

The High Time Resolution Universe Legacy Survey (HTRU)

Cherry Ng

July 20th 2011

Supervisors: Michael Kramer, David Champion
Max-Planck-Institut für Radioastronomie



MAX-PLANCK-GESellschaft

Max-Planck-Institut
für Radioastronomie

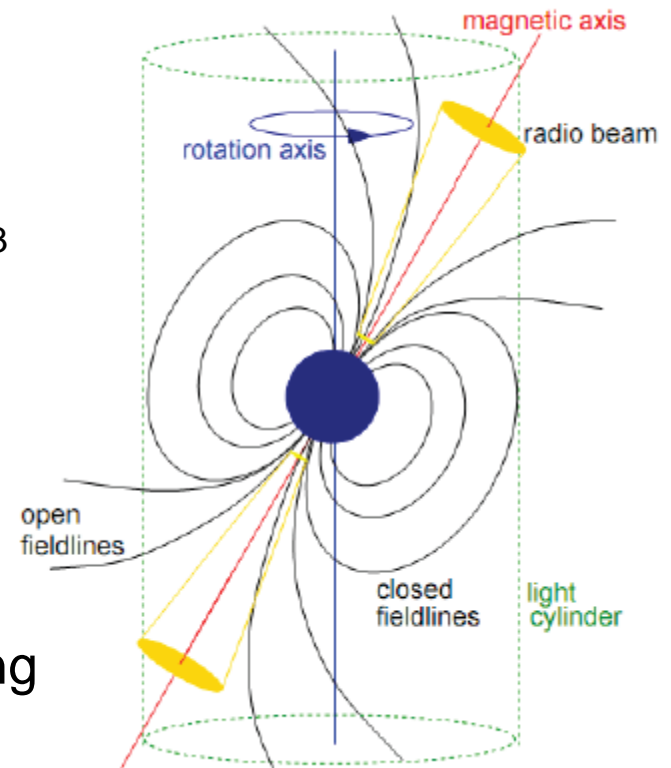


Outline

- ❖ Motivations – What is a pulsar? Why search?
- ❖ Standard search pipeline
- ❖ Previous survey – PMPS
- ❖ Specifications of HTRU
- ❖ Survey status & data processing
- ❖ Current work - RFI mitigation

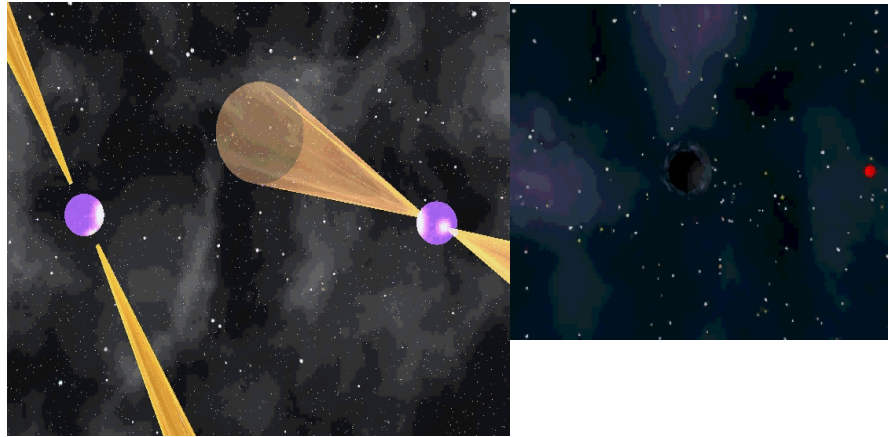
What is a Pulsar?

- Rapidly rotating Neutron Stars
- Radius ~ 10 km, mass ~ 1.4 solar mass
 - central densities: order of a billion tons / cm^3
- Highly magnetized ($10^8 \sim 10^{14}$ Gauss)
 - emit steep spectrum radio emission aligned with β -field axis
- Misalignment of rotational and β -axes → pulsing

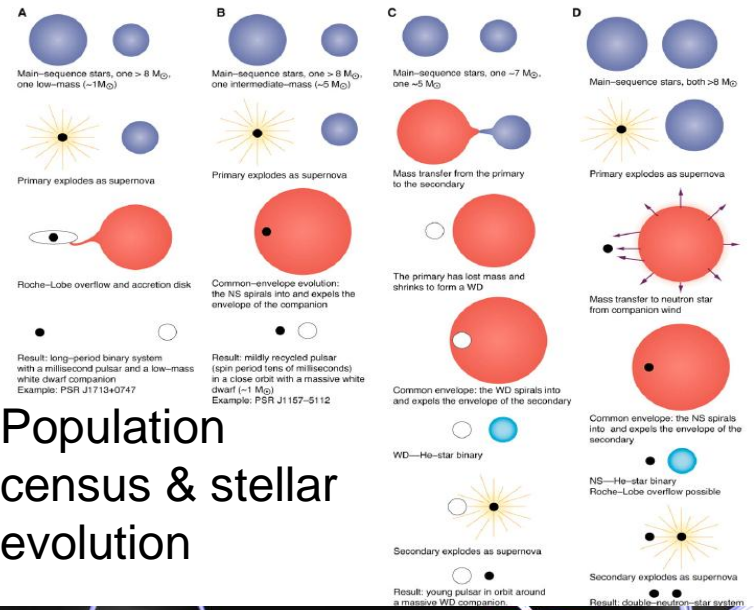


Large mass → consequent rotational stability → accurate clocks
→ many astrophysical applications!

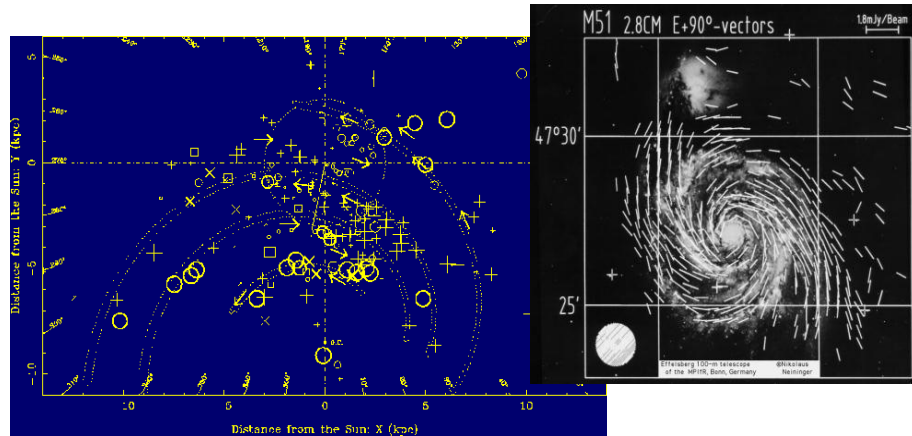
Motivations – Why search for pulsars?



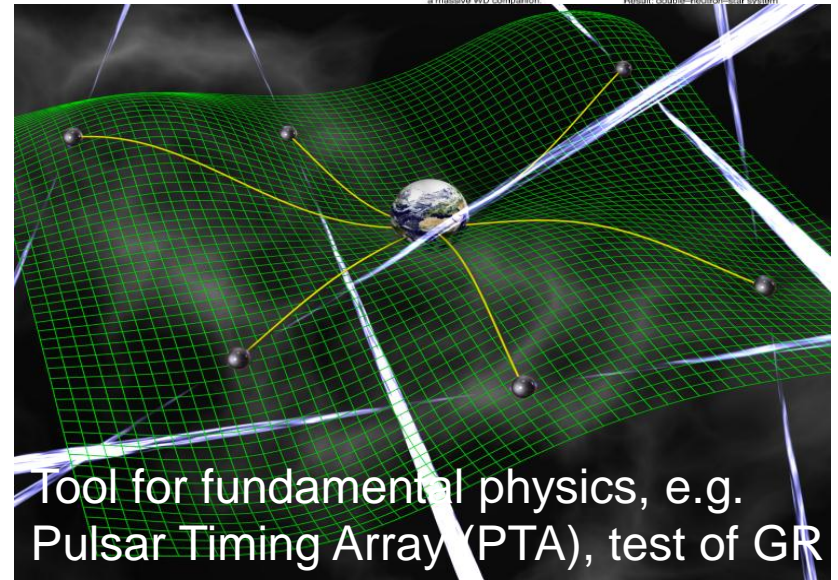
Exotic objects: NS-BH binaries, Planetary Systems, Glitches, RRATs, magnetars ...



