High cadence kurtosis based RFI excision for CHIME

by

Arash Mirhosseini

B.Sc., Leipzig University, 2016
M.Sc., Paris-Saclay University, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Astronomy)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled **High cadence kurtosis based RFI excision for CHIME** submitted by **Arash Mirhosseini** in partial fulfillment of the requirements for the degree of **Master of Science** in **Astronomy**:

- Mark Halpern, Physics and Astronomy (Supervisor)
- Richard Shaw, Physics and Astronomy (Examining committee member)

Abstract

This document describes the real-time Radio Frequency Interference detection system for the Canadian Hydrogen Intensity Mapping Experiment (CHIME). CHIME is a transit radio telescope located at Dominion Radio Astrophysical Observatory (DRAO) in Penticton, BC, and it is originally designed to map the large-scale structures in the redshift range $0.8 < z < 2.5$ by observing the 21-cm emission line of the neutral hydrogen atom. One of the common problems for astrophysical radio observations is Radio Frequency interference (RFI) from terrestrial sources such as TV stations, airplanes, cellphones, etc. RFI detection and mitigation is an essential part of any research in radio astronomy, because RFI contaminates the astrophysical data and reduces the sensitivity of the telescope to the faint sources. Since most of the RFI is non-Gaussian and lasts less than one second, we developed a real-time high cadence RFI excision system for CHIME which uses the fourth statistical moment (kurtosis) to detect non-gaussianity in the signal.

In this thesis, I introduce the algorithms for kurtosis based RFI excision that we have used in CHIME. The algorithms were tested and the results were compared with each other. I also discuss the effect of truncation of the samples in CHIME correlator on the spectral kurtosis estimates. I show that truncation bias causes the RFI system to flag bright celestial sources. I derive a correction for the truncation bias with a polynomial fitting and a cubic spline interpolation. Moreover, I evaluate the quality of the CHIME data taken between May 2019 and September 2020. I find that the RFI excision system can mitigate many types of RFI by excising less than 20% of the data (on average), from intermittent RFI caused by satellites or airplanes to static RFI, especially between 400 MHz and 500 MHz.

Lay Summary

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is a radio telescope in Penticton, BC, that is originally designed to probe the dark energy and make the largest 3D map of the large-scale structures in the observable universe. One of the common problems in radio astronomy research is the Radio Frequency Interference (RFI). RFI is a human made signal that interferes with the cosmic signal and degrades the quality of astronomical data. Since CHIME aims to measure very weak signals coming from deep space, we need to detect and mitigate RFI to increase the sensitivity of the telescope. In this thesis, I introduce the RFI excision system for CHIME and I evaluate its performance from May 2019 to September 2020. I find that the system is able to detect and mitigate different types of RFI and increase the sensitivity of the system.

Preface

This thesis is original, unpublished work by the author, Arash Mirhosseini, conducted as part of the CHIME collaboration, under supervision of Mark Halpern with input and ideas from Richard Shaw, as well as other members of the CHIME team. The RFI excision system for CHIME was previously designed and explained in Taylor et al. (2019) [1]. The contribution I made was to evaluate the performance of different excision algorithms with different thresholds, find and correct the bugs in the spectral kurtosis estimator (for example 4+4-bits truncation bias) and assess the data quality after RFI excision.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Twenty years ago, measurements of the distance to the Type Ia supernovae by two independent teams confirmed the discovery of accelerating expansion of the universe [2, 3]. The nature of this acceleration is still unknown and it is usually attributed to an enigmatic component of the universe: dark energy. Dark energy can be considered as a fluid with an equation of state $P = w\rho$, where P and ρ are the pressure and density of the fluid, respectively, and w is the equation of state parameter. The current observations are compatible with $w = -1$ which corresponds to a static density. More accurate observations are needed to impose tighter constraints on w as well as other cosmological parameters. Putting tighter constraints on the parameter w is an essential step towards understanding the nature of dark energy.

In this chapter, I review some basic concepts that are necessary for understanding the scientific goal of CHIME; a project whose primary goal is to characterize the dark energy by mapping the large scale structures between the redshift $0.8 \leq z \leq 2.5$ using the 21-cm intensity mapping technique.

1.1 Theoretical background

According to the cosmological principle, the Universe is homogeneous and isotropic at large scales. This principle is observationally supported by the quasi isotropy of the cosmic microwave background (CMB) and the homogeneity of the distribution of the galaxies at large scales. In the framework of General Relativity, symmetries of the cosmological principle can be used to define the metric for a homogeneous and isotropic universe, the Friedman-Lemaitre-Robertson-Walker (FLRW) metric:

$$ds^2 = dt^2 - a^2(t)\Big(\frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + sin^2\theta d\phi^2)\Big), \tag{1.1}$$

where $a(t)$ is the dimensionless scale factor, and $K = -1, 0$ or $+1$ is the spatial curvature parameter corresponding to an open, flat or closed universe, respectively. Then, the evolution of such a homogeneous and isotropic universe is described by the Friedman equations. They are derived from Einstein equations by substituting

the Ricci tensor (computed from FLRW metric) together with the stress-energy tensor for a perfect fluid into Einstein field equations:

$$H^2(z) \equiv (\frac{\dot{a}(t)}{a(t)})^2 = -\frac{K}{a^2(t)} + \frac{8\pi G}{3}\sum \rho \tag{1.2}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\sum(\rho + 3p), \tag{1.3}$$

where $H(z)$ is the Hubble constant at redshift z, and ρ and p are the density and pressure of different components of the Universe, respectively. The sum is taken over all components of the Universe, namely, matter, radiation and dark energy. The equation of state for the component x of the Universe is $p_x = w\rho_x$. The critical density is defined as the total density of a flat universe $\rho_c = 3H^2(z)/8\pi G$ (current critical density). The density parameter of the component x of the Universe is $\Omega_x = \rho_x/\rho_c$. The first Friedman equation (Eq. 1.2) for a flat universe can be written in terms of density parameters of different components of the Universe:

$$H^2(z) = H_0^2\Big(\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_{DE}\frac{\rho_{DE}(z)}{\rho_{DE}(z=0)}\Big), \tag{1.4}$$

where indices m, r, DE correspond to the matter, radiation and a generic form of dark energy.

The simplest form of dark energy is a cosmological constant. It turns out that a cosmological constant behaves like a fluid with static density ($\dot{\rho} = 0$). In this case, from continuity equation $p = -\rho$ and so equation of state parameter for the cosmological constant would be $w = -1$. Although current observations are compatible with $w = -1$, dark energy in the form of a cosmological constant is not the unique possibility.

1.1.1 Cosmological models

The simplest model which provides a good fit to a wide range of observational data is the standard ΛCDM model. It assumes a flat universe with the simplest possible form of dark energy, i.e. a cosmological constant (Λ), and Cold Dark Matter (CDM). A true cosmological constant has an equation of state parameter $w = -1$ which does not change in time. Standard ΛCDM model comprises 6 free parameters, including density parameters of baryons Ω_b and dark matter Ω_{DM}, optical depth at the time of reionization τ , current value of Hubble constant H_0, and primordial amplitude of scalar fluctuations A_s, and spectral index of scalar fluctuations n_s related to the inflationary epoch of the Universe.

Although this model can be well fitted to the data to find the parameters, it is not unique and extensions to the standard ΛCDM model are possible. Some extensions

to standard Λ CDM are models with dynamical dark energy (time-varying w), neutrino masses and additional relativistic particles. For example, if the dark energy is not a cosmological constant, then it is possible that the w parameter changes with time. In this case, a time-varying w parameter can be considered as a Taylor expansion of w at the first order in the scale factor:

$$w(t) = w_0 + [1 - a(t)]w_a. \tag{1.5}$$

Note that at low redshifts (i.e. $a(t) \approx 1$) we have $w(t) \approx w_0$ and only high redshift measurements are sensitive to w_a parameters.

A detailed list of extensions to the standard ΛCDM model is given in [4]. Because there is more than one model which can be fitted to the current observational data, one needs to combine different cosmological probes (CMB, BAOs, SNe Ia, weak lensing, etc.) to constrain cosmological parameters and invalidate some of these models.

1.2 Dark Energy probes

1.2.1 Supernovae Type Ia

Objects with known luminosity can be used as distance estimators. Luminosity distance is defined by $d_L^2 = L/4\pi f$, where L is the luminosity and f is the observed flux of the source. Hence, for an object with known luminosity, the problem of finding the distance is reduced to flux measurement. Such an object with a known luminosity is called a *standard candle*. Type Ia supernova (SN Ia) is an example of standard candles. Although SNe Ia show a dispersion of about 40% at their peak luminosity (which makes them not good distance estimator), it turns out that the dispersion is reduced if one considers the so called brighter-bluer and brighter-slower correlations. Brighter-bluer relation is the correlation between the rest frame color of the supernova and its maximum luminosity: bluer supernovae are brighter than redder ones. Brighter-slower relation, known as Phillips relation [5], is the correlation between the maximum luminosity of SNe Ia, and the decay rate of their light curve after reaching the peak luminosity: The light curve of a bright supernova decays slower than fainter ones. An empirical correction of the light curve using these two correlations reduces the peak luminosity dispersion of SNe Ia [6] and makes type Ia SNe a good distance estimator.

In a flat universe, the luminosity distance of a source observed at a redshift z is related to the other cosmological parameters θ_i, through:

$$d_L(z) = (1 + z) \times \int_0^z \frac{dz'}{H(z', \theta)}. \tag{1.6}$$

Hence, measurement of luminosity distance d_L will constrain the Hubble parameter at redshift z, and consequently, the density parameters through equation 1.4. Therefore, type Ia SNe are powerful probes of the cosmic expansion history and they are very sensitive to the w parameter in the equation of state of dark energy. After the discovery of the dark energy in the late 1990s, several second generation surveys have collected samples of a few hundred well-measured SNe Ia up to $z \sim 1$. The Supernova Legacy Survey (SNLS) used data taken as part of deep component of the five-year Canada-France-Hawaii survey (CFHT-LS). Using a rolling-search approach (i.e., repeatedly imaging the same sky patch), it targeted four one square degree fields during 5 to 7 consecutive lunations per year using four different broad-band filters g_M,r_M,i_M,z_M. During its 5 years of operation (mid-2003 to mid-2008) it delivered about 500 SNe Ia [7].

SDSS-II supernova survey used SDSS camera on SDSS 2.5 m telescope at the Apache Point Observatory (APO) and it searched for supernovae during the northern fall season of 2005 to 2007. Images were taken in *ugriz* SDSS passbands [8] with a typical cadence of once every four nights. The supernovae candidates were identified by a computing cluster at APO within 24 hours of data collection. Then, the spectroscopic follow up was performed using a dozen of telescopes. Details of the SDSS-II supernova survey are given in [9] and [10]. Out of 4607 candidates, 500 SN Ia were confirmed by spectroscopic follow up at a redshift $z < 0.5$. However, only 413 of them were used to constrain the cosmological parameters. Assuming a ΛCDM model, acceleration ($\Omega_M < 2_\Lambda$) was detected at a confidence of 3.1σ. Moreover, with a flat geometry $\Omega_\Lambda > 0$ at a confidence of 5.7σ is required and $\Omega_M = 0.315 \pm 0.093$. The details of the analysis is given in [11].

Joint Light Curve Analysis (hereafter JLA) was part of SNLS-SDSS collaborative effort that was initiated in 2010 to improve the previous analysis of SNLS and SDSS teams. The main goals of this collaborative effort were to improve the accuracy of photometric calibration of both surveys [12], to determine more rigorously the uncertainties in SNe Ia light curve models [13], and to include the full SDSS-II SNe Ia spectroscopic sample in cosmology analysis. JLA sample of SNe Ia which is used to improve the cosmological constraints includes 740 supernovae selected from the data of SDSS-II, SNLS, HST and several nearby experiments.

JLA sample is currently state-of-the-art collection of SNe Ia in the redshift range $0.01 < z < 1.2$ [14]. However, distances to supernovae with a redshift above 0.8 are not well measured due to photometric uncertainties. The next generation of surveys, including the Large Synoptic Survey Telescope (LSST) aim to increase the sample size of SNe Ia at higher redshifts.

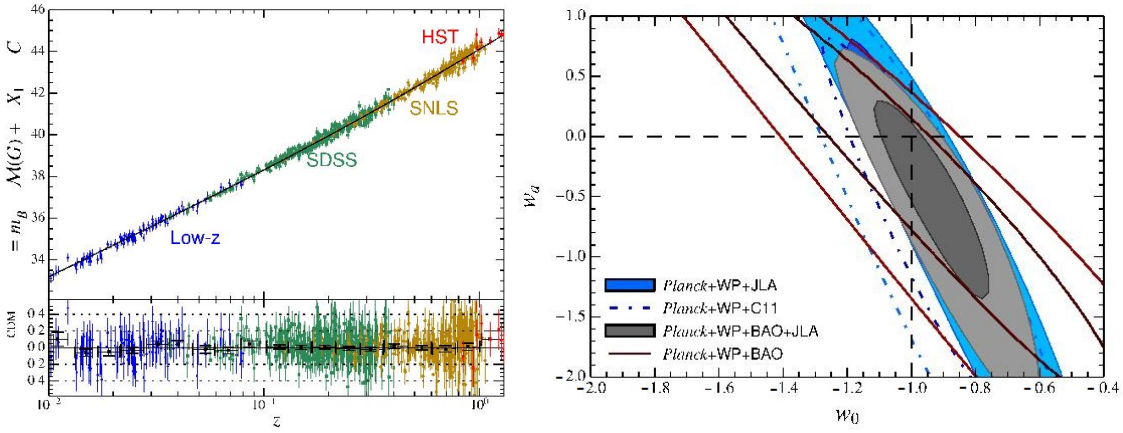

Figure 1.1: Left: Hubble diagram from the JLA sample (top) and Residuals from the best ΛCDM fit (bottom). Right: 68% and 95% confidence contours of w and w_a parameters for the flat $w-\Lambda$CDM model (figures taken from [14]).

1.2.2 Baryon Acoustic Oscillations

An alternative to the standard candles to constrain the cosmological parameters is the *standard rulers*. A standard ruler is an object or statistical feature whose intrinsic size is known and do not change with time. One example of standard rulers is the Baryon Acoustic Oscillations (BAOs). In the following, I briefly discuss the origin of BAOs.

The primordial universe was a hot and dense plasma of photons, baryons and dark matter, where baryons and photons were coupled together via Compton scattering. The initial overdensities of matter attract the surrounding baryons and dark matter. Because dark matter has no electromagnetic interaction, it only feels the gravitational pull from the overdensity. The baryons feel two competing forces: An inward gravitational force from the overdensity and a photon pressure outward. Baryon-photon fluid is compressed, the temperature increases, and the outward photon pressure increases. At some point, the outward force from photon pressure will be stronger than the compressing force of gravity, and thus the region will start to expand in the form of a sound wave. The speed of the sound wave in the plasma in the unit of the speed of light is:

$$C_s(z) = \frac{c}{\sqrt{3(1+R(z))}} \tag{1.7}$$

where $R(z)$ is the baryon to photon density ratio $R(z) = \rho_b(z)/\rho_r(z)$. At a redshift

of $z_{rec} \sim 1100$ the plasma is cool enough so that the electrons combine with protons and form neutral hydrogen. At this time photons decouple from baryon. The photon pressure is removed, leaving a spherical shell of baryons at a fixed radius from the initial overdensity. The characteristic comoving radius of the baryonic shell at the time of recombination is:

$$r_s = \int_{z_{rec}}^{\infty} \frac{C_S(z)dz}{H(z)} = \int_{z_{rec}}^{\infty} \frac{C_S(z)dz}{\sqrt{H_0(\Omega_m(1+z)^3 + \Omega_r(1+z)^4}}, \tag{1.8}$$

which is the comoving distance the baryonic shell has travelled until the time of recombination. Redshift at recombination z_{rec} is precisely determined by atomic physics, and the cosmological parameters in the above integral are measured by the CMB observations. These two fix the comoving radius of the baryonic shell at recombination. Planck has measured the acoustic scale at the time of recombination $r_s \approx 145$ Mpc [15]. Figure 1.2 shows the evolution of acoustic scale as a function of redshift.

Now that the characteristic scale of the BAO is known, it can be used as a standard ruler to estimate the distances and constrain the w parameter in the equation of state of dark energy. To use the BAO scale for distance measurements, we need to measure its observational scale. The angular diameter distance for a flat universe is:

$$d_A(z) = \frac{1}{1+z}\chi(z) = \frac{1}{1+z}\int_0^z \frac{cdz'}{H(z')}, \tag{1.9}$$

where $\chi(z)$ is the comoving distance. Since the comoving size of acoustic scale r_s is constant after recombination, the observed angular scale of BAO θ is related to the angular diameter distance via:

$$\theta = \frac{r_s}{D_M} = \frac{r_{s\perp}}{d_A(1+z)}, \tag{1.10}$$

where D_M is the *comoving* angular diameter distance, and $r_{s\perp}$ is the perpendicular component of r_s to the line of sight. Note that while perpendicular component of r_s in a *redshift slice* gives the angular diameter distance , the radial component $r_{s\parallel}$ along the line of sight corresponds to different redshifts and constrain the Hubble parameter through the Hubble's law:

$$r_{s\parallel} = \frac{\Delta z}{H(z)} \tag{1.11}$$

The signature of BAOs is imprinted in the two point correlation function of matter distribution. Two point correlation function (hereafter, 2PCF) quantifies the excess probability of finding a pair of galaxies at some redshift z that are separated

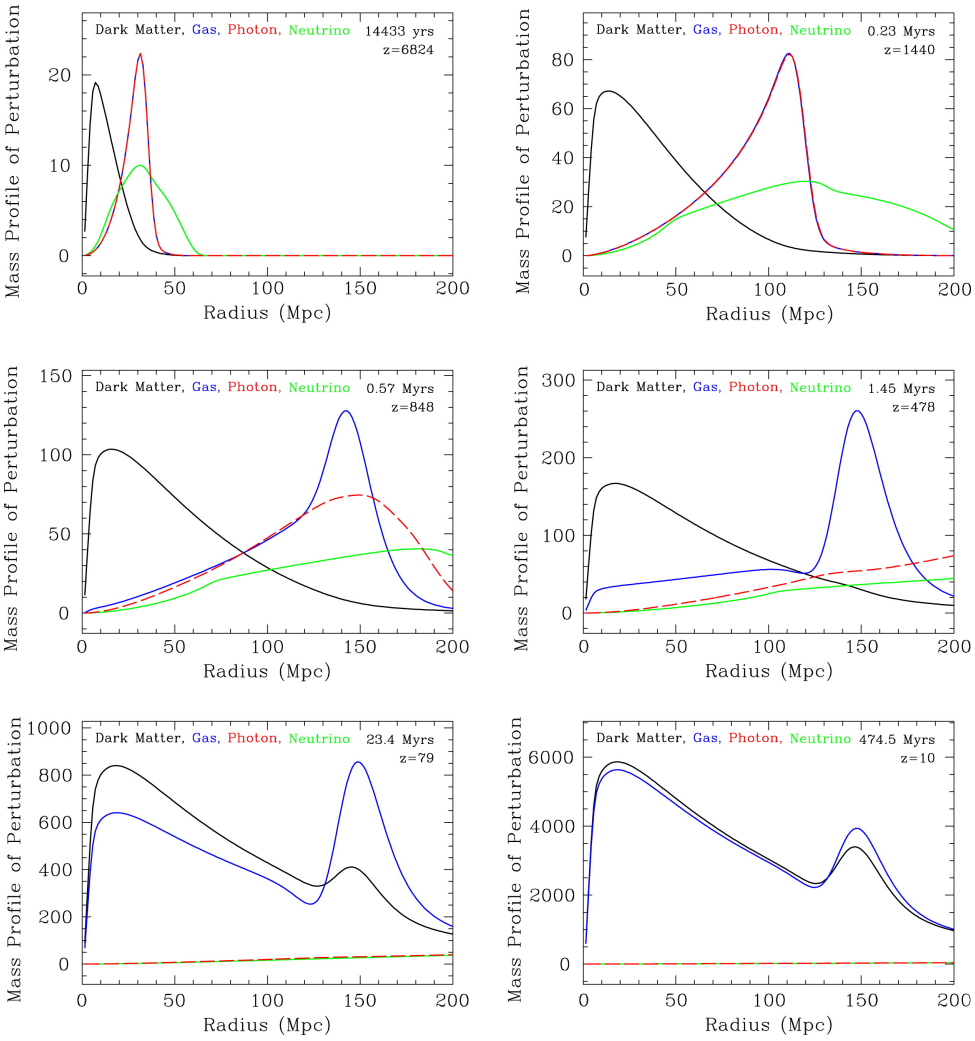

Figure 1.2: Figure shows the generation of BAO peak from initial overdensity. Dark matter overdensity sits where it is initially, as it has no electromagnetic interaction. Photons (red) and baryons (blue) move together until recombination ($z_{rec} \sim 1100$). After recombination photons free-stream and baryons attract surrounding dark matter gravitationally. In the end, there will be a mass concentration at a characteristic scale of $r_s \approx 148$ Mpc (figure from [16])

by a distance s compared to the case if they are uniformly random distributed. To do this, a random catalog (that is a random sample of galaxies) is constructed and the distribution of galaxies is compared with the data. A commonly used estimator for 2PCF $\zeta(s)$ is given by [17]:

$$\hat{\zeta(s)} = \frac{DD(s) - 2DR(s) + RR(s)}{RR(s)}, \tag{1.12}$$

where s is the comoving separation between two galaxies, $DD(s)$ and $RR(s)$ are the number of galaxy pairs with separation s in the real-real catalog (data) and random-random catalog, respectively, and $DR(s)$ is the number of galaxy pairs with a separation s between a galaxy in the real data and a galaxy in the random catalog. The BAO signal appears as a bump in the 2PCF.

