Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **501,** 4344–4358 (2021) doi:10.1093/mnras/staa3856
Advance Access publication 2020 December 16

H I intensity mapping with MeerKAT: 1/f noise analysis

Yichao Li [1]★ Mario G. Santos,[1,2] Keith Grainge,[3] Stuart Harper[3] and Jingying Wang [1]

[1]*Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, Belville 7535, South Africa*
[2]*South African Radio Astronomy Observatory (SARAO), 2 Fir Street, Observatory, Cape Town 7925, South Africa*
[3]*Department of Physics and Astronomy, Jodrell Bank Centre for Astrophysics, The University of Manchester, Manchester M13 9PL, UK*

Accepted 2020 December 10. Received 2020 November 27; in original form 2020 July 6

ABSTRACT

The nature of the time correlated noise component (the 1/f noise) of single dish radio telescopes is critical to the detectability of the H I signal in intensity mapping experiments. In this paper, we present the 1/f noise properties of the MeerKAT receiver system using South Celestial Pole tracking data. We estimate both the temporal power spectrum density and the 2D power spectrum density for each of the antennas and polarizations. We apply singular value decomposition to the data set and show that, by removing the strongest components, the 1/f noise can be drastically reduced, indicating that it is highly correlated in frequency. With two-mode subtraction, the knee frequency over a 20 MHz averaging is about 3×10^{-3} Hz, indicating that the system induced 1/f-type variations are well under the thermal noise fluctuations over a few hundred seconds time-scales. We also show that such cleaning on the time ordered data has very little impact on the 21-cm signal itself. The 2D power spectrum shows that the 1/f-type variations are restricted to a small region in the time–frequency space, either with long-wavelength correlations in frequency or in time. This gives a wide range of cosmological scales where the H I signal can be measured without further need to calibrate the gain time fluctuations. Finally, we demonstrate that a simple power spectrum parameterization is sufficient to describe the data and provide fitting parameters for both the 1D and 2D power spectrum.

Key words: methods: statistical – large-scale structure of Universe – cosmology: observations.

1 INTRODUCTION

A major goal of modern cosmology is to understand the formation and evolution of the cosmological large-scale structure (LSS), as well as the information it carries from the early Universe. In the past decades, cosmologists have traced the LSS fluctuations with wide field spectroscopic and photometric surveys of galaxies (Cole et al. 2005; Eisenstein et al. 2005; Anderson et al. 2014; Hinton et al. 2017). However, these surveys are often limited with either cosmologically small volumes or lower sampling density. Furthermore, detecting individual objects at high significance is time consuming.

Recently, the 21-cm emission line of neutral hydrogen (H I) hyperfine spin–flip transition, has been proposed as another cosmological probe of the LSS (e.g. Battye, Davies & Weller 2004; McQuinn et al. 2006; Pritchard & Loeb 2012). Instead of observing the H I emission line from individual galaxies, cosmologists proposed to measure the total H I intensity of the galaxies within large voxels, a technique known as H I intensity mapping (IM; Chang et al. 2008; Loeb & Wyithe 2008; Mao et al. 2008; Pritchard & Loeb 2008; Wyithe & Loeb 2008; Wyithe, Loeb & Geil 2008; Peterson et al. 2009; Bagla, Khandai & Datta 2010; Seo et al. 2010; Lidz et al. 2011; Ansari et al. 2012; Battye et al. 2013). Because of the low angular resolution requirement, an H I IM survey can be quickly carried out with single dishes and extended to very large survey volumes. The H I IM technique was explored with the Green Bank Telescope (GBT) by measuring the cross-correlation function between an H I IM survey and an optical galaxy survey (Chang et al. 2010). Later, the cross-correlation power spectrum between an H I IM survey and an optical galaxy survey was also reported with the GBT and Parkes telescopes (Masui et al. 2013; Wolz et al. 2017; Anderson et al. 2018), while the H I IM autopower spectrum remains undetected (Switzer et al. 2013). There are several planned H I IM experiments targeting the post-reionization epoch, such as the Tianlai project (Chen 2012), the Canadian Hydrogen Intensity Mapping Experiment (CHIME; Bandura et al. 2014), the Baryonic Acoustic Oscillations from Integrated Neutral Gas Observations (Battye et al. 2013), and the Hydrogen Intensity and Real-Time Analysis experiment (Newburgh et al. 2016). The SKA has also been proposed as a major instrument to probe cosmology using this technique (Bull et al. 2015; Santos et al. 2015; Square Kilometre Array Cosmology Science Working Group 2020). Recently, it was also proposed to have an H I IM survey with the newly built MeerKAT telescope in single-dish mode (Santos et al. 2017).

There are several challenges for H I IM power spectrum detection. The primary challenge is to remove the bright continuum radiation of the Milky Way and extragalactic galaxies. The continuum radiation foreground is known to have a smooth frequency spectrum and can be extracted by fitting the spectrum with low order polynomial functions (Mao 2012). However, due to instrumental effects, the smooth-spectrum assumption breaks down and the foreground signal leaks into higher order fluctuation modes. Several foreground cleaning methods have been proposed to try to address this (Alonso, Ferreira & Santos 2014; Wolz et al. 2015) and used in the analysis of GBT and Parkes H I IM survey (Switzer et al. 2015; Wolz et al. 2017).

★ E-mail: dr.yichao.li@gmail.com

H I IM measurements also require the receiver system to be stable. However, the receiver system noise is known to have time correlated fluctuations, the so-called 1/f-type noise (1/f noise). Such 1/f noise injects long-range correlations in time and leads to stripes in the final IM map. Since the measurements are performed in single dish mode (autocorrelation), they do not benefit from the suppression of 1/f noise afforded by interferometric measurements. The 1/f noise effect has been discussed in previous analyses of cosmic microwave background (CMB) experiments (Janssen et al. 1996). Several different destriping methods have been proposed and tested with the analysis of CMB data (Maino et al. 2002; Seiffert et al. 2002; Keihänen et al. 2004; Kurki-Suonio et al. 2009; Sutton et al. 2010).

The effect of 1/f noise on an H I IM survey has been analysed through simulations (Bigot-Sazy et al. 2015; Harper et al. 2018). In the case of H I IM, the data are collected across multiple frequency channels. However, the correlation of 1/f noise across frequency is currently not very well understood. In this work, we develop a 1/f noise power spectrum density estimator to extract the temporal and spectroscopic 1/f noise properties of the MeerKAT receiver system using astronomical observation data. We also apply singular value decomposition (SVD) to the data in order to reduce the 1/f-type fluctuations. The paper is organized as follows: Our power spectrum density analysis method and the 1/f noise model are introduced in Section 2; the details of observation data are given in Section 3, the SVD method is introduced in Section 4, a mask filling method is introduced in Section 5 to reconstruct the missing data due to the Radio Frequency Interference (RFI) flagging, the results are discussed in Section 6, and the conclusions are summarized in Section 7.

2 1/F NOISE POWER SPECTRUM DENSITY MODEL

2.1 Temporal power spectrum density

The time-ordered data (in arbitrary units) as a function of time t and frequency ν, $d(t, \nu)$, can be modelled as the input temperature, $T_{\rm in}(t, \nu)$, multiplied by the gain, $G(t, \nu)$:

$$d(t, \nu) = G(t, \nu)T_{\rm in}(t, \nu) + n(t, \nu), \tag{1}$$

where $n(t, \nu)$ represents the white noise term (a Gaussian variable uncorrelated in time and frequency). The input temperature can be expressed as $T_{\rm in} = (T_{\rm ext} + T_{\rm rx})$, where $T_{\rm ext}$ is the external temperature (sky, atmosphere and ground effects), convolved by the telescope primary beam, and $T_{\rm rx}$ is the receiver temperature. The gain, $G(t, \nu)$, refers to the gain of the amplifiers in the receiver.

There are several sources of temporal fluctuations in equation (1). First, we have the usual sky fluctuations. Since in these observations the telescope is fixed and pointing at the South Celestial Pole (SCP), we do not expect to see much variation, except as a result of point sources or our galaxy moving in and out of asymmetries in the primary beam (rising and setting could also be an issue but the contribution will be about 30° from the beam centre). With the MeerKAT beam model (Asad et al. 2019), as well as the sky model from PyGSM,[1] we found that the temperature fluctuation due to the beam asymmetries is at the ~ 0.1 per cent level over the 2.5 h of observation. Likewise, we do not expect fluctuations due to the atmosphere or ground pick-up during this observation time since the dish is fixed. Moreover, these contributions should be smooth in frequency and therefore removed in the cleaning process.

We then have instrumental fluctuations. There are changes in the receiver temperature, gain as well as the intrinsic white noise fluctuations mentioned above that should integrate down in time. The receiver temperature is dominated by the contribution from the first low-noise amplifier (LNA). This LNA is located within the cryostat whose physical temperature is relatively stable due to its large thermal inertia. The temperature injected by cryogenic LNAs is found to only have a weak dependence on physical temperature once cooled below 50 K. We therefore expect $T_{\rm rx}$ to be quite stable over the time-scales of our study. Also, the slow gain drifts (over hour scales) are expected to be calibrated out either through noise diodes or sky calibrators. We finally have the 'non-calibrated', correlated gain fluctuations that are the focus of this paper (the 1/f noise). This correlated noise has simple statistical properties (at least on time-scales $\lesssim$1 h), which we describe next.

In order to model the fluctuations, we define, $G(t, \nu) \equiv \bar{G}(\nu) + \delta G(t, \nu)$, $T_{\rm ext}(t, \nu) \equiv \bar{T}_{\rm ext}(\nu) + \delta T_{\rm ext}(t, \nu)$, and $T_{\rm rx}(t, \nu) \equiv \bar{T}_{\rm rx}(\nu) + \delta T_{\rm rx}(t, \nu)$, where we use an overbar to define the time-averaged quantities. We subtract and divide the data by its time average, $\bar{\rm d}(\nu)$, taking only the varying part for the rest of the analysis: $\delta_{\rm d}(t, \nu) \equiv \frac{{\rm d}(t,\nu)}{\bar{\rm d}(\nu)} - 1$. By dividing by $\bar{\rm d}(\nu)$ we also cancel out the frequency dependence both from the sky and instrument (the part that is stable in time, e.g. the bandpass). To first order we can then write

$$\delta_{\rm d}(t, \nu) \approx \frac{\delta T_{\rm ext}(t, \nu)}{\bar{T}_{\rm in}(\nu)} + \frac{\delta T_{\rm rx}(t, \nu)}{\bar{T}_{\rm in}(\nu)} + \frac{\delta G(t, \nu)}{\bar{G}(\nu)} + \frac{n(t, \nu)}{\bar{T}_{\rm in}(\nu)\bar{G}(\nu)}. \tag{2}$$

As described above, we expect fluctuations in the gain, $\frac{\delta G(t,\nu)}{\bar{G}(\nu)}$, to dominate this. However, our analysis does not make any distinction between the different components and we could as well assume that the fluctuations in the receiver temperature or even the sky, are included. In fact, as we will see later, we believe that the first mode is dominated by the sky contribution, possibly due to a small offset from the SCP. To that effect, we can claim that our results are upper limits as they can include more fluctuations than just the classical 1/f noise.

We are going to assume that we can model the 'non-calibrated' instrument fluctuations through a correlated Gaussian distribution (both in time and frequency). These can be characterized by the temporal power spectrum density function, $\hat{S}^t(f, \nu)$, which is estimated via the Fourier transfer of $\delta_d(t, \nu)$ along the time axis as

$$\hat{S}^t(f, \nu) = \left| \sqrt{\frac{\delta t}{N_t}} \sum_{p=0}^{N_t-1} \delta_{\rm d}(p\delta t, \nu) \exp[-2\pi i f p \delta t] \right|^2, \tag{3}$$

in which, f is the temporal frequency; δt is the time resolution of the data ($t = p\delta t$), and N_t the number of time samples.

