Full-Stokes Calibration for Breakthrough Listen GBT Observations

Mark Siebert
mas745@cornell.edu
Mentor: Emilio Enriquez
University of California, Berkeley — Cornell University

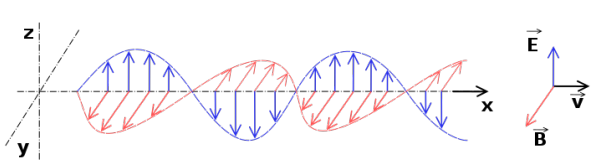
Summer 2018

Abstract

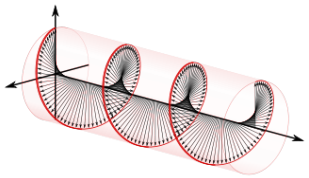
Polarization is an often overlooked property of electromagnetic radiation that can tell us a great deal about the underlying emission mechanisms of astronomical objects. It is critical for identifying stellar flare or potentially exoplanet radio emission with single dish radio telescopes, as it provides an important discriminant against radio frequency interference (RFI). Polarization information (in the form of the four Stokes parameters) is obtained through linear combinations of raw complex voltages from two orthogonal dipoles; however, Stokes parameters are highly sensitive to instrument error. In order to obtain accurate polarization information in radio observations, a calibration routine is crucial for identifying and correcting for this error. Currently, generating polarization information is not yet standard practice for Breakthrough Listen observations. Here I present a method for creating calibrated full-Stokes data products of Breakthrough Listen observations on the 100m Robert C. Byrd Green Bank Telescope using noise diode measurements and Mueller Matrix algebra. This method can be employed for any future GBT observations and could lead to a variety of promising astrophysical studies.

1 Introduction

In astronomical experiments one almost always seeks to measure different geometric properties of electromagnetic waves that are emitted from a particular object. The most important of these properties are the **amplitude** and **frequency** of the wave since these carry information about the intensity and energy per photon of emitted light, and thus provide a wealth of physical information about the source itself. These qualities of light can and have been used to do most astrophysical experiments; however, there is another valuable property of light that often goes overlooked by these experiments. **Polarization** refers to the orientation of a light wave in 3D space. The light wave in Fig 1a. is



(a) A linearly polarized EM wave.



(b) A circularly polarized EM wave

Figure 1: Examples of fully polarized light.

linearly polarized (i.e. $\vec{E}$ is confined to a single plane). The wave in Fig 1b. is circularly polarized where $\vec{E}$ comprises two orthogonal components that are oscillating $\frac{\pi}{2}$ out of phase.

Polarization information can offer clues into unique electromagnetic emission mechanisms, which can yield valuable physical properties of the source; however, in order to obtain polarization, one needs to measure the phase of the wave during an experiment. Because radio telescopes detect light by measuring the changing electric and magnetic fields directly (and thus retain phase information), they are the only instruments currently equipped to obtain polarization information from astronomical sources.

Currently, Breakthrough Listen uses about 20% of the available time on the 100m Robert C. Byrd Green Bank Telescope to search for technosignatures from intelligent civilizations. We are able to produce dynamic intensity spectra for a wide range of frequencies with incredibly detailed time resolution. However, there is no current method in place to obtain calibrated polarization information in addition to these Stokes-I data products. Over the course of this summer, I have developed a routine that can be used to create full-Stokes data products for any future Breakthrough Listen observations on GBT. This routine will be crucial to a variety of SETI experiments that could shed new light on the nature of life in the universe.

2 Motivation

Most of the light that we observe is unpolarized. That is, it consists of random incoherent combinations of all different polarizations at once. Some objects, however, do emit some degree of polarized light, and studying this polarization can tell us a lot about the object's physical properties.

Pulsars are fast rotating neutron stars that pulse with intrinsic and extremely consistent intensity profiles. These objects often also have intrinsic profiles in both linear and circular polarized light, which provide important information in understanding the nature of emission for these objects(1).

