7-17-2012

Noise Reduction in Millisecond Pulsar Timing

Michael Lam

Noise Reduction in Millisecond Pulsar Timing

By Michael Lam, Class of 2011

Advisor: Paul Demorest, National Radio Astronomy Observatory

Astronomers use pulsars as astronomical clocks to perform physical tests, such as in testing Einstein's Theory of General Relativity. As these tests require more precise clock time measurements, we run into problems associated with the pulses traveling through particles in space. Using data taken simultaneously at two radio telescopes, we looked to improve the precision of measurements taken of two pulsars. We looked for correlations between data sets using a variety of techniques to account for the error in our measurements. We report our results in the hopes that these methods can be used in future pulsar timings.

Introduction

Stars undergo a life cycle occurring over millions to billions of years with a variety of evolutionary tracks. At the end of their lives, stars become a new class of celestial object based upon their masses. An average mass star like our Sun will eventually become a Red Giant, expanding in size and shedding off its outer, gaseous layers, and leaving the burnt out core behind, forming what is known as a white dwarf. Larger mass stars between 1.4 and 2.0 times the mass of the Sun will form a Red Supergiant and eventually supernova. Left behind is what is called a neutron star. Much heavier stars will also supernova, the portion left behind collapsing under the force of gravity into a black hole. This paper will discuss the uses of a subset the middle class of stellar remnants: pulsars.

Neutron stars are composed of mostly neutrons surrounded in a thin shell of iron. Those that spin along a rotation axis are called pulsars. Pulsars viewed from Earth are analogous to cosmic lighthouses. Each has a magnetic axis offset from their rotation axis that shoots off radio beams. It is believed that all neutron stars that we see are pulsars since no non-rotating neutron stars have been discovered, though over the course of the age of the Universe many probably exist now. A similar offset phenomenon can be seen with Earth's axial tilt. Our rotation axis is offset from the plane of the solar system which causes seasons since once side of the Earth points directly toward the Sun on one side of its orbit (summer) and away on the other side (winter). Similarly, with pulsars, when the pulsar's radio beam sweeps along our line of sight we can see the radiation since it is pointing directly toward us. This happens
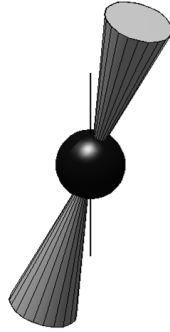
Figure 1: Pulsar Schematic

regularly as the pulsar spins and is called a pulse. Each pulse from a standard pulsar, one set spinning as a result of the supernova, occurs on the order of once per second.

The pulsars studied here are known as millisecond pulsars (MSPs). These pulse several hundred times per second (once every few milliseconds, or one-millionth of a second) and are believed to be the result of a pulsar and another object, usually a star, in the same stellar system. Mass moves from the star to the pulsar, increasing the pulsar's spin rate (period) as angular momentum is moved. This period slows down as energy is lost through the beam, decreasing the angular momentum of the pulsar, but the effect is very small (on the order of less than a nanosecond, or one-billionth of a second, per year). Because MSPs rotate so rapidly, emit regular pulses, and do not slow down much, they are the most accurate astronomical clocks known. Using them, we can test properties of physics, specifically as the signal passes through space, changes in some way, and reaches Earth.

The long-term goal of this project is to use pulsars as clocks to directly detect a phenomenon predicted by Einstein's Theory of General Relativity known as Gravitational Wave Radiation. General Relativity provides a geometric explanation for gravity in the context of space and time. Our universe contains three dimensions in space and one in time, interconnected in what is called spacetime. While we think of the universe as Euclidean in nature, that one can move straight in any six directions along three perpendicular axes (the x-y-z axes), spacetime is actually curved due to mass. In Figure 2, a mass is placed on a two dimensional analogy of our three spatial dimensions. The mass curves space and creates a gravitational potential well.
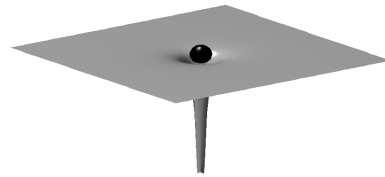

