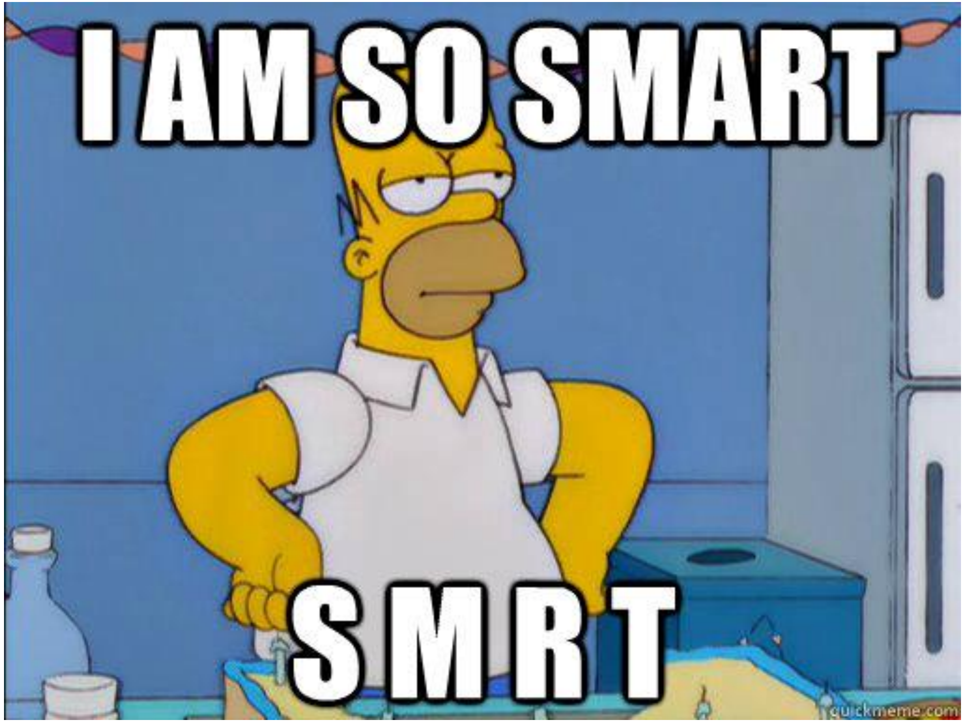
SMART

And so, the **Southern-sky MWA Rapid Two-metre pulsar survey** was conceived

- Only need *<100 observations* total (still a large ask given VCS data rates)
- *Better sensitivity* than previous Southern pulsar surveys at similar frequencies
- If all VCS observations can be stored, constitutes a *voltage record of the sky*
- Provides a *reference for SKA-Low* predictions and verification

SMRT... I mean, SMART!

Southern-sky **M**WA **R**apid **T**wo-meter pulsar survey



Current status

What we are doing right now.

How much data have we collected?

As of previous Project Meeting

- Collected 75% of observations needed to cover the visible sky
    - Dec range: +30 deg to SCP
    - RA range: <16hr and >21hr
    - Remaining observations will cover the Galactic Plane and Centre
        - Very interesting region, but also very difficult for low-frequencies!

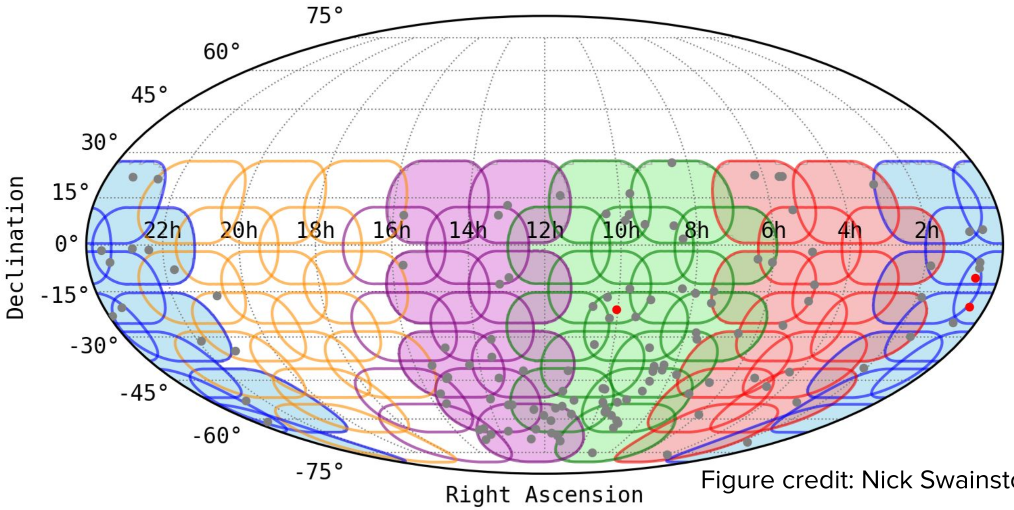


Figure credit: Nick Swainston

- 51 VCS observations = 2.15 PB of data
    - Observed between September 2018 and May 2021

How much data have we collected?

As of now

- Collected ~~75%~~ 100% of observations needed to cover the visible sky
  - **A voltage record of the Southern sky at 154 MHz**

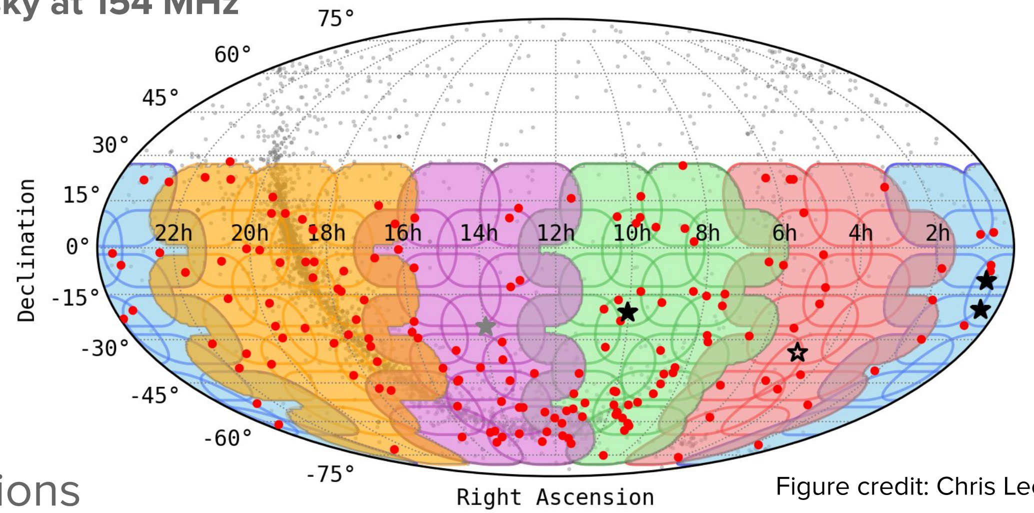


Figure credit: Chris Lee

- **3.6 PB** of voltage data
  - +10 TB of calibration observations (before & after each VCS obs.)

- **51 Legacy + 19 MWAX** observations
  - Legacy: observed between September 2018 and May 2021
  - MWAX: observed between January and July 2023

How much data have we processed?