The first detection of the BAO signal in large scale was made in 2005 by measuring the large-scale correlation function from a spectroscopic sample of luminous red galaxies from SDSS data release 3 in the redshift range of 0.16 to 0.47 by Eisenstein et al. [18]. More recent measurements of the BAO scale have been made by different surveys including WiggleZ Dark Energy Survey which used $\sim$ 130000 galaxies at redshift $z = 0.6$ [19], and SDSS-III BOSS survey of $\sim$ one million galaxies out to a redshift of 0.7 [20]. Moreover, BOSS detected the BAO signal using the Lyman α forest spectra of $\sim$ 48000 high redshift quasars across the redshift range $2.1 \leq z \geq 3.5$ [21].

Since the BAO scale is very large, a large volume of the universe must be probed to decrease the sample variance. Therefore, we need a large sample of galaxies to detect the BAO signal. For example, an order of 10^8 galaxies is required to approach the cosmic variance limit at redshift $z > 1$ [16]. Galaxy-redshift surveys are very expensive for this purpose: First, obtaining the redshift of each galaxy through high resolution spectroscopy for a large sample of galaxies is very time consuming. Moreover, high redshift objects are fainter and harder to detect and the measurements are dominated by shot noise. There is an alternative for the galaxy-redshift surveys which is discussed in the section 1.3.1.

1.2.3 Weak Gravitational lensing

Weak gravitational lensing is a powerful probe of total matter density in the Universe, without distinguishing between dark matter and baryonic matter. The intervening mass distribution between the galaxy and the observer distorts the image of the galaxy. Such small distortions contain rich information about the matter distribution on small and large scales and their evolution over time. Particularly, the dependence of the weak lensing signal on the angular diameter distance, and on the matter power spectrum makes it a great tool to probe the dark energy [23]. A

detailed overview of the basics of of weak gravitational lensing and its applications in cosmology is given in [24].

Two particularly powerful aspects of the method are that it is based on a geometrical observable, i.e. the distorted shapes of galaxy images, and that it is sensitive to the gravitational potential of structures, without distinguishing between baryonic and dark matter. Particularly, weak lensing signal is sensitive to the matter power spectrum over a redshift range and so it provides a measure of the growth rate of large scale structures. Because the growth rate depends strongly on the w parameter in the equation of state of the dark energy, cosmic shear can be used to study different dark energy models. It is also sensitive to the angular diameter distance (through ω factor in equation 9) and so it can also be used to determine angular diameter distances as a function of redshift and constrain w parameter with the distance-redshift relation. These two aspects make the cosmic shear measurements a great tool to probe the dark energy.

In the following, I quickly overview some of the ongoing weak lensing surveys and their results:

- **Kilo Degree Survey**: Kilo Degree Survey (KiDS) [25] is a wide-field optical imaging survey which covers 1500 square degree of the southern hemisphere sky started on October 2011. KiDS uses OmegaCam wide-field camera on 2.6-m VLT Survey Telescope (VST) of Paranal observatory in Chile. This survey was designed to measure the galaxy population out to redshift ~ 1 and to measure the effect of weak lensing by structures along the line of sight on galaxy shapes. Full cosmological analysis of 450 square degree of KiDS data (KiDS-450) including constraints on various cosmological parameters and extensions to the standard ΛCDM model are given in [26], [27] and [28]. Particularly, they found $w < -0.24$ With the KiDS data alone, but the constraint from the combined probe is $w < -0.73$ (all at $\%95$ CL).

- **Dark Energy Survey**: Dark Energy Survey (DES) is an optical galaxy survey conducted using the 570 Megapixel DECam instrument mounted on 4-m Blanco Telescope in Chile. DES is planned to map 5000 square degree of the sky and the main goal of this survey is to study the dark energy using several techniques, including weak lensing. DES collaboration have used year-one shape catalogs to study the shape of 26 million galaxies within the redshift range 0.2 to 1.3 over 1321 square degree of the sky and they constrained various cosmological parameters in ΛCDM model [29], particularly, the dark energy equation of state parameter $w = -0.95^{0.33}_{-0.39}$. They concluded that there is no disagreement between DES results and CMB data, and there is no evidence for a wCDM model with w deviating from -1.

- **HSC-SSP Survey**: Hyper Suprime-Cam Subaru Strategic program (HSC-SSP) is an imaging survey on the 8.2-m Subaru telescope at Mauna Kea Observatory in Hawaii. The goal of this survey is to probe the nature of dark energy and dark matter by various techniques, including weak lensing [30]. HSC-SSP is the deepest ongoing weak lensing survey. So although it's narrower than DES, but its S/N ratio is higher at high redshift, enabling a better measurement of the equation-of-state of dark energy at higher redshift. The cosmological constraints from cosmic shear power spectra with the first year data of HSC is recently published in [31]. Data Release 1 (DR1) of HSC survey was publicly released in February 2017 and it is based on data taken using 61.5 nights between March 2014 and November 2015 [32]. The w parameter in the equation of state of dark energy is not well constrained from shear analysis alone, $w = -1.37^{+0.43}_{-0.36}$. However, this constraint is still consistent with other observations and particularly it is compatible with $w = -1$ which corresponds to the simplest form of dark energy, a cosmological constant.

1.3 CHIME

1.3.1 21-cm intensity mapping

This section is based on Pritchard & Loeb (2012) [33].

An alternative to the galaxy-redshift surveys for the measurements of BAO scale is the **21-cm intensity mapping**. The basic idea of 21-cm intensity mapping is to use 21-cm emission line to map the hydrogen contents of the universe across some redshift range. 21-cm line arises from the hyperfine transition between the triplet and singlet levels in the ground state of neutral hydrogen atom. When such a transition occurs, a photon will be emitted (or absorbed) with a frequency $\nu = 1420\ \mathrm{MHz}$ which corresponds to a wavelength of $\lambda \sim 21\text{-cm}$. The probability of a spontaneous transition from triplet to singlet state in neutral hydrogen is very small and such transition can never be seen in the labs. However, the total number of neutral hydrogen atoms in the intergalactic medium is so large that the emission (or absorption) line can be observed by radio telescopes. Expansion of the universe changes the the observed frequency ν of 21-cm line. So the 21-cm emission or absorption from an object at redshift z will be observed at:

$$\nu = \frac{1420}{1+z}\mathrm{MHz}. \tag{1.13}$$

Therefore, there is no need for spectroscopy to obtain the redshift in this method. Instead, the frequency bin at which the sky is mapped in 21-cm directly gives the

redshift.

The intensity of photons emitted through the hyperfine transition of neutral Hydrogen atom is determined by the spin temperature T_s which is defined by the equation:

$$\frac{n_1}{n_2} = 3\exp\Big(-\frac{T_*}{T_s}\Big), \tag{1.14}$$

where $n1$ and $n2$ are the number densities of electrons in the triplet and singlet states respectively, 3 is the degeneracy ratio between the triplet and singlet levels, $T_* \approx 0.068\,\mathrm{K}$ is the temperature corresponding to the energy difference between the triplet and singlet states, and the spin temperature T_s is the excitation temperature of 21-cm line.

We want to use the 21-cm line as a probe of a hydrogen gas cloud with optical depth τ along the line of sight. The radiative transfer equation in the Rayleigh-Jeans limit for the gas along the line of sight is:

$$T_b = T_{CMB}e^{-\tau} + T_s(1-e^{-\tau}), \tag{1.15}$$

where T_b is the observed brightness temperature. I have assumed that in the background there are only CMB photons. The interpretation of this equation is easy: The CMB photons are exponentially attenuated by the gas cloud (hence we multiply T_{CMB} by $e^{-\tau}$) and the excitation temperature T_s releases 21-cm photons. Then, part of the 21-cm emission is absorbed by the gas.

The observed differential brightness temperature is:

$$\delta T_b = \frac{T_b - T_{CMB}}{1+z} = \frac{(T_s - T_{CMB})(1-e^{-\tau})}{1+z} \approx \frac{(T_s - T_{CMB})\tau}{1+z}. \tag{1.16}$$

Therefore, 21-cm signal is observable only if the the spin temperature deviates from the background temperature (otherwise $\delta T_b = 0$). We can see the 21-cm signal as emission against CMB if $T_s > T_{CMB}$ ($\delta T_b > 0$) or absorption if $T_s < T_{CMB}$ ($\delta T_b < 0$). In this case the 21-cm feature is seen as a spectral distortion to the background CMB and the diffuse distortion of the background can be studied in a similar way as the CMB anisotropies. Then, observations at different frequencies probe different redshift slices of the observable universe.

1.3.2 Radio Frequency Interference

A phenomenon which is problematic not only for 21-cm intensity mapping experiments, but for any research in radio astronomy is the Radio Frequency Interference (RFI) which is the subject of this thesis. RFI is an unwanted human made signal that interferes with the astronomical signal, degrades the quality of astronomical

data and leads to data loss. RFI is usually non-Gaussian and come from TV stations, cell phones, airplanes, satellites or any other activity which produces a radio signal in a frequency that the radio telescope is working. RFI can be considered as *radio pollution* in radio astronomy, similar to the *light pollution* for the optical observations.

Some radio frequencies that are very important for astronomical research are protected by regulations. For example, 1420 MHz (21-cm HI line). However, a radio telescope like CHIME covers a very broad bandwidth and such bandwidth cannot be fully protected. Therefore, radio astronomers have to deal with the RFI. There is no universal method for RFI mitigation. In the case of CHIME, we use fourth statistical moment (kurtosis) to detect non-gaussianity in a signal. The theoretical foundation of the kurtosis based RFI mitigation is given in the chapter 2.

1.3.3 CHIME instrument

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is a transit radio telescope which scans the northern sky across 400-800 MHz band. The telescope has no moving part and the sky is scanned over a sidereal day as it is drifting east to west. CHIME is located at Dominion Radio Astrophysical Observatory (DRAO) near Penticton, BC, Canada. It is originally designed to constrain the dark energy equation of state by measuring the BAO scale across the redshift range $0.8 \leq z \leq 2.5$. This is the redshift range where the dark energy began to dominate the total energy density of the Universe. We use the 21-cm emission line to map the hydrogen contents of the Universe over the required redshift range.

CHIME consists of 4 cylinders with axes oriented along the north-south direction, each with a length of 100 meters and a diameter of 20 meters. There are 256 dual polarization feeds along the focal line on each cylinder, i.e. 2048 inputs can receive the signal at the same time. The signal from the sky is focused by the cylinders onto the feeds and it is amplified by the low noise amplifier (LNA) which is connected to the output of the feed. The analog signal then is transmitted by a 50 m coaxial cable to a receiver hut, where it is further amplified by another set of amplifiers, filtered by a $400 - 800$ MHz bandpass filter and sampled every 1.25 ns and quantized to 8 bits. CHIME correlator has an FX design, i.e. the Fourier transform is done before spatial cross multiplication. The F-Engine is responsible for sampling the analog signal with 8 bits, Fourier transforming every 2048 time samples, and channelizing them by dividing the 400 MHz bandwidth into 1024 frequency channels. This process is done for all 2048 inputs. The real and imaginary parts of the complex-valued data from each frequency channel are then separately quantized to 4 bits (4-bits real and 4-bits imaginary). The resulting 2.56 $\mu\mathrm{s}$ samples from 2048 inputs fed into the X-Engine. There are 256 GPU nodes on X-Engine

Figure 1.3: CHIME telescope. It is an array of 4 cylinders in the east-west direction, with axes oriented in the north-south direction.

where each node processes the data 4 frequency bins. In other words, a GPU node does the cross multiplication of 2048 inputs for 4 frequency channels. Before the correlation operations, the real-time RFI excision system mask the RFI contaminated samples by generating the spectral kurtosis estimates every 0.65 ms for each frequency bin . Note that every 0.65 ms one kurtosis estimate per frequency bin is generated for the **whole array** (not for the individual inputs). Then the samples are integrated to 30 ms and the signal of each input is correlated with the signals of all other inputs to make N^2 correlation matrix. Correlation is the process of cross multiplication of each input with the complex conjugate of all other inputs. Then a second stage RFI excision is applied on 30ms frames to excise lower power, but longer duration RFI events. The theory of kurtosis based RFI excision and the details of the detection algorithms are described in chapters 2 and 3. After second stage of RFI excision, samples are integrated for ~10 seconds. The correlation product is called visibility, V_{ij}:

$$V_{ij}(t,\nu) = \langle E_i E_j^* \rangle, \tag{1.17}$$

where $V_{ij}(t, \nu)$ is the cross correlation of input i with input j at time t and frequency ν, E_i and E_j are the signals received by feeds i and j, respectively, and $\langle . \rangle$ stands for the average over 10 second. The correlation products together with some other information are written to various HDF5 files. In this thesis I evaluate the performance of the real time RFI excision system using the files containing autocorrelations (the correlation of one input with itself). The autocorrelation files include the autocorrelations of all inputs over 10 s, excision rate for 10 s samples. These files are written to an HDF5 file every $\sim$ 43 minutes.

1.3.4 RFI environment at DRAO

The mountains around the observatory at DRAO shield the site from RFI from nearby cities. But a significant portion of the CHIME frequency band is still contaminated by satellites, airplanes, wireless communication and TV broadcasting bands. This includes LTE band between 730 MHz and 755 MHz range, a few TV station bands between 500 MHz to 580 MHz, and a lot of RFI lines between 400 MHz to 500 MHz, including UHF repeaters around 450 MHz. These features are visible in the CHIME data (figure 1.4). Besides cell phone and TV station bands that are static in nature, there are many sources of intermittent RFI events such as satellites and airplane. Moreover, the atmosphere can be ionized by the meteors entering to it. The ionized region of the atmosphere acts as a reflector for the distant ground based RFI sources, such as distant TV stations. These scattering events typically appear as ~6 MHz wide bursts and last for a few seconds. One can see this type of RFI in the figure 1.4 between 480 MHz and 520 MHz around 00:15 PT.

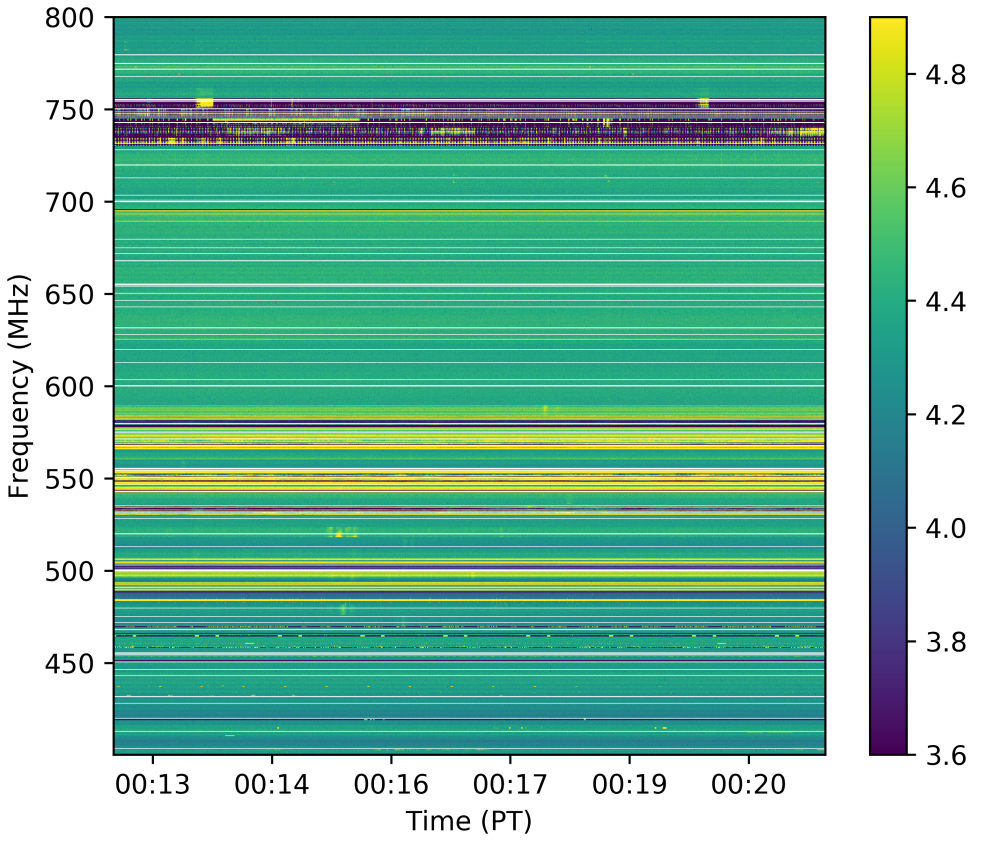

Figure 1.4: The autocorrelations of 30 ms samples for one feed. We can see various RFI features in this plot, namely LTE band from ~ 730 MHz to 755 MHz, TV station bands around 500-600 MHz, and repeating RFI spots between 400 MHz and 500 MHz. 6 MHz wide bursts from distant broadcast TV bands appear as blob around 00:15 PT in 480 MHz or 520 MHz. White lines correspond to the missing frequencies due to the malfunctioning of some of the GPU nodes.

