MANCHESTER 1824
The University of Manchester

SP-1879: Create set of PSS test vectors for the SPS pipeline

Lina Levin Preston
for the PSS Team
17 November 2021

Team PSS

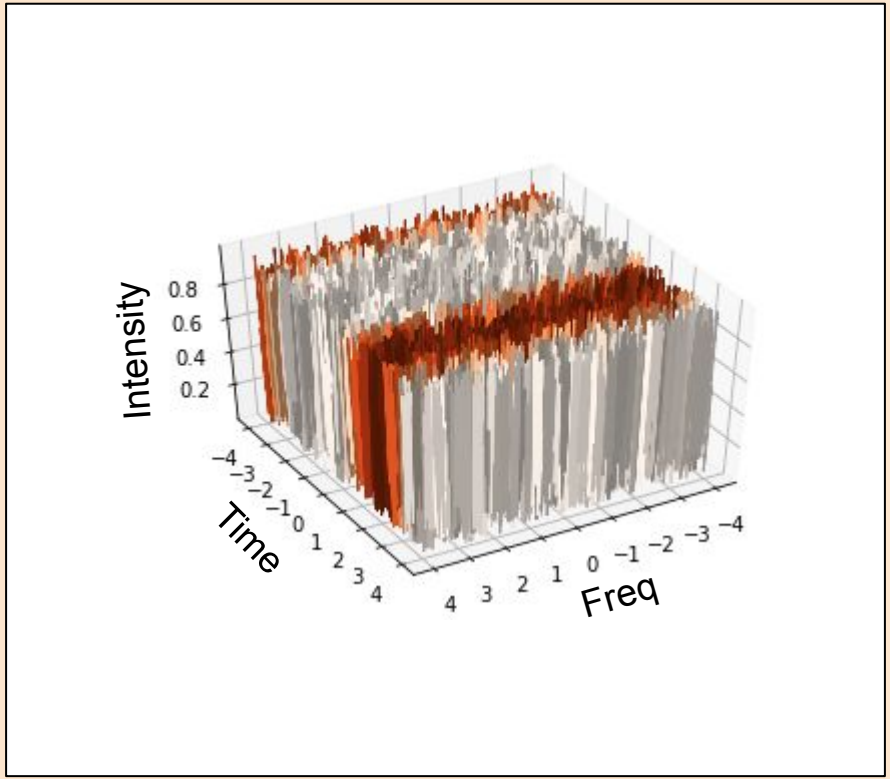
University of Oxford
Aris Karastergiou (PO)
Arun Naidu
Christopher Williams

Raman Research Institute
Prabu Thiagaraj

University of Manchester
Atul Ghalame
Michelle Hussey
Mitch Mickaliger
Vincent Morello
Mike Pearson (SM)
Lina Levin Preston
Benjamin Shaw
Ben Stappers (PO)

Test Vectors

- 4-D Data cube - signal intensity vs time, frequency, polarisation.
- Examples of real astrophysical signals (periodic (pulsars) and transient (FRBs, GP, RRATs)).
- SKA sized test vectors are up to 36GB based on current requirements.
- Crucial for
  - demonstrating pipeline performance
  - demonstrating compliance to system requirements
  - ad-hoc/unit/system/pipeline testing



Define SPS test vector parameters

Pulse parameters for the SPS test vector set:

- 2 pulse periods: 5, 8 s
- 10 DMs: 1, 100, 370, 500, 740, 1000, 1480, 2000, 2950, 3000 pc/cm^3
- 5 pulse widths: 100 μs, 1 ms, 10 ms, 100 ms, 1 s
- 3 S/N values: 20, 30, 50

Relevant requirements:

SKAO-CSP_Mid_PSS_REQ-13: CSP_Mid.PSS Single Pulse search
CSP_Mid.PSS shall, as part of the Pulsar Search, search for individual pulses with dispersion measures from 0 to 3000 pc cm^{-3} and with widths from 100 microseconds to 1 second.

SKAO-CSP_Low_PSS_REQ-14: CSP_Low.PSS Single Pulse search
CSP_Low.PSS shall, as part of the Pulsar Search, search for individual pulses with dispersion measures from 0 to 3000 pc cm^{-3} and with widths from 100 microseconds to 1 second.