- We have successfully conducted a *proof-of-concept “first-pass” survey*
  - Only analysed the **first 10 minutes** of each observation
  - **No RFI mitigation** (beyond forming tied-array beams)
  - Used a very **coarse dedispersion plan** (2500 trials)
  - Did only a **basic periodicity search** across each time series’ power spectrum

- Processed 80% of the Legacy observations
  - Corresponds to only 8% of total data volume
  - Only a small fraction of the produced candidates have been inspected! (<10% of the 8%...)
  - Still **found a handful of new pulsars**...

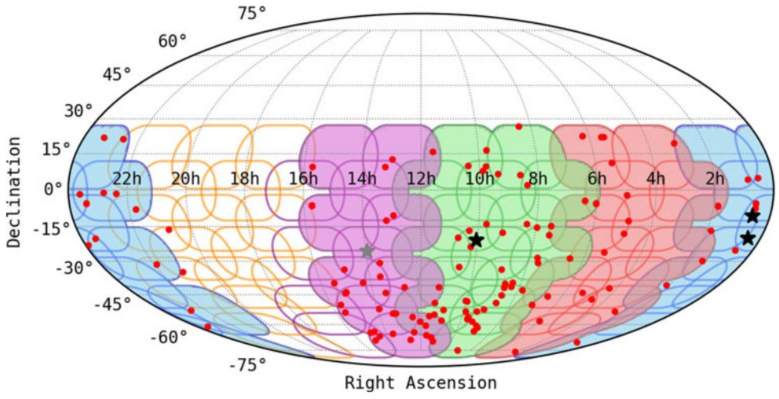


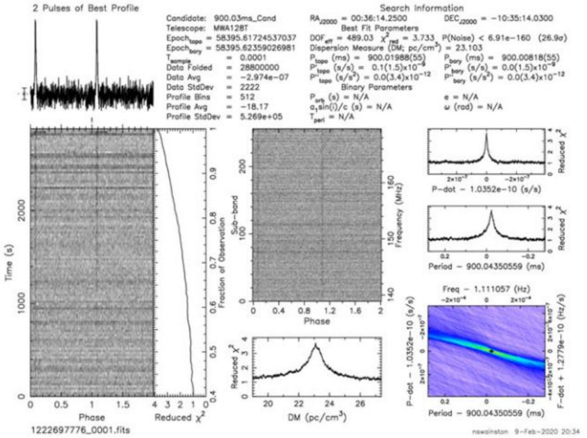
Fig. 1 in Bhat et al. 2023b

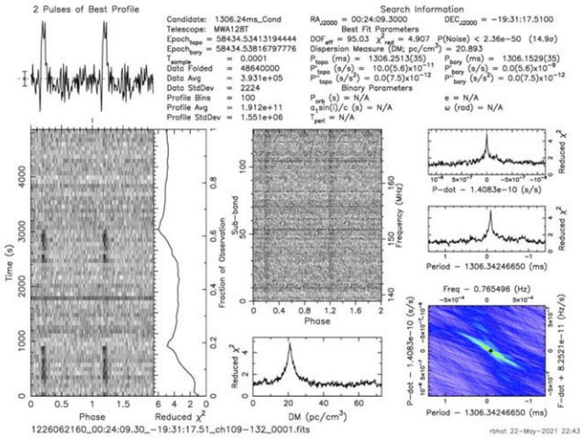
Pulsars we have discovered

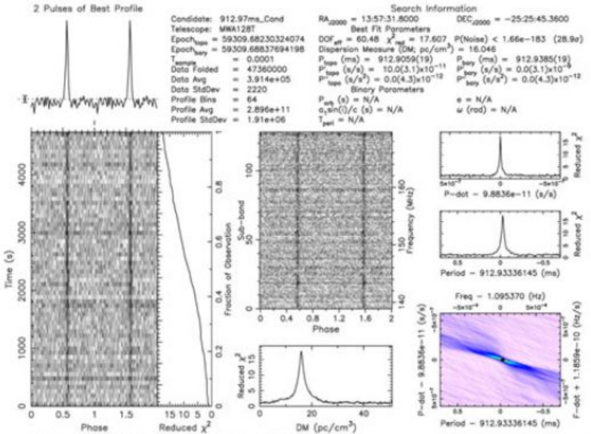
- Discovered:
  - **J0036-1033**; typical, but has a steep spectral index and low-luminosity
  - **J1002-2044**; typical, slightly steep spectral index, possibly another low-luminosity object
  - **J0452-3418**; newest addition, interesting emission phenomenology
    - *See Garvit Grover's talk for GMRT+MWA follow-up!*

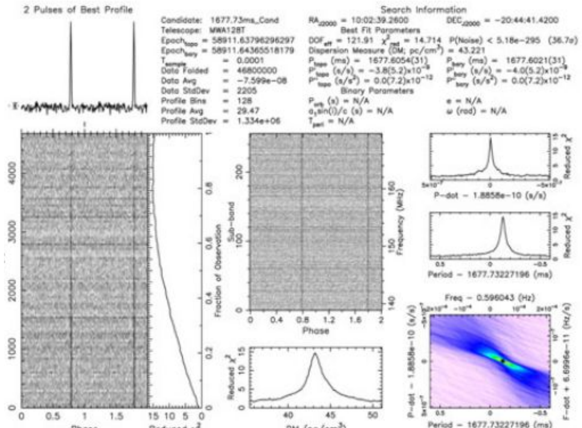
- Re-discoveries:
  - **J0026-1955**; a GBNCC candidate, initially detected in a grating lobe, sub-pulse drifting
    - *See Parul Janagal's poster for GMRT follow-up!*
  - **J1357-2530**; actually detection of mislabeled PSR J1358-2533 with vastly different P and DM!
    - GMRT imaging was required to nail down position and hence update the name