Fast Radio Bursts (**FRBs**) are mysterious transient radio events that have only been known about for a decade. The source of these events is still under discussion, but polarization data could be essential to identifying this. Michilli

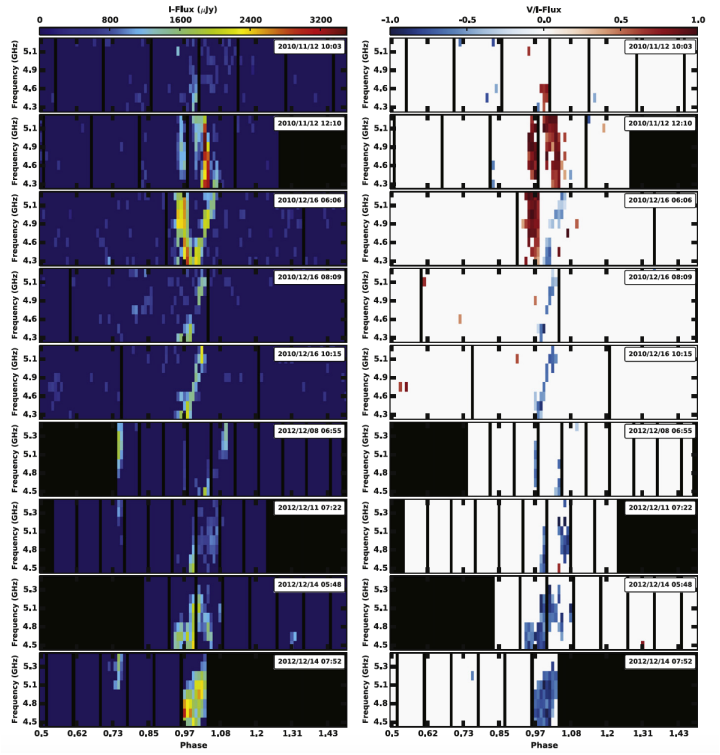


Figure 2: Multiple intensity (left) and circular polarization (right) measurements of 2M 0746+20 from Lynch et al. (2015)(4)

et al. (2018) show that FRB 121102 emits 100% linearly polarized light and use this to support the theory of this object being a highly-magnetized pulsar(2).

One particular method of radiation that is unique in its polarization is **electron cyclotron maser radiation**. This radiation results from electrons being accelerated in global magnetic fields, and it is 100% circularly polarized(3). Because flare stars are magnetized objects that eject large amounts of charged plasma regularly, circular polarization information can be used to glean information about them that would be impossible to obtain through intensity data alone. Figure 2 shows observations of electron cyclotron maser radiation from a very active brown dwarf. Lynch et al. (2015) uses these observations to construct a model of this star's magnetic field structure (4).

Furthermore, if a star has a planet with an intrinsic magnetic field, and this planet has a source of electrons such as a stellar wind, it will emit this circularly polarized light by the same mechanism(3). Not only would measuring this non thermal auroral emission be the first ever radio detection of an exoplanet, but it would also be the first confirmation of an exoplanet having its own magnetic field, which we know is crucial to life on Earth.

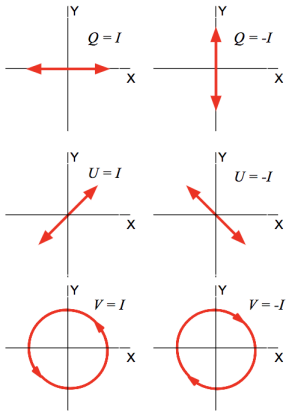


Figure 3: The Q, U, and V Stokes parameters

From a SETI perspective, there are more reasons to be interested in polarization data. For one, it provides a great discriminant against Radio Frequency Interference (RFI), which is something that has always plagued SETI science. Additionally, because there are so few polarized sources in the sky, the signal-to-noise ratio of polarized events is extremely high, which makes finding these sources very easy. Finally, because we are looking for intelligently engineered signals when we do SETI experiments, it is important to consider all possible ways a signal could be artificially generated. Although "polarization-modulated" signals are not a typical form of human communication, they would not be difficult to create and should still be considered when doing SETI science.

