

Application of wavelets for the detection of the Global 21cm signal from the Epoch of Reionisation

Akshatha K V

R V College of Engineering, Mysore Road, Bengaluru- 560059

Prof. Udaya Shankar N

Astronomy & Astrophysics, Raman Research Institute, Bengaluru- 560080

Abstract

According to the standard model of Cosmology, the expansion of the universe after the big bang resulted in its cooling and the drop in temperature allowed combination of neutrons and protons leading to the formation of atomic nuclei. The initial inhomogeneity of matter density resulted in the formation of the first stars in the universe, after 400 million years from the big bang. The radiation and heat from the first stars started ionising the matter, marking the beginning of the Epoch of Reionisation (EoR). The most promising approach so far for observing this early epoch of the universe is the 21cm hyperfine transition line of neutral hydrogen, red-shifted in the observer's frame to the frequencies in the range of 40-200 MHz. Recently an absorption profile centred at 78 MHz in the sky averaged spectrum was reported by researchers at ASU (Arizona State University, in their Experiment to Detect the Global Epoch of Reionisation Signature, called EDGES Bowman, et al, 2018). This absorption spectrum with a width of ~18 MHz and an amplitude of 530 mK is buried in foreground of the order of 2000K. We are exploring Wavelet transform technique to find the domain which gives the best signal to noise ratio of this absorption profile to its foreground. Initially, we studied the behavior of the signal to the Fourier transform (FT) technique. Then we explored the Continuous Wavelet Transform (CWT) to find a suitable technique for feature extraction in

spectral analysis and signal reconstruction. The Morlet, Mexican Hat (Ricker) and Symlet wavelets were explored in CWT. We also explored discrete wavelets like Daubechies and Haar. It was found that the absorption profile is very broad, making the available baseline in the foreground estimation too small for applying linear transformation techniques. In view of this, we explored the linear transformation technique by compressing the absorption profile in the frequency domain. We studied the efficacies of using FT, CWT and DWT techniques for extracting the Global 21cm feature in the presence of the foreground signal of sky averaged radio spectrum. We also implemented DWT and reconstructed the signal on GPU (GeForce GTX 1050) using PyCuda libraries.

Keywords or phrases: Absorption profile, Wavelet, Continuous Wavelet Transform, Discrete Wavelet Transform, Graphics Processing unit

Abbreviations

EoR	Epoch of Reionisation
CMBR	Cosmic Microwave Background Radiation
FFT	Fast Fourier Transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
EDGES	Experiment to Detect the Global EoR Signature
GPGPU	General-purpose Graphics Processing Unit
CUDA	Compute Unified Device Architecture
SNR	Signal to Noise Ratio
RFI	Radio Frequency Interference

1 INTRODUCTION

1.1 Background/Rationale

The Universe is believed to have begun in an extremely hot and dense state, followed by a sudden expansion after about 10^{-32} sec, and subsequent slow cooling and expansion. This aided for a gradual drop in temperature and density of the Universe, facilitating for the combination of neutrons and protons to form the atomic nuclei. After a few hundred thousand years, when the Universe was further cooled, these atomic nuclei captured the electrons to form singly ionized Helium, neutral Helium and neutral Hydrogen Ganjam, 2018 . This phase of the universe, widely known as Epoch of Recombination contributed to the Cosmic Microwave Background Radiation. When the first stars were formed, the heat and radiation ionized the neutral matter leading to the formation of bubbles of plasma surrounding the source *MIT Haystack Observatory: EDGES* . As the bubbles grew bigger, they started to overlap, leading to the formation of much larger bubbles, accelerating the ionisation process. This era of Universe is termed as Epoch of Reionisation. The most powerful probe to study the EoR signal so far is the redshifted Global 21-cm spectral line resulting from the spin-flip hyperfine transition of the hydrogen atom, red-shifted in the observer's frame to the frequencies in the range of 40-200 MHz Bowman, et al, 2018 .

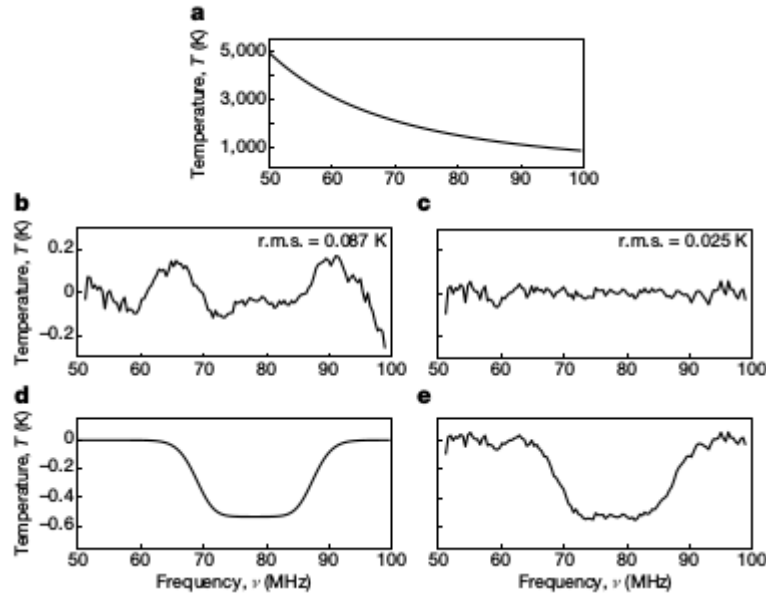


Fig 1 **a**, Measured spectrum for the referencedataset after filtering for data quality and radio-frequency interference. The spectrum is dominated by Galactic synchrotron emission. **b**, **c**, Residuals after fitting and removing only the foreground model (**b**) or the foreground and 21-cm models (**c**). **d**, Recovered model profile of the 21-cm absorption, with a signal-to-noise ratio of 37, amplitude of 0.53 K, centre frequency of 78.1 MHz and width of 18.7 MHz. **e**, Sum of the 21-cm model (**d**) and its residuals (**c**) Bowman, et al, 2018.

The 21-cm signal from EoR can be studied as fluctuations of temperature in the global signal varying as a function of frequency and can be seen as the absorption profile in the sky-averaged radio spectrum. Recently, an absorption profile centered at 78 MHz in the sky-averaged spectrum was reported by researchers at ASU (Arizona State University), in their Experiment to Detect the Global Epoch of Reionisation Signature, called EDGES Bowman, et al, 2018. This absorption spectrum has a width of ~ 18 MHz and an amplitude of 530 mK and is buried in the foreground of the order of 2000K. The reported profile was consistent with the expectations of the 21-centimeter signal induced by early stars, except that the amplitude of the profile was twice the predicted amplitude. This large variation in observed and predicted profile suggests that either the primordial gas was much colder than expected or the background radiation temperature was hotter than expected Bowman, et al, 2018.

In this summer project, an attempt is made to analyze the spectral data using Wavelet transform technique to find a contrast between the temperature dip, and

the foreground. As a second phase of the project, the wavelet transform technique was implemented on the GPU using PyCuda to explore and appreciate the properties and advantages of GPU.

1.2 Problem Statement

The EoR signal, seen as the absorption profile is of the order of mK, while the foreground of Global 21cm sky-averaged radio spectrum is of the order of 2000k. This fluctuation of mK is hidden in the foreground, making it challenging to detect. This signal is predicted to be extremely weak, spread over a wide frequency range from 40-200MHz and is buried in the foreground having cumulative effect of Galactic and Extragalactic foregrounds as well as Radio Frequency Interference Singh, et al, 2017. In order to extract this weak signal from enormous foreground, it is paramount to find the contrast between the temperature dip and the foreground itself. In this project, an attempt is made to find a domain which gives the highest signal to noise ratio. Since wavelet transform technique involves localisation in the time-frequency domain, this technique is explored to analyse the possible ways of extracting the Bowman dip from the foreground data.

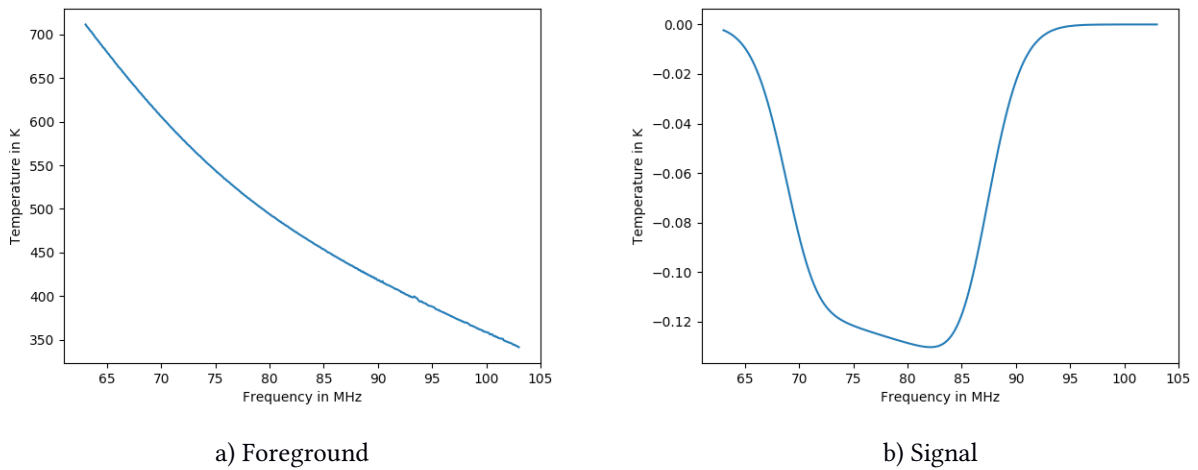


Fig 2 Comparision of magnitude of Foreground and EoR signal

For this project, the foreground, as measured by the instrument, data ranging from 63MHz to 103MHz was considered.