Chapter 2

Kurtosis-based RFI excision

In this chapter I review some definitions from statistics. Then the spectral kurtosis concept is introduced and spectral kurtosis estimator is derived. I will show that the statistical properties of the estimator can be used to detect non-Gaussian features of a signal in a Gaussian background. Most of this chapter is based on [34], [35], [36] and [1]. Note that in this chapter, I will use the words *antennas*, *receivers* and *inputs* interchangeably.

2.1 Statistics

2.1.1 Moments of a probability distribution

Moments of a probability distribution are the expectation values of a random variable to integer powers and they often give valuable information about the distribution of the random variable. The r^{th} moment of a probability distribution $P(x)$ of a random variable X is defined as:

$$E(x^r) = \int_{-\infty}^{\infty} x^r P(x) dx, \tag{2.1}$$

where x is the value of the random variable X. The first moment of a probability distribution is called the mean μ. The moments can also be derived from the so called *moment generating function*. Moment generating function of a random variable X is defined as:

$$MGF_X(t) = E(e^{tx}), \tag{2.2}$$

where t is a real-valued number. By expanding the exponential function and taking the expectation value, it can easily be shown that the moments of $P(x)$ are related to the derivatives of moment generating function at t=0:

$$E(x^r) = \frac{d^r}{dt^r} MGF_X(t)|_{t=0} \tag{2.3}$$

The central moment is the moment around the mean:

$$E[(x-\mu)^r] = \int_{-\infty}^{\infty} (x-\mu)^r P(x) dx. \tag{2.4}$$

Clearly, the second central moment is the variance σ^2 of $P(x)$. Lastly, one can define the normalized moment (commonly known as standardized moment) in the following way:

$$\tilde{\mu}_r = \frac{E[(x-\mu)^r]}{\sigma^r}, \tag{2.5}$$

where $\sigma^r = \left(\sqrt{E[(x-\mu)^2]}\right)^r$ is the $r^t h$ power of the standard deviation of $P(x)$.

The standardized moment facilitates the comparison of the shape of probability distributions. For example, third standardized moment (skewness) is a measure of asymmetry of a distribution about its mean. The quantity of interest for this work is the fourth standardized moment, known as kurtosis. Kurtosis measures the heaviness of the tails of a distribution. From equation 2.5 the *moment-based* kurtosis is:

$$\begin{aligned} \tilde{\mu}_4 &= \frac{E[(x-\mu)^4]}{\sigma^4} \\ &= \frac{E[(x-\mu)^4]}{(E(x^2)-\mu^2)^2}. \end{aligned} \tag{2.6}$$

One can evaluate the moment-based kurtosis for a Gaussian signal with mean zero, using the following Gaussian integrals:

$$E(x^4) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^4 e^{\frac{-x^2}{2\sigma^2}} = 3\sigma^4, \tag{2.7}$$

$$E(x^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2 e^{\frac{-x^2}{2\sigma^2}} = \sigma^2. \tag{2.8}$$

Therefore, the moment based kurtosis of a Gaussian signal with mean zero is 3.

2.1.2 Cumulant

An alternative to the moments of a probability distribution is the cumulant. Similar to the moments, cumulants can be used to characterize the statistical properties of a probability distribution. It is defined by the *cumulant generating function* $CGF(t)$ which is the natural logarithm of the moment generating function:

$$CGF(t) = \ln E(e^{tx}). \tag{2.9}$$

The power series of this function is:

$$CGF(t) = \sum_{r=1}^{\infty} k_r \frac{t^r}{r!}, \tag{2.10}$$

where the coefficients k_r are called cumulants. Using equation 2.10 it can be shown that cumulant of r^{th} order can be derived by taking the r^{th} order derivative of $CGF(t)$ and evaluating it at $t = 0$:

$$k_r = \frac{d^r}{dt^r} CGF(t)|_{t=0}. \tag{2.11}$$

The first three cumulants are same as the first three moments. However, one can show that the cumulant-based kurtosis for a real random variable x with mean zero is:

$$k_4 = E(x^4) - 3E^2(x^2), \tag{2.12}$$

which is different from equation 2.6. Using the Gaussian integrals given in equations 2.7 and 2.8 it can simply be shown that the cumulant-based kurtosis for a Gaussian signal with mean zero is zero.

2.2 Spectral kurtosis

Spectral kurtosis (SK) is a statistical tool which can be used to identify the non-Gaussian behaviour of a signal in frequency domain. A recent cumulant-based definition of the spectral kurtosis for a circularly symmetric random variable in the frequency bin k whose real and imaginary parts of the DFT have zero mean is given by [34]:

$$SK[X_k] = \frac{k_4(x_k, x_k^*)}{(k_2(x_k, x_k^*))^2} = \frac{E(|x_k|^4) - 2E(|x_k|^2)^2}{E(|x_k^2|^2)}. \tag{2.13}$$

Note that the expression for k_4 for a complex random variable is different from a real random variable (equation 2.12). Also, the spectral kurtosis is normalized by the second order cumulant k_2, which is same as the second order moment.

To evaluate the spectral kurtosis for a circularly symmetric Gaussian signal with mean zero, note that x_k=X_k+iY_k. So, we will have:

$$E(|x_k|^2) = E(X_k^2) + E(Y_k^2) = 2\sigma^2, \tag{2.14}$$

$$E(|x_k|^4) = E(X_k^4) + E(Y_k^4) + 2E(X_k^2 Y_k^2) = 8\sigma^4, \tag{2.15}$$

where I have used the Gaussian integrals (equations 2.7 and 2.8) and the fact that the real and imaginary parts of the x_k are independent, i.e. $E(X_k^2 Y_k^2) = E(X_k^2)E(Y_k^2)$. Therefore, the spectral kurtosis for a circularly symmetric Gaussian signal with mean zero is **zero**.

Although equation 2.13 gives the general definition for the spectral kurtosis of a circularly symmetric complex random variable with zero mean, it is much better to rewrite it in terms of total power P_k. This is because P_k is an observable quantity.

The total power P_k in a spectral frequency bin f_k is proportional to the absolute square of the frequency domain signal $|x_k|^2$. So, equation 2.13 in terms of power spectral density P_k is:

$$SK[X_k] = \frac{E(P_k^2) - 2E(P_k)^2}{E(P_k)^2}. \tag{2.16}$$

Moreover, the variance σ_k^2 and mean μ_k of power spectral densities are:

$$\sigma_k^2 = E(P_k^2) - E(P_k)^2, \tag{2.17}$$

$$\mu_k = E(P_k). \tag{2.18}$$

So, equation 2.16 can be rewritten in terms of σ_k^2 and μ_k parameters:

$$SK[X_k] = V_k^2 - 1, \tag{2.19}$$

where $V_k^2 = \frac{\sigma_k^2}{\mu_k^2}$ is the normalized uncertainty or the spectral variability of the signal. Equation 2.19 shows that the spectral kurtosis is equivalent to the spectral variability up to a constant. So, we can use the two terms interchangeably. The only difference between $SK[X_k]$ and V_k^2 for a Gaussian signal is that the spectral kurtosis of such a signal is zero from equation 2.13, but its spectral variability is one (equation 2.19). Since measuring the spectral variability is easier than spectral kurtosis and they are equivalent to each other, I use the term "spectral kurtosis" as a synonym for spectral variability in the rest of the thesis.

2.3 Spectral kurtosis estimator

Suppose that we have a set of n complex values x_k which represent the post Fourier-transform time-stream for a single frequency channel. The power spectral density estimator ($\widehat{P}_k$) is proportional to $|x_k|^2$. Based on the discussion in the previous section, the unbiased spectral kurtosis estimator is defined as:

$$\widehat{SK} \equiv \widehat{V}_k^2 = \frac{\widehat{\sigma}_k^2}{\widehat{\mu}_k^2}, \tag{2.20}$$

where $\widehat{\sigma}_k^2$ and $\widehat{\mu}_k$ are the unbiased variance and mean estimators of the $\widehat{P}_k$ for the frequency bin k, respectively:

$$\widehat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} \widehat{P}_{k,i} \tag{2.21}$$

$$\widehat{\sigma}_k^2 = \frac{1}{n-1}\sum_{i=1}^{n} (\widehat{P}_{k,i} - \widehat{\mu}_k)^2. \tag{2.22}$$

In order to simplify the final expression for the spectral kurtosis estimator, the following parameters are defined:

$$S_1 = \sum_{i=1}^{n} \widehat{P}_k \quad , \quad S_2 = \sum_{i=1}^{n} \widehat{P}_k^2. \tag{2.23}$$

Therefore, mean and variance estimators in terms of S_1 and S_2 are:

$$\widehat{\mu}_k = \frac{S1}{n} \quad , \quad \widehat{\sigma}_k^2 = \frac{1}{n(n-1)}(nS_2 - S_1^2). \tag{2.24}$$

And the spectral kurtosis estimator $\widehat{SK}$ would be:

$$\widehat{SK} = \Big(\frac{n}{n-1}\Big)\Big(n\frac{S_2}{S_1^2} - 1\Big) \tag{2.25}$$

The expected value of the SK estimator for a circularly symmetric Gaussian random variable must be one. But as it is shown by [36] this is not the case for the estimator defined by equation 2.25:

$$E(\widehat{SK}) = \frac{n}{n+1}. \tag{2.26}$$

Therefore equation 2.25 is a biased estimator. To get the unbiased estimator whose expectation value is one, one can rescale the estimator by multiplying it by $n+1/n$:

$$\widehat{SK} = \Big(\frac{n+1}{n-1}\Big)\Big(n\frac{S_2}{S_1^2} - 1\Big) \tag{2.27}$$

Nita & Gary (2010) [36] have shown that for x_i being drawn from a circularly-symmetric complex Gaussian distribution, SK estimator has the following statistical properties to the first order in n:

$$E(\widehat{SK}) = 1 \tag{2.28}$$

$$\text{Variance}(\widehat{SK}) = \frac{4}{n} + O(\frac{1}{n^2}). \tag{2.29}$$

Since the RFI events are mostly non-Gaussian [37], one can use the above properties to detect and mitigate the RFI. The mean of the spectral kurtosis estimator is invariant. So the sensitivity of the estimator to the RFI events is determined by its variance. To see the statistical properties visually, I generated 30000 SK estimates, each from 20000 complex random samples drawn from a circularly symmetric Gaussian distribution. Figure 2.1 shows the result of the simulation. The cumulative plot shows that half of the SK estimates are below 1 and the other

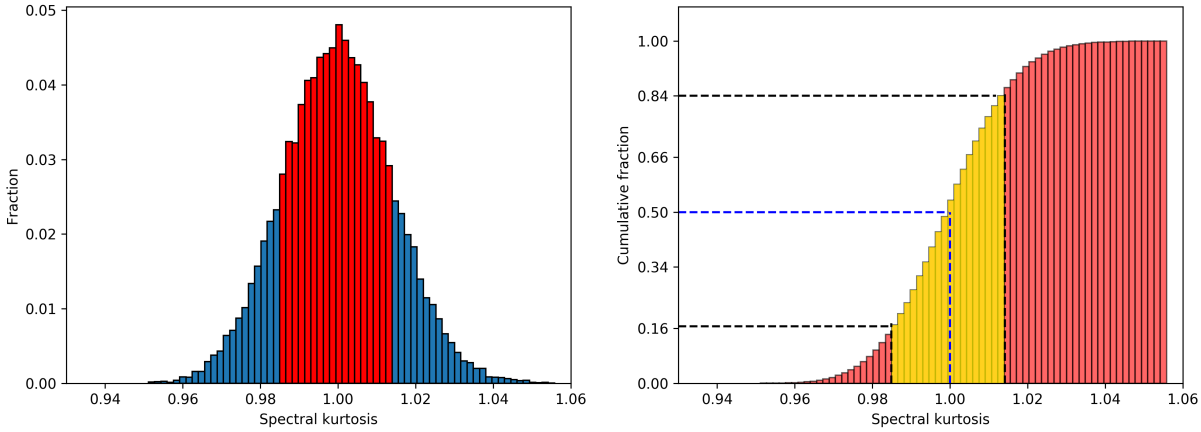

Figure 2.1: Left: Histogram of 30000 realizations of SK estimator. Each SK value is generated from 20000 complex random variables drawn from a Gaussian distribution. Red region indicates 1σ interval where σ is the theoretical standard deviation. Right: Cumulative histogram of the SK values. Yellow region corresponds to 1σ interval. The SK distribution is symmetric around 1, and %68 of the estimates lie within one standard deviation. These two confirm the derived mean and variance of the SK estimator.

half are above 1. This confirms that the mean value of the SK estimator is unity. Moreover, 68% of the SK estimates lie within 1σ (yellow region), where σ is the theoretical variance given by the equation 2.29. A non-Gaussian RFI changes the shape of the histogram, for example the tails of the histogram will be longer. So one can simply detect non-Gaussian part of the signal by setting a threshold, e.g. at 5σ, and mask all the samples whose spectral kurtosis go beyond the threshold.

2.4 Spectral kurtosis for a compact array

Equation 2.29 shows that the discrimination potential of the kurtosis estimator depend on the number of samples that are being used to generate the kurtosis estimates. To decrease the variance of the estimator one needs to integrate more samples in a frequency bin, which in turns decrease the cadence at which the kurtosis estimates are being produced. The output of CHIME data are 10-second samples. Our aim is to flag RFI at a very **short timescale** (i.e. high cadence), so that 10-second samples become RFI free. To increase the size of the input dataset

without increasing the timescale at which the kurtosis estimates are generated, one can combine the **simultaneous** measurements from multiple antennas in an interferometer. Combining measurements from different antennas requires the following two conditions: First, the array must be compact; i.e. the time it takes for a signal to go across the array must be smaller than the timescale at which the kurtosis estimates are generated. Second, receivers must provide independent samples. Correlated samples reduce the effective integration length and increase the variance of the estimator. I discuss these two conditions for CHIME in section 2.5.

One way to combine the measurements from different antennas is to take the average spectral kurtosis of all antennas. For a compact array with N independent antennas in which spectral kurtosis estimates are produced every n time-samples for each antenna, the average spectral kurtosis can be found in the following way: Using equation 2.27, the spectral kurtosis for the ith antenna is:

$$\widehat{SK}_i = \Big(\frac{n+1}{n-1}\Big)\Big(n\frac{S_{2,i}}{S_{1,i}^2} - 1\Big) \tag{2.30}$$

where $(S_1)_i$ and $(S_2)_i$ are defined by the equation 2.23 for the ith antenna. Then, spectral kurtosis estimator averaged over all antennas would be:

$$\widehat{SK}_{array} = \frac{1}{N}\sum_i^N \widehat{SK}_i \tag{2.31}$$

Using the fact that the mean of the constituent SK_i is unity, the average spectral kurtosis estimator for the whole array $\widehat{SK}_{array}$ is:

$$\Big\langle \widehat{SK}_{array} \Big\rangle = \frac{1}{N}\sum_i^N \Big\langle \widehat{SK}_i \Big\rangle = \frac{N}{N} = 1, \tag{2.32}$$

And since the total integration length is nN, the variance of the estimator is:

$$\text{Variance}(\widehat{SK}_{array}) = \frac{4}{nN} + O(\frac{1}{n^2N}) \tag{2.33}$$

where I used the the statistical properties of constituents SK_i from equations 2.28 and 2.29. Comparing equations 2.29 and 2.33, it is obvious that by combining the estimates from all antennas, the variance of the estimator is reduced by a factor of $1/N$.

There is an alternative formulation for the estimator to reduce the total number of operations. The idea is to redefine $S_{1,i}$ and $S_{2,i}$ parameters by normalizing them by the mean power of the ith receiver μ_i:

$$\mu_i = \frac{\sum_{j=1}^n P_{k,j}}{n} = \frac{S_{1,i}}{n} \tag{2.34}$$

Then, the new parameters S_1' and S_2' are defined by summing the normalized $S_{1,i}$ and $S_{2,i}$ over all the receivers:

$$S_1' = \sum_{i=1}^{N} \frac{S_{1,i}}{\mu_i} = \sum_{i=1}^{N} S_{1,i} \frac{n}{S_{1,i}} = nN, \tag{2.35}$$

$$S_2' = \sum_{i=1}^{N} \frac{S_{2,i}}{\mu_i^2} = \sum_{i=1}^{N} \frac{S_{2,i}}{S_{1,i}^2} n^2 \tag{2.36}$$

Using equation 2.30 one can rewrite S_2' in terms of individual $\widehat{SK}_i$:

$$\begin{aligned} S_2' &= n \sum_{i=1}^{N} \left(\frac{n-1}{n+1} SK_i + 1\right) \\ &= \frac{n(n-1)}{n+1} \sum_{i=1}^{N} SK_i + nN. \end{aligned} \tag{2.37}$$

Then the spectral kurtosis estimator will be:

$$\widehat{SK} = \frac{nN+1}{nN-1}\left(nN \frac{S_2'}{S_1'^2} - 1\right) \tag{2.38}$$

$$= \frac{nN+1}{nN-1}\left(nN \frac{S_2'}{(nN)^2} - 1\right) \tag{2.39}$$

This is obviously a biased estimator, because its expectation value is not unity. The unbiased estimator can be found by a simple re-scaling:

$$\widehat{SK} = \frac{n+1}{n-1}\left(\frac{S_2'}{nN} - 1\right) \tag{2.40}$$

One can check that the expectation value of this estimator is one and its variance is same as the one in equation 2.33. The advantage of this alternative formulation is that it requires slightly fewer number of operations to estimate the SK value [1].

2.5 Implementation for CHIME

As mentioned in the section 1.3.3, each of 256 X-Engine GPU nodes process four frequencies. The spectral kurtosis is estimated on each GPU node by accumulation of 256 time-samples from all inputs $N \sim 2048$ separately for each of the four frequencies . So, we would have one kurtosis estimate for each frequency channel every $256 \times 2.56\mu$s= 0.65 ms for the whole array. Since the expectation value and the variance of the SK estimator for a Gaussian population are known, one can

set a detection threshold to identify the samples that are contaminated by the RFI. Then signals from every feed for all 256 samples are removed when the SK value lies out of bounds. Different RFI detection algorithms are discussed in detail in the chapter 3. Note that the nominal variance in the case of CHIME ($n = 256$ and $N \sim 2048$) is:

$$\sigma^2 \approx \frac{4}{Nn} \approx 7.63 \times 10^{-3}. \tag{2.41}$$

As already mentioned, two conditions are needed for combining the samples from various antennas to increase the integration length: First, the cadence at which SK estimates are generated must be shorter than the time the signal needs to travel across the array. This condition is satisfied for CHIME: The maximum distance between the antennas is ≈ 100m, so any signal arrives at all antennas within $\approx 0.3\mu s$. This is much smaller than the cadence at which the kurtosis estimates are generated (0.65 ms). So the RFI appears simultaneous at all the antennas. The second condition is that the samples from different inputs (that are combined to increase the integration length) must be independent. This is normally the case unless there is a bright astrophysical radio source in the CHIME beam or an RFI. This issue is discussed in detail in the section 4.2.

Chapter 3

High Cadence Excision Algorithms

In this chapter, different excision algorithms that are used in CHIME pipeline are introduced, and the performance of the algorithms is evaluated by applying them to an offline dataset.