Figure 2: 2D Curved Space

Figure 3a shows the naïve view of space, without any mass, with straight lines connecting each of the points. Figure 3b adds a mass to the middle and shows how points in space curve towards it under its gravitational influence.
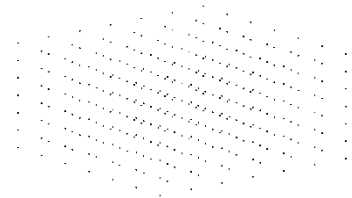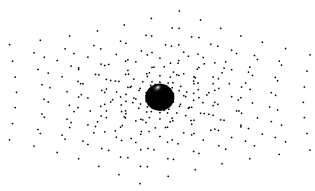

Figure 3a: 3D Flat Space

189

Figure 3b: 3D Curved Space

In 1974, Hulse and Taylor made the first indirect detection of this radiation using two pulsars orbiting each other. Figure 4 shows that as two massive objects orbit each other, gravitational waves are produced as a result of ripple-like disturbances in the curvature of spacetime. The two pulsars
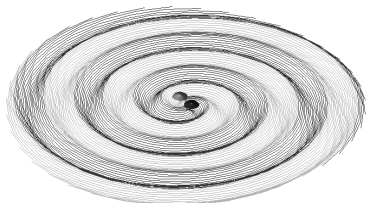

Figure 4: Spiral Gravitational Waves in 2D

in the Hulse and Taylor experiment had orbits that showed a slight decay, the two pulsars rotating toward their common center, a consequence of gravitational waves consistent with General Relativity. Over the past decade, lasers have been used to attempt to directly detect gravitational radiation passing through the beams though so far this has been unsuccessful. Recently, efforts have been made to utilize pulsars to accomplish a direct detection. By measuring extremely small shifts in the frequency (rate) of pulses as a result of this radiation, gravitational waves passing between the signal and Earth should be possible to detect using a sensitive enough detector. Currently, detectors are believed to be sensitive enough to detect these

shifts but random scatter in the measurements (noise) are overwhelming the signal. Systematic effects contribute to errors in these measurements. One source of this error comes from the particles in space known as the interstellar medium (ISM). Space is not a perfect vacuum but is filled with gas and dust particles, though on a significantly lower level than found on Earth. Plasma, or ionized gas, moves around through space, creating screens of particles moving together. This turbulent plasma in the ISM distorts the path of the signal much like the path of starlight changes as it passes through the Earth's atmosphere, causing commonly seen twinkling. While part of this is due to changes in the human eye, the rest relies on changes in the path of light. Light travels along different paths through different mediums, as evidenced by sticking a pencil into water and watching it appear to bend. The particles in the atmosphere are moving chaotically and shift the path of light.

The accuracy in these timing measurements is therefore limited by the ISM. This scattering effect must be removed to reduce noise levels to the point where we can even begin to attempt to detect gravitational wave radiation. Each pulse of the pulsar broadens and thus instead of a precise measurement of the peak, any fits from the regular timing models produce errors. This error contribution should be proportional to the amount of scattering time delay due to the ISM. Errors in pulse times of arrival (TOAs) are theorized to need to be reduced from roughly 100 nanoseconds (ns) to around 1

190

ns in order for gravitational wave radiation to be detected.

Observations

We present three nights of observations taken simultaneously at two radio telescopes, Arecibo Observatory and the Robert C. Byrd Green Bank Telescope (GBT), in mid-2008. The process by which radio telescope works is different from optical telescopes. Radio telescopes are large dish structures that collect radio waves into an antenna. These signals are stored as voltages, which correspond to the energy and phase of each radio wave. Optical telescopes gather light by sending it over mirrors and through lenses, leading it to a CCD camera where it counts individual photons. By retaining the phase of each radio wave, we are able to later electronically reconstruct the received light and discern something from its wave properties.

