

Building Models for Extended Radio Sources: Implications for Epoch of Reionisation Science

Cathryn M. Trott^{1,2,3,4} and Randall B. Wayth^{1,2,3}

¹International Centre for Radio Astronomy Research, Curtin University, Bentley 6845, Australia

²ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO), Curtin University, Bentley 6845, Australia

³ARC Centre of Excellence for All-Sky Astrophysics in 3 Dimensions (ASTRO 3D), Curtin University, Bentley 6845, Australia

⁴Email: cathryn.trott@curtin.edu.au

(RECEIVED August 24, 2017; ACCEPTED October 31, 2017)

Abstract

We test the hypothesis that limitations in the sky model used to calibrate an interferometric radio telescope, where the model contains extended radio sources, will generate bias in the Epoch of Reionisation power spectrum. The information contained in a calibration model about the spatial and spectral structure of an extended source is incomplete because a radio telescope cannot sample all Fourier components. Application of an incomplete sky model to calibration of Epoch of Reionisation data will imprint residual error in the data, which propagates forward to the Epoch of Reionisation power spectrum. This limited information is studied in the context of current and future planned instruments and surveys at Epoch of Reionisation frequencies, such as the Murchison Widefield Array (MWA), Giant Metrewave Radio Telescope and the Square Kilometre Array (SKA1-Low). For the MWA Epoch of Reionisation experiment, we find that both the additional short baseline uv -coverage of the compact Epoch of Reionisation array, and the additional long baselines provided by TGSS and planned MWA expansions, are required to obtain sufficient information on all relevant scales. For SKA1-Low, arrays with maximum baselines of 49 km and 65 km yield comparable performance at 50 MHz and 150 MHz, while 39 km, 14 km, and 4 km arrays yield degraded performance.

Keywords: (cosmology:) dark ages, reionization, first stars – instrumentation: interferometers – methods: observational – telescopes

1 INTRODUCTION

A sky model for an interferometer telescope is used for calibration, source deconvolution, and source subtraction. The sky model can be obtained: (1) externally, from prior observations with existing telescopes at the same frequency; (2) internally, via measurement with telescope itself; (3) with a combination of these. For existing experiments, the current suite of sky surveys provide the input sky models for calibration and source subtraction. These can be augmented with further surveys from the same, or upgraded versions, of existing telescopes. Complete and accurate sky models are crucial for calibrating data, and subtracting unwanted sources from the dataset. This is particularly important for Epoch of Reionisation (EoR) experiments, which aim to extract a weak signal from bright foreground contamination (Jacobs et al. 2016; Carroll et al. 2016; Line et al. 2017; Barry et al. 2016).

For the southern sky, the current suite of low-frequency sky surveys, used for building a sky model, include: the 74 MHz

Very Large Array Low Frequency Sky Survey redux (Lane et al. 2012), the MWA Commissioning Survey (MWACS, Hurley-Walker et al. 2014), MWA GLEAM (Wayth et al. 2015; Hurley-Walker et al. 2017), and GMRT TGSS (Intema et al. 2017). Cross-matching tools, such as the Positional Update and Matching Algorithm (PUMA, Line et al. 2017), combine these spatially and spectrally to provide the calibration sky model. Procopio et al. (2017) explored the impact of adding GMRT TGSS information to the calibration and source subtraction model for MWA EoR in the EoR1 field, finding that the additional double-source and extended source information was important for reducing bias. An extension of that work explores the direct impact of imprecisely-modelled extended sources on the MWA EoR experiment, and we undertake that work here.

Source modelling is typically performed for physical insight into the source itself (e.g., spectral structure of radio lobes to understand their energetics) and often includes multi-wavelength information to constrain physical models. Such studies typically rely on image-plane maps (e.g., Perley,

Dreher, & Cowan 1984; Salter et al. 1989; Castelletti et al. 2007; Braun 2013; McKinley et al. 2015; Procopio et al. 2017, and references therein). Perley et al. (1984), for example, studies the multi-frequency spatial structure of Cygnus A. The image data are represented on a discretised grid at a set of frequencies, and this forms a representation of the underlying information, some of which has been lost. This is a commensurability problem, whereby continuous data are represented discretely. Braun (2013), for example, then uses the spatial power in these frequency slices to estimate the structural properties of the source. A subset of the full data has therefore been used for the analysis.

Image reconstruction and analysis tools are designed with minimal information loss considerations, but these are imperfect. Performing analysis in the measurement plane offers the greatest ability to preserve information, but can be computationally and algorithmically challenging compared with the image plane. In this work, we consider the *full* information content of the data, taking this as the best possible outcome (most optimistic), while appreciating that using a segmented image-plane representation will likely degrade the result. In Section 2, we explore the impact of discretisation more formally.

The SKA aims to be the world's foremost radio telescope, and aims to make major scientific advances across a range of programs. The EoR and Cosmic Dawn (CD) experiment (Koopmans et al. 2015) is envisaged to be one of the most challenging, and is one of the SKA High Priority Science Objectives. Prior to its construction and commissioning, existing and past facilities are providing the sky models for the current generation of southern hemisphere EoR experiments, such as the MWA (Tingay et al. 2013), the Precision Array for Probing the Epoch of Reionization (PAPER)¹ (Parsons et al. 2010), and Hydrogen Epoch of Reionization Array (HERA)² (DeBoer et al. 2017). Compared with the SKA, current low frequency sky surveys have poorer sensitivity and spatial resolution, limiting the ability of these surveys to accurately measure the spatial and spectral structure of complex sources.

For SKA1-Low, with its long baselines and good snapshot *uv*-coverage, we expect that SKA1-Low itself will form the primary sky model. Therefore, the sky model will be well-sampled on modes measured by SKA, and not well-measured on scales that are not. We note that, in general, the application to calibration of an incomplete sky model formed exclusively by the same instrument being calibrated, will likely produce 're-substitution bias', where the performance of a calibration procedure is over-estimated (similar to the re-use of a training set for the real dataset in machine learning).

In this paper, we discuss the impact of the current and future suite of sky surveys at EoR frequencies to enable EoR and CD science, with particular reference to the EoR

and CD power spectrum. In the current era, we study the availability of sufficient information from existing and imminent sky surveys to measure the values for the parameters describing the spatial and spectral structure of extended sources. For the future, we discuss the implications for the planned SKA1-Low array configuration to execute its ambitious EoR/CD program. We use data information content, as the basis for quantitatively assessing the performance of the instrument under a set of defined designs. We begin with the original MWA Phase I design, and subsequently add TGSS, and the additional baselines of the upgraded MWA, and show whether the EoR science experiment is biased by the incomplete information available to correctly represent extended sources. Attention is then turned to the future SKA1-Low array, and the more challenging EoR/CD experiments proposed. We use these results to inform the smallest array maximum baseline required to execute these experiments.

2 METHODOLOGY

Calibration of a radio interferometer requires estimation of the unknown complex gain parameters of each station. It typically relies on the fitting of data to a model of the sky signal plus instrument, allowing freedom in the complex gain parameters to perform the least-squares fit. Successful calibration therefore requires good knowledge of the received sky signal and the instrument (station locations, sky response, etc.). This is partially true for purely redundant arrays (e.g., PAPER, HERA), where sky information is used for initial calibration estimates and breaking degeneracies (e.g., Omnicol; Ali et al. 2015).