The idea for the high cadence excision is the following: The mean value of CHIME visibilities are archived with 10 s cadence. But transient RFI events are much shorter than 10 seconds. They usually last from a few milliseconds to 1-2 seconds. This means that a tiny fraction of a 10 second sample is contaminated by the RFI, but they are so bright that they saturate the whole sample. Therefore, performing the excision at a high cadence can detect and remove those short events. So instead of removing the whole 10-second sample which is saturated by the RFI, we remove only the part which is being contaminated. Then the output of 10 second sample which is free from RFI will be written on the disk.

3.1 Excision algorithms

As mentioned in the section 2.5, one SK estimate per frequency bin is generated every 0.65 ms by combining the measurements from different antennas. In chapter 2 I showed that the expectation value and the variance of the SK estimator are known for a Gaussian signal: the mean value of the estimator is invariant $\langle SK \rangle = 1$ and the variance is $\sigma^2 \sim (0.0027)^2$ for CHIME. In general, a non-Gaussian signal causes the SK estimator to deviate largely from the expected value with respect to its variance. So a non-Gaussian RFI can be detected by setting a threshold. If the SK estimate of the signal exceed that threshold, it will be flagged as RFI. In the following I will describe different excision algorithms.

3.1.1 Single Stage

In the single stage algorithm, we mask all 0.65 ms samples whose SK value exceed $1 \pm n\sigma$, where n is the detection threshold. We set $n = 5$, because according to [36], if the SK estimates are generated from Gaussian signals the distribution of

SK estimates is also Gaussian. This means that the probability of the SK estimate of a Gaussian signal being beyond 5σ limit is less than 10^{-6}. So a sample whose SK value is beyond 5σ is most probably RFI.

3.1.2 Excision on 30 ms frames (EOF)

Suppose that we have ~500000 SK values that are estimated from a Gaussian signal. Now I add a few values (representing RFI) that change the shape of the Gaussian. The situation is shown in the figure 3.1. Single stage algorithm is only sensitive to the powerful RFI events whose SK values exceed 5σ limit. But it is insensitive to the lower power RFI events. Lower power RFI appears as non-Gaussian features that are closer to the mean value; for example the bumps around 2σ in the figure 3.1. Obviously, if we lower the threshold of the single stage algorithm to detect lower power RFI, the Gaussian part of the signal will be lost as well. So, how can we make an algorithm which is sensitive to those features without losing too much of data?

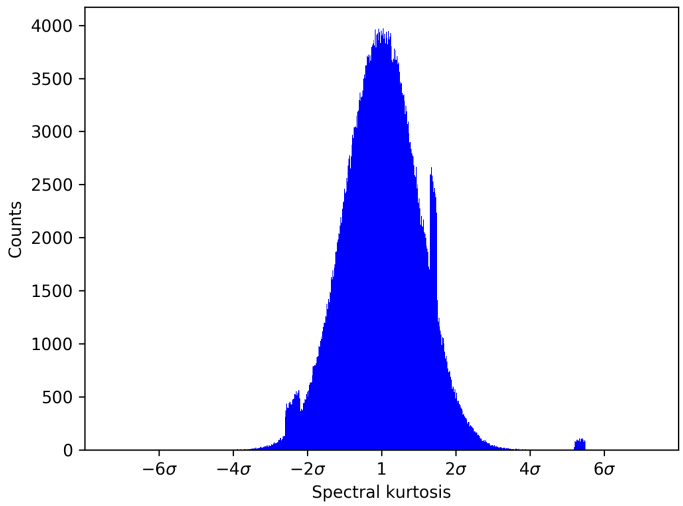

Figure 3.1: Histogram of ~ 500000 random numbers drawn from a Gaussian distribution with mean 1 and some variance σ^2. These values represent the RFI-free SK values. RFI events change the shape of the RFI-free SK distribution. So, a few non-Gaussian features are added to the the distribution to represent the RFI. The single stage algorithm is sensitive to the features beyond 5σ, while EOF is sensitive to lower power, but longer duration RFI.

One way is to make a frame by accumulating M individual 0.65 ms SK estimates, and if more than a few percent of the samples in the frame exceed some threshold, mask the whole frame. In the CHIME pipeline, M=48 individual samples are collected to make a 30 ms frame (48×0.65ms. $\sim 30\mathrm{ms}$). Now, if more than a few percent of the samples in the 30 ms frame exceed some threshold $1 \pm n\sigma$, the whole frame is masked. So, in this algorithm a set of 2 parameters (n, f) defines the RFI detection threshold, where f is the fraction of the samples in a 30 ms frame whose SK values exceed $1 \pm n\sigma$.

To choose appropriate values for n and f parameters, I generated random numbers from a Gaussian distribution whose statistical properties are similar to the free-RFI SK estimator for CHIME (i.e. a normal distribution with mean 1 and variance $(0.0027)^2$). Then, 48 of the samples are integrated to form a $30\mathrm{ms}$ frame. A set of thresholds with different (n, f) parameters are applied to the Gaussian 30 ms frames, and the probability for the frames to pass the EOF algorithm without being masked is computed. For example, for $(n = 2, f = 0.3)$, I compute the probability that 30 ms frames are not flagged by the EOF algorithm, that is, less than 30% of the samples in a frame (whose size is 48) exceed $1 \pm 2\sigma$. We try to keep this probability high enough, for example around 99.9999%. In this case, only 1 in a million of Gaussian frames will be masked. This probability is equivalent to the probability for a Gaussian random variable to fall inside 5σ interval. The result is shown in the figure 3.2 in terms of the probability in the unit of σ interval. We choose the parameters so that they are equivalent to 7σ threshold.

3.1.3 Two stage

The two stage algorithm, is nothing but the combination of the single stage and EOF algorithms. In this algorithm, all the individual 0.65 ms samples whose SK value exceed $1 \pm 5\sigma$ are removed. So the first stage of excision is similar to the single stage algorithm. Then, 48 samples are integrated to form a 30 ms frame and a second stage excision with the EOF algorithm is performed on these frames.

In comparison to the EOF, two stage algorithm is more sensitive to the very short RFI. Suppose that there are less than 13% samples in a 30 ms exceeding 3σ, in this case EOF does not remove any of those samples, but the first stage excision of the two stage algorithm does remove those samples. It is also more sensitive to RFI whose SK values are closer to the mean value. In other words, two stage algorithm is both sensitive to the powerful, short time-scale RFI and to the lower power, but longer lasting RFI.

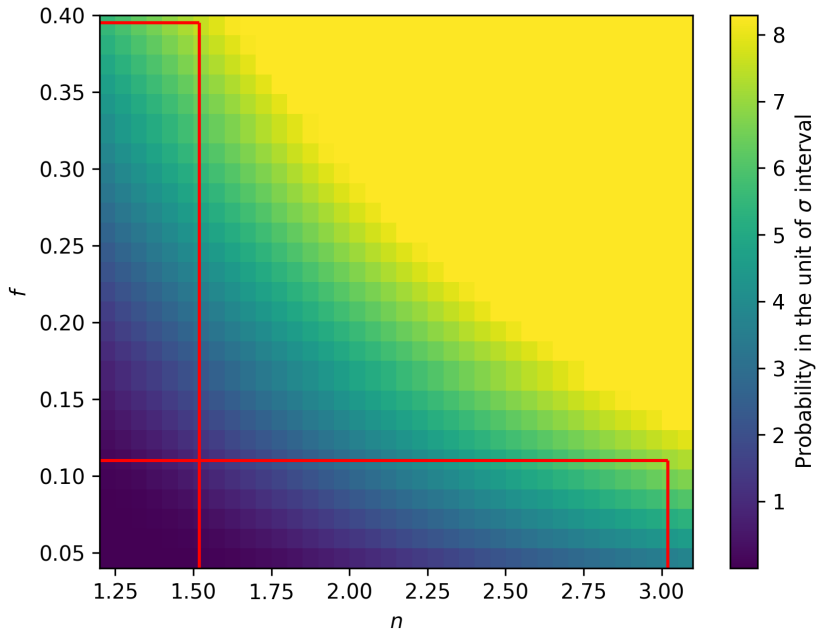

Figure 3.2: The probability for a Gaussian 30 ms frame to pass the EOF algorithm without being removed for different values of (n, f) parameters. For example it is very probable for a 30 ms frame to not to be masked by EOF algorithm with $(n, f) = (3, 0.4)$. In this case the probability for the frame to pass the algorithm is more than 8σ. The two red lines show two thresholds that we have used used for RFI detection between December 2019 and February 2020 using EOF algorithm.

3.2 Evaluation of excision algorithms

3.2.1 Gaussianity test

Gaussianity test is our main tool to assess the quality of the data and to see how close are data are to being Gaussian distributed. It can be used to compare the gaussianity of the data before and after the RFI excision. Note that the signal received by an antenna is mainly composed of three components: astrophysical sky signal, thermal noise and possible RFI. By thermal noise I mean the noise generated by the receiver itself, CMB radiation, galactic background radiation, atmospheric emission, and the receiver noise. Thermal noise and the sky signal are both Gaussian with mean zero, however, the thermal noise varies in a much shorter time scale. The RFI

is usually non-Gaussian. Since sky is slowly varying in time, one can take the difference between the neighboring time samples in autocorrelation products to remove the astrophysical sky contribution:

$$\Delta V_{ii}(t_n, \nu) = V_{ii}(t_{n+1}, \nu) - V_{ii}(t_n, \nu), \tag{3.1}$$

where V_{ii} is the autocorrelation of input i, and n is an integer representing the time sample. Suppose that there is no RFI. Then what remains in ΔV_{ii} is the thermal noise component. As already mentioned, thermal noise is Gaussian and the fluctuation in thermal noise can be reduced by accumulating more samples, i.e. increasing the integration time Δt in a bandwidth B. Then, variance of ΔV_{ii} normalized by average autocorrelation is:

$$\frac{1}{2} Var(\frac{\Delta V_{ii}}{\overline{V}_{ii}}) = \frac{1}{N}, \tag{3.2}$$

where $\overline{V}_{ii}$ denotes the average autocorrelation of neighboring samples and $N = \Delta t \times B$ is the total number of samples in bandwidth B accumulated over time Δt. The 1/2 factor in front of the variance is because we are taking the difference of two random variables that are drawn from the same Gaussian distribution. A large deviation from this relation is a signature of non-Gaussian RFI. Gaussianity test is the process of checking our data to see if the equation 3.2 is satisfied. I define the Gaussianity test value GT in the following way

$$GT = \sqrt{Var(\frac{\Delta V_{ii}}{\overline{V}_{ii}}) \frac{\Delta t B}{2}}. \tag{3.3}$$

Note that for $GT = 1$ we get the equation 3.2 which is for a Gaussian noise. A large deviation from $GT = 1$ shows that the data is non-Gaussian.

The Gaussianity test is our tool to evaluate the RFI excision system performance and to compare different excision algorithms. For each excision algorithm, I run the gaussianity test on all all the frequencies (and a single input) and by definition, the frequency channels whose GT value is smaller than 2 are considered to be *good* frequency channels, i.e. nearly free from RFI. Therefore, we can compare the number of *good* frequency channels for different algorithms and see which algorithm can detect and mitigate the RFI more effectively.

3.2.2 Offline test

To evaluate the performance of excision algorithm, the variance and SK values of 0.65ms samples for a single input were recorded on August 29, 2019, from 17:00

to 18:00 PT, and on October 10, 2019, from 00:00 to 10:00 PT. Different excision algorithms were applied to the data. Since the mean value of the voltage is zero ($\langle v \rangle$=0)), variance of the voltage gives the autocorrelation of the input:

$$Var(x) = \langle x^2 \rangle - \langle x \rangle^2 = \langle x^2 \rangle, \tag{3.4}$$

where x is the measured voltage. The average of SK values for samples drawn from a circularly-symmetric complex Gaussian distribution is 1. The standard deviation of the SK values are estimated by

$$\sigma \approx \sqrt{\frac{4}{Nn}} \approx 0.0027, \tag{3.5}$$

where $n = 256$ is the number of time samples (2.56 µs cadence) that are collected to generate spectral kurtosis every 0.65 ms, and $N \sim 2048$ is the number of inputs that are used to estimate the SK values. Note that SK values are generated by combining the signals from all the healthy inputs, and the variance files contain the rms of the voltage for a single input. The files are processed in the following way:

1. The excision algorithm is applied to the SK files. For example in the single stage scenario, an upper and lower limit at $1 \pm 5\sigma_{SK}$ is used to mask the SK outliers.

2. This mask is applied to the variance values to remove the samples exceeding the upper and lower limits.

3. The resulting variances are averaged over ~ 8 seconds.

4. Fraction of the excised data (those that exceed the upper and lower limits) are computed for every 8 seconds samples. I call this quantity the excision rate.

5. Those samples that are affected by malfunctioning of the GPU nodes are detected and removed.

6. Variances and excised rate of 8 s time samples are reported.

Unflagged-variances (i.e. no excision) are reported in the same way by doing steps 3 to 6. To see the effect of excision algorithm I run the Gaussianity test on flagged and unflagged data (∼1 hour, 8 second samples) and compare the results. Figure 3.3 show the results of Gaussianity test on the data with the single stage algorithm with a threshold at 5σ (left) and with EOF algorithm with $(n, f) = (1.5, 0.4)$, and compare both with the case with no RFI excision. Each data point corresponds to one frequency channel and the color shows the fraction of excised

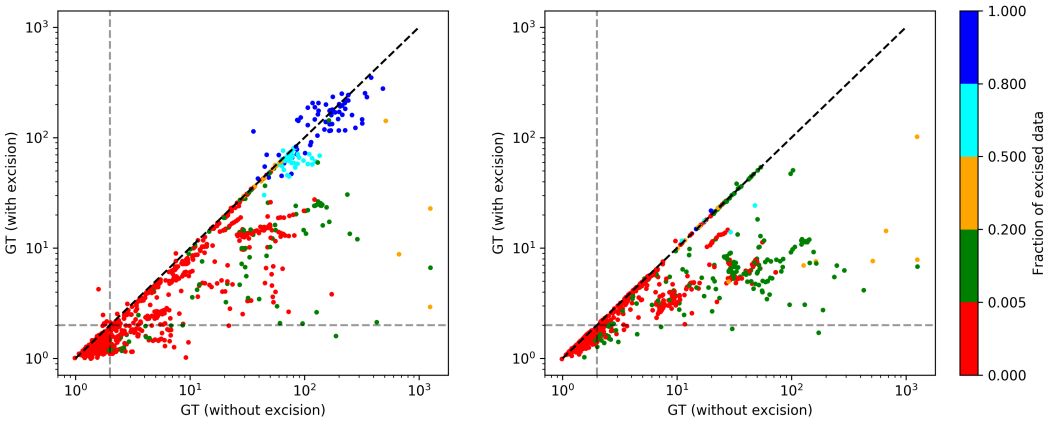

Figure 3.3: Left: Result of gaussianity test without any RFI excision (horizontal axis) and with the single stage algorithm (vertical axis). Right: Gaussianity test with EOF algorithm using $(n, f) = (1.5, 0.4)$. Each data point corresponds to one frequency bin and the color shows the average excised rate at that frequency. In other words, the figure shows three things for each frequency bin: GT value before RFI excision, GT value after applying the excision algorithm, and the average excised rate over $\sim$ 1 hour. The color bar is the same for both figures. Most of the data points are below the diagonal line. This means that Excision algorithm has improved the Gaussianity of those frequencies.

data in an 8 s sample. In figure 3.3, anything below the diagonal line has become more Gaussian after the excision. Note that there is no data point above the diagonal line, except a few that correspond to LTE and TV station bands (mostly blue and cyan points, see figure 3.4)). Before excision, $\sim$ 15% of the data points had a GT value less than 2. After excision, this increases to $\sim$ 34% of frequency channels (those that are below the horizontal gray line). In other words, we have recovered $\sim$ 19% of frequency channels that were initially contaminated by RFI. In most of the frequency channels we can increase the quality of the data by excising **less than 0.5%** of time samples (red points).

Figure 3.4 shows the average excised rate as a function of frequency, and color shows the difference in GT value between unflagged and flagged data . Large positive difference shows an improvement on gaussianity of the data. Cyan and green points are UHF repeaters around 450 MHz. These frequencies are cleaned by excising $< 25\%$ of the samples and their GT values are decreased by a factor of 100 or more. These two figures together show the potential of high cadence

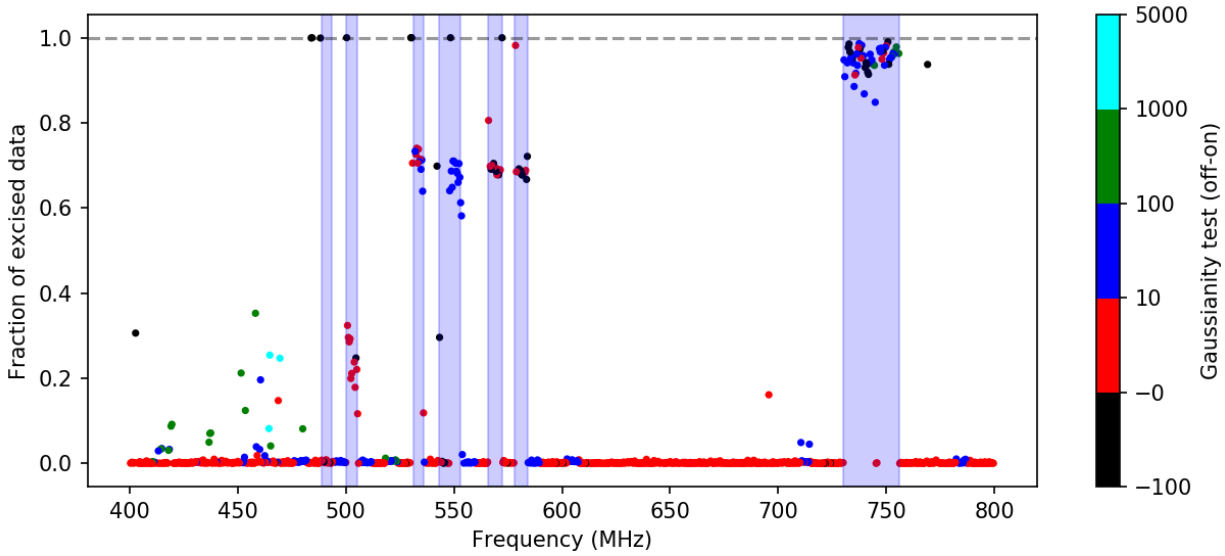

Figure 3.4: Average excision rate for 8 s samples over 1 hour as a function of frequency. 600-700 MHz band is the cleanest part of the spectrum. There are many RFI features in 400-500 MHz band that. The kurtosis based RFI excision system detected those RFI and some frequency channels are cleaned by excising only $\sim 20\%$ of the samples in the integrated 8 second sample.

kurtosis based RFI excision.