Observations at each site lasted just over one hour in both L (centered 1410 MHz) and S (2650 MHz) frequency bands with a 64 MHz bandwidth, described later in the timing analysis section. Two different backends were used: ASP, which recorded TOAs, and the VLBA Mark5 recorder, which records raw voltage samples for measurements of intensity with respect to time of observation and frequency. We ran post-processing on the samples to obtain any time/frequency resolution needed for data reduction, provided in a data cube known as a dynamic spectrum.

The observations consisted of measurements of the pulsars B1937+21 and B1713+07, chosen because of their usefulness in conducting physical tests.

1937+21 was the first MSP discovered (Backer et al. 1982) and is currently the second fastest MSP observed. 1713+07 also has a high flux density, as well as a shallow spectrum and sharp pulse profile peak, allowing for accurate pulse measurements over a large range of frequencies. These specifications make these prime candidates in these tests.

Data Reduction

The data was reduced in three stages. We began by converting the raw data into a readable format. This was then processed using a combination of pulsar timing analysis packages. The PSRCHIVE package (http://psrchive.sourceforge.net) contains a library of data reduction algorithms specific to pulsar .fits files. Dr. Demorest made several additions and extensions to these packages. We first made Gaussian fits of the pulse profiles as a comparison model using paas. Using pat, we next generated TOAs that were provided to the program TEMPO. This fit timing models to the TOA data, residuals were calculated, and lastly converted into ASCII output by print_resid.

The next portion of the project involved calibrating the data and modifying the parameter files for the two pulsars using pac and pam respectively. We then ran psrflux to obtain the integrated flux density, analogous to the intensity of the source, with respect to frequency and time for each observation, which is also explained further in the analysis section. With all of the data processed, we wrote a variety of python scripts to analyze the data using the scipy/numpy packages and viewed with the matplotlib (pylab) package. The first

191

program developed was a graphical interface for comparing and correlating the timing residuals generated by TEMPO. The second was used to perform our analysis of each of the data cubes.

Timing Delay Analysis

First, we tried to determine the statistical correlation between the timing residuals of both observatories for each pulsar and frequency band. We would optimally expect a random scatter around the residuals. Any trends in the residuals in one set of observations would imply some sort of systematic effect, though similar trends in both might indicate external effects on the signal. Unfortunately, computing the cross correlation, explained below, yielded no discernable results.
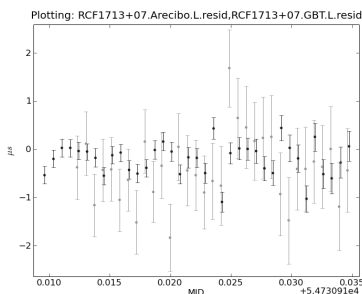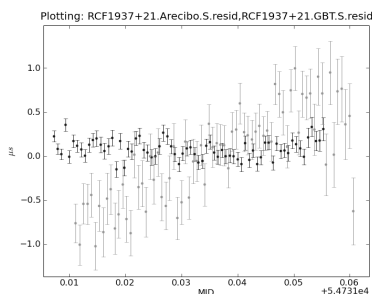
The process of correlation is often used in signal analysis to determine patterns within a signal or between signals. It involves taking one discrete signal and convolving it with itself or another, yielding the autocorrelation or cross correlation respectively. Figure 6 demonstrates the process of convolution, in which one of the signals is shifted along the time axis and the overlap or area between each signal is calculated. At each time shift, or lag, this value is computed. The final result is normalized between -1 and 1. In autocorrelation, since the signal is being convolved with itself, at a time lag of zero, the entire signal will overlap itself and this value is entirely correlated.