In this work, the Fisher Information is used to quantify the ability of a given array configuration to estimate the sky model parameters for a generalised extended source, in the presence of a realistic background of point sources. We use the residual model uncertainties as a measure of the error (uncertainty) in the model, and propagate these uncertainties into visibilities measured by the telescope for EoR science. We further propagate these uncertainties into the EoR power spectrum, yielding a measure of the error (power bias) due to the incomplete source model. In studying different SKA array configurations, we focus on removal of outer stations to reduce the maximum baseline, but do not re-locate stations to the array core. For the generalised extended source, we are careful to model structure on all scales of relevance for the SKA, in order to robustly and generally assess calibration performance (scales from the size of the PSF to the FOV). We apply the same model throughout, for the SKA and also the MWA and TGSS-derived results. This is to ensure consistency, and to test the ability of precursor instruments to form a sky model of relevance for the future SKA. For the MWA and TGSS, the existence of power on scales much finer than those available to their baselines, will test whether unmeasured power on small scales impacts the estimation of longer wavemodes, of relevance to the EoR.

¹ <http://eor.berkeley.edu>

² <http://reionization.org>

2.1. Approach

We use the Fisher Information, and then the Cramer-Rao Bound (Kay 1993, CRB) to quantify the information available in our calibration data to estimate the values of the spectral and spatial parameters describing a generalised extended source, embedded within measurement noise, and a realistic sky of extragalactic point sources. The calibration model is obtained from the sky survey used to construct the sky model.

The Fisher Information computes the amount of information a given dataset (with a particular probability distribution function of noise, here modelled to include radiometric and background sources) contains about the values of a model parameter, for a pre-defined source model. In general, measurements that vary rapidly with a varying parameter value have the ability to estimate its value precisely. In contrast, no change of the expected measurement with a varying parameter value means that there is no information available to estimate that parameter. The CRB takes into account degeneracies between model parameters and correlations in the data, and represents the estimation performance of an ideal estimation algorithm.

For complex-valued data, embedded within generalised Gaussian noise with covariance, \mathbf{C} , and an expected signal vector $\vec{\mu} = \vec{\mu}(\vec{\theta})$ with parameters $\vec{\theta}$, the Fisher Information Matrix has the following elements:

$$[I]_{ab} = \left(\frac{\partial \vec{\mu}}{\partial \theta_a} \right)^\dagger \mathbf{C}^{-1} \left(\frac{\partial \vec{\mu}}{\partial \theta_b} \right), \quad (1)$$

where ab are two elements of the parameter vector. In general, the data covariance can also be a function of unknown parameters, but here we assume we have full knowledge of the per-visibility properties of the radiometric noise:

$${}_{th}\mathbf{C}(u, v; v, v') = \left(\frac{2kT_{\text{sys}}}{A_{\text{eff}}} \right)^2 \frac{1}{\Delta v \Delta t} \delta(v - v') \text{Jy}^2, \quad (2)$$

and the point source covariance (Trott et al. 2016) as a function of spectral channels v, v' and Fourier mode $\vec{u} = u, v$:

$${}_{fg}\mathbf{C}(\vec{u}; v, v') = \frac{\alpha}{3 - \beta} \frac{S_{\text{max}}^{3-\beta} \pi c^2 \epsilon^2}{S_0^{-\beta} D^2} \frac{1}{v^2 + v'^2} \times \exp\left(\frac{-|\vec{u}|^2 c^2 f(v)^2 \epsilon^2}{4(v^2 + v'^2) D^2} \right) \text{Jy}^2, \quad (3)$$

where $\epsilon = 0.42$ converts an Airy disk to a Gaussian characteristic width, D is the station diameter, and $f(v) = (v - v')/v_0$. The point source model is represented by a broken power-law, characterised by parameters α, β , such that in a sky area $d\mathbf{l}$:

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

$$= \alpha \left(\frac{v}{v_0} \right)^\gamma \left(\frac{S_{\text{Jy}}}{S_0} \right)^{-\beta} dS d\mathbf{l}. \quad (5)$$

We use values of $\alpha = 4100 \text{Jy}^{-1} \text{sr}^{-1}$, $\beta = 1.59$, and $\gamma = -0.8$ at 150 MHz (Intema et al. 2011). We assume that our sky model is formed from 15 min (1 h) of data for GMRT (MWA),

reducing the radiometric noise component with respect to the confusion, and for spectral resolution commensurate with the EoR experiment (100 kHz). (The GMRT TGSS survey used 15-min per field, split over 3–5 pointings.) Observations are assumed to be of a field at declination -27° , centred on RA=0.

The CRB yields the uncertainties, and correlations, for each parameter value. If we assume that these uncertainties are then embedded within the sky model for that extended source, we can propagate these errors into the science data measured by the telescope (the measured visibility dataset). This is achieved using a standard Jacobian and the Fisher Information, I , such that

$$\mathbf{C}_V(u, v; v) = \mathbf{J}^\dagger \mathbf{I}^{-1} \mathbf{J}, \quad (6)$$

where \mathbf{J} is the matrix of derivatives of the visibility measured at u, v , and channel v . We further propagate from measured visibilities to the EoR power spectrum:

$$\Delta P(k_\perp, k_\parallel) = (\mathcal{F}_v^\dagger \mathcal{W}^\dagger \mathbf{C}_V \mathcal{W} \mathcal{F}_v) \delta(k_\parallel - k'_\parallel, k_\perp^2 - u^2 - v^2), \quad (7)$$

where \mathcal{F}_v is the Fourier Transform operator along the spectral direction, \mathcal{W} is a spectral taper function that aims to reduce spectral leakage, and the delta-function extracts the variance estimates (power) from the covariance matrix as well as identifying the perpendicular k modes with the L_2 -norm of the angular Fourier modes ($k_\perp^2 = u^2 + v^2$). Herein, we employ a Blackman taper.

Therefore, we can assess the error introduced into our science data by an incomplete extended source model produced from a given array configuration.

Fourier Transform of the data to the image plane, and discretisation into surface brightness pixels, transforms the visibility covariance matrix (equation 6) into a degraded image covariance, such that

$$\mathbf{C}_i(l_i, m_i; v) = \mathcal{D}_i^\dagger \mathcal{F}^\dagger \mathbf{J}^\dagger \mathbf{I}^{-1} \mathbf{J} \mathcal{F} \mathcal{D}_i, \quad (8)$$

where \mathcal{D}_i and \mathcal{F} denote the discretisation operator (continuous-to-discrete) and spatial Fourier operator (discrete-to-continuous), respectively, and the Jacobians act on the data used to estimate the source properties. Inadequate discretisation sampling compared with scales and shapes of the underlying source components, will degrade the quality of the estimation, and this can be quantified by studying the mean-squared-error (MSE) between the actual underlying source structure and the discretised representation. This can be seen most simply by considering a Fourier Transform back to the visibility plane (to estimate scale sizes), where a discrete-to-discrete transform couples pixel properties into the data covariance matrix. At this point, both information loss and bias in the estimates, are possible.

2.2. Generalised extended source model

We want to form the most general extended source, in order to provide a fair basis for comparison. We aim to build a model with multi-scale structure that encases the spatial

Table 1. Parameter values as a function of scale, k_i . $\mathcal{N}(\mu, \text{var})$ denotes a Gaussian-distributed random number with mean μ and standard deviation $\sqrt{\text{var}}$.

