Recovering Dispersion Measure Fluctuations from the Interstellar Medium

Calvin Tsing Shing Chan, 1004096630
University of Toronto, Scarborough

December 2020

Abstract

Understanding the dispersion measure (DM) fluctuations from pulsar signals is important for correcting pulsar timing but also important for probing turbulent and lensing structures in the interstellar medium (ISM). In the case of lensing structures, most studies have only focused on short timescales, which translates to small scale structures in the ISM. There has been little work done to date on detecting lensing over long timescales. Part of the reason is because of missing time samples in the data collected. This work aims on reconstructing the missing time samples for both the intensity and the DM fluctuations to see if we can find a correlation between the two using simulations and a real data set. In simulations, we find that it should be possible to find a correlation between the two signals while in the real data set, we find potential correlation over large timescales. This would be one of the first detections of lensing from large scales in the ISM which are currently not well understood.

1 Introduction

In pulsar astronomy, a quantity that is often used is the Dispersion Measure (DM) of a pulsar. Its importance manifests itself when we are observing interstellar medium (ISM) passing by pulsar emissions. Since ISMs have varying electron density, as the pulsar emission passes through it, it interacts with the emission resulting in the signal being dispersed. This dispersion varies with the frequency range or bandwidth used to observe the pulsar signal [1]. The DM is defined as the integrated column density of free electrons along a line of sight

$$\mathrm{DM} = \int_0^d n_e(l)\mathrm{d}l, \tag{1}$$

where n_e is the free electron density along the line of sight l and d is the distance from the observer to the pulsar in question [3].

The DM of a signal is often measured as fluctuations through time or in other words, as a function of time. These measurements are mostly used for pulsar timing experiments, *e.g.* gravitational wave timing [3], however, they can also provide information about the ISM. More specifically, the derivatives of the DM can provide information on lensing caused by the ISM as it moves pass the pulsar emissions as well as the turbulent structure of the ISM [4]. Although these observational data sets collected from telescopes are often contaminated by noise and missing sample points. Working with these data sets can very much obscure analyse and results in an experiment. In this project, we attempt to overcome this contamination by applying reconstruction algorithms built by my supervisor in order to reduce the noise and fill in the missing sample points to recover a DM fluctuation signal. In doing so, we are able to analyze the DM data further, particularly, we look for possible signs of lensing in the ISM over long timescales.

Previously, most studies have only considered this type of lensing over small timescales. This may range from a few minutes to many pulses spanning a few days. However, here we are considering lensing events over timescales of up to half a decade. As such, this makes it incredibly more difficult to interpret the data for results and findings.

In the next section, we discuss the theoretical astrophysics of our interpretation and the relevant mathematics. We follow up by applying our interpretation to simulated data sets to have a better understanding of the results. We then move on to a real data set for further analyses.

2 Background & Theory

DM fluctuations through time or DM time series, is a tool that can provide information on the lensing and turbulent scales of the ISM that is passing by pulsar emissions [5]. Assuming that the ISM is passing by the pulsar emission at an approximately constant velocity, we can convert the DM timescale into spatial scale over the ISM. This means that the DM fluctuations through time

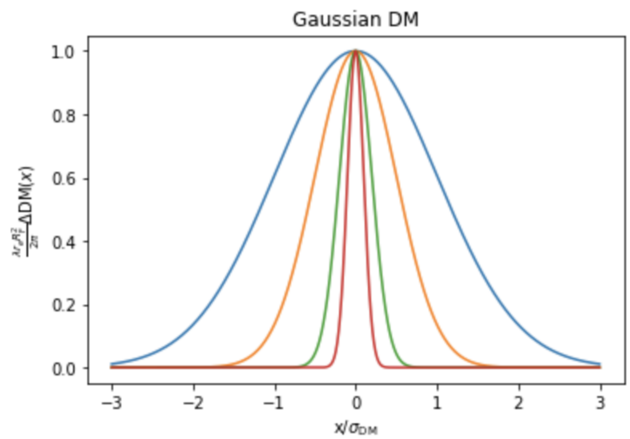

Figure 1: A dispersion measure fluctuation following a Gaussian function with the widths being scaled with respect to a fixed size, σ_{DM}. Each of these curves produce different lensed intensities, for example for the magnification for each curve, see the corresponding colours in Figure 2.

is determined by the different individual spatial scale fluctuations in the ISM. Furthermore, the derivatives of the spatial scales can provide information on possible signs of lensing. In theory, this would also provide information on the turbulent structure of the ISM, however, we will not analyze them in this project.

