Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **510,** 3838–3848 (2022) https://doi.org/10.1093/mnras/stab3706
Advance Access publication 2022 January 5

Epoch of reionization parameter estimation with the 21-cm bispectrum

Catherine A. Watkinson [1]★ Bradley Greig [2,3] and Andrei Mesinger [4]

[1]*School of Physics and Astronomy, Queen Mary University of London, Mile End Road, London E1 4NS, UK*
[2]*School of Physics, University of Melbourne, Parkville VIC 3010, Australia*
[3]*ARC Centre of Excellence for All-Sky Astrophysics in 3 Dimensions (ASTRO 3D)*
[4]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

Accepted 2021 December 15. Received 2021 December 7; in original form 2021 February 3

ABSTRACT

We present the first application of the isosceles bispectrum to MCMC parameter inference from the cosmic 21-cm signal. We extend the MCMC sampler 21CMMC to use the fast bispectrum code, BIFFT, when computing the likelihood. We create mock 1000-h observations with SKA1-low, using PYOBS21 to account for uv-sampling and thermal noise. Assuming the spin temperature is much higher than that of the cosmic microwave background, we consider two different reionization histories for our mock observations: fiducial and late-reionization. For both models we find that bias on the inferred parameter means and 1σ credible intervals can be substantially reduced by using the isosceles bispectrum (calculated for a wide range of scales and triangle shapes) together with the power spectrum (as opposed to just using one of the statistics). We find that making the simplifying assumption of a Gaussian likelihood with a diagonal covariance matrix does not notably bias parameter constraints for the three-parameter reionization model and basic instrumental effects considered here. This is true even if we use extreme (unlikely) initial conditions which would be expected to amplify biases. We also find that using the cosmic variance error calculated with Monte Carlo simulations using the fiducial model parameters while assuming the late-reionization model for the simulated data also does not strongly bias the inference. This implies we may be able to sparsely sample and interpolate the cosmic variance error over the parameter space, substantially reducing computational costs. All codes used in this work are publicly available.

Key words: methods: statistical – intergalactic medium – cosmology: theory – dark ages, reionization, first stars.

1 INTRODUCTION

The Square Kilometre Array[1] (SKA) aims to detect the high-redshift 21-cm line of neutral hydrogen. It is projected to produce high precision maps at a wide range of redshifts. These maps can be used to infer the properties of early generations of stars and galaxies as they influence the intergalactic medium (IGM) via coupling, heating, and ionizations (Dewdney 2016). The phase change in the Universe's ionization state induced by the latter process is called the Epoch of Reionization (EoR).

Numerous studies have predicted great benefits from using higher-order statistics, such as the bispectrum in our analysis of such data sets. For example, Shimabukuro et al. (2016), Majumdar et al. (2017), Watkinson et al. (2019b), Hutter et al. (2019), and Gorce & Pritchard (2019) show that, due to the non-Gaussian nature of the signal, additional information is contained in higher-order statistics, which unlike the power spectrum, are sensitive to non-Gaussian structure in a data set. In particular, Shimabukuro et al. (2016) perform a Fisher forecast and find that using the equilateral bispectrum in addition to the power spectrum substantially shrinks the credible limits of the parameters of a three-parameter EoR model compared to those resulting from using the power spectrum alone.

Furthermore, it appears that the error due to instrumental noise is not as large as one might naively expect; see for example, Yoshiura et al. (2015), Watkinson et al. (2019b), and Trott et al. (2019). This is because Gaussian distributed noise has zero bispectrum so that it is only the statistical fluctuations of the noise bispectrum that contributes to the measured error on the bispectrum (Yoshiura et al. 2015).

The Fisher analysis of Shimabukuro et al. (2016), while an important first step towards understanding the improvements gained in performing parameter estimation with the bispectrum, likely underestimates the credible limits associated with each parameter. This is because a parameter's covariance matrix is only accurately described by the inverse of the Fisher matrix if the errors on the measured quantities are perfectly Gaussian (i.e. the likelihood surface is Gaussian around the maximum likelihood point), which is not a given for even the 21-cm power spectrum. It has also been shown that the covariance predicted by a Fisher forecast, by the Cramer–Rao bound, provides the smallest possible attainable error, i.e. it provides a lower limit (Fisher 1935; Cramér 1946; Rao 1945; Tegmark, Taylor & Heavens 1997). In this paper, we take the work of Shimabukuro et al. (2016), a step further by adding the isosceles bispectrum (in which we include a wide range of triangle configurations in addition to the equilateral) within a Monte Carlo Markov Chain (MCMC)

★ E-mail: catherine.watkinson@gmail.com
[1]The Square Kilometre Array http://www.skatelescope.org/ and https://astronomers.skatelescope.org/wp-content/uploads/2016/05/SKA-TEL-SKO-0000002_03_SKA1SystemBaselineDesignV2.pdf

© 2022 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

parameter estimation framework, building on the established 21-cm MCMC code 21CMMC (Greig & Mesinger 2015a, 2017b; Park et al. 2018).