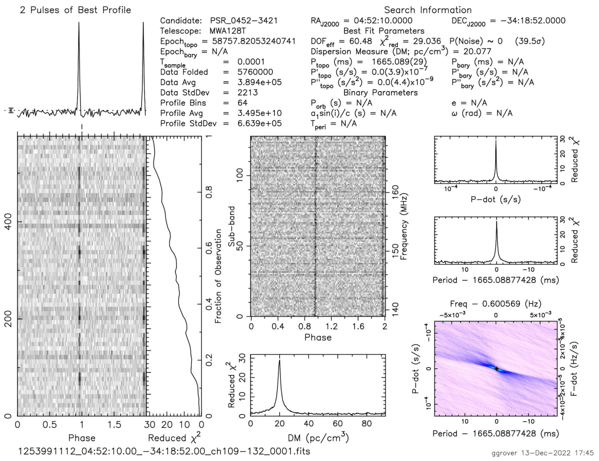
Pulsars we have discovered









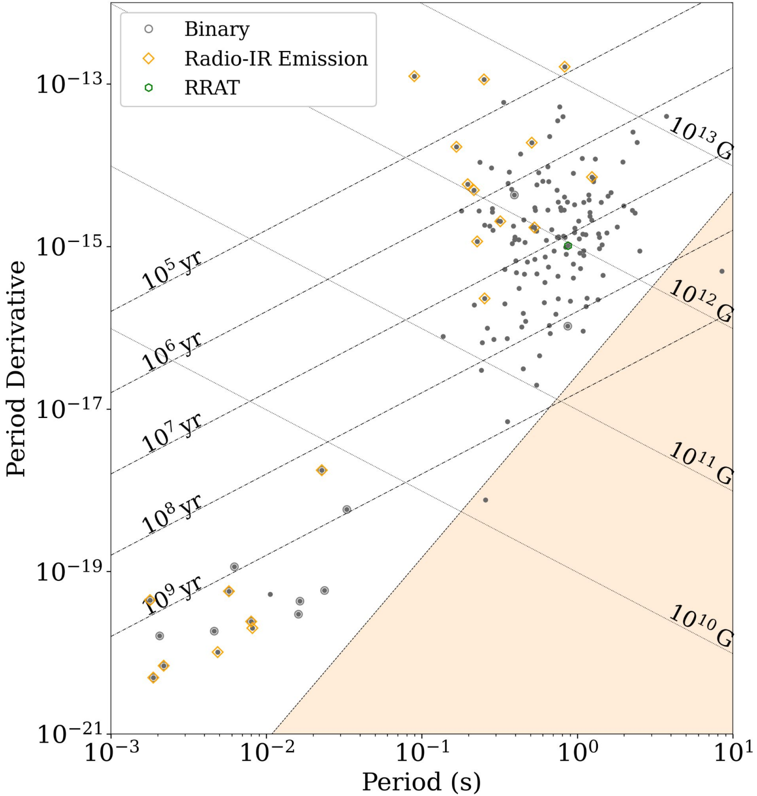


Other pulsars we have detected with SMART

We have also **detected many known pulsars** in SMART observations

- Pulsars: 170
- Millisecond pulsars: 15*
- Binary pulsars: 15*

Many of these are the **first detections below 300 MHz**!



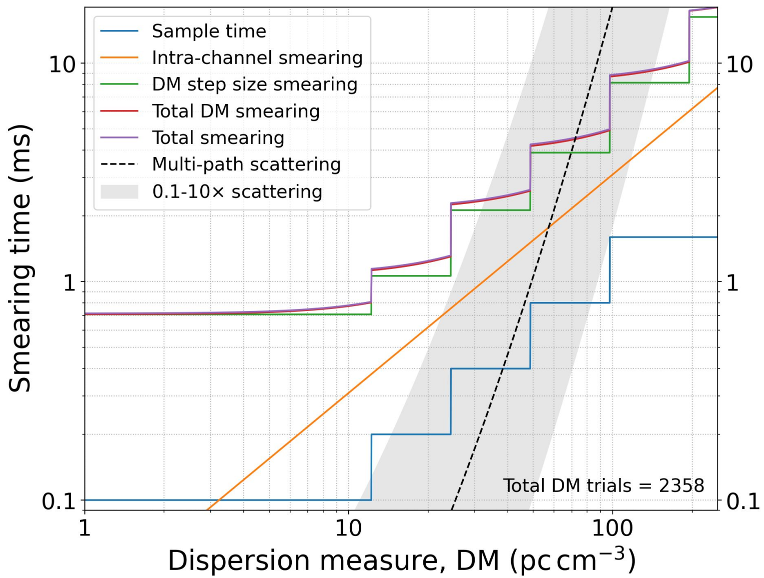
Credit: Chris Lee

*Not mutually exclusive.

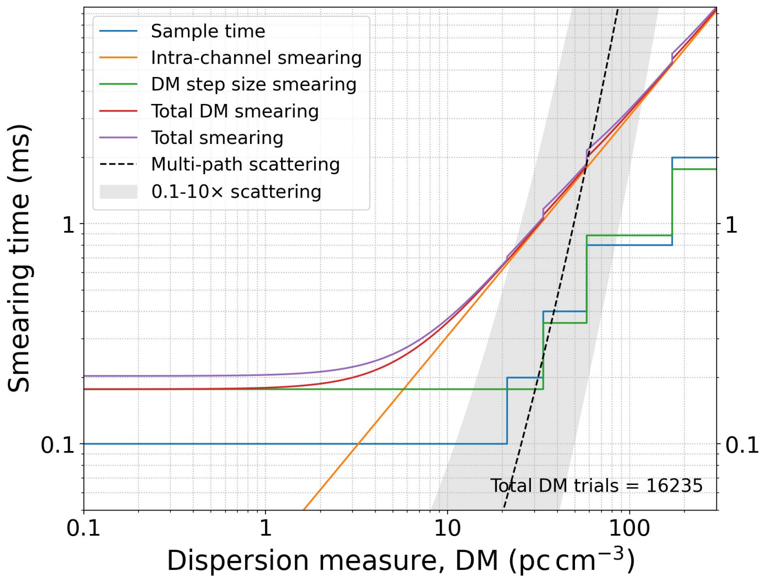
Planning the next stage of SMART processing

Optimising dedispersion plans

- Must make informed decisions based on instrumental limits
  - **Reduce artificial data degradation** by increasing DM precision
  - **Cost/benefit compromise** in terms of data generation/degradation