Population census & stellar evolution

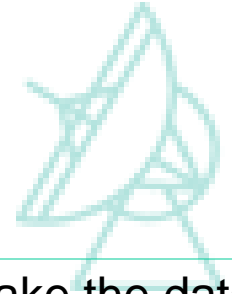
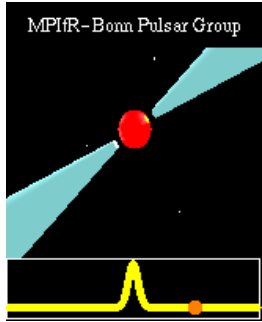


Galactic probe: ISM, Magnetic field



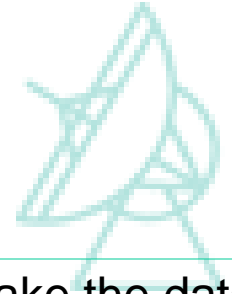
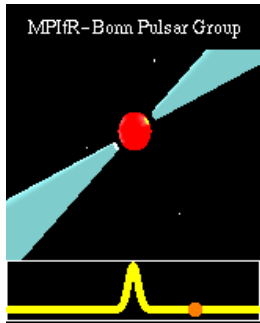
Tool for fundamental physics, e.g. Pulsar Timing Array (PTA), test of GR

Standard search - Sample pipeline



Take the data

Standard search - Sample pipeline

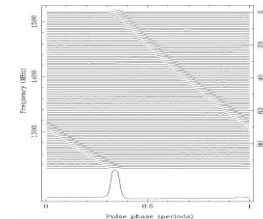


Take the data

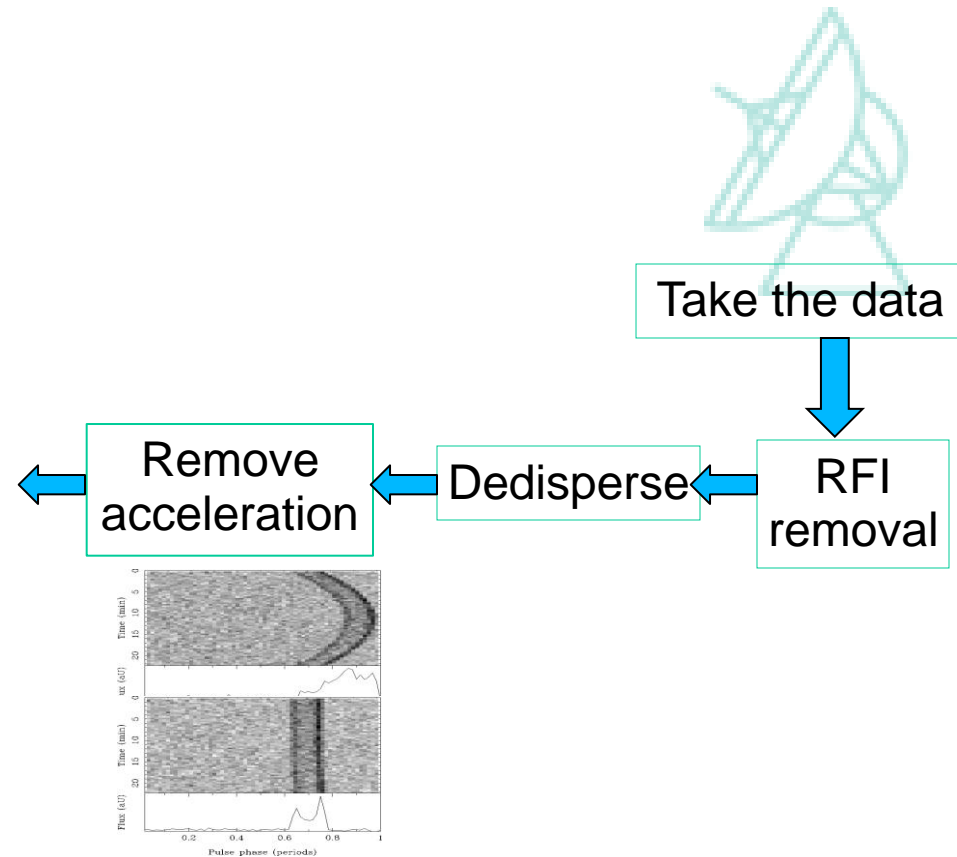
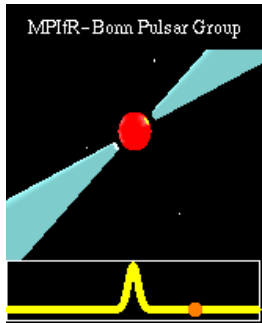