Section 2 describes our bispectrum likelihood and the methods used to simulate instrumental effects and measure the bispectrum. In Section 3, we look at an idealized case with no instrumental effects or sample variance to see the maximal achievable improvement to the parameter constraints when combining the bispectrum and power spectrum. In Section 4.1, we compare analytic approximations to the sample-variance error with the true sample-variance error calculated using Monte Carlo (MC) methods. We will show in this section that assuming a sample-variance error that is a fixed percentage of the statistics in any given bin is a very poor approximation, as is propagating the power-spectrum sample-variance error on to the bispectrum assuming Gaussianity. In Section 4.2, we present our main analysis that include instrumental effects (*uv* sampling & noise) and sample variance. We will show in this section that using the bispectrum in combination with the power spectrum reduces the bias (and in some cases the credible intervals) on all parameters relative to that of the power-spectrum only analysis. This is true regardless of how likely is the realization of the "true" universe (i.e. if the initial conditions are outliers) or its reionization history.

2 INCLUSION OF INSTRUMENTAL EFFECTS AND BISPECTRUM LIKELIHOOD TO 21CMMC

For the purposes of this analysis, we modify the latest version of 21CMMC : an MC sampler of 21CMFASTv3 (a PYTHON-wrapped, seminumerical simulation of the 21-cm signal at high redshifts; Murray et al. 2020). 21CMMC can be downloaded from https://github.com/21cmFAST/21CMMC, and is detailed in: Greig & Mesinger (2015a; which describes the first implementation that used a three-parameter model for reionization), Greig & Mesinger (2017b; which extends sampling to parameters responsible for heating and Lyman-α coupling effects), and Park et al. (2018; which introduces mass dependence to the star formation rates and escape fraction of ionizing radiation, as well as luminosity functions). The latest version of 21CMMC has the option of using either the EMCEE or Multinest samplers; here we use EMCEE which is an Affine-invariant, openMP-parallelized MCMC sampler (for more details see https://emcee.readthedocs.io/en/stable/) (Goodman & Weare 2010; Foreman-Mackey et al. 2013).

21CMFASTv3 is a standalone code for computing 3D realizations of the 21-cm signal and its component fields. Sampling Gaussian initial conditions, it uses Lagrangian perturbation theory to generate density and velocity fields (e.g. Bernardeau et al. 2001); then using a combination of excursion set (Furlanetto, Zaldarriaga & Hernquist 2004) and light cone integration, it generates ionization and temperature fields. We refer the interested reader to Mesinger & Furlanetto (2007) and Mesinger, Furlanetto & Cen (2010) for details, as well as to the extensive documentation associated with the code itself available at https://github.com/21cmfast/21cmFAST.

For this demonstrative work, we use the simplest, three parameter reionization model (as described in Greig & Mesinger 2015a), and assume the spin temperature exceeds the cosmic microwave background temperature. We also compute our summary statistics from coeval cubes, instead of light cones.[2] These choices keep the analysis time to a minimum facilitating the ability to experiment with different aspects of the analysis while still being informative. In future work, we will relax these assumptions.

The parameters that we vary in our analysis are:

(i) $\zeta = f_{\rm esc}\, f_*\, N_{\gamma/b}\, (1 + n_{\rm rec})^{-1}$ which is the ionizing efficiency of galaxies. Here, $f_{\rm esc}$ is the escape fraction of ionizing photons, $N_{\gamma/b}$ is the number of ionizing photons produced per baryon in stars, and $n_{\rm rec}$ is the cumulative number of IGM recombinations per baryon. This is assumed to be a constant, and a region is deemed to be ionized if the collapsed fraction within that region is greater than or equal to ζ^{-1}. Increasing ζ therefore speeds up the EoR.

(ii) $T_{\rm vir}$ is the minimum virial temperature needed for haloes to host star-forming galaxies (determined by cooling and feedback mechanisms that allow star formation). Smaller $T_{\rm vir}$ means star formation is possible in lower-mass haloes that are less biased. Thus reducing $T_{\rm vir}$ results in an earlier EoR, characterized by smaller, more uniformly distributed cosmic H II regions.

(iii) $R_{\rm max}$ defines the maximum distance a photon can travel in an ionized IGM before it encounters a recombined atom. This effective parameter can loosely be related to a characteristic mean free path (c.f. Furlanetto & Oh 2005 and Sobacchi & Mesinger 2014). As $R_{\rm max}$ is only relevant when it is smaller than the typical H II region size, reducing it extends the late stages of the EoR without impacting the early stages.