Parameter	Value (scale i)
a_i (Jy/bm)	$1/\sqrt{1+i}$
l_i (arcmin)	$\mathcal{N}(0, \Delta l^2 = 10^2)$
m_i (arcmin)	$\mathcal{N}(0, \Delta m^2 = 10^2)$
σ_i (rad)	$1/k_i$
γ_i	$\mathcal{N}(-0.8, \Delta \gamma^2 = 0.02^2)$

modes accessible to the Baseline Design SKA1-Low, corresponding to scales of the synthesized beam (10 arcsec at 150 MHz) to a fraction of the field-of-view (0.08°). We build a model with a complete set of angular scales, spaced evenly in the range $k = [570\text{--}17\,000] \text{ rad}^{-1}$, and construct a source based on the summation over a series of Gaussians, each with five parameters: central location l_i, m_i (rad), peak brightness S_i (Jy/beam for SKA-Low beam at 150 MHz), characteristic scale σ_i (rad), and spectral index γ_i . This corresponds to 57 angular scales ($k = [570\text{--}17\,000] \text{ rad}^{-1}$, sampling evenly at a spacing of half of the lowest k -mode), yielding 285 total parameters. The parameters for each scale are generated via a Gaussian-distributed random sampling, as described in Table 1, and fixed thereafter for each array configuration. The Fourier-space expectation of the signal (μ in Equation 1) is given by

$$\mu(\vec{u}, \nu; \vec{\theta}) = \sqrt{2\pi} \sum_{i=1}^{57} a_i \sigma_i^2 \left(\frac{\nu}{\nu_0} \right)^{\gamma_i} \times \exp(-2\pi i \vec{u} \cdot \vec{l}_i) \exp(-2|u|^2 \pi^2 \sigma_i^2), \quad (9)$$

where the two exponentials encode the Fourier kernel, and the multi-scale Gaussian, respectively, and $\vec{\theta} = [a, l, m, \sigma, \gamma]$.

Figure 1 shows an image of the complete model. Note that although a base source peak brightness of 1 Jy/beam has been chosen for the source, the amplitude here is not relevant to the final error introduced into the visibilities. This is because brighter sources can be estimated more precisely, and weaker sources less precisely, with a linear scaling with peak brightness. The propagation of error back into the visibilities also scales linearly with peak brightness, and these scalings cancel (Trott, Wayth, & Tingay 2012). Nonetheless, the peak brightness refers to an SKA-Low beam at 150 MHz, and remains consistent for all models. Therefore, each extended source with these angular scales being estimated for the sky model will contribute this error.

2.3. Extended sources within the field

The procedure described above yields the power bias from a single extended source in the sky model for a given field. If we now consider the total number of extended sources in a given field, we can estimate the total power bias in the EoR power

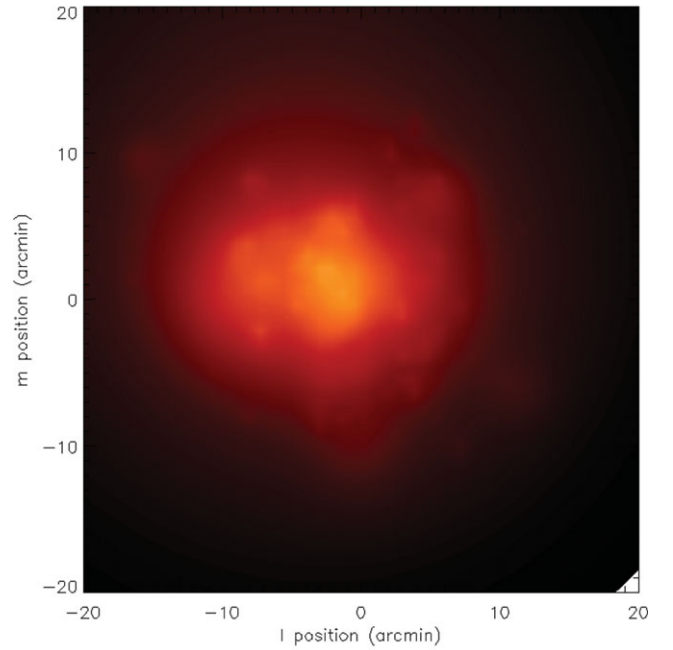


Figure 1. Image of the extended source produced using the model described (lowest spectral channel).

spectrum to be $N_{\text{src}} \Delta P(k_{\perp}, k_{\parallel})$, where N_{src} is the number in the field. This implies that the extended sources considered here all have the same power spectrum of spatial and spectral structure, although each individual source will have a different realisation of the parameter values. This ensures we are considering a consistent set of sources.

There is no complete low-frequency census of extended sources, due to the limited sensitivity and spatial resolution under study in this work. Results from the high-resolution FIRST (1.4 GHz) (Becker et al. 1994) and TGSS catalogues (Intema et al. 2017) can be used to estimate the number of sources per unit area of a given angular scale and flux density. Procopio et al. (2017) showed that the brightest sources were most important to model because these tended to be closer and therefore more extended.

We use the study of Windhorst, Mathis, & Neuschaefer (1990) at 1.4 GHz to estimate the 150 MHz and 50 MHz distributions, assuming a spectral index of -0.8 between the bands and no structural evolution. They find that the fraction of sources larger than size ψ (arcsec) at flux density, S (mJy), can be described by

$$h(\psi) = \exp[-\ln 2(\psi/\psi_m)^{0.62}], \quad (10)$$

where $\psi_m = 2.0 S_{1.4}^{0.3}$ arcsec is the median size at 1.4 GHz. We scale the flux densities to 150 MHz to find the corresponding median size for low frequencies. Figure 2 shows this fraction as a function of flux density and size. Strictly this distribution is applicable for flux densities below 1 Jy: the extended, close radio galaxies (e.g., Fornax A, Centaurus A, Pictor A), which have much larger flux densities, have their own individual spatial and spectral distributions. Assuming a point source

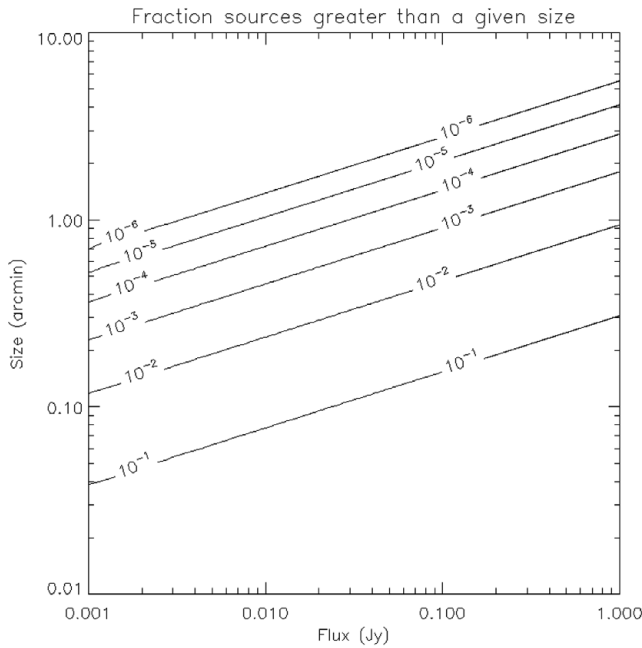


Figure 2. Fraction of sources with angular extent greater than a given size, as a function of source flux density, for 150 MHz (adapted from Windhorst et al. 1990).

number count distribution, a field-of-view and this fraction, the number of sources larger than a given angular size and greater than some flux density limit can be approximated as

$$N(> \psi, S_0) = \int_{S_0}^{\infty} \frac{dN}{dS} dS \Omega h(\psi) = \int_{S_0}^{\infty} \alpha S^{-\beta} \Omega h(\psi), \quad (11)$$

where $\alpha = 4000(\nu/150)^{-0.8} \text{ Jy}^{-1} \text{ sr}^{-1}$, $\beta = \{1.59 (S < 1 \text{ Jy}), 2.5 (S > 1 \text{ Jy})\}$, parametrise the number count distribution (Intema et al. 2011), and Ω is the field-of-view (steradians). Figure 3 shows the associated contour plots for the MWA 150 MHz, SKA 150 MHz, and SKA 50 MHz experiments. At the flux density confusion limits of these experiments, we would expect ~ 1 source in each of the experiments, of scale larger than ~ 1 arcmin. Most of the weaker sources have angular extents of 10 s of arcseconds, and these have fewer critical parameters for estimation. Given these estimates, we focus on the bright, distributed nearby radio galaxies in this work, and consider a single source of importance in the field.

