Spectral Calibration Requirements of Radio Interferometers for Epoch of Reionisation Science with the SKA

Cathryn M. Trott[1,2] and Randall B. Wayth[1,2,3]

[1]International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6103, Australia
[2]ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO), Redfern NSW 2016, Australia
[3]Email: cathryn.trott@curtin.edu.au

Abstract

Spectral features introduced by instrumental chromaticity of radio interferometers have the potential to negatively impact the ability to perform Epoch of Reionisation and Cosmic Dawn (EoR/CD) science. We describe instrument calibration choices that influence the spectral characteristics of the science data, and assess their impact on EoR/CD statistical and tomographic experiments. Principally, we consider the intrinsic spectral response of the antennas, embedded within a complete frequency-dependent primary beam response, and instrument sampling. The analysis is applied to the proposed SKA1-Low EoR/CD experiments. We provide tolerances on the smoothness of the SKA station primary beam bandpass, to meet the scientific goals of statistical and tomographic (imaging) of EoR/CD programs. Two calibration strategies are tested: (1) fitting of each fine channel independently, and (2) fitting of an nth-order polynomial for each ∼1 MHz coarse channel with $(n+1)$th-order residuals ($n = 2, 3, 4$). Strategy (1) leads to uncorrelated power in the 2D power spectrum proportional to the thermal noise power, thereby reducing the overall sensitivity. Strategy (2) leads to correlated residuals from the fitting, and residual signal power with $(n+1)$th-order curvature. For the residual power to be less than the thermal noise, the fractional amplitude of a fourth-order term in the bandpass across a single coarse channel must be <2.5% (50 MHz), <0.5% (150 MHz), <0.8% (200 MHz). The tomographic experiment places constraints on phase residuals in the bandpass. We find that the root-mean-square variability over all stations of the change in phase across any fine channel (4.578 kHz) should not exceed 0.2 degrees.

Keywords: (cosmology:) dark ages, reionization, first stars – techniques: interferometric – instrumentation: interferometers

1 INTRODUCTION

Detection of the signal from the Epoch of Reionisation (EoR), and estimation of the spatial properties of neutral hydrogen from the early Universe, are challenging experiments, requiring high fidelity data with known statistical properties (Koopmans et al. 2015; Parsons et al. 2010; van Haarlem et al. 2013; Bowman et al. 2013). For these, instrumental effects that bias the signal can have a large impact on the success of the experiment, and interpretation of results. Statistical experiments, such as measurement of the spatial power spectrum of neutral hydrogen (H I) brightness temperature fluctuations, implicitly rely on datasets that have been observed with a spectrally smooth instrument, in order to associate spectral structure with intrinsic spatial fluctuations. Similarly, H I tomography (imaging as a function of fine spectral channel) aims to detect and measure weak spectral lines from H I clouds, thereby requiring smooth instrumental

spectral characteristics in order to not bias the signal. Further, phase residuals in the calibration of stations reduces imaging dynamic range, destroying the detectability of weak signals (Perley 1999).

Low-frequency aperture array interferometers, such as the Square Kilometre Array (SKA-Low), use the electronic combination of individual element antennas in an 'aperture array', to form a single beam on the sky. This set of beam-formed elements forms a single station. These stations are then used as the primary aperture with which to form cross-correlations for interferometric visibilities. The advantages of aperture array telescopes include flexible and dynamic allocation of elements to a station, flexible, and dynamic assignment of weights to alter the beam shape (apodisation), and durability due to the lack of moving parts. Such trade-offs are being studied for the SKA (Grainge 2014; Mort et al. 2016). The disadvantages include direction-dependent primary beam shapes, strong off-zenith

instrumental polarisation, and mutual coupling between elements (Sutinjo et al. 2015).

Calibration of radio interferometers is crucial to obtaining science-quality data. For precision quantitative experiments, accurate and precise calibration is required for reliable results to be achieved and claimed. Calibration is the process by which the electronic gain amplitude and phase is set for each receiving element, such that the sky position and source flux density of any sky source is correct. Calibration of interferometric arrays implies measurement and setting of the complex gain parameters for each station, as constructed from the gains of each element. Direction-independent, but frequency-dependent bandpass and phase calibration is required for each station that will form an interferometric baseline. The calibration process for a low-frequency interferometer typically involves simultaneously solving for station direction-independent and -dependent complex-valued gains (possibly including ionospheric effects) using a sky model and primary beam model (Mitchell et al. 2008; Yatawatta et al. 2008; Kazemi, Yatawatta, & Zaroubi 2013; Tasse et al. 2013). Principally, each antenna will have a bandpass that must be calibrated. The simplest model is to treat each frequency channel as being independent, but other approaches can be used if characteristics of the bandpass shape are known. In general, we are separating the direction-independent bandpass fitting from the direction-dependent beam fitting, but these may be combined to exploit the expected characteristics of the instrument (e.g., Yatawatta 2015, where a smooth polynomial is used to regularise the bandpass solution in a full direction-dependent calibration). In these more advanced approaches, instrument responses that fail the underlying smoothness tests (e.g., rapid changes in the bandpass shape) will still leave residual bandpass structure, and yield potential biases in the final calibrated data. Here, we try to take a general approach to fitting of the bandpass, under the assumption of a smooth response in frequency.

Understanding the instrument spectral response is especially important for EoR and Cosmic Dawn (CD) experiments (Trott et al. 2016; Offringa et al. 2016; Barry et al. 2016). Probing an emission line over cosmological volume is achieved by detecting and measuring the line as a function of redshift, and hence as a function of frequency. For statistical experiments where the spatial structure of H I is probed by a Fourier-like transform along frequency (obtaining information about the characteristic size scale of neutral and ionised regions), unmodelled instrumental spectral structure that is localised in frequency space contaminates the entire signal space after Fourier transform, thereby biasing results. Likewise, for imaging experiments, which aim to directly detect and map individual ionised bubbles, spectral structure can mimic these structures. Further, residual phase in the data smears the signal from other sources in the field, reducing the contrast (dynamic range) between the background and the bubble of interest.

In addition to the *imprecision* of gain calibration parameters due to limited information, there is the potential for *inac-*

curate calibration if an input sky model, or input instrument model is incomplete or incorrect. The latter is determined by the bandpass shape of the station antennas and electronics, and may be element- and time-dependent. Early testing of the SKA low-frequency aperture array element dipoles (SKALA) show strong spectral features (de Lera Acedo et al. 2015). These features have the potential to leave residual signal in the data if calibration is not carefully performed. In this work, we describe two common calibration methods, and explore the effects of imprecise and inaccurate calibration for SKA-Low for the planned EoR experiments.

2 EXPERIMENTS AND THEIR SCIENCE REQUIREMENTS

The EoR/CD program includes three major experiments:

1. Power spectrum: The underlying data for the power spectrum estimation will be visibilities, which have been averaged to ∼100 kHz and 5 s. Simplistically, these visibilities are gridded and integrated, and then squared and normalised by volume to form a power spectrum (units: K^2 Mpc^3). Tolerances for spectral gradients in the post-calibration bandpass will be computed on this basis, using the station calibration timescale to define the information available for calibration.
2. Tomography: The imaging data consist of visibilities averaged to ∼100 kHz and integrated.
3. 21-cm Forest: The data will be high spectral resolution image cubes, with fine spectral channels (∼4 kHz).

Table 1 lists the relevant parameters of the data and experiments, as used in this analysis.

The current SKA Level 1 requirement for station complex gain calibration is 600 s[1]. This timescale will be used to determine the temporal tolerance of the bandpass stability. The spectral resolution allows 65 536 fine channels across the 300-MHz bandwidth[2], yielding 4.578-kHz channels.

3 APPROACH

We use the array baseline design[3], incorporating 94 super-stations of six logical ∼33-m stations arranged in a floral configuration. This configuration provides a large amount of sensitivity to angular modes on the sky corresponding to the ∼33 m intra-superstation baseline, and twice this length.

---

[1] SKA System Requirement SYS_REQ-2635
[2] SKA System Requirement SYS_REQ-2148
[3] Described in Document SKA-SCI-LOW-001 (SKA1-LOW Configuration)

Table 1. Spectral and temporal parameters of EoR/CD experiments, as described in Koopmans et al. (2015). [1] Averaging to match an integer number of fine channels.