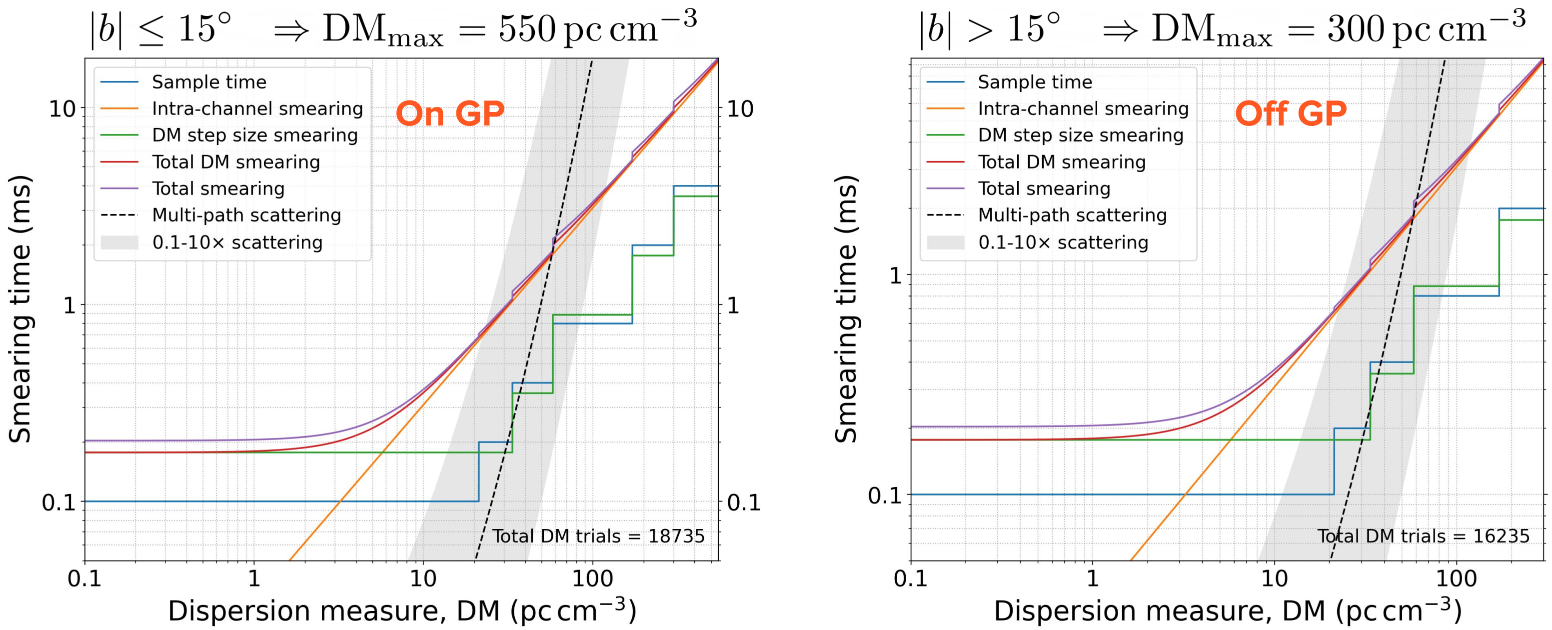


7-8x trials



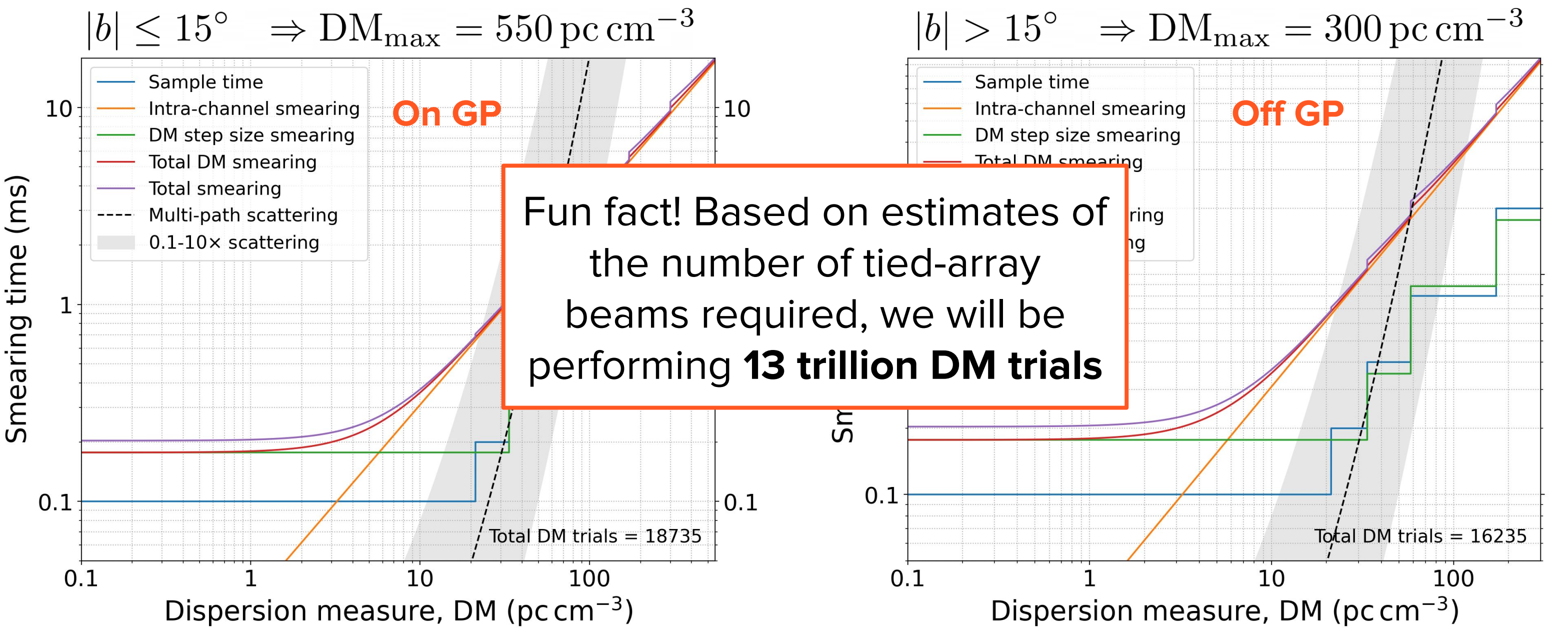
Optimising dedispersion plans

- Use two **different plans based on Galactic latitude**
  - DM rapidly increased towards the Galactic plane



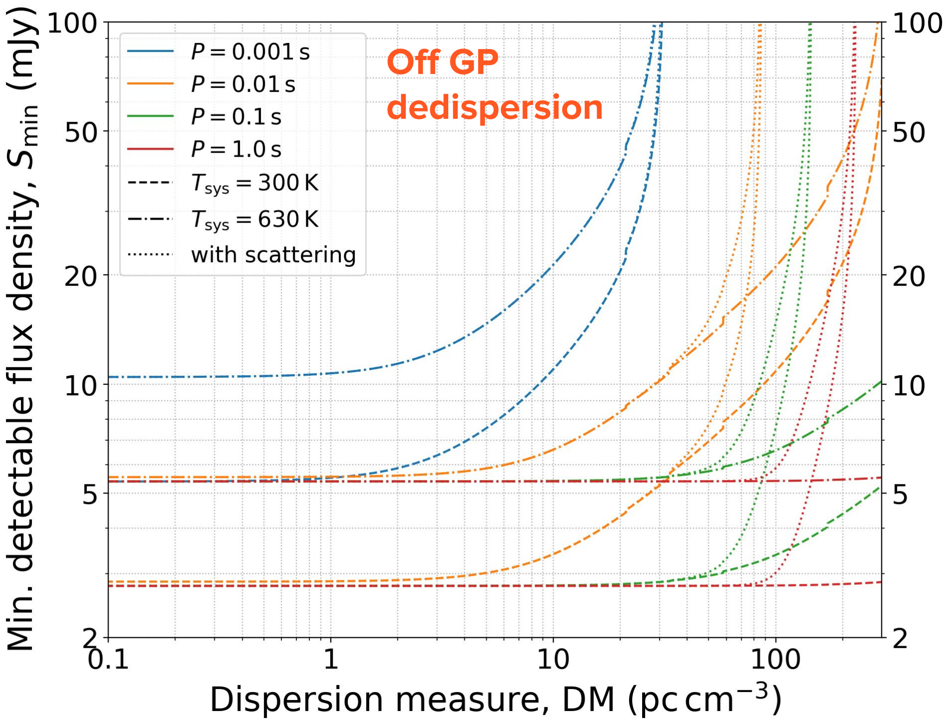
Optimising dedispersion plans

- Use two **different plans based on Galactic latitude**
  - DM rapidly increased towards the Galactic plane