2.4. Telescope arrays

2.4.1. Murchison widefield array

The first three years of MWA EoR observations form the basis for all published limits to date, and use the Phase I MWA configuration with 128 tiles. The GLEAM survey uses the same array for its observations. The GMRT TGSS survey is a re-processing of the GMRT 150 MHz sky survey with new methods for instrument and ionospheric calibration (Intema et al. 2017). The additional sensitivity of the GMRT, and the longer baselines, provides higher spatial resolution,

but the small number of dishes (30) limits the additional uv -coverage. The upgraded MWA will provide additional surface brightness sensitivity and spatial resolution with its 256 tiles. The zenith-pointed uv -coverage for these are displayed in Figure 4 (left). GMRT clearly adds information at high spatial resolution, but even with rotation synthesis, the coverage is sparse and the larger angular scales are not significantly improved.

GMRT has different system temperature, latitude, and dish sensitivity (effective area) compared with the MWA, and these are taken into account in the analysis. Earth Rotation synthesis is considered when computing the uv -coverage, assuming an EoR field-of-interest located at declination -27° (MWA latitude). The GMRT TGSS survey parameters (15 min integration time per field) are used as the basis for the sky model formed from these different hybrid arrays.

2.4.2. SKA1-Low configurations

We trial four array configurations, which are all subsets of the Baseline Distribution. Table 2 describes the number of stations, and maximum baseline, for each. Figure 4 (right) shows the zenith snapshot uv -coverage of each at 150 MHz. We also considered a 4 km baseline array, which effectively corresponds to an array with EoR-science scales. Additionally, we consider a variation on the Baseline Design, whereby the inner clusters of six stations are unpacked for improved instantaneous imaging performance (Jones et al. 2016).

The station effective areas, and sky temperature as a function of frequency, are taken from the SKA1-Low system description (L0 Requirements). We compute the EoR/CD power spectrum error at two prime frequencies of interest (150 MHz, $z = 8.6$; 50 MHz, $z = 27$), and compare to typical expected 21 cm cosmological power spectra, from 21cmFAST simulations (Mesinger, Furlanetto, & Cen 2011). Standard conversions are undertaken from $\text{Jy}^2 \text{ Hz}^2$ to $\text{mK}^2 \text{ h}^{-3} \text{ Mpc}^3$ (Morales & Hewitt 2004).

3 RESULTS

We form the power biases due to extended sources requiring estimation for formation of the sky model, and compare these with expected 21 cm power spectra obtained from 21cmFASTv2 (reionisation via faint galaxies).

3.1. MWA

The MWA EoR experiment aims to detect the 21 cm signal through a power spectrum at redshifts, $z = 6.5 - 9.0$ ($\nu = 137-197$ MHz). We take a nominal lower frequency of 150 MHz and estimate the power bias due to the three hybrid arrays. Figure 5 displays the signal-to-power bias ratio (SNR).

Addition of the small scales from TGSS does not provide a substantial improvement in performance on EoR scales, and this hybrid array and the original MWA 128-tile array display contrast ratios of order 0.1 in the lowest portion of the

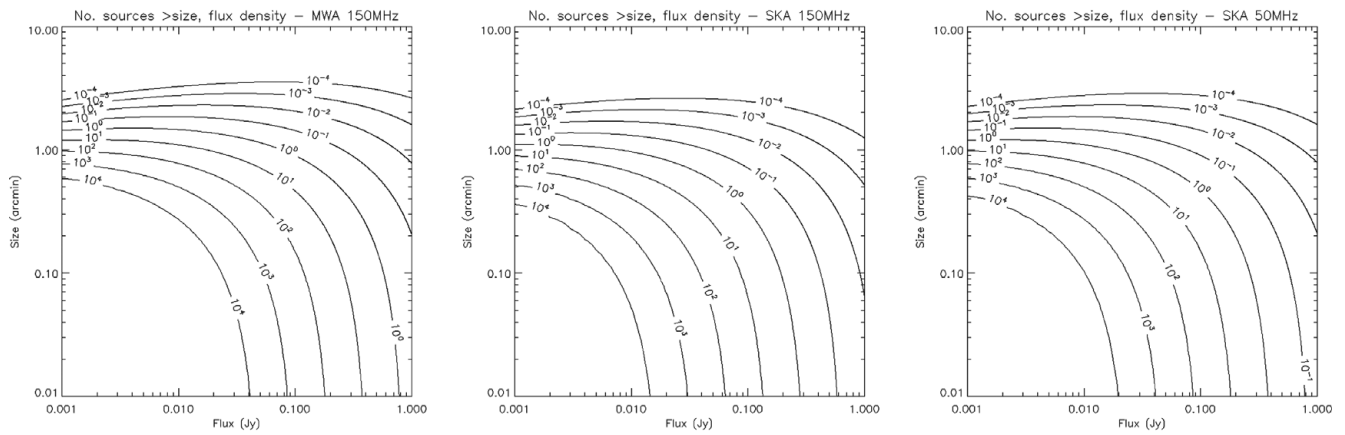


Figure 3. Number of sources with angular extent and flux density greater than given values for the MWA 150 MHz (left), SKA 150 MHz (centre), and SKA 50 MHz (right) experiments.