Experiment	$\Delta\nu$	Δt	Product	Total time	Δ_T^2 ($k = 0.03$ Mpc^{-1})
Power spectrum (54 MHz)	109.8 kHz[1]	5 s	Visibilities	1 000 h	0.2 mK^2
Power spectrum (84 MHz)	109.8 kHz[1]	5 s	Visibilities	1 000 h	0.02 mK^2
Power spectrum (143 MHz)	109.8 kHz[1]	5 s	Visibilities	1 000 h	0.005 mK^2
Experiment	$\Delta\nu$	Δt	Product	Total time	ΔT
Tomography (>50 MHz)	100 kHz	1 000 h	Image cube	1 000 h	1 mK
Experiment	$\Delta\nu$	Δt	Product	Total time	ΔS
21-cm absorption (>50 MHz)	4.5 kHz	1 000 h	Image cube	1 000 h	0.2 mJy

3.1 Power spectrum

We compute the thermal noise uncertainty for the experiments described in Table 1, for the 2D (cylindrical) power spectrum. The *uv*-plane resolution is defined by the field of view (FOV) of the telescope at a given frequency. We choose an experiment bandwidth of 10 MHz around three central frequencies: 50 MHz ($z \sim 25$), 150 MHz ($z \sim 8.5$), and 200 MHz ($z \sim 6.1$). This thermal noise estimate provides the reference level, below which, spectral features may be tolerated.

We compare the thermal noise and bias tolerance to the signal in the power spectrum due to calibration errors and residuals, using a statistical model of the point source population. The following approach is taken:

1. Compute the signal available from the sky to perform calibration, using a statistical model of unresolved point sources, as observed by an interferometer with a frequency-dependent primary beam. This model serves two purposes: (a) as a reference signal of the power in the sky, and its statistical signature in power spectrum parameter space; (b) as the source population to which calibration is applied (i.e., the sky model), and from which residual signal stems after application of an incorrect calibration model.
2. Compute the precision with which calibration is achievable with an ideal estimator for two calibration schemes, based on the statistical sky model, and propagate uncertainties into the power spectrum.
3. Compute the residual signal (accuracy) from the sky model due to unfitted spectral structure in the calibration process, and propagate into the power spectrum.

3.2 Tomography

We compute the loss in dynamic range in an image cube due to residual phase gradients across the fine frequency channels. We compute the residual gradient in phase that may be tolerated such that an intrinsic 1 mK brightness temperature fluctuation in a 100-kHz spectral channel is detectable.

4 POWER SPECTRUM

4.1 Signal due to calibration uncertainties and residual spectral structure

One possible calibration approach for EoR will be composed of a set of steps to model and remove sky signal, potentially including a direct subtraction of sources through a Global Sky Model (GSM). This step occurs after calibration of each station's bandpass, and assumes accurate and precise calibration. There are two sources of residual spectral signal due to calibration: (1) In the case of bandpass calibration *uncertainties*, residual noise-like sky signal will remain in the data from fitting of the bandpass calibration parameters using the information available in the sky; (2) residual sky signal power and structure will remain from any spectral curvature terms that are not accounted for in the calibration model. Subtraction of the GSM will leave these residuals in the data due to incomplete and imprecise calibration of the complex station gain parameters.

The model and approach used for bandpass calibration will affect the calibration precision and the residual signal. Conceptually, there is a trade-off between precision and complexity of the model. For example, a polynomial fitted in amplitude and phase to each coarse channel of order N will yield a handful of parameters with precision dependent on the information available in the data obtained over the calibration timescale, and leave residual signal with spectral curvature of order $>N$. The parameter estimates are also likely to be correlated in frequency, thereby correlating the fine frequency channels. Conversely, each fine channel can be independently fitted, using a single amplitude and phase estimate. This would yield uncorrelated uncertainties between fine channels, but requires the fitting of more parameters, thereby reducing signal to noise.

In this work, we explore both precision and accuracy. We compute the signal introduced by imprecise bandpass calibration, leading to additional noise in the data. We then consider gain amplitude and phase residuals, whereby the calibration model is incorrect, and spectral structure remains. The former introduces additional signal due to noise-like fluctuations in the visibilities. The latter introduces the additional signal into the power spectrum due to residual signal from the sources

in the sky. In all cases described, we are computing signal power that is *additive* to thermal noise and sky power in the power spectrum space.

We take two common calibration approaches—low-order polynomial fitting and individual fine-channel fitting—and compute the precision with which these parameters can be estimated and their correlations using the Cramer–Rao Bound (Kay 1993). The parameters of importance are the complex-valued gains for each station (or, equivalently the amplitude and phase) over a solution interval in frequency. The uncertainties on these parameters, which may be correlated, are propagated back into the bandpass solution, and through to the power spectrum, to yield the *calibration uncertainty power*. Remaining spectral curvature is then propagated into the power spectrum to yield the *residual spectral power*. The former sets the noise floor for the calibration method, while the latter informs the spectral gradient tolerance for the bandpass solution by determining how much additional bias signal is introduced into the experiment.

In all cases, we are considering real sky signal that remains in the data, and yields signal in the power spectrum that is additive to the thermal noise power. We also do not specify an actual procedure for performing the calibration. Rather, the approach is valid for any general unbiased estimator, hence represents the best possible result that can be obtained using the proposed models.

4.2 Formalism

We use a statistical model of the instrument, which allows us to propagate spectral signal into the power spectrum parameter space. This model includes all instrumental spectral features, including a frequency-dependent primary beam, and instrumental layout (configuration). The covariance between frequency channels ν and ν', at angular wavenumber u is given generically by Trott et al. (2016):

$$\boldsymbol{C}_{\mathrm{FG}}(\nu, \nu'; u)$$
$$= \rho(\nu; \nu'; u) \int_0^\infty B(l; \nu) B(l; \nu') J_0\left(2\pi(ul)(\nu - \nu')\right) l dl \quad \mathrm{Jy}^2. \tag{1}$$

Here, $B(l, m; \nu)$ describes the frequency-dependent primary beam response to the sky at radial position $l = |\mathbf{l}|$, ρ contains the frequency correlations at spatial wavemode u, and the Bessel function performs the Fourier Transform to wavenumber space at mode u ($\propto k_\perp$). This expression is appropriate for a circularly-symmetric primary beam, and allows the propagation of any spectral function into the power spectrum. It is then propagated into the 2D power spectrum parameter space according to a Fourier transform:

$$\langle P_k(k_\perp, \eta) \rangle = \mathcal{F}^\dagger \boldsymbol{C}_{\mathrm{FG}} \mathcal{F}, \tag{2}$$

with relevant cosmological co-ordinate conversions at a given frequency (redshift) to convert the line-of-sight wavenumber,

η (Hz^{-1}), to $k_\parallel$ (Mpc^{-1}), and the angular wavenumber u to $k_\perp$.

4.2.1 Reference signal power from the sky

To estimate the power due to astrophysical sources in the sky, we use a statistical model for extragalactic point sources, based on a parametric source counts function and a Poisson random position distribution across the sky. The model used, and the framework for it, are described in Trott et al. (2016) and Trott & Tingay (2015). The number of sources of a given flux density in a unit area of sky is Poisson-distributed ($\mathcal{P}()$):

$$N(S, S + dS) dS \, d\mathbf{l} \sim \mathcal{P}(\langle N \rangle), \tag{3}$$

where $\langle N \rangle$ is the expected number, given parametrically by

$$\langle N(S, S + dS) \rangle(\nu) = \frac{dN}{dS}(\nu) \, dS \, d\mathbf{l} \tag{4}$$

$$= \alpha \left(\frac{\nu}{\nu_0}\right)^\gamma \left(\frac{S_{\mathrm{Jy}}}{S_0}\right)^{-\beta} dS \, d\mathbf{l}. \tag{5}$$

We use values of $\alpha = 4100 \, \mathrm{Jy}^{-1}\mathrm{sr}^{-1}$, $\beta = 1.59$ and $\gamma = -0.8$ at 150 MHz, in line with recent measurements (Intema et al. 2011; Gervasi et al. 2008). This Poisson statistical model is propagated into the power spectrum, using a full frequency–frequency covariance to describe the spectral correlations, and instrument sampling to produce the 'wedge' effect of foreground contamination.

