

Pulsar post processing tools 3pt.exe and 3pt-plot.py

This software tool can be used to extract pulsar signals from 8bit unsigned IQ data captured with a RTL dongle or equivalent.

First the data has to be captured with the tool rtl_sdr.exe as described in

<http://parac.eu/projectmk17.htm>

example; rtl_sdr -s 2e6 -f 419.0e6 -n 2e8 dump.bin, meaning;

-s 2e6; sample rate 2MS/s

-f 419.0e6; frequency 419MHz

-n 2e8; number of samples to process

dump.bin; name of the file

Next the pulsar has to be extracted from that dump.bin and that can be tricky.

From the many tools available we have chosen to write our own.

Short description of the method

Short description: 3pt.exe chops the I&Q file into pulsar periods, performs FFT and writes them as spectrum files to disk.

The python script reads the spectrum files and folds them plots many different plots and finally the multi plot graph.

User instructions

-Put the 3pt.exe, the IQ file the 3pt.cnf and the python plot script in one directory.

-Open the config file, edit it, and save it.

-Run 3pt.exe.

-When finished a number of period files will be written:

--spectrum files contain 10 spectrum rows and 1 sum row each.

--2 band pass and a correction curve files; FYI.

-The files can be examined with a spread sheet program or with the 3pt python plot script.

Long description of the method:

The 3pt tool takes the parameters from a configuration file 3pt.cnf.

Here you can set all the needed parameters and next run 3pt.exe.

The first data strings taken from the I&Q file are used to calculate the pass band of the radio.

This pass band is used for to correct to make the pass band flat in frequency response.

This pass band file is also written to disk (FYI). Also you will notice the 3pt.log file.

Next the scanning of the I&Q file starts again and a chunk of samples is taken to fill a bin and the spectrum amplitudes are corrected with the bandpass correction.

Next the spectrum is split up into 10 channels and the dispersion correction is applied.

The 11th channel is used to add all the 10 same bins together in a sum file.

Next all the channel values are individually appended to the next bin. When all the bins in one period are determined, then one period file with 11 channels is written to disk.

This is done for all the periods.

You can read these files in Excel or Libre office, but you also can use the python script 3pt-plot.py.

By writing the period files to disk (like with CFRAD2), the advantage is that you can quickly select another range of periods to fold.

In that way you can scan the entire IQ file to check if the pulsar is strong in one section and weak in another section of that captured file.

The Python script was written in Python 2.7. Convert with 2to3 or online to Python 3 if needed.

The config file

```
0.-----first part calculation-----
1.pulsar name = B0329+45 SGRT abcd
2.pulsar period time [s]= 0.7145032439188
3.Sample rate [Sps]= 2000000.0
4.search optimal sr(+) or pt(-) [ppm] (default step:+-0.05, 0:off)=0
5.DM=26.78
6.center freq[ MHz]= 404.0
7.IQ input file name =dump-sgrt01.bin
8.number of output bins= 200
9.spectrum file name without extension=abc-
10.process period from=0
11.process period to (0:all)= 0
12.clipping & RFI filtering (0:no,1:yes)=0
13.clipping & RFI filtering level =0
14-----second part folding, plotting-----
15.plot period from = 0
16.plot period to (0:all)=10
17.fold period from =2
18.fold period to (0:all)=600
19.exclude channels (0:None else 1..10 csv)=2,3
20.date, time= 1234,1234
21.observatory, location=SGRT-PT
```

Config file description

Both programs use the info given after the "=" token.

The 3pt.exe only uses the config info from the first part. The 3pt.py script uses all the config info.

1. Any name can be given. Also extra info can be added; this will be displayed in the multiplot.

2. The period time can be calculated very precise using for instance TEMPO, or online:
http://neutronstar.joataman.net/technical/psr_topo_freq_calc_estimate_200522a.html.

3. The sample rate given must be accurate and stable. You can obtain that by using the test instruction

```
rtl_sdr -s 2e6 -f 419.0e6 -n 200 test.bin
```

in a command window. As a result the real sample rate is displayed.