3 Calibration Procedure

3.1 Obtaining Stokes Parameters

In radio astronomy, polarization information comes in the form of the four Stokes Parameters. Together, these four numbers (I, Q, U, and V) completely describe the intensity and polarization of detected light. Stokes I quantifies the total flux density at a certain time and frequency, and Figure 3 shows how the other three Stokes parameters describe polarization.

The four Stokes parameters are obtained by taking various algebraic combinations of the real and imaginary parts of the raw data from two orthogonal dipoles. Heiles (2001) offers the following equations(5):

$$I = E_x E_x^* + E_y E_y* \tag{1}$$

$$Q = E_x E_x^* - E_y E_y* \tag{2}$$

$$U = E_x E_y^* + E_x^* E_y \tag{3}$$

$$-V = E_x E_x^* - E_x^* E_y \tag{4}$$

Although the negative sign in V is not present in Heiles (2001), I have included it to follow the convention that positive Stokes V corresponds to right-hand Circularly polarized light.

Breakthrough Listen's standard reduction routine after GBT observations, GPUSPEC, only produces Stokes I products from raw data files. However, there is another reduction routine, rawspec, which can be used to output cross polarization filterbank files. The values reflected in rawspec's output are $E_x E_x^*$, $E_y E_y^*$, $Re(E_x E_y^*)$, and $Im(E_x E_y^*)$ (arranged in this order in the second axis of the 3-D data array). Although these aren't the Stokes parameters themselves, Eqs. 1-4 above can be rearranged to get:

$$I = E_x E_x^* + E_y E_y* \tag{5}$$

$$Q = E_x E_x^* - E_y E_y* \tag{6}$$

$$U = 2 * Re(E_x E_y^*) \tag{7}$$

$$V = -2 * Im(E_x E_y^*) \tag{8}$$

It is worth noting that rawspec has an option to produce full-Stokes filterbank files instead of the crosspols, but there is currently an error in this routine as it leaves out the factor of 2 in Eqs. 7 and 8 above. Because of this, all of the calibration code outlined below acts on cross polarization rawspec outputs.

3.2 Necessary Measurements

To fully calibrate a GBT observation, there are four needed observations:

1. A one minute scan ON a calibrator source of known flux density with flickering noise diode
2. A one minute scan OFF the calibrator source (pointing at blank sky) with flickering noise diode
3. A one minute scan ON the target with flickering noise diode
4. The main observation ON the target (any duration and no noise diode)

Three of these four observations make use of the noise diode: a piece in the electronics chain that injects a signal into both feeds that is in phase and has constant power. At GBT, this noise diode "flickers" ON and OFF with a period of 40 ms. In testing calibration, we discovered that the flickering does not begin until 1s into the scan, and for this first second the noise diode is on the whole time. This effect has been accounted for in my calibration routine by rejecting the first second of scans 1-3 described above.

Additionally, the first two scans in the sequence outlined here serve only to calibrate the noise diode, which is said to be constant to within 30 days of a given observation. By this way, it would also be possible to fully calibrate an

observation with only scans 3 and 4, provided scans 1 and 2 had been done in the weeks prior. We have not currently tested the consistency of the noise diode's flux density to justify this "30 day" claim, but it would no doubt be a simple test using the scripts presented here.

3.3 Flux Calibration

The method employed to flux calibrate is largely the same as that used in the PSRCHIVE pulsar calibration software, which is outlined in van Straten et al. (2012)(6). The first step in both flux and polarization calibration is extracting the power spectra of the first three scans for noise diode ON and noise diode OFF. This requires extracting and averaging the ON and OFF levels from the flickering pattern of the noise diode.

To flux calibrate, first noise diode data for the first two scans are folded and integrated over the coarse channels (Fig. 5). This produces four spectra values that correspond to H_{on}, H_{off}, L_{on}, L_{off} in Eqs. 2-4. These equations are then applied to get the noise diode coarse channel spectrum, C_o, and the system temperature, S_{sys}.