The frequency–frequency covariance at angular wavenumber $u = |x| c / \nu_0$, for a Gaussian-shaped primary beam of a station of diameter D is given by (Trott et al. 2016)

$$\boldsymbol{C}_{\mathrm{FG}}(\nu, \nu') = \frac{\alpha}{3 - \beta} \left(\frac{\nu}{\nu'}\right)^{-\gamma} \frac{S_{\max}^{3-\beta}}{S_0^{-\beta}} \frac{\pi c^2 \epsilon^2}{D^2} \frac{1}{\nu^2 + \nu'^2}$$
$$\exp\left(\frac{-u^2 c^2 f(\nu)^2 \epsilon^2}{4(\nu^2 + \nu'^2) D^2}\right), \tag{6}$$

where $\epsilon = 0.42$ converts an Airy disk to a Gaussian characteristic width, and $f(\nu) = (\nu - \nu')/\nu_0$. This model describes the covariance for a full extragalactic point source population within the FOV of an SKA1-Low station. Equation (6), propagated according to equation (2), then contains the expected signal in the 2D parameter space due to sources within the FOV, for a perfectly calibrated instrument. For the SKA, we consider a 12-MHz bandwidth for each experiment (16 coarse channels), and an FOV associated with a 35-m station at each frequency. This yields a power spectrum parameter space (at 150 MHz) of $k_\parallel = [0.01, 1.60]h \, \mathrm{Mpc}^{-1}$, $k_\perp = [0.02, 2.07]h \, \mathrm{Mpc}^{-1}$.

The top-left panels of Figure 1–3 display the power due to unresolved point sources for the experiments described in Table 1 and a 2D, cylindrically averaged parameter space $(k_\perp, k_\parallel)$. All unresolved sources with a flux density below 5 Jy remain in the data. This is to ensure the model is statistically consistent (i.e., brighter sources are rare and not accurately characterised by this statistical model). The unresolved point source model describes the total sky power as observed by
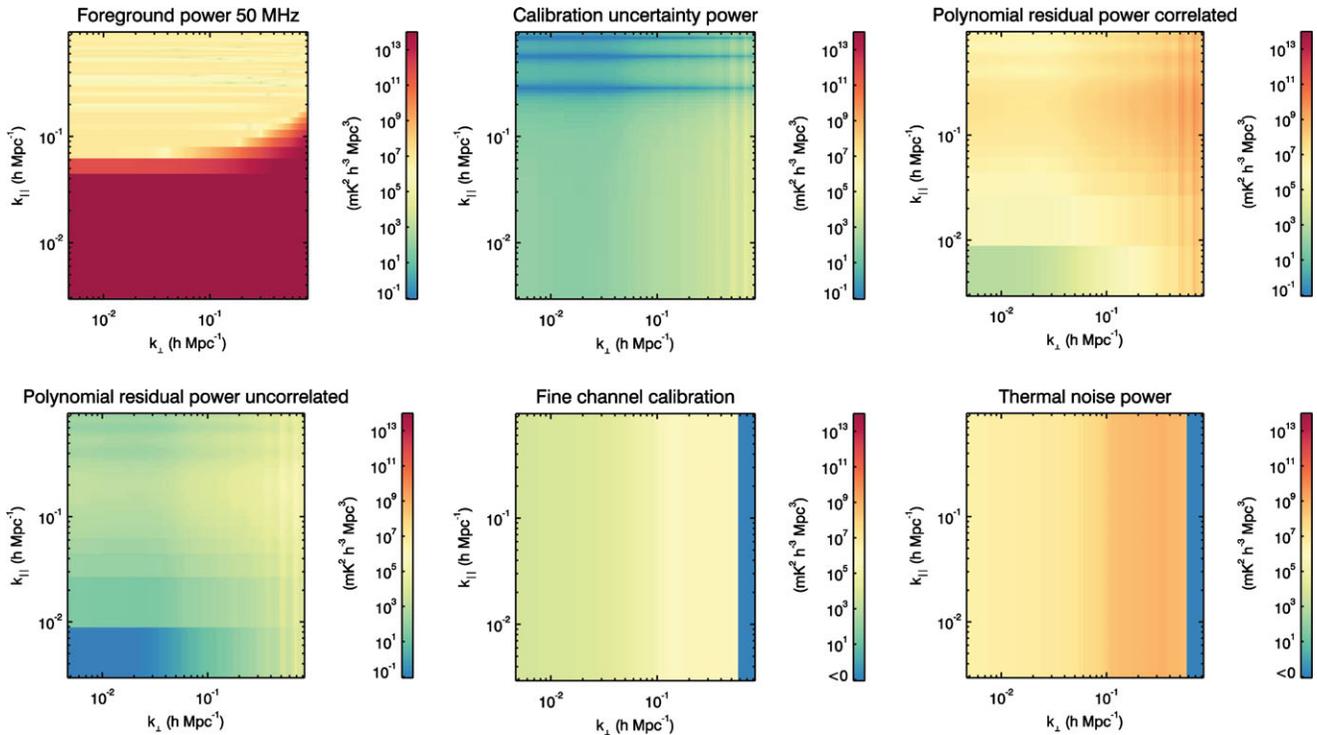
Figure 1. 50 MHz: (Top-left) Reference unresolved point source model, propagated into the power spectrum; (top-centre) the calibration uncertainty power for a third-order polynomial fit; (top-right) residual unresolved point source power in the most pessimistic calibration model (scaled to the tolerance level); (bottom-left) the residual unresolved point source power due to residual fourth-order curvature (most optimistic model); (bottom-centre) the uncertainty power due to calibration of each fine channel independently; (bottom-right) thermal noise power.

the instrument, with a frequency-dependent primary beam. At its core is a frequency–frequency covariance matrix that describes all the frequency correlations of signal in the data. It is this covariance matrix that can be used to approximate spectral bandpass features, and their effects in the EoR power spectrum. The full reference point source model described above provides the complete signal due to sources in the sky without any signal subtraction.

4.3 Calibration uncertainty

Both calibration schemes (polynomial and fine-channel fitting) are expected to have temporally uncorrelated uncertainties between calibration cycles, yielding a factor of $T_{obs}/T_{cal} = 1\,000\,\text{h}/600\,\text{s} = 6\,000$ reduction in calibration uncertainty power over the full experiment. It is often implicitly assumed that individual stations will have uncorrelated *errors*, yielding a stochastic reduction by a factor equivalent to the *uv*-sampling over the 4-h nightly track (we term this the 'optimistic' case). These are taken into account in the modelling. In particular, the estimation of parameters for each station uses visibilities for the whole array, yielding a factor of $(N_{ant} - 1)/2$ improvement in power. In the case where short baselines are omitted from the calibration (to avoid large-scale structure biasing the procedure), fewer baselines are available and the improvement will be reduced. For SKA1-

Low, there are many short baselines in the core, and within each of the superstations. Of the 158 000 baselines, >4 000 are shorter than 50 wavelengths at 150 MHz, effectively reducing the improvement from measurements from a factor of $(N_{ant} - 1)/2 \simeq 280$ to $\sim$270. This has, therefore, only a minor impact.