If we only have the white noise fluctuations, we can write

$$S^t(f, \nu) \equiv \langle \hat{S}^t(f, \nu) \rangle = \frac{\delta t}{N_t} \sum_{p=0}^{N_t-1} \left\langle \delta_d^2(p\delta t, \nu) \right\rangle \tag{4}$$

$$= \delta t \frac{\sigma_n^2}{\bar{T}_{\rm in}^2(\nu)\bar{G}^2(\nu)} = \frac{\delta t}{\bar{T}_{\rm in}^2(\nu)} \frac{\bar{T}_{\rm in}^2(\nu)}{\delta t \delta \nu} = \frac{1}{\delta \nu}, \tag{5}$$

where σ_n^2 is the white noise variance and $\sigma_n^2/\bar{G}^2(\nu) = \frac{T_{\rm in}^2}{\delta\nu\delta t}$ for a frequency bin width of $\delta\nu$ and integration time of δt (e.g. Wilson, Rohlfs & Hüttemeister 2009). Please note that, since $\delta_{\rm d}$ is normalized with the system temperature, the power spectrum density is normalized with $T_{\rm sys}^2$. The power spectrum in the time direction from the extra

[1] git@github.com:telegraphic/PyGSM.git

1/f-type noise component can then be modelled as (Harper et al. 2018)

$$S^t_{\rm fn}(f,\nu) = \frac{1}{\delta\nu}\left(\frac{f_k}{f}\right)^{\alpha}, \tag{6}$$

where α is the spectral index of the noise. This enforces that for large f the 1/f noise power spectrum goes to zero and the overall power spectrum becomes dominated by white noise. Our model for the full temporal power spectrum density function is then

$$S^t(f,\nu) = \frac{A}{\delta\nu}\left(1+\left(\frac{f_k}{f}\right)^{\alpha}\right), \tag{7}$$

where A ($\sim$1) is as a free parameter fit together with α and f_k.

2.2 2D power spectrum density

The 1/f noise can potentially be correlated in frequency. We need therefore to consider a 2D power spectrum to fully describe its statistics. This 2D power spectrum density can be estimated by Fourier transforming the observed time-ordered data along the time and frequency axes:

$$\hat{S}(f,\tau) = \left|\sqrt{\frac{\delta t\delta\nu}{N_t N_\nu}}\sum_{p=0}^{N_t-1}\sum_{k=0}^{N_\nu-1}\delta_d \exp\left[-2\pi i\left(fp\delta t+\tau k\delta\nu\right)\right]\right|^2, \tag{8}$$

in which f is the temporal frequency and τ is the spectroscopic frequency (the Fourier conjugate in the frequency domain). In this case, if we only have white noise fluctuations, $\hat{S}(f,\tau)=1$. Following Harper et al. (2018), we then build an empirical 2D power spectrum density model as

$$S_{\rm fn}(f,\tau) = F(f)H(\tau), \tag{9}$$

where we assume that the correlations in time and frequency are separable and only a function of $|t-t'|$ and $|\nu-\nu'|$. In reality, the system can be more complex than this, especially if we have non-linearities and ultimately we need to check the validity of our model through the data itself as we will see later. In the equation above, $F(f)$ describes the temporal correlation power spectrum,

$$F(f) = \frac{1}{\delta\nu}\left(\frac{f_{\rm k}}{f}\right)^{\alpha}, \tag{10}$$

with $f_{\rm k}$ the knee frequency defined at the frequency resolution of $\delta\nu$). $H(\tau)$ is the spectroscopic correlation power spectrum density, which can be modelled as

$$H(\tau) = \left(\frac{\tau_0}{\tau}\right)^{\frac{1-\beta}{\beta}}, \tag{11}$$

where β specifies the amount of correlation across frequencies and $\tau_0 = 1/(N_\nu\delta\nu)$. Combining the white noise term, the 2D power spectrum model can be expressed as

$$S(f,\tau) = A\left(1+\frac{1}{K\delta\nu}\left(\frac{f_k}{f}\right)^{\alpha}\left(\frac{\tau_0}{\tau}\right)^{\frac{1-\beta}{\beta}}\right), \tag{12}$$

in which, $K=\int {\rm d}\,\tau\,{\rm sinc}^2(\pi\delta\nu\tau)\left(\frac{\tau_0}{\tau}\right)^{(1-\beta)/\beta}$. The derivation of equation (12) is shown in the appendix. $A\sim 1$ due to the normalization with $T^2_{\rm sys}$. In our analysis, A is set as an overall amplitude parameter that can be constrained by the observation data together with f_0, α, and τ_0.

The knee frequency $f_{\rm k}$ as a function of frequency resolution is an important consideration for the LSS correlation signal on the largest scale sizes. For example, if we are interested in line-of-sight scales of $\sim 100\,{\rm Mpc}\,h^{-1}$, at 900 MHz (i.e. $z\sim 0.6$), this corresponds to frequency scales of $\sim$ 25 MHz. Depending upon the knee frequency, at the $\sim$ 25 MHz frequency resolution there is the potential to detect the 1/f noise more significantly than at lower values of the frequency resolution. The knee frequency at two different frequency resolutions, $\delta\nu$, $\delta\nu'$, is related via

$$\lg f_k = \lg f_{k'} + \frac{1}{\alpha}\lg\left(\frac{K\delta\nu}{K'\delta\nu'}\right). \tag{13}$$

The derivation is shown in the appendix. We test the shift of the knee frequency with simulated time-ordered data. Fig. 1(a) shows the waterfall plots of the simulated time-ordered data with different frequency correlation properties. As $\beta\rightarrow 0$, the 1/f noise becomes fully correlated over the frequency band. As $\beta\rightarrow 1$, the frequency correlation length is reduced and the 1/f noise between different frequencies becomes independent (down to the frequency resolution).

The corresponding temporal power spectrum of the simulated data is shown in Fig. 1(b). The black solid curve shows the power spectrum with 0.1 MHz frequency resolution, which is the raw frequency resolution of the simulation, while the red curve shows the power spectrum after averaging over 10 frequency channels. The dashed curves show the simulation with only 1/f noise (set white noise level to 0). The horizontal lines indicate the white noise level with frequency resolution of 0.1 MHz (blue) and 1 MHz (green), respectively. The cross-points with the vertical lines indicate the knee frequency at the corresponding frequency resolution, which is estimated with equation (13). The white noise floor, as expected, is reduced by one order of magnitude after frequency averaging. However, the 1/f noise level behaves differently with different β values. In the case of $\beta=0$, the 1/f noise is fully correlated over the frequency band. The level of 1/f noise power spectrum does not change with averaging frequency channels, but the white noise does. The different behaviour between 1/f noise and white noise results in a higher knee frequency value at lower frequency resolution. With β increasing, the 1/f noise behaves more like the white noise. In the case of $\beta=1$, the 1/f noise if fully uncorrelated between frequencies and the power spectrum level is reduced by one order of magnitude as well. In this case, the knee frequency does not change with frequency resolution.

2.3 Parameter fitting

The parameters that characterize the 1/f noise can be constrained by fitting the model against the measured noise power spectrum. We build the χ^2 function both for temporal and the 2D power spectrum density function:

$$\chi^2_t = \sum_f \frac{\left(\langle\hat{S}^t(f)\rangle_\nu - S^t(f)\right)^2}{\sigma^2_{\hat{S}^t}}, \tag{14}$$

$$\chi^2 = \sum_{f,\tau}\frac{\left(\hat{S}(f,\tau)-S(f,\tau)\right)^2}{\sigma^2_{\hat{S}}}, \tag{15}$$

in which, $\langle\,\rangle_\nu$ represents the average over the frequency channels and $\sigma_{\hat{S}^t}$, $\sigma_{\hat{S}}$ are the estimated errors of the temporal and 2D power spectrum density, respectively. The errors of the temporal power spectrum density are estimated via the standard deviation of the power spectrum density using different frequency channels:

$$\sigma^2_{\hat{S}^t} = \frac{1}{N_\nu}\left(\langle(\hat{S}^t(f))^2\rangle_\nu - (\langle\hat{S}^t(f)\rangle_\nu)^2\right), \tag{16}$$

where N_ν is the number of frequency channels.

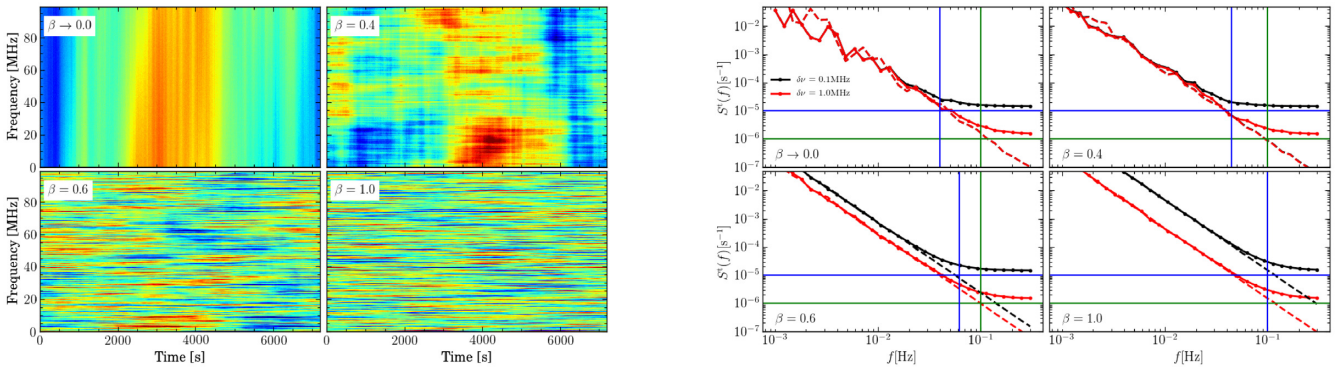

Figure 1. (a) The waterfall plot of the simulated time-ordered data using the model of equation (12) as the input power spectrum. The simulation uses 1/f noise parameters: $f_k = 0.1$ Hz at frequency resolution $\delta\nu = 1$ MHz, $\alpha = 2.5$ and different values of β as shown in each panel. $\beta = 1$ corresponds to uncorrelated 1/f noise across frequency, while $\beta \to 0$ is fully correlated (one can still see a fluctuation since we cannot simulate $\beta = 0$ exactly). (b) The temporal power spectrum density of the simulated time-ordered data as shown in Fig. 1(a). The black-dashed curve shows the power spectrum of pure 1/f noise simulation (set white noise equals to 0) with frequency resolution of 0.1 MHz and the solid curve shows the simulation with white noise added; The red dashed/solid curve show the same simulations with frequency resolution reduced to 1 MHz. The horizontal lines indicate the white noise level with frequency resolution of 0.1 (blue) and 1 MHz (green), respectively. The cross-points with the vertical lines indicate the knee frequency at the corresponding frequency resolution, which is estimated with equation (13). For $\beta = 1$ (fully uncorrelated), we expect the 1/f noise power spectrum to be inversely proportional to the frequency resolution.

The error on the 2D power spectrum density is estimated with jackknife samples (Efron 1982), which is widely used in LSS surveys. Following the jackknife method, the total time-ordered data is evenly divided into $N = 200$ sub-time blocks. By dropping one sub-time block at a time, we estimate the 2D power spectrum density using the remaining $N - 1$ sub-time blocks and obtain N realizations as the jackknife samples. The errors are then estimated from the jackknife samples as

$$\sigma^2_{\hat{S},\mathrm{JK}} = \frac{N-1}{N}\sum_{k=1}^{N}\left(\hat{S}^k(f,\tau) - \langle\hat{S}^k(f,\tau)\rangle_k\right)^2, \tag{17}$$

where $\hat{S}^k(f,\tau)$ is the 2D power spectrum density estimated using the kth jackknife sample and $\langle\,\rangle_k$ represents the average over the jackknife samples. The jackknife errors are then used in the parameter fitting.

We constrain the free parameters by minimizing the χ^2 function, using the publicly available Markov Chain Monte Carlo algorithm `emcee` (Foreman-Mackey et al. 2019).

3 OBSERVATIONS AND DATA REDUCTION

Details on the MeerKAT telescope can be found in Jonas & MeerKAT Team (2016), Camilo et al. (2018), and Mauch et al. (2020). In order to characterize the 1/f-type fluctuations of the system noise, we need a constant input signal and a long-duration observation. Both requirements can be satisfied by tracking the SCP over several hours. Two SCP data sets are used in our analysis. One was collected in 2016 with a few antennas; the other was collected in 2019 using the majority of the MeerKAT array.