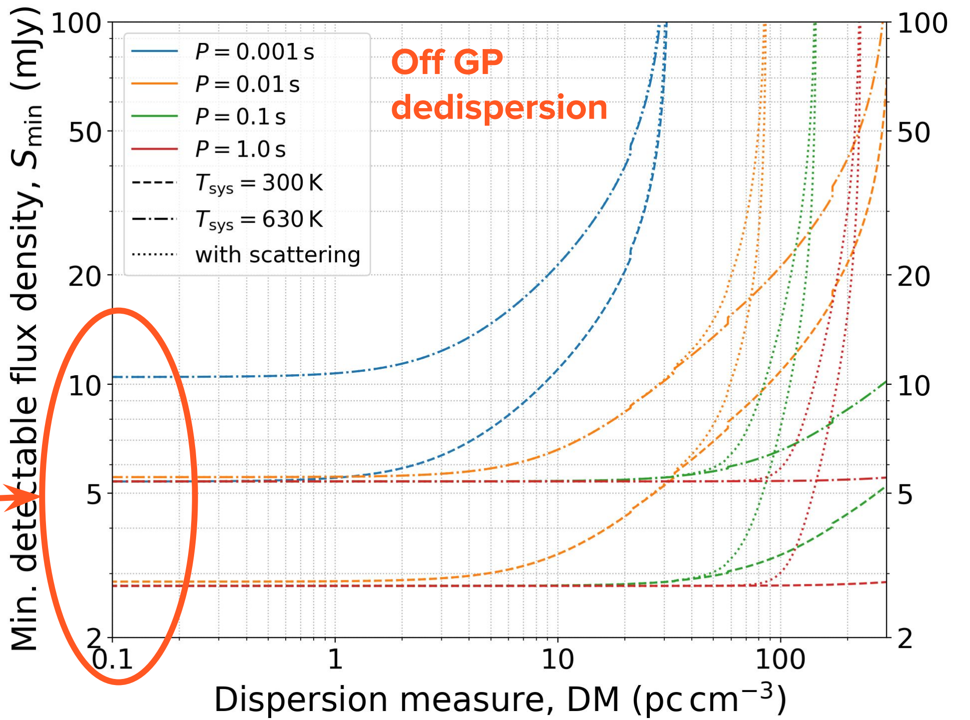
Updating our sensitivity estimates

- The dedispersion plan directly impacts theoretically achievable sensitivity
  - Also depends on pulsar period, pulse duty cycle, dwell time, and the assumed gain and system temperature



Updating our sensitivity estimates

- The dedispersion plan directly impacts theoretically achievable sensitivity
    - Also depends on pulsar period, pulse duty cycle, dwell time, and the assumed gain and system temperature



Even at a DM of 0.1 pc cm^{-3}, our minimum effective time resolution is about 250 μs, thus our minimum detectable flux density is worse than the “ideal” (for periods <15 ms).

Need *semi-coherent* dedispersion!

Bringing our database and web-app online

- Database to manage processing stages and outputs
    - Developed through ADACS (special thanks to James Tocknell)
- REST API framework linked to a web application
    - Programmatic interactions possible during Nextflow tasks (and beyond)
    - Website with documentation, survey status and candidate ranking portal
- Currently live, but still ironing out some issues
- Eventual location for SMART data product releases and hosting

https://apps.datacentral.org.au/smart/

How should we *actually* search for pulsars?

- For our entire *first-pass processing*, we used about **300 kSU***
- For a *naive second-pass*, need about **250 kSU per observation**
  - That's 70 x 250 kSU = 17500 kSU!
  - A very-well graded ASTAC proposal may achieve 500-600 kSU
  - So, SMART would take… **23 years to complete**
- Need to make difficult **decisions about the scope of "second-pass"**
  - Is the proposed dedispersion plan absurd? What about RFI mitigation?
  - Can we afford any acceleration searches? Can we afford single-pulse searches?
  - Do we only process a portion of the sky? Which part? (High-|*b*| is easier, but fewer pulsars…)

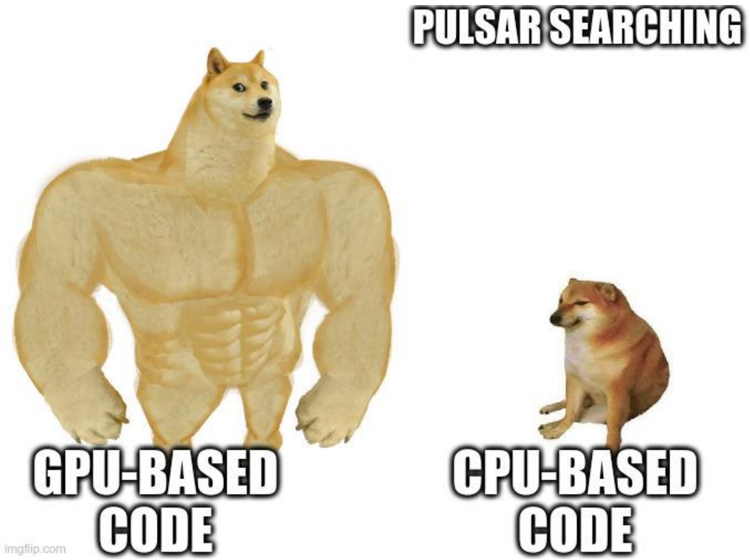
*On OzSTAR, **1 kSU** is about **1000 CPU-hours** (or 125 GPU-hours)

Looking to the future

Accelerating our search software

The *searching methods* are now our **biggest computational bottleneck**

- Just like our beamforming code, moving to **GPU-accelerated techniques** will be **critical in the long-run** to efficiently search such a large data volume



+ Actually it's also about Intelligent algorithmic design, efficient resource use, etc., etc.

Reprocessing of SMART data

SMART data are voltages, retains *maximum flexibility* for various processing tasks

- Relies on being able to **archive the SMART data** in perpetuity
    - Phase 1 ARDC; 2018-2019 campaigns; allocation of 1081 TB (~40% of sky).
    - Phase 2 ARDC; 2020-2021 campaigns; allocation of 977 TB (~40% of the sky)
    - *??????????????; 2023 campaign; Galactic Plane coverage (final ~20% of the sky)*

`https://asvo.mwatelescope.org/collections`

- Historically, **pulsar surveys have been processed many times over**
    - Motivated by new technologies, algorithmic developments, and/or phenomena
    - FRBs, RRATs, new MSPs and binaries…
- Next **generation telescope survey data *cannot* be kept** for any extended time
    - Will force creative solutions that will also aid in reprocessing SMART data!

