Chapter 6

Effect of Frequency Correlated Bandpass Errors

In this chapter, we study the effect of bandpass errors and their impact on power spectrum estimation of redshifted 21- cm signal from the epoch of reionization. Similar to the method followed in the previous chapters, we derive the analytical expressions of bias and variance of the multifrequency angular power spectrum (MAPS) (Datta et al., 2007) in the presence of frequency correlated gain errors, along with the time-correlation of the gains, bright foreground, and thermal noise. This completes the analytical framework for estimating bias and variance of the power spectrum for time and frequency-correlated gains.

6.1 Time and Frequency-Dependent Gains

The measured visibility $\tilde{V}(\vec{U}_i)$, at an observation frequency ν, for the i^{th} baseline $\mathbf{U}_i$ at a time t can be given as (Hamaker et al., 1996; Taylor et al., 1999)

$$\tilde{V}(\vec{U}_i,\nu) = \tilde{\mathbb{G}}_i(t,\nu)\tilde{V}_i^S(\vec{U}_i,\nu) + \tilde{N}_i(\vec{U}_i) \tag{6.1}$$

where $\tilde{\mathbb{G}}_i(t,\nu)$ is the time and frequency dependent baseline-based complex gain for the i^{th} baseline; $\tilde{N}_i(\vec{U}_i)$ is the thermal noise and $\tilde{V}_i^S$ is the sky visibility (see Chapter 3). In most practical cases we can write $\tilde{\mathbb{G}}_i(t,\nu)$ as

$$\tilde{\mathbb{G}}_i(t,\nu) = \mathrm{G}_i(t)\tilde{\mathfrak{B}}_i(\nu). \tag{6.2}$$

Here $\mathrm{G}_i(t)$ represents the gain, and $\tilde{\mathfrak{B}}_i(\nu)$ is the bandpass, for the i^{th} baseline.

We can write both the time and frequency-dependent part of the gain $\mathrm{G}_i(t)$ and $\tilde{\mathfrak{B}}_i(\nu)$, as arising from the individual gains of the pair of antennae A and B, in terms of antenna-based gains as

$$\tilde{\mathrm{G}}_i(t) = \langle \tilde{g}_A(t)\tilde{g}_B^*(t)\rangle \; ; \quad \tilde{\mathfrak{B}}_i(\nu) = \langle \tilde{b}_A(\nu)\tilde{b}_B^*(\nu)\rangle, \tag{6.3}$$

where $\tilde{g}_A(t)$, $\tilde{g}_B(t)$ and $\tilde{b}_A(\nu)$, $\tilde{b}_B(\nu)$ are the gain and bandpass response of the individual antenna A and B; angle brackets represent the average over the integration time.

6.1.1 Bandpass Errors

In Chapter 3, we discussed the ionospheric effects and the instrumental response to the electric field signal coming from the sky; these effects are collectively termed as the gain. The gains depend both on time and frequency of observation. The origin of time and frequency dependence of the gains and various calibration strategies to solve the antenna-based gains are also discussed.

The frequency-dependent part of the gain $\mathfrak{B}_i(\nu)$ is solved using bandpass calibration. We observe a bright flat-spectrum source to correct the frequency variation of amplitude and phase with the frequency. Bandpass solutions are obtained for each channel or averaged over a few nearby channels depending on the signal-to-noise of observation. Along with the frequency dependence of the gains, bandpass calibration also helps correct the slow variation of bandpass response with time. Good bandpass calibration is key to detecting

and accurately measuring spectral features. Bandpass calibration can also be the limiting factor in a dynamic range of continuum observations. There are various limiting factors for accurate bandpass calibration, such as the unavailability of a good bandpass calibrator source in the sky during the observation, errors in the complex response of bandpass filters, etc. These introduce residual errors in the bandpass solutions. The presence of RFI or bad instrumental response in a few observation channels may force to the removal of the whole data for those channels. In cases where observed data for a few channels are unavailable, the bandpass solutions for those frequencies are obtained by interpolation from nearby channels. Alternatively, the bandpass solutions are obtained using a polynomial model. This introduces the frequency correlation in the residual bandpass errors.

6.1.2 Modelling the Frequency Correlated Gain Grrors

Before using the data for any scientific purpose, calibration is performed. We assume that the primary and self-calibration have been done for time-dependent gain, and bandpass calibration has also been done to correct the frequency response. After the calibration, there remain residual gain errors. At this stage of data processing, only residual gains and residual bandpass for each antenna contribute to the gain term $\tilde{G}_i(t)$ and $\tilde{\mathfrak{B}}_i(\nu)$ respectively. For the time-dependent gains, we use the same gain error model as discussed in Chapter 5. Just for revision, we write it here again. The gain from an individual antenna A is given as

$$\tilde{g}_A(t) = [1 + \delta_{AR}(t) + i\delta_{AI}(t)] \tag{6.4}$$

where $\delta_{AR}(t)$ and $\delta_{AI}(t)$ stands for the real and imaginary part of the residual gain error. The variance σ^2_{bC} and the normalized two-point correlation $\xi_{bC}(\tau)$ of the time-dependent residual gain from antenna A as

$$\sigma^2_{AC} = \langle \delta^2_{AC} \rangle, \quad \xi_{AC}(\tau) = \langle \delta_{AC}(t)\delta_{AC}(t+\tau) \rangle / \sigma^2_{AC}, \tag{6.5}$$

where C is to be read as R or I for real and imaginary parts of the gain. Here, for simplicity, we assume that the statistical properties of the gains for all the antennae in the array are the same. The normalized two-point correlation $\xi(\tau)$ of the residual gains is a function of (time) delay τ and its value is unity at zero delay. We model it as

$$\xi(\tau) = \exp\left[-\frac{\tau^2}{2T_{corr}^2}\right] \tag{6.6}$$

where T_{corr} gives the correlation time of the residual gains.

We model the frequency-dependent bandpass response for the individual antennae as

$$\tilde{b}_A(\nu) = [1 + b_{AR}(\nu) + ib_{AI}(\nu)] \tag{6.7}$$

where $b_{AR}(t)$ and $b_{AI}(t)$ stands for the real and imaginary part of the residual bandpass error.