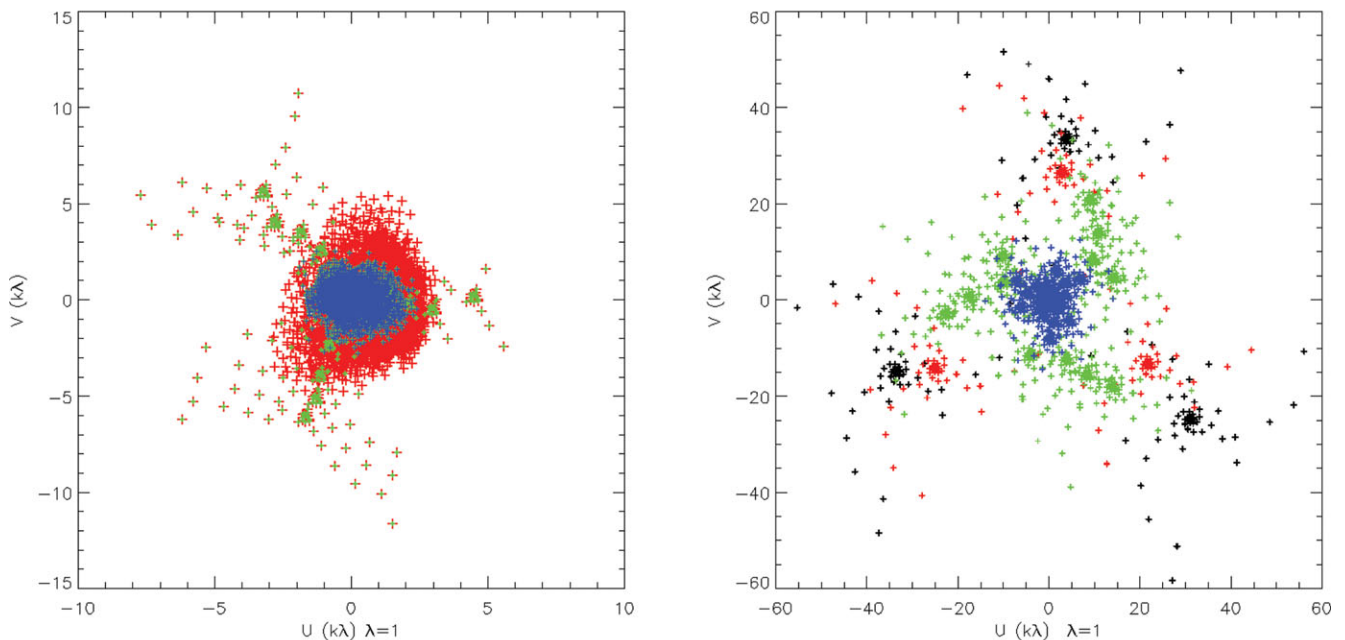


Figure 4. (Left) uv coverage for a zenith snapshot pointing for the three MWA arrays: MWA Phase I (blue), +GMRT (green), MWA Phase III + GMRT (red). (Right) uv coverage for a zenith snapshot pointing for the four SKA1 arrays considered: black (max. baseline 65 km), red (49 km), green (39 km), blue (14 km), ($\lambda = 1$ m).

Table 2. The four arrays considered.

N_{stn}	Max. baseline (km)
512	65
494	49
476	39
400	14

EoR window. Extending the array to 256 tiles and including TGSS increases the uv -coverage on all scales, leading to an improvement and contrast ratios exceeding unity across the EoR window.

PASA, 34, e061 (2017)
doi:10.1017/pasa.2017.57

The improvement in parameter estimation can be explored to understand the relative importance of uv -coverage and maximum baseline in measuring the source parameters. Figure 6 display results for the five source parameters as a function of angular scale of the feature (σ). Each plot shows the ratio of estimation performance (square-root of CRB) for the extended MWA (256 tiles + TGSS) relative to the original MWA128 and MWA128 + TGSS. Also displayed are histograms of the u (solid) and v (dashed) distributions for each array.

It is clear that the additional long north-south baselines for the MWA256 + TGSS offer improved performance on small scales for estimating the m -position of each source feature.

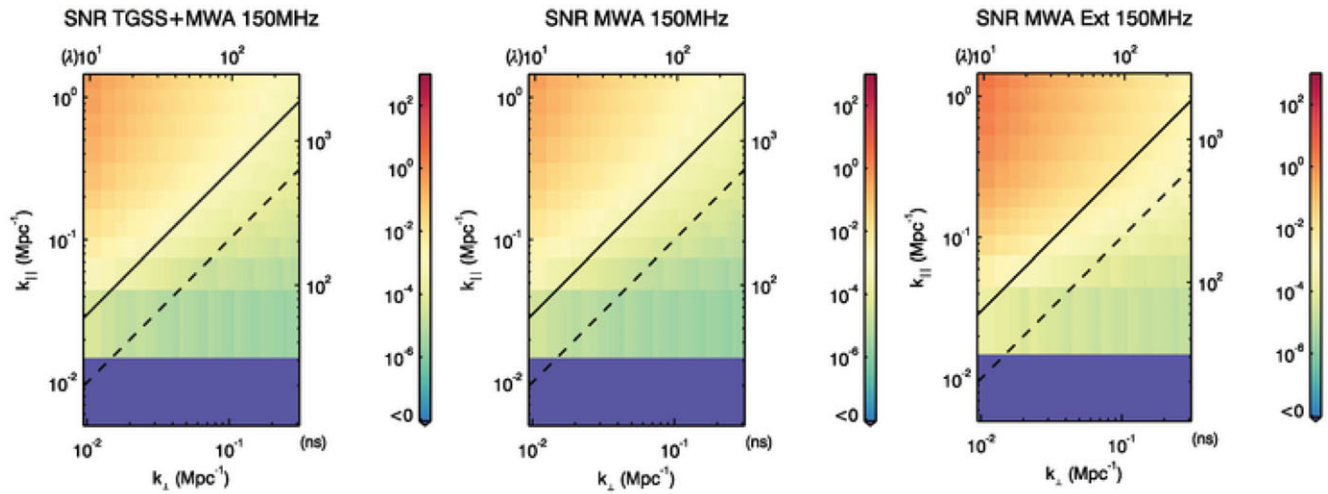


Figure 5. Signal-to-noise (contrast) ratios of a typical 21 cm cosmological signal to the power error introduced by one extended source in the field, for the three MWA-based hybrid array configurations considered, and a 10 MHz bandwidth experiment centred at 150 MHz ($z = 8.6$).

It is also clear that the increased number of short baselines from the MWA256 hexagonal arrays improves estimation of large scale features. Addition of the TGSS long baselines aids the precision. The clear and intuitive conclusion is that for sources with information on a range of angular scales, both long baselines and good short baseline coverage are required in both dimensions. Particularly for EoR scales, excellent short baseline uv -coverage is crucial.

3.2. SKA1-Low

Figures 7 and 8 display the signal-to-noise ratios (power contrast ratios) for the four arrays considered. The dashed and solid black lines denote the first sidelobe and horizon limits for the expected leakage of foreground power (the ‘wedge’). At 150 MHz, all four arrays yield $\text{SNR} > 10$ across most of the EoR Window (area outside of the horizon wedge). The Blackman taper performs well to suppress foreground leakage, but at the expense of a broader DC term, and residual error leaks into the EoR window at $k_{\perp} \approx 0.02$, $k_{\parallel} \approx 0.1 \text{ Mpc}^{-1}$ for all arrays. The relative strength of the 21 cm signal to the sky temperature yields acceptable performance for the 49 km and 65 km arrays, while degradation is evident for 39 km and 14 km. *Notably, the 49 km and 65 km arrays yield comparable results (power ratio ≈ 1.004).*

At 50 MHz, the system temperature is higher, foregrounds are brighter, and the signal is weaker. Therefore, the performance is degraded for all arrays, relative to 150 MHz. Both the 14 km and 39 km arrays yield a low SNR detection across a large region of the EoR window, while 49 km and 65 km yield good performance ($\text{SNR} > 10^3$). Again, the 49 km and 65 km arrays yield acceptable performance (high contrast ratios).

The final metric of interest for 21 cm studies, where the cosmological signal is expected to be isotropic, is the spherically-averaged (1D) power spectrum. To remove the bulk of the

foreground extended source bias, we consider line-of-sight scales larger than $k_{\parallel} = 0.1$ (smaller spatial scales). In doing so, foregrounds are reduced when averaging spherically, but we also lose spatial modes. Figure 9 displays these 1D profiles, obtained directly from the power on each measured baseline at each frequency (i.e., not obtained from 2D, but considering the original 3D distribution). The same conclusions can be drawn about the relative merits of each array at the lowest frequency (50 MHz). The different structural properties stem from the interaction of the Blackman–Nuttall spectral taper and each array’s spectral covariance profiles. Fundamentally, it is the progression in k -mode of the additional leaked power as the arrays become more compact, that is of concern. Here, we also provide results from the 39 km maximum baseline, but using the v7 proposed array of Jones et al. (2016) with unpacked clusters (green dashed). Here, the addition of uv information on intermediate scales provides some improvement relative to the Baseline Design, and can recover information lost from cutting from 49 km to 39 km maximum baselines.