Then the same folding and integrating method is applied to the third scan (on the target with flickering noise diode). The absolute on-target gain is then solved by taking

$$g_{target} = \frac{C_o}{H_{on} - L_{on}} \tag{9}$$

Here g_{target} is in $\frac{Jy}{count}$. Finally, the coarse channel data in the main observation is multiplied by the gain values and the system temperature is subtracted.

3.4 Polarization Calibration

There are many different ways of calibrating the Stokes parameters, and they all have varying accuracy. We chose to apply the most rudimentary calibration routine, that is, solving the Mueller matrix for the amplifier chain, because this method only requires the one minute noise diode observation on the target.

The Mueller matrix in a linear basis for the electronics chain is derived in Heiles (2001) (Fig. 5)(5).

Here, ΔG, corresponds to a difference in gain between the X and Y feeds, whereas ψ corresponds to a difference in phase. To solve the terms of this Mueller matrix, Using the third noise diode scan (on the target), the Stokes parameters are derived using Eqs. 5-8. Now, all four Stokes parameters are folded the same way as in flux calibration to get full-Stokes spectra for noise diode ON and OFF. By taking $ON - OFF$ for these spectra, we obtain the measured full-Stokes noise diode spectrum (Fig. 6).

Here we see a sinusoidal pattern in U and V which results from the fact that the constant path difference the two amplifier chains causes a frequency dependant phase difference. The elements of the Mueller matrix can be obtained from here because we assume that the noise diode is a "linear" signal at a 45^o

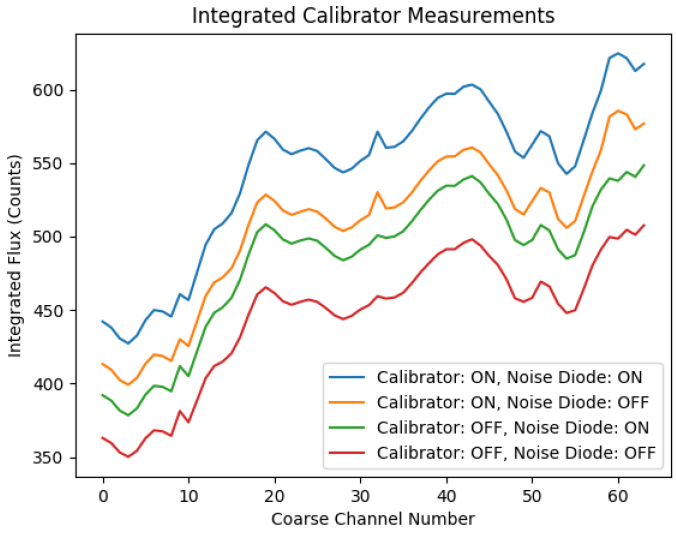


Figure 4: Power spectra from the first two calibration scans.

$$\mathbf{S_{out}} = \mathbf{M} \cdot \mathbf{S_{in}}$$

$$\mathbf{M_A} = \begin{bmatrix} 1 & \frac{\Delta G}{2} & 0 & 0 \\ \frac{\Delta G}{2} & 1 & 0 & 0 \\ 0 & 0 & \cos\psi & -\sin\psi \\ 0 & 0 & \sin\psi & \cos\psi \end{bmatrix}$$

Figure 5: The Mueller Matrix for the electronics chain from Heiles (2001)

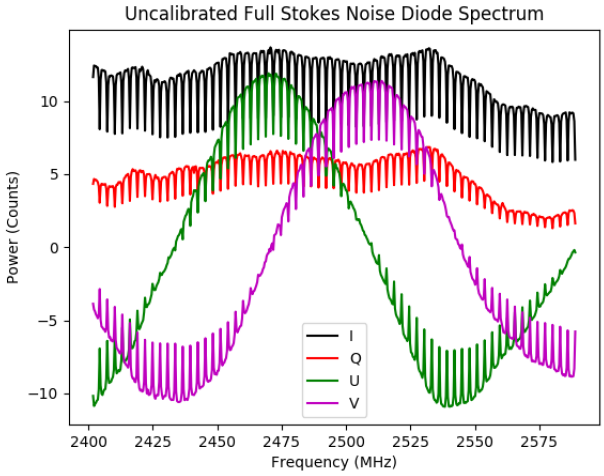


Figure 6: Example of a measured full-Stokes noise diode spectrum.

angle to the feeds ($I = U$ and $Q = V = 0$). ΔG was thus calculated for each frequency bin by taking $\frac{I_{off}}{Q_{off}}$ and ψ was calculated by taking $\arctan(\frac{V}{U})$. These values are then averaged over each coarse channel, so the calibration could be applied to data with any frequency resolution.