1.3 Objectives of the Research

- To study the Fourier and Wavelet transform of the EoR signal and the foreground and to verify if they can be used as a tool for extracting the signal from the foreground
- To attempt to find domain of highest signal-to-noise ratio for foreground and the EoR signal
- To understand and analyse both continuous and discrete wavelet transform technique for the case of foreground data and the temperature dip.
- To understand and appreciate the scientific computational capabilities of the GPU and study it for the case of spectral data.

2 LITERATURE REVIEW

Mathematically, wavelet transform of a signal $s(t)$ with respect to a mother wavelet $g(t)$ is,

$$S(\tau, a) = \frac{1}{2\pi} \sqrt{a} \int S(\omega) G^*(a\omega) e^{j\omega\tau} d\omega \quad \text{Equation (i)}$$

where $S(\omega)$ is a Fourier transform of the signal, $a > 0$ is a dilation parameter, τ is a dilation parameter and $G^*(\omega)$ is a complex conjugate of the Fourier transform of $g(t)$ Suvichakorn, et al, 2008.

In the discrete wavelet transform, the scale parameter is always discretized to integer powers of 2, 2^j , $j=1,2,3,\dots$, so that the number of voices per octave is always 1. The DWT provides a sparse representation for many natural signals. In other words, the important features of many natural signals are captured by a subset of DWT coefficients that is typically much smaller than the original signal. This *compresses* the signal. Thus, DWT always gives the same number of coefficients as the original signal, but many of the coefficients may be close to zero in value. As a result, it is often practical to throw away those coefficients and still maintain a high-

quality signal approximation Vijendra, 2016. This forms the basis of signal compression.

In feature extraction of Heart rhythms, wavelet transform was used for noise removal and peak detection. The major source of noise, in this case, were muscle contraction and movement of the patient while recording ECG. The Heart Rate Variability, HRV signals were obtained after pre-processing, which were then analysed in time and frequency domains. Wavelets have been utilized and performed well for signal processing of ECG signals and then the HRV signal is generated and analyzed on the basis of features extracted Gautam, et al, 2017.

Wavelet transform technique has also been used widely in feature extraction in signal processing. Sea surface elevation and ocean wave spectrum, both significant in marine science research and ocean engineering were calculated using the wavelet transform algorithm. The wave scalogram obtained from the data collected by ultrasonic wave gages on the Cigu pile station from the Taiwan Strait was used to analyse instant nonlinear features at short-time duration Wu, et al, 2010.

DWT has been effectively used in feature extraction of ultrasonic flaw data. The signal was decomposed into detail and approximation coefficients using successive high pass and low pass filters respectively. The approximation coefficients obtained from level-1 of DWT operation was subjected to level-2 DWT operation giving sets of approximate and detail coefficients of level-1 coefficients. The authors considered four levels of decomposition for various feature extraction such as Mean, Variance, Mean of Energy, Maximum and Minimum amplitude among others Shodhganga.

DWT is a non-redundant transform. It was developed so there would be a one to one correspondence between the information in the signal domain and the transform domain. This tight correspondence in the DWT allows signal reconstruction Jorgensen, et al, 2009 .

As a second phase of the project, discrete wavelet transform was implemented on a GPU. A graphics processing unit (GPU) is similar to a computer's CPU but is designed specifically for performing the complex mathematical and geometric calculations that are necessary for graphics rendering. However, it is interesting to note that GPUs are widely used in Scientific Computations. Some of the fastest GPUs have more transistors than the average CPU. GPUs are highly parallel and architecture sensitive Andreas Klockner, 2009 and are capable of handling as many threads as 10 times number of cores Vincent Hindriksen, 2017. For a computational activity in GPU, Host to Device and Device to Host memory transfer is essential. For this, GPU receives strings of data and instructions through PCI-Express from CPU.

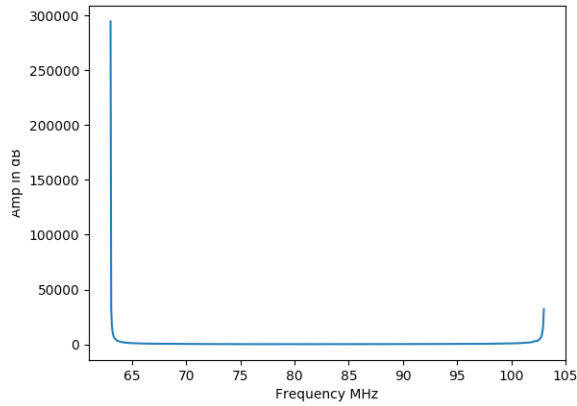
3 METHODOLOGY

3.1 Fourier Transforms

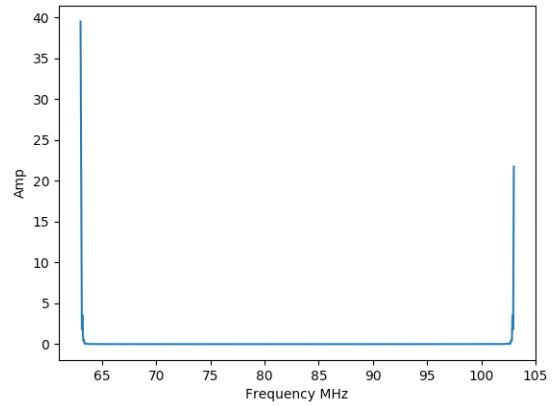
As an initial exercise, the foreground and the signal were analysed in Fourier domain. A real-valued function or the analytic signal gives a Hermitian symmetric signal or complex-valued function in Fourier domain. Since the foreground and the signal both are analytic signals, a complex-valued function is obtained in the Fourier domain. The amplitude is thus calculated as

$$Amplitude = \sqrt{|Real Part|^2 + |Imaginary Part|^2} \quad \text{Equation (ii)}$$

and is plotted as a function of frequency.



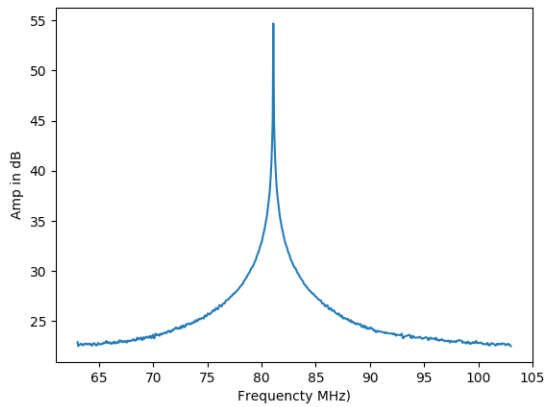
a) FFT of Foreground



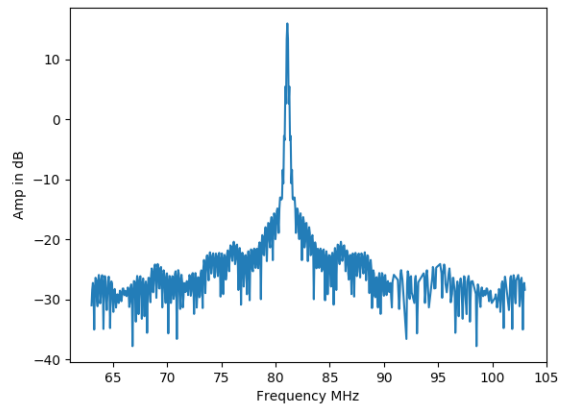
b) FFT of Dip

Fig 3 Foreground and Signal in fourier Domain

It can be clearly seen that the FFT of foreground and Dip vary on a wide dynamic range. Therefore, the FFT results were plotted on a log scale. Since FFT of both foreground and the dip is Hermitian symmetric, the results were shifted to the center of the band.



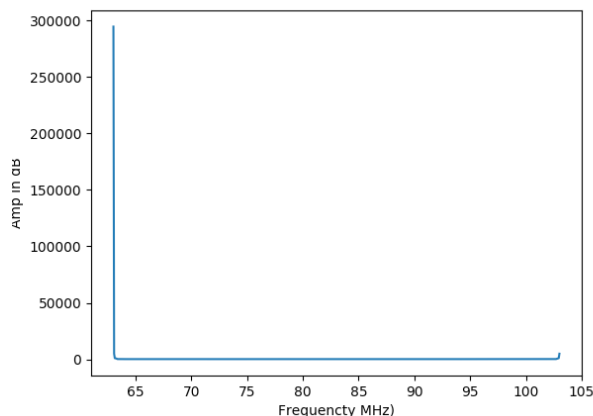
a) Shifted FFT of foreground on log scale



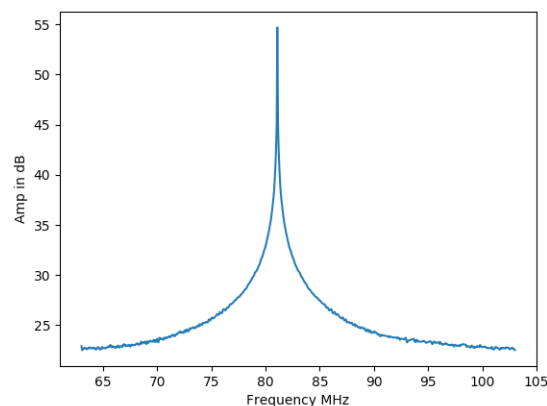
b) Shifted FFT of Signal on log scale

Fig 4 Shifted FFT of foreground and Signal on log scale

A similar approach was followed for the sum of foreground and signal. FFT was computed, shifted to the center and was plotted on a log scale.