2016 SCP Data The data in 2016 (SCP16) were observed with three of the MeerKAT antennas, named M017, M021, and M036, pointing at the SCP. The observation started with a 20 Hz sampling rate for 3.5 min (Experiment ID 20160922-0004), followed by 1 Hz sampling rate for 2 h (Experiment ID 20160922-0005). The frequencies range from 856 to 1711.791 MHz, with 4096 frequency channels and 0.209 MHz frequency resolution.

2019 SCP Data The SCP tracking data in 2019 (SCP19) was observed on April 24 (Experiment ID 20190424-0024) using ∼60 antennas over 2.5 h. The data were taken with a 0.5 Hz sampling rate. The frequency range and resolution for the data in 2019 are the same as in 2016.

Fig. 2 shows the frequency spectrum of the SCP tracking data, averaged over the observation time. The top panel shows the spectrum of all three antennas used in the SCP16. The bottom panel shows the spectrum of the first six antennas used in the SCP19 observation. The amplitude shown here is the uncalibrated raw detector output power. The scattering of the amplitude across antennas is mainly due to the different digital gain settings. We use the relatively RFI-free frequency range between 1294.8672 and 1503.8516 MHz for the rest of the analysis but exclude the H I signal from our galaxy at 1420MHz.

As discussed in Section 2, we are interested in the fluctuations around the time average, $\delta_{\mathrm{d}}(t,\nu) = \frac{\mathrm{d}(t,\nu)}{\bar{\mathrm{d}}(\nu)} - 1$. We then need to normalize the data by its time average. Note, however, that we use median values instead of the mean values to avoid time-varying RFI. We expect the receiver temperature as well as most external sources to be constant during the observation time since the telescope is fixed and observing the SCP. Therefore, this normalization should calibrate out the telescope bandpass as well as most spectral features from the sky, ground pick-up and atmosphere. The remaining time and frequency fluctuations in $\delta_{\mathrm{d}}(t,\nu)$ are expected to be from 1/f and white noise. However, some fluctuations can still be present from sources rising and setting and due to primary beam asymmetries. The waterfall plots of the normalized data are shown in Fig. 3.

Three antennas from SCP16 at 1 Hz sampling rate are shown in the top left-hand, top right-hand, and bottom left-hand panels of Fig. 3, while one antenna from SCP19, as an example, is shown in the bottom left-hand panel. In each panel, the two polarizations are shown in the top and bottom sub-panels. The colour scale is restricted to run between the mean value plus/minus two times the root mean square (rms) of the data shown. The amplitude varies over both frequency and observation time. The variation of SCP16 data is around 0.3 per cent of the mean, T_{sys}, which is much less than the SCP19 data. The SCP19 data show strong variations in time while showing strong correlations across frequency. Such frequency-

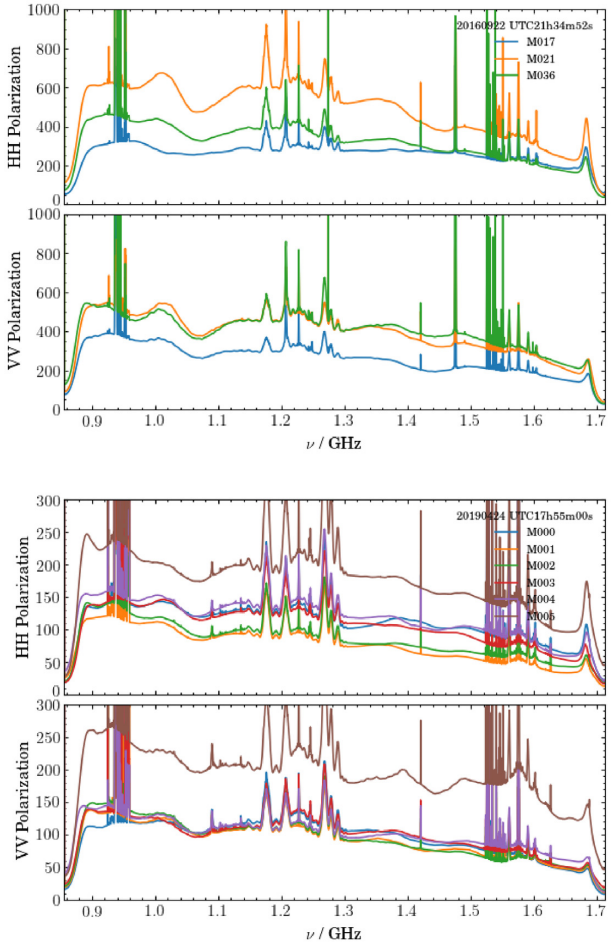

Figure 2. The frequency spectrum of SCP tracking data. The top panel shows the spectrum of the SCP16 20Hz data; the bottom panel shows the SCP19 data. In each panel, the top/bottom sub-panels show the two polarizations. The *y*-axis is the uncalibrated raw detector output power.

correlated variations are synchronous across different antennas, which indicates a larger overall environment variation during the 2019 observation.

4 TIME-ORDERED SVD

We apply SVD to the time-ordered data in order to extract its strongest components. If the 1/f noise is strongly correlated in frequency, we expect the SVD will be able to remove it while keeping the H I signal unaffected on the scales of interest. Foregrounds should also be removed in this process. SVD corresponds to the following matrix decomposition:

$$\mathbf{D} = \mathbf{U}\Lambda\mathbf{V}^{\dagger}, \tag{18}$$

where $\mathbf{D}$ is the data matrix with shape of $n_t \times n_\nu$; the columns of $\mathbf{U} = \{u_0, ..., u_{n_t}\}$ and $\mathbf{V} = \{v_0, ..., v_{n_\nu}\}$ are the temporal and spectroscopic modes and $\mathbf{\Lambda}$ is a $n_t \times n_\nu$ diagonal singular value matrix. The symbol † denotes the conjugate transpose. Both the temporal and spectroscopic modes are sorted according to their singular values, and the first N_m modes are subtracted:

$$\mathbf{D}^\mathrm{c} = \mathbf{D} - \mathbf{U}\Lambda'\mathbf{V}^{\dagger}, \tag{19}$$

where $\mathbf{\Lambda}'$ equals $\mathbf{\Lambda}$ with diagonals beyond N_m set to 0. It can be further expressed as

$$\mathbf{D}^\mathrm{c} = \mathbf{D} - \sum_{i=0}^{N_\mathrm{m}} u_i(u_i^{\dagger}\mathbf{D}) \tag{20}$$

$$= \mathbf{D} - \sum_{i=0}^{N_\mathrm{m}} (\mathbf{D}v_i)v_i^{\dagger}, \tag{21}$$

where u_i, $u_i^{\dagger}$, v_i, and $v_i^{\dagger}$ are the elements of $\mathbf{U}$, $\mathbf{U}^{\dagger}$, $\mathbf{V}$, and $\mathbf{V}^{\dagger}$ respectively.

The singular values of the SCP data sets are shown in Fig. 4, normalized with the first (largest) singular value. The first five singular values are shown on the left of the black-dashed line and the rest are shown on the right. It is clear that after the first five modes there is little variation in the amplitude. Therefore, we restrict the analysis in this paper to the first five modes. The temporal and spectroscopic modes for SCP16 and SCP19 are shown in Figs 5 and 6, respectively. In each panel, the two polarizations are shown in the upper/lower sub-panels with the blue/green colours and the temporal/spectroscopic modes shown in the left/right sub-panels, respectively. The areas shown in yellow are the re-filled values due to the RFI flagging, and the orange curves are the Wiener filtered smooth curves. We will discuss our mask-filling strategy in Section 5.

The 20 and 1 Hz data of SCP16 have significant differences in the singular values, as well as the singular modes. The left-hand panel of Fig. 4 shows the singular values of all three antennas of both the 20 and 1 Hz data of SCP16; the singular values of the 1 Hz data decrease more quickly than those of the 20 Hz data. This is because the 20 Hz data have much shorter observation time than the 1 Hz data and the SVD modes are dominated by the system noise. This difference can also be seen with the SVD modes in Fig. 5. The two panels of Fig. 5 show the SVD modes of the antenna M017 of SCP16 data, where the 1 Hz data are on the left and the 20 Hz data are on the right. We can see that, at least for the first three modes, both for the temporal and spectroscopic modes, the 1 Hz data have clear variation shapes; but the modes of 20 Hz data are mostly noise dominated.

The singular values of the data of SCP19 are shown in the right-hand panels of Fig. 4. Only the first six antennas are shown here as examples; the rest of the antennas have the same trend as these first six antennas. The SVD modes of the SCP19 data are shown in the Fig. 6, in which the left-hand and right-hand panel show the modes of two different antennas as examples. Similar to the SCP16 1 Hz data, with long enough observation time, the system variations both along time and frequency are well represented with the first several singular modes.

5 POWER SPECTRUM IN THE PRESENCE OF RFI FLAGGING

The gaps due to the RFI flagging result in significant window function effects in the final power spectrum. To reduce the impact of this effect in our analysis, we fill the missing data with values reconstructed with the SVD modes. The mask-filling strategy is described below.

First of all, some data are removed across either whole frequency channels or whole time samples. Frequency channels that are contaminated by some known narrow band RFI are fully masked for the whole observation time. The H I emission line of the Milky Way, which is in our selected frequency band, is also removed. On the other hand, some short time duration occasional RFI, for example, due to transiting satellites, are masked across the whole frequency band. By ignoring the masked frequency channels and time stamps, the rest of the data are merged into a continuous frequency-time matrix. SVD is applied to this merged data.

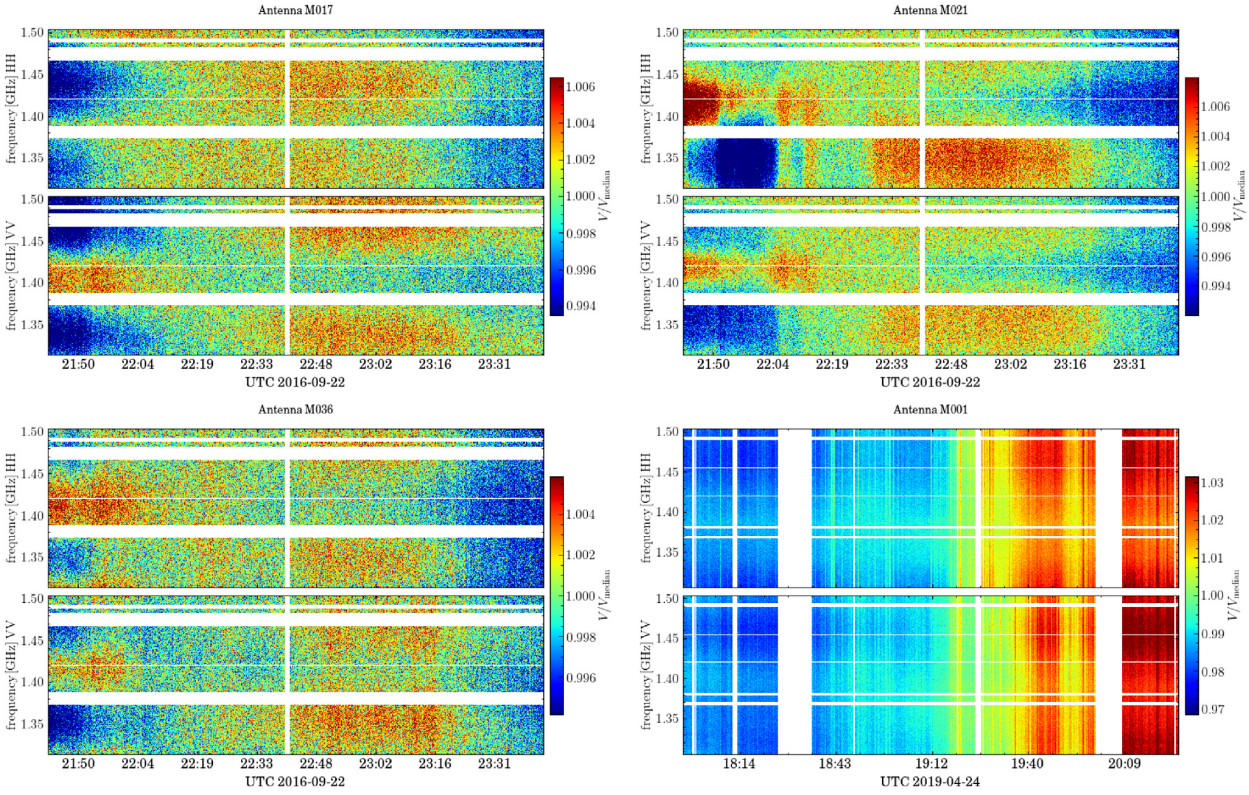

Figure 3. The waterfall plots of the SCP tracking data. The SCP16 1 Hz data observed by antenna M017, M021, and M036 are shown in the top left-hand, top right-hand, and bottom left-hand panels, respectively; The SCP19 data observed by M001, as an example, is shown in the bottom right-hand panel (0.5 Hz sampling rate). The frequency range is truncated between 1294.8672 and 1503.8516 MHz, which is relatively RFI-free. The power amplitude of each frequency is normalized with the median values along the time axis. The colour scale is restricted between the mean value of the shown data plus/minus two times of the rms of the data.