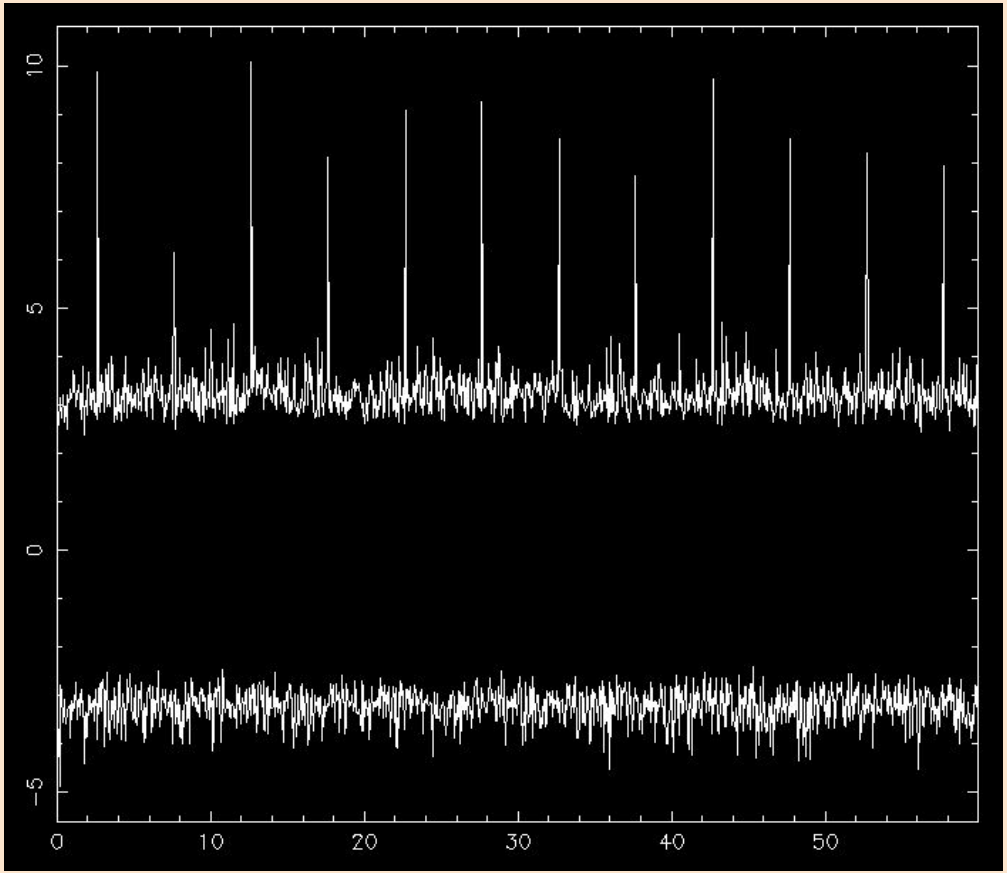
Create SPS test vector set

```
[llevin@dokimi plots]$ ls -1 /skatvnas3/doc_root/testvectors/SPS-MID/
latest
SPS-MID_747e95f_0.125_0.000125_1000.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1000.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1000.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_100.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_100.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_100.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1480.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1480.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_1480.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_2000.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_2000.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_2000.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_2950.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_2950.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_2950.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_3000.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_3000.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_3000.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_370.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_370.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_370.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_500.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_500.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_500.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_740.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_740.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.000125_740.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.00125_1000.0_0.0_Gaussian_20.0_123123123.fil
SPS-MID_747e95f_0.125_0.00125_1000.0_0.0_Gaussian_30.0_123123123.fil
SPS-MID_747e95f_0.125_0.00125_1000.0_0.0_Gaussian_50.0_123123123.fil
SPS-MID_747e95f_0.125_0.00125_100.0_0.0_Gaussian_20.0_123123123.fil
```

The SPS test vector set:

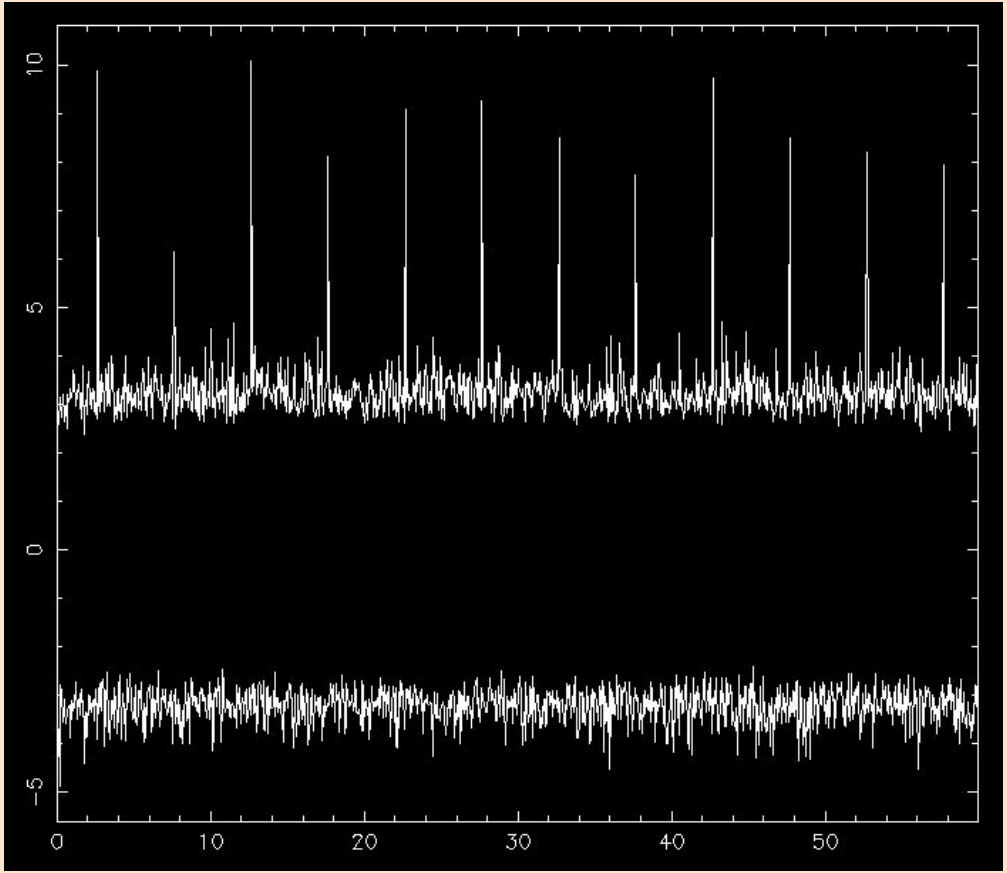
- 300 test vectors:
  - 60 second integration time
  - 4096 frequency channels
  - 64 µs time resolution
  - 8-bit samples.
- 3.84 GB/file = 1.152 TB

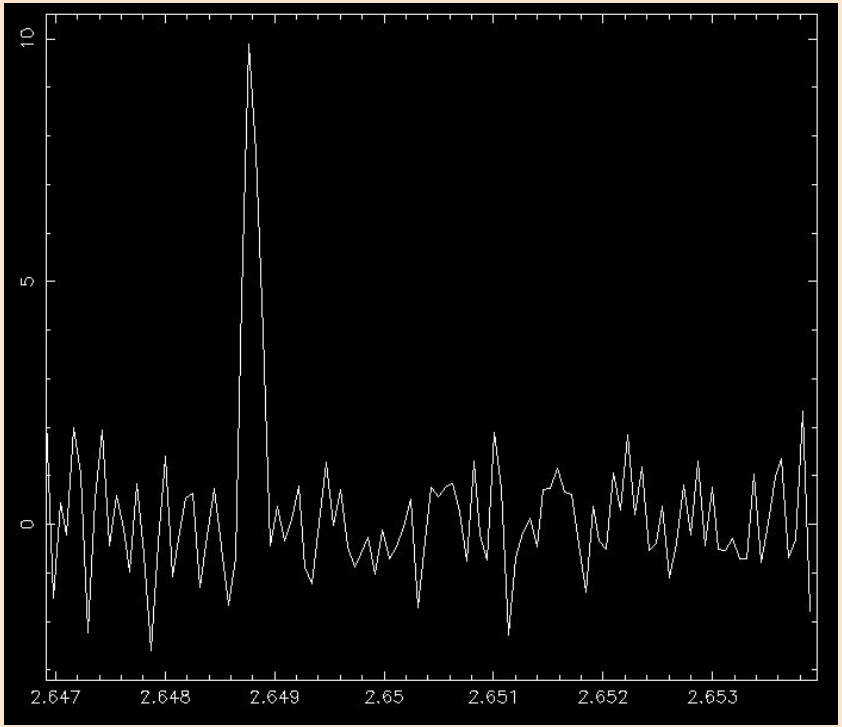
Test Vector Generator:
https://gitlab.com/ska-telescope/pss-test-vector-generator/