To maintain stability a tcxo or a GPS disciplined oscillator is advised.

4. When the period time or the sample rate are not exactly known, you can search them. Change the default value '0' to '-0.05' to search for an optimal period time or to '0.05' to find an optimal sample rate. Pay attention to the sign.

For instance when the set sample rate is 2000000Sps, then with 0.005 the application will vary the sample rate from

$2000000 + 0.005 * 2000000 = 1999980$ to 2000010 Sps in 10 steps and record the results. The result files are written to disk and can be read with notepad or the python script.

The calculation can take a long time..up to many hours or days, so set the app search not over 200 process periods max.(line 11)

Choose a low bin value; like 100, because else the routine that tries to find the maximum peak, can mistakenly select a RFI peak.

5. The de-dispersion measure (DM) is given here.

6. Give the center observation frequency; this is used together with the DM to calculate the channel delay.

7. Give the name of the I/Q file.
8. Give the number of output BINs. A value between 8 and 2000 can be used; but normally it is in the range 100-400. see also the notes.
9. Give a name for spectrum output file without extension. You can change the name every time you re-calculate the IQ file with other parameters.
10. Give the period number the processing has to start with. Normally this is 0. see also the notes.
11. Give the number of the period to end with. So for a session of 0 to 1000, fill in 1000 here. see also the notes.
12. Clipping can be activated in line 12 and trimmed in setting 13.
 - 0: no clipping.
 - 1: use Hanning window when performing FFT.
 - 2: use clipping in the form of $y=x^B$.
 - 3: use Hanning and clipping together.
13. clipping level value is "B" in $output=input^B$.

When "B"=0.5, then the output y is a \sqrt{x} , so it is a gradual limitation of the amplitude. "B" may be set from >0 to 1.
14. --second part
All settings in this part have no influence on part 1 and are used by 3pt.py only.
15. plot from period.
16. plot to period.
17. fold from period
18. fold to period
19. exclude frequency channels (0:None, else 1..10 csv)=2,3,8
You can try to exclude a channel when this channel does not contribute to a higher signal to noise ratio.
In the multiplot, plot nr 5 gives the signal strength of the strongest bin. When that level is low you can decide to exclude that channel.
Re-run the script with the edited config file and see if the s/n ratio is improving.
20. not used; just displayed in the multiplot
21. not used; just displayed in the multiplot

Notes

-1. The period time of the pulsar can be calculated with TEMPO like described here

http://www.k5so.com/documents--and-downloads/download-tempo_calc.html

Or you can use the on line calculator here

http://neutronstar.joataman.net/technical/psr_topo_freq_calc_estimate_200522a.html

-2. The Dispersion Measure or DM can be found here

https://www.atnf.csiro.au/people/pulsar/psrcat/proc_form.php?version=1.63&startUserDefined=true&c1_val=&c2_val=&c3_val=&c4_val=&sort_attr=jname&sort_order=asc&condition=&pulsar_names=&ephemeris=short&coords_unit=raj%2Fdecj&radius=&coords_1=&coords_2=&style=Long+with+last+digit+error&no_value=*&size=3&x_axis=&x_scale=linear&y_axis=&y_scale=linear&state=query&table_bottom.x=46&table_bottom.y=21

or here

<http://www.jb.man.ac.uk/research/pulsar/Resources/eprn/browser.html>

-2b. Optimal period/sample rate searching.

There are parameters which influence the folding of the pulsar signals. These are period time, sample rate and DM.

The period time can be calculated very precise using for instance TEMPO or online:

http://neutronstar.joataman.net/technical/psr_topo_freq_calc_estimate_200522a.html.

The DM and the observing radio frequency must be known so a correct de-dispersion can be performed.

The sample rate should be known, but often is not accurately known. Also temperature drift shifts the sample rate. Best is to use a GPS disciplined oscillator.

So, if the standard x-tal in the dongle is used, the sample rate set is not exactly the sample rate executed.

To find the correct sample rate, a search for the best result can be done.

In the config file set the search value (in row 4) from 0 (=no search) to 0.05 (=0.05 pro mil) of the set sample rate to vary.