Similar to the time-correlated gain errors, we assume the bandpass errors to be Gaussian random and quantify them with their variance and two-point correlation functions. We use the assumption that the statistical properties of the bandpass for all the antennas in the array are the same, and the real and imaginary parts of the bandpass errors from a given antenna and residual bandpass errors from the different antennae are not correlated. We denote the variance of the bandpass σ_{bC}^2 and the normalized two-point correlation $\xi_{bC}(\tau)$ of the residual bandpass for antenna A as

$$\sigma_{bC}^2 = \langle b_{AC}^2 \rangle, \quad \xi_{bC}(\Delta\nu) = \langle b_{AC}(\nu) b_{AC}(\nu + \Delta\nu) \rangle / \sigma_{bC}^2, \tag{6.8}$$

where subscript b is for the bandpass. The normalized two-point correlation of the residual bandpass is a function of frequency separation $\Delta\nu$. Using the assumption that the normalized two-point correlation functions for all antennae are the same irrespective of

being of real or imaginary parts, we denote it for residual bandpass error by $\xi(\Delta\nu)$ and model it using a Gaussian function.

$$\xi(\Delta\nu) = \exp\left[-\frac{\Delta\nu^2}{2\nu_c^2}\right] \tag{6.9}$$

where ν_c is the correlation frequency of the residual bandpass. In reality, the normalized-two-point correlation function can be rather complicated and needs to be estimated for the particular telescope and observation settings and then to be used to calculate the effect of residual bandpass for the given observation.

It has been shown in various works such as Barry et al. (2016); Ewall-Wice et al. (2017); Trott et al. (2016) that for spectrally smooth bandpass solutions, the power spectral bias from residual gain errors will not be significant. However, a departure from the spectrally smooth behavior of the frequency-dependent gains will result in complex residual gain errors in frequency, and contaminate the otherwise free EoR window. Although the existing experiments enforce the spectral smoothness of the frequency-dependent gains, here we do not strictly make any such assumption for the development of the framework to estimate the residual effects. The presented framework uses a simple Gaussian model for the frequency-dependent gain errors, but it is not limited to this model and can be applied to any model whether it assumes spectral smoothness of the gains or not.

6.2 Analytical Estimates of Bias and Variance

This section finds an analytical estimate for the bias and variance in the power spectrum estimator for a time and frequency-dependent gain error model. Along with the time-correlation of the gains discussed in Chapter 4 and 5, we incorporate the effect of frequency correlation of the gains and, using a model for the same, provide an analytical estimate for both bias and variance of the angular power spectrum in the presence of both time

and frequency-dependent residual gain errors and noise for the angular power spectrum estimator. Further, we use these analytical expressions to study various effects in the power spectrum estimation methods using visibility correlations. We use the visibility correlation-based power spectrum estimator discussed in Chapter 2 to estimate the power spectrum. Here we consider a multichannel observation and correlate nearby visibilities from two different channels to estimate the angular power spectrum of the sky brightness distribution. The multifrequency angular power spectrum (MAPS) (Datta et al., 2007) quantifies the statistics of the sky brightness distribution as a function of angular multipole (ℓ) and frequency (ν). This assumes that the sky brightness temperature fluctuations are Gaussian random in nature, statistically homogeneous, and isotropic on the sky. Under the assumption that the signal is ergodic, the visibility correlations at nearby baseline give estimates of angular power spectrum as $C_{2\pi U}(\Delta\nu) \sim \langle \tilde{V}(\mathbf{U},\nu)\tilde{V}(\mathbf{U}',\nu+\Delta\nu)\rangle$. The angular multipole ℓ and baseline U are related as $\ell = 2\pi U$. The 3D-Power Spectrum $P(k)$ fails to quantify the evolving statistical properties of the signal, while MAPS fully quantifies the second-order statistics and readily determines the reionization history of the universe (Mondal et al., 2018, 2019, 2020). This motivates us to choose MAPS over 3D-Power Spectrum.

6.2.1 Type of Baseline Pair Fractions

In Chapter 4 and 5, we consider the observations centered at a single channel of frequency and discuss how we can get a pair of nearby baselines in a given uv-grid where the visibilities are measured and correlated. Every visibility is measured by correlating the electric fields from a pair of antennas. Radio interferometers simultaneously measure visibilities at many baselines and frequency channels. The visibility V($\mathbf{U}$, ν) is a function of the baseline, which changes with time. It is recorded for every pair of antennae at different frequency channels in the frequency range of observation. Hence, for a given

antenna pairs A and B, at a time t, at a frequency of observation ν, the visibility is measured at a baseline $\vec{U}_{AB}(t)$; we write this visibility as $\tilde{V}_{AB}(t,\nu)$. We only correlate the visibilities at two different frequency channels to estimate the angular power spectrum. Correlating the visibilities from the same frequency channel shows a sudden jump in the power (Pal et al., 2021). We do not consider the correlation of the visibilities measured at the same frequency channel. The availability of different frequency channels increases the total number of visibility correlations by many folds. For nearby baseline correlation, we can get a pair of baselines in the following ways:

- **Type 1** Correlation of the visibilities measured by the same antenna pair, at different times, and different frequencies, i.e. $\langle \tilde{V}_{AB}(t,\nu)\tilde{V}^*_{AB}(t',\nu')\rangle$. The fraction of such baseline pairs is n'_1.
- **Type 1A** Correlation of the visibilities measured by the same antenna pair, at the same time, and different frequency, i.e. $\langle \tilde{V}_{AB}(t,\nu)\tilde{V}^*_{AB}(t,\nu')\rangle$. The fraction of such baseline pairs is n'_{1A}.
- **Type 2** Correlation of the visibilities measured by antenna pairs having one antenna in common, measured at the same time, and different frequency, i.e. $\langle \tilde{V}_{AB}(t,\nu)\tilde{V}^*_{AC}(t,\nu')\rangle$. The fraction of such baseline pairs is n'_2.
- **Type 3** Correlation of the visibilities measured by antenna pairs having one antenna in common measured at different times, and different frequency $\langle \tilde{V}_{AB}(t,\nu)\tilde{V}^*_{AC}(t',\nu')\rangle$. The fraction of such baseline pairs is n'_3.
- **Type 4** Correlation of the visibilities measured by antenna pairs having no antenna in common measured at any time. $\langle \tilde{V}_{AB}(t,\nu)\tilde{V}^*_{CD}(t',\nu')\rangle$. The fraction of such baseline pairs is n'_4.