In order to find signs of lensing, it is possible to analyze the intensity fluctuation of the signal through time as the ISM passes by the pulsar emission. From the intensity, we can look for places that have a particular spike in the intensity, effectively known as caustics. Caustics is the phenomenon where light rays that are reflected or refracted by a curved surface or object converge together, causing a major increase in the intensity. In our context, this is exactly where the lensing event happens. The interaction between the pulsar emission and the ISM acts as the object that causes a refraction of the pulsar emission. This means that by observing for regions that have a sharp increase in intensity, it will also be the place for a lensing event to occur. However, as stated earlier, because the data sets collected by telescopes are often contaminated by noise and missing data points, we must validate whether the data we are using is accurate. To do this, we cross-correlate the derivatives of the DM data with the intensity of the signal.

In the simplest case of lensing in a DM signal, we can assume that the signal in spatial scales is composed by a Gaussian function,

$$\Delta\mathrm{DM}(x) = \mathrm{e}^{-\frac{(x-x_0)^2}{2\sigma_{\mathrm{DM}}^2}}. \tag{2}$$

It is given that the intensity of the signal at a point in time is represented by the following equation

$$I(t) = \frac{I_0}{\left|\frac{d^2}{dx^2}\phi(x)\right|} \tag{3}$$

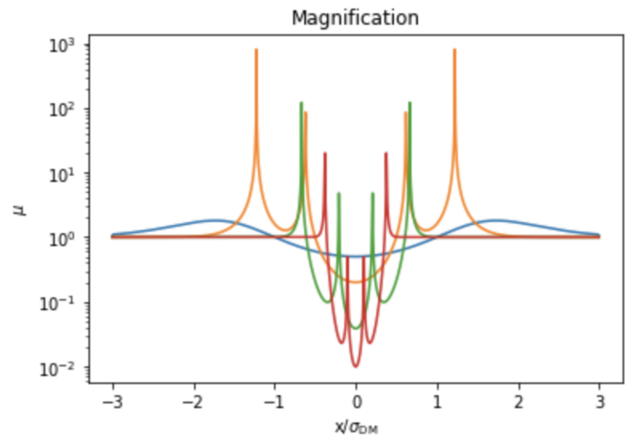

Figure 2: A plot for the magnification as a function of position for each of the curves in Figure 1. Note the spikes correspond to caustics. This shows that there is a strong relationship between the scale and strength of DM fluctuation and the magnification structure.

where ϕ equals,

$$\phi(x) = \frac{x^2}{2} - \Delta\text{DM}(x) \tag{4}$$

(Note that spatial scales can be converted into timescales by knowing the relative velocity between the pulsar and the ISM). Therefore, for an ideal signal we should have

$$\left|1 - \frac{d^2}{dx^2}\Delta\text{DM}(x)\right| = \frac{I_0}{I(x)}. \tag{5}$$

By setting our initial intensity $I_0 = 1$, we have

$$\left|1 - \frac{d^2}{dx^2}\Delta\text{DM}(x)\right| = \frac{1}{I(x)}. \tag{6}$$

We expect that the DM fluctuations will leave an imprint as magnification in the intensity depending on the spatial scale of the DM fluctuation (see Figures 1 and 2). Effectively, this is the relation that we are cross-correlating in our project. Since we have data on the DM signal and intensity, by cross-correlating the two, we are able to figure out whether there is lensing happening over the long timescales. Furthermore, we are able to map out the correlation coefficient as a function of input signal-to-noise ratio (ISNR). By doing so, we would then approximately know what the ISNR is required to have an affirmative finding for lensing.

The cross-correlation coefficient we are calculating is known as the Pearson correlation coefficient, it measures the strength of the relationship and the direction between two variables. In this situation, it is measuring how much of the DM signal and the intensity signal overlap, following equation (5).