Just as a correlation can be computed in one dimension, it can also be computed in higher dimensions. This technique will be used in the next step of analysis but will be shown here in Figure 7. In a data cube, the two axes on the bottom can be anything, with the z-axis representing some type of intensity. For each x lag and each y lag, the overlap is again computed over the entire domain of the signal.

Figure 5a: Possible correlation

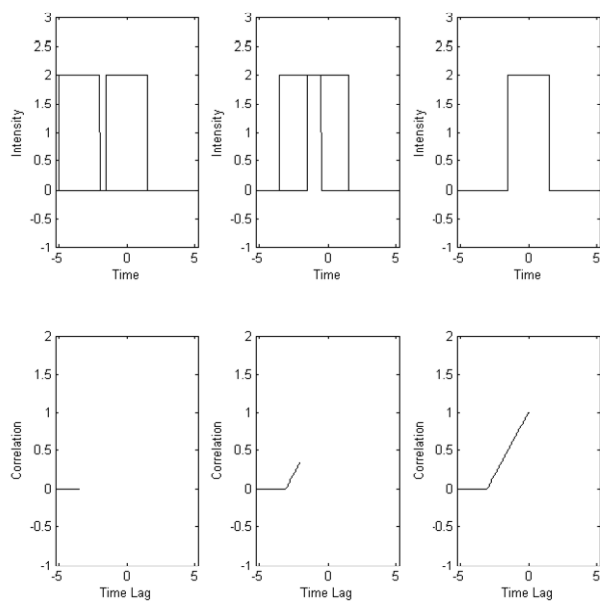Figure 5b: Systematic effect in one telescope

192

Figure 6a: Example of autocorrelation. As one signal is passed over another, there is more overlap. In the case of a square pulse, the overlap increases linearly. The peak is shown at a time lag of zero, when the signal overlaps itself.
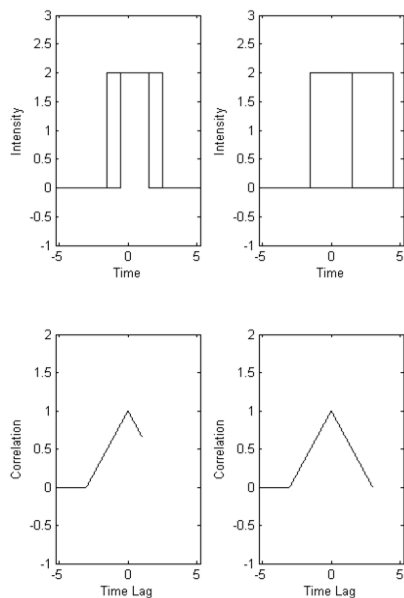

Figure 6b: As a square pulse is convolved at positive time lags, since it is a symmetric function, the correlation will also be symmetric.
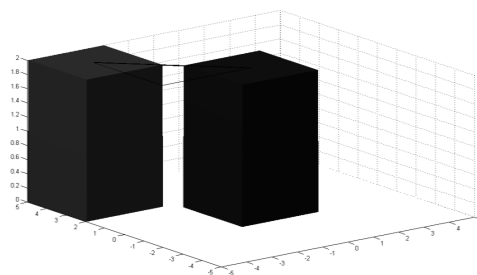

Figure 7a: Autocorrelation, computed in two dimensions. Thus, the signal is moved around in both the x- and y-axes.
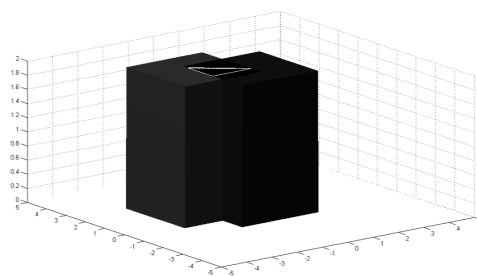

Figure 7b: The two signals have some overlap just as in the one-dimensional case. Instead of one axis lag, there are now two. These do not necessarily have to be in the same units.