The noise is uncorrelated between any two measurements considered for the above cases. In addition to the baseline pair fractions discussed in Chapter 4 and 5, the presence of

different frequency channels introduces a new type of baseline pair. The visibility pairs measured by the same antenna pair, at the same time, due to being at different frequencies of observation, will also provide a visibility correlation free from noise bias; we defined this by typing n'_{1A}. The other types of visibility correlations are the same as those defined previously, except that now two visibilities are from different channels. For any uv-grid in the baseline plane in which nearby baseline correlations are performed to estimate the power spectrum, all types of baseline pairs can contribute to the total estimate of the power spectrum. Here we assume that visibility measurements for all the channels are falling in the grids, and all the channels in the grid have a similar number of baselines. Under these assumptions, the new baseline pair fractions are related to the older ones, as defined in Chapter 5. Let N_b be the number of baselines in one channel, and N_c be the total number of channels. The new baseline pair fractions are given as

$$\begin{aligned} n'_1 &= n_1 \frac{N_b - 1}{N_b} \sim n_1 \\ n'_{1A} &= \frac{1}{N_b} \\ n'_2 &= n_2 \frac{N_b - 1}{N_b} \sim n_2 \\ n'_3 &= n_3 \frac{N_b - 1}{N_b} \sim n_3 \\ n'_4 &= n_4 \frac{N_b - 1}{N_b} \sim n_4 \end{aligned} \tag{6.10}$$

where n_is are the older baseline pair fractions defined in Chapter 5. Note that the baseline pair fractions n'_i are estimated in a uv-grid; these depend on the baseline of the uv-grid $\vec{U}_g$ through the antenna configuration of the telescope and source position for observation.

6.2.2 Bias and Variance

In this formalism, we incorporate the effect of both residual gain and bandpass errors and estimate the bias and variance of the power spectrum. The methodology we use in this thesis can be used for any high dynamic range observation. Our interest has been towards detecting the power spectrum of redshifted 21- cm signal from EoR. The foregrounds dominate the radio sky in the desired frequency range of observations. The sky visibility $\tilde{V}^S(\vec{U}_i,\nu)$ has contribution from the redshifted 21- cm signal $\tilde{V}^{HI}(\vec{U}_i,\nu)$ as well as the foregrounds $\tilde{V}^F(\vec{U}_i,\nu)$. A brief discussion of the various components of foregrounds and their removal methods is presented in Chapter 2. We use the similar assumptions made in Chapter 4 and 5 that the foreground is well estimated and use the subtraction method to mitigate its effect on the sky visibility. These foreground subtracted visibilities are later used to estimate the redshifted 21-cm power spectrum. We correlate the visibilities from two different channels present in the grid to estimate the multifrequency angular power spectrum. The foregrounds are highly correlated across a large range of frequencies, while the H I 21- cm signal is highly uncorrelated (Bharadwaj and Ali, 2005; Santos et al., 2005). With the increase in frequency separation, the correlation of the H I signal would fall rapidly while that of the foregrounds be significant. As we are using the foreground subtracted visibilities to estimate the bias and variance, the effect of bright foregrounds manifests itself through the residual gain errors. We demonstrated earlier in Chapter 4 that the bias and excess variance in H I 21- cm power spectrum arise due to time-correlated gain errors; the frequency-correlated residual bandpass errors will further enhance the bias and variance. We denote the bias in the angular power spectrum estimate as $\mathscr{B}_{C_l}$ and its

variance as $\sigma^2_{C_l}$, and these are [*]:

$$\mathscr{B}_{C_l}(\ell,\Delta\nu) = \left[(n_{13}\chi + n_{12})\Sigma_2^{\delta} + (n_{13}+n_{12})\kappa_2\right]\frac{C_l(\ell,\Delta\nu)}{N_d} \tag{6.11}$$

$$\sigma^2_{C_l}(\ell,\Delta\nu) = \left[\frac{P^2_{HI}(\Delta\nu)}{N_G} + \frac{N_2 P_{HI}(\Delta\nu)}{N_B N_d} + \frac{N_2^2}{2N_B N_d^2}\right] + \left[2(\Sigma_2^{\delta}+\Sigma_2^{b})\frac{N_2}{N_B N_d^2}C_l(\ell,\Delta\nu)\right] + \left[2(n_{13}\chi+n_{12})\Sigma_2^{\delta} + (n_{13}+n_{12})\kappa_2)^2 + 4(\Sigma_2^{\delta}+\Sigma_2^{b})^2\frac{C_l(\ell,\Delta\nu)}{N_G N_d^2}\right]$$

$$\text{with } n_{13} = 2n_1' + n_3' \;;\; n_{12} = 2n_{1A}' + n_2' \;;\; \Sigma_2^{\delta} = \sigma^2_{\delta R} + \sigma^2_{\delta I} \;;\; \Sigma_2^{b} = \sigma^2_{bR} + \sigma^2_{bI}$$

$$\kappa_2(\Delta\nu) = \Sigma_2^{b}\xi(\Delta\nu)$$

where $C_l(\ell,\Delta\nu)$ is the angular power spectrum of the foregrounds, $\ell = 2\pi U$ and $\Delta\nu$ being the angular multipole and frequency separation respectively. The angular power spectrum of the foreground is much higher than the redshifted 21-cm signal; hence, the contribution to bias from the 21-cm signal can be ignored as they are much smaller than the signal. The quantities N_G and N_B are the number of independent estimates of the angular power spectrum and the total number of visibility correlations in a baseline annulus, and N_d is the number of days of observation. The quantity χ gives the integrated effect of the time-correlated residual gain errors over the uv-grids defined in Chapter 5 (see section 5.2.2, eqn 5.10). The average effect of frequency correlation of the bandpass errors, defined by normalized two-point correlation function $\xi(\Delta\nu)$ over a uv-grid is given by κ_2. As one estimate of $C_\ell(\Delta\nu)$ involves two channels only, it is directly proportional to $\xi(\Delta\nu)$. Both the real imaginary component of the bandpass errors contribute to it equally, hence it is given as $\kappa_2(\Delta\nu) = \Sigma_2^{b}\xi(\Delta\nu)$ with $\Sigma_2^{b} = \sigma^2_{bR} + \sigma^2_{bI}$.