3 Simulation

In order to have a better understanding of how well the reconstruction algorithms work and its impact on the recovered DM for predictive signs of lensing, we create a simulated DM signal and analyze it in the same way as a real data analysis. First, we simulate the DM by assuming the signal follows a Gaussian random field $f(k)$ in the frequency domain, k. From there, the probability that $f(k)$ is at a given point q is given by

$$P(f(k) = q) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \mathrm{e}^{-\frac{q^2}{2\sigma_k^2}}. \tag{7}$$

Next, we assume the standard deviation follows a power law, $\sigma_k = k^{-\beta}$, where the spectral index, $\beta = 5/3$. This would imply that the amount of scattering decreases with an increase in k. Having set a visual of the signal in the frequency domain, we can then calculate the ΔDM by performing a Fourier transform of it,

$$\int f(k)\mathrm{e}^{2\pi i x k}\mathrm{d}k. \tag{8}$$

We can keep in mind that the time coordinates of our measurements are proportional to the spatial coordinates for when the ISM passes through the pulsar emission, *e.g.* $x = vt$, where v is the relative velocity of the pulsar to the ISM and t is the pulse time.

Having created the simulated ground truth DM signal, we then replicate measurements taken by a telescope by adding noise into the ground truth signal. The amount of noise added is determined by the ISNR that is set for the simulation following the relation

$$\sigma = \|\mathbf{x_{true}}\| \frac{10^{-\frac{ISNR}{20}}}{\sqrt{2M}}, \tag{9}$$

where $\mathbf{x_{true}}$ is the simulated ground truth vector and M is the number of measurements (the dimension of $\mathbf{x_{true}}$). The signal that we perform reconstruction on is effectively these measurement data. We compare the simulated DM signals with different sampling ratios. In each sampling ratio, we examine how close the reconstructed DM signal is with the ground truth signal as well as the simulated measurement samples. We see that with increasing sampling ratios, the reconstruction becomes closer to the ground truth DM signal. For further understanding, see Figure 3, 4, and 5 for examples of reconstructing DM signals when the model is under sampling, match sampling and oversampling respectively.

After simulating the DM signal, we move on to simulating the lensed intensity signal give by the relation,

$$I = \mu I_0 \tag{10}$$

where I_0 is the initial intensity of the signal and

$$\mu = \frac{1}{1 - \frac{\lambda r_e R_{\mathrm{F}}^2}{2\pi}\partial_x^2 \Delta\mathrm{DM}}, \tag{11}$$

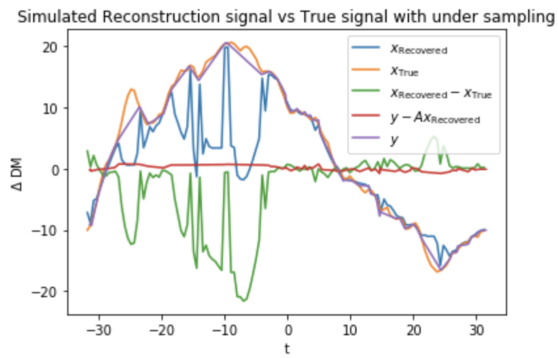

Figure 3: A plot comparing the ground truth, measurement and reconstructed signal from under sampling, along with the residuals of the reconstruction. Due to the lack of sample points, the reconstructed signal is unable to mimic the ground truth signal. We assumed arbitrary units for the DM signal and time. For improved reconstruction models, see Figures 4 and 5.

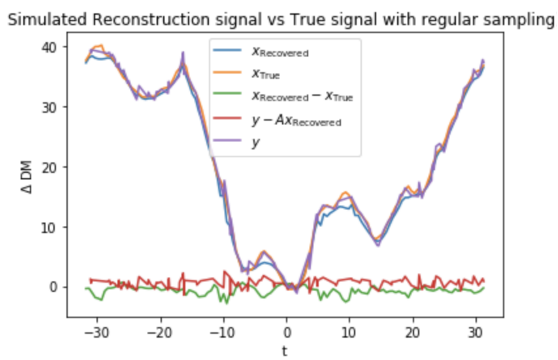

Figure 4: A plot comparing the ground truth, measurement and reconstructed signal with regular match sampling, along with the residuals of the reconstruction. By increasing the sampling ratio to match the grid size used, the reconstructed model is able to further replicate the ground truth signal. For better reconstruction model, see Figure 5.