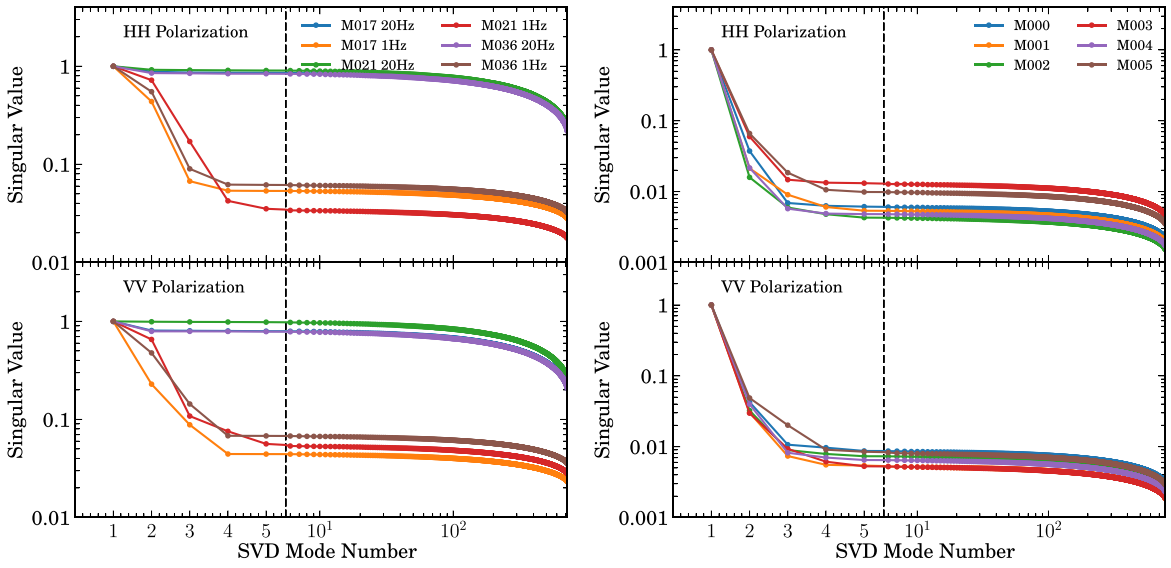

Figure 4. The singular values of each of the SVD modes, normalized to the first singular value. The black-dashed line indicates the first five singular values. The left-hand panel shows the singular values of the SCP16 data, both for 1 and 20 Hz sampling rate, observed with different antennas. The right-hand panels show the singular values of data from April 24, SCP19. The results of the first six antennas are shown in different colours. For the 20 Hz data, the changes in amplitude are small since the modes are dominated by noise.

We take the resulting first five temporal and spectroscopic modes, which have larger singular values than the noise modes, and fill the masked regions with linear interpolation. The interpolation is applied for each of the temporal and spectroscopic modes individually. However, if the singular modes are noise dominated, such as the first several modes of SCP16 20 Hz data, the interpolation is ignored and the masked region is filled with the mean value of the singular mode. Then, we apply a Wiener filter to each of the masked-filled

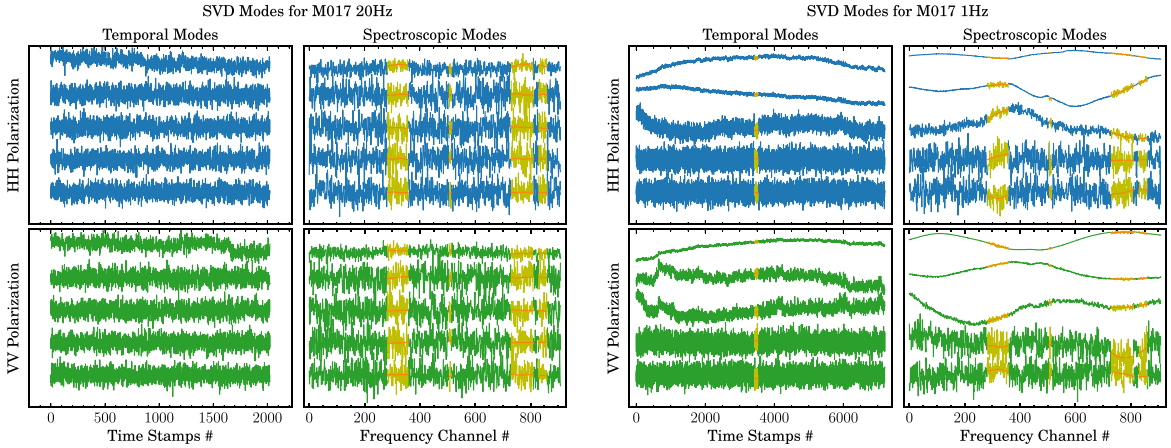

Figure 5. The singular modes of each of the first five SVD modes, of antenna M017, as an example, of the SCP16 data, with results of 20 Hz data on the left-hand panel and 1 Hz data on the right. The HH and VV polarization are shown in the top and bottom sub-panels, respectively; and the temporal and spectroscopic modes are shown in the left-hand and right-hand sub-panels, respectively. The areas shown in yellow are the re-filled values due to the RFI flagging, and the orange curves are the Wiener filtered smooth curves. We discuss our mask-filling strategy in Section 5).

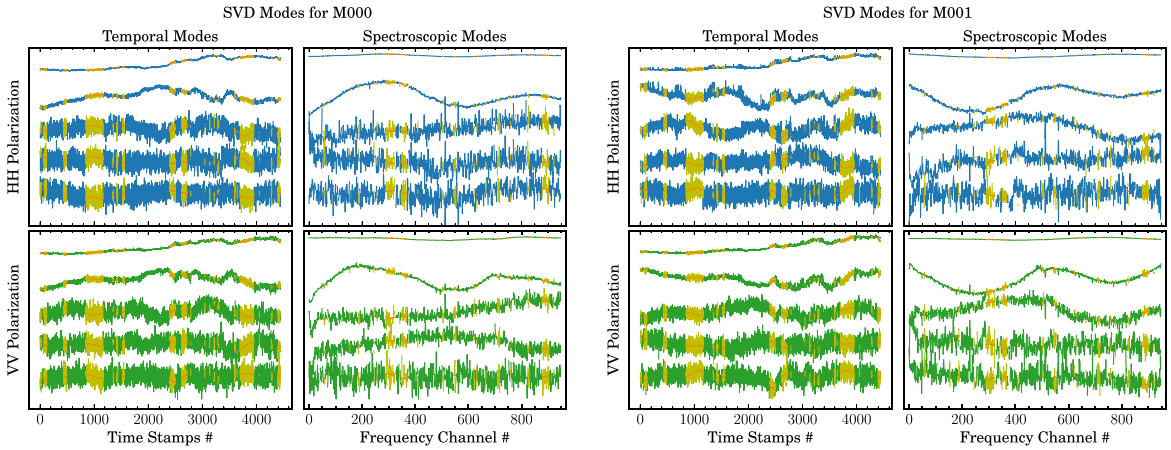

Figure 6. Same as Fig. 5 but shown with the modes of antenna M000 (left) and M001 (right) of the SCP19 data. The two polarizations are shown in the top and bottom sub-panels; and the temporal and spectroscopic modes are shown in the left-hand and right-hand sub-panels, respectively.

singular modes to make the filled values smoothly connect with the unmasked region. The smoothed filling values are shown with the orange curves in Figs 5 and 6.

We then subtract the Wiener-filter-smoothed singular modes from the five modes, estimate the rms of the residuals for each mode, and add random noise to the filling values according to the residual rms of each mode. The noise-added filling values are shown in yellow in Figs 5 and 6. Using only these five modes we construct a new data set (through equation 18) that now has values in the missing gaps. We could be tempted to use these values to fill the flagged gaps in our original data. However, this reconstructed data with the first five modes still has noise missing (from the remaining modes). In order to account for this, we subtract this new data set from the original data set and estimate an overall noise rms using the non-flagged part of the data. We then add random noise to the new data set using this rms and use its values to fill the flagged gaps in the original data. The original masked data and masked filled data of SCP19 (antenna M001) are shown in the top left-hand and top right-hand panels of Fig. 7, respectively.

Finally, we perform SVD a second time to the mask-filled data. The waterfall plots with the first one, two, and five modes removed are shown in the bottom left-hand, bottom middle, and bottom right-hand panels of Fig. 7. The corresponding first five modes are shown in Fig. 8. The power spectrum estimation discussed in Section 6 is performed with the mask-filled data.

6 RESULTS AND DISCUSSION

6.1 The temporal power spectrum density

We now focus on the analysis of the power spectrum along the time domain and how it compares to our model. The temporal power spectrum is estimated by Fourier transforming the data along the time axis. Before the power spectrum estimation, we reduce the frequency resolution down to ~ 20 MHz by averaging every 100 frequency channels. The frequency averaging can reduce the white noise level and shift the knee frequency to the high end of f. However, as we discussed before, the shift of the knee frequency is also dependent on the frequency correlation of the 1/f noise.

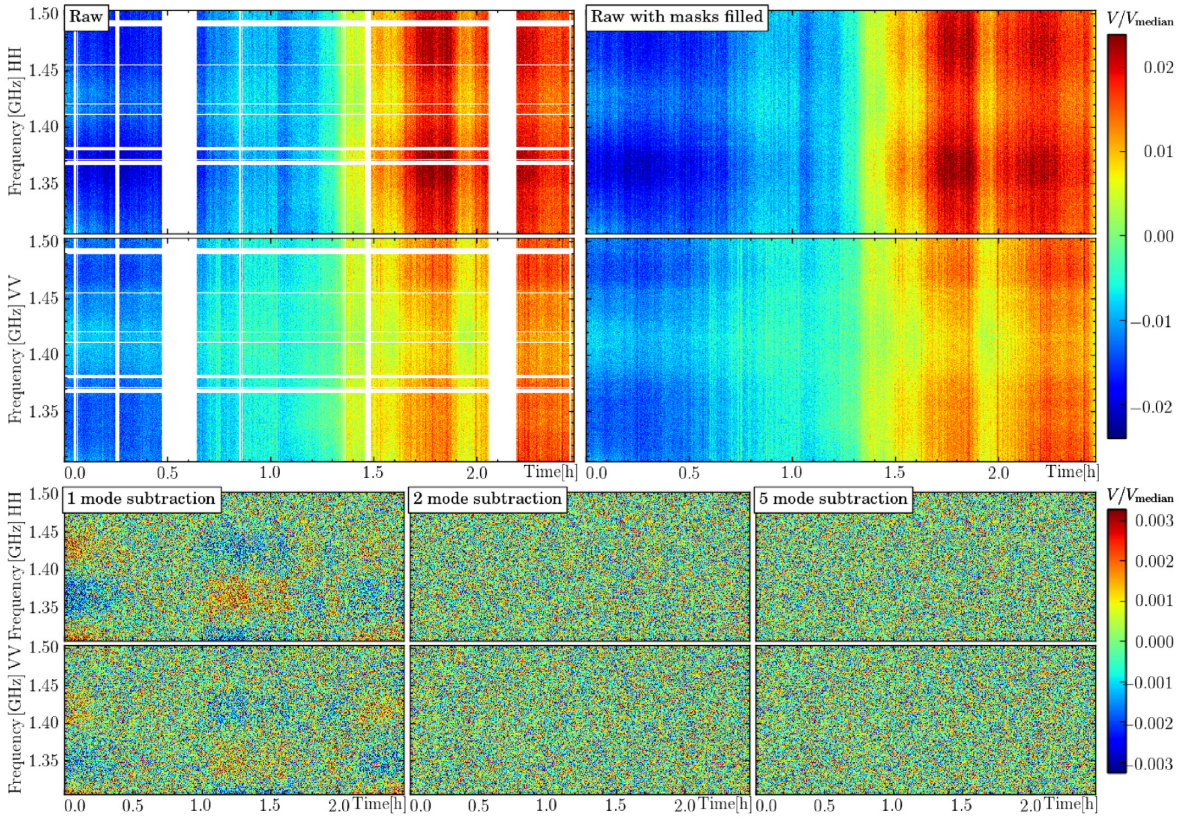

Figure 7. The top left-hand panel shows the waterfall plots of the mean-subtracted data before SVD mode subtraction; The top right-hand panel shows the same data with the RFI masks filled; The bottom left-hand, bottom middle, and bottom right-hand panels are the waterfall plots of the mask filled data with one, two, and five SVD modes subtracted. All these data are the SCP19 observation with antenna M001.