We assume that the noise, residual gain, and bandpass errors are not correlated between the observations done on different days. Note that bias and variance change with baseline and

[*]An analytical path almost similar to Appendix B is used to arrive at these expressions, we do not present it explicitly here to avoid repetition.

hence with ℓ. Apart from N_2, Σ_2^{δ} and Σ_2^{b}, all the other factors in the expressions for bias and variance depend on ℓ through U. Here N_2 is the noise correlation defined in terms of per visibility mean square system noise σ_N as $N_2 = \langle N_i N_j^* \rangle = 2\sigma_n^2 \delta_{ij}$. In radio interferometric observations, noise in each visibility σ_N observed in a given frequency channel depends on the source equivalent flux density (SEFD), frequency width of the channel $\Delta\nu_c$ and integration time for each visibility $\Delta\tau$, and given as (Thompson et al., 1986)

$$\sigma_N = \frac{\text{SEFD}}{\sqrt{2\Delta\nu_c \Delta\tau}}. \tag{6.12}$$

As we are correlating the visibilities across different channels, the number of independent samples of noise correlation would be different and will depend on the frequency separation $\Delta\nu$. As we will consider the nearby channels, there will be more number of channel pairs and hence independent samples of noise correlations across the bandwidth (BW) of observation. With increasing frequency separations, the number of independent samples of noise correlations will fall; e.g., for $\Delta\nu \sim BW$, only one channel pair will be available. The noise correlation as a function of $\Delta\nu$ is

$$N_2(\Delta\nu) = \frac{\sigma_N^2}{(BW - \Delta\nu)/\Delta\nu_c} \tag{6.13}$$

The bias mainly originates from the time and frequency correlation of the gain and bandpass errors. Hence, the major contribution comes from the baseline pair fractions of type 1 and 3. The baseline pair of type 4 does not contribute to the bias and variance, as the gains from different antennae are not correlated. The terms kept in the first square bracket in the expression for $\sigma_{C_l}^2$ represent the variance in the absence of any residual error when the power spectrum estimator is unbiased. The terms in the second and third square brackets arise due to residual gain and bandpass errors and depend on the gain and bandpass properties of the instrument.

6.3 Analysing the Results: Effects of Time and Frequency-Dependent Gain Errors

In Chapter 5, we studied the effect of time-correlated gain errors and estimated the bias and variance of the power spectrum for the uGMRT baseline configuration at 150 MHz. We assume the gains are independent of frequency and perform correlations in single channels. This work includes the frequency-dependent bandpass errors and estimates the bias and variance. Here we use the SKA-Low 1 baseline configurations to analyze our analytical results. SKA-1 Low is an aperture phased array located at the Murchison Radio-astronomy Observatory (MRO) site in Western Australia (Koopmans et al., 2015; Mellema et al., 2013). It is currently in the construction phase and will be operational in the low-frequency range (50 MHz - 350 MHz) of the SKA bandwidth. It will be composed of 512 stations with the largest baseline $\sim$ 65 km., each station consisting of 256 log-periodic dual-polarized antennas.

6.3.1 Baseline pair fractions for SKA-1 Low

We estimate the baseline pair fractions for SKA- 1 Low baseline configuration at 150 MHz. As SKA-1 Low contains 512 stations, estimating the baseline pair fractions for SKA-1 Low is computationally very challenging. First, we estimate the baseline pair fractions for a single channel of observation as defined in Chapter 5 (we may call it old baseline pair fractions), and from those, we calculate the baseline pair fractions for multichannel observation (new baseline pair fraction) as discussed in section 6.2.1. We use the integration time of 4 seconds and a grid size of 0.004 $k\lambda$. In figure 6.1 we show the variation of new baseline pair fractions for SKA-1 Low with the baseline **U**, derived from old baseline pair fractions, for 8 hours of observation. Here we see that baseline pair fractions of type 4 dominate over others, and that of type 1A and 2 is very small. We

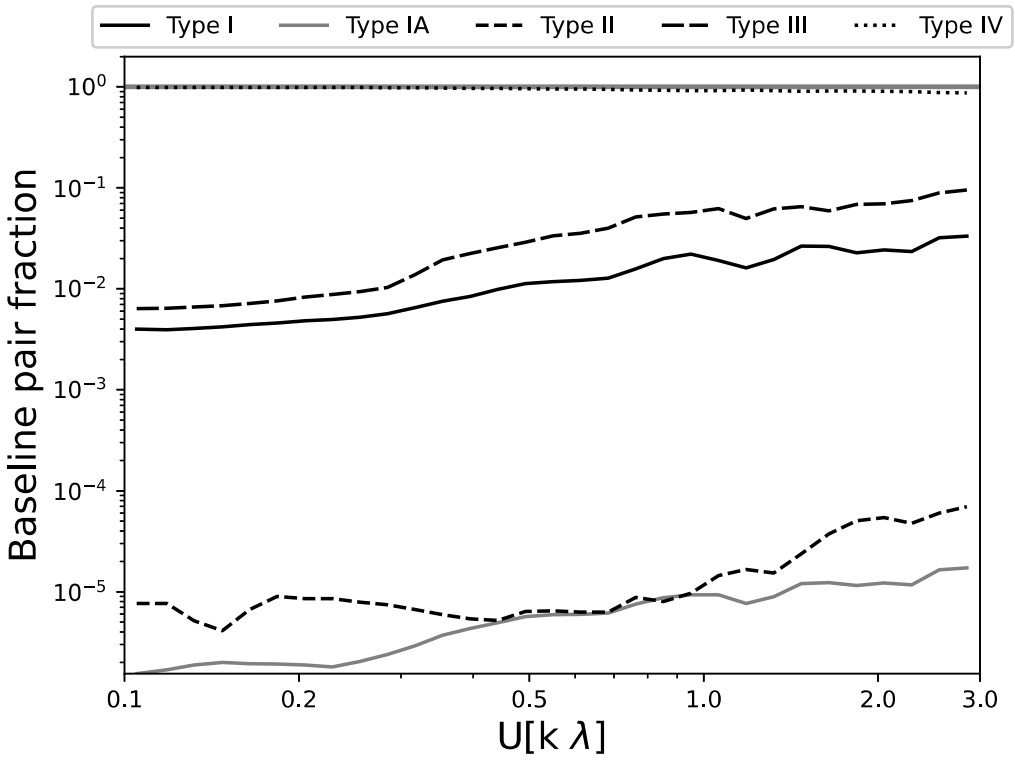

Fig. 6.1 Variation of baseline pair fractions with baseline U for SKA-1 Low, 8 hours of observation. Integration time is 4 sec and uv-grid size is 0.004 $k\lambda$. The fraction of baseline pairs is dominated by type 4 for the above-used parameters.

restrict ourselves only up to the baseline $\mathbf{U} = 3\ k\lambda$, as the 21- cm signal falls rapidly after 1 $k\lambda$ (discussed in Chapter 4 and 5).