-3. If in the config file the "to" period is set to 0, then all the periods will be processed.

This is useful when the number of period is not yet known.

-4. Any number of bins can be set but advised is a number between 10 and 1000. For B0329+54 optimal is 108 bins.

For optimal bin size see also http://neutronstar.joataman.net/technical/epoch_folding.html

If after processing and plotting it appears that the pulsar is located close to a border, you can change that by recalculation.

For instance when the pulsar peak is located at 10% of the bin scale and you want it at 50%, then fill in the number of bins as xyz.4.

If the peak shows the first time at 85% of the bins location, then fill in the bins xyz.65 to shift it to the center. Recalculate the period files.

-5. By writing the period files to disk (like with CFRAD2), the advantage is that you can quickly select another range of periods to recalculate.

In that way you can scan the entire set to check if the pulsar is strong in one section and weak in another section of the captured file.

This can occur when you have done a passive scan, where the pulsar sweeps through the beam of the antenna.

The disadvantage is that the number of files can be huge.

-6. The data IQ files can be huge, so to start a post processing session you can begin with processing only the first 10 periods and writing them to disk.

Now you can examine the 10 files and decide how to proceed.

-7. If you have a really big IQ file to process, you can decide to use all your PC processor cores for parallel calculation.

For instance when the first session is from 0 to 1000, the second session could be from 1000 to 2000 etc. So for the second session fill in the start from value '1000', and fill in the line 'process to' with 2000.

The procedure is to edit the config file; process from 0, process to 1000. Start 3pt.exe.

Next modify the config file; 'process from 1000' and 'process to 2000' and save the file. Copy and rename 3pt.exe to 3pt2.exe. Run 3pt2.exe etc.

Your PC will be more loaded and the fan speed will increase. You can repeat this procedure as many times as you and your PC cores like.

-8. It appears that when using the RTL dongle, the data of the first two periods is mis-formed. It is advised to skip those and not add them and use them for pulse formation.

You can do that by setting 'fold period from' to 2 or 3 in the config file.

-9. The 10 spectra channels are useful for checking if the pulsar is indeed present.

Because of the de-dispersion correction, the channels are shifted in time. When looking at a bump in one of the channels, the same bump must be in line with all the channels for a confirmation of the pulsar.

If the bumps are shifted (and the correct DM and frequency have been set) then the bump is not a pulsar.

-10. When starting a session, be sure that the exact sample rate is known and stable. Remember that sample rate and pulsar period are connected. For a chosen sample rate, the product of $Sps * period\ time = constant$.

You also can find that constant in the log file. So when you are confident that the calculated period time is correct but you are not confident of the sample rate, then change the SPS and examine the s/n ratio of the pulsar peak value.

Now you can use the calibrated sample rate also for other pulsar measurements.

Also; it is not useful to perform a very long measuring session when the sample rate is not stable. The pulsar peaks will be smeared and folding them give no enhancement.

-11a. -You could want to plot individual period files; this can be useful when you have a strong signal.

Also this is useful to notice time jumps in your signal caused by PC interrupts.

-11. When doing pulsar analysis, it is useful to also plot the frequency channels, like with this tool, so you can judge the quality of them.