RFI
removal

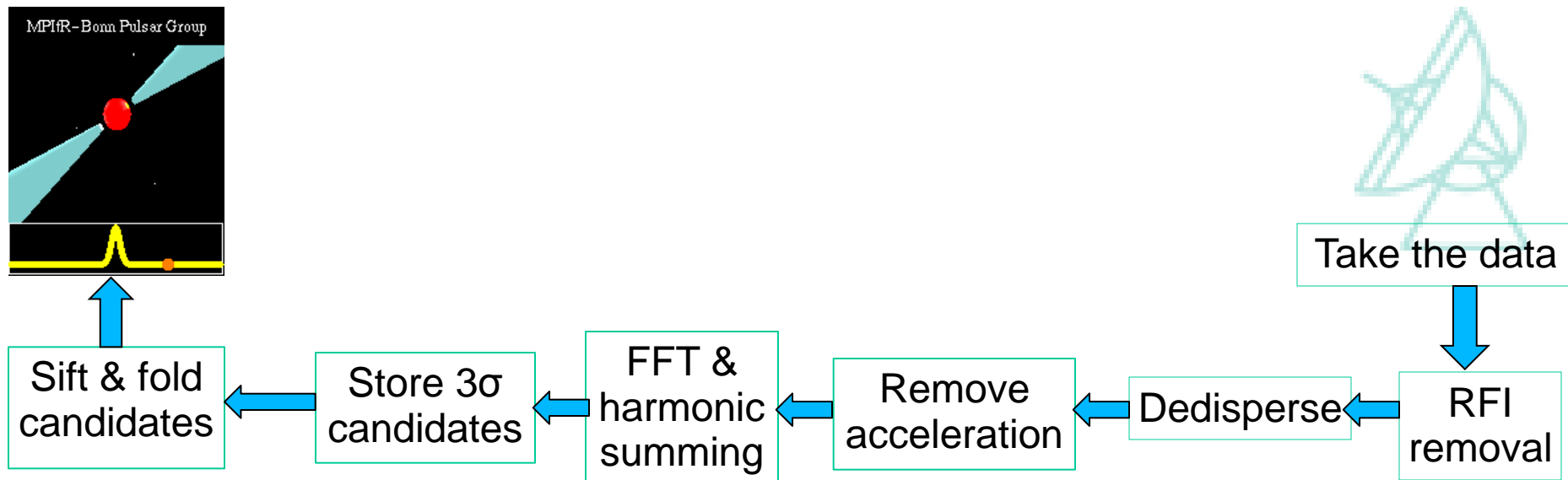
Dedisperse



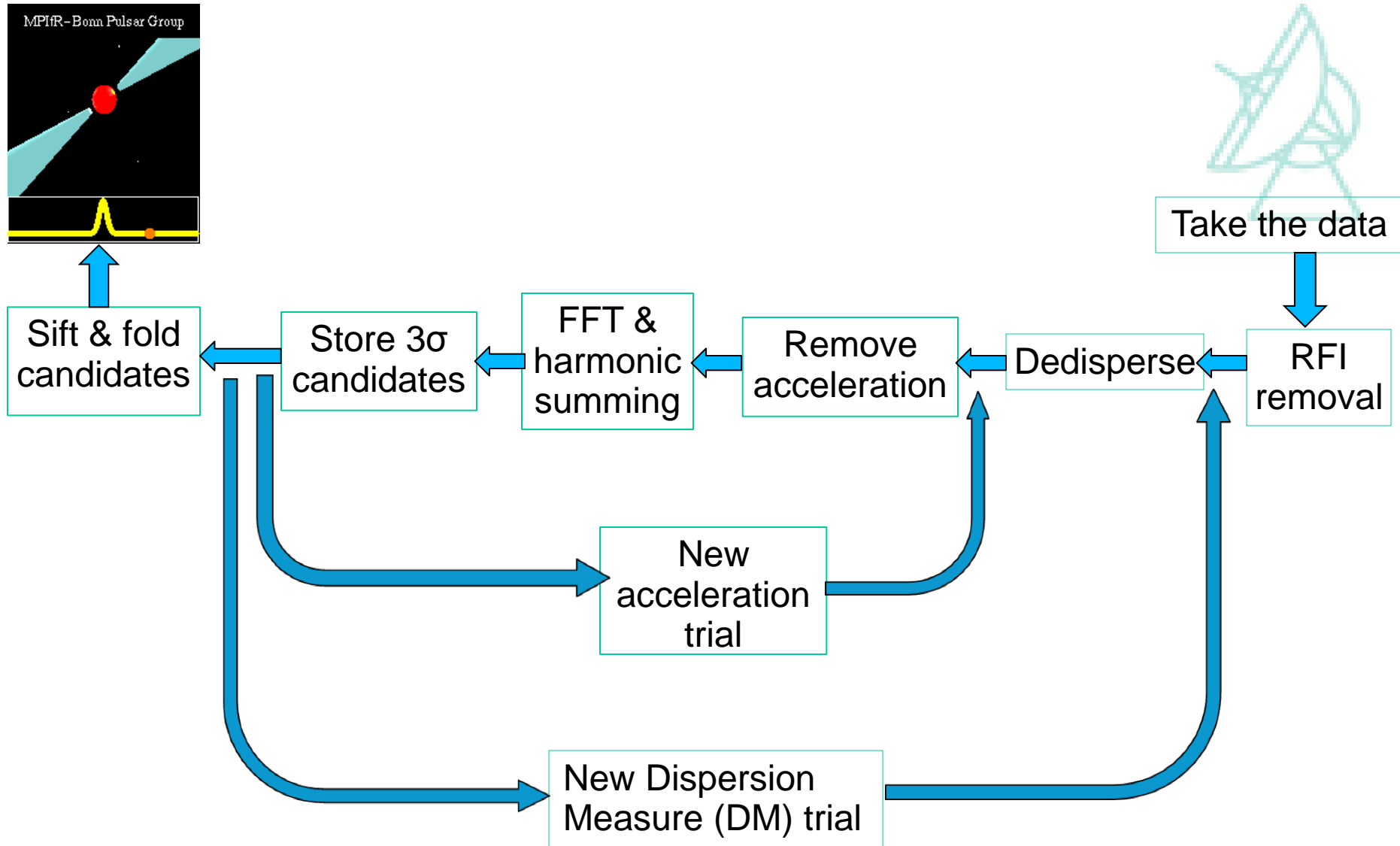
Standard search - Sample pipeline



Standard search - Sample pipeline

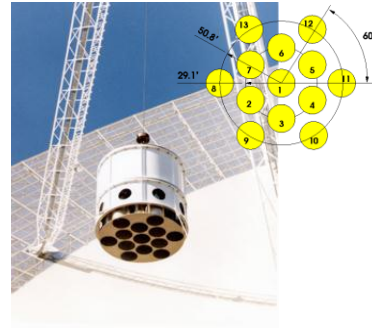


Standard search - Sample pipeline



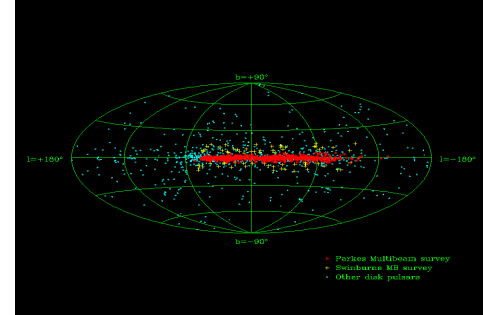
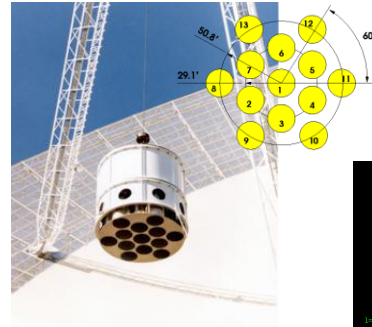
Previous survey - PMPS

- commenced in 1997, completed in 2003
- 20-cm Multi-beam receiver



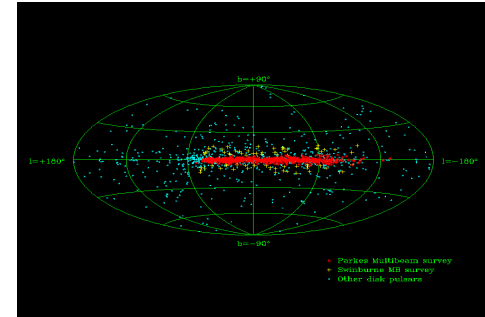
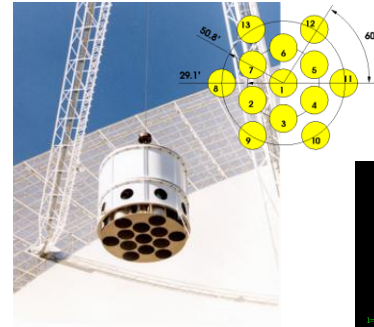
Previous survey - PMPS

- commenced in 1997, completed in 2003
- 20-cm Multi-beam receiver
- Discoveries:
 - ~ 60% of all PSR known, 6 DNSs, Double-PSR (from sister survey), first RRATs ...



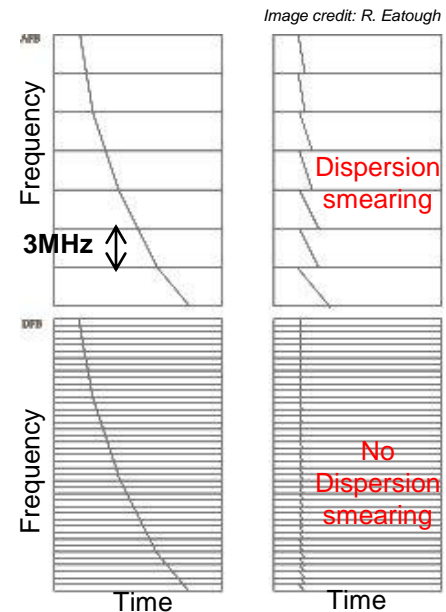
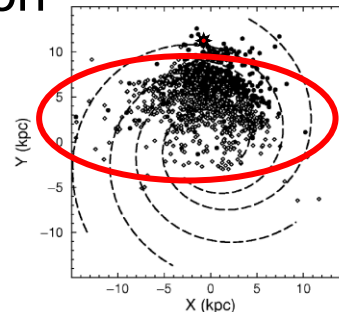
Previous survey - PMPS

- commenced in 1997, completed in 2003
- 20-cm Multi-beam receiver
- Discoveries:
 - ~ 60% of all PSR known, 6 DNSs, Double-PSR (from sister survey), first RRATs ...



Limitations:

- Modest sampling time (250 μ s)
- Low freq resolution (3MHz)
 - dispersion smearing & low time resolution
 - miss high DM MSPs
- Only Southern Hemisphere
 - latitude $> 38^\circ$ N virtually unsearched

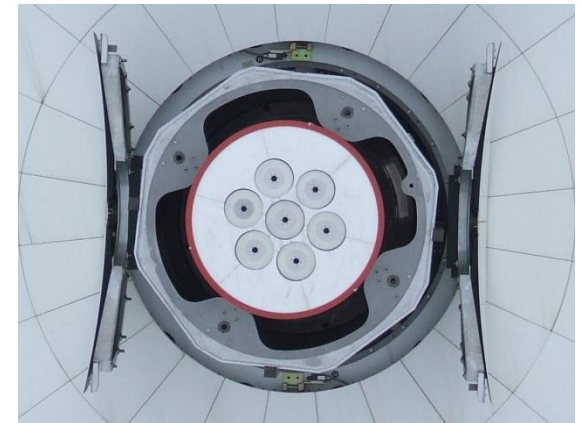


Specifications of HTRU

	Northern Survey	Southern Survey
Start date:	Summer 2010	Early 2008
Telescope:	Effelsberg-100m	Parkes-64m
Sky coverage:	$\delta > 0^\circ$	$\delta < +10^\circ$
Integration time:	Low-lat: 1500 s Mid-lat: 180 s High-lat: 90 s	Low-lat: 4300 s Mid-lat: 540 s High-lat: 270 s
Receiver:	7-beam 1.4-GHz receiver	13-beam 1.35-GHz receiver
Backend:	Pulsar Fast Fourier Transform Spectrometer (PFFTS)	Berkeley-Parkes-Swinburne Recorder (BPSR)
Bandwidth:	300MHz	340MHz
No. of channels:	512	1024
Freq resolution:	0.58MHz	0.39Mhz
Time resolution:	54 μ s	64 μ s
No. sky pointings:	~ 180,000	~ 43,000
Data sizes:	~ 5 petabytes	~ 1 petabyte



Pulsar Fast Fourier Transform Spectrometer (PFFTS)



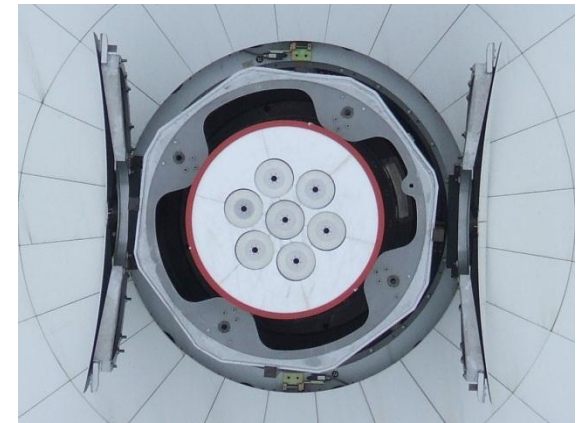
New 7-beam 21-cm primary focus receiver at Effelsberg

Specifications of HTRU

	Northern Survey	Southern Survey
Start date:	Summer 2010	Early 2008
Telescope:	Effelsberg-100m	Parkes-64m
Sky coverage:	$\delta > 0^\circ$	$\delta < +10^\circ$
Integration time:	Low-lat: 1500 s Mid-lat: 180 s High-lat: 90 s	Low-lat: 4300 s Mid-lat: 540 s High-lat: 270 s
Receiver:	7-beam 1.4-GHz receiver	13-beam 1.35-GHz receiver
Backend:	Pulsar Fast Fourier Transform Spectrometer (PFFTS)	Berkeley-Parkes-Swinburne Recorder (BPSR)
Bandwidth:	300MHz	340MHz
No. of channels:	512	1024
Freq resolution:	0.58MHz	0.39Mhz
Time resolution:	54 μ s	64 μ s
No. sky pointings:	~ 180,000	~ 43,000
Data sizes:	~ 5 petabytes	~ 1 petabyte



Pulsar Fast Fourier Transform Spectrometer (PFFTS)



New 7-beam 21-cm primary focus receiver at Effelsberg

High time & freq resolution previously unachievable

→ transient sky on timescale down to 10s of μ s

→ Higher freq resolution for negation of IS dispersion

HTRU Observing Strategy

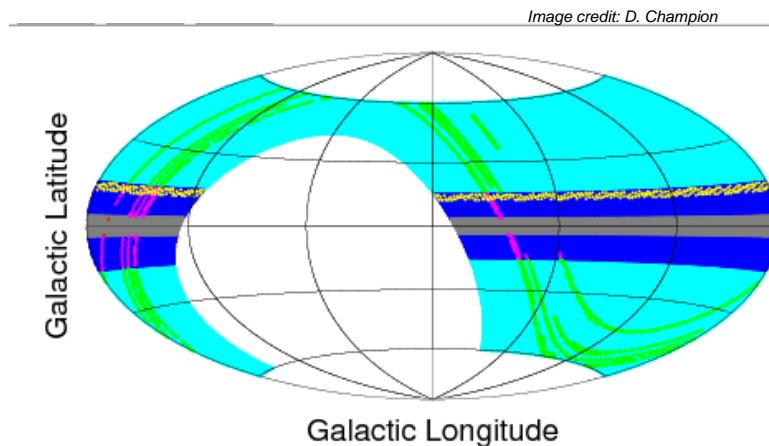
To optimise pointing times for different regions of the Galaxy, the survey will be split into 3 distinct observing regions:

- High-lat ($|b| > 15^\circ$):
 - sub-millisecond PSRs and transients
- Mid-lat ($3.5^\circ < |b| < 15^\circ$):
 - Survey for MSPs as timing array sources
 - large number of MSPs separated by variety of angular distances
- Galactic plane ($|b| < 3.5^\circ$):
 - Deep survey - longest observation per pointing
 - ultimate goal: discovery of PSR-BH binaries

Current Survey Status

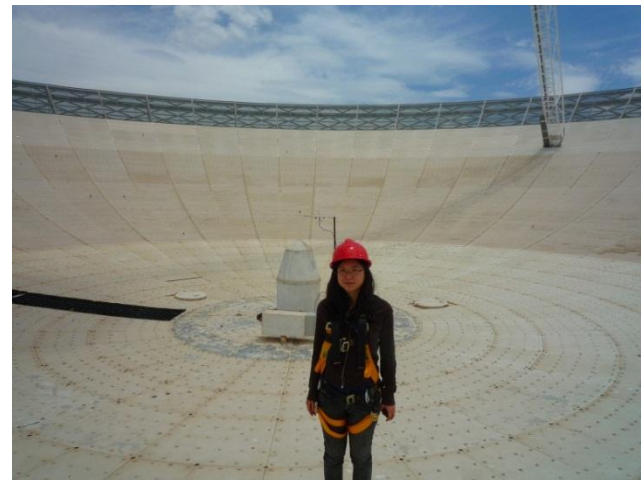
HTRU-North:

- Effelsberg pilot observations in 2009
 - System tests by observing unidentified Fermi point sources
 - Jan 2010 Effelsberg's first ever MSP
- Survey observations underway & data processing just begun in last few months



HTRU-South:

- > 95% mid-lat & >30% high-lat been observed
- Currently 88 pulsars discoveries
 - 15 new MSPs, 5 at high DMs
- First low-lat observation >12% observed



Data processing: Computational power challenges

```
#!/usr/bin/python
import os
import sys, os.path, fileinput
import re
import pprint
import random

SOURCE = 'g'
PREFIXES = 'n'
D1 = '0'
D2 = '5'
UNIT = 'Y'
DirPath = '/usr/local/src/ncopy/vinproc-3.0.3/bin/'
AlwaysKill = '/usr/local/src/ncopy/vinproc-3.0.3/bin/'

os.system('rm -f *')
for i in range(1000):
    seq = seqv[i]
    if [seq == "D1"]:
        D1 = seqv[i+1]
    if [seq == "D2"]:
        D2 = seqv[i+1]
    if [D1 == D2]:
        PREFIXES = 'Y'
    if [D1 != D2]:
        SOURCE = 'Y'
    if [seq == "H"]:
        MARK = seqv[i+1]

def getfilesinDirList():
    os.system('ls -lR * |> /dev/null')

def md5(linefil):
    print 'Deduping all '+linefil+' -d '+cmd+' -c '+cmd+' -g 1000000 -k '+AlwaysKill
    os.system(DirPath+'dedupers_all'+linefil+' -d '+cmd+' -c '+cmd+' -g 1000000 -k '+AlwaysKill)
    print 'Deduping completed'

def hrt(linefil):
    print 'mdk '+linefil+'.000000.tba and protecting against'
    os.system(DirPath+'mdk '+linefil+'.000000.tba -a -dr')
    os.system('rm -f '+linefil+'.000000.drs')
    os.system('cpoc '+linefil+'.000000_01.gpc -d -P')
    os.system('sv dump '+linefil+'.000000_01.drs')
    print 'Specimen dumped to file '+linefil+'.000000_01.drs'
    os.system('h '+linefil+'.000000_d1.sh')

def vmstat(linefil):
    print 'Creating back-up copy of file '+linefil+' to '+linefil+'.orig'
    os.system('cp '+linefil+' '+linefil+'.orig')
    print 'cd file '+linefil+' using FPU rank '+MARK
    os.system(DirPath+'HMD '+linefil+' -k '+MARK)

def detail(linefil):
    print 'Deduping all for '+linefil+' is DM range: '+cmd+' '+cmd
    os.system(DirPath+'dedupers_all'+linefil+' -d '+cmd+' -c '+cmd+' -g 1000000 -k '+AlwaysKill)

def hmr():
    f = open('HMD.txt', 'r')
    lines = f.readlines()
    f.close()
    for line in lines:
        lines = lines[:-1]
        data = line.strip('\n')
        print 'mdk '+line+' and protecting against'
        os.system(DirPath+'mdk '+line+' -a -dr -a')
```

```
os.system('rm -f '+linefil+'.sh')
os.system('rm -f '+linefil+'.log')

def hmr(linefil):
    os.system('cat '+linefil+'.pd > '+linefil+'.drt+cmd+'+'+cmd+'.pr')
    os.system(DirPath+'hmr '+linefil+'.drt+cmd+'+'+cmd+'.pr')

def dpr(D1, D2, linefil):
    print 'Running dpr on candidate with period: '+D1, 'mark: '+D2, 'on file: '+linefil
    ll
    #ls -lR * |> /dev/null
    os.system('dpr -c '+D1+' -d '+D2+' -l 1000 -b -D -R -A -O '+linefil+' -p '+D1+' -a '+D2+' '+linefil)

def dpr(linefil):
    print 'Running dpr on archive file'
    os.system('dpr -c D2 -m 16 -g '+linefil+' -p '+D1+' -a '+D2+' -d '+D2+' '+linefil)

os.system('ls -lR * |> /dev/null')
f = open('HMD.txt', 'r')
lines = f.readlines()
f.close()

if PREFIXES == 'Y':
    print 'Finding multi-beam HRT in Fourier Space'
    os.system('cat '+linefil+'.log')
    for line in lines:
        lines = lines[:-1]
        MD5(linefil)
        hrt(linefil)

if SOURCE == 'Y':
    print 'Running dpr, deduping, finding, and deduping'
    os.system('cat '+linefil+'.log')
    for line in lines:
        lines = lines[:-1]
        hrt(linefil)
        MD5(linefil)
        getfilesinDirList()
        hmr()
    for line in lines:
        lines = lines[:-1]
        MD5(linefil)

if MARK == 'Y':
    print 'Using period of candidate'
    os.system('cat '+linefil+'.log')
    for line in lines:
        lines = lines[:-1]
        data = line.strip('\n')
        data = line.readlines()
        f.close()
        lines = lines[1:]
        for i in range(len(data)):
            data2 = data[i].split()
            cmd2 = data2[0].split()
            cmd3 = data2[2].split()
            cmd4 = data2[1].split()
            if cmd4 > 0: # All other settings OK threshold, user back later.
                D1 = cmd2
                D2 = cmd3
                hmr(linefil)
                dpr(linefil)

print 'dpr completed'
```

Data processing: Computational power challenges

Parkes low-lat (70min deep pointings) observation:

- Size of 1 raw filterbank data file: 17Gb
 - Number of pointings ($N_{\text{pointings}}$) = 1230
 - Number of beams = 13
- Total raw data: $S_{\text{fil}} = 17\text{Gb} \times N_{\text{beam}} \times N_{\text{pointings}} \sim 265\text{Tb} !$

```
Friday February 25, 2011
DataProc-ParkesNew5.py
1/1
```

Data processing: Computational power challenges

Parkes low-lat (70min deep pointings) observation:

- Size of 1 raw filterbank data file: 17Gb
 - Number of pointings ($N_{\text{pointings}}$) = 1230
 - Number of beams = 13
- Total raw data: $S_{\text{fil}} = 17\text{Gb} \times N_{\text{beam}} \times N_{\text{pointings}} \sim 265\text{Tb} !$
- Size of 1 time series $\sim 65\text{Mb}$
 - ~ 1350 dispersion measure (DM) trials each → $65\text{Mb} \times 1350 \sim 85\text{Gb}$
- Total time series: $S_{\text{tim}} = 85\text{Gb} \times N_{\text{beam}} \times N_{\text{pointings}} > 1\text{Pb} !$

Data processing: Computational power challenges

Parkes low-lat (70min deep pointings) observation:

- Size of 1 raw filterbank data file: 17Gb

- Number of pointings ($N_{\text{pointings}}$) = 1230

- Number of beams = 13

→ Total raw data: $S_{\text{fil}} = 17\text{Gb} \times N_{\text{beam}} \times N_{\text{pointings}} \sim 265\text{Tb} !$

- Size of 1 time series $\sim 65\text{Mb}$

- ~ 1350 dispersion measure (DM) trials each $\rightarrow 65\text{Mb} \times 1350 \sim 85\text{Gb}$

→ Total time series: $S_{\text{tim}} = 85\text{Gb} \times N_{\text{beam}} \times N_{\text{pointings}} > 1\text{Pb} !$