We make the assumption that the power spectrum and bispectrum measurements are independent (from each other and between each k bin for the power spectrum or triangle configuration for the bispectrum). We also assume independence of these statistics at each redshift. This allows us to approximate the total likelihood using a simple sum over χ^2 values. Specifically, we take $\ln \mathcal{L}(\theta|d) = -\sum_{ij}(d_{ij} - m_{ij})/(2\,\sigma_{ij})$, where the indices denote redshift and statistical bins, i.e. each ij corresponds to the measurement of a single power spectrum or bispectrum bin (from the data d_{ij} or model m_{ij}) at one of the redshift bins under consideration. For the main results of this paper, we pre-compute σ_{ij} by forward simulating the fiducial model, each time varying the initial seed of the simulation to account for the sample variance error, and including a random realization of instrumental noise. The standard deviation we use in this study is calculated using 2000 such Monte Carlo (MC) samples of the power spectrum and bispectrum in each bin (although it is worth noting that the error estimate has mostly converged by 1000 iterations).

We ignore the contribution to the power spectrum and bispectrum for k modes that fall outside of the range $0.1 \leq k \leq 1.0$ cMpc^{-1}. The lower k cut is motivated by avoiding modes that are likely to suffer from corruption due to foreground leakage, and the upper cut excludes modes that will suffer from the effects of shot noise (Greig & Mesinger 2015a). For the bispectrum, this means that if any one of the three k-vectors that form a given triangle configuration fall outside of this range, then the configuration is excluded from our likelihood calculation.

We set our fiducial model parameter values as $\zeta = 30.0$, $\log T_{\rm vir} = 4.7$, and $R_{\rm max} = 15$. We also consider a late reionization model with $\zeta = 17.0$, $\log T_{\rm vir} = 5$, and $R_{\rm max} = 10$. We initialize the core of 21CMMC to simulate coeval cubes at $z = [6.3, 7, 8, \text{and } 9]$, chosen to sample a range of ionized fractions, with our redshifts corresponding to $x_{\rm H\,I} = [0.13, 0.33, 0.62, 0.79]$ for our fiducial model and $x_{\rm H\,I} = [0.70, 0.80, 0.89, 0.94]$ for our late reionization model. Note that our

[2]A coeval cube is a data cube that has been simulated using a fixed cosmological time throughout. A light cone data set is one in which the simulated epoch evolves with frequency (or redshift), i.e. each slice along the z-axis represents a different cosmological time.

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

late reionization model is not picked as a realistic model, it is selected somewhat arbitrarily to provide a test case that is quite different to the fiducial model.[3] We use the same prior ranges as Greig & Mesinger (2015a), i.e. $10 \leq \zeta \leq 30$, $4 \leq T_{\rm vir} \leq 6$ and $5 \leq R_{\rm max(bubble)} \leq 20$. Our coeval cubes are 128^3 and $(256\ {\rm Mpc})^3$ in dimension, chosen to keep both sample variance and analysis time to acceptable levels (Iliev et al. 2014; Kaur, Gillet & Mesinger 2020).

Note, performing forward modelling of the full intended survey area for EoR science with the SKA is not possible. For example, the smallest intended survey footprint is $\sim$100 deg.^{2} (Koopmans et al. 2015) which at 150 MHz corresponds to $\sim$1.5 Gpc. Instead, for parameter inference, we typically focus on simulating close to the primary field of view, which for the SKA is $\sim$3 deg ($\sim$480 Mpc at 150 MHz). Nominally, we would choose our mock observation to be of this scale, however, in this work we are interested in estimating the true covariance of the sample variance uncertainty, which requires MC sampling a large number of realization. As such, we restrict our investigations to simulations of side length 250 Mpc.

Importantly, this implies that our estimate of the sample variance will be larger than that expected from the SKA. However, the impact of this can simply be interpreted as our inferred constraints being considered conservative estimates. Further, the primary focus of this work is investigating the improvements in our constraining power, following the inclusion of the bispectrum rather than performing a forecast for the SKA.

2.1 BIFFT – a fast code for measuring the bispectrum

The bispectrum is defined as the Fourier transform of the three-point correlation function (which measures excess probability as a function of three points in real space). It can be written as

$$(2\pi)^3 B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\delta^{\rm D}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) = \langle \Delta(\boldsymbol{k}_1)\Delta(\boldsymbol{k}_2)\Delta(\boldsymbol{k}_3)\rangle, \quad (1)$$

where $\delta^{\rm D}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)$ is the Dirac-delta function. Accordingly, the bispectrum is a function of three k vectors that form a closed triangle, often referred to (as we will from here on) as a triangle configuration. It is necessary to perform some kind of averaging when measuring the bispectrum to beat down statistical noise. As is common in bispectrum and power spectrum analysis, we choose to perform spherical averaging, i.e. our bispectrum measurements are functions of triangle shape and size only, not orientation.

The bispectrum is the lowest-order polyspectra that is sensitive to non-Gaussian information, or structure, in a data set. For a nice description of the real-space structures that different k-space triangle configurations are sensitive to see Lewis (2011), Watkinson et al. (2019b), and Hutter et al. (2019; see in particular fig. 1).