6.3.2 Foreground Model and Other Observational Details

For the analysis presented in this paper, we use the point source foregrounds and adapt the model of foregrounds from Santos et al. (2005). We use the foreground subtraction method and estimate the angular power spectrum from foreground subtracted visibilities. A detailed discussion on foregrounds and the model adapted is presented in Chapter 2. For convenience, here again, we present the model of the multi-frequency angular power

spectrum for different components of the foreground

$$C_\ell(\nu_1,\nu_2) = A\left(\frac{1000}{\ell}\right)^{\beta}\left(\frac{\nu_f}{\nu_1}\right)^{\alpha}\left(\frac{\nu_f}{\nu_2}\right)^{\alpha} I_l(\nu_1,\nu_2) \tag{6.14}$$

where A is the amplitude and β and α (spectral index) are the power law indices for the ℓ. For the point source component of foregrounds the value of the parameters is $A = 57.0\ mK^2,\ \beta = 1.1,\ \alpha = 2.07,$ and $\zeta = 1.0$ (see table 2.1). Although we choose a point source model, the methodology presented here can be applied to any foreground model, given that the foregrounds are well modeled and their effects can be directly subtracted.

We fix the parameters for time-dependent gain error models based on our analysis presented in Chapter 5. The longest possible integration time for SKA-1 Low array avoiding the time smearing effect is ~ 7 seconds, so we choose integration time to be 4 seconds. The correlation time of the time-correlated gains is fixed at 4 seconds, the same as the integration time. For the uncertainty in the real and imaginary part of the time-correlated gain errors σ_R and σ_I, we choose a relatively moderate value 0.1 % for both, and denote it with σ_g, i.e. $\sigma_R = \sigma_I = \sigma_g = 0.1\%$. The uv-grid size ΔU within which visibility correlations are performed is kept at 0.004 $k\lambda$. The width of the uv-grid is chosen to be smaller than the maximum width up to which visibilities remain correlated and large enough to contain sufficient visibility measurements. We consider an observation using SKA-1 Low baseline configuration with a bandwidth of 8 MHz, centered at 150 MHz. There are 128 channels across the bandwidth of observation with a channel-width $\Delta\nu_c$ = 62.5 kHz. In figure 6.2 we show the variation of the bias $\mathscr{B}_{C_l}$ and standard deviation σ_{C_l} of the angular power spectrum for one day of observation, plotted with the frequency separation $\Delta\nu$. We keep the bandpass error σ_b = for 0.1 and 1.0 %, and frequency correlation length ν_c = 0.5 and 1.0 MHz. As mentioned above, the integration

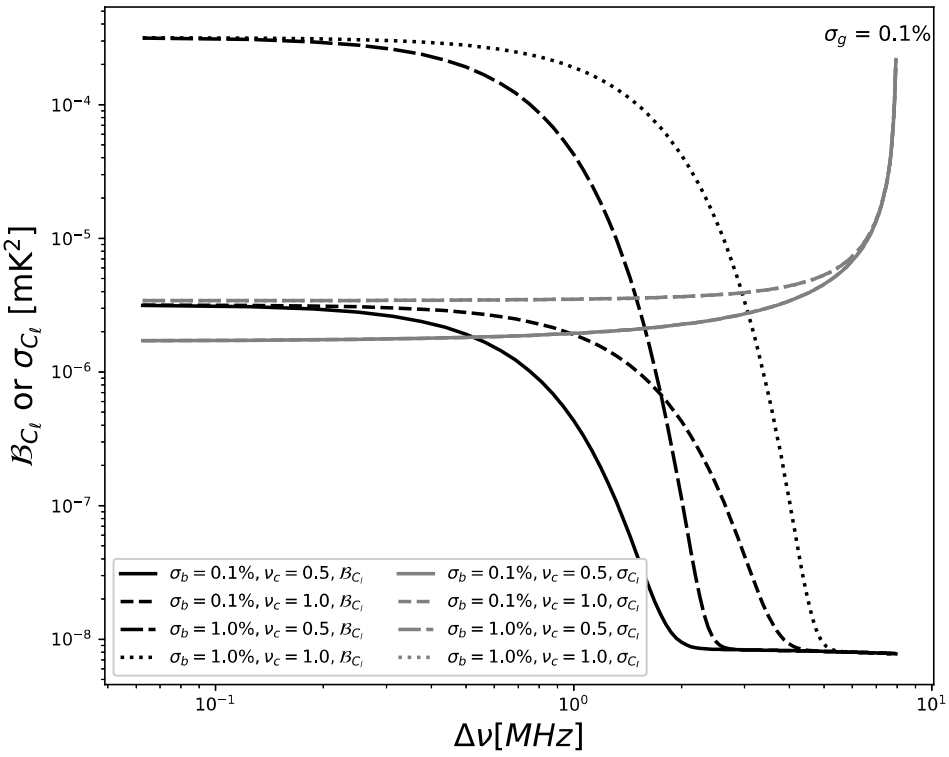

Fig. 6.2 Variation of $\mathscr{B}_{C_l}$ (black) and σ_{C_l} (gray) with frequency separation $\Delta\nu$ for 0.1 and 1.0 % bandpass error (σ_b) and 0.5 and 1.0 MHz of correlation length in frequency (ν_c). Integration time and correlation time are 4 sec, and uv-grid size is 0.004 $k\lambda$. The gain error value σ_g is kept constant at 0.1%. The bandwidth of observation is 8 MHz. The total observation time is 8 hours (1 day of observation).