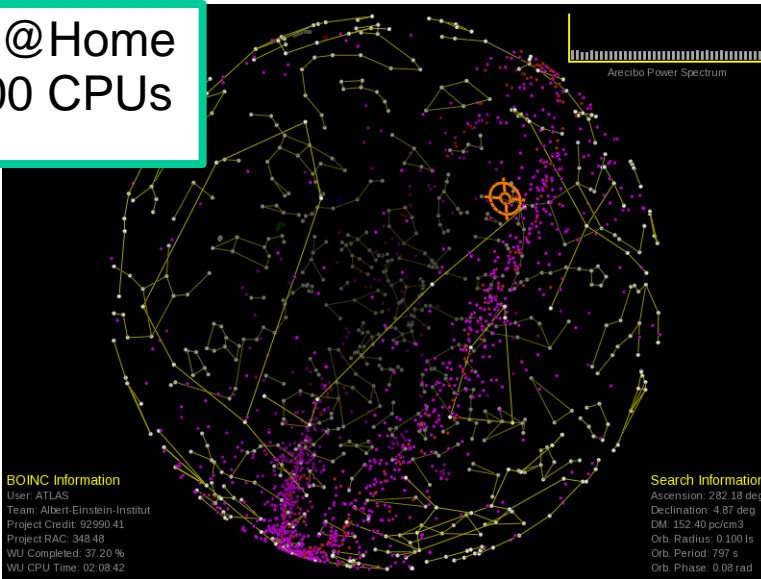
- Time needed for 1 beam standard search on 1 computer ~ 3 days

→ Total time for analyzing all data:

$$T_{\text{total}} = 3 \text{ days} \times N_{\text{beam}} \times N_{\text{pointings}} \sim 131 \text{ years} !$$

Data processing: Computational power challenges

Einstein@Home
~ 200000 CPUs



Jordell Bank –
HYDRA Pulsar Searching
Cluster
~ 1200 CPUs

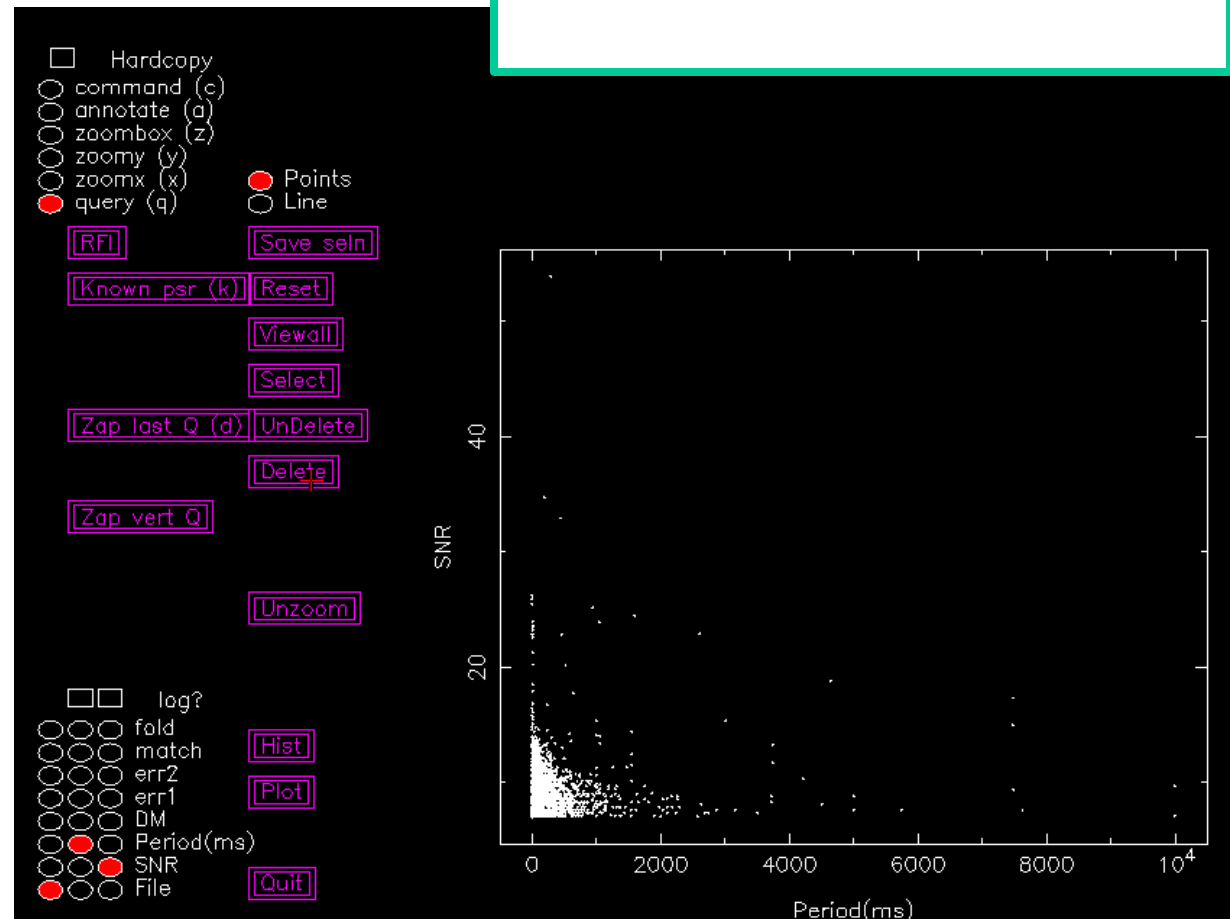
Image credit: The University of Manchester

Hannover ATLAS computer cluster
~ 8000 CPU



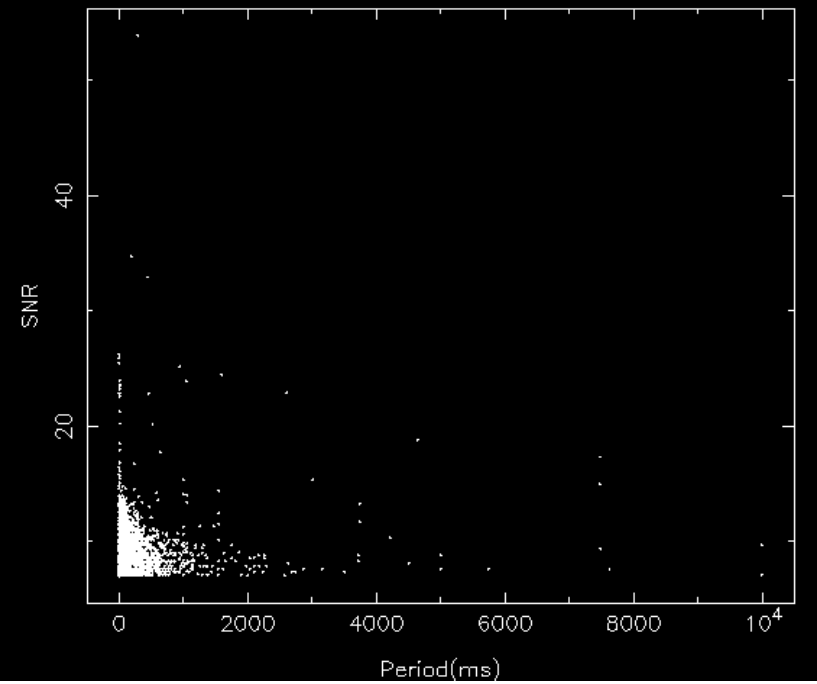
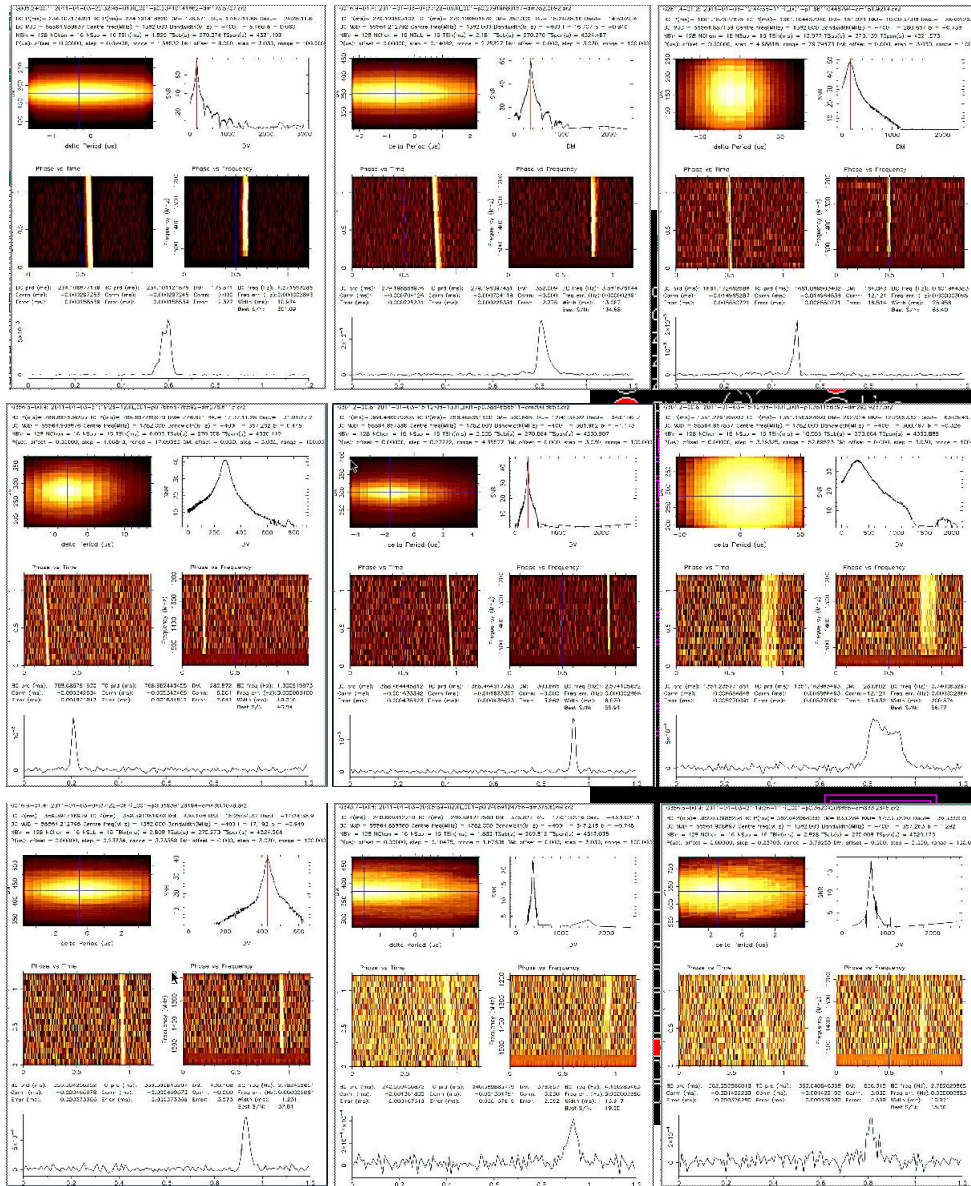
Data processing: Candidates viewing