Due to computational limitations, the bispectrum is often overlooked in forward-modelling frameworks. Naively, it requires multiple nested loops to find the k-space pixels that form closed triangles of the desired shape and size. However, there are methods that make the calculation tractable for many applications. One of these is to use Fast-Fourier Transforms to enforce the Dirac-delta function in equation (1; Scoccimarro 2015; Sefusatti et al. 2016). BIFFT is a PYTHON package that wraps a C implementation of the Fourier-transform bispectrum method, described in Watkinson, Majumdar & Pritchard (2017a) and publicly available from https://bitbucket.org/caw11/bifft. It is very fast, taking only a few seconds per triangle configuration on a MacBookPro (2.3 GHz i9 intel core, 16Gb RAM) for a data cube of size 256^3. This method is extensively described in Watkinson et al. (2017a); Watkinson, Trott & Hothi (2021c).

Throughout we will normalize out the amplitude of the bispectrum to isolate the non-Gaussian information:

$$b(k_1, k_2, k_3) = \frac{B(k_1, k_2, k_3)}{\sqrt{(k_1\, k_2\, k_3)^{-1}\ P(k_1)\, P(k_2)\, P(k_3)}}. \quad (2)$$

Equation (2) is commonly applied in signal processing, see for example Hinich & Clay (1968), Kim & Powers (1978), Hinich & Messer (1995), and Hinich & Wolinsky (2005). It has also been argued by Brillinger & Rosenblatt (1967) that equation (2) is the preferred normalization for the bispectrum. In the context of the 21-cm signal, this was explored at length by Watkinson et al. (2019b). Primarily, this preference comes from the fact that it removes the correlations with the power spectrum meaning it better isolates the non-Gaussian information. Equally, since it is not modulated by the power spectrum amplitude, it is less susceptible to statistical noise which is preferred for parameter inference. Consequently, since this normalized bispectrum is not a direct function of the power spectrum, it is easier to justify its linear combination with the power spectrum in our likelihood function.

2.2 *uv* sampling and noise generation with PYOBS21

In order to carry out our investigation we wrote PYOBS21 (which can be used as a bolt-on module for 21CMMC or 21CMFASTV3) to apply *uv* sampling and add Gaussian random noise (with standard deviation based on 21CMSENSE calculations) to a 21-cm brightness-temperature coeval simulation.[4]

The established code 21CMSENSE outputs the noise and sample-variance error of the spherically averaged power spectrum as a function of k. PYOBS21 relies on an adapted version of calc_sense.py from 21CMSENSE, which instead outputs a file containing the k_x, k_y, k_z (in cMpc^{-1}) corresponding to the instrument's *uv* sampling and bandwidth associated with the simulation dimensions, along with the noise power spectrum associated with each *uv* sample. 21CMSENSE is described extensively in Pober et al. (2013a) and Pober et al. (2014b). We assume optimistically that foregrounds are fully removed and assume a track scan mode of operation. On the first call to PYOBS21, a maskfile of the same dimension as the 21CMFASTV3 simulation is created containing the noise power in each pixel (the noise in pixels that are repeat samples are combined coherently using inverse-covariance weighting) and zeroed where there are no *uv*-samples. Once the *uv*-noise maskfile is written to file, PYOBS21 accesses it each time it is called, zeroes any unsampled pixels in the cosmological simulation, and adds a random sample of Gaussian noise to each pixel (based on the noise power in the corresponding *uv*-noise maskfile pixel).

By working in simulation co-ordinates (i.e. cMpc) and creating the uv-noise maskfile on the first call, PYOBS21 is very quick, making it suitable for MC calculations, including calculating instrumental error on any statistic (that is in itself also relatively quick to compute). Note, this approach is an approximation as it ignores the evolution of the *uv* sampling along the line of sight. It also ignores the effect of the primary beam, effectively assuming the field size is small enough

[3]The ionized fractions we quote are for our 'standard' seed, which we discuss in Section 4.1.

[4]PYOBS21 can be used for light cone data if it is chunked into cubes, but since PYOBS21 assumes a fixed redshift in translating the *uv* sampling of the instrument to simulation co-ordinates, it is not the ideal tool for use with light cones.

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

to not be affected by this (which for the box sizes simulated here is not unreasonable) or that the primary beam has been corrected for. The SKA noise level produced by this PYOBS21 (using the central region from the current design for the SKA-Low phase 1 telescope model and assuming 1000-h observation time) is consistent with that predicted by Mellema et al. (2013) and Koopmans et al. (2015). The SKA1-Low details and antenna locations used for our noise calculations are based on the latest SKA configuration coordinates[5] (central region) and Dewdney (2016).

3 PARAMETER RECOVERY USING THE ISOSCELES BISPECTRUM FOR AN IDEALIZED CASE

In this work, we only consider the isosceles configuration as a function of angle between k1 and k2, and for a range of scales. Our range of isosceles triangles span shapes from squeezed to stretched, and should therefore be able to pick up a large range of non-Gaussian structures in the 21-cm maps. We refer the reader to section 3 of Watkinson et al. (2019b) and Lewis (2011) for discussions of the types of structures that various configurations are sensitive to, as well as to the results of Majumdar et al. (2017) for verification that the isosceles configuration captures key features of reionization maps.