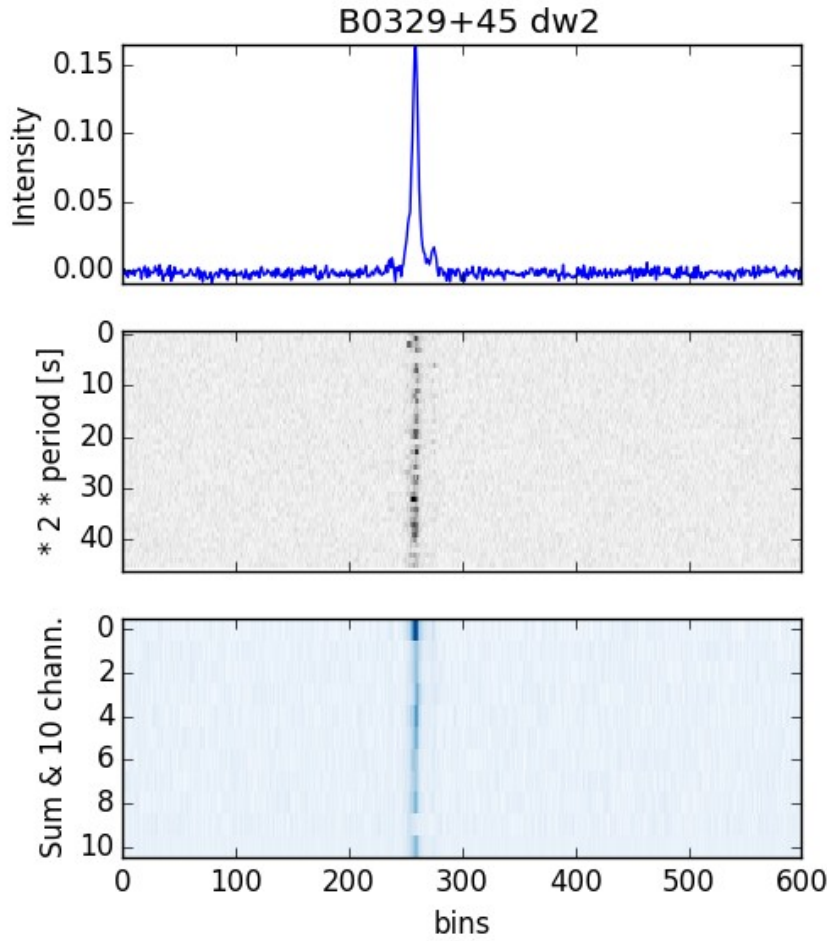
When one or more channels displays severe RFI you can decide not to use that channel and so improve the S/N ratio.

-12. When experimenting with other settings, it is advised to alter the output spectrum file name, because else they will be overwritten.

-13. Pulsar B0329+54 has two different mode states; normal or abnormal. They can be calculated also. You have to fill in the pre- and post bump bin ranges.