To further evaluate the excision algorithms, the parameters of the excision algorithm are changed and the results are compared. The results for three algorithms are shown in the tables 3.1, 3.2 and 3.3. It is clear that two-stage scenario can recover more frequency channels without excising too much data. The results are summarized in figure 3.5. For comparison, note that 288 frequency channels passed the gaussianity test before excision and 387 passed the test using the single stage scenario (at 5σ) while 0.15% of the data are excised. Using two-stage algorithm with $p = 0.4$, $n = 1.4$, 22 more frequency channels pass the gaussianity test with excision of another 0.25%

n	**# of frequencies with** $GT < 2$	**Excised fraction (%),** $GT < 2$	**Excised fraction (%), all frequencies**
5	387	0.15	11.63
4	394	0.19	12.35
3.5	398	0.31	12.83
3	400	0.81	13.69
2.5	408	2.58	15.65
2	408	7.38	20.27

Table 3.1: the results of single stage excision for different values of parameter n on August 29 between 5:00 and 6:00 PM. Second column shows the number of frequency channels having $GT < 2$ after excision. Third and fourth columns show the average fraction of excised data for the frequency channels that passed the test and for all the frequency channels, respectively. As expected, we recover more frequency channels by decreasing the threshold value at the expense of loosing more data. Note that without any RFI excision 288 frequency channels have passed the Gaussianity test.

n	f	**# of frequencies with** $GT < 2$	**Excised fraction (%),** $GT < 2$	**Excised fraction (%), all frequencies**
3	0.13	348	0.35	15.37
2	0.25	341	0.19	15.75
1.5	0.4	327	0.14	15.83

Table 3.2: Results of excision with EOF algorithm. Third column shows the number of frequency channels having $GT < 2$ after excision. The number of good frequency channels ($GT < 2$) is smaller than the single stage algorithm. This is because the single stage algorithm is more sensitive to the very *short* RFI events where only a few samples exceed the detection threshold. EOF algorithm is sensitive to the longer RFI, when more than a few percent of the samples in 30 ms exceed the threshold, but it is insensitive to the very short RFI. The meaning of the third and fourth columns is similar to the ones in the table 3.1.

f	n	**# of frequencies,** $GT < 2$	**Excised fraction(%),** $GT < 2$	**Excised fraction(%), all frequencies**
0.4	1.5	406	0.23	15.69
	1.4	409	0.39	16.24
	1.3	414	1.18	17.41
	1.2	417	4.16	20.54
	1.1	419	12.75	28.59
0.5	1.5	401	0.17	14.85
	1.4	403	0.18	15.05
	1.3	405	0.20	15.33
	1.2	410	0.29	15.77
	1.1	413	0.85	16.72

Table 3.3: The results of gaussianity test for the two stage algorithm. I used 5σ threshold for the first stage, and different values of (n, f) for the second stage. The meaning of the third and fourth columns is similar to the ones in the table 3.1.

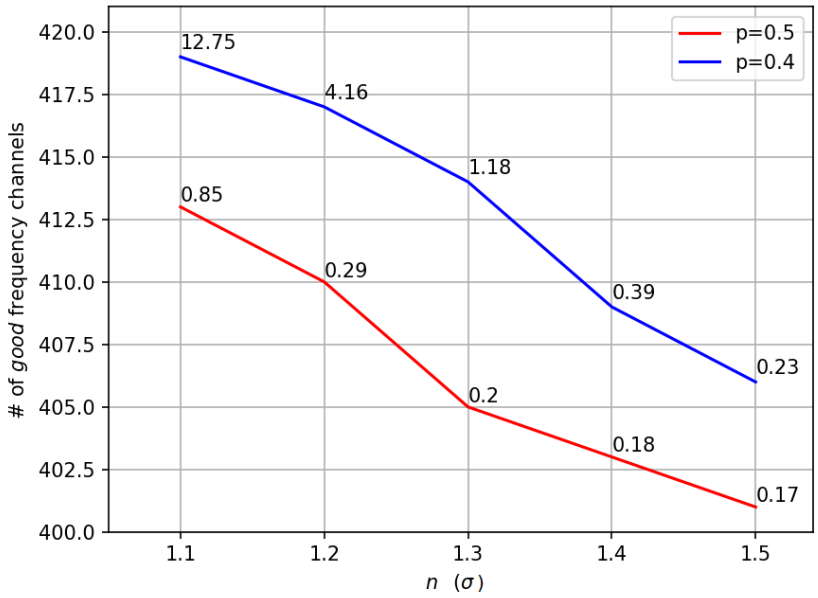

Figure 3.5: A comparison between two excision algorithms with different f parameters. The number above data points shows the average percentage of excised data in *good* frequency channels.

3.2.3 Comparison of different excision algorithms

In this section, single-stage, two-stage and EOM algorithms are applied to the 10 hour dataset recorded between 12:00 and 10:00 AM on October 11. Note that in the single stage scenario $n = 5$ is used, and in the two-stage case I used the fixed values $n_2 = 1.4$ and $f = 0.4$ as the excision parameters. Then, the gaussianity test is performed on 8-s samples every ~ 1.5 hour and those frequencies whose *GT* value is less than 2 are labeled as *good* frequencies. The fraction of excised data averaged over *all* frequency channels for the single-stage and two-stage excision algorithms are $\sim 12\%$ and 17%, respectively. The average excised fraction for *good* frequencies is $\sim 0.2\%$ for the single-stage and $\sim 0.5\%$ for the two-stage algorithm. Average excised fraction in all frequencies for EOM algorithm is $\sim 17\%$ and $\sim 0.5\%$ for good frequency channels (similar to two-stage).

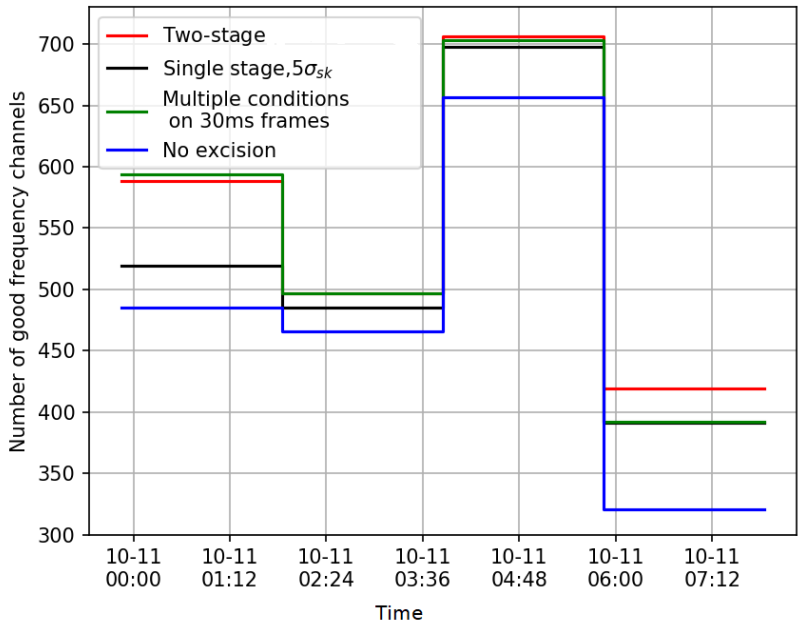

Figure 3.6: The performance of single stage (black), EOF (green) and two stage (red) algorithms over ~ 8 hours. In general, two stage and EOF outperform the single stage. The difference between two stage and EOF is not significant because this test was performed at night, when there are no strong source of transient RFI. If there are powerful, short time scale RFI, the first stage of two stage algorithm can remove it, while EOF is not able to do so.

Chapter 4

SK estimator biases

The SK estimator and its statistical properties discussed in chapter 2 are derived from floating point Gaussian samples. Moreover, the value of the variance of the estimator for the CHIME array (equation 2.41) is based on the assumption that all the samples from different inputs are independent. These are all idealistic assumptions. The samples from which the SK value is estimated are not in the form of floating point values, because they are truncated to 4 bits before transferring to the GPUs. Furthermore, we might have a very bright celestial signal, e.g. the Sun, that is measured by all of the antennas. In this case the signals measured at different inputs are not totally independent. In this chapter I discuss the effect of 4+4-bits truncation and the presence of a common mode signal on SK estimates and on the variance of the estimator, respectively. I also introduce a correction for the truncation bias through a python simulation.

4.1 4+4 bits truncation

Spectral kurtosis estimator is sensitive to the shape of the random variable distribution. In our case, the random variable is the voltage and if there is no RFI contamination, it follows a Gaussian distribution with mean zero. So according to the spectral kurtosis estimator properties discussed in the previous section, its expected value must be unity with a known variance. However, if for some reason the shape of the distribution changes, it could change the spectral kurtosis estimates.

As I mentioned in the section 1.3.3, the real and imaginary parts of the voltage are separately truncated to 4-bits. Suppose that we have a Gaussian background noise with no RFI contamination. If the signal is truncated, then the shape of the signal could change. This is because our samples are not floating point values anymore after the 4+4-bits truncation (4 bits real and 4 bits imaginary), but they are integers in the interval [-7,7] (see figure 4.1). In this case, our SK estimates will deviate from the expectation value and the RFI detection algorithm will mask the signal. This is obviously not desired, as we don't want to mask the celestial signal. I made a simulation in python to see the effect of 4+4-bits truncation on spectral kurtosis estimates:

- A set of $nN = 256 \times 2048$ complex random variable are drawn from a circularly symmetric Gaussian distribution with a known rms.
- Real and imaginary parts of the random variable are separately truncated to 4-bits. This means that they are rounded to their nearest integer and bounded between [-7,7]. All the values greater than 7 or smaller than -7 are replaced by 7 or -7, respectively. This is shown in the figure 4.1.
- The spectral kurtosis with and without truncation are estimated.
- This process is repeated 6000 times. To see the effect of truncation, the mean and the variance of 6000 spectral kurtosis estimates with and without truncation are compared.

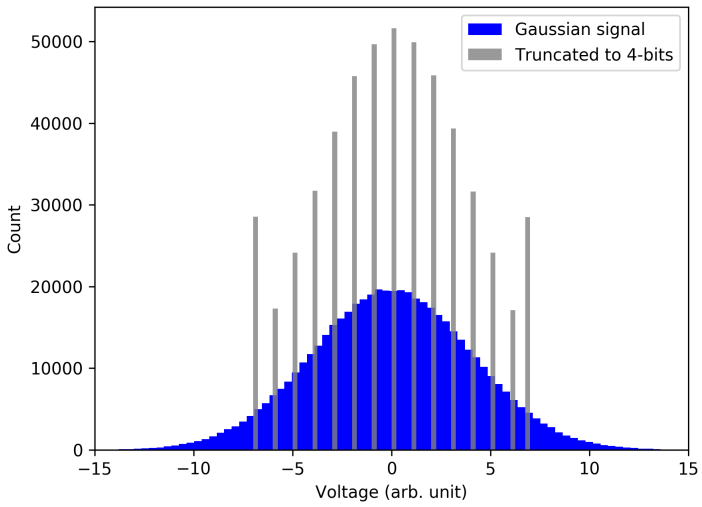

Figure 4.1: Effect of 4-bits truncation on the shape of a Gaussian signal with an rms of ~ 5.5.

Looking at the figure 4.1, one can see that 4+4-bits truncation changes the shape of the distribution significantly if the rms is high. Since kurtosis is a measure of the shape of the distribution, truncation introduces a bias in the estimator, because the spectral kurtosis estimates from truncated samples deviates from the expectation value. This is shown in the figure 4.2. The SK estimates generated from digitized samples tends to be smaller than those generated from non-digitized samples. So truncation introduces a **negative bias** in the SK estimates. This is expected, because

at high rms, voltage samples beyond [-7,7] are replaced by -7, 7 which reduces the S_1 and S_2 estimates, and S_2/S_1^2 in the equation 2.27. For high rms values, the effect of bounding of the samples to [-7,7] interval on S_2/S_1^2 estimate is more than rounding them to the nearest integer. For lower rms values (rms<2.2), there are not many samples beyond [-7,7], but rounding of the samples to the nearest integer slightly reduces the S_2/S_1^2 estimates. So truncation always leads to a negative bias in the SK estimates.

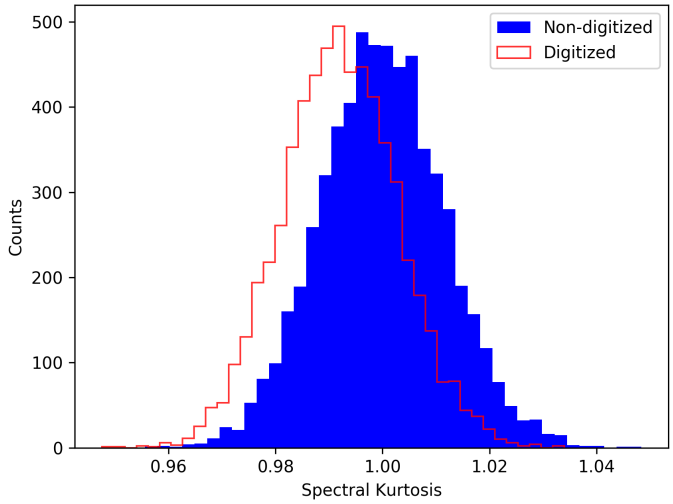

Figure 4.2: Effect of 4+4-bits truncation on SK estimates. The SK estimates are generated from 6000 Gaussian dataset with rms $\sim$ 2.9. The SK estimates tend to be smaller after truncation.

Since the effect of truncation on the shape of a distribution is more significant for high rms values (as there will be more sample with a value exceeding the interval [-7,7]), we should expect a correlation between the rms and the bias in the SK estimate. Figure 4.3 shows the relationship between the truncation bias and the rms of **truncated** samples in terms of nominal σ for CHIME. Note that the average rms of the samples in the absence of bright sources in the sky is around 2.2. So the SK estimator is always biased, although the bias is less than 0.5σ when there is no bright source in the sky.

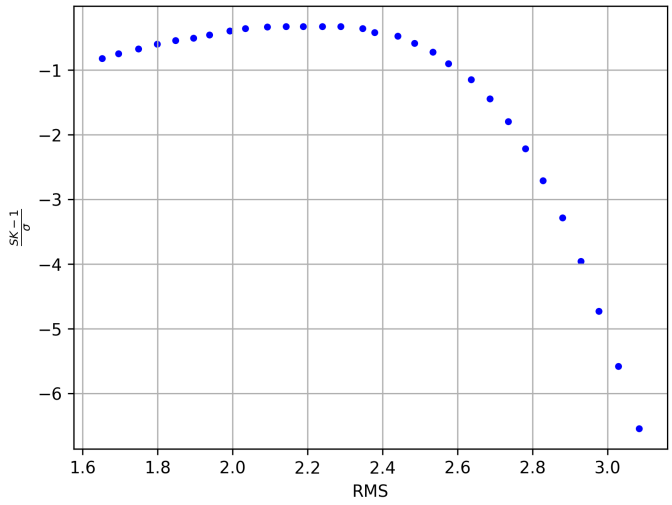

Figure 4.3: Bias of the SK estimator as a function of the rms of truncated samples. Each data point shows the deviation of the average SK value (over 6000 Gaussian dataset) from unity for a given rms.

4.1.1 Effect of truncation bias on CHIME data

On January 2020, we noticed an increase in the excision rate during the transit of Cygnus-A (hereafter Cyg-A) and Cassiopeia-A. This was obviously not desirable, as the RFI system should never excise the astrophysical signals. To understand this feature, spectral kurtosis values were dumped during Cyg-A transit on February 2020. The goal was to study the spectral behaviour of the SK values during the transit of bright sources, and to see how much the SK values deviates from the expectations. Figure 4.4 shows the deviation of SK the values from unity during the Cyg-A transit. In the ideal case, we shouldn't see the celestial sources in such a plot, as they are Gaussian and SK estimator is not sensitive to them. However, we can see the trace of Cyg-A as a column which is darker than its neighbourhood around 9:29 AM. The lower part of the plot is darker during the transit, which means that the negative SK fluctuations occur more at lower frequencies. In order to see how much data are being flagged by the RFI system at lower frequencies, the average excised rate between 400MHz and 500 MHz is estimated. To find the average excised rate, an extreme threshold $(n, f) = (1.5, 0.4)$ was applied on $30\mathrm{ms}$ frames (see section 3.1.2). Then the number of samples exceeding our threshold is averaged between 400 MHz and 500 MHz, and it is normalized by the total number

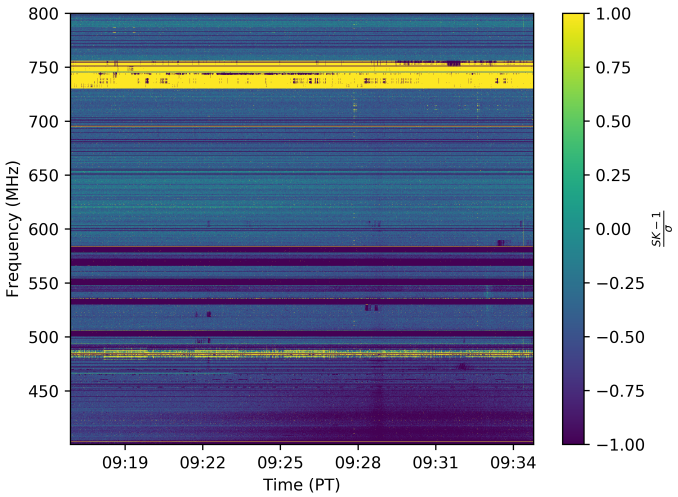

Figure 4.4: Bias of SK values in terms of the nominal σ for CHIME. The SK values are more negatively biased during the transit of Cyg-a which is around ~ 9:29 AM. Note that lower frequencies are more negatively biased than the higher frequencies.

of samples in a 2-seconds frame. We can see this effect in the figure 4.5. The increases in the excision rate at the transit time of the Cyg-A is about %4. In other words, our kurtosis based RFI detection system mitigate %4 of useful data at lower frequencies during the transit of Cyg_A.

On the other hand, we know that Cyg-A is brighter at lower frequencies and I also showed that 4+4-bits truncation introduces a negative bias in the SK estimator. This negative bias increases as the rms of the data increases. So, the observed correlation between the transit of Cyg-A and SK deviation from the expected value (which is higher at lower frequencies) is a hint that the increased excision rate during the transit of bright celestial sources can be explained by the truncation bias. So correcting the estimator for the truncation bias should solve this issue.

4.1.2 Correction for truncation bias

The truncation bias is defined as the deviation of the SK value from unity. As discussed in section 4.1, the truncation bias depends on the rms of the samples from which the SK value is estimated. Our goal in this section is to derive the truncation bias as a function of rms. Suppose that we have a set of Gaussian data with mean zero and known rms. If there is a table to relate the truncation bias with the rms, the SK estimator can be corrected for the truncation bias for a given rms. To do so, I simulated 1000 SK estimates, where each estimate is generated from $Nn = 2048 \times 256$ complex samples drawn from a circularly symmetric Gaussian distribution with some rms. Then, I take the average of 1000 SK realizations, the rms of the truncated distribution is changed and the average SK value is re-estimated. The simulation steps are similar to what described in the section 4.1. The result is shown in the figure 4.6 (right).