Finally, the Stokes parameters of the main observation are obtained and the inverse of the Mueller matrix is applied to each coarse channel to fully correct for the gain and phase differences.

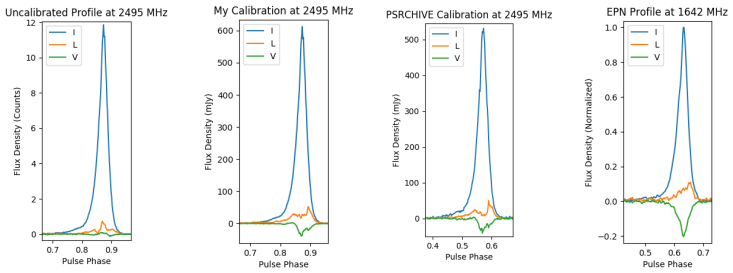
4 Testing Calibration

This calibration procedure was tested by comparing its results with those obtained by using the pre-packaged PSRCHIVE calibration routine. This routine requires the same four observations, and also only corrects for the errors in the amplifier chain. However, this routine can only be used for high time-resolution pulsar data whereas the one developed here can be applied to any filterbank resolution. Nevertheless, PSRCHIVE provides a great way to test the accuracy of this method.

Two sets of pulsar observations were used to test calibration for two different frequency bands. J0953+0755 was observed at S-band, and J1932+1059 at X-band. For each observation, the three one-minute noise diode measurements were taken using the quasar 3C249.1 as a calibrator source. Although data was taken for the whole frequency band, calibration was only applied on a single compute node because the full observation is too much memory for Python without integrating the tools into blimpy.

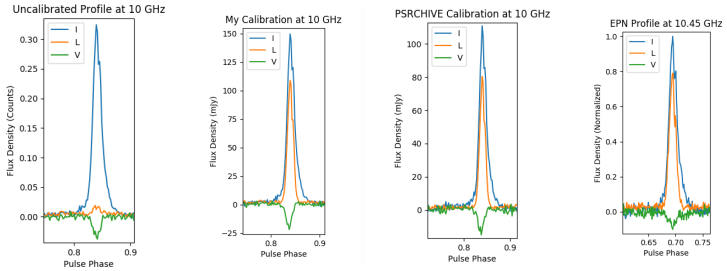
Pulse profiles were obtained by folding the calibrated filterbank files using dspsr and outputting them with pdv. From here, peak fluxes were obtained and average flux densities were calculated by integrating over the pulse (using a 3σ threshold for "on-pulse") and dividing by the total number of time bins. Profiles and flux values were then compared with those published on the *EPN Database of Pulsar Profiles* and previous studies of these pulsars.

4.1 Results



(a) (b) (c) (d)

Figure 7: B0950+08 Pulse Profiles



(a) (b) (c) (d)

Figure 8: B1929+10 Pulse Profiles

B0950+08	Frequency (GHz)	Average Flux Density (mJy)	Peak Flux Density (mJy)	Peak % Linear Polarization	Peak % Circular Polarization
My Calibration	2.495	23.9 ± .892	577	8.59	-6.42
PSRCHIVE Calibration	2.495	23.4 ± 1.68	534	9.48	-7.37
Jankowski et al. (2017)* / EPN** Database	2.495* 1.642**	30.8 ± 24.1*		10.8**	-20.3**

Figure 9

B1929+10	Frequency (GHz)	Average Flux Density (mJy)	Peak Flux Density (mJy)	Peak % Linear Polarization	Peak % Circular Polarization
My Calibration	10.05	2.80 ± 1.15	136	72.9	-14.7
PSRCHIVE Calibration	10.05	2.25 ± 1.37	111	72.4	-13.5
Seiradakis et al. (1995)* / EPN** Database	10.71* 10.45**	1.82*		79.0**	-10.2**