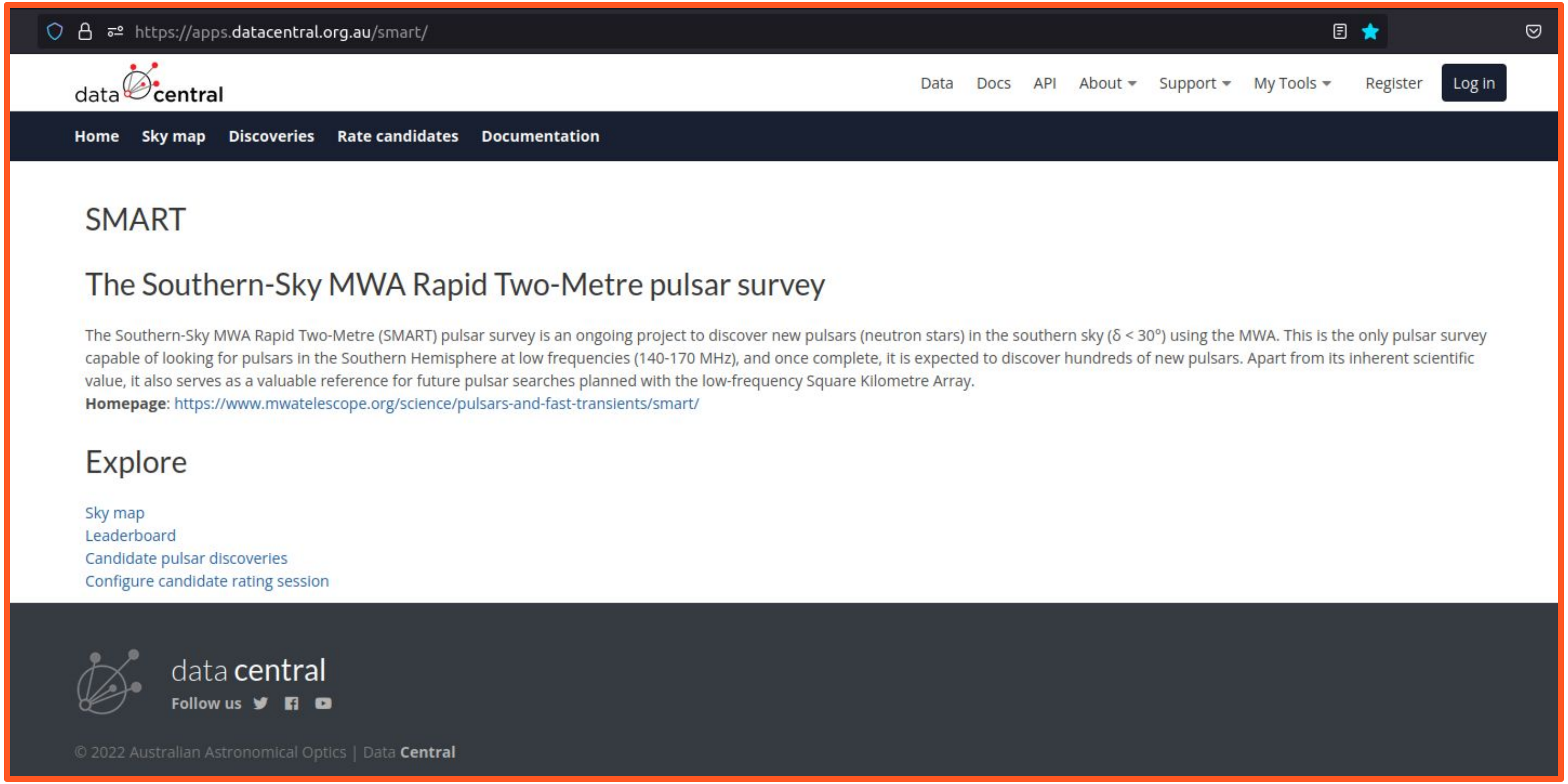
SMART team: R. Bhat, B. Meyers, N. Swainston, S. McSweeney, M. Xue,, M. Sokolowski, S. Dai, S. Kudale, W. van Straten, R. Shannon, S. Tremblay, ...

Summary

- **Completed data collection**
  - 3.6 PB of VCS observations
  - Voltage record of the visible sky
- **Critical decisions about next processing steps**
  - The reality is we *can't* naively scale
  - Innovation & development required!
- **Handful of pulsars already!**
  - Promising for total survey yield
  - On track for >100 pulsars

Additional slides

DataCentral database and web-app support



https://apps.datacentral.org.au/smart/

data central

Data Docs API About Support My Tools Register Log in

Home Sky map Discoveries Rate candidates Documentation

SMART

The Southern-Sky MWA Rapid Two-Metre pulsar survey

The Southern-Sky MWA Rapid Two-Metre (SMART) pulsar survey is an ongoing project to discover new pulsars (neutron stars) in the southern sky ($\delta < 30°$) using the MWA. This is the only pulsar survey capable of looking for pulsars in the Southern Hemisphere at low frequencies (140-170 MHz), and once complete, it is expected to discover hundreds of new pulsars. Apart from its inherent scientific value, it also serves as a valuable reference for future pulsar searches planned with the low-frequency Square Kilometre Array.
Homepage: https://www.mwatelescope.org/science/pulsars-and-fast-transients/smart/

Explore

Sky map
Leaderboard
Candidate pulsar discoveries
Configure candidate rating session

data central
Follow us

© 2022 Australian Astronomical Optics | Data Central

Practicality of search methods

Simulations testing FFT vs. FFA searching for SMART-like data suggests

- FFT searching should be limited to P < 10 seconds
  - Red noise dominates, can mitigate but information fundamentally lost
- FFA searching most effective at P > 1 second
  - Red noise removal in time domain, but still removing potential information
  - Efficiency grows for large P due to re-use of data in-memory

Proposed strategy

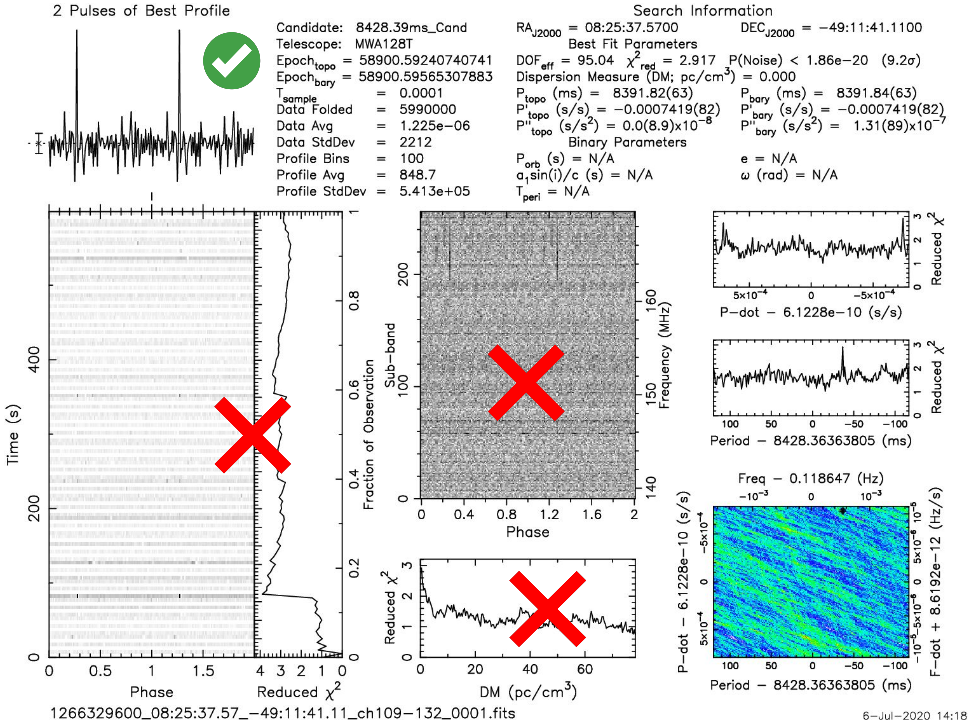
- FFT for: **1 ms < P < 2 s**
- FFA for: **1 s < P < 120 s**