Once we have SK estimates as a function of rms, we can fit a polynomial to it to find the bias for a given rms. The bias for a given rms is:

$$\text{bias} = \text{SK}_{\text{fitted}} - 1 \tag{4.1}$$

and the SK estimate corrected for the truncation bias is:

$$\text{SK}_{\text{corr}} = \text{bias} + \text{SK}, \tag{4.2}$$

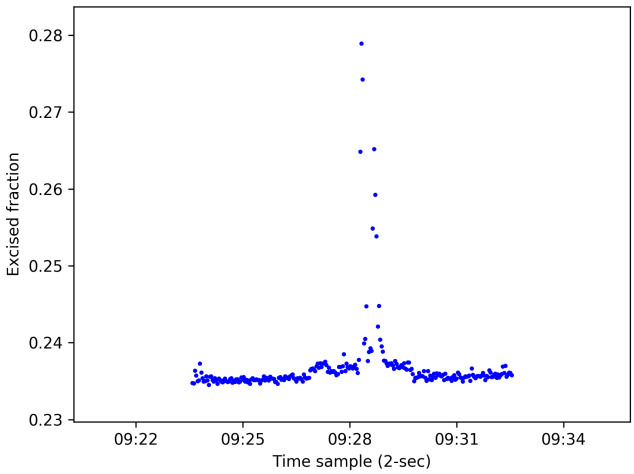

Figure 4.5: The excised rate during Cyg-A transit . Each data point represents the fraction of excised samples in a 2 seconds frame averaged between 400 MHz and 500 MHz.

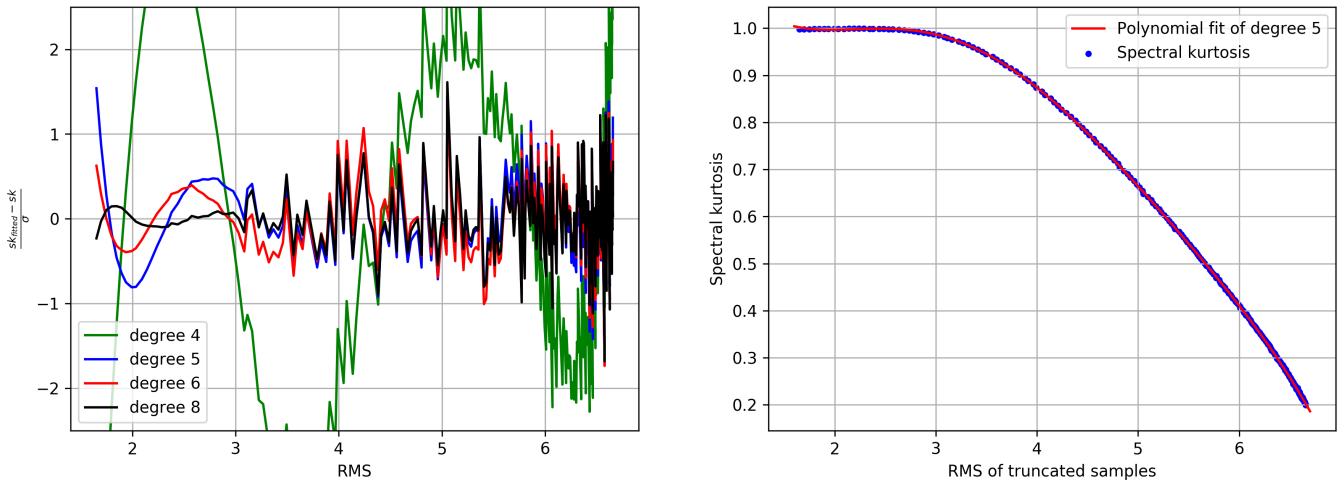

Figure 4.6: Left: Fitting residuals in terms of nominal σ as a function of rms for different polynomials. Right: Spectral kurtosis estimated from truncated samples as function of the rms of the truncated samples.

where SK is the spectral kurtosis measured by the estimator given in the equation 2.40. In other words, we continue to use equation 2.40 to estimate the SK value. This is a biased estimate. So we add one more step to correct the truncation bias: Having the instantaneous rms of the samples, we evaluate the bias for that rms using a polynomial, and then it is added to the measured SK value. The result is the SK estimate corrected for the truncation bias.

The major source of error in this process is the polynomial fit. Figure 4.6 (left) shows the residuals of the fitting in terms of the nominal sigma for 4 different polynomials. The residuals are the difference between the fitted SK value (using polynomial) and the simulated SK estimates. The lowest order polynomial which gives a good fit within the range of $\pm 0.5\sigma$ is of degree 5. Polynomials of degree 8 and more gives a much better fit with a lower residual, but fitting with a polynomial of degree 5 is just enough to test the system and to see whether our idea for correcting the truncation bias works as expected.

4.1.3 Results of truncation bias correction on CHIME data

We recorded the SK values together with the variances of 0.65 ms samples simultaneously to correct the SK estimates for the truncation bias. Because the mean of

the random variable x is zero, variance directly gives the rms:

$$Var(x) = \left\langle x^2 \right\rangle = \text{rms}(x). \tag{4.3}$$

Using instantaneous rms values and the polynomial function I derived for the fitting (of degree 5), one can estimate the instantaneous truncation bias every 0.665 ms. Adding the truncation bias to the measured SK value gives the unbiased SK estimate. The bias of the SK estimator before and after the correction for truncation bias is shown in the figure 4.7. Cyg-A transit is around 4:45 AM in these figures. I applied the threshold $(n, f) = (1.5, 0.4)$ to 30 ms frames. White spots are the 30 ms frames exceeding the threshold. We can still see the trace of the Cyg-A on the SK estimates after the correction, however, the bias is getting smaller and most of the samples do not exceed the threshold after the correction. Moreover, the SK estimator is still sensitive to the other RFI events including LTE band, TV station bands, repeating RFI around 470 MHz, and transient RFI.

The effect of truncation is expected to be more clear at lower frequencies because the sky is getting brighter at those frequencies during the transit of the galactic plane. So, zooming into the lower frequencies can give us a better sense of the performance of the SK estimator with and without the truncation bias correction. The average excision rate for 2-sec frames between 400 MHz and 500 MHz before and after correction are shown in the figure 4.8. The bump around 4:45 in the figure 4.8 (left figure) corresponds to the Cyg-A transit, showing that the excised rate increases at lower frequencies without bias correction. The bump disappears after correction with a polynomial of degree 5. We can still see a $\%2$ increase in the excised rate at the transit time, however, this is a much better than $\%10$ increase before correction.

Although the correction with a polynomial of degree 5 improves the SK estimator, for better results we need higher order polynomials or a different fitting method. An alternative to the polynomial fitting is the cubic spline interpolation. It turns out that this method gives a much better result. Figure 4.9 shows the SK deviation from the expected value in terms of nominal σ after correcting the SK values with a cubic spline function during the transit of Cyg-A. The SK values are much closer to unity in this case, and the signature of Cyg-A in the plot is disappeared. Moreover, the SK values are much more uniform across the spectrum. The excision rate does not change with the cubic spline function, but in general it gives a much better fit to the SK v.s rms plot. Currently, we are using the polynomial fitting in the pipeline.

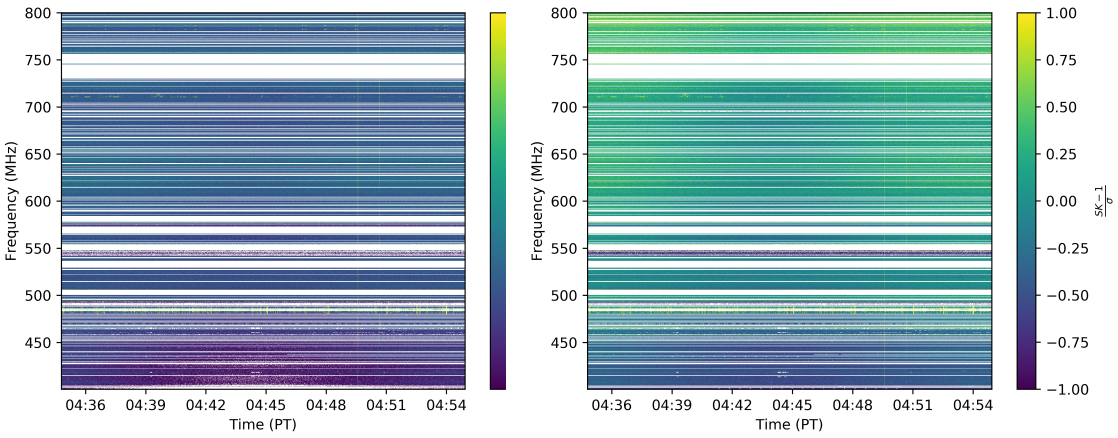

Figure 4.7: The bias of SK estimates in terms of nominal σ for CHIME without (Left) and with (Right) the correction for truncation bias.Cyg-A transit is around 4:45 PT in this plot. Each data point is the average SK value over 30 ms and the white spots are 30 ms frames that are masked by EOF algorithm with $(n, f) = (1.5, 0.4)$. The color scale is the same in both plots.

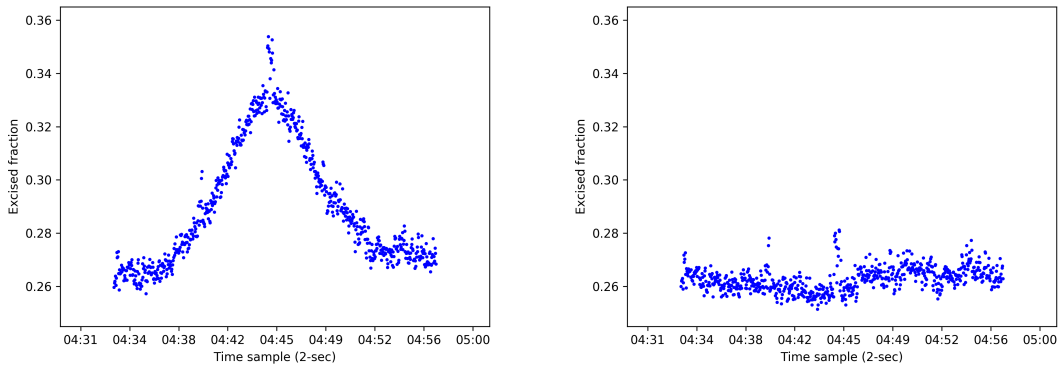

Figure 4.8: Left: The excised fraction of the samples in an 2 second frame, averaged between 400 MHz and 500 MHz during Cyg-A transit ($\sim$ 4:45 PT) without truncation bias correction. Right: The same quantity as in the left figure, but corrected for truncation bias with a polynomial of degree 5. The bump corresponding to Cyg-A disappeared, i.e. the RFI system no longer excise the Cyg-A.

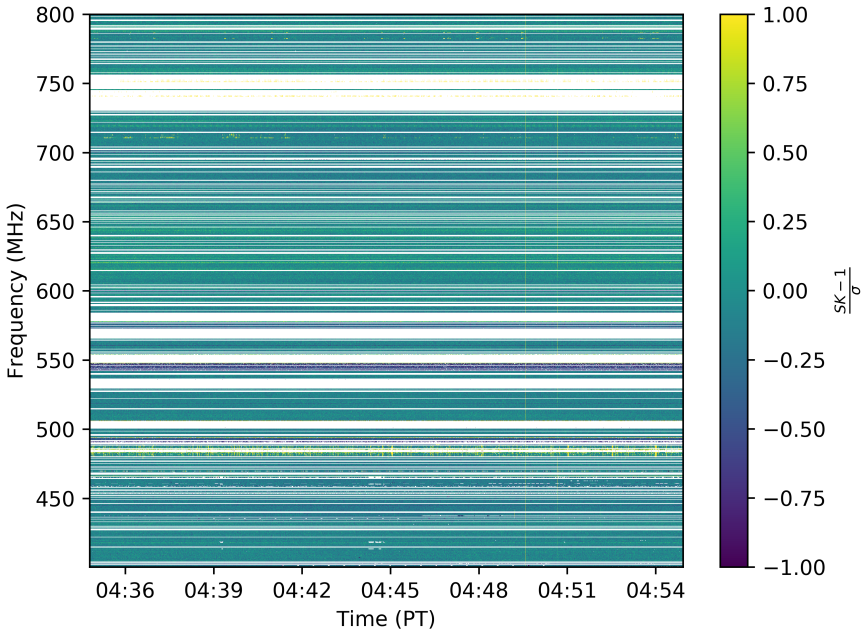

Figure 4.9: Deviation of SK values from unity after correcting the kurtosis values for truncation bias with a cubic spline function during Cyg-A transit around $\sim 4:45PT$. The bias is closer to zero for all frequencies compared to the figure 4.7.

4.2 Common mode signal

We derived the variance of the SK estimator for a compact array in the section 2.4 assuming that the signals received by different receivers are independent. In general this is not the case, as the sky signal contributes to all the receivers. So there will be common mode signal and this reduces the total number of independent samples, and increases the variance of the estimator. In other words, the sensitivity of the kurtosis-based RFI mitigation system which depends on the variance of the SK estimator is decreased in presence of a common mode signal.

A simulation can reveal the effect of common mode signal on the variance of SK estimator. Suppose that the receiver noise is described by the matrix *N*. A set of 256×2048 complex samples is selected randomly from a Gaussian distribution with mean zero and variance 1. I put these samples in a 2 dimensional matrix *N*. This matrix has two axes: time axis which is represented by 256 samples; and receiver axis representing 2048 inputs. Each row is independent from the other one. This matrix represents the receiver noise of each input for 256 time samples. The component N_{ij} represents the signal received by the i^{th} receiver at time j. So each row is the time-stream for a single receiver (with a length of 256). Note that the noise of different receivers are independent from each other, and each of them follows a circularly symmetric Gaussian distribution.

$$N = \begin{bmatrix} N_{1,1} & N_{1,2} & ... & N_{1,256} \\ N_{2,1} & N_{2,2} & ... & N_{2,256} \\ ... & ... & ... & ... \\ N_{2048,1} & N_{2048,2} & ... & N_{2048,256} \end{bmatrix} \tag{4.4}$$

Then, a common mode signal is added to the noise. It is described by the matrix C whose shape is same as matrix *N*. The only difference is that *C* follows a circularly symmetric Gaussian distribution along the time axis, but it is the same over all of the receivers:

$$C = \begin{bmatrix} C_{1,1} & C_{1,2} & ... & C_{1,256} \\ C_{2,1} & C_{2,2} & ... & C_{2,256} \\ ... & ... & ... & ... \\ C_{2048,1} & C_{2048,2} & ... & C_{2048,256} \end{bmatrix} = \begin{bmatrix} C_1 & C_2 & ... & C_{256} \\ C_1 & C_2 & ... & C_{256} \\ ... & ... & ... & ... \\ C_1 & C_2 & ... & C_{256} \end{bmatrix}$$

Then the signal is:

$$S = \sigma_N N + \sigma_C C \tag{4.5}$$

where σ_N is the noise amplitude and σ_C is the common mode amplitude. Then, the SK value of the signal is estimated, and the variance of the estimator for a set of 1000 dataset is calculated. This process is repeated for different values of σ_C. The fractional common mode amplitude shows how much of the amplitude of the signal is dominated by the common mode:

$$\sigma_{\text{frac}} = \frac{\sigma_C}{\sigma_{total}} = \frac{\sigma_C}{\sqrt{\sigma_N^2 + \sigma_C^2}}. \tag{4.6}$$

Figure 4.10 shows that the effect of the common mode signal on the variance of the SK estimator is negligible if the fractional common mode amplitude is less than 0.3. The signal is dominated by the common mode when $\sigma_{\text{frac}} \approx 1$. In this case there are only $n = 256$ independent time samples and the variance converges to $4/n \approx 0.015$. Since the increase in the common mode amplitude reduces the number of independent samples, one can define the effective number of receivers using the fact that for a compact array $N \approx \frac{4}{n \times Var(SK)}$. When the fractional common mode amplitude is unity the effective number of antennas becomes 1, as expected. This is shown in the figure 4.11. Red horizontal line in both figures corresponds to a signal that is totally dominated by a common mode.

Note that the common mode signal does not alter the mean of the estimator, because mean of the SK estimator is an invariant quantity and does not depend on the integration length. Therefore common mode is not counted as a bias. But it reduces the sensitivity of the SK estimator to the RFI as it reduces the number of independent samples.

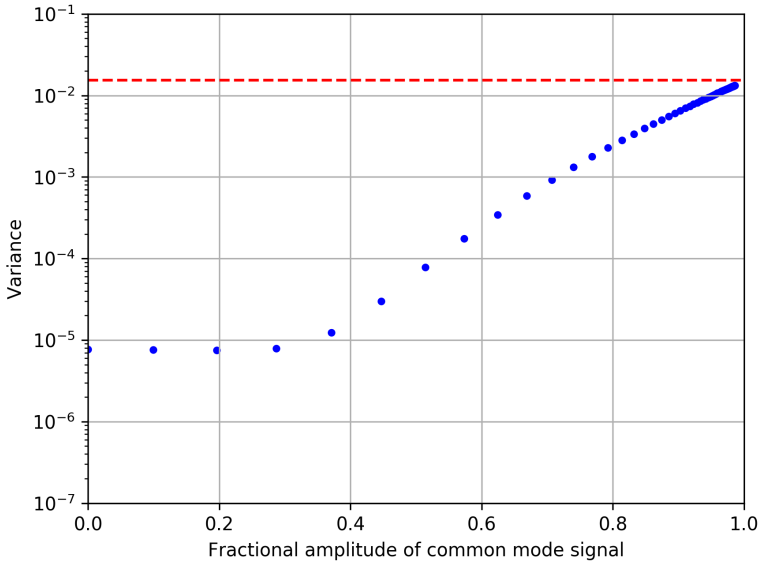

Figure 4.10: Variance of the SK estimator in presence of a common mode signal. The variance reaches to 10^{-2} if the signal is totally dominated by the common mode signal, as there are only $n = 256$ independent time samples.

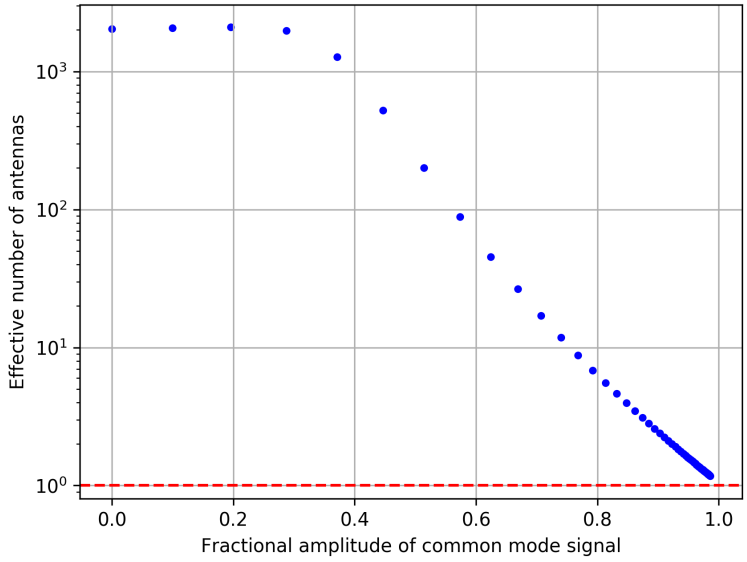

Figure 4.11: Number of effective antenna as a function of fractional common mode amplitude.

Chapter 5

Results

In this chapter, I summarize the results of the real-time RFI excision from May 2019 to September 2020. In this period we have used all three excision algorithms described in the chapter 3. The final part of this chapter is a discussion about the reliability of the Gaussianity test for the evaluation of the RFI system, the performance of the system during the solar transit, and a brief discussion about the effect of polarized RFI on the sensitivity of the RFI system.

5.1 May to November 2019: Single stage algorithm

We started to use the kurtosis-based RFI detection system in May 2019. From May to November 2019, single stage algorithm with a detection threshold of 5σ was used. The estimator was not corrected for the truncation bias in this period.