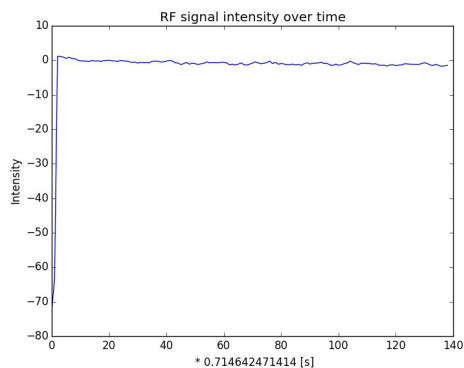
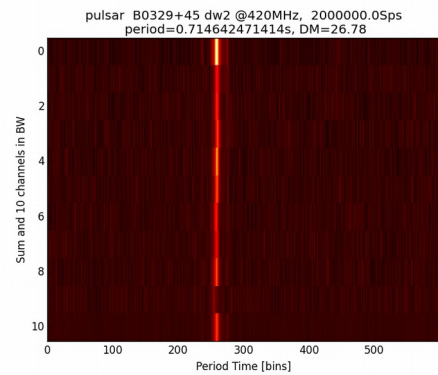
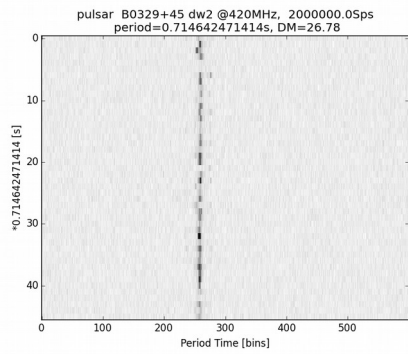
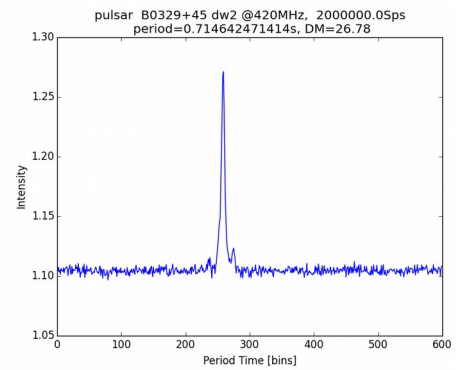
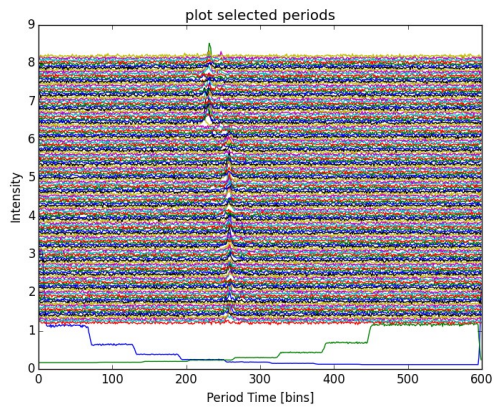
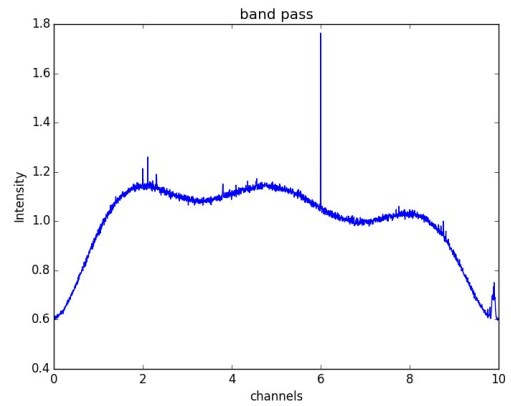
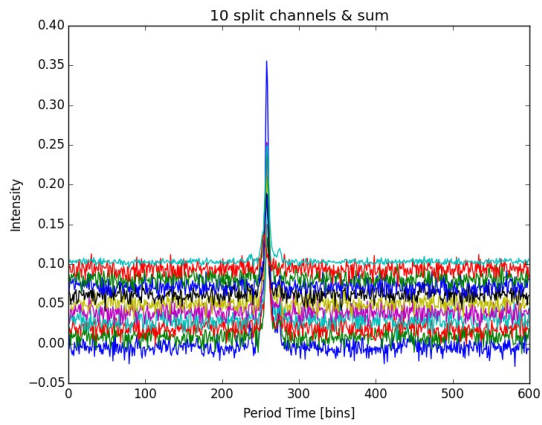
The answer for the example file, or the complete dump2000a.bin file, should be $R=0.4$; this is the normal mode.