- 122 processed beams
- 16085 candidates
- Plot here: SNR vs Period



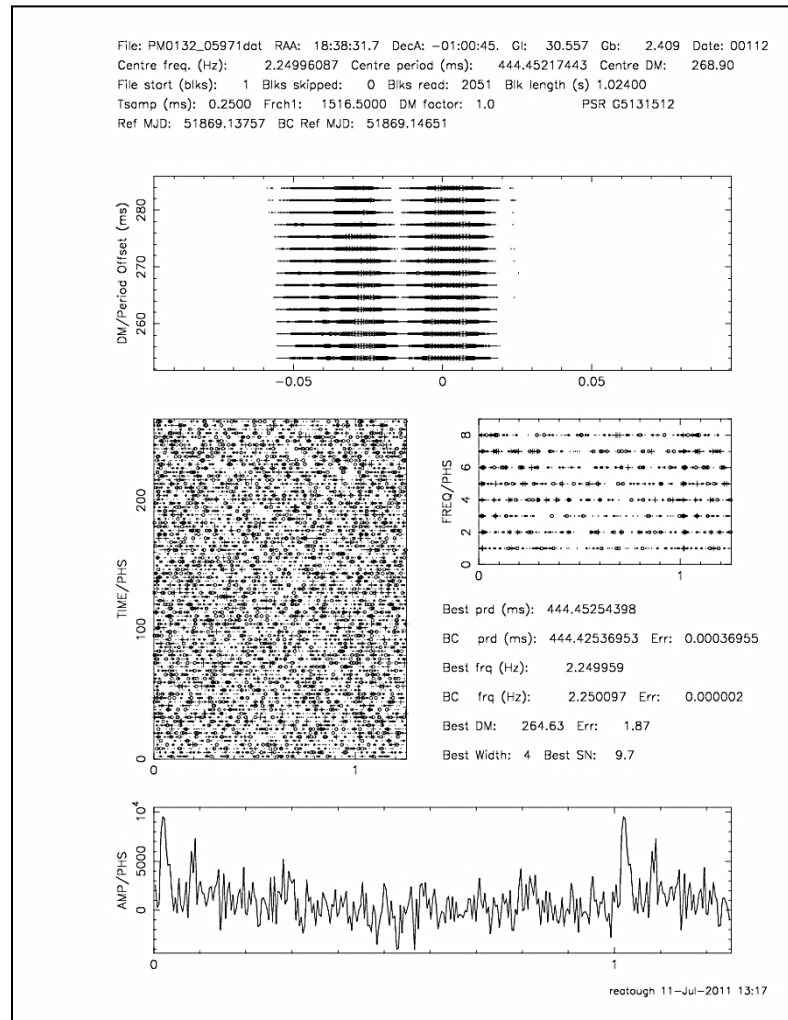
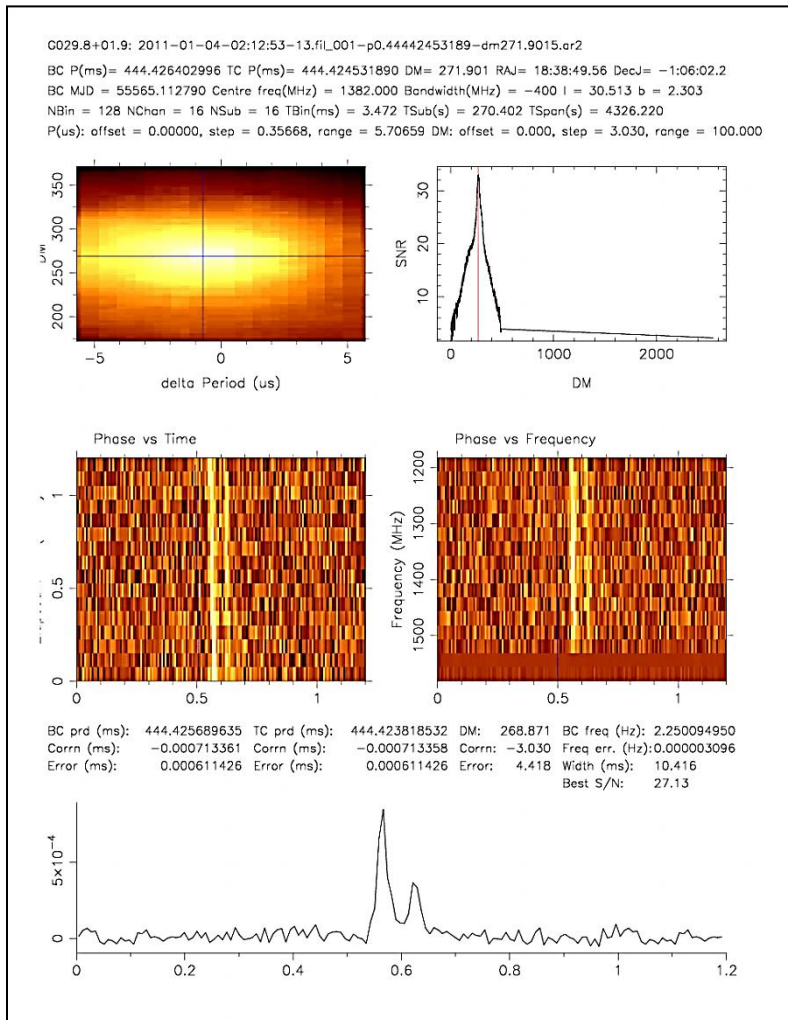
Data processing: Candidates viewing

- 122 processed beams
- 16085 candidates
- Plot here: SNR vs Period
- Re-detection of known pulsars



Data processing: 2 New Pulsars!

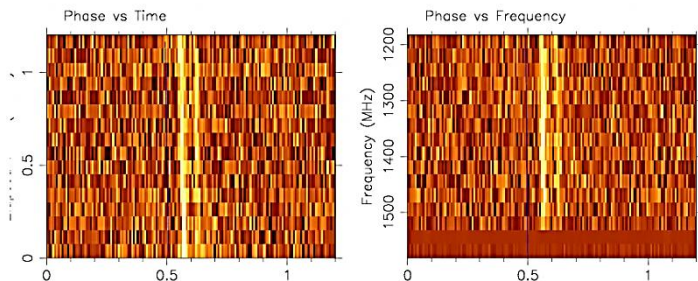
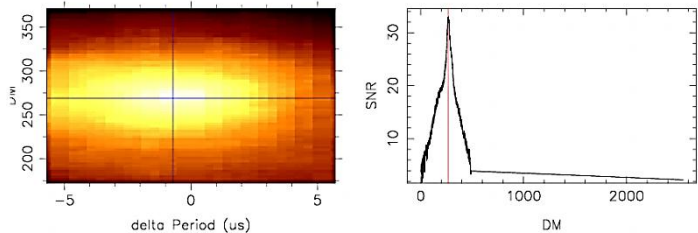
New Pulsar 1: J1838-0106



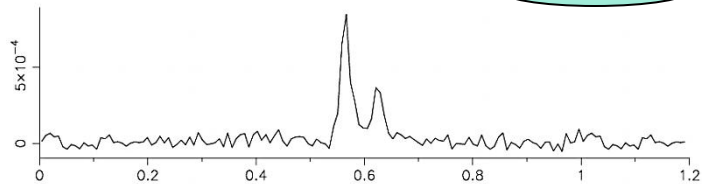
Data processing: 2 New Pulsars!