Figure 10

4.2 Discussion

We were able to fully calibrate these pulsar observations very close to that done by PSRCHIVE. Although our calibration was consistently high in peak/average flux density, this discrepancy could result from the fact that our method calibrates each coarse channel, whereas PSRCHIVE calibrates each frequency bin separately. Our fractional polarizations were similar those calculated by PSRCHIVE to 1%, and these were relatively close to those published on the EPN Database. Many of the differences from these profiles could be explained by the fact that they are taken at different frequencies, and pulsars are known to show evolution (especially in L and V) over large frequency ranges. For example, B0950+08 shows much lower circular polarization than the EPN profile which was taken almost 1 GHz below our observations. At about 4 GHz though, the EPN profile shows Stokes V that is even lower than ours. Thus, it would make sense that between the two EPN profiles, one would receive about 10% circular polarization like we did.

These results indicate that our method of calibration is accurate enough to be applied to future observations with the Green Bank Telescope.

5 Outline of calib_utils Python Modules

The python scripts used to calibrate the observations above are currently located in a folder called "calib_utils" in my fork of the blimpy python package. Hopefully, in the near future, this will be integrated into a newer version of blimpy. Here I will list the methods contained in each of the three modules and describe each module's functionality. Methods in **bold** denote ones that need to be called to calibrate future measurements.

5.1 Module: fluxcal.py

- foldcal
- integrate_chans
- integrate_calib
- get_calfluxes
- get_centerfreqs
- **diode_spec**
- **calibrate_fluxes**

This module provides tools used to flux calibrate an observation given the four scans listed in Section 3.2. **diode_spec** takes in two rawspec filterbank files (scans ON and OFF the calibrator), details (flux density, frequency, and spectral index) about the calibrator source to obtain C_o and T_{sys}. **calibrate_fluxes**

Takes in rawspec filterbank files of scans 4 and 3, as well as C_o and T_sys, and uses these to write a new flux-calibrated filterbank file with the ".fluxcal" extension.

5.2 Module: stokescal.py

- get_stokes
- convert_to_coarse
- phase_offsets
- gain_offsets
- apply_Mueller
- **calibrate_pols**
- fracpols
- write_stokefils
- write_polfils

This module provides tools used to calibrate the Stokes parameters for an observation given scans 3 and 4 described in Section 3.2. To calibrate, simply call **calibrate_pols** with the names of two rawspec cross polarization filterbank files corresponding to scans 4 and 3. This function will calculate ΔG and ψ using "gain_offsets" and "phase_offsets", then apply the inverse Mueller matrix to the main observation using "apply_Mueller". It writes a new filterbank file with the extensions ".SIQUV.polcal".

5.3 Module: calib_plots.py

- get_diff
- plot_Stokes_diode
- plot_calibrated_diode
- plot_phase_offsets
- plot_gain_offsets
- plot_diode_fold
- plot_fullcalib
- plot_diodespec

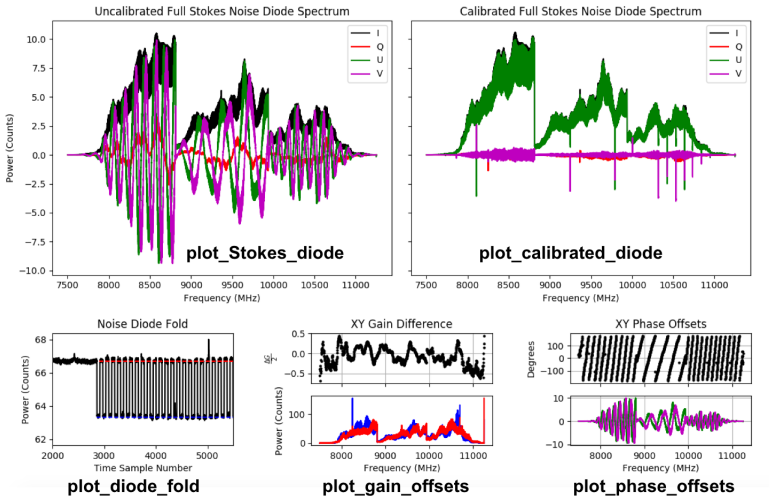


Figure 11: Output of "plot_fullcalib" with annotations denoting the outputs of other plotting functions within "calib_plots"