In this section, we compare the parameter constraints achieved when using the isosceles bispectrum (for $k_1 = k_2 = [0.12\,\mathrm{cMpc}^{-1}, 0.3\,\mathrm{cMpc}^{-1}, 0.7\,\mathrm{cMpc}^{-1}, 0.98\,\mathrm{cMpc}^{-1}]$) and for $\theta/\pi = [0.01, 0.05, 0.1, 0.2, 0.33, 0.4, 0.5, 0.6, 0.7, 0.85,$ and $0.95]$ (where θ is the internal angle to $\boldsymbol{k}_1 + \boldsymbol{k}_2$), the power spectrum, and a combination of the two statistics.[6] To do so, we assume a best case scenario of negligible instrumental effects, perfect foreground removal, and negligible sample variance (i.e. the only source of error is our 15 per cent modelling uncertainty). In practice, this involves a running analysis on the raw coeval cubes produced by 21CMFASTv3 and assuming the same random seed for the data and model. We also include a modelling uncertainty of 15 per cent of the modelled statistics, as is default in 21CMMC).

The corresponding corner plot for the three parameter model is shown in Fig. 1. Darker/lighter shading encloses 68 per cent/95 per cent of the credible limits. Different colors indicate different statistics used for computing the likelihood: (i) bispectrum is shown with grey; (ii) power spectrum is shown with red; and (iii) bispectrum + power spectrum is shown with blue.

Under these idealized conditions, the power spectrum only (pspec-only) statistic results in tight, unbiased constraints, which can be seen in the bottom of Fig. 1, where we plot the marginal statistics, i.e. the marginalized posterior's mean ± the 68 per cent upper and lower credible limits. As in Greig & Mesinger (2015a), we see a moderate degeneracy between the ionizing efficiency and the virial temperature. This is because both parameters affect the timing of reionization; for example, both a high virial temperature and a low ionizing efficiency will delay and slow the progress of reionization. The epoch of heating, ignored in this exploratory work, should break this degeneracy (e.g. Greig & Mesinger 2017b).

[5]The SKA antenna positions we use are given by the central region antenna positions of https://astronomers.skatelescope.org/wp-content/uploads/2016/09/SKA-TEL-SKO-0000422_02_SKA1_LowConfigurationCoordinates-1.pdf

[6]For all the statistics we consider, we disregard contribution from any k modes that fall outside of the range $k_f < k < k_{\mathrm{nyq}}$, where $k_f = 2\pi/L$ is the fundamental k scale and L is the length of a side of the simulation, and $k_{\mathrm{nyq}} = 1.0/2.0{*}N{*}k_f$, where N is the resolution on a side. A consequence of these restrictions is that not all θ bins will be included for larger values of k_1.