4.3.1 Polynomial fitting

One approach to bandpass calibration assumes that the bandpass may be fitted with a low-order polynomial function. This has the advantage of estimating only a handful of parameters over the bandwidth. In this work, we take a reasonable approach of fitting for each coarse channel independently, but using three contiguous coarse channels to perform the fit. The real and imaginary components are fitted as separate polynomials. We assume that a *n*th-order polynomial ($n = 2, 3, 4$) is fitted over these three coarse channels independently to each of the real and imaginary components of the visibilities, with $N_{fine} = 168$ fine channels of $\nu_{fine} = 4.578$ kHz within a coarse band ($\Delta\nu_{coarse} = 769.043$ kHz). Each fit therefore uses $3 \times 168 = 504$ datapoints, with 600 s of data and the full sky power, and there are two fits (the information for which is contained in twice as many datapoints when counting the real and imaginary components of the complex visibilities). Figure 4 shows this calibration scheme. For the third-order
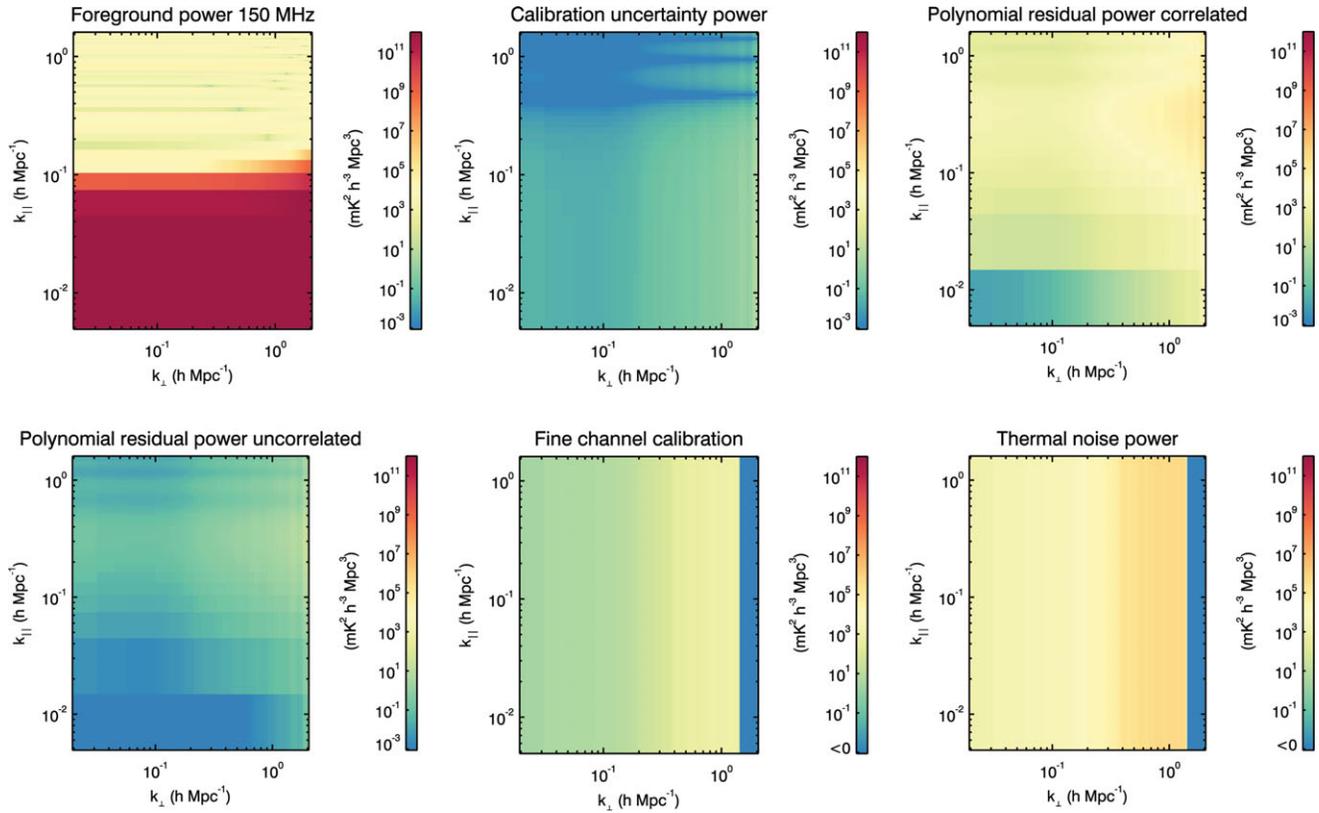
Figure 2. 150 MHz: (Top-left) Reference unresolved point source model, propagated into the power spectrum; (top-centre) the calibration uncertainty power for a third-order polynomial fit; (top-right) residual unresolved point source power in the most pessimistic calibration model (scaled to the tolerance level); (bottom-left) the residual unresolved point source power due to residual fourth-order curvature (most optimistic model); (bottom-centre) the uncertainty power due to calibration of each fine channel independently; (bottom-right) thermal noise power.

polynomial, we fit four parameters, A, B, C, D, such that

$$S_{r,i}(\nu) = A(\nu - \nu_0)^3 + B(\nu - \nu_0)^2 + C(\nu - \nu_0) + D, \qquad (7)$$

where r, i index real and imaginary, and ν_0 is defined as the centre of the coarse channel. The Fisher Information Matrix and the subsequent propagation of the correlated errors back into the signal, are independent of the actual values of A, B, C, and D (due to the fact that they are linear), thereby making this a general third-order model. The second- and fourth-order fits have one fewer or one more parameter. The amplitude, $\mathcal{A}$, of the calibration solution is given, as usual, by

$$\mathcal{A}(\nu) = \sqrt{S_r(\nu)^2 + S_i(\nu)^2}, \qquad (8)$$

and the phase, Φ, by

$$\Phi(\nu) = \operatorname{atan}\left(\frac{S_i}{S_r}\right). \qquad (9)$$

4.3.2 Fine-channel fitting

An alternative calibration method is to estimate the amplitude of each fine frequency channel independently. This has the advantage of removing any residual correlation between channels, but, requiring estimation of 168 parameters over a single coarse channel, requires sufficient information to be

available on the calibration timescale (i.e., sufficient signal to noise in 600 s to obtain a precise estimate). The consequent calibration precision will be lower for more parameters, but spectral correlations will be removed, and *no residual curvature will remain, except within a fine channel.*

4.4 Residual spectral power—amplitude residuals

The fitting of a low-order polynomial (order n) over three coarse channels leaves residual spectral signal of, at least, order $n + 1$ in each coarse channel. This residual signal is correlated across fine frequency channels, and hence generates residual power, with structure in the power spectrum parameter space. It is also subject to Runge's Effect, which occurs when fitting polynomials to regularly discretized data (see Section 4.6).

For each coarse channel for each of the real and imaginary components, we assume there remains a residual signal with unity mean (relative bandpass):

$$S_{r,i,\mathrm{res}}(\nu) = S_{\mathrm{mn}}\left(\frac{\nu - \nu_c - \xi_{r,i}}{\Delta\nu_{\mathrm{coarse}}}\right)^{n+1}, \qquad (10)$$

where ν_c and $\Delta\nu_{\mathrm{coarse}}$ are the frequency offset and coarse channel bandwidth, respectively, and S_{mn} is a normalisation
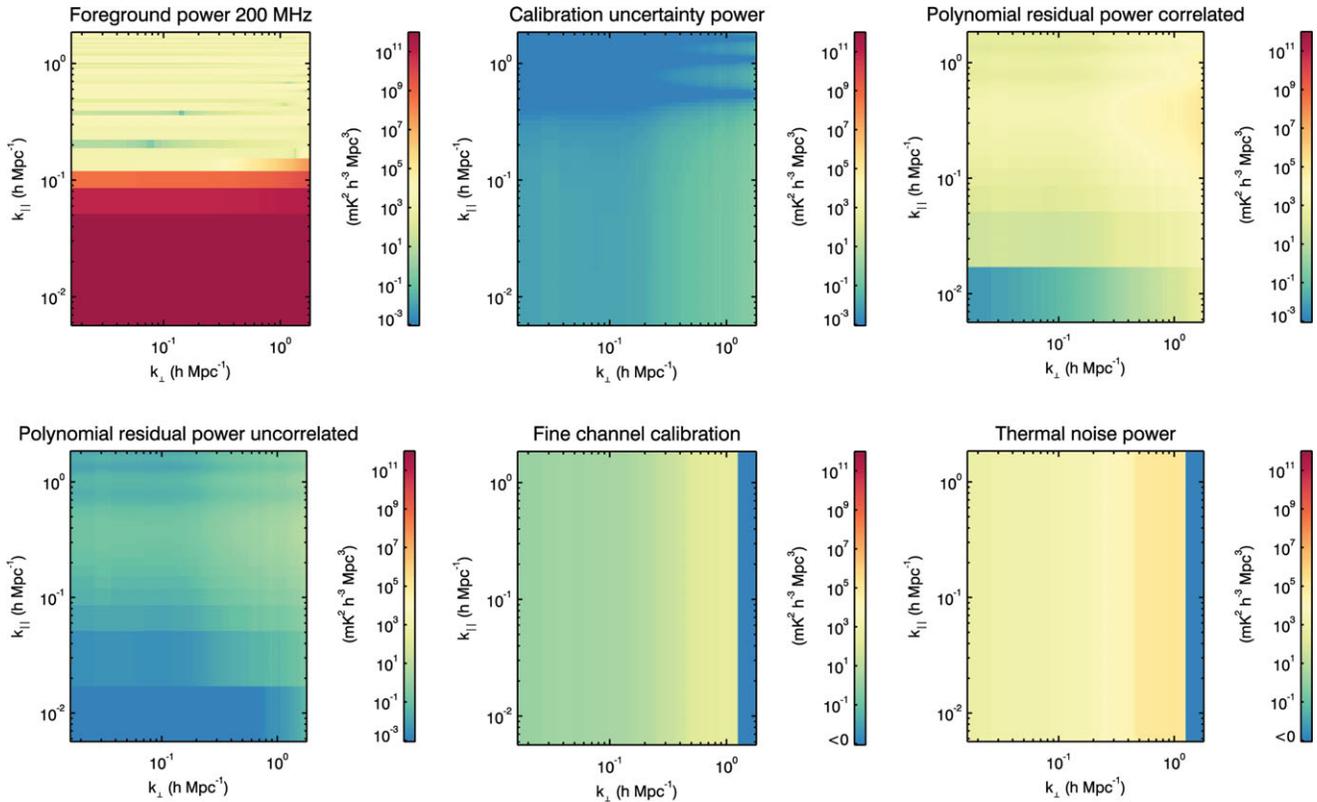
Figure 3. 200 MHz: (Top-left) Reference unresolved point source model, propagated into the power spectrum; (top-centre) the calibration uncertainty power for a third-order polynomial fit; (top-right) residual unresolved point source power in the most pessimistic calibration model (scaled to the tolerance level); (bottom-left) the residual unresolved point source power due to residual fourth-order curvature (most optimistic model); (bottom-centre) the uncertainty power due to calibration of each fine channel independently; (bottom-right) thermal noise power.