To compare the quality of the data before and after the RFI excision visually, I plotted the stacked autocorrelations normalized by the median across time for each frequency. Stacked autocorrelations are the autocorrelations averaged over all inputs. Figure 5.1 shows the waterfall plot of autocorrelations before and after we start to use the kurtosis based RFI excision system. I zoomed in the range 400 MHz and 500 MHz, because this is the range that we see many repeating RFI lines. One can see that there are a few RFI lines in April 2019 (without RFI excision) that are absent in May 2019 (after RFI excision). Most of these lines are not transient, for example the repeating line around 468 MHz.

To be more quantitative, I plotted the average excision rate as a function of frequency. Figure 5.2 shows the excision rate in each frequency bin averaged over 1.5 hours. The shape of this plot is very similar to the figure 3.4. This shows that the excision algorithm is working as expected. Most of the RFI lines appearing between 400 and 500 MHz, can be recovered by removing less than 20% of the data. On a 10-second cadence, this repeating line appears every 20 seconds. But the high cadence RFI excision system can clean the frequency without excising too much of the data. Figure 5.2 shows that this frequency can be recovered by excising only 8% of the data in a 10 second sample at this frequency.

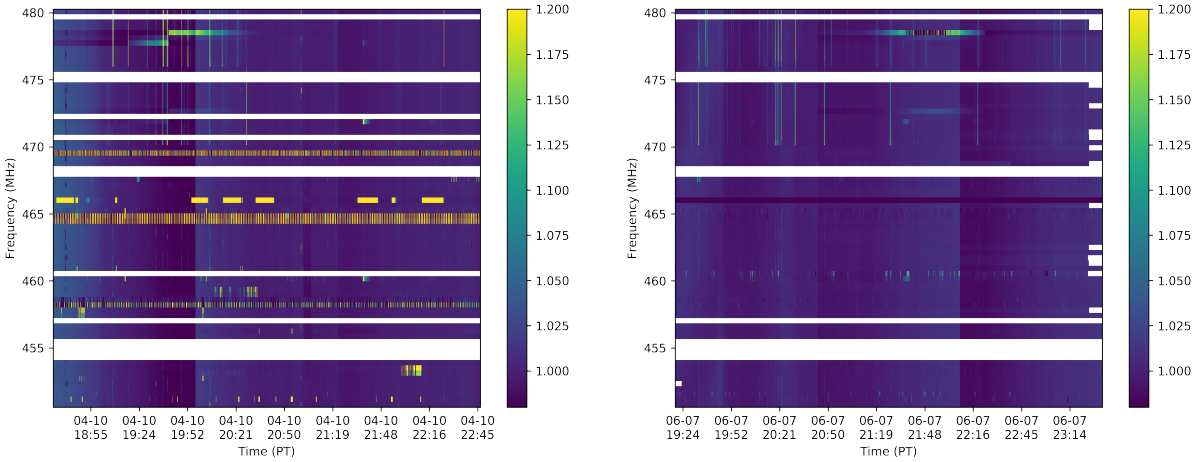

Figure 5.1: Left: Normalized autocorrelations in April 10, 2019 without any RFI excision. Right: Normalized autocorrelations in June 7, 2019 after deploying high cadence RFI excision system with single stage algorithm. Most of the repeating RFI features disappeared.

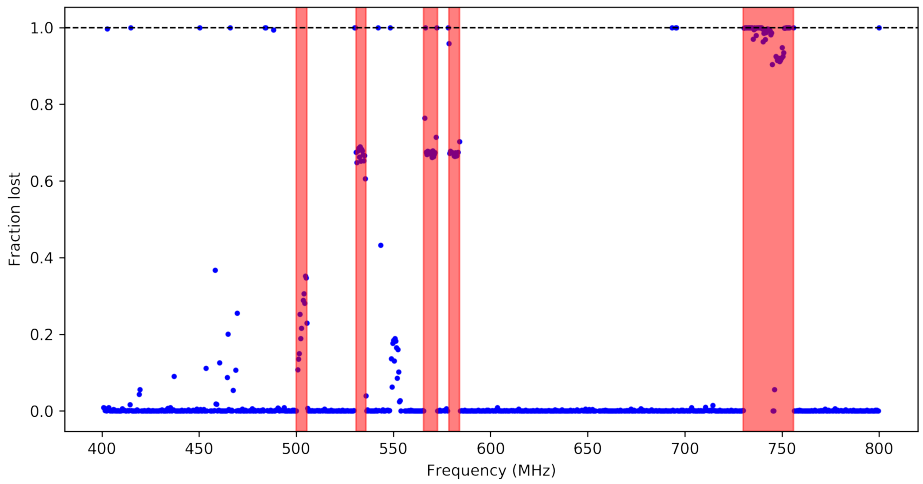

Figure 5.2: Excised fraction of 10 s samples averaged over 1.5 hours as a function of frequency.

5.1.1 Moving blob

In July 2019, we detected a mysterious signal that was being removed by the RFI system. The signal was repeating every day with a regular pattern . Preliminary evaluations showed that the signal appears around 480 to 490 MHz and it is locked in sidereal time, i.e. it appears at the same time in each sidereal day, just like stars. Being locked in sidereal time was a hint that the signal source is not ground based. However, tracking the signal over a month showed that the *blob* is slowly drifting in sidereal time. Figure 5.3 shows how the *blob* moves back in time. Each diagonal line corresponds to the excised rate at 483.98 MHz for 24 hours. We can see the increase in the excised rate, two times per day, at specific right ascensions. Note that CHIME is a transit radio telescope. So the right ascension of an object is equal to the local sidereal time when the object is in the CHIME beam. The question was whether this is a true RFI or the RFI system is removing a celestial source by mistake. There was no bright radio source at those right ascensions. Finally, with the help from Scott Tilley, an amateur satellite observer, it turned out that the signal is coming from a Russian Meridian satellite and the RFI system did a good job in detecting the RFI.

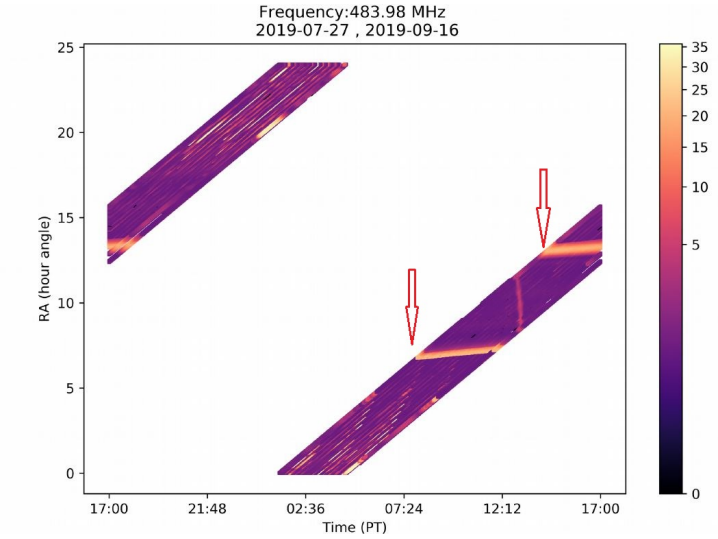

Figure 5.3: The signature of the Meridian satellite is visible in the plot of excision rate of 10 s samples (in percent) as a function of time around frequency ~484 MHz. Each diagonal line corresponds to 24 hours. The blob appears every day, but it slowly drifts in right ascension.

5.2 December 2019 to February 2020: Excision on 30 ms frames with multiple thresholds

This section summarizes the overall performance of the RFI excision system from December 2019 to February 2020. During this period the EOM algorithm with two thresholds $(n, f) = \{(1.5, 0.4), (3, 0.13)\}$ was used. This algorithm was initially tested on SK/variance dataset in October 2019 (see section 3.2.2) .

Figure 5.4 shows the excised fraction of 10 s samples (averaged over all frequencies) for 3 consecutive days. One can see that the excision rate averaged over all frequencies is around 16%, as expected from the analysis in chapter 3 for the EOF algorithm. We can also see 3 bumps which appear every day. Two of these bumps are already identified as the meridian satellite (see section 5.1.1). These bumps are moving back in time every day. These three bumps have some similar characteristics: They are almost locked in sidereal day, and if one plots the excised fraction as a function of frequency and time, they are all appearing as blobs between 480 and 490 MHz (figure 5.8). However, unlike the previous blobs that are drifting slowly in RA, the new blob seems to be tightly fixed in RA, and it also seems to be more broadband. Figure 5.8 shows that the new blob appears in 480-490 MHz and 400-440 MHz at the time of Cyg-A transit. Therefore, it is obvious that the new blob has a different origin than the other two. In fact, the RFI system is excising the Cyg-A. This is not desirable and as already mentioned in chapter 4, it is due to the 4+4-bits truncation. The effect of 4+4-bits truncation was not very clear with the single stage algorithm because the detection threshold was at 5σ.

To see how the excision work in different frequencies, I plot the average excised fraction of 10 s samples at different frequencies over 1.5 hours. Figure 5.5 shows that the LTE and TV station bands are completely masked by the EOF algorithm. However, the RFI lines around 450 MHz are more excised than what the single stage algorithm does. Note that removing more samples around 450 MHz by the EOF algorithm does not necessarily mean that more RFI is removed. Higher excision rate in those frequencies is mainly due to the fact that the EOF algorithm performs the excision on 30 ms frames, while single stage algorithm masks the individual outliers (0.65 ms).

Another quantity of interest is the number of good frequency channels throughout the day. This is shown in the figure 5.6. Remember that good frequency channels are those channels that pass the Gaussianity test (i.e. those with GT <2). The maximum number of good frequency channels ($\sim$ 700) occurs every night between 2:00 and 4:00 AM. The minimum number occurs in the morning between 8:00 AM and 12 PM. These numbers are consistent with the results of our test on sk/variance dataset.

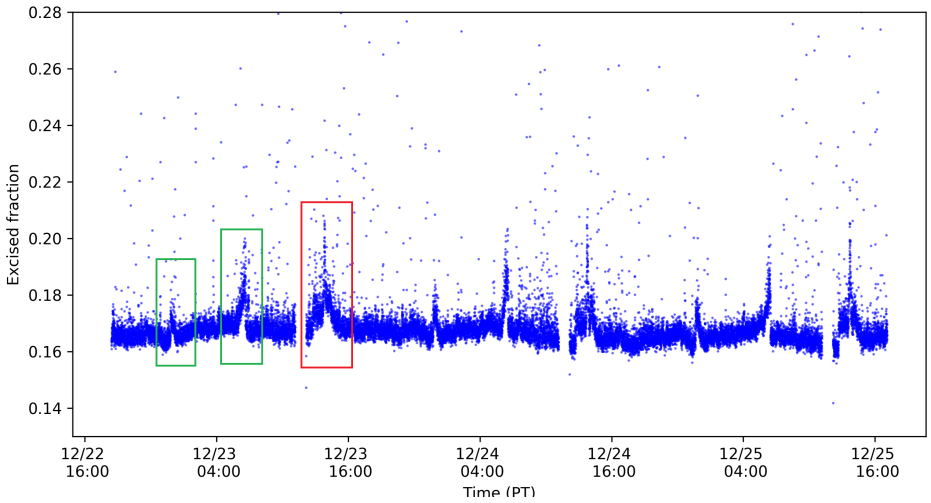

Figure 5.4: Excised rate of 10 s samples averaged over all frequencies for 3 consecutive days. The red box is the Cyg-A transit time and two green boxes correspond to the meridian satellite. The excision rate increases every day during Cyg-A transit. This is due to 4+4-bit truncation.

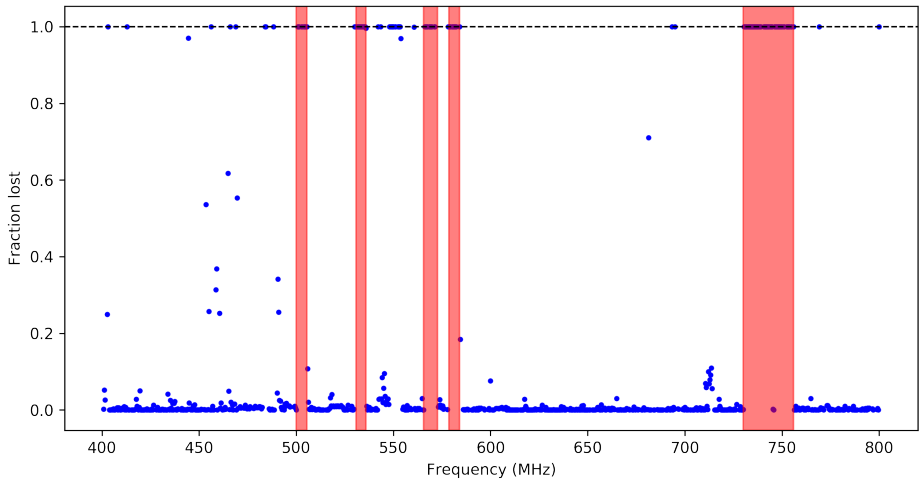

Figure 5.5: Average excised rate over 1.5 hours as a function of frequency for a typical day in February 2020. The excision algorithm completely masks the LTE and TV station bands.

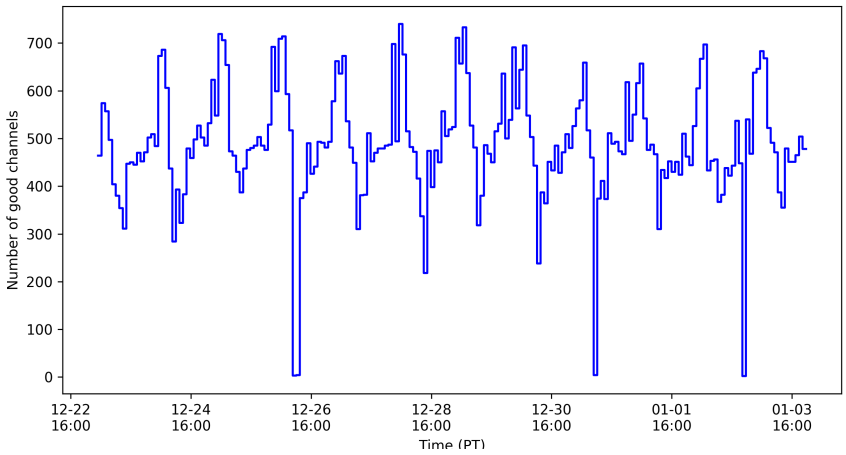

Figure 5.6: Number of good frequency channels reported every 1.5 hours for 12 consecutive days. We can see a repeatable pattern every day. Note that there are zero good frequency channels during solar transit, as the RFI system flags the Sun. This issue is discussed in section 5.4.2

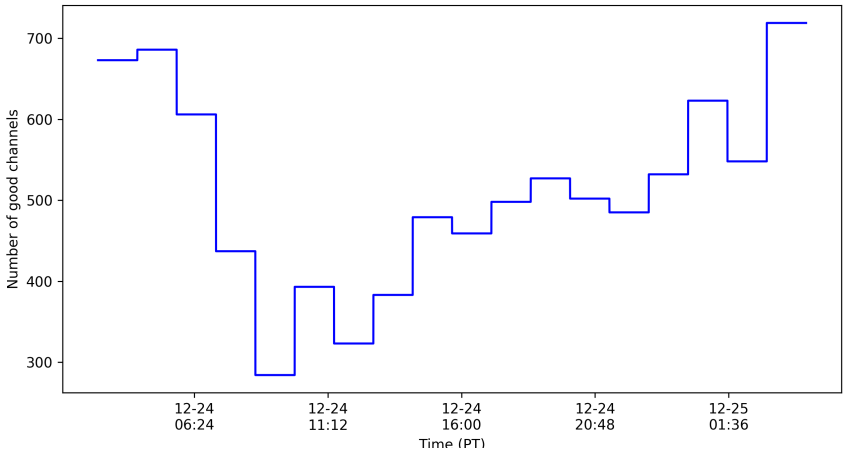

Figure 5.7: Change of the number of good frequency channels in 24 hours. The number of good frequency channels is more at night and it goes down in the morning, when there are more RFI at site.

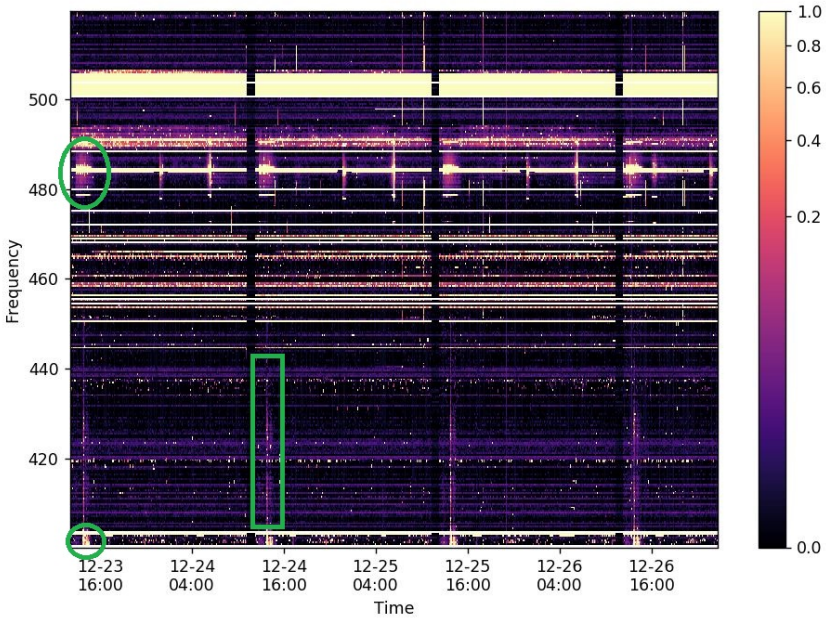

Figure 5.8: Excised fraction as a function of time and frequency. The green circles and rectangle show the blob which is fixed in RA (which turned out to be Cyg-A).

5.3 March to September 2020: Two stage algorithm

As of March 2020, we are using two stage excision algorithm: In the first stage, all individual 0.65 ms samples whose kurtosis exceed 5σ limit are masked. The second stage excision is performed on 30 ms frames whose more than 13% of their samples exceed 3σ. In June 2020, we corrected the SK estimator for truncation bias with a fifth degree polynomial as described in chapter 4. The number of good frequency channels as a function of time for 6 consecutive days in July 2020 is shown in the figure 5.11. The maximum number of good frequencies is a little lower than the winter run . This could possibly be due to the new RFI bands appearing around 700 MHz and the change of excision algorithm and the threshold values. Figures 5.9 and 5.10 show the effect of truncation bias correction on excision rate during the transit of Cyg-A. Since the truncation bias is more severe in lower frequencies during the Cyg-A transit, I averaged the excision rate over 400 and 430 MHz. The bump in the figure 5.9 corresponds to the time of Cyg-A transit. This bump is disappeared in June when the SK estimator was corrected for the truncation bias.