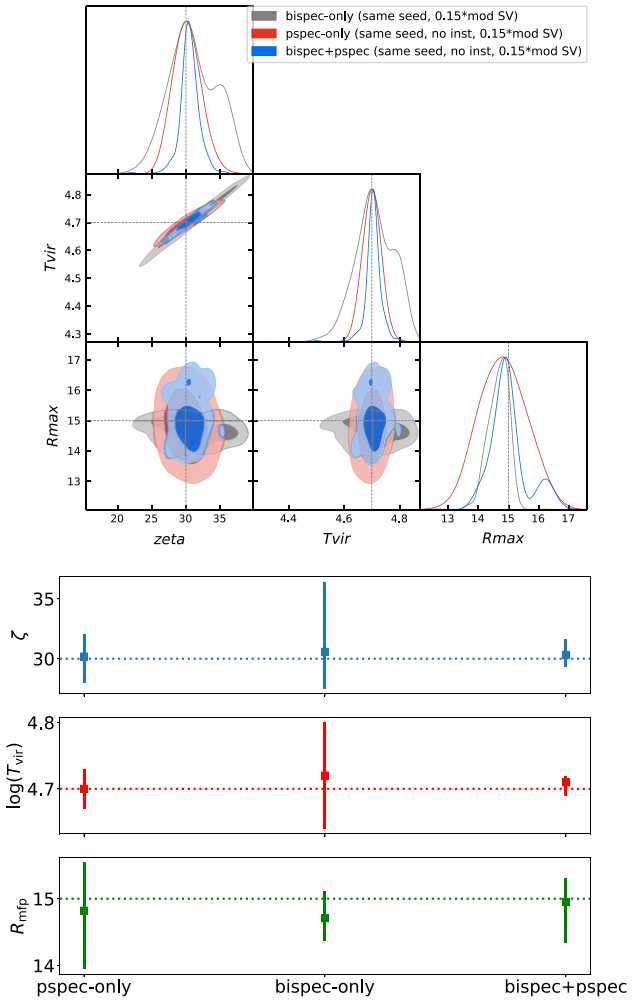

Figure 1. Corner plot (top) for a likelihood based on the spherically averaged isosceles bispectrum (bispec-only; grey), power spectrum (pspec-only; red), and power spectrum + bispectrum (bispec+pspec; blue). The bottom plot shows the mean ±68 per cent credible intervals for each parameter. All assume a best-case scenario of no instrumental effects or foregrounds and use the same random seed for our models and data. In this and all the figures that follow, our simulations have dimensions of 128^3 pixels and $(256\,\mathrm{cMpc})^3$ and redshifts simulated are $z = [6.3, 7, 8,$ and $9]$. For our bispectrum likelihood, we use the isosceles triangle configurations for 11 linearly spaced θ bins and for $k_1 = k_2 = [0.12\,\mathrm{cMpc}^{-1}, 0.3\,\mathrm{cMpc}^{-1}, 0.7\,\mathrm{cMpc}^{-1}, 0.98\,\mathrm{cMpc}^{-1}]$ (where θ is the internal angle to $\boldsymbol{k}_1 + \boldsymbol{k}_2$). We see the power spectrum in such a case does a good job of constraining the data but constraints are improved by the inclusion of the bispectrum.

We find the pspec-only statistic generally results in tighter constraints than the bispectrum only (bispec-only) statistic. This tells us that even in the idealized scenario, the amplitude of the signal is more informative than the non-Gaussian information alone (at least for our fiducial model). However, the credible intervals of R_{max} are reduced by a factor of 0.47 relative to the pspec-only case (see also Shaw, Bharadwaj & Mondal 2020). This is because R_{max} (by applying a hard limit beyond which photons from a source will cease to be effective at ionizing the IGM) induces structural features, to which the bispectrum is particularly sensitive.

When we combine the bispectrum with the power spectrum, the additional information from the non-Gaussianities in the maps greatly reduces the degeneracies of the credible limits for all the

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

parameters. This corresponds to a shrinkage of the credible intervals by a factor of 0.70, 0.50, 0.60 for ζ, $\log(T_{\rm vir})$, and $R_{\rm max}$, respectively (with respect to those of pspec-only). Although we note that the marginalized posterior mean is closer to the truth for both ζ and $\log(T_{\rm vir})$ for the pspec-only case. This degree of improvement is roughly in agreement with Shimabukuro et al. (2016), who perform a Fisher analysis using 21CMFASTv2, in which they consider the sensitivity levels of LOFAR and MWA. Although, even in the best case scenario of a perfect observation, the degree of improvement is not as extreme as the Fisher analysis suggests. This is understandable, since the inverse of the Fisher matrix only provides an estimate of smallest achievable credible limit.

4 THE IMPACT OF SAMPLE VARIANCE AND INSTRUMENTAL EFFECTS

4.1 Modelling the sample-variance error

A major challenge to performing parameter estimation with 21-cm data and simulations is correctly accounting for sample variance. Even at the level of the power spectrum, this is difficult as the error due to sample variance is dependent on the 21-cm signal itself, and therefore the model parameters. This makes it a great challenge to model the sample-variance error using MC simulations as we have here. One would need to effectively sample the full model parameter space (which for the current most complex 21CMMC model consists of 17 astrophysical parameters, see Qin et al. 2020) at each point performing at least several hundred, ideally thousands of simulations with different initial conditions. This would realistically require the use of a machine-learning interpolation procedure to make this tractable. You would also need to decide *a priori* how you are going to chop up your light cone to measure your statistics as a function of redshift (necessary to effectively capture the evolution of the signal with a redshift using such summary statistics), or store all the simulations to avoid being locked into any such choice (not a terribly practical option). It is therefore interesting to consider whether we might be able to approximate the sample-variance error for the bispectrum using a similar analytic approach to that for the power spectrum.

Assuming the signal is Gaussian, an estimate for the power spectrum sample-variance error is given by $\Delta^2_{\rm sv}(k) = \Delta^2_{21}(k) = k^3/(2\pi^2)P_{21}(k)/\sqrt{N(k)}$, where $P_{21}(k)$ is the 21-cm brightness-temperature power spectrum and $\sqrt{N(k)}$ is the number of times a particular mode has been sampled. Similarly, we can calculate the theoretical bispectrum sample variance error assuming it is Gaussian distributed (as is often done in the case of Gaussian noise) as,

$$[\Delta_{\rm sv} B(k_1, k_2, k_3)]^2 = k_f^3 \frac{n_{123}}{V_{123}} \Delta_{\rm sv} P(k_1)\, \Delta_{\rm sv} P(k_2)\, \Delta_{\rm sv} P(k_3), \qquad (3)$$

in this expression $k_{\rm f} = 2\pi/L$ is the fundamental k scale, $V_{123} \approx 8.0\pi^2\, k_1\, k_2\, k_3\, (s\, k_{\rm f})^3$ gives the number of fundamental triangles in units of $k_{\rm f}^3$, $s\, k_{\rm f}$ is the binwidth, and $n_{123} = 1, 2, 6$ for general, isosceles, and equilateral triangle configurations, respectively (Scoccimarro et al. 1998b; Scoccimarro, Sefusatti & Zaldarriaga 2004a; Liguori et al. 2010). We assume $s = 1$ to obtain the maximum possible estimate for the theoretical sample-variance contribution to the bispectrum using this approximation.