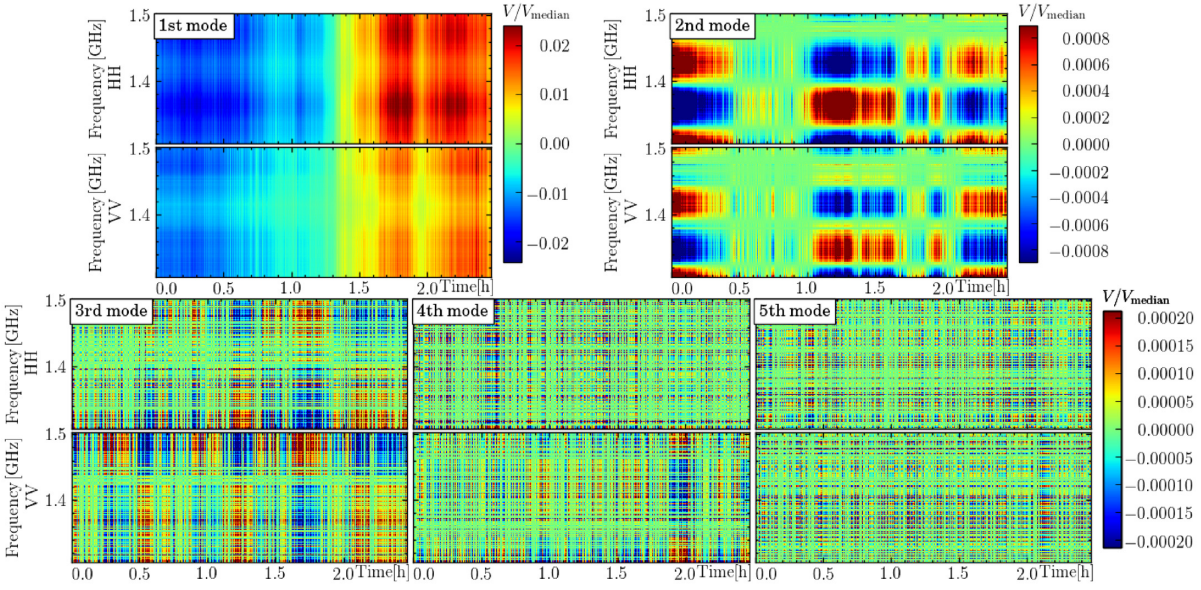

Figure 8. The waterfall plots of the first five SVD modes from the mask filled data. All these data are the SCP19 observation with antenna M001.

The temporal power spectrum density results of SCP16 data observed with the three antennas are shown in Fig. 9, and the results of the first three antennas of SCP19 data are shown in Fig. 10. The two polarizations are shown in the left-hand and right-hand subpanels. We do not show the raw data as it is mostly dominated by external sources (e.g. sky and ground pick-up). This should be very smooth in frequency with our observation and most of it should be removed with the first SVD mode. The results with one- ,two-, and five-mode subtraction are shown in different colours as labelled in the legend. The errors of the power spectrum density are estimated via the variance over different frequency channels. Significant 1/f-type noise in the power spectrum is visible in all the plots. After SVD

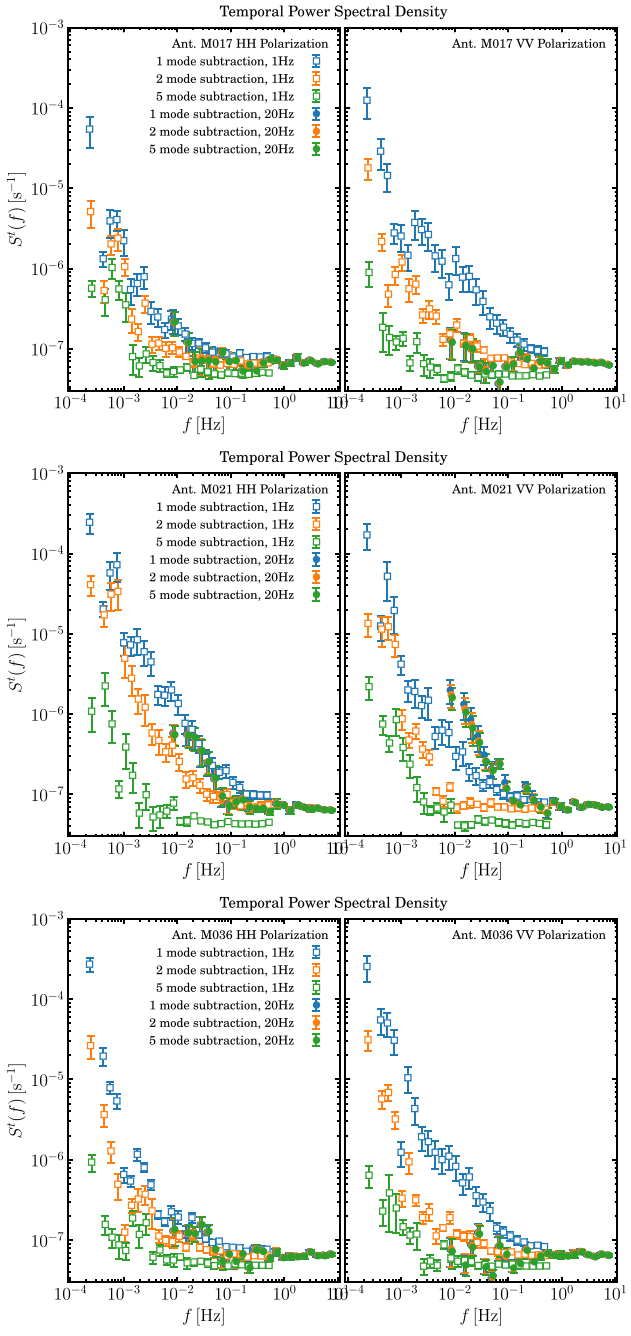

Figure 9. The temporal power spectrum density of the SCP16 data. The results of the three antennas are shown in different panels. In each panel, the results of two polarizations are shown in the left-hand and right-hand subpanels; The results with one-, two-, and five-mode subtraction are shown in different colours as labelled in the legend. The errors of the power spectrum density are estimated via the variance over different frequency channels.

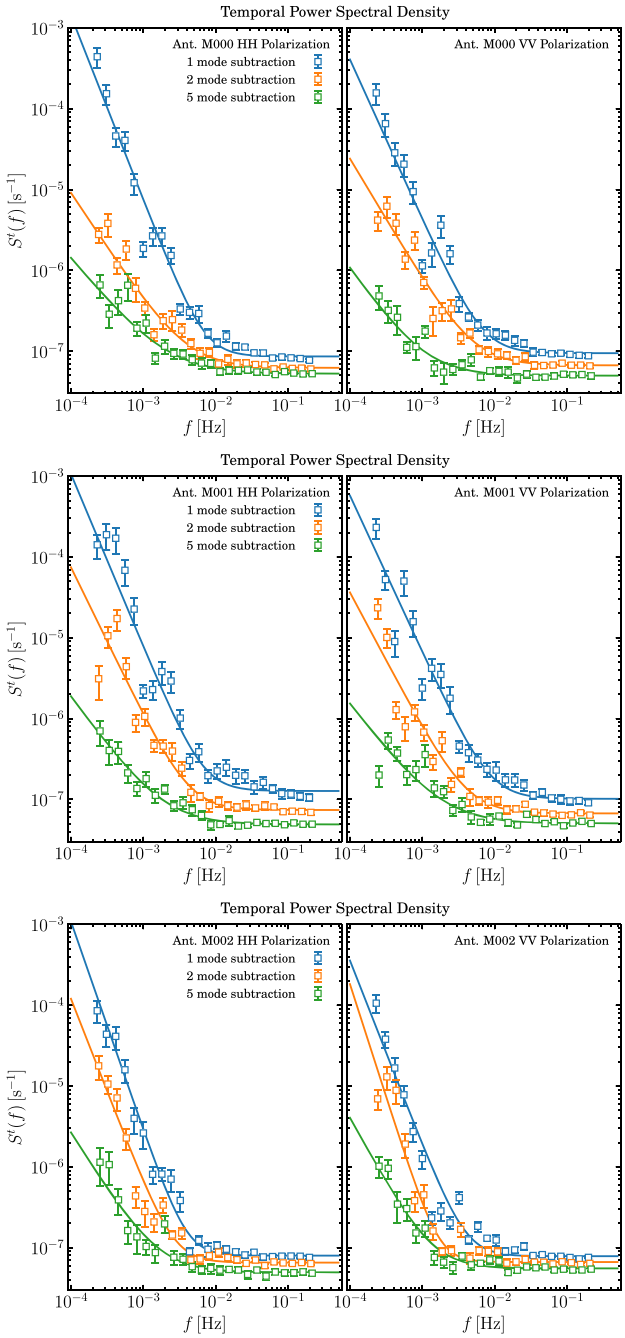

Figure 10. Same as Fig. 9 but for SCP19 data sets. The results of the first three antennas are shown in different panels. The solid lines are the fitted 1/f noise temporal model using equation (7).

modes subtraction, the 1/f-type noise power is reduced and flat white noise dominates the power spectrum over a wide frequency range, with a clear knee frequency visible. This indicates that most of the 1/f noise is correlated in frequency.

For the SCP16 data, the results with 1 Hz sampled data are shown with empty markers and the 20 Hz sampled data with filled markers. Both show good agreement of the white noise floor (Fig. 9). This level is also consistent with the theoretical value in equation (7) given by $1/\delta\nu$, which at $\delta\nu = 20$ MHz corresponds to $\sim 5 \times 10^{-8}$. The noise floor is also slightly reduced with the five modes subtraction. Note, however, that it is still equal or above the predicted theoretical white noise value. Mostly likely, this noise floor reduction is due to the removal of correlated modes in frequency but that are fluctuating on short time-scales.