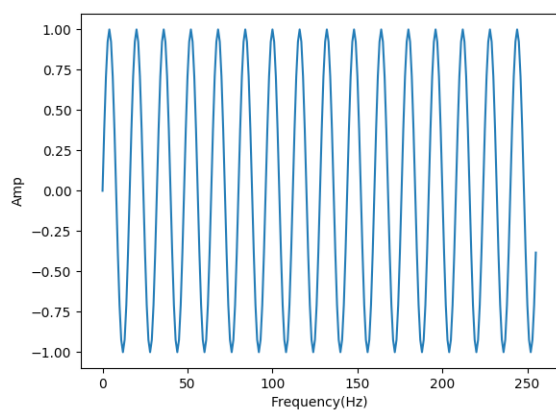
a) FFT of Sum of foreground and signal



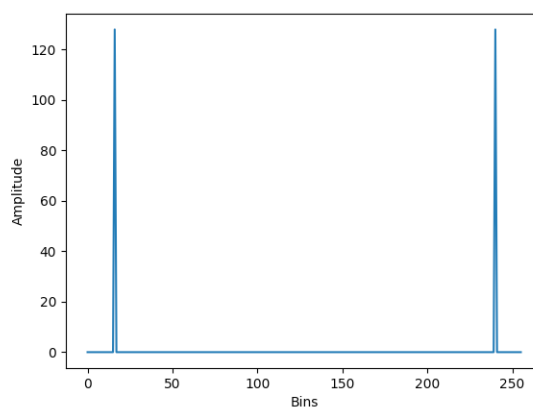
b) Shifted FFT of sum of foreground and signal on log scale

Fig 5 FFT of sum of foreground and signal on linear and log scale

The purpose is to find the domain with best signal to noise ratio. For example, FFT of sine wave is a delta peak in frequency domain. Now, consider a noisy sine wave. Extracting the information in presence of noise in time domain is difficult. Taking Fourier transform of noisy sine gives two peaks with a small Gaussian noise. The sine wave is distinctively seen in Fourier domain.



a) Pure Sine wave



b) FFT of Pure Sine wave

Fig 6 Sine wave and its Fourier transform

Gaussian white noise (random noise at all frequencies) with mean=0 and standard deviation=1 was added to the above sine wave.

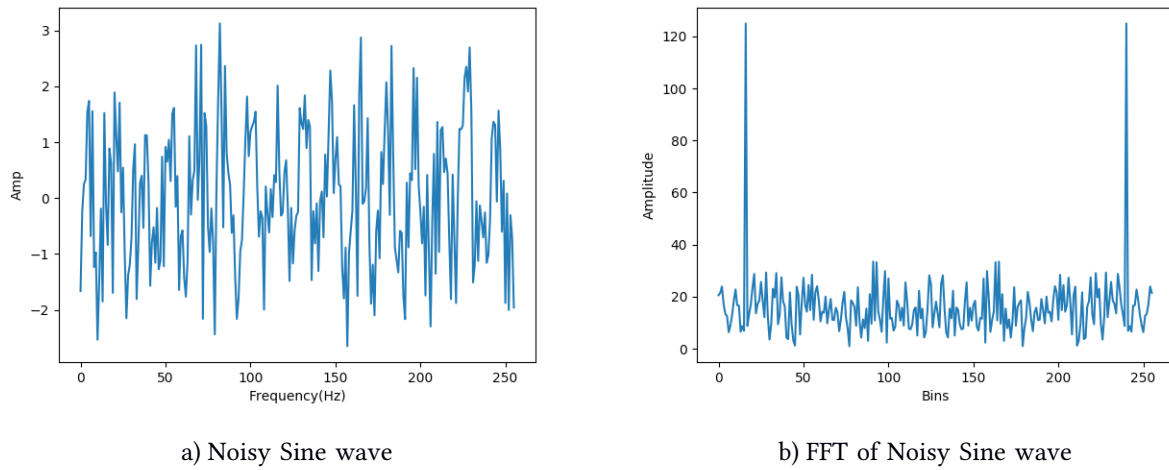


Fig 7 Noisy sine wave and its Fourier transform.

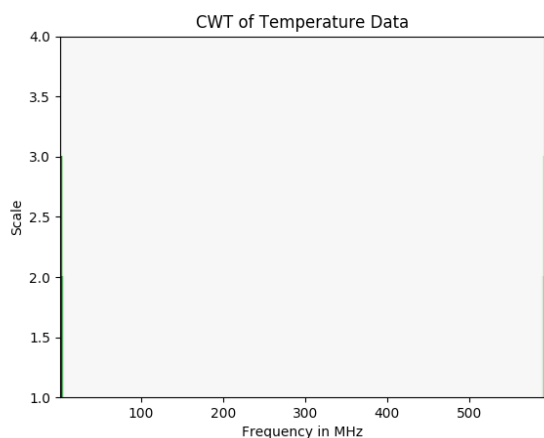
In case of Fourier transform of a noisy sine wave, the basis functions are sine and cosine functions, orthogonal to the sine wave. Therefore when FFT was computed, characteristic delta functions were obtained. Noisy sine in time domain is indistinguishable, due to low signal to noise ratio. Transforming it to frequency domain gives characteristic delta peaks with amplitude of ~ 128 (for 256 point FFT) and noise of amplitude 1, making SNR very high. From Parsival's theorem, power in time and frequency domain is identical. Thus a domain which gives highest signal to noise ratio is preferred.

3.2 Continuous Wavelet Transforms

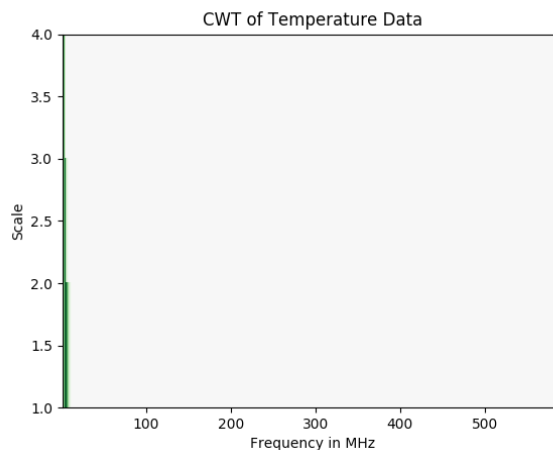
In order to study the foreground and the temperature dip using wavelet transform technique, a suitable mother wavelet is to be selected. Results of two important wavelets, the Morlet and the Mexican hat (or Ricker) wavelet were compared. Morlet Wavelet has been used in signals containing Lorentzian function, described by full width half maximum and Voigt function (obtained by convolution of two broadening mechanisms, one of which would produce gaussian profile) Suvichakorn, et al, 2008.

3.2.1 Comparision of Morlet and Ricker Wavelets

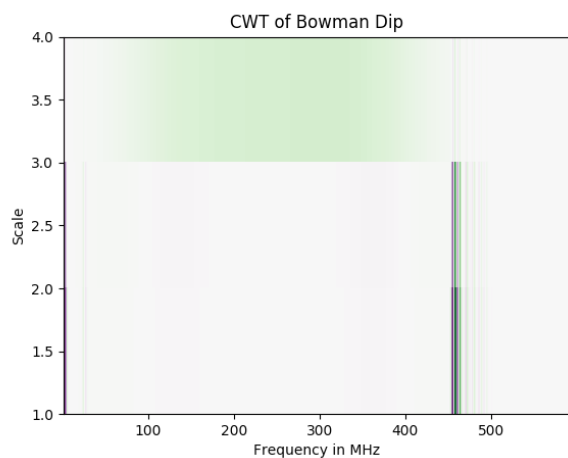
The wavelet transform was computed using SciPy library in Python. Wavelet transform of the temperature data, Bowman dip, Sum of temperature data and dip, difference of CWT coefficients of sum of dip and foreground & the foreground was computed. Each of these results were compared for both the wavelets.