and correlation time are 4 sec, and the uv-grid size is 0.004 $k\lambda$. Throughout this analysis, gain error value σ_g and BW is kept constant at 0.1% and 8 MHz, respectively. The bias $\mathscr{B}_{C_l}$ is plotted with black color lines, and σ_{C_l} is plotted with gray lines. Here we see that for SKA-1 Low array, both $\mathscr{B}_{C_l}$ and σ_{C_l} increases with increase in the value of bandpass error σ_b, although the increment in σ_{C_l} is not as prominent as in $\mathscr{B}_{C_l}$. For the given set of parameters, σ_{C_l} is mostly dominated by the system noise and gain, and bandpass error terms only become effective at their larger values. The bias $\mathscr{B}_{C_l}$ also increases with an

increase in the frequency correlation length ν_c. This is because the bandpass errors remain correlated for the larger frequency scales and contribute more to the bias. At very small frequency separations $\Delta\nu$, this is not visible as at smaller $\Delta\nu$, the effect of gain errors will dominate through the strongly correlated foregrounds. On the other hand, σ_{C_l} remains the same for different frequency correlation lengths ν_c; this is because σ_{C_l} is dominated by the terms in the first and second square brackets. The bias $\mathscr{B}_{C_l}$ for both values of σ_b initially remains approximately constant at lower frequency separations $\Delta\nu$ and then falls sharply as $\Delta\nu$ increases. This is due to the decorrelation of the foreground signal and the bandpass errors with $\Delta\nu$. At larger $\Delta\nu$, it again becomes constant for all the parameter values, this constant level of bias is due to the time-correlated gain error properties σ_g and χ. In contrast to the $\mathscr{B}_{C_l}$, σ_{C_l} increases sharply at larger $\Delta\nu$, because at larger $\Delta\nu$, the number of independent samples of noise correlations becomes very small.

Figure 6.3 shows the variation $\mathscr{B}_{C_l}$ and σ_{C_l} with bandpass errors for a set of parameters $\Delta\nu$ and ν_c given in the legend. Similar to the figure 6.2, here we see that both $\mathscr{B}_{C_l}$ and σ_{C_l} increases monotonically with σ_b, except for $\Delta\nu = 4.0$ MHz. At this large value of frequency separations, the foreground signal, as well as the bandpass errors both, decorrelates to smaller values. Again the σ_{C_l} doesn't changes with ν_c and it remains the same for both values of ν_c. As expected, $\mathscr{B}_{C_l}$ at $\Delta\nu = 1.0$ MHz is much larger than its value at $\Delta\nu = 1.0$ MHz. For smaller σ_b the σ_{C_l} is more for $\Delta\nu = 4.0$ MHz compared with $\Delta\nu = 1.0$ MHz due to noise dominance, but as σ_b increases σ_{C_l} for both the values of $\Delta\nu$ becomes similar. This shows that at larger σ_b the effect of bandpass errors starts dominating.

We compare the results for SKA-1 Low and uGMRT in figure 6.4, where we have shown the $\mathscr{B}_{C_l}, \sigma_{C_l}$ for SKA and uGMRT for different days of observation. To check, we fix the gain and bandpass error to 0.1%, and $\nu_c = 0.5$ MHz. For SKA, the quantities are shown in the legend with $\mathscr{B}_{SKA}$ (black solid curve), σ_{SKA} (solid gray curve), while for the uGMRT, with $\mathscr{B}_{GM}$ (black dashed curve), and σ_{GM} (gray dashed curve). For comparison,

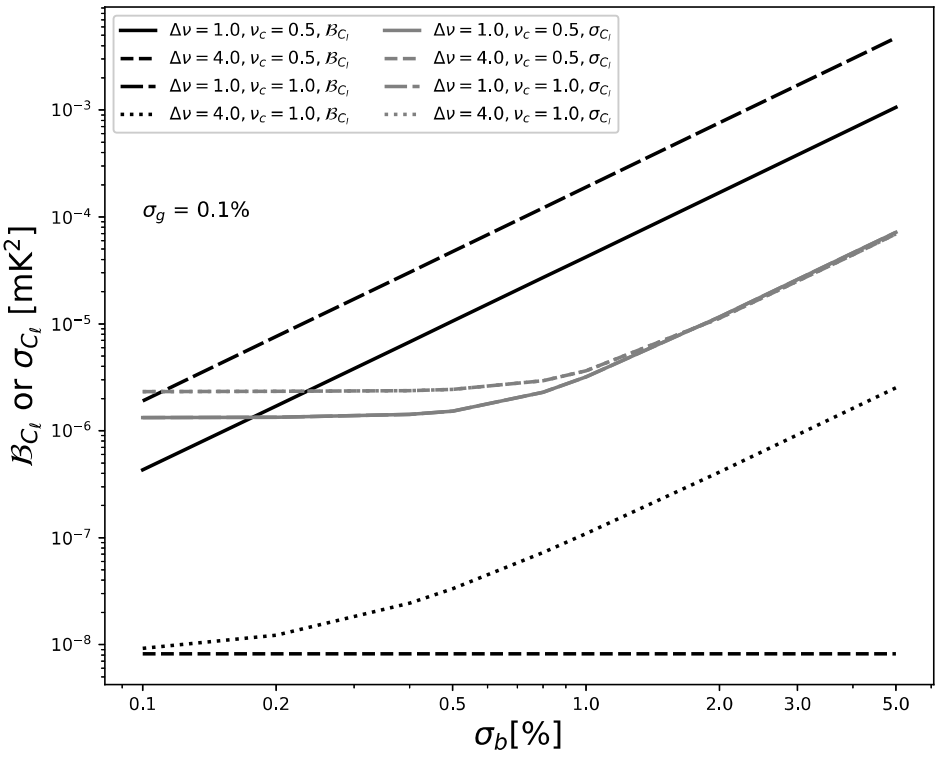

Fig. 6.3 Variation of $\mathscr{B}_{C_l}$ (black) and σ_{C_l} (gray) with bandpass error σ_b, for 0.5 and 1.0 MHz of correlation length in frequency (ν_c) and frequency separation $\Delta\nu$ = 1.0 and 4.0 MHz. The total observation time is 1 day.