Example 1



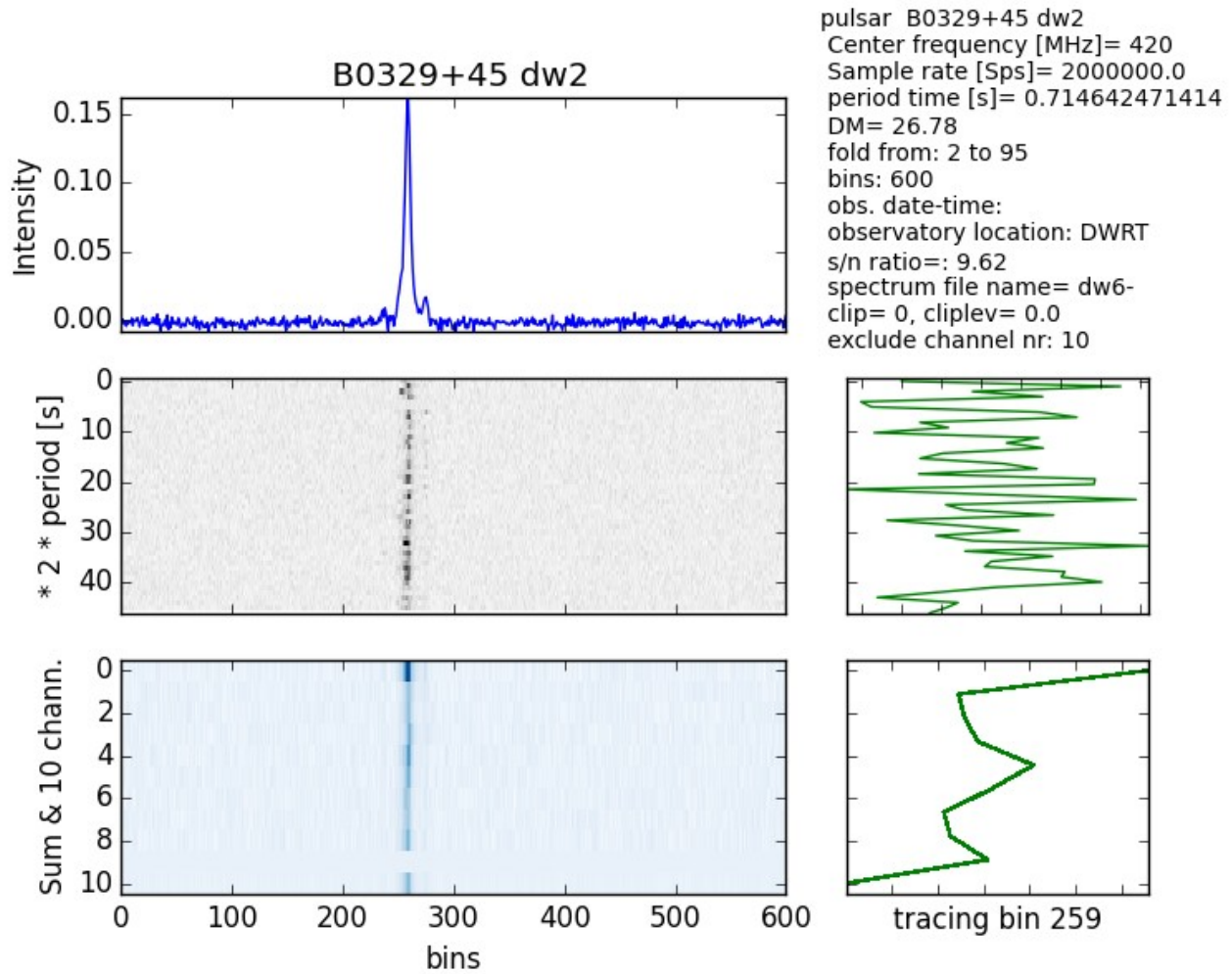
pulsar B0329+45 dw2
Center frequency [MHz]= 420
Sample rate [Sps]= 2000000.0
period time [s]= 0.714642471414
DM= 26.78
fold from: 2 to 95
bins: 600
obs. date-time:
observatory location: DWRT
s/n ratio=: 8.182
spectrum file name= dw6-
clip= 0, cliplev= 0.0
exclude channel nr: none

The individual plots are also given.