We also notice that the SVD mode subtraction does not work well with antenna M021 for the 20 Hz sampled data, especially for the VV polarization. As shown in Fig. 4, the singular values of this antenna with 20 Hz sampled data are barely reduced. With 1 Hz

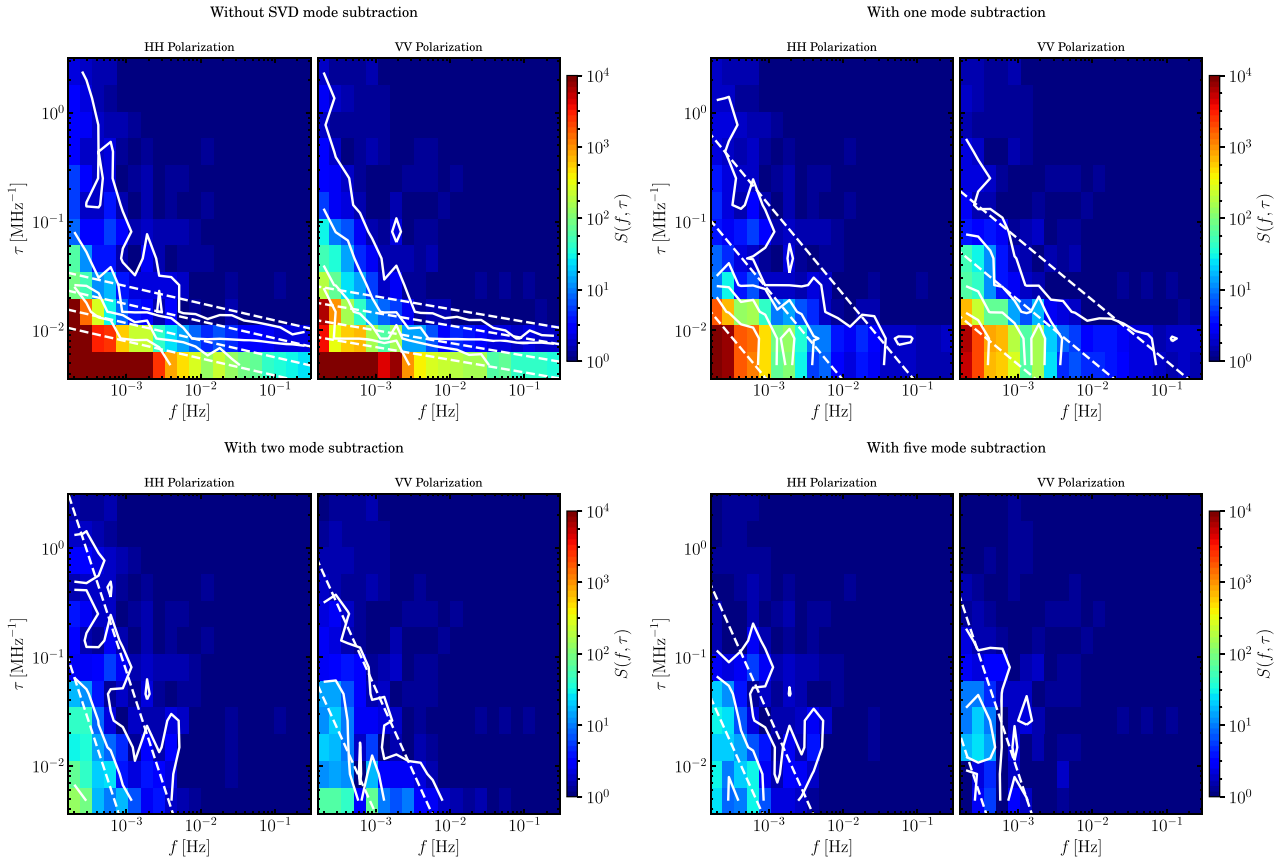

Figure 11. The 2D power spectrum density estimated with SCP19 data. From the top left to the bottom right, each panel shows the result with no SVD modes subtraction, one mode, two modes, and five modes subtraction. The two polarization are shown in the left-hand/right-hand sub-panels. All the plots are truncated with the same colour scale. The white contours shows the levels 2, 10, 10^2, and 10^3. and the dashed contours show the fitted power spectrum model at the same levels.

sampled data, M021 has larger singular values for the second and third modes compared with the other two antennas. By looking at the waterfall plot in Fig. 3, M021 has more fluctuations than the other two antennas. This indicates that the system of M021 was quite unstable during the observation of SCP16.

The SVD mode subtraction works well for the SCP19 data, as shown in Fig. 10. We only show the results of the first three antennas because the other antennas have similar behaviors. Comparing to SCP16 data sets, the f upper bond of SCP19 is limited by the lower time sampling resolution. The solid lines in Fig. 10 are the fitted 1/f noise temporal model using equation (7). Again, we do not fit to the raw data here as that would require a more complex model, possibly with a running power law. Once we remove one or more modes the fit using equation (7) works quite well. Removing two modes is quite conservative and we expect it to be done for most data analysis.

After the subtraction of the first singular mode, the 1/f type correlation in time is highly reduced and the knee frequency is reduced below 10^{-2} Hz. A clear white noise floor is shown at the high-f end and the noise floor is $\sim 5 \times 10^{-8}$, which is consistent with both the SCP16 data, as well as the model prediction. With additional singular modes subtraction, the knee frequency is further reduced. Again, the noise floor is slightly reduced with five-mode subtraction.

6.2 2D power spectrum density

The 2D power spectrum density is estimated by Fourier transforming the data along both the time and frequency axes. The results for SCP19 are shown in Fig. 11 for one antenna (M000) as an example. From top left-hand to the bottom right-hand panels, it shows the 2D power spectrum with zero, one, two, and five SVD modes subtracted, as labelled in the title of each panel. The results for the two polarizations are shown in the left-hand and right-hand subpanels, respectively. The white contour shows the levels 2, 10, 10^2, and 10^3. The dashed contours show the fitted 2D power spectrum model at the same power spectrum levels as the measurements. Since we are analysing the power spectrum of $\delta_{\rm d}$ (equation 12), the 2D power spectrum of the white noise should be at a level of 1.

The top left-hand panel of Fig. 11 shows the results of the data without singular mode subtraction. The power spectrum is peaked at the low-τ end, which indicates a strong correlation across frequency channels. Such strong frequency correlation results in a smooth frequency spectrum, which is clearly shown in the waterfall plots (top panels of Fig. 7). The power spectrum has a tail below $f \sim 10^{-2}$ extending to higher τ. Such a tail structure indicates other 1/f components, which are less correlated across frequency channels. However, the model is only fitted to the component with strong frequency correlation as it is the dominant one and the model assumes a single spectrum index. This is also the reason why the temporal power spectrum model cannot fit the data without singular mode subtraction as discussed in Section 6.1.

The strong frequency correlated component can be removed by subtracting the first singular mode. The waterfall plots of the first singular mode are shown in the top panel of Fig. 8, which have structures consistent with the raw data. The 2D power spectrum of

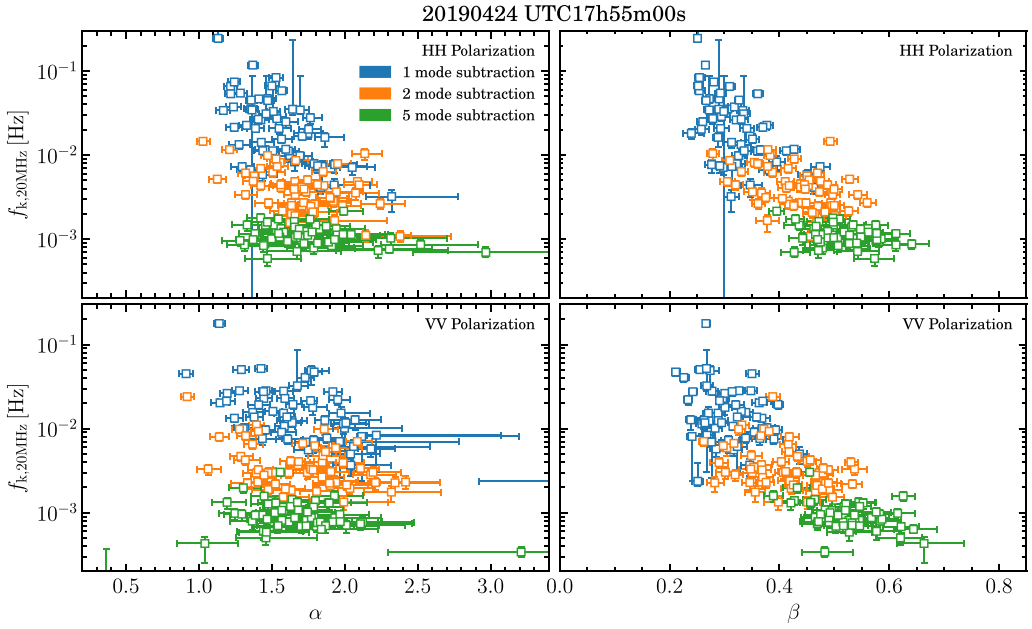

Figure 12. The fitted $f_{k,\,20MHz}$ versus α (left-hand panels) and β (right-hand panels) for each of the antennas. Top/bottom panels show the results of HH/VV polarizations. The results with 1, 2, and 5 modes subtracted are shown with different colours.

the data with first mode subtracted is shown in the top right-hand panel of Fig. 11. Comparing to the top left-hand panel, the strong frequency correlated component is subtracted out.

The first singular mode has consistent fluctuations between polarizations, as well as different dishes, which indicates that it is due to the environmental variations. The sky fluctuation is supposed to be strongly correlated across frequencies and could be the main contribution to the first mode. Its residuals could even remain in the second mode. However, the fluctuations of the first mode along time are ~ 1 per cent of the system temperature, which is higher than what we would expect from the beam asymmetries. After some investigation, we found that, during the observation, dishes were tracing the celestial point with Dec $= -90^\circ$ of epoch of $J2000$, which is about 6 arcmin off the SCP of current epoch. Such minor offset pointing could potentially cause large enough sky fluctuations during the 2.5-h observation to explain this first mode. We do not quote 1/f values without SVD subtraction given this sky contamination of the first mode and warn that there could also be some residuals in the second mode.

With more singular modes subtraction, the 1/f-type power spectrum is highly reduced and the knee frequency is also reduced to lower values. The 2D power spectrum model, equation (12), is fitted to the measurements for each antenna and polarization with a differing number of SVD modes subtracted, individually. The fitting parameters $f_{k,\,20MHz}$, α and β are shown in Fig. 12, where each square marker shows the fitting value for one antenna. The results for the two polarizations are shown in the left-hand and right-hand panels and the results with 1, 2, and 5 singular mode subtraction are shown in different colours as labelled in the legend. The median value of the fitting parameters, as well as the rms, over all dishes are listed in Table 1.

The fitting value for $f_{k,\,20MHz}$ is continuously reduced with additional singular mode subtraction. With 2 modes subtraction, the knee frequency at 20 MHz is reduced to around 3×10^{-3} Hz, which indicates that the system 1/f-type variations are well under the thermal noise fluctuation over $\sim 3 \times 10^2$ s. The time-scale can be even longer with five modes subtraction.

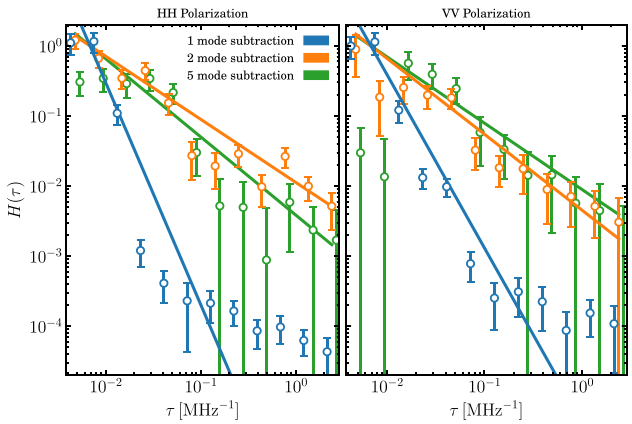

Figure 13. The spectroscopic correlation power spectrum density, $H(\tau)$, of SCP19 data. The results for two polarizations are shown in the left-hand and right-hand panels and the results with 1, 2, and 5 mode subtraction are shown with different colours. The errors are estimated with the jackknife samples. The solid lines show the corresponding best-fitting $H(\tau)$ model.