that sets the relative bandpass amplitude across the channel to equal unity. The offset $\xi_{r,i}$ translates the real and imaginary component residuals with respect to each other, and is a feature of a dipole antenna where the real and imaginary components are frequency dependent. The mismatch occurs when combining voltages from antennas with slightly different characteristics (e.g., due to wind load or temperature), such that a sharp feature arises where there is a rapid change in the voltage response. This form allows a spectral feature in the bandpass amplitude at ν_c, with an associated phase gradient with characteristic width $|\xi_r - \xi_i|$ (see Section 4.5). The residual signal, equation (10), is used to define the intra-channel correlation matrix, $\rho(\nu, \nu')$ (channels within each coarse channel are correlated, while inter-channel correlations are zero). This matrix is used, along with the sky point source signal model according to equation (1), to define the power in the 2D power spectrum due to the fractional residual signal in the bandpass.

We consider an optimistic and a pessimistic model to describe the temporal and inter-station correlation of residual spectral curvature:

- **Optimistic model:** The most optimistic model for the residual signal is when each station's bandpass shape is different, and varies on timescales smaller than the

calibration timescale. This leads to uncorrelated errors between stations and over time, yielding a stochastic signal that reduces with time.

- **Pessimistic model:** The pessimistic model assumes that the station bandpass spectral shapes are the same across all stations (as described by the model dipole bandpass), and vary slowly in time. This model assumes complete correlation of residual errors between all stations and over a full 4-h nightly track, thereby increasing the residual signal power by a factor of $(N_{\mathrm{ant}} \times t_{\mathrm{track}}/t_{\mathrm{calibration}})$ compared with the optimistic scenario.

In order to set the tolerance level, δ, for the $n + 1$-order polynomial residual, we scale the pessimistic model residual power until the amplitude of the thermal noise exceeds the residual across the full range of angular wavenumbers, $k_\perp$. *This defines the maximum fractional bandpass error over a coarse channel.*

4.5 Residual spectral power—phase residuals

Fitting of a nth-order polynomial to the real and imaginary components of the complex gain leaves at least $(n+1)$th-order residuals in both the components. For the third-order fit, the
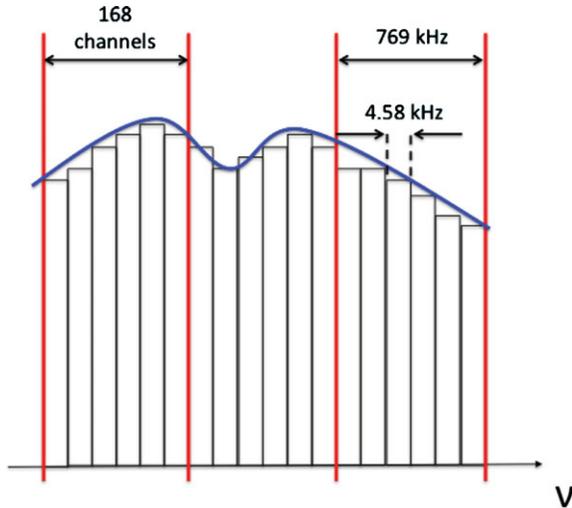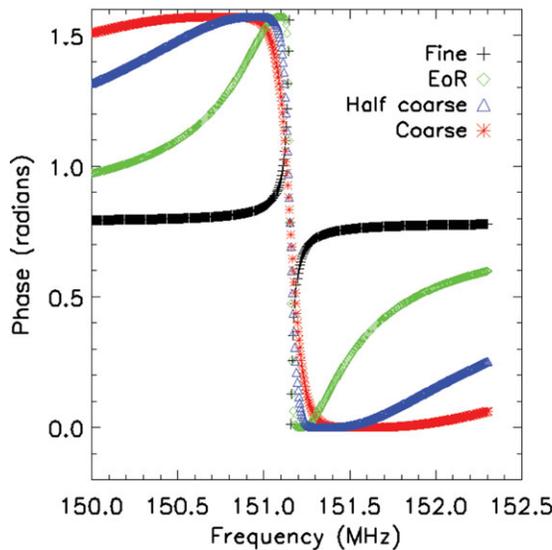
Figure 4. Schematic figure showing a smooth fit over three contiguous coarse channels, where the fit is performed on a fine-channel basis. The fitting parameters derived from these three coarse channels are used to calibrate the central channel only. Each vertical bar denotes a single fine channel, and the red lines denote coarse channel band edges.

mismatch between the real and imaginary components yields a residual phase, according to

$$\phi(\nu) = \mathrm{atan} \frac{A_i}{A_r} \frac{(\nu - \nu_i - \xi/2)^4}{(\nu - \nu_r + \xi/2)^4}, \tag{11}$$

where, $\xi = \xi_r - \xi_i$, and in the simplest case of equal amplitudes and turning points ($A_r = A_i$, $\nu_i = \nu_r$), has the form:

$$\phi(\nu) = \mathrm{atan} \frac{(\nu - \nu_i - \xi/2)^4}{(\nu - \nu_i + \xi/2)^4}. \tag{12}$$

Here, ξ is the frequency mismatch between the real and imaginary components. This form encapsulates the features of a dipole antenna where there is a frequency-dependent change in the real and imaginary components of the voltage response. Such features are observed in early testing of the SKALA antenna (de Lera Acedo et al. 2015).

Figure 5 (left) shows four representative phase plots as a function of frequency using equation (12), where the mismatch in frequency is equivalent to a single fine channel ($\xi = 4.578$ kHz; 'Fine'), 24 fine channels (an EoR spectral channel; $\xi = 109.87$ kHz; 'EoR'), one-half of a coarse channel ($\xi = 384.6$ kHz; 'Half coarse'), and a single coarse channel ($\xi = 769.10$ kHz; 'Coarse'). Each symbol denotes a single fine channel. The smaller the mismatch ξ, the steeper the phase feature. However, when the data are averaged to the EoR spectral resolution, mismatches finer than the channel resolution, 109.87 kHz, are smoothed.

We average each fine channel to the EoR spectral resolution, before forming the frequency–frequency correlation matrix, and propagating into the power spectrum. The errors are computed for a single coarse channel, with the remainder of the band showing no phase residuals. This therefore corresponds to the best-case scenario.