Alternatively, we can estimate the true sample variance using an MC approach in which we vary the initial-conditions (random seeds) and random-noise realizations assuming the fiducial model parameters. The ratio of the MC sample-variance error to that calculated using equation (3) is plotted in Fig. 2 (where solid line correspond to $z = 6.3$, dot–dashed to $z = 7$, dotted with triangles to $z = 8$, and dashed with circles to $z = 9$). It is clear that this approximation is orders of magnitude lower than the true sample variance for this box size and resolution. It is also clear that there is no clean connection between this theoretical sample variance and the true sample variance.

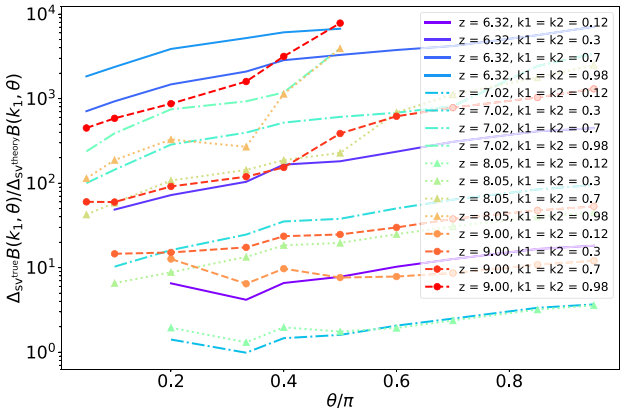

Figure 2. Ratio of the sample variance on the bispectrum as measured from brute force repeat simulation to that measured from theory assuming the signal is Gaussian. Solid line correspond to $z = 6.3$, dot–dashed to $z = 7$, dotted with triangles to $z = 8$, and dashed with circles to $z = 9$. The Gaussian assumption for the sample variance is unable to even qualitatively capture the features we see in the simulated sample variance.

Qualitatively, there are two main trends represented here. First, as reionization proceeds the non-Gaussianity increases owing to the percolation of the ionized regions. As such, the ratio of the sample variance estimates increases in amplitude with decreasing redshift (as the Gaussian approximation deviates further from the truth). Secondly, the amplitude increases for increasing $k_1 = k_2$ as the non-Gaussianity is more prevalent on small scales due to the bi-modality of the 21-cm signal from patchy reionization. Theoretically, we could improve on this approximation by adding the trispectrum contribution to each of the individual power spectrum sample variance terms in equation (3), where this trispectum term accounts for the non-Gaussianity of the 21-cm signal. However, in doing so, we still will not entirely account for all possible non-Gaussian contributions.

For the power spectrum, correctly accounting for this non-Gaussianity in the covariance has a non-negligible effect on the resulting parameter constraints (provided large-scale measurements are limited by thermal noise; Mondal, Bharadwaj & Majumdar 2016, Shaw, Bharadwaj & Mondal 2019b; Shaw et al. 2020). For the bispectrum, the sample variance dominates over the thermal noise contribution for a wider range of k values, thus it is more imperative to include the full sample variance than the approximation above. Therefore, for the rest of the paper, we will use the MC-estimated error estimates.

4.2 Parameter constraints using Monte Carlo-simulated error term

The initial conditions of our Universe can impact the outcome of our parameter estimation. To quantify this, we choose a "standard" and an "extreme" model for our mock observations used for parameter inference. Specifically, we use two different random seeds that exhibit minimal and maximal χ^2 from the mean of the signal, selected from among ~50 different realizations. In the analysis of this section,