It is important to understand where the differences in the arrays are contributing to the differences in the performance. Intuitively, sources that are extended, but have power on angular scales smaller than that for the EoR, will not contribute power bias into the EoR parameter space. Similarly, sources that have power on large angular scales, but no power on small scales, contribute to EoR baselines, but each of the 39 km, 49 km, and 65 km arrays sample these scales equally. Therefore, we hypothesise that it is sources that have both large and small-scale structures that couple inability to measure small scales into the estimation of the larger scales. To test this hypothesis, we perform the same analysis at 50 MHz with two additional sources: one source with power only above 1 arcmin scales (broad source), and one source with power only on scales of 10 arcsec–36 arcsec (compact source). When comparing the 39 km and 65 km arrays for their power bias

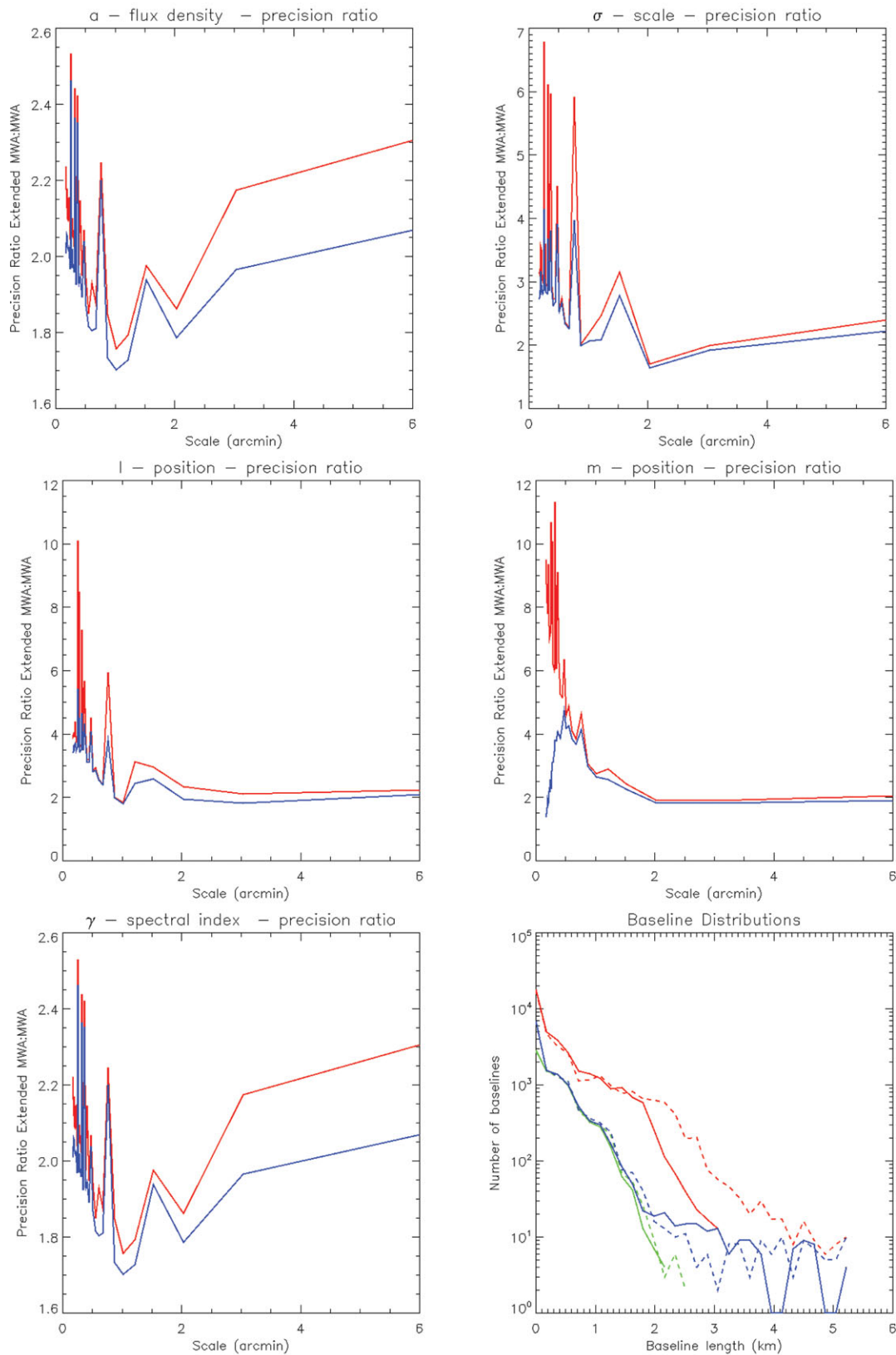


Figure 6. (Top, middle, bottom left) Ratio of estimation performance (precision) for extended MWA (256 tiles + TGSS) relative to the original MWA128 (red) and MWA128 + TGSS (blue), as a function of scale of source feature. (Bottom right) Histograms of baseline distributions in u (solid) and v (dashed) directions for MWA256 + TGSS (red), MWA128 + TGSS (blue), and MWA128 (green).