In phase, the correlation matrix contains complex phase terms that imprint phase structure. Figure 5 (right) plots the ratio of power for each value of ξ, compared with a flat profile (zero phase), $\frac{|P_\xi|}{P_{\mathrm{flat}}}$, for a chosen cut through the 2D parameter space. The different phase structure imprints features at harmonics of the coarse bandpass, convolved with the spectral shape of the phase error. The phase errors *bias* the cosmological signal by re-distributing power in LOS modes, but do not contribute power because the power spectrum
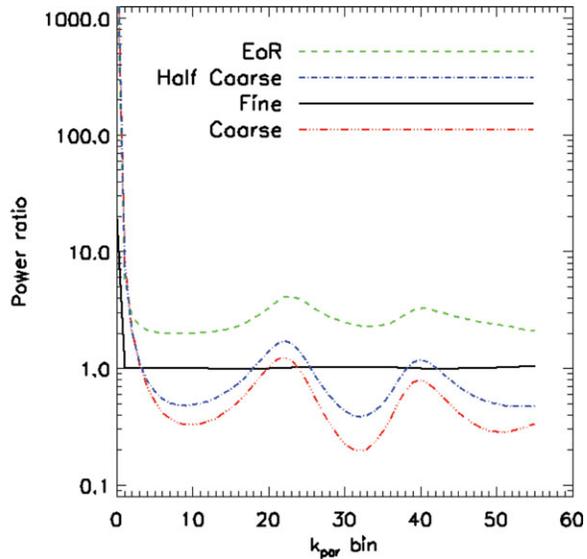

Figure 5. (Left) Phase plots as a function of frequency using equation (12), where the mismatch in frequency is equivalent to a single fine channel (4.578 kHz; 'Fine'), 24 fine channels (an EoR spectral channel; 109.87 kHz; 'EoR'), one-half of a coarse channel (384.6 kHz; 'Half coarse'), and a single coarse channel (769.10 kHz; 'Coarse'). (Right) Ratio of power to a flat phase profile, $\frac{|P_\xi|}{P_{\mathrm{flat}}}$, for a single value of $k_\perp$, and four values of the frequency mismatch, ξ.
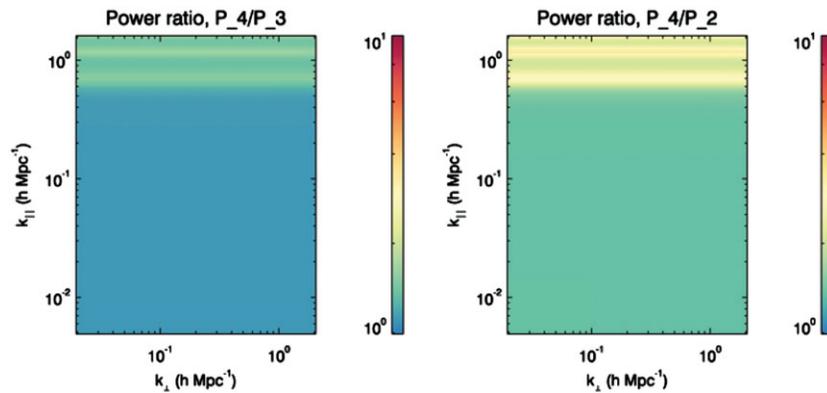
Figure 6. Ratio of residual power after polynomial fitting, where the ratios compare a fourth-order fit to a third-order fit (left), and a fourth-order fit to a second-order fit (right). Runge's phenomenon is visible at high $k_\parallel$ where fitting residuals at the box edges are larger for higher order fits.

natively destroys phase information, by squaring the complex visibilities.

4.6 Results

Figure 1–3 display the 2D power spectra for each frequency, for the following: (top-left) the reference unresolved point source model; (top-centre) the calibration uncertainty power for a third-order polynomial fit; (top-right) the residual unresolved point source power due to residual fourth-order curvature (pessimistic model, scaled to reach thermal noise level); (bottom-left) the residual unresolved point source power due to residual fourth-order curvature (optimistic model); (bottom-centre) the uncertainty power due to calibration of each fine channel independently; and (bottom-right) the thermal noise power.

The strategy to calibrate each fine channel independently leads to no residual curvature (above the fine-channel level) and no correlations, but is less precise than estimating the polynomial parameters because more independent parameters needs to be fitted, and therefore there are fewer available degrees of freedom. For bandpass responses that are smooth, this strategy costs a larger amount than the polynomial fitting, in the optimistic case.

In each experiment, the calibration uncertainty power due to fitting a third-order polynomial over three contiguous coarse channels yields structured power in the power spectrum space. This power is less than the thermal noise, but has the potential to cause low-level bias in the derived science results due to its shape. This is particularly true at low frequencies where the calibration uncertainty power is comparable to the thermal noise. The regular, horizontal lines of increased power in $k_\parallel$ are due to the correlation of uncertainties across a given coarse channel, but with independence between channels. This leads to a block-diagonal structure to the correlation matrix, and the features observed in the power spectrum. It is also a consequence of Runge's phenomenon

(e.g., see Dahlquist & Bjorck 2003), whereby fitting a polynomial to regularly space datapoints across a box leads to high-frequency residuals at the box edges. These residuals increase in amplitude for higher order polynomial fits (for the same data spacing, as is the case here). This effect is visible in both the calibration uncertainty power (order n fitting) and the residual power (order $n+1$ fitting). We can further demonstrate this by computing the ratio of residual power for different polynomial orders. Figure 6 displays the ratio of the power in the fourth-order fit to the second- and third-order fits. The additional power at high frequency (high $k_\parallel$) is visible for the higher order fits.

This structure in $k_\parallel$ has the potential to yield bias in the results. As compared to the uncertainties (additional power) due to limited information, as computed by the Cramer–Rao bound, bias can add signal power that may be mistaken for cosmological power, in a statistical estimate. Despite potentially yielding a better fit to the data, the additional structured power from higher order polynomials may be more problematic than using a fine-channel fitting procedure, where the residual power is higher, but is flat across the parameter space. This choice is subject to the experiment itself, where the power at the location of the structures in $k_\parallel$ relative to the cosmological signal power will determine the best approach (i.e., if the cosmological signal is expected to peak in the same region where the residual power gradient is large, then the bias will be increased). This is model-, bandwidth-, and redshift-dependent for a given telescope.