we also have plotted the redshifted 21- cm signal angular power spectrum with the solid green curve. We adopted the 21- cm signal power spectrum from Bharadwaj and Ali (2005). Here the signal is assumed to be correlated up to 0.3 MHz. The dashed green curve shows the $3-\sigma$ level of the 21- cm signal. The integration and correlation time for uGMRT SKA-1 Low is 16 and 4 seconds, respectively. We see that for $3-\sigma$ detection of 21- cm angular power spectrum, uGMRT will require 3200 days of observation, while with SKA-1 Low, it will need only 64 days with the given set of parameters used here. Though these results are susceptible to the instrument's gain properties, and in actual observation, the results may vary, it is clear from here that SKA-1 Low will require $\sim$ 50

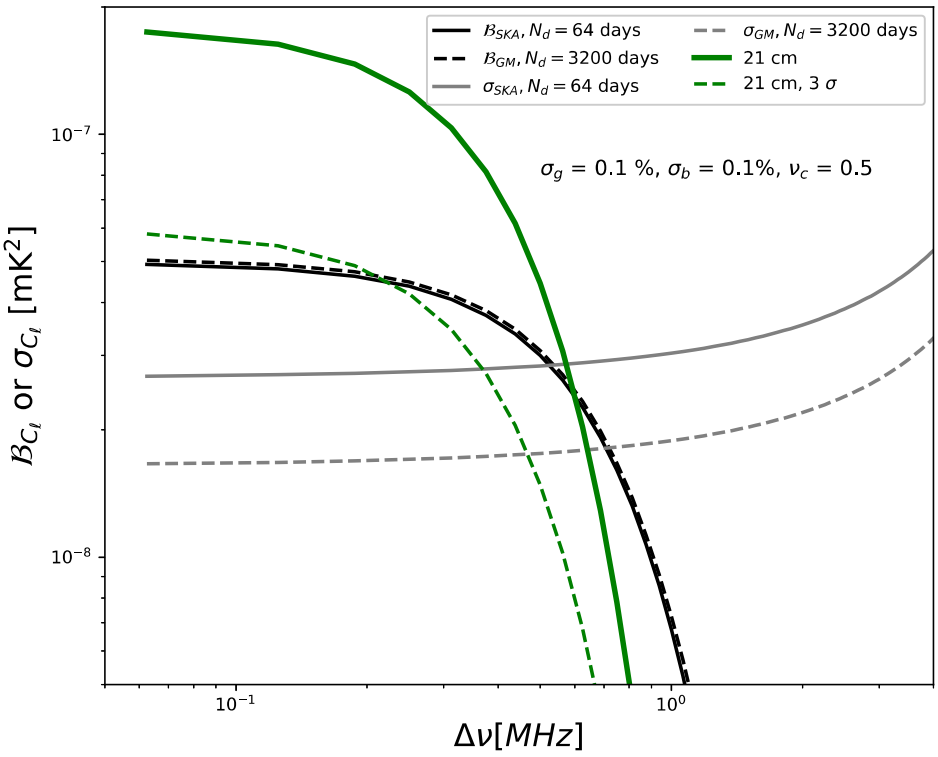

Fig. 6.4 Bias and Variance for SKA-1 Low ($\mathscr{B}_{SKA}, \sigma_{SKA}$) and uGMRT ($\mathscr{B}_{GM}, \sigma_{GM}$) with $\Delta\nu$ for $\sigma_g = \sigma_b = 0.1$, and $\nu_c = 0.5$ MHz. Here we compare the performance of SKA-1 Low and uGMRT for their optimum correlation and integration time.

times less observation time. This is primarily because of the baseline configuration and larger collecting area of the SKA-1 Low.

Figure 6.5 is the same as figure 6.2 where we again show the variation of $\mathscr{B}_{C_l}$ and σ_{C_l} for SKA-1 Low, but here we have not included the contribution from the baseline pairs of type 1 and 3. As the majority of the bias was coming through the aforementioned baseline pair types due to the correlation in the gain and bandpass, we remove such baseline pairs from the estimation of $\mathscr{B}_{C_l}$ and σ_{C_l}. This significantly reduces the bias $\mathscr{B}_{C_l}$. However, being dominated by the contribution from noise bias and the gain and bandpass errors

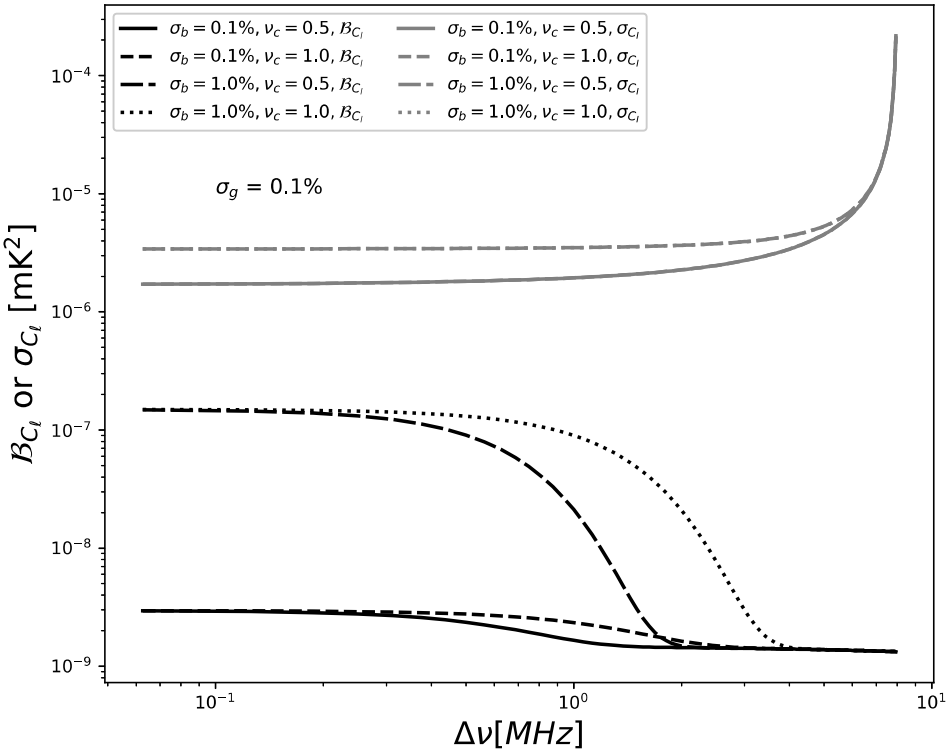

Fig. 6.5 Variation of $\mathscr{B}_{C_l}$ (black) and σ_{C_l} (gray) for SKA-1 Low with frequency separation $\Delta\nu$ for $\sigma_b = 0.1$ and 1.0 % and $\nu_c = 0.5$ and 1.0 MHz, same as figure 1 but after excluding the effect of gain and bandpass errors coming through the baseline pairs of type 1 and 3.

through σ_g and σ_b (terms in the first and second square bracket in the expression of σ_{C_l}), the uncertainty σ_{C_l} doesn't reduce with the removal of correlation effects.