Current state of searching software

- Almost all CPU-based, or GPU-based is developed for higher frequencies
    - A big factor for MWA is the dispersion sweep is enormous which requires equivalently enormous GPU resources to process optimally
    - **AstroAccelerate** (developed for SKA) notionally works on MWA-like data, but performance is drastically reduced due to memory limitations on all but the biggest of GPUs
    - **Peasoup** (Swinburne group) also developed for higher-frequency surveys, probably cannot handle MWA-like data without significant changes (which likely decreases efficiency)
    - **GPU-PRESTO**(?) has been on the table for years, but there is no supported/well-maintained version and no plan towards it in the near-future

Include RFI mitigation

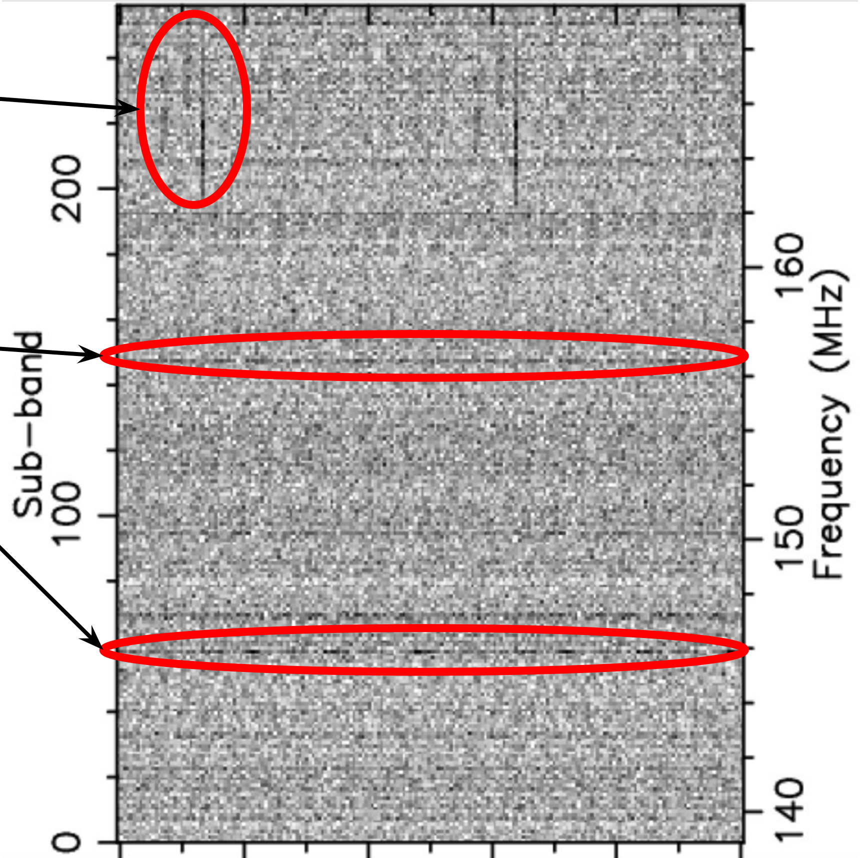
- Clear examples in shallow pass where RFI causes grief
    - While the observatory site is exquisitely RFI quiet, we still see impulsive and periodic varieties
    - RFI can mimic single pulses *and* periodic pulsar signals - they are by far the majority of candidates
    - Exploring options of how to deal with these in an effective and computationally feasible way
        - RFI “excision” can happen after candidate creation, too



RFI in pulsar data - an example

Impulsive interference
(band-limited, 7 MHz wide)

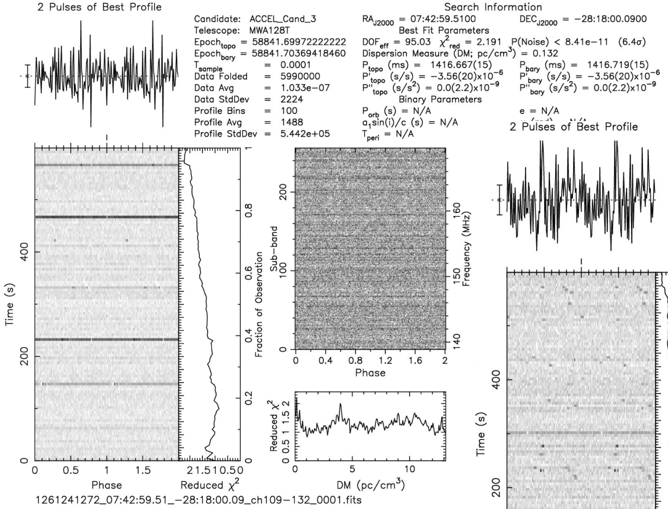
Narrow band, periodic interference (maybe not pure tone, but periodic enough to cause trouble)

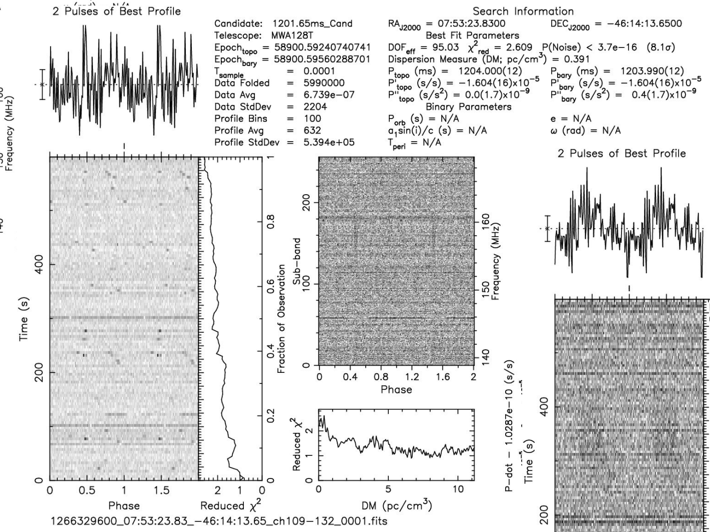


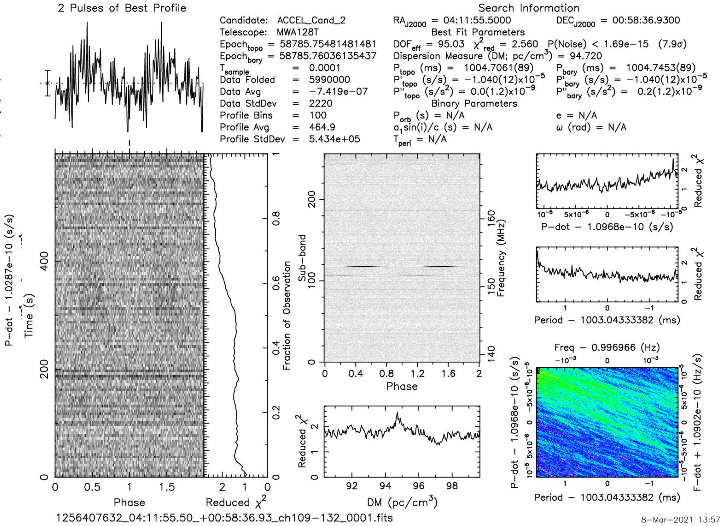
Common periodic signals at zero DM leak into other DM trials because they are much brighter than typical pulsars.