The remainder of our analysis was done with the dynamic spectrum of a pulsar, or its intensity with respect to both frequency of the light waves and time of observation. The frequency of a wave is how fast it cycles per unit time. One Hertz (Hz) is equal to one cycle per second. Two sample sine waves are shown in Figure 8. For any
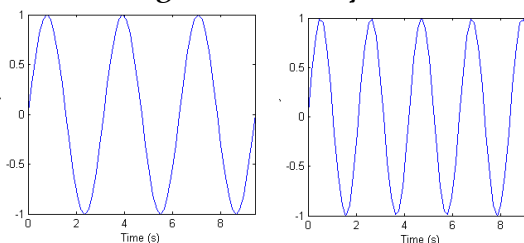

Figure 8: Two waves. Compared to the first, the second has a shorter wavelength (peak-to-peak distance) and a larger

193

frequency since it cycles more quickly in the same amount of time.

wave, the amount of energy associated with it is proportional to the square of its height, or amplitude. Thus, the dynamic spectrum is a data cube representing the intensity of light waves of a given frequency observed at each time interval. Figure 9 shows a typical dynamic spectrum. Without the effects of
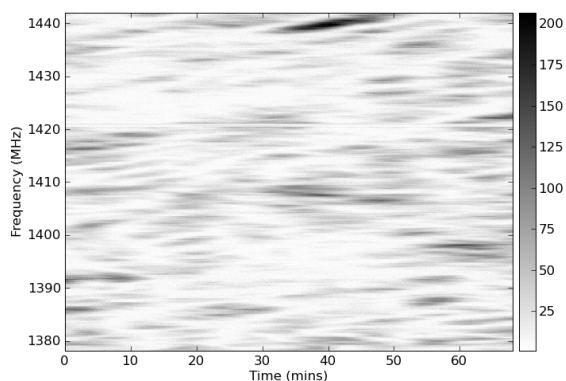

Figure 9: Dynamic spectrum for B1937+21 at Arecibo in L band on Julian Day (JD) 2454730. Darker indicates greater intensity. This is value is in calibrator units, where the approximate average intensity (flux) level of 10.0 mJy is equal to 22 calibrator units (1 Jy = 10^{-26} W/(m^2*Hz)).

the ISM, this would appear roughly flat in nature but because of the scintillation effects, observed intensity shifts both in frequency and in time.

Frequency channels were clipped around the 1420 MHz area as hydrogen gives off strong radiation at this frequency. Channels ruined by radio frequency interference generated by humans were also clipped. Traditional analysis with the dynamic spectrum involves computing the autocorrelation

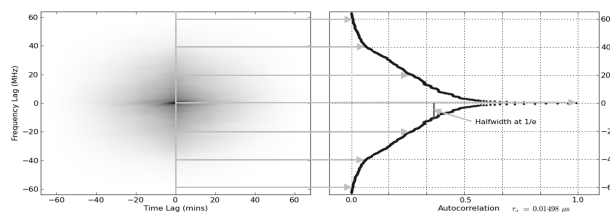function (ACF) of the dynamic spectrum in two dimensions. At the

Figure 10: ACF (left) for the same data set as Figure 9 and the profile at zero time lag (right). The halfwidth at 1/e is shown.

values of zero time lag for the ACF, the scattering time delay is the halfwidth at 1/e, which is a strong function of the strength of ISM scattering.

We used three other methods using the dynamic spectrum like the one shown in Figure 1 to determine the timing delay due to the ISM. Channels around the 1420 MHz hydrogen line, as well as channels ruined by radio frequency interference, were clipped in each method. The traditional method involves computing the autocorrelation (ACF) of the dynamic spectrum. At the values of zero time lag for the ACF, the scattering time delay is the halfwidth at 1/e, which is a strong function of the strength of ISM scattering.