New Pulsar 1: J1838-0106

G029.8+01.9: 2011-01-04-02:12:53-13.fil_001-p0.44442453189-dm271.9015.ar2
 BC P(ms)= 444.426402996 TC P(ms)= 444.424531890 DM= 271.901 RAJ= 18:38:49.56 DecJ= -1:06:02.2
 BC MJD = 55565.112790 Centre freq(MHz) = 1382.000 Bandwidth(MHz) = -400 l = 30.513 b = 2.303
 NBin = 128 NChan = 16 NSub = 16 TBin(ms) = 3.472 TSub(s) = 270.402 TSpan(s) = 4326.220
 P(us): offset = 0.00000, step = 0.35668, range = 5.70659 DM: offset = 0.000, step = 3.030, range = 100.000

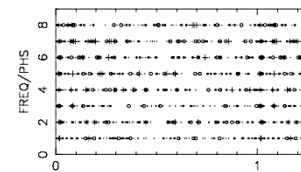
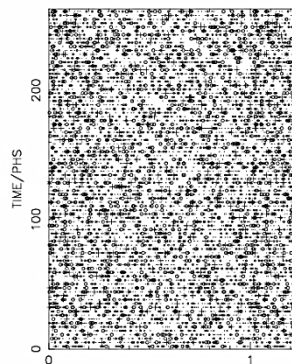
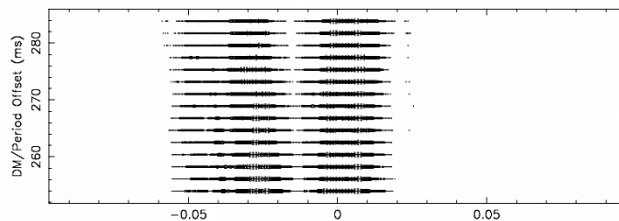


BC prd (ms): 444.425689635 TC prd (ms): 444.423818532 DM: 268.871 BC freq (Hz): 2.250094950
 Corrn (ms): -0.000713361 Corrn (ms): -0.000713358 Corrn: -3.030 Freq err. (Hz): 0.000003096
 Error (ms): 0.000611426 Error (ms): 0.000611426 Error: 4.418 Width (ms): 10.416
Best S/N: 27.13

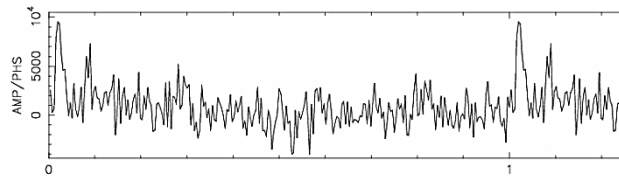


- Tobs(HTRU) ~ 70 min ~ 2 x Tobs(PM)
- Expected SN(HTRU) -> at least $\sqrt{2}$ x SN(PM)

Ref MJD: 51869.13757 BC Ref MJD: 51869.14651



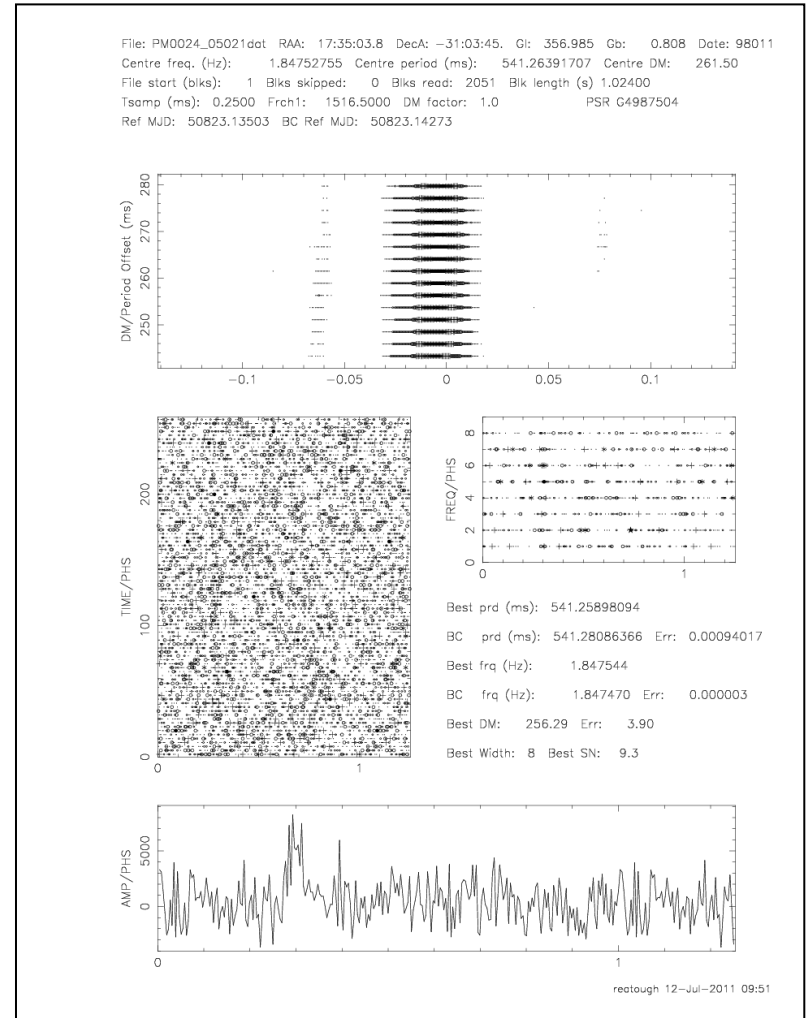
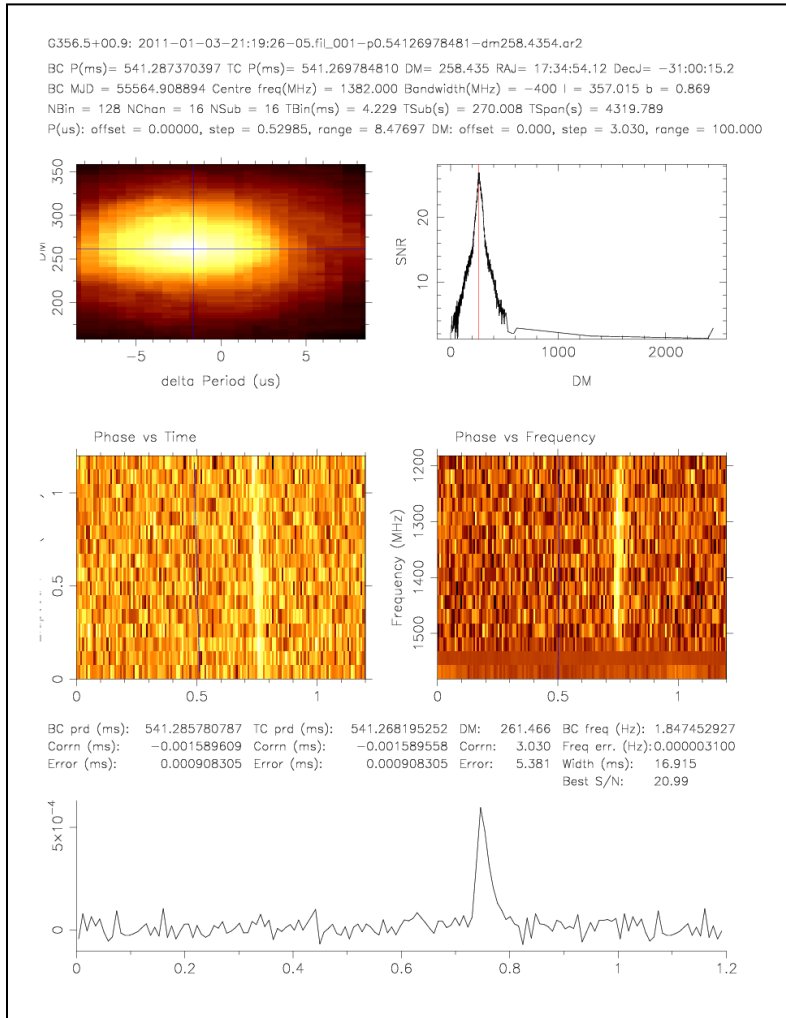
Best prd (ms): 444.45254398
 BC prd (ms): 444.42536953 Err: 0.00036955
 Best freq (Hz): 2.249959
 BC freq (Hz): 2.250097 Err: 0.000002
 Best DM: 264.63 Err: 1.87
Best Width: 4 Best SN: 9.7



reotough 11-Jul-2011 13:17

Data processing: 2 New Pulsars!

New Pulsar 2: J1734-3100

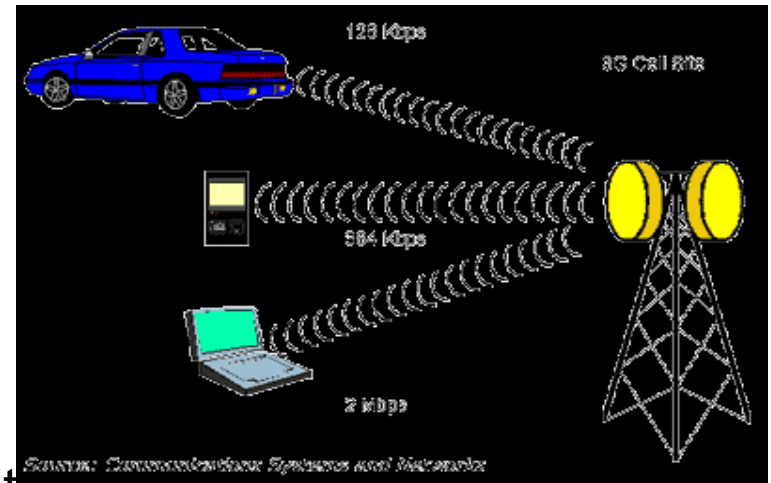


RFI mitigation

- RFI is terrestrial in origin ($DM=0 \text{ cm}^{-3} \text{ pc}$)
 - not travelled through ISM
 - signals not dispersed across bandwidth
 - power peak at $DM=0 \text{ cm}^{-3} \text{ pc}$
- Increasingly deteriorating RFI environment
 - RFI mitigation crucial
- RFI shows up in many beams (say >3)
 - Celestial sources only appear in 1 beam (or if strong, up to say 2 beams)