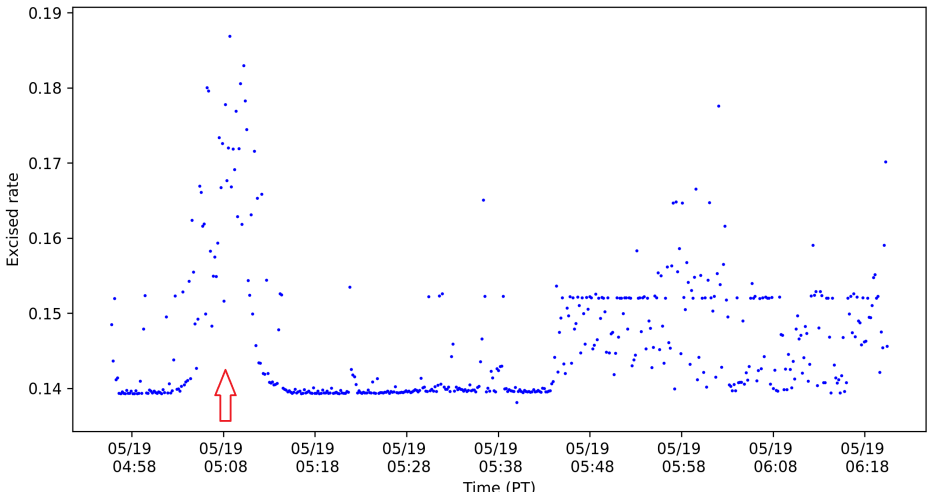

Figure 5.9: Excised rate of 10 second samples averaged between 400 and 430 MHz on May 19, 2020; without truncation bias correction. Red arrow shows the Cyg-A transit time.

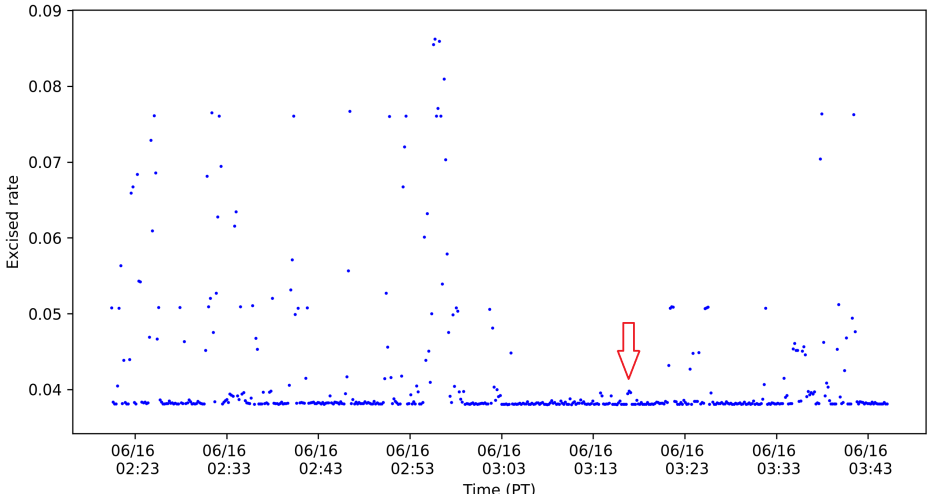

Figure 5.10: Excised rate of 10 s samples averaged between 400 and 430 MHz on June 16, 2020. The SK estimator is corrected for the truncation bias. The The bump corresponding to the Cyg-A transit disappeared.

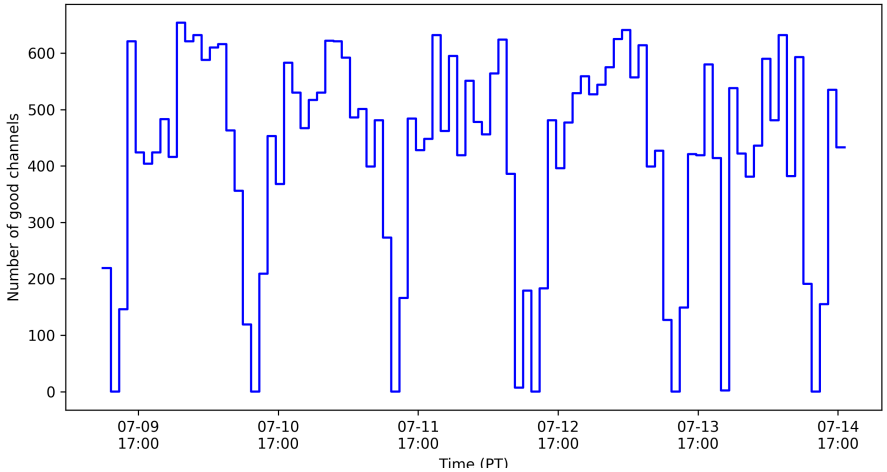

Figure 5.11: Number of good frequencies from July 9, 2020 to July 14,2020. Two stage algorithm was used in this period with 5σ threshold on individual 0.65 ms samples and $(3\sigma, 0.13)$ on 30 ms frames. Zeros correspond to the solar transit, when the RFI excision system is off.

5.4 Discussion

5.4.1 Reliability of Gaussianity test

In this analysis, the aim of gaussianity test was to compare the quality of the data before and after the excision by identifying those frequency channels which are less contaminated by RFI. To see whether the gaussianity test selects only such channels, the standard deviation of the data over an hour or so can be used. The RFI should appear as a spike in the standard deviation plot. So, if any RFI-contaminated channel pass the gaussianity test, we must see it as a spike in the plot of standard deviation versus frequency.

Figure 5.12 is the standard deviation of the frequency channels that have passed the gaussianity test. Clearly, there are a few outliers: These frequencies have passed the gaussianity test, but they have a large standard deviation and indeed they are RFI.

Why the outliers pass the gaussianity test? Are they gaussian-like RFI? Not necessarily. In the gaussianity test, we subtract the sky contribution by taking the difference between neighbouring samples. But if the RFI is stable in time (over ~ 10 seconds), the subtraction removes the RFI and the frequency channel might pass

the test. For example, figure 5.13 shows the histogram of one of the outliers before and after subtraction of neighbouring samples. After subtraction, the distribution looks more Gaussian. This behaviour is not so much common, but it should be taken into account.

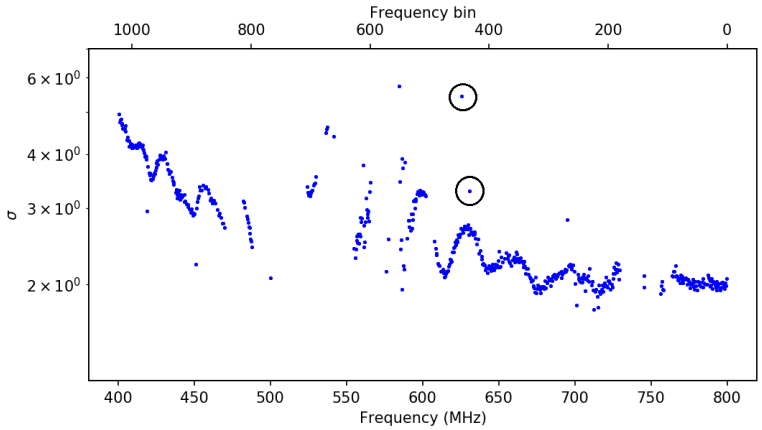

Figure 5.12: Standard deviation of the data between 3:00 and 4:00 AM for those channels which pass the gaussianity test. Two of the outliers are shown with black circles.

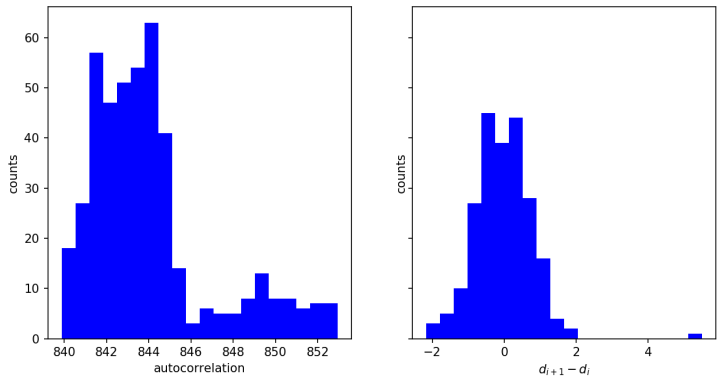

Figure 5.13: Left: Histogram of the autocorrelations one of the outliers at frequency 419.14 MHz. Right: Histogram after subtracting the neighbouring samples.

5.4.2 Solar transit issue

The RFI excision rate is extremely high during the solar transit. Ideally this should not happen, because the Solar data are Gaussian and the RFI system should not remove the Gaussian signal. However, the two effects mentioned in the chapter 4 affect the SK estimates and its variance. The truncation bias is very significant for the sun, as it is the brightest object in the sky. The average rms of the solar data is greater than 4. According to the simulations explained in the chapter 4, the bias of the SK estimates due to the truncation effects at rms~ 4 is more than -40σ . This means that all the excision algorithms (single stage, EOM or two stage) will excise almost all of the solar data. However, this issue persists even after the correction of SK estimates for truncation bias. This is most probably due to the common mode signal: the common mode amplitude approaches to the unity in the case of solar transit. This reduces the number of independent samples from which the kurtosis estimates are generated. So the variance of the estimates will increase and the SK values exceed the detection thresholds. To avoid loosing solar data, the RFI system must be turned off during the solar transit.

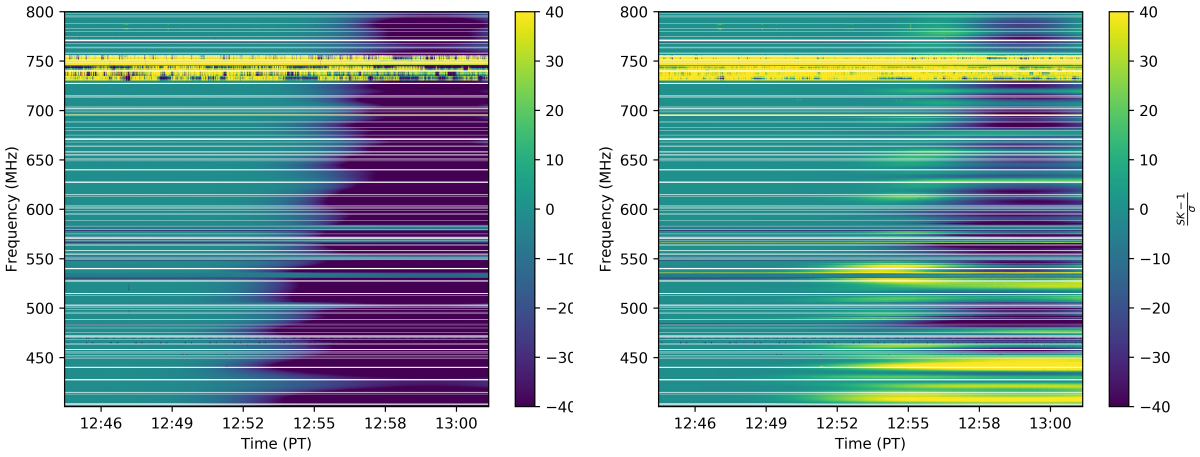

Figure 5.14: Deviation of SK values from unity in terms of nominal σ for CHIME during solar transit (around 13:00). Left: Without truncation bias correction SK values have a very high negative bias. Right: Since SK values sill highly deviate from expected value after truncation bias correction with a fifth degree polynomial, the Sun is still being flagged by the RFI system. This might be explained by the fact that the variance of the estimator is increased when the common mode signal dominates.

5.4.3 RFI polarization

Running the Gaussianity test on all of the inputs reveals that in a few frequencies, one polarization is more Gaussian than the other. This is possibly due to the polarized RFI events. Since kurtosis is estimated by combining the time samples from **all** of the inputs, it can lead to a false estimate of RFI at some frequencies . The following figures show the result of gaussianity test over different inputs. It is clear that one polarization is more non-Gaussian than the other one. One way to deal with the polarized RFI is to estimate the kurtosis for different polarizations separately. This means that instead of 2048 inputs, we will have 1024 inputs to combine. However, reducing number of inputs increases the variance of the estimator, and decreases the sensitivity of the excision algorithms to the RFI. Since it is important to have the highest possible number of samples to keep the variance of the SK estimator as low as possible, we ignore this effect for now.

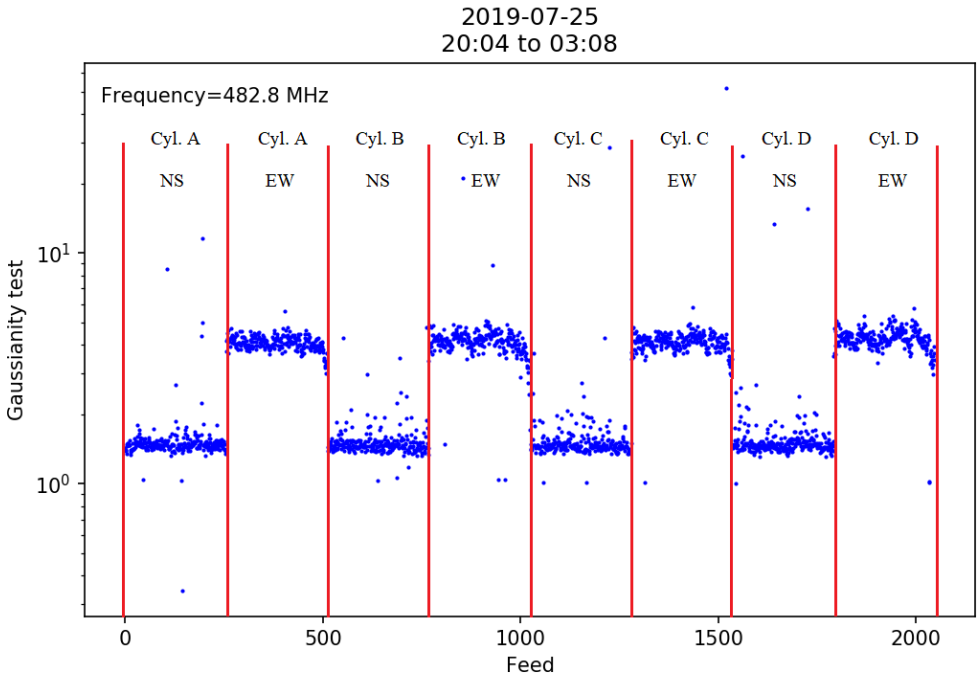

Figure 5.15: Result of the Gaussianity test for all feeds at a single frequency. The NS and EW corresponds to two different polarizations. Note that the first 256 inputs on each cylinder have a different polarization than the next 256.

Chapter 6

Summary and future work

In this thesis I evaluated the performance of the high cadence kurtosis based RFI excision system for CHIME. In general, RFI excision system is able to detect and flag many types of RFI, from repeating RFI lines between 400 MHz and 500 MHz to intermittent RFI from satellites and airplanes. A good example for this is the excision of the Meridian satellite signal. Moreover, the RFI system detects and masks the static RFI bands such as LTE and TV station bands. It is also possible to recover a few frequency channels that were totally non-Gaussian by excising less than %20 of millisecond samples in a 10 second sample. Before excision, these frequency channels were totally contaminated by the RFI. I also discussed the effect of quantization on SK estimates and I corrected the SK estimator for 4+4 bits truncation bias using polynomial fitting and cubic spline interpolation.

Although the RFI detection system detects powerful non-Gaussian RFI events, it must be improved in many ways. First of all, we have to find new detection thresholds and design new algorithms to increase the sensitivity of the RFI system; for example to excise 6 MHz TV channels more effectively. Moreover, truncation bias correction with a polynomial fitting must be replaced by the cubic spline interpolation or by a polynomial of higher degree, as I showed that the cubic spline function gives a much better correction for the truncation bias, and fitting residuals for a polynomial of degree 8 are much smaller than those for 5^{th} degree polynomial. Another issue with the RFI algorithms that must be fixed is the solar transit issue. The number of independent samples from different inputs is reduced during the solar transit. Therefore, the variance of the SK estimator increases and our detection algorithms flag the Sun. To avoid loosing the solar data, we turn off the RFI system during the solar transit. Currently, there is no resolution for this problem, but another potential improvement of the system includes designing an algorithm which does not flag the Sun.

Bibliography

[1] Taylor J., Denman N., Bandura K., Berger P., et al., 2019, JAI, 8, 1940004

[2] Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

[3] Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., et al. 1998, ApJ, 507, 46

[4] Planck Collaboration, Ade P. A. R., Aghanim N., Arnaud M., et al., 2015, A&A 594, A13

[5] Phillips, M. M., 1993, ApJ, 413, L105

[6] Hamuy M., Phillips M. M., Suntzeff N. B. , et al. 1996, AJ, 112, 2398

[7] Regnault, N., Conley, A., Guy, J., et al. 2009 A&A, 506, 999

[8] Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

[9] Frieman, J. A., Bassett, B., Becker, A., et al. 2008, AJ, 135, 338

[10] Sako, M., Bassett, B., Becker, A., et al. 2008, AJ, 135, 348

[11] Sako, M., Bassett, B., Becker, A. C., et al. 2014, arXiv:1401.3317

[12] Betoule, M., Marriner, J., Regnault, N., et al. 2013, A&A, 552, A124

[13] Mosher, J., Guy, J., Kessler, R., et al. 2014, ApJ, 793, 16

[14] Betoule. M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22

[15] Planck Collaboration, Aghanim N., Akrami Y., Ashdown M., et al., 2020, A&A, 641, A6

[16] Weinberg D. H., Mortonson M. J., Eisenstein D. J., et al., 2013, PhR, 530, 87

[17] Landy, S. D. Szalay, A. S., 1993, ApJ, 412, 64

[18] Eisenstein D. J., Zehavi I., Hogg D. W., et al., 2005, ApJ, 633, 560

[19] Blake C., Davis T., Poole G. B., et al., 2011, MNRAS, 415, 2892

[20] Anderson L., Aubourg É., Bailey S., et al., 2014, MNRAS, 441, 24

[21] Busca N. G., Delubac T., Rich J., et al., 2013, AA, 552, A96

[22] Furlanetto S. R., Oh S. P., Briggs F. H., 2006, PhR, 433, 181

[23] Peacock J. A., Schneider P., Efstathiou G., et al. 2006, arXiv:astro-ph/0610906

[24] Bartelmann M., Maturi M., 2017, SchpJ, 12, 32440

[25] de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A.,, Experimental Astronomy, 2013, 35, 25

[26] Joudaki S., Mead A., Blake C., et al., 2017, MNRS, 471, 2, 1259–1279

[27] Joudaki S., Mead A., Johnson A., et al., 2018, MNRS, 474, 4, 4894–4924

[28] Martinet N., Schneider P., Hildebrandt, H., et al., 2018, MNRS, 474, 1, 712–730

[29] Troxel M. A., MacCrann N., Zuntz J., Eifler T. F., et al., 2018, Phys. Rev. D 98, 043528

[30] Aihara, H., Arimoto, N., Armstrong, R., et al., 2018, PASJ,70, SP1, S4

[31] Hikage C., Oguri M., Hamana T., et al., 2019, PASJ, 71, 2, 43

[32] Aihara, H., Armstrong, R., Bickerton, S., et al., 2018, PASJ, 70, S8

[33] Pritchard J. R., Loeb A., 2012, RPPh, 75, 086901

[34] V.D. Vrabie, P. Granjon, C. Servière, "Spectral kurtosis: from definition to application", IEEEEURASIP Workshop on Nonlinear Signal and Image Processing, Grado, Italy, June 8-11, 2003.

[35] Nita G. M., Gary D. E., Liu Z., et al., 2007, PASP, 119, 805

[36] Nita G. M., Gary D. E., 2010, PASP, 122, 595

[37] Fridman P. A., Baan W. A., 2001, AA, 378, 327