Our next method was to take the power (secondary) spectrum of the dynamic spectrum. This is performed by taking the square of the Fourier transform. A simple, one-dimensional Fourier transform takes a signal and determines the strength of different frequency waves throughout it. Figure 10 shows two signals, one a simple sine wave, one a half-sinc wave. For the sine
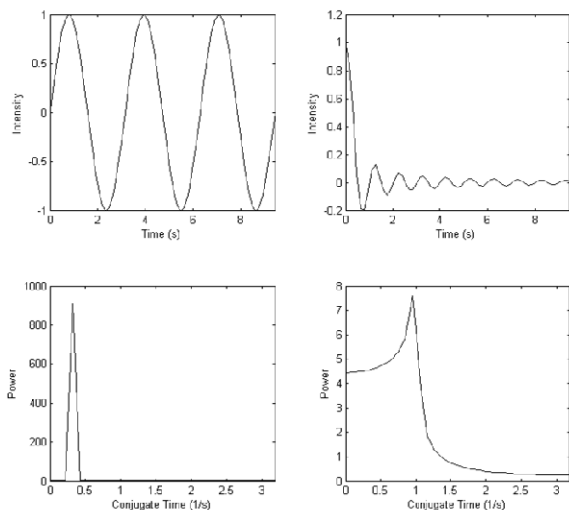
194

Figure 10: Sine wave (top left) with the square of the Fourier transform (bottom left). Half-sinc wave (top right) with its square Fourier transform (bottom right)

wave, since there is only one component frequency present, there is only power at this point in the frequency axis. In actuality, this is the conjugate time axis, which is proportional to the frequency and the rate of sampling for a discrete sample. For the half-sinc wave, smaller frequencies fit the curve better, with a clear peak. At larger frequencies, there is less power present. This is similar in nature to how convolution fits the curves, where different sinusoidal waves are analogously convolved with the signal.
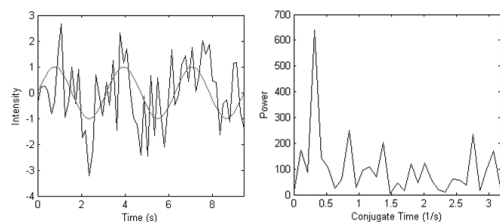

Figure 11: A noisy signal with an underlying sine component. Sinusoidal waves are fit to the curve as shown (left). The power associated with this wave is placed on the conjugate time axis (right).

There is some power associated with higher frequency waves, notably around 2, 3, and 4.5 times the base frequency

The power spectrum is the squared, two-dimensional Fourier transform of the dynamic spectrum. A 2D Hanning window was applied to the data to prevent edge effects present in Fourier transform computations. This smoothes the power spectrum so that the edges hit zero, changing the data cube into a weighted data two-dimensional sine wave for better fitting. Otherwise, in order to fit the edges of the cube, smaller and smaller frequencies would obtain power to fit it properly, creating unstable ringing.

In this two-dimensional case, the frequency axis transforms to a conjugate time axis and the time (observation) axis transforms to a conjugate frequency axis. This conjugate frequency axis is not the frequency of specific light waves but rather the rate of intensity changes over the course of an observation. The same applies to the conjugate time axis, which can be thought of as the time delay axis. Looking back at the dynamic spectrum (Figure 9), this tries to find the characteristic scales of the scintillation in both axes. Figure 12 shows the power spectrum. Note that power is concentrated close to around zero conjugate
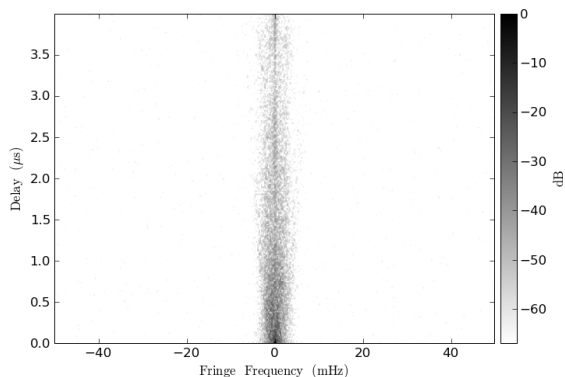
195

Figure 12: Secondary spectrum for the dynamic spectrum in Figure 9. Power is displayed logarithmically and is scaled to the highest point.

