Pulsar Data Reduction
2021 Winter Observer Training Workshop
Overview

- PSRFITS format used for all standard pulsar observing modes
  - Slightly different data structure depending on online processing mode ("search" vs "fold" vs "cal" modes)
  - Baseband recording modes available upon special request and coordination; we will not cover these here
- GBO does not maintain an independent pulsar data reduction package
- All GBO pulsar data can be analyzed using community-supported packages
  - PRESTO (especially good for searching for pulsars)
  - PSRCHIVE (especially good for fold-mode data)
  - PSRFITS UTILS (useful for search mode → fold mode)
  - DSPSR (useful for search mode → fold mode; also for processing baseband data)
Getting Set Up

• To access pulsar data analysis tools, run the following command

   source /home/pulsar_rhel7/pulsar.presto3.bash

• You can wrap this in a function in your .bash_profile. Just add the following lines

   psrenv() {
       source /home/pulsar_rhel7/pulsar.presto3.bash
   }

• Then just type ‘psrenv’ when you open a new terminal
  – Must be done every time you open a new Linux terminal
• Wrapping in a function is necessary to prevent problems with VNC (remote desktop viewer)
PSRFITS Overview

• Conforms to FITS standard; compatible with astropy (though other packages provide more functionality)

• “Search” mode, aka filterbank mode
  – High time resolution spectra
  – Can record 1 (summed) or 4 (self/cross product) polarizations
  – Organized into “sub-integrations”

• “Fold” mode
  – Use a pulsar ephemeris to calculate precise rotational phase for each sample
  – Results in data cube of pulse phase vs freq. vs time
  – Must record 4 polarization products

• “Cal” mode
  – Specialized fold-mode for observing pulsed noise diode
PRESTO

- C and Python-based data reduction package
- Lots of functionality for searching for, finding, folding, and timing pulsars
  - Limited flux calibration routines
  - No polarization calibration
- Lots of additional Python utilities for pulsar astronomy
- https://www.cv.nrao.edu/~sransom/presto/
- Help menus for tools are fairly complete
- Tutorial available from Scott Ransom
  - https://www.cv.nrao.edu/~sransom/PRESTO_search_tutorial.pdf
- Another tutorial available from Ryan Lynch
  - Just ask!
Common PRESTO Commands

- `rfifind` – Identify and “mask” RFI based on statistical outliers
- `prepdata` – Dedisperse, sum frequencies, and create time series
- `prepsubband` – Similar to `prepdata`; make lots of time series at once
- `accelsearch` – Find pulsars (including binaries)
- `single_pulse_search.py` – Find pulsars, RRATs, FRBs through single pulses
- `prepfold` – Fold a candidate or known pulsar
- `get_TOAs.py` – Get pulse times of arrival
PSRCHIVE

- C++ tools for working with fold-mode data
- Tools for viewing, editing, RFI-zapping, flux/polarization calibration, extracting TOAs
- Online documentation is only partially complete
- Help menus for each program can answer common questions
- Some user-generated recipes for things like calibration
- Python wrapper available for some basic analysis
Common PSRCHIVE Commands

- pav – Lots of different ways to plot data
- paz and pazi – Automatic and interactive RFI removal
- pam – Lots of ways to manipulate data (e.g. downsample)
- psredit – View/edit header information
- vap – Another way to view header information
- fluxcal – Make flux calibration solutions
- pac – Flux/polarization calibration
- pat – Extract TOAs
- pas – Work with standard timing templates
- psrsMOOTH – De-noise standard timing templates
- psrplot – Advanced plotting
PSRFITS UTILS

• Mostly used to fold or reduce search-mode data
• https://github.com/demorest/psrfits_utils
• Two main useful commands
  – `fold_psrfits` – Folding search-mode data (then it can be analyzed with PSRCHIVE)
  – `psrfits_subband` – Reduce time/frequency/polarization resolution (useful for decreasing data volume)
DSPSR

• Useful for folding search-mode data and working with baseband data
  – Can also make search-mode data in PSRFITS and SIGPROC filterbank format from baseband data
  – Can also connect to data buffer for online observing (e.g. VLA)
• [http://dspsr.sourceforge.net](http://dspsr.sourceforge.net/)
• Help menu is pretty descriptive
• After folding, use PSRCHIVE
GBO Pulsar Data

- All data are written to a mass file store (/lustre/gbtdata)
- Data further organized by project code and session number
  - e.g. /lustre/gbtdata/TGT21A_501_05
- Data further organized by mode (VEGAS for incoherent dedispersion, VEGAS_CODD for coherent dedisersion)
  - e.g. /lustre/gbtdata/TGBT21A_501_05/VEGAS_CODD

- Please do not analyze your data in these directories!!
  - These areas are archived
  - Create symlinks to files in /home/scratch/<username> or /lustre/pulsar/<username>
Merging VEGAS Data

- Depending on observing mode, data from VEGAS may need to be “merged”
- In coherent dedispersion modes, one sampler is connected to 1, 2, or 8 processing computers (depending on bandwidth)
  - Each computer processes full, \( \frac{1}{2} \), 1/8th total bandwidth
  - Data written to subdirectories by bank (A, B, C, D, E, F, G H)
Merging VEGAS Data

- Depending on observing mode, data from VEGAS may need to be “merged”

- In incoherent dedispersion mode, multiple samplers can be used to cover up to 6 GHz (depends on receiver and IF system limits)
  - Data still written to sub-directories by bank
Merging VEGAS Data

- Depending on observing mode, data from VEGAS may need to be “merged”

- Data will be merged for you within ~24 hours
  - You can still do spot-checks of data from each bank (they are all valid PSRFITS files in and of themselves, just with a fraction of the total bandwidth)

- Search-mode data merged manually
  - Data volumes are large, so per-bank data needs to be deleted after manual checks for data integrity
  - Only fully merged data will be left
  - If you don’t see it within 24 hours, email your project friend
Support

• GBT pulsar support scientists have many years of experience working with pulsar data

• We can answer most questions about data reduction, and the third-party packages

• Can answer some questions about installing on your own computer

• Package-maintainers are also very helpful and responsive
The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.