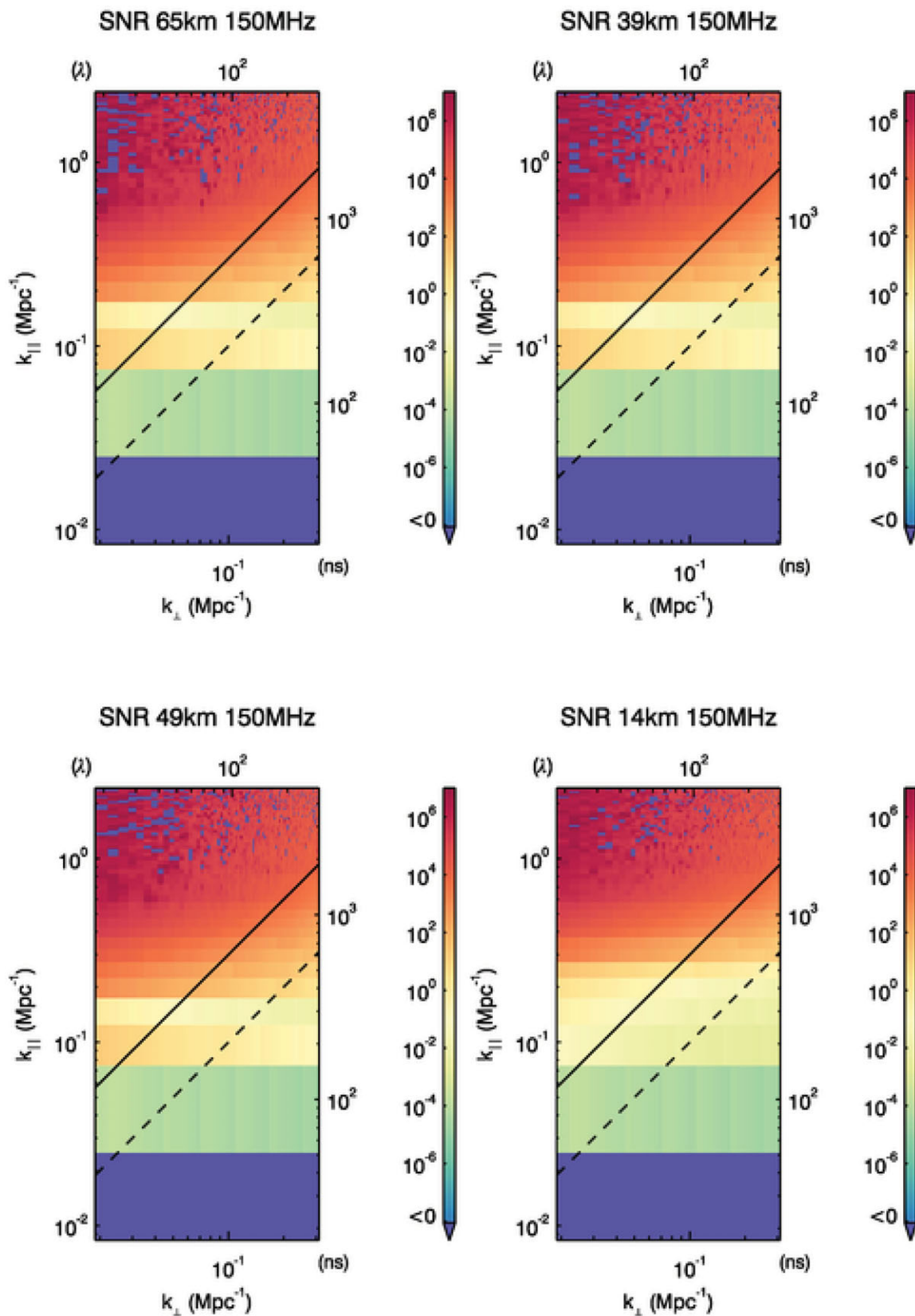


Figure 7. Signal-to-noise (contrast) ratios of a typical 21 cm cosmological signal to the power error introduced by one extended source in the field, for the four array configurations considered, and a 10 MHz bandwidth experiment centred at 150 MHz ($z = 8.6$).

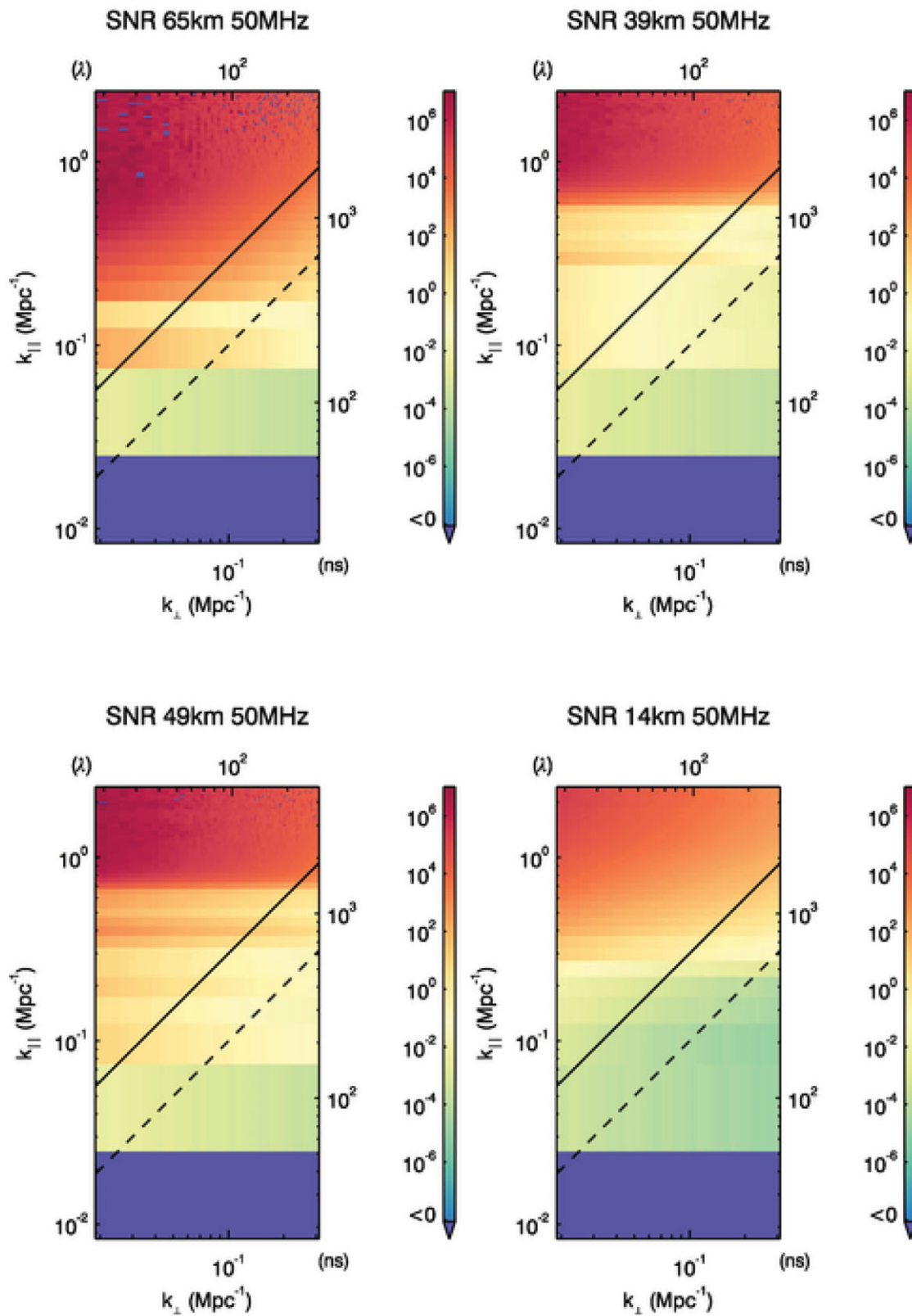


Figure 8. Signal-to-noise (contrast) ratios of a typical 21 cm cosmological signal to the power error introduced by one extended source in the field, for the four array configurations considered, and a 10 MHz bandwidth experiment centred at 50 MHz ($z = 27$).

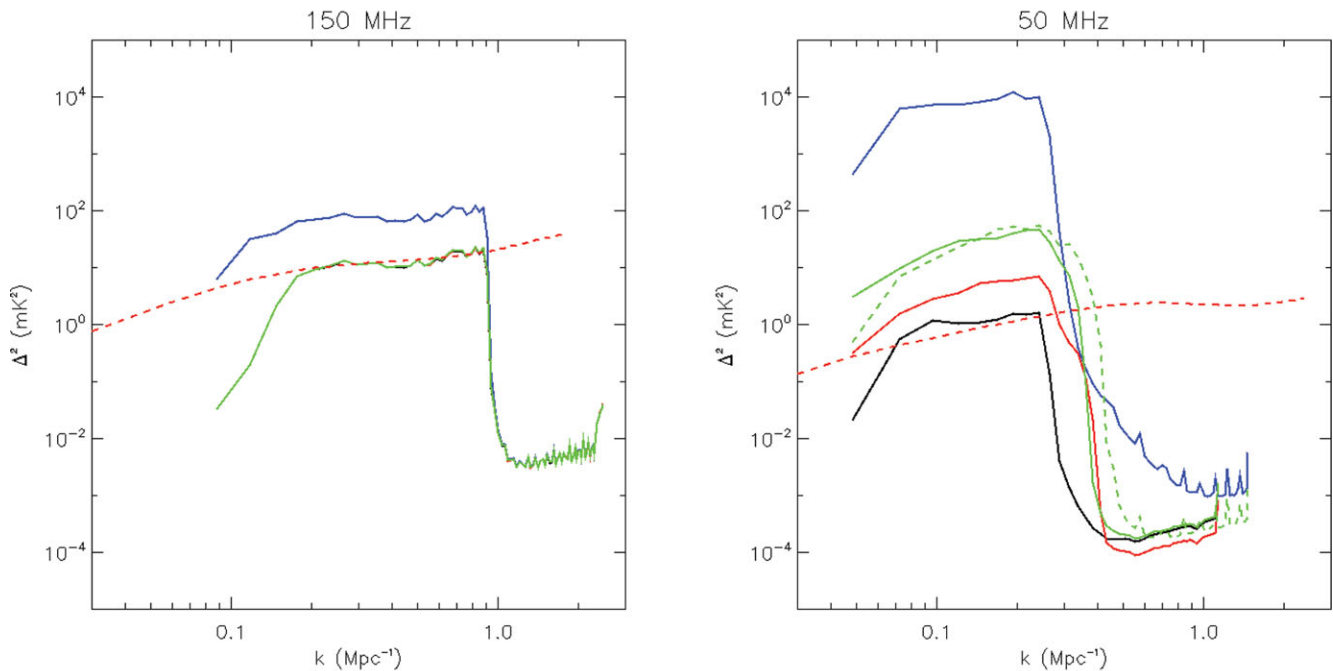


Figure 9. (Left) 150 MHz spherically-averaged power spectrum and simulated 21cmFAST 21 cm power for comparison (dashed). 14 km (blue), 39 km (green), 49 km (red), 65 km (black). Note that the black, red and green are overlapping. (Right) Same but for 50 MHz, including an alternative array configuration with 39 km baselines (green, dashed).

in the EoR power spectrum, we find that both yield comparable performance (power ratio approximately unity). This supports the hypothesis that sources that have multi-scale power are those that show differences in estimation precision for the different arrays.