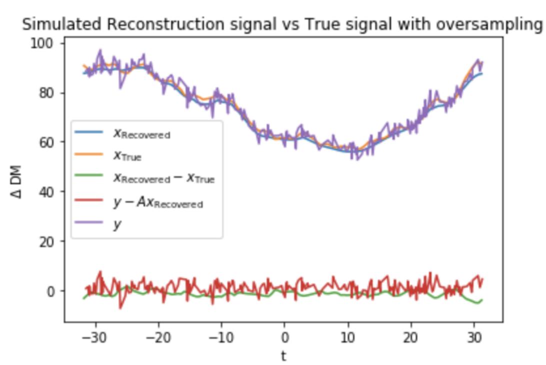

Figure 5: A plot comparing the ground truth, measurement and reconstructed signal with oversampling, along with the residuals of the reconstruction. By oversampling the measurements, the reconstructed model is at its best in reproducing the ground truth signal relative to the different sampling ratios. This shows that the higher the sampling ratio, the greater the reconstruction.

is the intensity magnification due to lensing. Here, λ is the wavelength of the signal, r_e is the classical electron radius and R_{F} is the Fresnel scale [4]. For simplicity, we set the initial intensity of the signal, $I_0 = 1$, which reduces the intensity relation to $I = \mu$. Looking at μ we can see that the second derivative of the DM (with respect to spatial coordinates) is related to the amount of lensing there is as the ISM passes by the pulsar emissions. For the purposes of the simulation, we chose the factor next to ΔDM in μ to be 1.

Finally, we want to see if the DM and the intensity signal are correlated. To do this, we define the variables

$$\eta = \frac{1}{I} \quad \text{and} \quad \kappa = \frac{1}{\mu} \tag{12}$$

to see if the inverse of the intensity and the second derivative of the DM are correlated. This is essentially the relation we have in the previous section with an extra group factor in front of the DM term. We perform this comparison using time sample measurements and therefore we need to standardize our measurements first. For a sample point j in the measurements, we have

$$\bar{\eta_j} = \frac{\eta_j - m_\eta}{\sigma_\eta} \tag{13}$$

where the mean is defined as

$$m_\eta = \frac{1}{N} \sum_{j=1}^{N} \eta_j \tag{14}$$

and the standard deviation is

$$\sigma_\eta = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (\eta_j - m_j)^2}. \tag{15}$$

We can keep in mind that κ follows a similar standardization in the measurements. We then calculate the cross-correlation coefficient as

$$\rho_{\bar{\eta}\bar{\kappa}} = \sum \bar{\eta_j} \bar{\kappa_j}. \tag{16}$$

In the case of simulations, we find that if there are no noise and regular sampling, the cross-correlation coefficient should be 1, since the intensity of the signal is proportional to μ. See Figure 6 for examples of the reconstructed DM fluctuations and intensity fluctuations and the corresponding standardized signals for both the ground truth and reconstructed signals.

However, because not all of the sample points DM_j and I_j overlap throughout the measurements, along with the addition of noise and gaps due to missing sample points, it is not guaranteed to have a cross-correlation coefficient of 1, even in simulations. One of the main goals of this project is to remove the measurement noise and fill in the gaps of the missing sample points in DM_j and