This module provides plotting tools that can be used to show how the calibration process will be carried out given a certain noise diode scan. "plot_fullcalib" can be used to consolidate all needed information into one figure (Figure 11), and it gives the user enough information to know if the Stokes calibration will be done correctly. This function is most useful when applied to the on-target noise diode scan, as this is the scan that is used for polarization calibration. The last plotting method, "plot_diodespec", uses the first two calibration scans to calculate and plot the noise diode's power spectrum.

5.4 Steps for Full Calibration

Here I will list out the basic steps needed to fully calibrate a Breakthrough Listen observation on GBT. These steps can be used for any band, and any frequency resolution, provided the four necessary scans are present as rawspec cross polarization outputs.

1. Call 'plot_fullcalib(Scan 3)' and examine plots
    - Make sure diode fold is correct: red lines should be on top of the ON time samples while blue should be on OFF.
    - Make sure gain and phase offsets are normal. Gain difference should be less than one and phase offsets should be nearly linear
    - Make sure the noise diode calibrates correctly. That is, $I \approx U$ and Q, U are nearly 0.
2. Call 'calibrate_pols(Scan 4, Scan 3)'

- Produces a '.SIQUV.polcal.fil' file (Scan 4 polcal)

3. Call 'calibrate_pols(Scan 3, Scan 3)'
   - Produces a '.SIQUV.polcal.fil' file (Scan 3 polcal)
4. Call 'dspec,Tsys = diode_spec(Scan 1, Scan 2, flux density, frequency, spectral index)'
5. Call 'calibrate_fluxes(Scan 4 polcal, Scan 3 polcal, dspec, Tsys, fullstokes=True)'
   - Produces a '.SIQUV.polcal.fluxcal.fil' file. This is the final calibrated product.

5.5 UPDATE: Circular Feed Calibration with Stokescal

As of Fall 2018, stokescal.py has functionality to perform full Stokes calibration in a circular feed basis in addition to the linear dipole calibration described in previous sections. This is necessary because many of the receivers at GBT (including X-band, which is frequently used in standard Breakthrough Listen observations) output circular polarizations instead of linear.

To transform all necessary calculations to a circular basis, we use the following equations:

$$E_r = \frac{E_x + iE_y}{\sqrt{2}} \tag{10}$$

$$E_l = \frac{E_x - iE_y}{\sqrt{2}} \tag{11}$$

With these, eqs. 5-8 become:

$$I = E_r E_r^* + E_l E_l^* \tag{12}$$

$$Q = 2 * Re(E_r E_l^*) \tag{13}$$

$$U = -2 * Im(E_r E_l^*) \tag{14}$$

$$V = E_r E_r^* - E_l E_l^* \tag{15}$$

Because the rawspec routine outputs dynamic spectra in the order:

1. Power in feed 1 ($E_a E_a^*$)
2. Power in feed 2 ($E_b E_b^*$)
3. Real part of correlated product ($Re(E_a E_b^*)$)
4. Imaginary part of correlated product ($Im(E_a E_b^*)$)

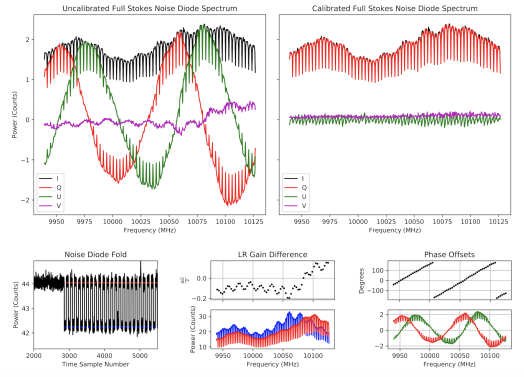


Figure 12: Output of "plot_fullcalib" for circular dipole calibration.