=> RFI mitigation algorithm:

```
if ((DM==0) && (no_beam >3))
```



HTRU Challenges: RFI mitigation (Time and freq channel)

- Step 1: time-domain RFI

- remove impulsive RFI
- apply zap mask generated during observation
- replace with noise

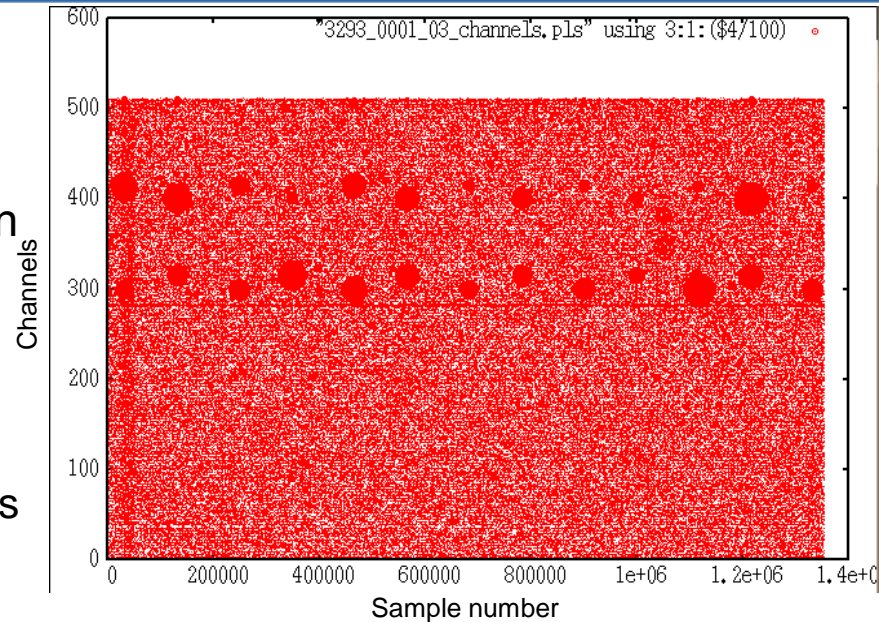
- Step 2: frequency-domain RFI

- Remove “static” (always bad) filterbank channels
- excise channels with excess power

- search each channel for excess power
- remove contaminated channels

- Step 3: fourier-domain RFI

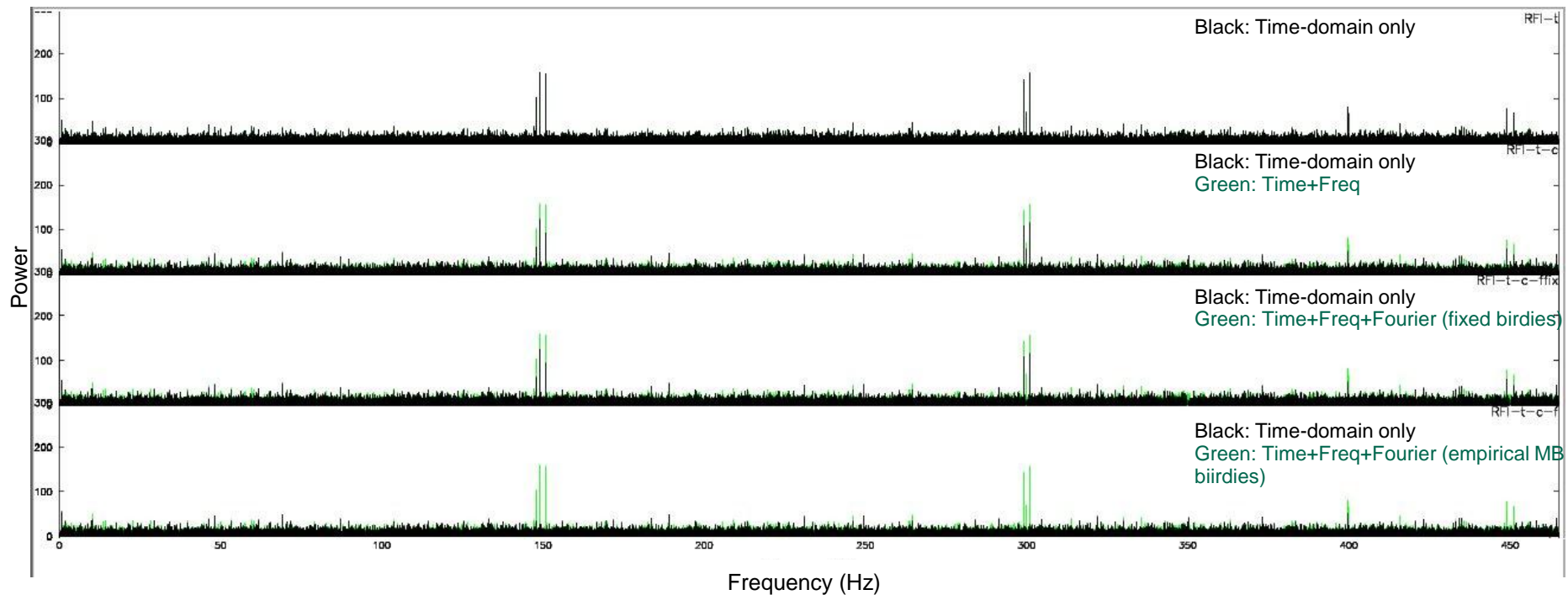
- look into Fourier space
- remove periodic, RFI with multi-beam occurrence (“birdies”)



HTRU Challenges: RFI mitigation (Fourier space)

Parkes deep pointing (Tobs=4300 sec)

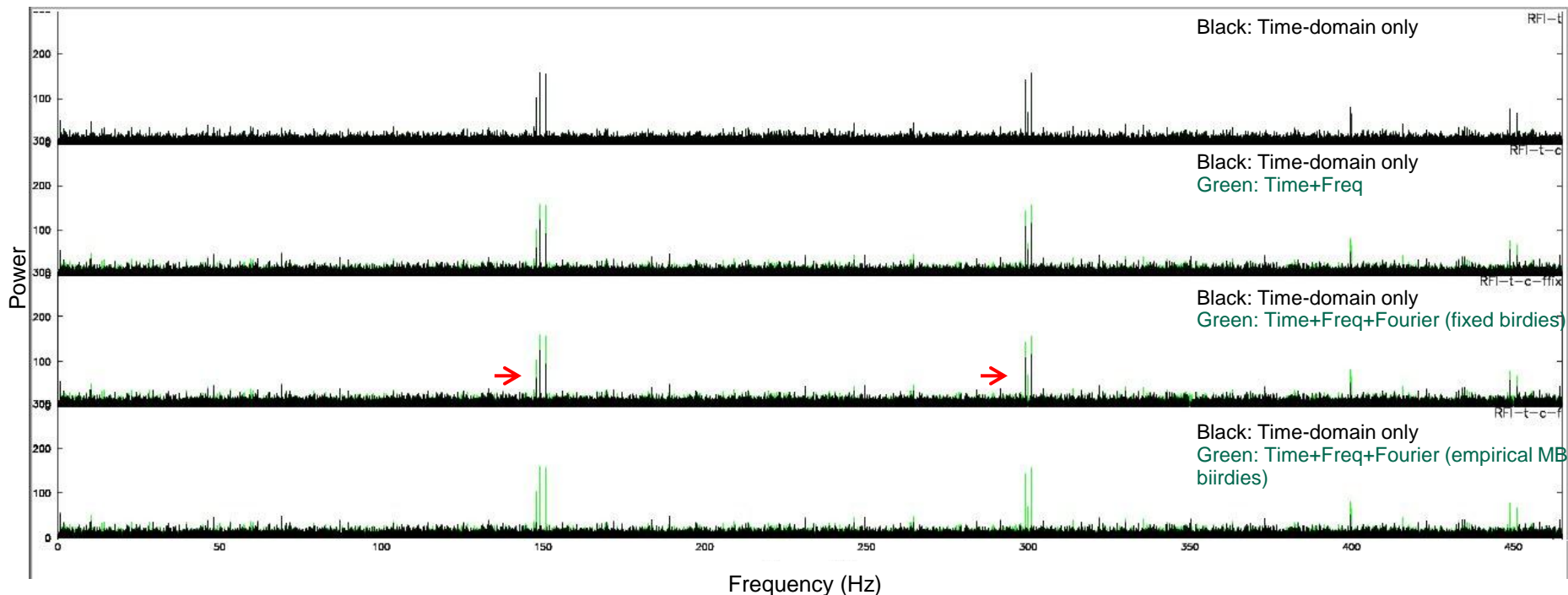
- Central beam (beam 01)
- Dispersion measure (DM) = 0
- Harmonic fold = 1
- Zoomed in fourier freq range 0 – 465 Hz



HTRU Challenges: RFI mitigation (Fourier space)

Parkes deep pointing (Tobs=4300 sec)

- Central beam (beam 01)
- Dispersion measure (DM) = 0
- Harmonic fold = 1
- Zoomed in fourier freq range 0 – 465 Hz



- Changing RFI environment.
- Fixed birdies list inadequate.

Conclusion

- ❖ HTRU-South galactic plane (low-lat) deep-pointings processing commenced with Hydra computer cluster since June 2011
- ❖ ~ 150 beams processed (~ 1% of all low-lat beams)
- ❖ 2 New Pulsars discovered
- ❖ Current work: better RFI mitigation
- ❖ Next stage: develop new algorithms for binary search

Thank you