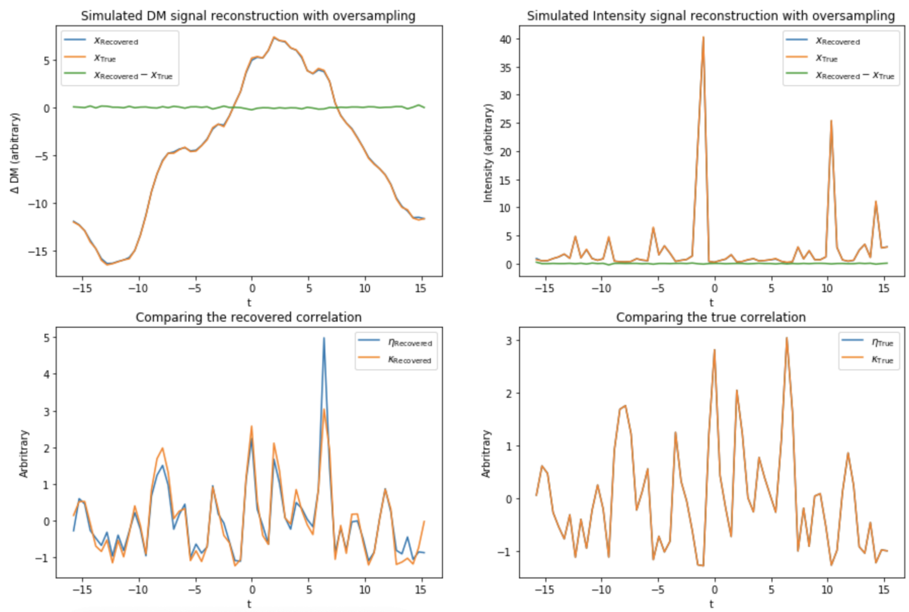

Figure 6: These plots are a comparison of the reconstructed models with the ground truth signal for a DM and intensity signal as well as the cross-correlation between the DM and intensity signal for reconstructed signals versus the ground truth signals. Evidently, the reconstruction algorithms are capable of reproducing a signal that closely resembles the ground truth with a sufficient sampling ratio. We assumed arbitrary units for the elapsed time.

I_j using the reconstruction algorithms. In doing so, it may help increase the cross-correlation, $\rho_{\bar{\eta}\bar{\kappa}}$, where it was lost due to missing time samples or noise in the second derivative of the DM signal. See Figure 7 for an example of how the ISNR after oversampling impacts the cross-correlation. We find that by oversampling the reconstructed model, we expect the cross correlation to be stronger for lower noise levels.

4 Real Data

For the real data sets, we performed the same analysis as the simulated one. We applied the reconstruction algorithms to the measurements in order to better reduce the noise and fill in any missing sample points. We then compared the DM and intensity data through a cross-correlation in order to find possible signs of lensing events. The data used for this analysis is from the pulsar, J0358+5413. The DM and intensity measurements were taken with the Low-Frequency Array (LOFAR) radio telescope located in the Netherlands. The measurements were taken at a frequency of 149 MHz with 95 MHz bandwidth. The DM data was taken from March 9, 2014 to January 7, 2020, while the intensity data was taken from July 5, 2014 to April 15, 2019. Since some of the data span over different periods, we cropped the two signals and only used the data that overlapped in

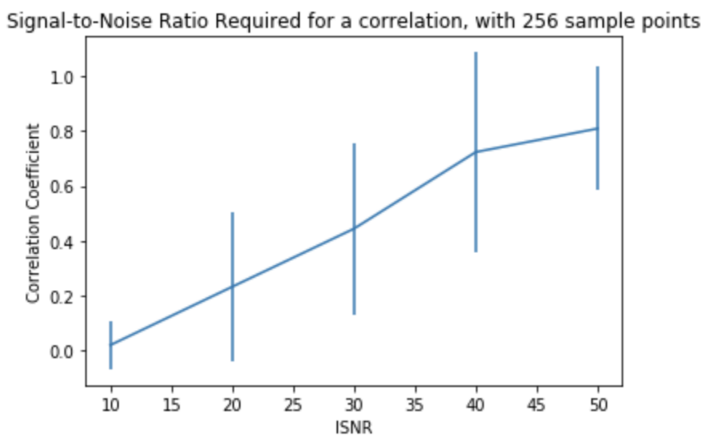

Figure 7: A plot relating the relative ISNR required for an approximate correlation value. As the plot shows, there is a positive relation between the amount of ISNR from a signal to the amount of correlation there is between a DM and intensity signal. This means that with better signal, the reconstruction model is able to further tighten the discrepancy between the DM signal and intensity signal.

time. The data processing is performed by Julian Donner at the Max Planck Institute for Radio Astronomy (MPIfR). Further details on the data reduction and processing are outlined in this paper, Donner et al. (2019) [2].