An alternate way of presenting the 3D power spectrum in literature is the cylindrical averaged power spectrum where $P(k_\perp, k_\parallel)$ is the quantity of interest (e.g. Datta et al. (2010)). As discussed earlier, the foreground signal in $(k_\perp, k_\parallel)$ plane is contained inside a wedge, and one uses the rest of the plane for the estimation of the power spectrum. We consider the bias in the MAPS with $\sigma_b = 1\%$ and $\nu_c = 0.5$ MHz (black dot-dashed line in figure 6.2), with the gain error value $\sigma_g = 0.1\%$. The bias corresponding to the cylindrical averaged power spectrum is shown in the figure 6.6. The values are normalized to the

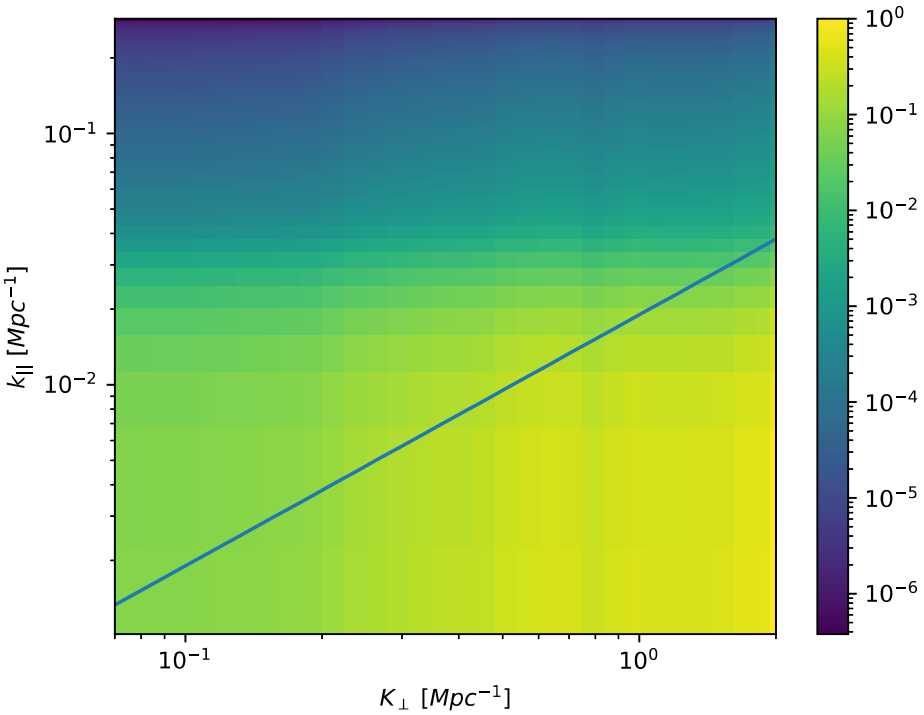

Fig. 6.6 This figure shows the bias in the cylindrical averaged power spectrum estimate for $\sigma_b = 1\%$ and the frequency correlation width $\nu_c = 0.5$ MHz in $(k_\perp, k_\parallel)$ plane. The gain error value σ_g is kept at 0.1%. The bias is normalized to the highest value. The blue line shows the wedge boundary.

highest value of the bias in this plane. The blue line corresponds to the wedge boundary. Note that from the foreground avoidance discussions, one expects the foreground to be contained below the blue line. We observe here that the time and frequency-correlated residual gain errors bleed the effect of foreground to the region which otherwise is expected as foreground free.

6.4 Discussion and Conclusion

This chapter discusses the effect of frequency-correlated bandpass error in the angular power spectrum estimation of the redshifted 21- cm signal in the presence of bright

foregrounds. We considered a point source foreground model and used the visibility-based foreground subtraction method for foreground removal. We consider a time and frequency-dependent gain error model, where both time and frequency-correlated gain and bandpass errors are present. Complementary to the analytical framework discussed in Chapter 5, we derive the expressions for bias and variance of the angular power spectrum for the gain model discussed in section 6.1. Further, we analyze our analytical results, and the major findings of our analysis are as follows:

- We present an analytical framework to estimate the bias $\mathscr{B}_{C_l}$ and variance $\sigma^2_{C_l}$ of the angular power spectrum estimates in the presence of bright foregrounds by modeling the time and frequency correlated gain and bandpass errors.

- The bias $\mathscr{B}_{C_l}$ and variance $\sigma^2_{C_l}$ of the angular power spectrum depend on the antenna gain and bandpass properties like variance and correlation of the gain errors, foreground, and array configuration.

- The $\mathscr{B}_{C_l}$ and σ_{C_l} increases with an increase in the residual gain and bandpass errors and the time and frequency correlation. $\mathscr{B}_{C_l}$ decreases at larger frequency separation $\Delta\nu$ while σ_{C_l} shoots up at larger $\Delta\nu$ which is due to the larger noise correlations.

- The majority of the bias originates from the time and frequency correlation of the gain and bandpass errors. Removing the contribution from baseline pair fractions of type 1 and 3 to exclude the correlation effect reduces the bias significantly. This can be a possible remedy to the problem.

- The SKA-1 Low performs much better than the uGMRT, and it promises to detect the 21- cm signal power spectrum in ~ 50 times less observation time than that of the uGMRT. For ska-1 Low baseline pairs of type 4 dominates for which the effect of gain errors is zero. Due to the larger collecting area and baseline contribution of

the SKA-1 Low array, the effect of time and frequency-correlated gains will have lesser effects.

The methodology followed in this framework is general and can be used for any high dynamic range observations. Here we are particularly interested in detecting the 21-cm signal power spectrum. The bias and variance depend on the foreground model, the array configuration of the telescope, and the antenna gain and bandpass properties; therefore, results will vary with these. For a telescope of interest, the bias and variance can be estimated prior to the observation, given the observation parameters and the above properties are known.

****