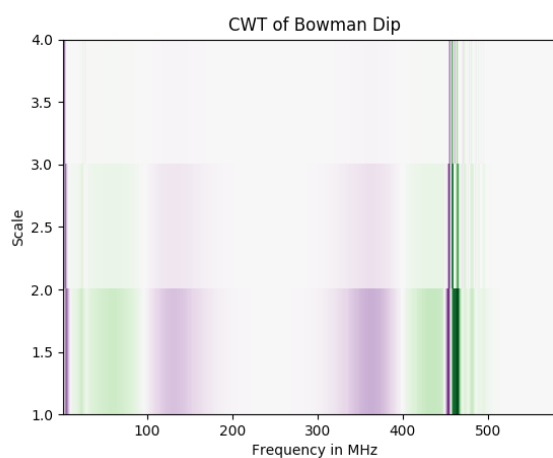
a) CWT of Foreground using Morlet



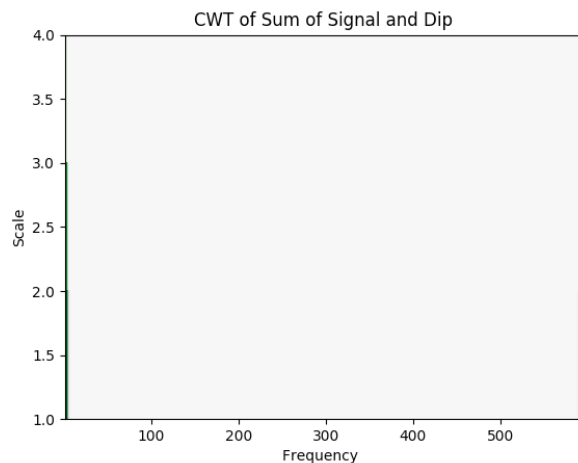
b) CWT of Foreground using Ricker



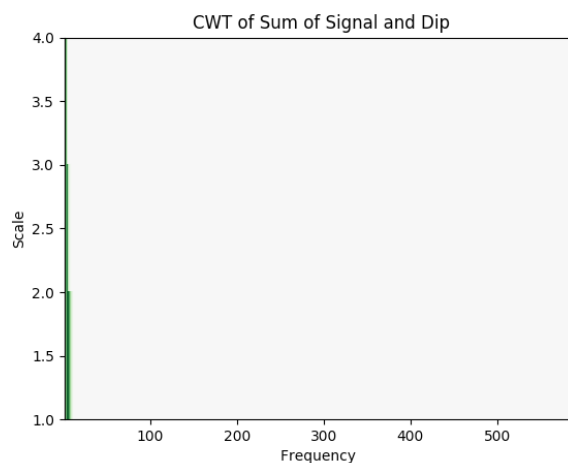
c) CWT of Signal using Morlet



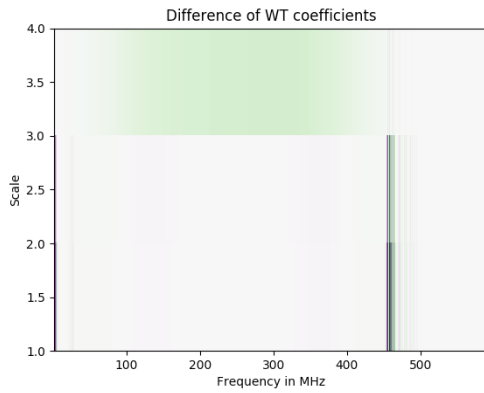
d) CWT of Signal using Ricker



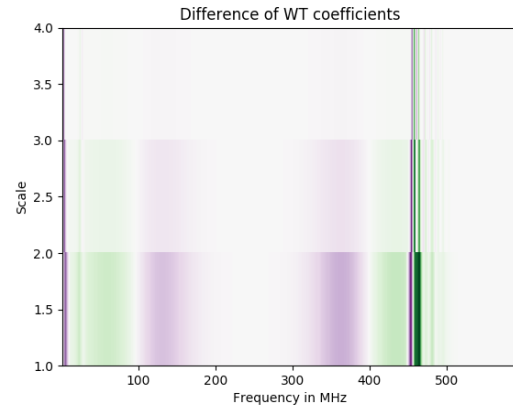
e) CWT of sum using Morlet



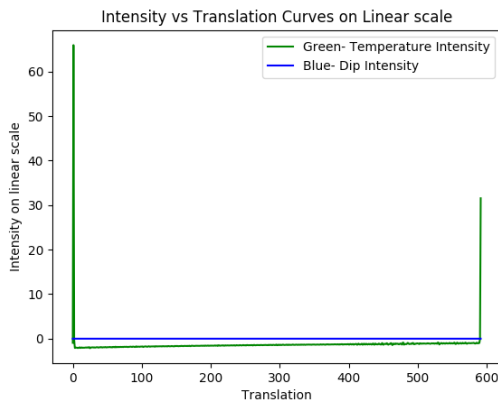
f) CWT of sum using Ricker



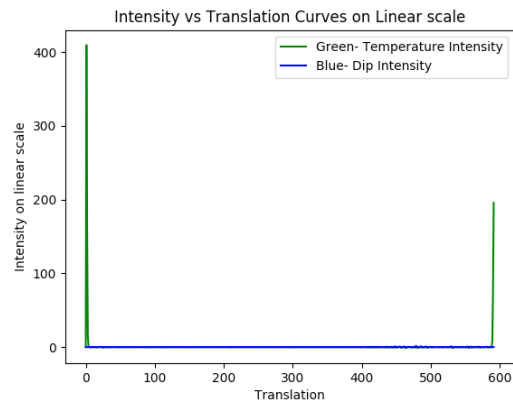
g) Difference in coefficients foreground+signal & foreground for the case of Morlet wavelet



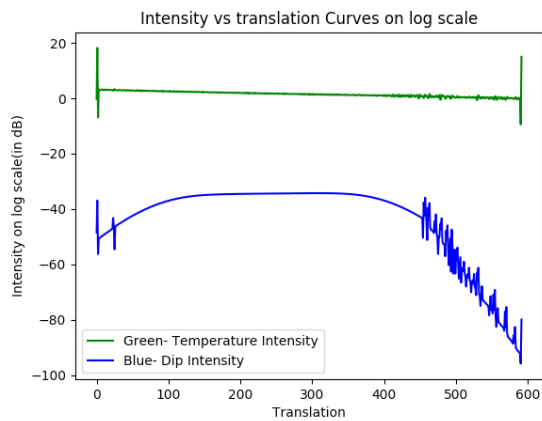
h) Difference in coefficients foreground+signal & foreground for the case of Ricker wavelet



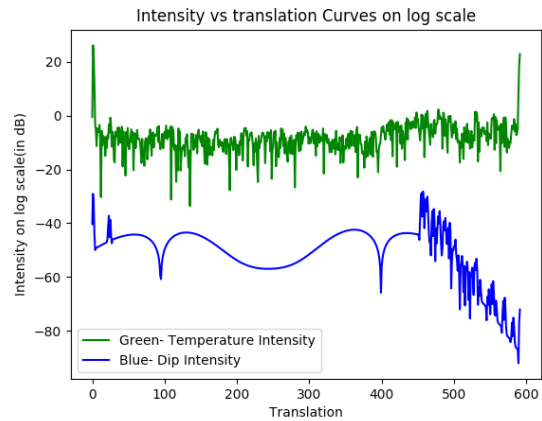
i) Scale 1 Intensity curve of Signal using Morlet wavelet



j) Scale 1 Intensity curve of Signal using Ricker wavelet



k) Scale 1 Intensity curve of Signal on log scale using Morlet wavelet



l) Scale 1 Intensity curve of Signal on log scale using Ricker wavelet

Fig 8 Comparison of Morlet and Ricker wavelets

In the scalogram, *Fig 8(a to h)*, values on X-axis correspond to the 592 data points(ranging from 63-103MHz), while that on Y-axis corresponds to scales ranging from 1 to 4. For the CWT of foreground using Morlet and Ricker wavelets, there is visible intensity on the scalogram at higher and lower frequency, while there are no significant values in the mid- frequency range.

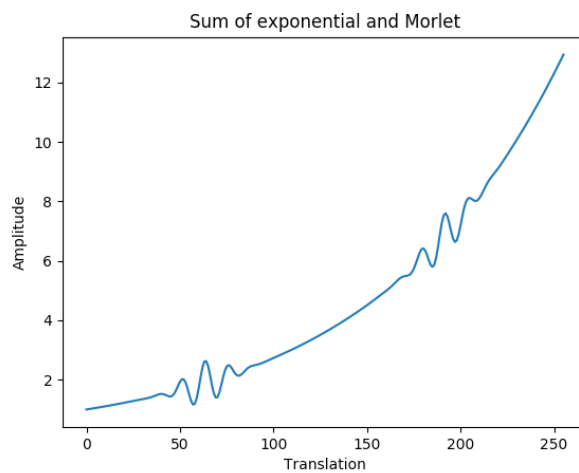
From the above comparison table, it can be observed that, the wavelet transform coefficient for Ricker wavelet in the translation range of 0 to 500 ranges between - 0.00527 and $1.51e-09$ with the minimum of $1.51e-09$ at (591,2) and maximum of 0.00234 at (464, 1.16) , while for the Morlet wavelet, it ranges between -0.000126 and $-4e-06$ with the minimum of $-2.45e-07$ at (206, 2.16) and maximum of 0.00162 ,at (457, 1.6). It can be seen from corresponding intensity plots of scale 1 that the *Ricker wavelet gives better results for the given data*. Thus, for the rest of the analysis, Ricker wavelet was considered.