In Figure 8, we compare the measured and reconstructed signal for the DM and intensity data from the pulsar. We can see that the measured signal is contaminated by noise and missing sample points, however after reconstruction we are able to fill in those gaps and missing data points. In the process, we are taking into account the absolute uncertainties of the data, the weights are proportional to $\frac{1}{uncertainty}$, while the closeness of the fit is determined by the root-mean-square (rms) of the uncertainty.

After identifying η and κ in the intensity and DM data respectively, we standardized the measurements and cross-correlated them. We plotted the entire signal in Figure 9 and calculated the Pearson correlation coefficient. It turns out, for the entire signal the correlation coefficient, $\rho_{\bar{\eta}\bar{\kappa}} \cong 0.35$. However, we noticed that certain areas of the plot seem to correlate more than others and therefore we attempted to find a time period with the best correlation. In Figure 10, we constrained the correlation to the area that we found to have the best correlation coefficient value. We find the value in this region to be, $\rho_{\bar{\eta}\bar{\kappa}} \cong 0.63$. Undoubtedly, we see that there is a correlation between the DM signal and the intensity signal with certain regions being stronger. Evidently, this shows that there are potential lensing events happening over long periods of time that we may be unaware of.

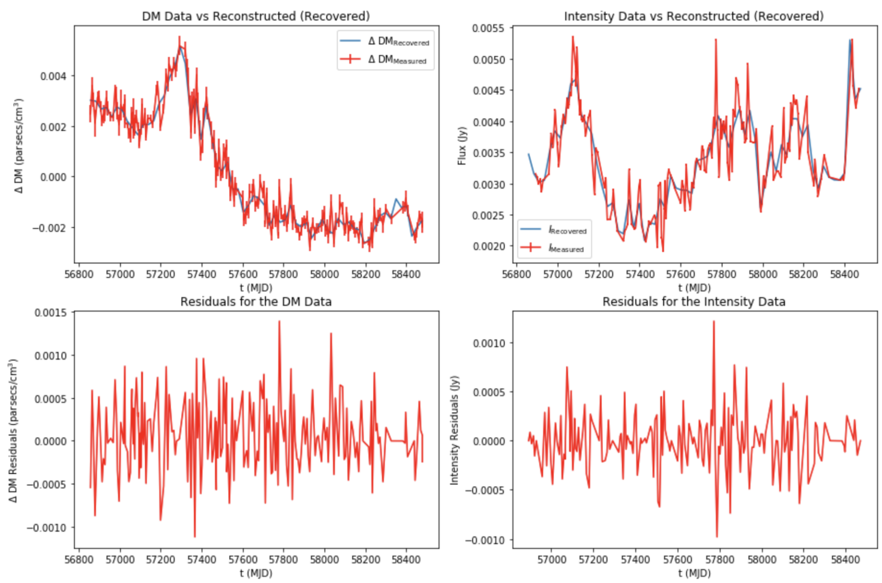

Figure 8: These plots show the DM and intensity measurements taken from the pulsar, J0358+5413 along with the reconstructed model and the residuals associated. In general, the reconstructed model follows closely with the measurements. In the residuals, there does not seem to be any obvious signs of long timescale structures left out. Most of the fluctuations are more likely caused by short timescale, and thus small spatial scale structures.

5 Discussion/Conclusion

Starting with the simulations, we find that even if the DM and intensity signals are simulated, a perfect correlation is not guaranteed. Noise as well as sampling ratios play a vital role in determining the correlation between the two signals. This in turn affects the findings of long scale lensing events in the ISM as it passes by pulsar emissions. Up-to-date, there has not been well known studies that have attempted to measure lensing events due to ISM on large timescales, mostly due to missing sampling points. However, from the simulations we see that it is possible to fill in the missing sampling points with the algorithms provided by my supervisor, Dr. Luke Pratley. This led to the recovery of the real data set of the pulsar, J0358+5413, we find that the signals correlate on large timescales with certain regions correlating better than others, This project concludes that the reconstruction algorithms can help recover missing sample points in the data set as well as reducing noise. There is also potential for lensing events to occur in the ISM over long timescales. However, further research is needed in order to better understand the mechanism behind these events.