Choosing the circular feed option in stokescal.py (using feedtype='c') combines rawspec output arrays in ways different from the linear basis to produce the stokes parameters (e.g. $V = (output1) - (output2)$ instead of $V = -2 * (output4)$).

In addition to the change in definition of the Stokes parameters with respect to rawspec data products, converting to a circular basis also changes the Mueller matrix used for calibration. Like before, we consider error in the amplifier chain to be characterized by a power gain difference (ΔG) and a phase difference (ψ) between the two feeds. For circular dipoles, the Mueller matrix becomes:

$$\boldsymbol{M_A} = \begin{bmatrix} 1 & 0 & 0 & \Delta G/2 \\ 0 & cos\psi & -sin\psi & 0 \\ 0 & sin\psi & cos\psi & 0 \\ \Delta G/2 & 0 & 0 & 1 \end{bmatrix}$$

From this we can see that there is now sinusoidal coupling between Stokes Q and U (instead of U and V in the case of linear feeds). We follow the same calibration routine as before, again assuming the noise diode signal is in phase and equal in power for the two antennae. This means $2Re(E_aE_b*) = I$ and $2Im(E_aE_b*) = 0$ but now, this implies the ND signal is pure Q (instead of pure U in the linear case). We calculate ΔG and ψ using these assumptions, and if done correctly, the calibrated noise diode plot should show $U \to 0$, $V \to 0$, and $Q \to I$ (like in fig. 12).

To make use this circular dipole calibration functionality, use the parameter *feed_type* within stokecal.py and calib_plots.py. When this parameter is set to 'c', the calculation of the Stokes parameters and Mueller matrix is done as described in this section. Other than this detail, the steps for full calibration are the same as described in Section 5.4.

6 Questions?

If you have any questions about the code or how to use it, feel free to contact me:

Mark Siebert
Undergraduate - Cornell University Astronomy Department
mas745@cornell.edu

References

[1] J. Dyks, "The origin of radio pulsar polarisation," *ArXiv e-prints*, May 2017.

[2] D. Michilli, A. Seymour, J. W. T. Hessels, L. G. Spitler, V. Gajjar, A. M. Archibald, G. C. Bower, S. Chatterjee, J. M. Cordes, K. Gourdji, G. H. Heald, V. M. Kaspi, C. J. Law, C. Sobey, E. A. K. Adams, C. G. Bassa, S. Bogdanov, C. Brinkman, P. Demorest, F. Fernandez, G. Hellbourg, T. J. W. Lazio, R. S. Lynch, N. Maddox, B. Marcote, M. A. McLaughlin, Z. Paragi, S. M. Ransom, P. Scholz, A. P. V. Siemion, S. P. Tendulkar, P. van Rooy, R. S. Wharton, and D. Whitlow, "An extreme magneto-ionic environment associated with the fast radio burst source FRB 121102," *Nature*, vol. 553, pp. 182–185, Jan. 2018.

[3] A. D. Maps, T. S. Bastian, and A. J. Beasley, "A Search for Radio Emission from Nearby Exoplanets," in *American Astronomical Society Meeting Abstracts #229*, vol. 229 of *American Astronomical Society Meeting Abstracts*, p. 146.35, Jan. 2017.

[4] C. Lynch, R. L. Mutel, and M. Güdel, "Wideband Dynamic Radio Spectra of Two Ultra-cool Dwarfs," *ApJ*, vol. 802, p. 106, Apr. 2015.

[5] C. Heiles, "A Heuristic Introduction to Radioastronomical Polarization," in *Single-Dish Radio Astronomy: Techniques and Applications* (S. Stanimirovic, D. Altschuler, P. Goldsmith, and C. Salter, eds.), vol. 278 of *Astronomical Society of the Pacific Conference Series*, pp. 131–152, Dec. 2002.

[6] W. van Straten, P. Demorest, and S. Oslowski, "Pulsar Data Analysis with PSRCHIVE," *Astronomical Research and Technology*, vol. 9, pp. 237–256, July 2012.