We need to identify and mask these signals ("birdies") *before searching*!

Other RFI examples - the horrifying truth



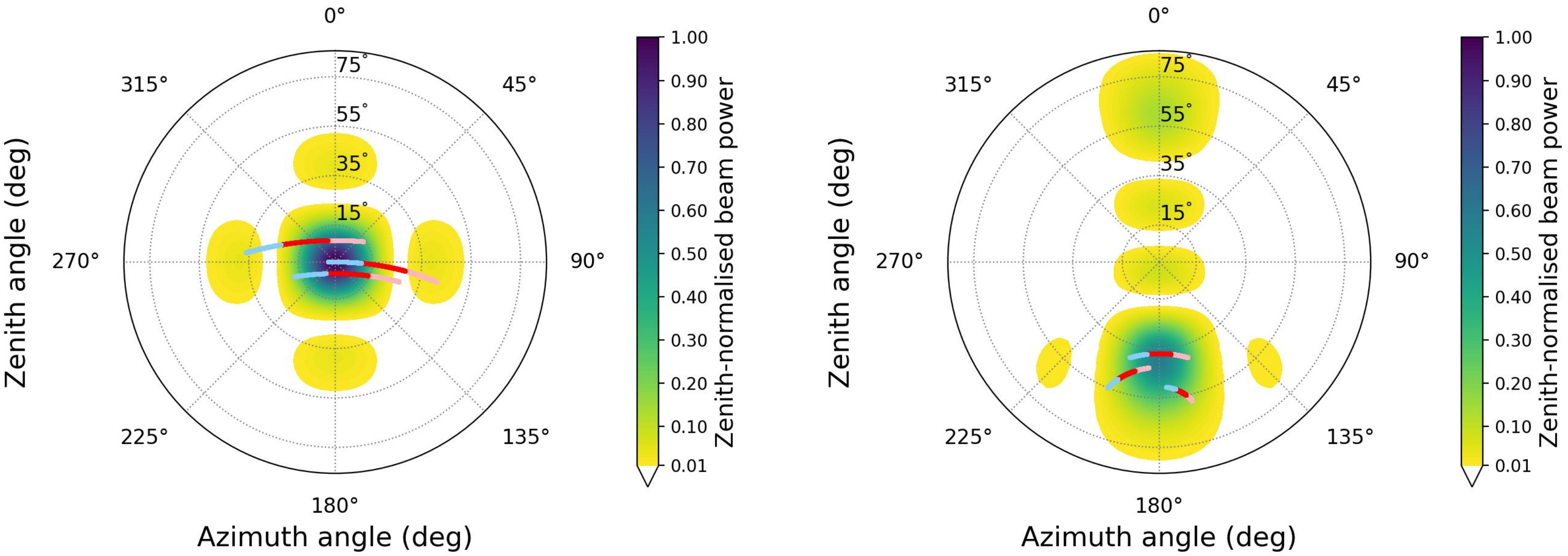




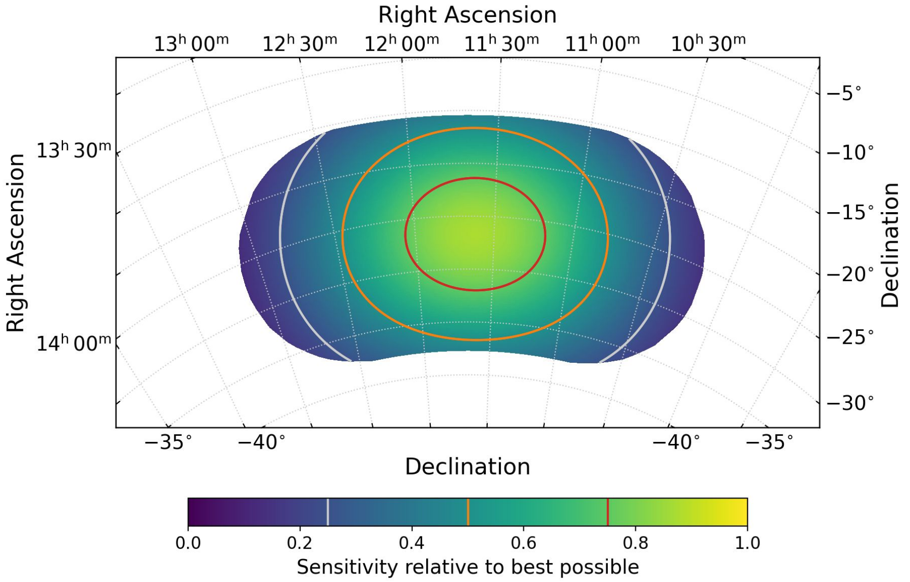
Consider realities of our observing strategy

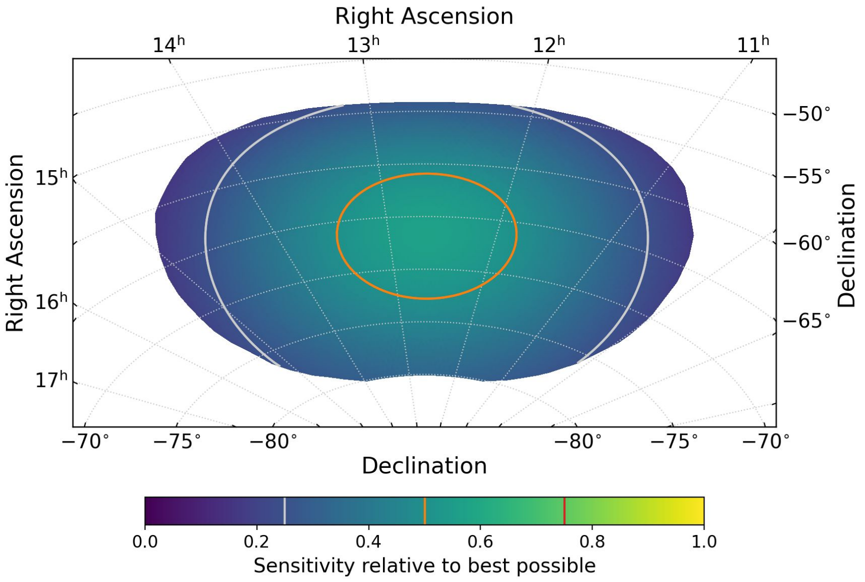
- How does a static primary beam combined with a sky-tracking tied-array beam impact sensitivity?
  - SMART is unique among pulsar surveys - inform considerations for SKA-Low!

1. Computationally inefficient to beamform all sky-positions for full 80 mins
   - Why bother forming tied-array beams when source is in a primary beam null?
2. Tracking position for full duration may *decrease* sensitivity
   - Consequence of 1. is that we essentially just add noise to the data when searching
   - But… longer tracks also improve chances of detecting bright single pulses and long-period pulsars
3. Sensitivity will be position dependent
   - Primary beam sky positions (Alt/Az) *in addition to* target sky position (RA/Dec) affects sensitivity



Observing near vs. far-from zenith: beam pattern & source traces





Observing near vs. far-from zenith: relative sensitivity