Table 2 lists the relative fractional bandpass error tolerable in a $(n+1)$th-order polynomial residual, as applied to the plots, where δ is the fractional reduction in power values. Each tolerance is approximately 1%, but higher order fits have more stringent constraints due to the additional structured power at high $k_\parallel$. Note that these results are appropriate when the underlying shape is well fitted by an $(n+1)$th-order polynomial, such that the amplitude of
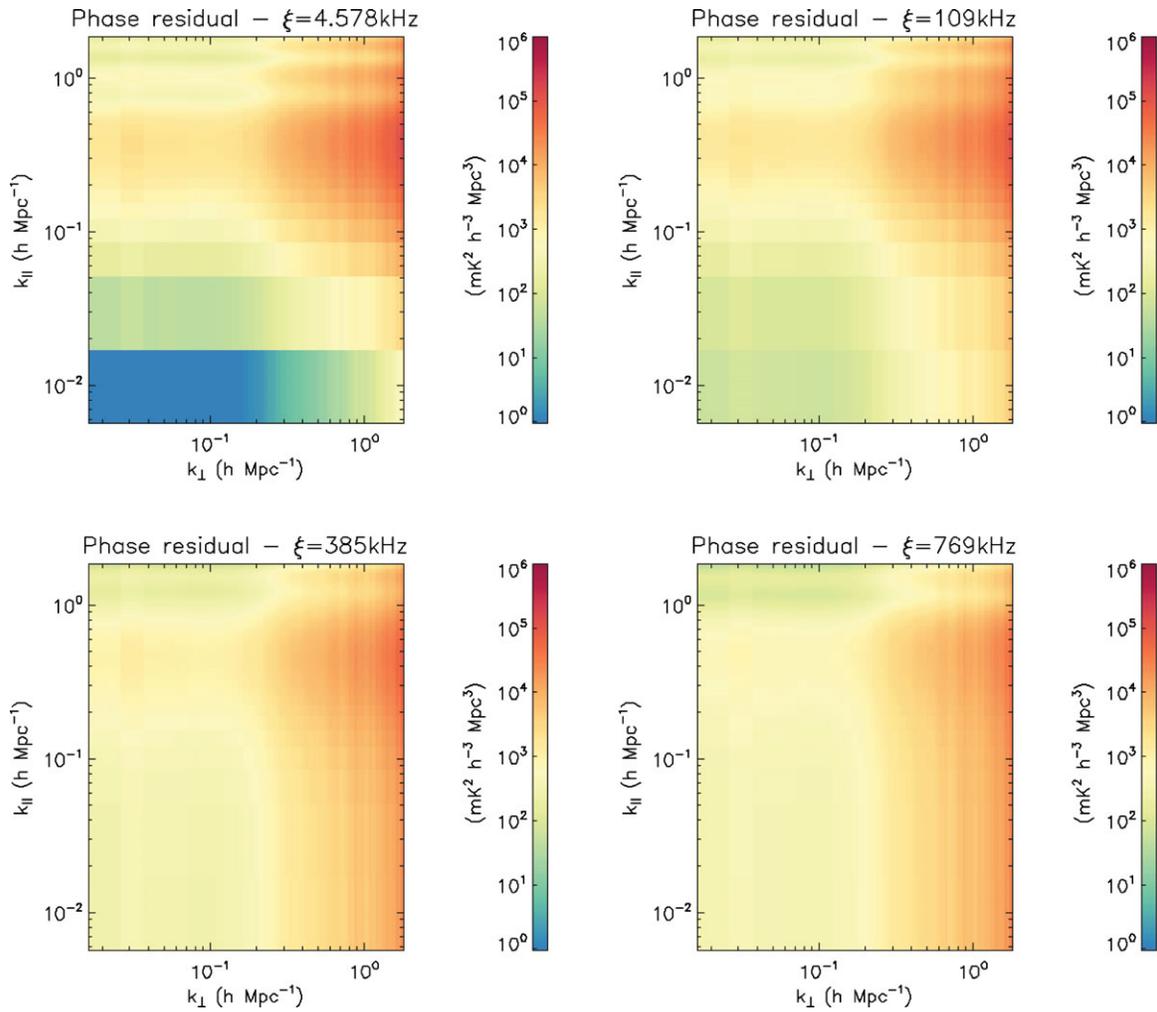
Figure 7. 2D power from phase residuals on sky point sources at 200 MHz, scaled by the amplitude tolerance, $\delta=0.008$, for four values of the frequency mismatch, ξ. The overall scale of power is not markedly changed, but the power is distributed differently in $k_\parallel$. (a) Fine. (b) EoR. (c) Half a coarse channel. (d) Coarse.

Table 2. Derived tolerances for each experiment such that the residual power due to $(n+1)$th-order curvature in the bandpass is less than the thermal noise.

Experiment	$\delta_{n=2}$	$\delta_{n=3}$	$\delta_{n=4}$
50 MHz	0.027	0.025	0.019
100 MHz	0.011	0.010	0.008
150 MHz	0.006	0.005	0.004
200 MHz	0.009	0.008	0.006

terms with $> (n + 1)$th-order are small compared with the $(n + 1)$th order.

The impact of phase residuals described in Section 4.5 is plotted in the 2D power spectrum in Figure 7. Here, the same fractional reduction in power values, δ, as described in Table 2 is applied to the sky point source model, in or-

der to scale the uncertainty power to that deemed tolerable from amplitude residuals. Frequency mismatches below the EoR channel averaging bandpass (109.8 kHz) are smoothed, while much larger mismatches imprint a slower gradient structure. The largest differential is obtained for a frequency mismatch of the same scale as the spectral resolution (4.578 kHz), while the power distribution is most different from flat for a mismatch corresponding to the coarse channel bandwidth.

When the amplitude and phase residuals are combined, the phase errors introduce a small power bias difference in the low k-modes. Figure 8 displays the spherically averaged (1D) power from the amplitude and phase residuals, for each of four different values for ξ. Smoother transitions in the spectrum (larger ξ) yield larger bias at low k, with factors of two reduction in power. There is, therefore, a bias in the power. The phase gradient tolerance derived from this work is weak, due to the implicit loss of phase information in the
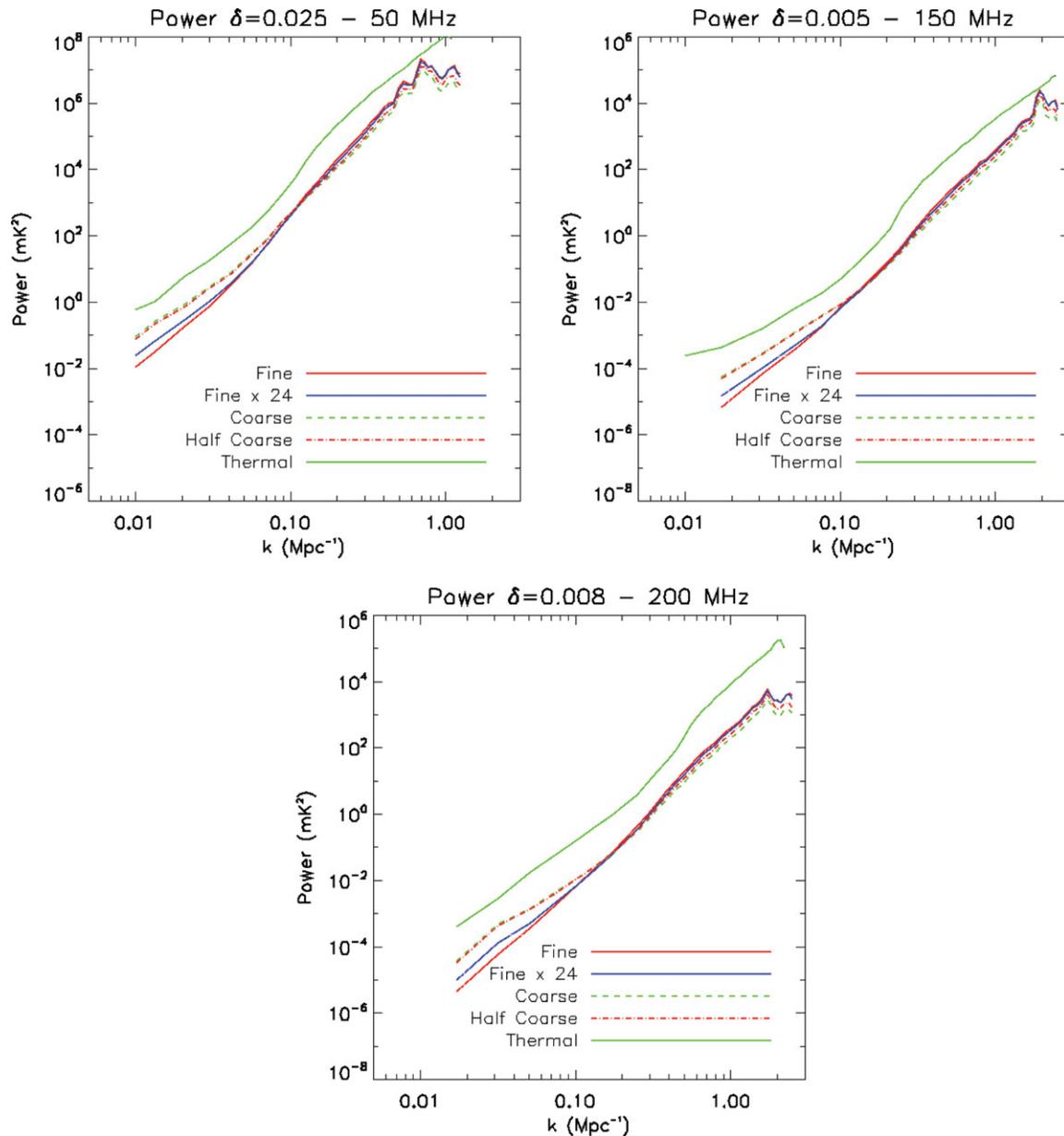
Figure 8. 1D (spherically averaged) power spectra for each value of the frequency mismatch, ξ, displaying the power bias introduced by a given combination of δ and ξ. Also shown is the thermal noise level. (a) 50 MHz. (b) 150 MHz. (c) 200 MHz.

power spectrum upon squaring of the complex-valued data. A much more stringent gradient is set by the tomographic experiment, where phase is retained, and this is explored in the next section. As was shown for the 2D case in Figure 6, the structures at high $k_\parallel$ yield biases at large k.