Close observation of *Fig 2 and Fig 8(l)* shows that

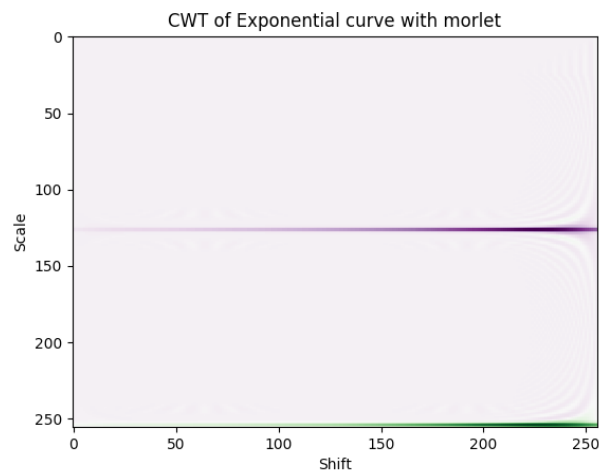
1. the temperature dip and the foreground are in the same frequency range (dip is not sharp)
2. the foreground is not smooth and has few RFIs
3. there is an edge effect on either side of the CWT coefficients.

In order to overcome the above problems, either the foreground data set is to be extrapolated over a wide range of frequency or the dip is to be compressed. In the extrapolation method, if the polynomial is of a lower order, there would be a larger error between actual foreground and the polynomial. However, if the polynomial is of the higher order, it would have higher degrees of freedom and the polynomial tends to deviate from the foreground. Thus, the idea of extrapolating the foreground was dropped. The latter method did not appreciably reduce the edge effect. However, higher order (higher scale) CWT gave lesser edge effects (*Fig10*).

It is well understood that, if a particular frequency data in a signal is to be found, the wavelet should look like the signal we are looking for, and that the basis of the transform must be orthogonal. For everything else but the identical signal, the output will be zero Lathi, 2009. This was verified by adding a Morlet wavelet to an exponential function and finding morlet wavelet transform of the data over a large scale ranging from 1 to 256. A clear intensity peak was observed in the scalogram of the CWT coefficients.



a) Morlet wavelet added to an exponential signal



b) Scalogram of the CWT coefficients

Fig 9 Feature extraction using Morlet wavelet

Wavelet transform technique is based on the match filtering concept. The idea of match filtering worked well when morlet function was to be found from an exponential signal, where the feature to be extracted was of the order of the signal itself. The scale variable provides analysis in the frequency domain: a compressed version of the wavelet function corresponds to the high-frequency components of the original signal while the stretched version corresponds to the low-frequency components; and the translation variable provides analysis in the time domain R. E. J. Yohanes, et al, 2012.

The scalogram gives 3-dimensional intensity plot of Wavelet transform coefficients. To study these coefficients, intensity vs translation of CWT of Foreground and the Signal was plotted corresponding to each scale ranging from 1 to 6.

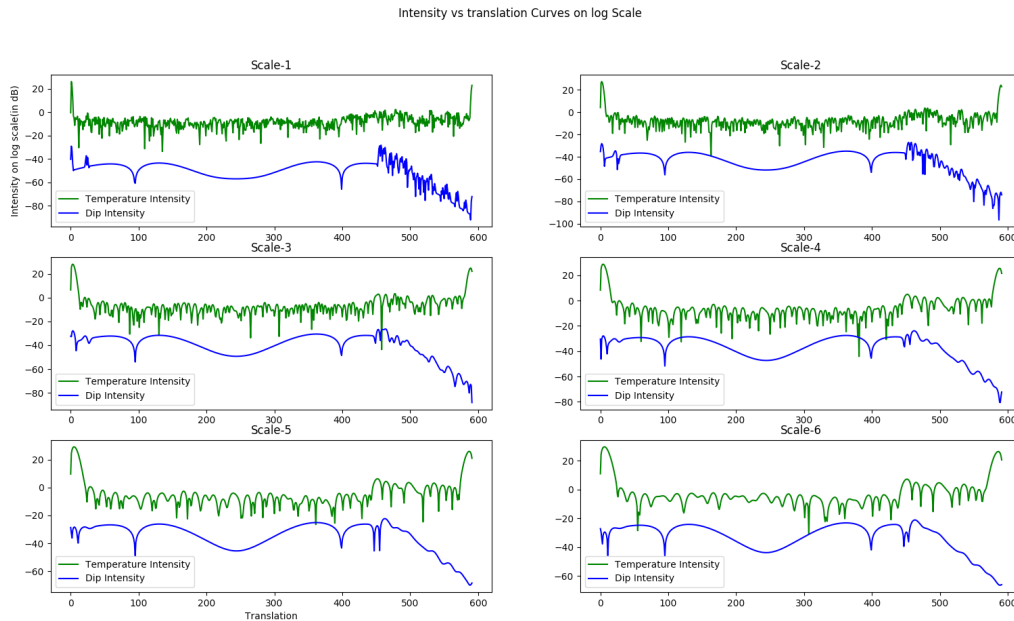


Fig 10 CWT of Foreground and the Signal using Ricker wavelet over the scale 1 to 6 (as described in step 2)

In order to find the contrast between the signal and the foreground, the following procedure was followed:

1. Compute CWT of the Signal and the foreground
2. Compute and plot the coefficients on a logarithmic scale(Consider absolute values, as logarithm can be computed only for positive values)

3. Compute FFT of data obtained in 2
4. Plot the results on a log scale to cover the Dynamic range

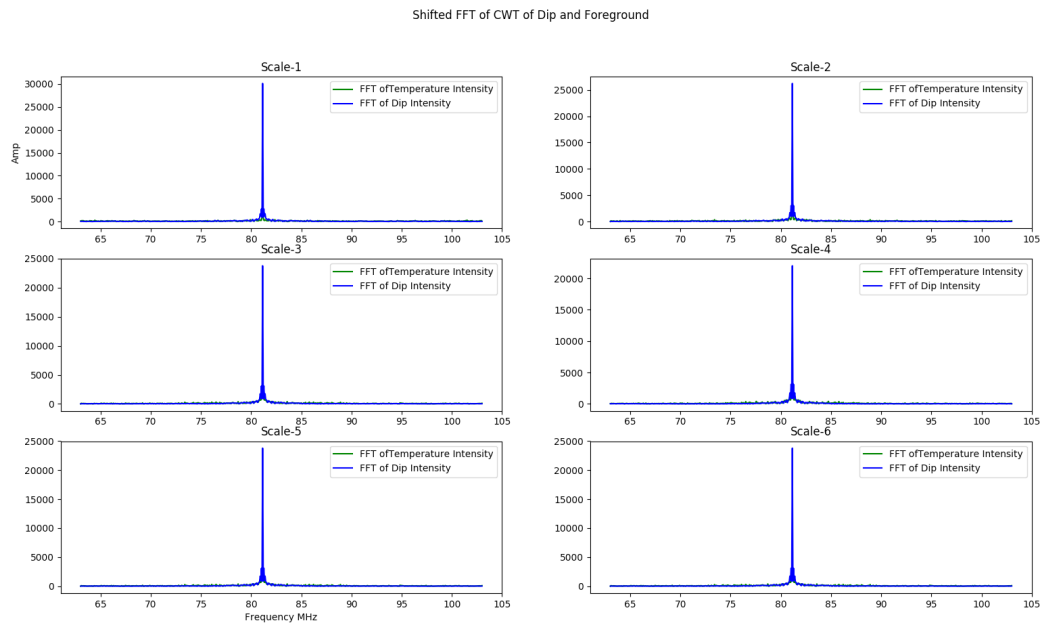


Fig 11 FFT of CWT coefficients (*as described in step 3*)

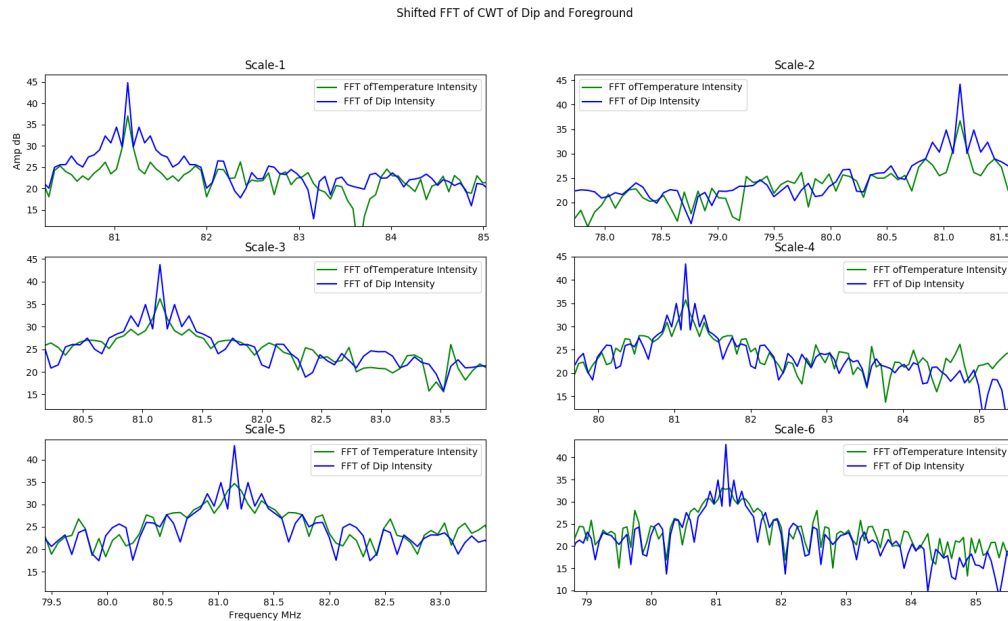


Fig 12 FFT plotted on a log scale, maximised near the peak (*as describes in step 4*)

It can be clearly seen that the dip shoots up at about 81.8MHz. Although this method gives an excellent contrast between the signal and the foreground, in real-time data,

the Dip is uncertain and is not known beforehand. Thus, this cannot be a generalised solution.