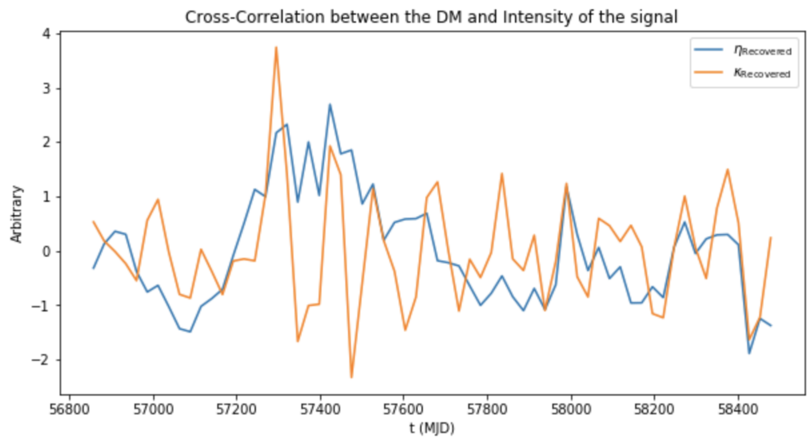

Figure 9: A cross-correlation between the DM and intensity signal (defined as κ and η respectively) after standardization, measured from the pulsar J0358+5413. It is evident that some areas are more correlated than others. The correlation coefficient is calculated to be approximately, $\rho_{\bar{\eta}\bar{\kappa}} \cong 0.35$.

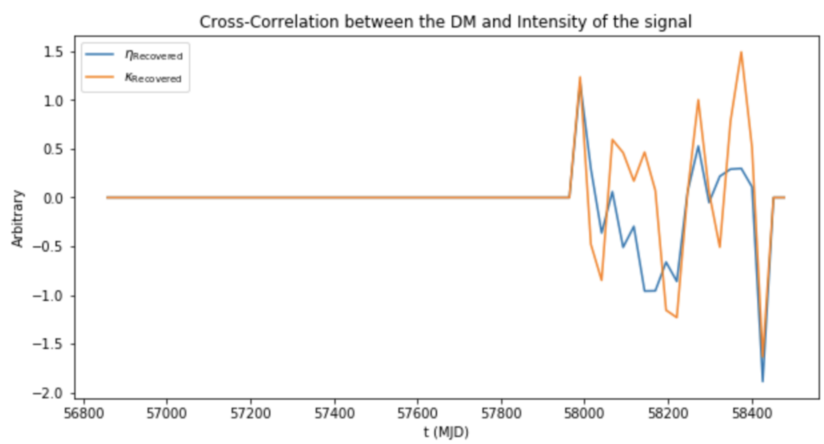

Figure 10: The cross-correlation between the DM and intensity signal of a specific time frame with higher correlation coefficient. By constraining the plot to a region of higher correlation, we found a further increase in the correlation coefficient value, $\rho_{\bar{\eta}\bar{\kappa}} \cong 0.63$.

References

[1] J. M. Cordes, R. M. Shannon, and D. R. Stinebring. "Frequency-dependent dispersion measures and implications for pulsar timing". In: *The Astrophysical Journal* 817(1).16 (2016).

[2] J. Y. Donner et al. "First detection of frequency-dependentm time-variable dispersion measures". In: *Astronomy & Astrophysics* 624 (2019).

[3] M. L. Jones et al. "THE NANOGRAV NINE-YEAR DATASET: MEASUREMENT AND INTERPRETATION OF VARIATIONS IN DISPERSION MEASURES". In: *The Astrophysical Journal* 841 (2018).

[4] F. X. Lin et al. "Discovery and modelling of broad-scale plasma lensing in black-widow pulsar J2051-0827". In: *Submitted to Monthly Notices of the Royal Astronomical Society* (2020).

[5] R. Main et al. "Pulsar emission amplified and resolved by plasma lensing in an ecplipsing binary". In: *Nature* (2018).