In order to further check the quality of the 2D fit, we decided to extract the spectroscopic power spectrum density, $H(\tau)$ in equation (9), from the estimated 2D power spectrum. To get a better signal to noise, we use a weighted average across the f bins:

$$H(\tau) = \frac{1}{W} \sum_{f_{\rm min}}^{f_{\rm max}} w(f) K \left(\frac{S(f,\tau) - A}{F(f)} \right) \tag{22}$$

in which, A, $f_{\rm k}$, and α use the best-fitting values. $W = \sum w(f)$ is the normalization factor and we take $w(f) = (F(f)/K)^2$ in order to down-weight the low-amplitude (noisy) contributions. The results with 1, 2, and 5 mode subtraction for one antenna (M000) are shown in Fig. 13 in different colours. The left-hand and right-hand panels show the results of two polarizations, respectively. The solid lines show the

Table 1. The median values of fitting parameters across all dishes. The errors are the rms of the fitting values across all dishes. The median values of reduced χ^2 across all dishes are listed in the last column, where d.o.f. = 236 and 21 are the degree of freedom for 2D and 1D power spectrum density, respectively.

	lg $f_{k,20\mathrm{MHz}}$		α		β		χ^2/d.o.f.	
	HH	VV	HH	VV	HH	VV	HH	VV
2D power spectrum density								
1 mode subtraction	-1.67 ± 0.39	-1.87 ± 0.34	1.51 ± 0.24	1.61 ± 0.43	0.31 ± 0.05	0.30 ± 0.05	1.453	1.151
2 mode subtraction	-2.40 ± 0.25	-2.52 ± 0.26	1.65 ± 0.52	1.74 ± 0.39	0.45 ± 0.06	0.41 ± 0.07	1.037	1.047
5 mode subtraction	-2.96 ± 0.11	-3.07 ± 0.17	1.56 ± 0.64	1.47 ± 0.73	0.51 ± 0.05	0.52 ± 0.08	2.660	1.996
1D power spectrum density								
1 mode subtraction	-2.24 ± 0.19	-2.22 ± 0.18	2.28 ± 0.34	2.24 ± 0.41	–	–	3.930	3.954
2 mode subtraction	-2.40 ± 0.21	-2.40 ± 0.21	1.83 ± 0.40	1.98 ± 0.38	–	–	2.123	2.214
5 mode subtraction	-2.78 ± 0.14	-2.85 ± 0.21	1.28 ± 0.26	1.47 ± 0.37	–	–	1.259	1.218

corresponding best-fitting $H(\tau)$ model. The errors are estimated with the jackknife samples. As discussed, we assume a single power-law power spectrum density for the spectroscopic correlations, which does not seem enough with 1 mode subtraction. With 2 mode subtraction, however, the model fits the measurements quite well. The mean reduced χ^2 of the 2D power spectrum fitting, averaged over the measurements of all antennas, are (1.037, 1.047) for (HH, VV) polarizations. The power spectrum density starts deviating from a power law at the low-τ end with five-mode subtraction, as the 1/f spectroscopic correlation is oversubtracted.

6.3 Impact on H I signal

The temporal and spectroscopic wavenumber, f and τ, are related to the cosmological scales k by

$$\tau = \frac{\nu_0}{\nu_{\rm obs}^2}\frac{c}{H(z)}\frac{k_\parallel}{2\pi}, \quad f = \frac{k_\perp \chi(z) u}{2\pi}, \quad \text{and } k^2 = k_\parallel^2 + k_\perp^2, \tag{23}$$

in which, k is in units of $\mathrm{Mpc}^{-1}h$; $\nu_0 = 1420\mathrm{MHz}$ is the rest-frame H I emission line frequency, c is the speed of light, $\nu_{\rm obs}$ is the observing frequency, and u is the scanning speed. We assume the fiducial cosmology parameters from Planck Collaboration VI (2018; $h = 0.6736$, $\Omega_{\rm m} = 0.3153$ and $\Omega_\Lambda = 0.6847$).

Assuming an observation at 900 MHz with scanning speed of 5 arcmin s^{-1}, the corresponding cosmological scales projecting to the $f - \tau$ space are shown with the white-dotted lines in Fig. 14. We see that there is a large region in the $f - \tau$ space that will be available for the H I measurements. In particular, even without any mode subtraction, most contamination, either because of 1/f noise or foregrounds is constrained to a region of low f or low τ. The available scales are enough to probe Baryon Acoustic Oscillations and even primordial non-Gaussianity at $k \sim 0.01\mathrm{Mpc}^{-1}h$ (Fonseca, Maartens & Santos 2017). In order to probe smaller $k_\perp$ modes we will need to increase the survey speed or rely on extra calibration such as with noise diodes.

We have shown that through SVD mode subtraction we are able to push the 1/f noise contamination to large time-scales. One possible concern is if such cleaning can also remove the H I signal. In order to investigate the signal loss fraction of the SVD mode subtraction, we build a transfer function:

$$T(f,\tau) = \frac{S^{\rm c}(f,\tau)}{S(f,\tau)}, \tag{24}$$

where $S(f, \tau)$ is the 2D power spectrum estimated using the simulated time-ordered data, $\mathbf{D}_{\rm sim}$, which includes only the H I signal and $S^{\rm c}(f, \tau)$ is the power spectrum of the same time-ordered simulation with SVD mode subtraction:

$$\mathbf{D}^{\rm c}_{\rm sim} = \mathbf{D}_{\rm sim} - \sum_{i=0}^{N_{\rm m}} u_i (u_i^\dagger \mathbf{D}_{\rm sim} v_i) v_i^\dagger. \tag{25}$$

To generate the time-ordered simulation, we simply assume the observation to be a drift-scan targeting at the celestial equator. This would be more typical of an actual 21-cm survey (observing the SCP would give a signal constant in time). We use the temporal

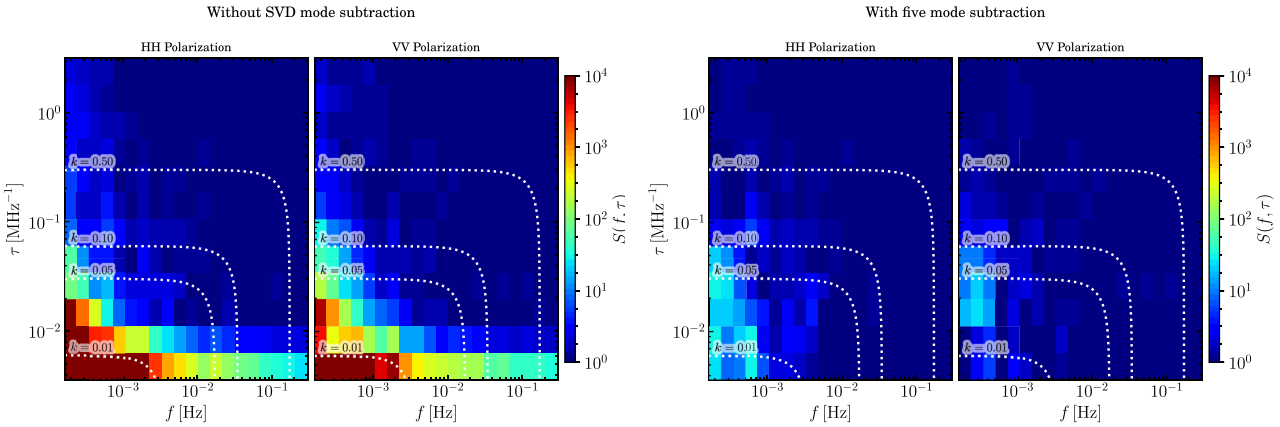

Figure 14. Same as Fig. 11 but only for the results with no SVD mode subtraction (left) and five-mode subtraction (right). The white-dotted lines indicate the corresponding cosmological scales projecting to the $f - \tau$ space, assuming observations at 900 MHz with scanning speed of 5 arcmin s^{-1}. k is in units of $\mathrm{Mpc}^{-1}h$.

and spectroscopic modes, u_i and v_i, that are obtained from the SVD decomposition of the combined simulated and real time-ordered data, $\mathbf{D}_{\rm sim} + \mathbf{D}_{\rm real}$. This way we can test the expected effect on the H I signal of such decomposition.

We found that the projection of the simulated time-ordered data to the temporal and spectroscopic mode is close to 0:

$$u_i^{\dagger} \mathbf{D}_{\rm sim} v_i \sim 0. \quad (26)$$

Therefore, the signal loss is negligible [e.g. $T(f, \tau) \sim 1$]. With five SVD modes removed, the signal loss is less than 3 per cent and only at large scales where $f \sim f_{\rm min}$ and $\tau \sim \tau_{\rm min}$. This can be understood because the cleaning is done in the time-ordered data while the H I signal 'lives' in map space. Although the projection of the cosmological scale k to the f-τ space depends on the scanning speed and frequency range, given the negligible signal loss in the full f-τ space, the time-ordered cleaning will have a negligible impact on the H I signal with most scanning strategies.

7 SUMMARY AND CONCLUSIONS

In this work, we measured the power spectrum density of the 1/f noise for the MeerKAT receiver system. The analysis is performed with SCP tracking data to avoid sky variations. Two SCP tracking data sets are used in this analysis. We find a relatively RFI-free frequency range from 1313.6758 to 1461.8457 MHz. Absolute flux calibration is ignored in our analysis as the data are normalized with the time-averaged system temperature for each frequency channel.

We apply SVD to the data and determine how effective removing the first several principal components is on suppressing the time–frequency correlated noise. The results show that indeed, the 1/f noise can be drastically reduced by removing the first few SVD modes. Moreover, the correlation features are shown to be well described by the proposed noise model with just a few parameters once the first two SVD modes are removed. With two mode subtraction, the data averaged over 20 MHz has a knee frequency ($f_{\rm k}$) of $\sim 3 \times 10^{-3}$ Hz, indicating that the system 1/f-type variation is well under the thermal noise fluctuation over a few hundred seconds time-scales. This increases further with five modes removed. The results from this analysis, along with the described noise power spectrum model, can be used in realistic noise simulations for MeerKAT and extended to SKA1-Mid.

The 2D power spectrum shows that the 1/f noise is constrained to a small region of either low τ or low f, e.g. large-scale correlations in time or in frequency. This provides many scales where the H I signal can be probed without contamination. With scanning speeds of 5 arcmin s^{-1}, 10^3 s time-scales would correspond to $\sim 80^{\circ}$, which is enough for our cosmological purposes. Longer time-scales can be achieved using noise diodes. Our calibration plan is to use celestial sources for calibration on time-scales $\sim$ 1.5 hour (absolute flux and bandpass calibration) and noise diodes for shorter time-scales. We can then apply a conservative SVD cleaning on the time-ordered data (e.g. two modes) in order to remove most of the 1/f noise contamination. These modes correspond to large-frequency scales where we expect the 21-cm correlations to be negligible.

We further tested the effect of the SVD subtraction on the 21 cm itself. According to the transfer function of the signal loss fraction, the proposed cleaning has negligible impact on the H I signal. Any further residual noise can be included in the map making process which will allow for correlated noise in frequency.

The 1/f noise has been a substantial challenge to precision cosmology with single dish telescopes in the past and, if it is not carefully treated, has been shown to be detrimental to future H I IM experiments. We have demonstrated here a methodology that can be used to effectively suppress 1/f noise in single dish H I IM observations that should preserve the cosmological signal. In future work, we plan to demonstrate the effectiveness of this technique to extract the 21-cm signal both on simulated and real MeerKAT data.

ACKNOWLEDGEMENTS

We are grateful to Phil Bull, Clive Dickinson, and Jonathan Sievers for very useful discussions. YL and MGS acknowledge support from the South Africa National Research Foundation (NRF) through the South African Research Chairs Initiative (SARChI) Grant No. 84156 and the South African Radio Astronomy Observatory (SARAO) group grant. The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. We acknowledge the use of the Inter-University Institute for Data Intensive Astronomy (IDIA) and Ilifu computing facilities.

DATA AVAILABILITY

Available on request.