3.3 Discrete Wavelet Transforms

The important feature to be noted in DWT coefficients is that it always gives the same number of coefficients as the original signal, unlike CWT, which gives M-by-N matrix of N samples for an N-length signal and M number of scale. Also, in DWT, many of the coefficients may be close to zero in value, making it convenient to throw away those coefficients and still maintain a high-quality signal approximation. The CWT is a highly redundant transform. There is significant overlap between wavelets at each scale and between scales. It also requires much larger computational resources than that in DWT *Wavelet Toolbox, MathWorks*.

3.3.1 Reconstruction of Signal from the foreground data

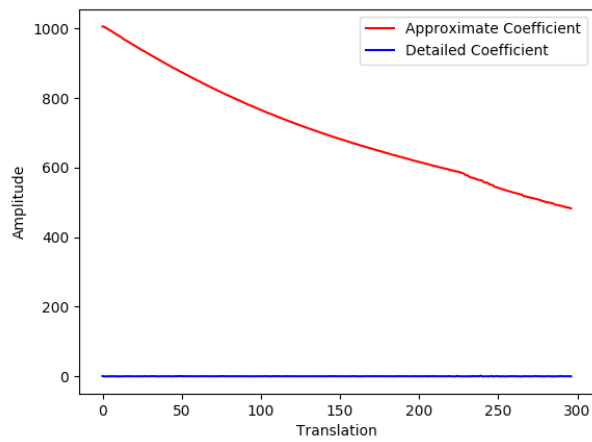
In order to complete the feature extraction, signal reconstruction is necessary. Inverse CWT can be implemented, but it is often the case that the reconstruction is not perfect. Because the CWT is a redundant transform, there is no unique way to define its inverse *Wavelet Toolbox, MathWorks*. The inverse CWT requires an infinite number of coefficients. Due to the fact that CWT takes in non-base 2 values also, there can be many values in an octave, unlike in DWT, which has only one value in an octave. This makes CWT an overlapping transform with 2D coefficient matrix while DWT gives one-dimensional coefficients corresponding to each data points.

The DWT provides a perfect reconstruction of the signal upon inversion. This implies that DWT coefficients of a signal can be used to synthesize an exact reproduction of the signal within numerical precision. The outputs of the DWT algorithm are **detailed coefficients**, corresponding to high-frequency detail signals, and **approximation coefficients**, corresponding to the coarse approximation of the original signal in time domain Lee, et al, 1994. They represent the degree of

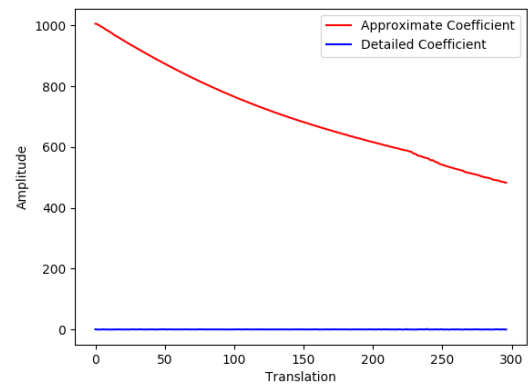
correlation between the analyzed signal and the wavelet function at different instances of time, making it suitable for feature extraction in spectral analysis.

The procedure followed for reconstruction of Signal using DWT is:

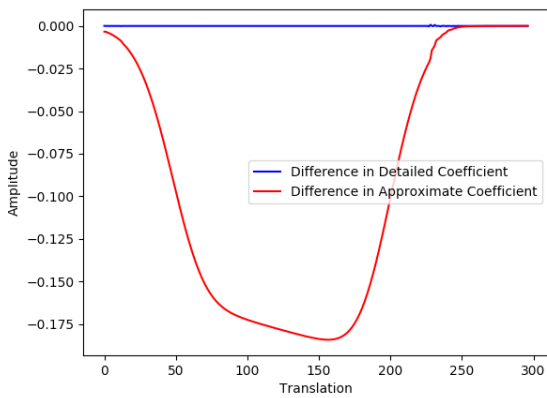
1. Compute DWT of Foreground
2. Compute DWT of foreground+dip
3. Find the difference between DWT coefficient of (2) and (1), both detailed and approximate
4. Compute IDWT of Difference



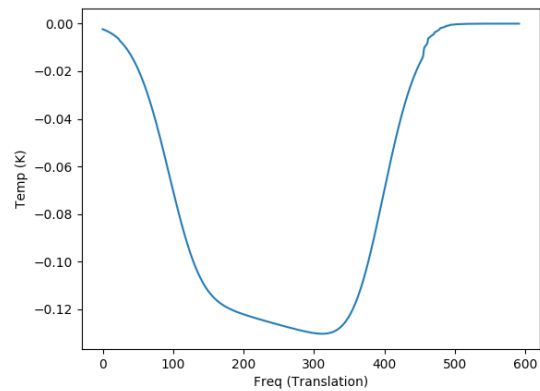
a) DWT of Foreground using Daubechies2 wavelet



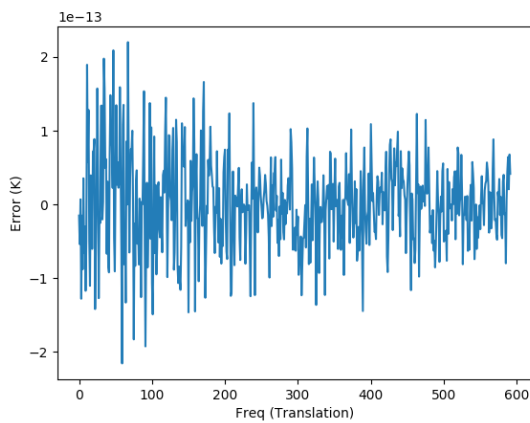
b) DWT of Foreground+Signal using Daubechies2 wavelet



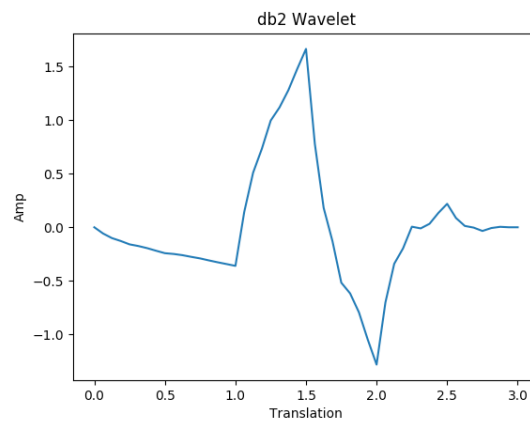
c) Difference in Coefficients from (b) and (a)



d) Reconstructed signal



e) Error in the reconstructed signal



f) Daubechies 2 wavelet

Fig 13 Signal reconstruction using db2 wavelet in DWT

The above results are obtained from asymmetric, orthogonal Daubechies 2 wavelet. Considering both approximate and detailed coefficients, the dip can be satisfactorily reconstructed. The error in the reconstructed signal is observed to be in the order of 10^{-13} , which is an acceptable margin of error in signal reconstruction. Discrete wavelets like Haar, Daubechies 2 and Symlet 2 wavelets were also used in the signal reconstruction. All of them gave similar results.

The DWT analysis windows are fixed in both the time and scale directions, so if the coefficients are plotted, a grid of boxes that start out large at one end of the scale axis and end as smaller ones will be obtained Jorgensen, et al, 2009. To verify the details of the DWT, second level decomposition is applied to the obtained approximate coefficients. The results were similar, upholding the fact that the approximate coefficients give similar plot as that of the input signal while detailed coefficients are much smaller.

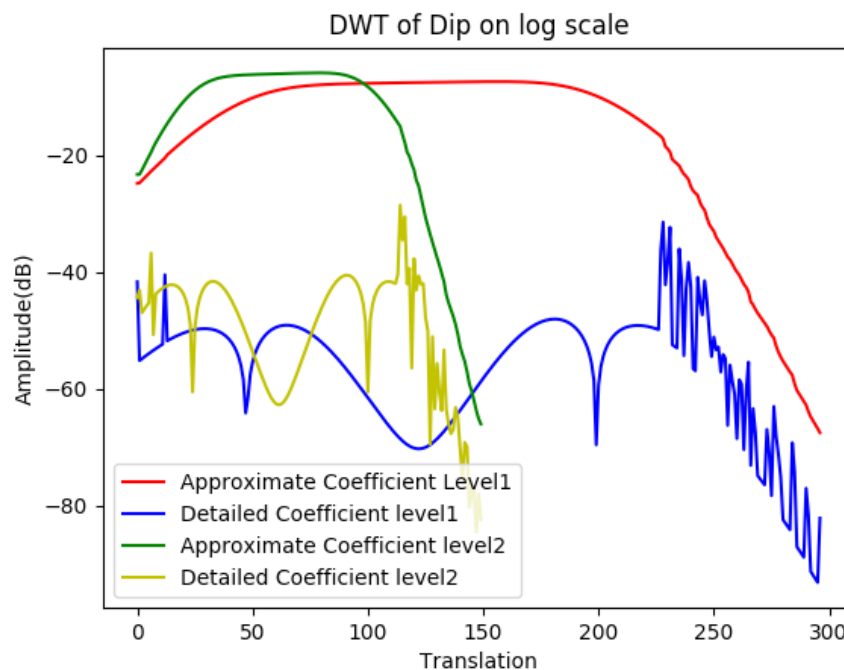


Fig 14 Level 1 and Level 2 DWT of Signal on a log scale

The difference in wavelet transform coefficients of the sum of foreground and signal & the foreground gives the signature of the Dip.