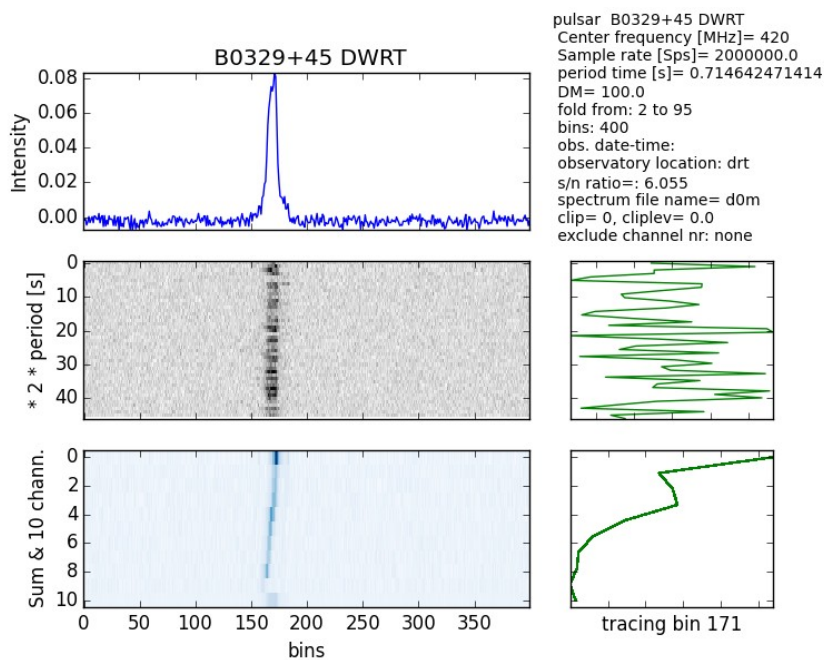
Example 2

Here channel 10 is excluded and the s/n ratio is improving.



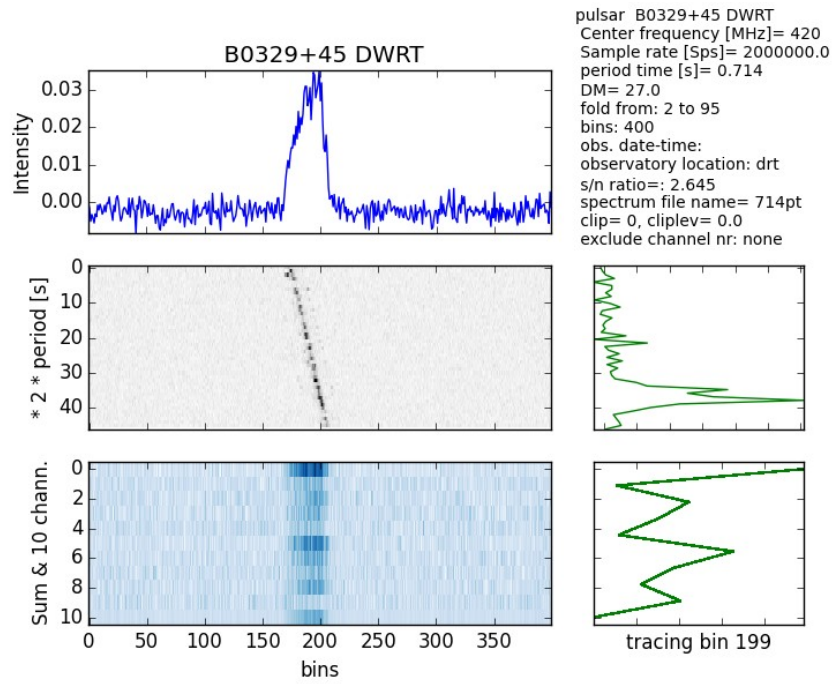
Example 3

When given the wrong DM, the frequency channels are not aligned anymore.



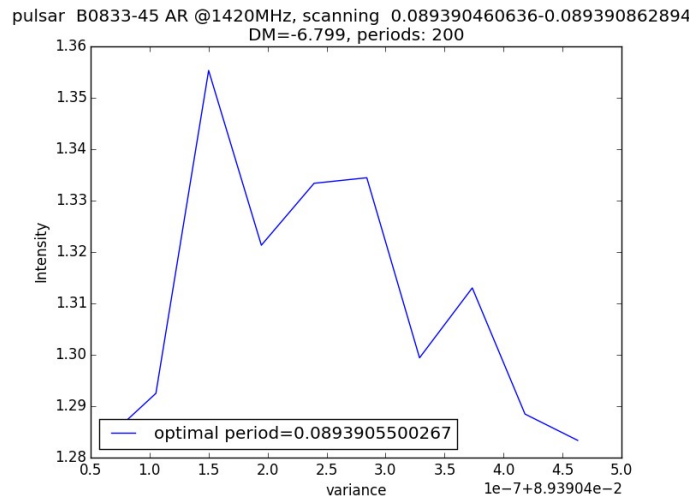
Example 4.

When the pulsar period given is wrong, then the periods plot is not aligned anymore.



Example 5.

Plot of a searched optimal period time



Michiel Klaassen

parac.eu

July 2020