REFERENCES

Alonso D., Ferreira P. G., Santos M. G., 2014, MNRAS, 444, 3183
Anderson L. et al., 2014, MNRAS, 441, 24
Anderson C. J. et al., 2018, MNRAS, 476, 3382
Ansari R. et al., 2012, A&A, 540, A129
Asad K. M. B. et al., 2019, preprint (arXiv:1904.07155)
Bagla J. S., Khandai N., Datta K. K., 2010, MNRAS, 407, 567
Bandura K. et al., 2014, Proc. SPIE Conf. Ser. Vol. 9145, Ground-based and Airborne Telescopes V, SPIE, Bellingham, WA, p. 914522
Battye R. A., Davies R. D., Weller J., 2004, MNRAS, 355, 1339
Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, MNRAS, 434, 1239
Bigot-Sazy M. A. et al., 2015, MNRAS, 454, 3240
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
Camilo F. et al., 2018, ApJ, 856, 180
Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, Phys. Rev. Lett., 100, 091303
Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, Nature, 466, 463
Chen X., 2012, 2nd Galileo-Xu meeting, International Journal of Modern Physics: Conference Series. p. 256
Cole S. et al., 2005, MNRAS, 362, 505
Efron B., 1982, The Jackknife, the Bootstrap and Other Resampling Plans, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia
Eisenstein D. J. et al., 2005, ApJ, 633, 560
Fonseca J., Maartens R., Santos M. G., 2017, MNRAS, 466, 2780
Foreman-Mackey D. et al., 2019, J. Open Source Softw., 4, 1864
Harper S. E., Dickinson C., Battye R. A., Roychowdhury S., Browne I. W. A., Ma Y. Z., Olivari L. C., Chen T., 2018, MNRAS, 478, 2416
Hinton S. R. et al., 2017, MNRAS, 464, 4807
Janssen M. A. et al., 1996, preprint (astro-ph/9602009)
Jonas J., MeerKAT Team, 2016, MeerKAT Science: On the Pathway to the SKA, Proceedings of MeerKAT Science. p. 1
Keihänen E., Kurki-Suonio H., Poutanen T., Maino D., Burigana C., 2004, A&A, 428, 287
Kurki-Suonio H., Keihänen E., Keskitalo R., Poutanen T., Sirviö A. S., Maino D., Burigana C., 2009, A&A, 506, 1511
Lidz A., Furlanetto S. R., Oh S. P., Aguirre J., Chang T.-C., Doré O., Pritchard J. R., 2011, ApJ, 741, 70
Loeb A., Wyithe J. S. B., 2008, Phys. Rev. Lett., 100, 161301

Maino D., Burigana C., Górski K. M., Mand olesi N., Bersanelli M., 2002, A&A, 387, 356
Mao Y., Tegmark M., McQuinn M., Zaldarriaga M., Zahn O., 2008, Phys. Rev. D, 78, 023529
Mao X.-C., 2012, ApJ, 744, 29
Masui K. W. et al., 2013, ApJ, 763, L20
Mauch T. et al., 2020, ApJ, 888, 61
McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815
Newburgh L. B. et al., 2016, HIRAX: A Probe of Dark Energy and Radio Transients, SPIE, Bellingham. p. 99065X
Peterson J. B. et al., 2009, Astro2010: The Astronomy and Astrophysics Decadal Survey, Science White Papers, no. 234
Planck Collaboration VI, 2018, A&A , 641, A6
Pritchard J. R., Loeb A., 2008, Phys. Rev. D, 78, 103511
Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901
Santos M. et al., 2015, Proc. Sci.,. Cosmology from a SKA HI intensity mapping survey, SISSA, Trieste, PoS(AASKA14)019
Santos M. G. et al., 2017, preprint (arXiv:1709.06099)
Seiffert M., Mennella A., Burigana C., Mand olesi N., Bersanelli M., Meinhold P., Lubin P., 2002, A&A, 391, 1185
Seo H.-J., Dodelson S., Marriner J., Mcginnis D., Stebbins A., Stoughton C., Vallinotto A., 2010, ApJ, 721, 164
Square Kilometre Array Cosmology Science Working Group, 2020, Publ. Astron. Soc. Aust., 37, e007
Sutton D. et al., 2010, MNRAS, 407, 1387
Switzer E. R. et al., 2013, MNRAS, 434, L46
Switzer E. R., Chang T. C., Masui K. W., Pen U. L., Voytek T. C., 2015, ApJ, 815, 51
Wilson T. L., Rohlfs K., Hüttemeister S., 2009, Tools of Radio Astronomy, Springer-Verlag, Berlin, Germany
Wolz L. et al., 2015, Proc. Sci., Foreground Subtraction in Intensity Mapping with the SKA, SISSA, Trieste, PoS(AASKA14)035
Wolz L. et al., 2017, MNRAS, 464, 4938
Wyithe J. S. B., Loeb A., 2008, MNRAS, 383, 606
Wyithe J. S. B., Loeb A., Geil P. M., 2008, MNRAS, 383, 1195

APPENDIX A:

A1 Relation between temporal power spectrum density and 2D power spectrum density

The temporal power spectrum density can be expressed as

$$S_{ij}(f) = \langle \delta(f, \nu_i)\delta^{\dagger}(f, \nu_j)\rangle = \int \mathrm{d}\,\nu\,\mathrm{d}\,\nu'\,\phi_i(\nu)\phi_j^{\dagger}(\nu')\xi(f, \nu_{ij}) \tag{A1}$$

in which, $\nu_{ij} = \nu_i - \nu_j$, $\phi_i(\nu)$ is the spectroscopic window function and $\delta(f, \nu_i) = \int \mathrm{d}\,\nu\,\phi_i(\nu)\delta(f, \nu)$; $\xi(f, \nu_{ij})$ is related to the 2D power spectrum density, $S(f, \tau)$, via inverse Fourier transform,

$$\xi(f, \nu_{ij}) = \int \mathrm{d}\,\tau S(f, \tau) \exp\left[2\pi i \tau \nu_{ij}\right]. \tag{A2}$$

Substituting equation (A2) to equation (A1), the temporal power spectrum density can be further expressed as

$$S_{ij}(f) = \int \mathrm{d}\,\tau\,\phi_i(\tau)\phi_j^{\dagger}(\tau)S(f, \tau), \tag{A3}$$

where $\phi_i(\tau) = \int \mathrm{d}\,\nu\,\phi_i(\nu)\exp[-2\pi i\tau\nu]$ is the Fourier transform of the spectroscopic window function. If we ignore the cross-correlation between frequencies, the diagonal term of the temporal power spectrum density is

$$S_i(f) = \int \mathrm{d}\,\tau\,\phi_i^2(\tau)S(f, \tau). \tag{A4}$$

A2 The white noise level

If we use a top-hat window function with width of $\delta\nu$, the Fourier transform of the top-hat window function, $\phi_i(\tau)$, can be expressed with a sinc function:

$$\phi_i(\tau) = \mathrm{sinc}(\pi\delta\nu\tau), \tag{A5}$$

where the window function is normalized with $\int \mathrm{d}\,\tau\phi_i^2(\tau) = \int \mathrm{d}\,\tau\phi_i(\tau) = 1/\delta\nu$. Substituting the 2D white noise power spectrum model (the first term of equations 12), (A4) becomes

$$A\int \mathrm{d}\,\tau\,\phi_i^2(\tau) = \frac{A}{\delta\nu}, \tag{A6}$$

which is consistent with the white noise term of 1D power spectrum density model equation (7).

A3 The knee frequency conversion between frequency resolutions

We first model the 2D 1/f noise power spectrum density with f_0 at arbitrary frequency resolution:

$$S(f, \tau) = A\left(\frac{f_0}{f}\right)^{\alpha}\left(\frac{\tau_0}{\tau}\right)^{\frac{1-\beta}{\beta}}. \tag{A7}$$

Substituting the noise model into equation (A4),

$$A\left(\frac{f_0}{f}\right)^{\alpha}\int \mathrm{d}\,\tau\phi_i^2(\tau)\left(\frac{\tau_0}{\tau}\right)^{\frac{1-\beta}{\beta}} = AK\left(\frac{f_0}{f}\right)^{\alpha}, \tag{A8}$$

where K is

$$K = \int \mathrm{d}\,\tau\,\mathrm{sinc}^2(\pi\delta\nu\tau)\left(\frac{\tau_0}{\tau}\right)^{\frac{1-\beta}{\beta}}. \tag{A9}$$

The 1D power spectrum model assumes the 1/f noise is reduced with factor of $\delta\nu$, which indicates that the 1/f noise is uncorrelated across frequencies. That equivalent to set $\beta = 1$ in the 2D power spectrum density model:

$$\begin{aligned} AK\left(\frac{f_0}{f}\right)^{\alpha} &= A\left(\frac{f_0}{f}\right)^{\alpha}\int \mathrm{d}\,\tau\,\mathrm{sinc}^2(\pi\delta\nu\tau)\left(\frac{\tau_0}{\tau}\right)^{0} \\ &= \frac{A}{\delta\nu}\left(\frac{f_0}{f}\right)^{\alpha}. \end{aligned} \tag{A10}$$

Comparing with the second term of 1D power spectrum density model equation (7), we have $f_0 = f_k$. This indicates that, if the 1/f noise is fully uncorrelated across frequencies, the knee frequency is constant with any frequency resolution.

On the other hand, if $\beta \to 0$, the spectroscopic spectrum density index is approaching to $+\infty$ and the spectroscopic power spectrum density model becomes a Dirac delta function, $H(\tau) \to \delta^{\mathrm{D}}(\tau)$. In this case, we have

$$\begin{aligned} AK\left(\frac{f_0}{f}\right)^{\alpha} &= A\left(\frac{f_0}{f}\right)^{\alpha}\int \mathrm{d}\,\tau\,\mathrm{sinc}^2(\pi\delta\nu\tau)\delta^{\mathrm{D}}(\tau) \\ &= \frac{A}{\Delta\nu}\left(\frac{f_0}{f}\right)^{\alpha}, \end{aligned} \tag{A11}$$

where $\Delta\nu = 1/\tau$. When $\tau = 0$, $\Delta\nu \to +\infty$, indicating the infinity frequency bandwidth.

However, the measurements are always limited within finite frequency bandwidth. If we write the integrals in terms of discrete sums, equation (A4) is expressed as

$$S_i(f) = \sum_{p=0}^{N_\nu - 1}\phi_i^2(p\tau_0)S(f, p\tau_0)\tau_0, \tag{A12}$$

where τ_0 is the minimal spectroscopic frequency interval and related to the minimal frequency interval ν_0 via $\tau_0 = 1/(N_\nu \nu_0) = 1/\Delta\nu$, where $\Delta\nu$ is the full frequency bandwidth. The equation (A11) becomes

$$A\left(\sum_{p=0}^{N_\nu-1}\mathrm{sinc}^2(\pi\delta\nu p\tau_0)\tau_0\delta_p^D\right)\times\left(\frac{f_0}{f}\right)^\alpha = \frac{A}{\Delta\nu}\left(\frac{f_0}{f}\right)^\alpha. \quad \text{(A13)}$$

Comparing with the 1D power spectrum density model, we have

$$\frac{1}{\Delta\nu}\left(\frac{f_0}{f}\right)^\alpha = \frac{1}{\delta\nu}\left(\frac{f_k}{f}\right)^\alpha, \quad \text{(A14)}$$

$$\lg f_0 = \lg f_\mathrm{k} - \frac{1}{\alpha}\lg\frac{\delta\nu}{\Delta\nu}. \quad \text{(A15)}$$

When $\delta\nu = \Delta\nu$, we have $f_0 = f_\mathrm{k}$, which indicates that f_0 is the knee frequency at full frequency bandwidth, which is corresponding to the minimal spectroscopic frequency interval τ_0.

In the case of $0 < \beta < 1$, the shift of f_0 with frequency resolution is dependent on the frequency correlation properties. The relation of f_k between different frequency resolutions, $\delta\nu$ and $\delta\nu'$, is expressed as

$$\lg f_{\mathrm{k},\delta\nu} = \lg f_{\mathrm{k},\delta\nu'} + \frac{1}{\alpha}\lg\left(\frac{K\delta\nu}{K'\delta\nu'}\right). \quad \text{(A16)}$$

Replacing f_0 with f_k, we have equation (12):

$$S(f,\tau) = A\left(1+\frac{1}{K\delta\nu}\left(\frac{f_k}{f}\right)^\alpha\left(\frac{\tau_0}{\tau}\right)^{\frac{1-\beta}{\beta}}\right). \quad \text{(A17)}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.