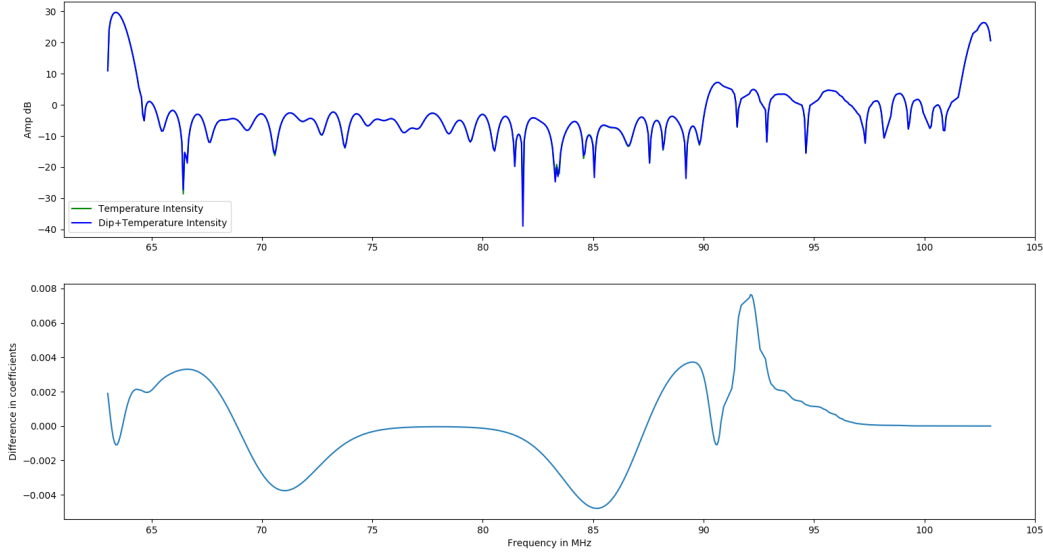


Fig 15 Signature of the Signal in the wavelet domain

This signature is clearly due to Dip, as the effect of the foreground has been removed. But when there is variation in the foreground, this signature may not remain the same. Thus finding a basis function which makes the signal stand out in the foreground is essential.

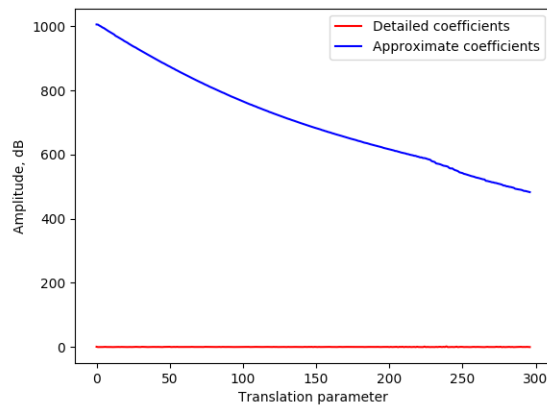
3.4 Computing with GPU

As a second phase of the project, DWT was implemented on a GPU available in the lab, *Nvidia GeForce GTX 1050*. A GPU uses special programming tool to help it analyze and use data. In this project, PyCuda which gives Pythonic access to Nvidia's CUDA parallel computation API was used. The *pycuda* module was used to fetch the details of GPU being used. It was also a verification if the GPU is active in computation. The various device properties that were verified were:

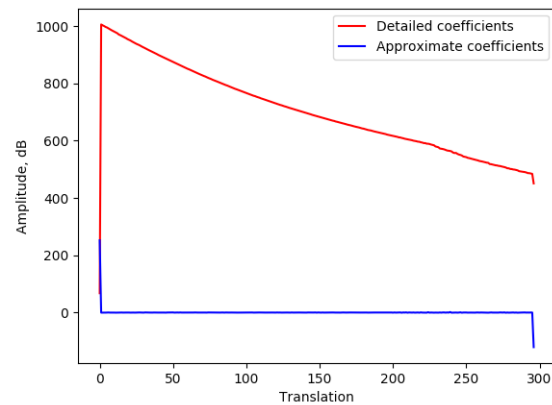
Table 1 Device Properties for GeForce GTX 1050

	A	B
1	CUDA Driver Version / Runtime Version	9.1 / 9.1 C
2	CUDA Cores	128
3	GPU Max Clock rate	1569 MHz (1.57 GHz)
4	Memory Bus Width	128-bit
5	Total number of registers available per block	65536
6	Warp size	32
7	Maximum number of threads per multiprocessor	2048
8	Maximum number of threads per block	1024
9	Max dimension size of a thread block (x,y,z)	(1024, 1024, 64)
10	Multiprocessors	5

To implement DWT on GPU, PyCuda's PyGASP package was used. It provides with the `dwtCuda` function for the computation of DWT using GPU. DWT was implemented on both CPU and GPU and the results were compared.



a) Level-1 DWT of foreground using db2 wavelet computed in CPU



b) Level-1 DWT of foreground using db2 wavelet computed in GPU

Fig 16 Comparison of CPU and GPU results

The error in the DWT coefficients was calculated. It can be seen that the GPU and CPU give similar results.

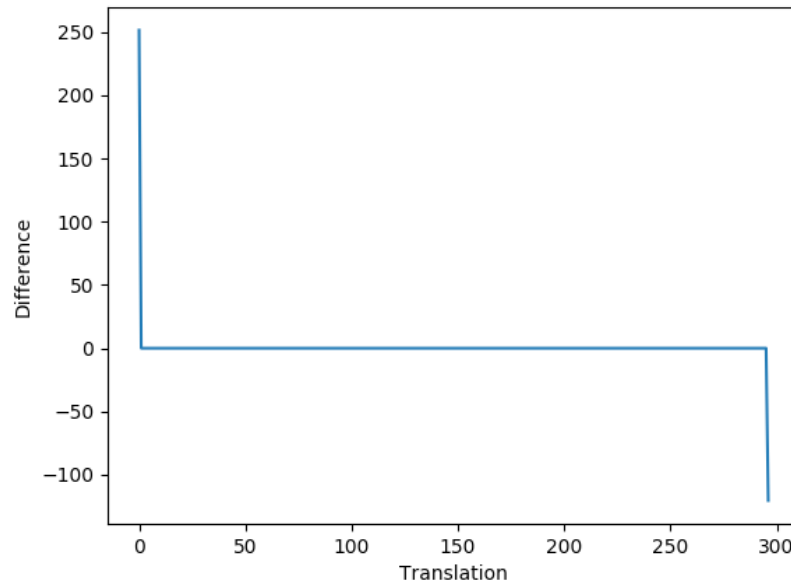


Fig 17 Difference in DWT coefficients as calculated in CPU and GPU

When the time duration of the operation was seen, it was noted that

CPU Time= 117 μ sec

GPU time= 1901 μ sec

Enabling *nvprof* to see the detailed time scale of computation for the DWT of foreground using db2 wavelet, GPU activity of about 51 μ sec was observed.

```

==9909== NVPROF is profiling process 9909, command: python dwt.py
CPU Time= : 0:00:00.000117
GPU time= 0:00:00.001901
==9909== Profiling application: python dwt.py
==9909== Profiling result:

```

Type	Time(%)	Time	Calls	Avg	Min	Max	Name
GPU activities:	41.44%	22.622us	14	1.6150us	1.2800us	2.3040us	[CUDA memcpy HtoD]
	28.72%	15.679us	3	5.2260us	5.0560us	5.3120us	cuDWT
	24.50%	13.376us	7	1.9100us	1.8560us	1.9840us	[CUDA memcpy DtoH]
	5.33%	2.9120us	1	2.9120us	2.9120us	2.9120us	cuIDWT
API calls:	67.50%	244.49ms	1	244.49ms	244.49ms	244.49ms	cuCtxCreate
	28.74%	104.09ms	1	104.09ms	104.09ms	104.09ms	cuCtxDetach
	1.81%	6.5399ms	20	326.99us	7.4980us	1.5089ms	cuMemAlloc
	1.63%	5.9144ms	20	295.72us	7.0590us	1.0546ms	cuMemFree
	0.09%	339.87us	14	24.276us	8.4750us	52.235us	cuMemcpyHtoD
	0.08%	287.50us	1	287.50us	287.50us	287.50us	cuModuleLoadDataEx
	0.06%	202.70us	7	28.957us	19.263us	40.731us	cuMemcpyDtoH
	0.04%	159.54us	1	159.54us	159.54us	159.54us	cuModuleUnload
	0.04%	145.13us	4	36.283us	23.176us	46.714us	cuLaunchKernel
	0.00%	6.9360us	2	3.4680us	1.0410us	5.8950us	cuCtxPopCurrent
	0.00%	6.0740us	4	1.5180us	807ns	1.8970us	cuModuleGetFunction
	0.00%	5.8690us	4	1.4670us	1.0910us	1.9030us	cuFuncSetBlockShape
	0.00%	4.2300us	2	2.1150us	365ns	3.8650us	cuCtxPushCurrent
	0.00%	3.0820us	2	1.5410us	1.3280us	1.7540us	cuCtxGetDevice
	0.00%	3.0220us	3	1.0070us	264ns	2.2450us	cuDeviceGetCount
	0.00%	1.4870us	2	743ns	656ns	831ns	cuDeviceGet
	0.00%	1.0650us	3	355ns	254ns	492ns	cuDeviceGetAttribute
	0.00%	945ns	1	945ns	945ns	945ns	cuDeviceComputeCapability

Fig 18 Profiling of the GPU activity in the computation of DWT

This could be because of the fact that the GPU is configured for the floating point operations, while the data given is a double precision array. However, Signal reconstruction was done in GPU using DWT. The procedure followed was the same as in the case of CPU explained in the section 3.3.1.