(labeled fringe) frequency. With better sampling resolution, parabolic arcs can be seen more clearly toward the edges of this structure, showing the relationship between time delay and frequency shift in the ISM.

On the secondary spectrum, we calculated the noise levels by taking two boxes in the upper corners and subtracting the mean value to make this zero. We integrated the total power across each time delay, producing the projected secondary spectrum in Figure 13. In the next step, we
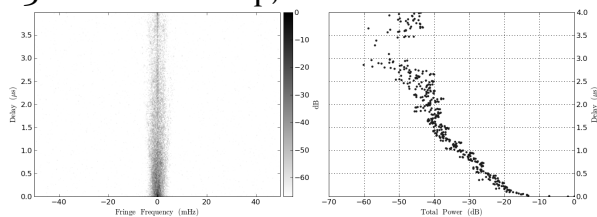

Figure 13: Secondary spectrum (left) and its renormalized projection (right), the sum of the power for each time delay. Increasing delay is on the vertical axis with total power increasing to the right.

calculated the running first moment (cumulative delay), defined as

$$\Gamma_1(\tau) = \int_0^\tau \frac{t f(t)\, dt}{f(t)\, dt},$$

by Hemberger and Stinebring (2008). This can be thought of as taking the amount of power up to each delay point in the projected secondary spectrum and accumulating it. As the time delay increases, the power goes down and thus less is accumulated at each subsequent point. Therefore, we reach an asymptotic limit, as
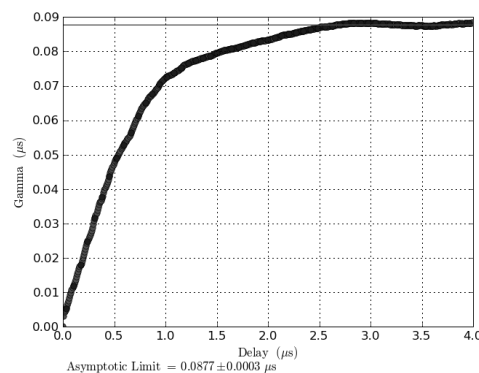

Figure 14: Cumulative Delay.

shown in Figure 14. This limit estimates the scattering time delay. Fitting this curve produced undesired effects, so the asymptotic limit was calculated by running a window filter over the curve, averaging every five points together for a smoother fit, and then calculating the discrete derivative. As the slope approached zero within certain tolerances, the remaining points were averaged together to find the asymptotic limit.

Our final method involved taking one-dimensional Fourier transforms along each time slice of the dynamic spectrum. A similar procedure was performed to calculate the asymptote of the cumulative delay of each of the Fourier slices. Figure 15 shows this method. This produced curves that

196

showed variations on timescales we were hoping to relate to the timing residuals. Initial attempts to correlate the two did not provide any meaningful results. Unfortunately, we were unable to finish our analysis of this section.
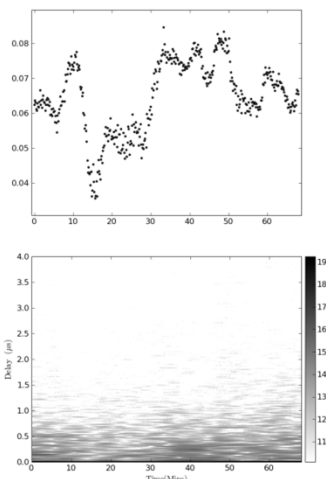

Figure 15: 1D Fourier Slices (bottom) of the dynamic spectrum and the cumulative delays for each time slice (top). This shape did not seem to relate to the shape seen in the corresponding timing residuals.