3.3. Discussion and interpretation

The key message of this work is that a sky model generated from a given array, and subsequently used as ‘truth’ (the reference model for a source for calibration and subtraction), has inherent errors, due limitations of that array, and that those errors will propagate forward into science data products. Here, we considered a generalised multi-scale Gaussian source (extended source), and used mock SKA1-Low and existing and upcoming MWA + TGSS array configurations to estimate their model parameters, in the presence of a realistic sky with structured sidelobes from other sky sources. We applied these errors to model EoR data, and quantified the bias in power in the EoR 2D and 1D power spectra due to these incomplete source models. Compared with previous work in this area, the use of a generalised extended source model, the inclusion of other confusing sources in the field (classical and sidelobe confusion), and the direct equating of modes measured by a survey to input calibration model ‘truth’ parameters, expands on existing studies.

For the MWA EoR experiment, we find that both the additional short baseline uv -coverage, and the additional MWA256 + TGSS long baselines, are required to obtain suf-

ficient information on all relevant scales, for a high signal-to-noise ratio detection.

For SKA1-Low, we find that arrays with maximum baselines of 49 km and 65 km yield comparable performance at 50 MHz and 150 MHz, while 39 km, 14 km, and 4 km arrays yield degraded performance. This is particularly true at CD (low) frequencies, where SKA1 is aiming to be transformational and new. We therefore conclude that 49 km maximum baselines are sufficient to form the sky and calibration model for EoR/CD science, *but 39 km baselines, corresponding to removal of two clusters from each spiral arm, yield degraded results and threaten high-redshift Cosmic Dawn science.* This is not true for the same longest baseline but an array with more intermediate scales (i.e., the v7 design). Such alterations of the inner array may be used to alleviate some of the degradation caused by removing the outer stations. We additionally find that it is multi-scale sources, which have power on large and small scales, that are those that yield different power bias for the 39 km and 65 km arrays. It is these sources for which the sky model estimation will yield differences depending on array design.

4 SUMMARY AND FORWARD LOOK

The broad design of SKA1-Low is relatively fixed, with an expectation of an aperture array interferometer of $\sim 130\,000$ dipoles collected into ~ 500 stations, spanning tens of kilometres, and with a densely-packed core of 50–60% of the collecting area within 2–3 km (Dewdney 2016). This broad

model provides the exceptional surface brightness sensitivity and wide frequency coverage to address the exciting science goals of the observatory. The specific details of station location and maximum array baseline are under discussion. To contribute to that discussion, here we study the impact of incomplete sky models of extended sources on EoR and CD science for MWA and SKA. The recommendations for the design of SKA1-Low are (1) an unpacking of inner station clusters; (2) a minimum longest baseline of 50 km. Combining these recommendations may improve performance but at an overall reduced cost.

ACKNOWLEDGEMENTS

CMT thanks Robert Braun and Ben McKinley for useful discussions. The Centre for All-Sky Astrophysics (CAASTRO) is an Australian Research Council Centre of Excellence, funded by Grant CE110001020. The Centre for All-Sky Astrophysics in 3D (ASTRO 3D) is an Australian Research Council Centre of Excellence, funded by Grant CE170100013. This research has made use of NASA's Astrophysics Data System. CMT is supported under the Australian Research Council's Discovery Early Career Researcher funding scheme (project number DE140100316). This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. We acknowledge the International Centre for Radio Astronomy Research (ICRAR), a Joint Venture of Curtin University and The University of Western Australia, funded by the Western Australian State government.

REFERENCES

- Ali, Z. S., et al. 2015, *ApJ*, **809**, 61
- Barry, N., Hazelton, B., Sullivan, I., Morales, M. F., & Pober, J. C. 2016, *MNRAS*, **461**, 3 135
- Becker, R. H., White, R. L., & Helfand, D. J. 1994, in ASP Conf. Ser. Vol. 61, *Astronomical Data Analysis Software and Systems* III, eds. D. R. Crabtree, R. J. Hanisch, & J. Barnes (San Francisco: ASP), 165
- Braun, R. 2013, *A&A*, **551**, A91
- Carroll, P. A., et al. 2016, *MNRAS*, **461**, 4151
- Castelletti, G., Dubner, G., Brogan, C., & Kassim, N. E. 2007, *A&A*, **471**, 537
- DeBoer, D. R., et al. 2017, *PASP*, **129**, 045001
- Dewdney, P. 2016, SKA Memo Series, SKA-TEL-SKO-0000002
- Hurley-Walker, N., et al. 2014, *PASA*, **31**, e045
- Hurley-Walker, N., et al. 2017, *MNRAS*, **464**, 1146
- Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, *A&A*, **598**, A78
- Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., & Lal, D. V. 2011, *A&A*, **535**, A38
- Jacobs, D. C., et al. 2016, *ApJ*, **825**, 114
- Jones, M., Mort, B., Dulwich, F., Wayth, R., Abeywickrema, S., & Bolton, R. 2016, SKA Engineering Change Proposal, www.skatelescope.org
- Kay, S. M. 1993, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Upper Saddle River: Prentice-Hall)
- Koopmans, L., et al. 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 1
- Lane, W. M., Cotton, W. D., Helmboldt, J. F., & Kassim, N. E. 2012, *RaSc*, **47**, RS0K04
- Line, J. L. B., Webster, R. L., Pindor, B., Mitchell, D. A., & Trott, C. M. 2017, *PASA*, **34**, e003
- McKinley, B., et al. 2015, *MNRAS*, **446**, 3478
- Mesinger, A., Furlanetto, S., & Cen, R. 2011, *MNRAS*, **411**, 955
- Morales, M. F., & Hewitt, J. 2004, *ApJ*, **615**, 7
- Parsons, A. R., et al. 2010, *AJ*, **139**, 1468
- Perley, R. A., Dreher, J. W., & Cowan, J. J. 1984, *ApJ*, **285**, L35
- Procopio, P., et al. 2017, *PASA*, **34**, 33
- Salter, C. J., et al. 1989, *A&A*, **220**, 42
- Tingay, S. J., et al. 2013, *PASA*, **30**, 7
- Trott, C. M., Wayth, R. B., & Tingay, S. J. 2012, *ApJ*, **757**, 101
- Trott, C. M., et al. 2016, *ApJ*, **818**, 139
- Wayth, R. B., et al. 2015, *PASA*, **32**, 25
- Windhorst, R., Mathis, D., & Neuschaefer, L. 1990, in ASP Conf. Ser. Vol. 10, *Evolution of the Universe of Galaxies*, ed. R. G. Kron (San Francisco: ASP), 389