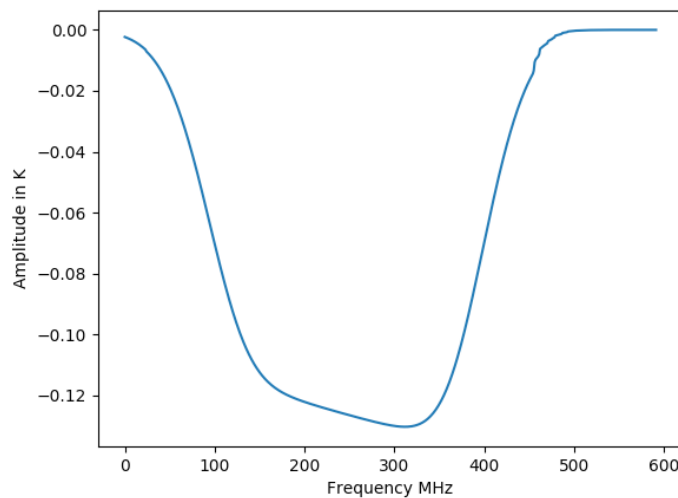


Fig 19 Reconstructed Signal using DWT module in PyGASP

4 CONCLUSION AND RECOMMENDATIONS

When the signal is a sinusoid buried in the noise, analysis of the Fourier transform technique, very clearly indicated that a sinusoid shows up with a higher signal to noise ratio in the frequency domain. However, if the signal is of the nature of the predicted EoR signal in the frequency domain, Fourier transform to the other domain in no way improves the SNR. Hence a simple Fourier analysis is not helpful in detecting EoR signal buried in the foreground. We wanted to investigate and find out whether there are wavelets which are more suitable for separating out EoR signal from the foreground. Thus we started with CWT and found certain disadvantages (as described in section 3.3.1). So we moved on to DWT and explored its properties. We found that by applying the procedure (as described in Section 3.2.1), the contrast of the EoR signal could be significantly improved. However, it has not been verified if this method is independent of the foreground and noise. As the foreground changes with the direction of observation, this method cannot be generalized. Therefore, in an overall sense, one can continue with different foregrounds and verify if it works with all foregrounds. It is to be noted that, if a particular frequency data is in the signal and it is to be found, the wavelet should be similar to the signal we are looking for. Therefore, in order to extract the signal, a custom wavelet similar to the signal and of the comparable magnitude forming orthogonal basis should be developed. Alternatively, we can also take up the job of constructing a set of wavelets which would remove all the foreground but remain orthogonal to the EoR signal.

The results from the GPU were as seen in the CPU, except for the time duration of computation. This time duration of computation can be reduced considerably with various optimization techniques and by upgrading to higher end GPU. However, successful signal reconstruction suggests that GPUs can be used as a potential computational tool in spectral analysis.

ACKNOWLEDGEMENTS

First and foremost I express my deep sense of gratitude to the Indian Academy of Sciences (IASC-INSANASI) for giving me an opportunity to carry out this project in the Summer Research Fellowship Programme-2018.

I owe my heartiest gratitude to respected Prof. Udaya Shankar N, Raman Research Institute, Bangalore for the continuous support, motivation, and immense knowledge. I could not have imagined having a better advisor and mentor for the project.

My sincere thanks to Prof. R Subrahmanyam for his valuable guidance and encouragement throughout the project.

My sincere thanks to Srivani K S, Girish B S and Somshekhar R for their valuable guidance, consideration and always inspiring and supporting me during the internship.

It is my privilege to express regards to Dr. Bhuvaneshwar Babu and Dr.Karthik Shastry for providing the letter of recommendation to the Indian Academy of Sciences.

I thank Shruti Bhatporia, Mugundhan V, Saurabh Singh, Jishnu Nambissan T and Sindhu Gaddam for their continuous guidance, encouragement, moral and academic support throughout my project.

I also thank my institute R V College of Engineering, Bengaluru for allowing me to pursue this internship.

I express my deepest sense of gratitude towards my beloved parents, fellow interns and friends for their constant support and encouragement.

APPENDICES

The various libraries used in this project are:

1. Numpy : A fundamental package for scientific computing with Python, mainly used for numerical operations of arrays. (<http://www.numpy.org/>)
2. Matplotlib: A Python 2D plotting library used extensively in this project for visualising the results. (<https://matplotlib.org/>)
3. FFT package in Numpy: Discrete Fourier Transform package of numpy used for computation of FFT. (<https://docs.scipy.org/doc/numpy-1.13.0/reference/routines.fft.html>)
4. SciPy: A scientific computation tool in Python. The package **scipy.signal** was used extensively for the computation of CWT, using available wavelets. (<https://docs.scipy.org/doc/numpy-1.13.0/reference/routines.fft.html>)
5. PyWavelet: Package used for computation of DWT of data using a wide range of available wavelets. (<https://pywavelets.readthedocs.io/en/latest/>)
6. PyGASP: Package used for computation of DWT and IDWT on GPU. This package contains several transformations used for signal processing, along with associated tools(<https://pypi.org/project/PyGASP/>).

References

1. Bowman, Judd D., et al. "An absorption profile centred at 78 megahertz in the sky-averaged spectrum." *Nature* 555.7694 (2018): 67.
2. Ganjam, Samskruthi. (2018)., "Satellite Mission for the study of Epoch of Reionisation", submitted as B.Tech thesis to Ramaiah Institute of Technology, Bangalore
3. <https://www.haystack.mit.edu/ast/science/epoch/>
4. Singh, Saurabh, et al. "SARAS 2: a spectral radiometer for probing cosmic dawn and the epoch of reionization through detection of the global 21 cm signal." *arXiv preprint arXiv:1710.01101* (2017).
5. Suvichakorn, Aimamorn, and Jean-Pierre Antoine. "Analyzing NMR spectra with the Morlet wavelet." *Signal Processing Conference, 2008 16th European*. IEEE, 2008.
6. Vijendra, V., and Meghana Kulkarni. "ECG signal filtering using DWT haar wavelets coefficient techniques." *Emerging Trends in Engineering, Technology and Science (ICETETS), International Conference on*. IEEE, 2016.

7. Gautam, Desh Deepak, V. K. Giri, and K. G. Upadhyay. "Feature extraction of HRV signal using wavelet transform." *Convergence in Technology (I2CT), 2017 2nd International Conference for*. IEEE, 2017.
8. Wu, L. C., C. C. Kao, and D. J. Doong. "Application of the wavelet transform and inverse wavelet transform to analyze the ocean wave signals from data buoys." *OCEANS 2010 IEEE-Sydney*. IEEE, 2010.
9. [shodhganga.inflibnet.ac.in/bitstream/10603/26360/12/12_chapter 7.pdf](http://shodhganga.inflibnet.ac.in/bitstream/10603/26360/12/12_chapter%207.pdf)
10. Jorgensen, Palle ET, and Myung-Sin Song. "Comparison of discrete and continuous wavelet transforms." *Encyclopedia of Complexity and Systems Science* (2009): 1-30.
11. Andreas Klöckner., "GPU Metaprogramming using PyCUDA: Methods & Applications", Division of Applied Mathematics Brown University Nvidia GTC · October 2, 2009
12. <https://streamhpc.com/blog/2017-01-24/many-threads-can-run-gpu/>
13. Lathi, B. P. *Principles of Signal Processing and Linear Systems*. ser. 9780198062288, Oxford International, 2009.
14. R. E. J. Yohanes, W. Ser and G. b. Huang, "Discrete Wavelet Transform coefficients for emotion recognition from EEG signals," 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, 2012, pp. 2251-2254. doi: 10.1109/EMBC.2012.6346410
15. <https://in.mathworks.com/help/wavelet/gs/continuous-and-discrete-wavelet-transforms.html>
16. Lee, Daniel TL, and Akio Yamamoto. "Wavelet analysis: theory and applications." *Hewlett Packard journal* 45 (1994): 44-44.

Source

1. Fig 1: Bowman, Judd D., et al. "An absorption profile centred at 78 megahertz in the sky-averaged spectrum." *Nature* 555.7694 (2018): 67.