Results

Using both the values provided by the cumulative delays and the ACFs, we determined delays on the order of 1 - 100 ns as shown in Figure 16 for two nights at
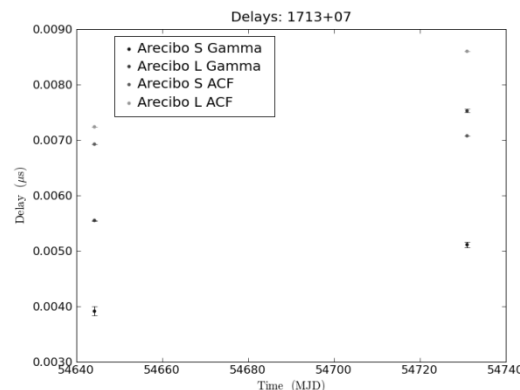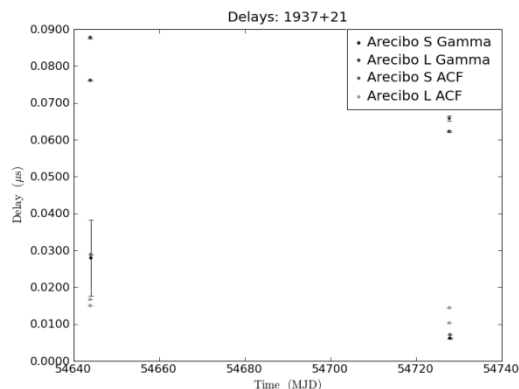

Figure 16a: Scattering Time Delays for both pulsars at Arecibo Observatory. Large error on the one point at top on the left is attributed to bad radio frequency interference producing poor fits.

Arecibo. We have improved our noise estimations by about a factor of 10. We can see that we are severely limited by the amount of data we have. With only three nights of data, and only two fully analyzed, we cannot get a good estimate of any trend present. Lack of time forced us to pause our analysis of GBT observations and limit our attempts at the 1D Fourier slices method. One of the important steps we need to take next will be in determining the scaling factor between the scattering time delay and the physical delay error due to the ISM. Once this number is approximated, we can subtract this amount globally from pulsar timing measurements. We hope that with longer observations and higher frequency resolutions (providing finer conjugate time resolution) we can determine finer structure due to the ISM and reduce our noise levels even further. After testing these numerical analysis techniques, we see that it is feasible to continue this line of research to better pulsar timing and provide an appropriate model that can be made to subtract this

197

noise and reduce timing residual levels needed to directly detecting gravitational wave radiation.

Acknowledgments

I would like thank my advisor Paul Demorest, the director of the NRAO Charlottesville summer students Jeff Mangum, and all of the radio astronomers who both helped me this summer and opened the door to a new realm of astronomy.

References

Backer, D., et al., 1982, Millisecond Pulsar in 4C 21.53, IAU Circ. No. 3743 (1982).

Cordes, J.M., 2002, Pulsar Observations I. -- Propagation Effects, Searching Distance Estimates, Scintillations and VLBI, in Stanimirovic, et al, ed., Single-Dish Radio Astronomy: Techniques and Applications, The Astronomical society of the Pacific, San Francisco, p. 227-250.

Hemberger, D.A., and Stinebring, D.R., 2008, Time Variability of Interstellar Scattering and Improvements to Pulsar Timing, ApJ, 674: L37-L40.

Hulse, R. A. & Taylor, J. H. 1975, ApJ, 195, L51

Lorimer, D., and Kramer, M., Handbook of Pulsar Astronomy, Cambridge, 2005.

Maron, O., Kijak, J., and Wielebinski, R., 2004, Observations of millisecond pulsars at 8.35 GHz, A&A, 413: L19-L22.

Splaver, E.M., et al., 2005, Masses, Parallax, and Relativistic Timing of the PSR J1713+0747 Binary System, ApJ, 620: 405-415.