Input parameters for vector:

- DM = 100 pc/cm^3
- Period = 5 s
- Pulse width = 1 μs
  => <2 bins
- S/N = 50
  => single pulse S/N ~14


10
5
0
-5
0
10
20
30
40
50


10
5
0
2.647
2.648
2.649
2.65
2.651
2.652
2.653

Validate SPS test vectors

Validation steps:

- Run SPS parts of PSS Gold Standard matlab code
  - RCPT: read in test vector data
  - RFIM: remove interference (mask channels)
  - DDTR: de-disperse data at given DM steps
  - SPDT: search each time series for single pulses
  - SPSIFT: remove duplicate single pulse detections
- Compare output detection lists with input parameters

PSS Gold Standard MatLab model:
https://gitlab.com/SKA-TDT/tdt-matlab-models

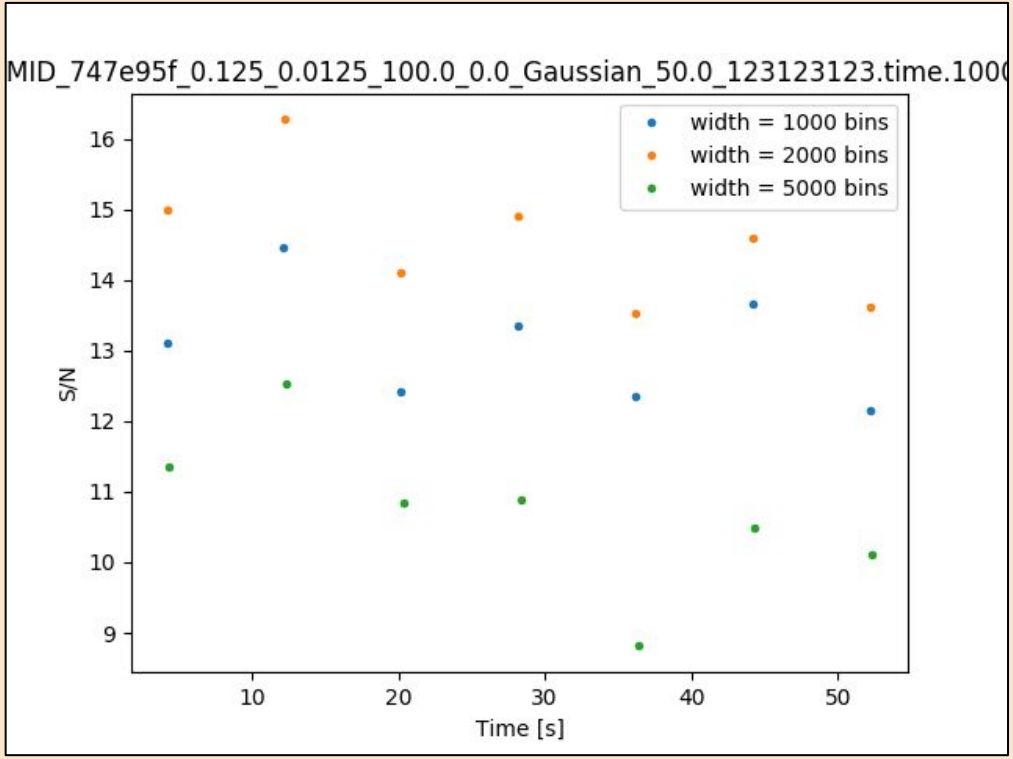
PSS Signal Model document:
https://www.dropbox.com/s/sv86txc0zkwmx7f/SKA-TEL-CSP-0000085_5_PSSSignalModel_Mickaliger_2018-09-04%20-%20signed.pdf?dl=0

Validate SPS test vectors

Input parameters for vector:

- DM = 100 pc/cm^3
- Period = 8 s
- Pulse width = 100 ms
  => 1563 bins
- S/N = 50
  => single pulse S/N ~18

Summary document
https://www.overleaf.com/read/xmrjnskdjfqg



Future work

- We have a plan for how to implement the end-to-end testing framework, with many features already identified.

- Next steps:
  - Implement reverse proxy on University of Manchester machines to enable test vectors to be accessed remotely
  - End-to-end testing will be using pytest, need to develop classes for running Cheetah and parse the output
  - Set up and implement SPS tests in pss-pipeline
  - Further down the line we will also implement tests for FDAS pipeline