5 TOMOGRAPHY

The tomographic experiment is a new addition to the EoR observational arsenal, not previously accessible with lower sensitivity instruments. It is a challenging experiment that uses deep image cubes to detect brightness temperature fluc-

tuations corresponding to ionised structures (e.g., bubbles) in the CD and EoR. It is therefore a direct detection experiment, rather than a statistical experiment, necessitating high image dynamic range and precision calibration. We assess phase residual impact at 150 MHz, and at a wavenumber, k, where the array filling factor is close to unity.

There are several approaches to assess the image dynamic range due to phase residuals, with a thorough treatment and rules-of-thumb presented in Perley (1999). Here, we wish to consider the impact of a residual phase gradient across a 100-kHz channel, with uncorrelated gradients between stations. To perform this analysis, we consider the maximum phase

change across a fine channel. We define the phase within a fine channel for station α to be

$$\phi_\alpha(\nu) = \epsilon_\alpha \nu, \tag{13}$$

with a visibility at ν measured between two stations of

$$V'_{\alpha\beta}(\nu) = V_{\alpha\beta}(\nu) \exp\left(-2\pi i(\epsilon_\alpha - \epsilon_\beta)\nu\right). \tag{14}$$

Integration of a signal across a channel leads to signal decorrelation according to the integral of this expression, such that the visibility degradation is given by

$$\frac{V'}{V} = \frac{1}{2\pi\Delta\epsilon\Delta\nu}\left[\sin(2\pi\Delta\epsilon\Delta\nu) + i(\cos(2\pi\Delta\epsilon\Delta\nu) - 1)\right], \tag{15}$$

where $\Delta\epsilon$ and $\Delta\nu$ are the phase gradient differences between stations and channel width, respectively. One can immediately observe that the real part of this expression is a sinc function, which is <1 when $\Delta\epsilon \neq 0$. This expression describes the fractional flux density loss of a coherent signal due to decorrelation across the bandwidth of a channel $\Delta\nu$.

We can now extend this analysis for a single visibility to explore the impact in an image slice, by defining the phase gradient for each station to be a random value, which is Gaussian distributed with characteristic width $\sigma_\epsilon = \epsilon$. We proceed by forming the *uv*-sampling in a snapshot at zenith, with unity-amplitude visibilities corrupted according to equation (15). We then Fourier transform these gridded, corrupted visibilities to the image plane, and compare the peak of the point spread function (PSF) to that for a gridded set of unity-amplitude visibilities (i.e., the instrument sampling function, leading to the instrumental PSF). We define the tolerance to be the characteristic phase gradient, σ_ϵ, where the reduction in dynamic range causes the residual from the brightest image source to exceed the expected bubble signal strength. A phase residual of 1 radian across a channel leads to almost complete decorrelation of signal.

5.1 Results

All sampled *uv* points in the core are used to estimate the snapshot PSF of the array, matching the expected bubble size to the array resolution. At 150 MHz, the longer baselines provide resolution at the 10-s arcseconds level, while the densely packed core provides a wider PSF base of $\sim$10 arcmin. We compute the characteristic phase gradient, σ_ϵ on each station bandpass fine channel such that the dynamic range degradation destroys the cosmological signal. We take a pessimistic approach (where the phase residual on each channel is fixed over the full experiment) and an optimistic approach (where the phase residual is uncorrelated in time between calibration cycles, yielding an increase in dynamic range and weaker constraints).

On scales of the SKA1-Low core (20 arcmin), a 1-mK brightness temperature fluctuation corresponds to a $\sim$0.02-mJy source, increasing to 0.2 mJy for degree scales (Koop-

Table 3. Characteristic phase residual across a fine channel, $\sigma_\epsilon \Delta\nu$, for a loss in dynamic range that would destroy the cosmological signal, due to phase decorrelation.

Scale	$\sigma_\epsilon \Delta\nu$ (degrees) Pessimistic	$\sigma_\epsilon \Delta\nu$ (degrees) Optimistic
20′	0.04	0.2
60′=1°.	0.2	1.3

mans et al. 2015). For a 5-Jy source in the field, these levels corresponds to a 0.02/5 000=10^{-5} (0.2/5 000=10^{-4}) fractional residual flux density (noise term in the image) across the array in a 4.58-kHz channel for the pessimistic case. Loss in dynamic range of this amplitude would erase any high-redshift signal, and sets the tolerance level. Any 1-degree bubble is statistically likely to have a 1-Jy source within its synthesized beam, and will therefore be affected by this effect. Table 3 shows the derived parameters for each experiment. The constraints are therefore stringent in either the pessimistic or optimistic case for bubbles on tens of arcminute scales.

The tolerance allows for a phase gradient that yields a phase difference that does not exceed a residual of 0.2 (1.3) degrees per fine channel in the pessimistic (optimistic) case, on scales of 1°. This calculation assumes that the phase residual is uncorrelated between two different antennas.

6 SUMMARY

This work takes a signal estimation theoretic approach to calibration of low-frequency radio interferometers. Specifically, it computes the tolerances on bandpass amplitude and phase residuals for SKA-Low to achieve the goals of the EoR and CD experiments, using calibration precision and accuracy. The additional power due to imprecise and inaccurate bandpass calibration in the power spectrum is defined to not exceed the thermal noise power expected for the experiment. Similarly, the imaging dynamic range is defined to not destroy brightness temperature fluctuation detections for the tomography experiment. These tolerances provide quantitative specifications for the intrinsic spectral response of SKA-Low antennas. The choice of calibration approach depends on the shape of the antenna bandpass, whereby the requirement to fit higher order polynomials may lead to increased signal power bias. In this case, a fine-channel fitting model, while yielding larger uncertainties across the parameter space, also yields a smooth response in frequency, thereby producing a potentially less-biased solution.

REFERENCES

Barry, N., Hazelton, B., Sullivan, I., Morales, M. F., & Pober, J. C. 2016, preprint (arXiv:1603.00607)

Bowman, J. D., et al. 2013, PASA, 30, 31

Dahlquist, G., & Bjorck, A. 2003, Numerical Methods (New York: Dover Publications)

de Lera Acedo, E., Razavi-Ghods, N., Troop, N., Drought, N., & Faulkner, A. J. 2015, ExA, 39, 567

Gervasi, M., Tartari, A., Zannoni, M., Boella, G., & Sironi, G. 2008, ApJ, 682, 223

Grainge, K. 2014, SKA Memo 151

Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., & Lal, D. V. 2011, A&A, 535, A38

Kay, S. M. 1993, Fundamentals of Statistical Signal Processing: Estimation Theory (New Jersey: Prentice-Hall)

Kazemi, S., Yatawatta, S., & Zaroubi, S. 2013, MNRAS, 430, 1457

Koopmans, L., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 1

Mitchell, D. A., Greenhill, L. J., Wayth, R. B., Sault, R. J., Lonsdale, C. J., Cappallo, R. J., Morales, M. F., & Ord, S. M. 2008, ISTSP, 2, 707–717

Mort, B., Dulwich, F., Razavi-Ghods, N., de Lera Acedo, E., & Grainge, K. 2016, preprint (arXiv 1602.01805)

Offringa, A. R., et al. 2016, MNRAS, 458, 1057

Parsons, A. R., et al. 2010, AJ, 139, 1468

Perley, R. A. 1999, in ASP Conf. Ser., Vol. 180, Synthesis Imaging in Radio Astronomy II, eds. G. B. Taylor, C. L. Carilli, & R. A. Perley (San Francisco: Astron. Soc. Pac.), 275

Sutinjo, A. T., et al. 2015, ITAP, 63, 5433

Tasse, C., van der Tol, S., van Zwieten, J., van Diepen, G., & Bhatnagar, S. 2013, A&A, 553, A105

Trott, C. M., et al. 2016, ApJ, 818, 139

Trott, C. M., & Tingay, S. J. 2015, ApJ, 814, 27

van Haarlem, M. P., et al. 2013, A&A, 556, A2

Yatawatta, S. 2015, MNRAS, 449, 4506

Yatawatta, S., Zaroubi, S., de Bruyn, G., Koopmans, L., & Noordam, J. 2008, preprint (arXiv 0810.5751)