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

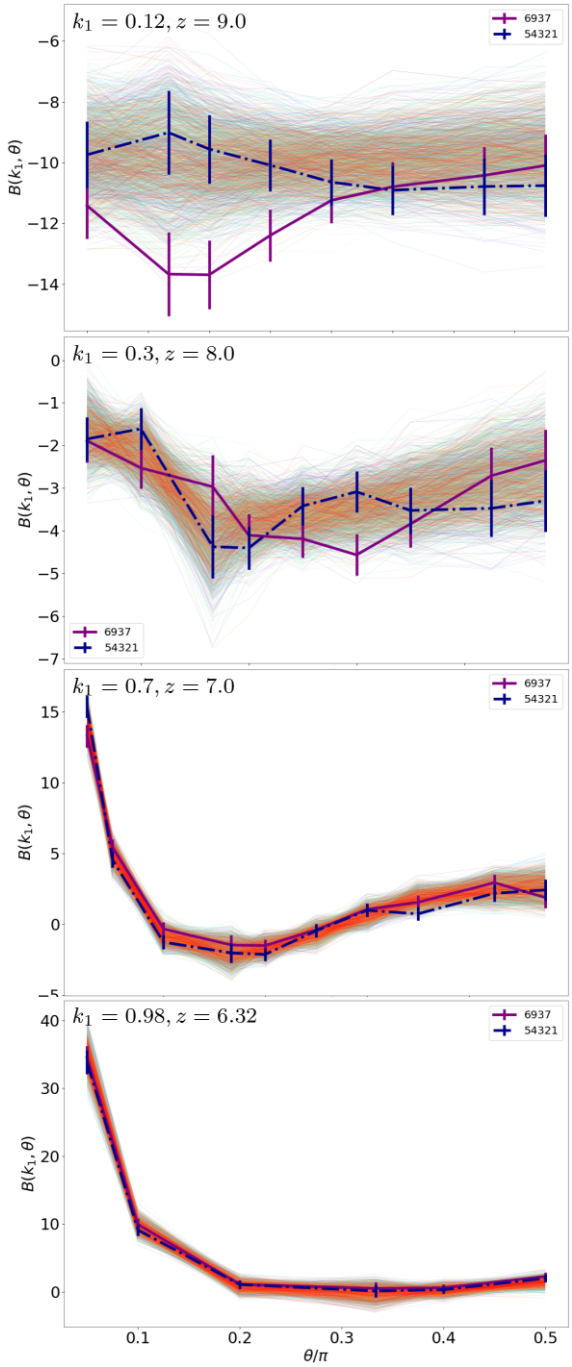

Figure 3. Here, we plot with thin lines all 2000 bispectra used in estimating the error due to sample variance for our simulation dimensions. The plots from top to bottom correspond to $k_1 = [0.12, 0.3, 0.7, 0.98]$ and $z = [9.0, 8.0, 7.0, 6.3]$. We overplot the two random seeds used in our parameter estimation analysis chosen from about 50 trial runs to minimize (54321) and maximize (6937) the reduced χ^2 between them and the mean of the distribution of the thin lines shown by the thin lines in the plot.

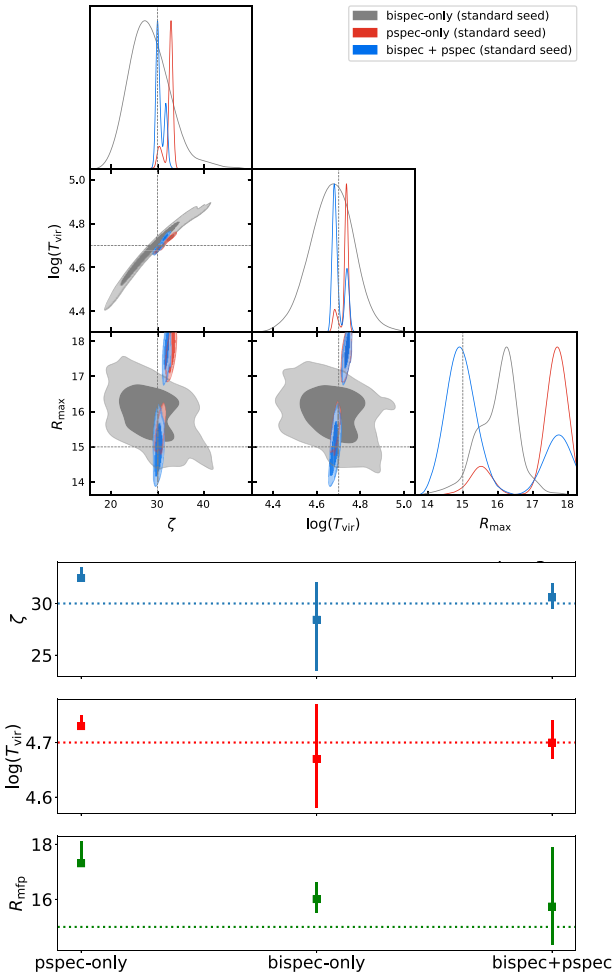

Figure 4. Corner plot (top) of credible intervals when using mock observed data generated using the fiducial model and the standard seed and the bispec-only (grey contours), the pspec-only (red contours), and bipsec+pspec (blue contours) as summary statistics in the likelihood. The blacked-dashed lines indicate the parameter values used to generate the mock observed data sets for each model. The bottom plot shows the mean ± 68 per cent credible intervals for each parameter. All include the effects of SKA-LOW (phase 1) *uv* sampling and noise, as well as sample variance, which we model the associated standard deviation using MC methods and using the parameters of the fiducial model. While all cases contain the truth within their 95 per cent credible intervals, the posterior probability mass for the pspec-only case is concentrated in a different region of model parameter space, resulting in biased marginal statistics.

we use the MCMC-estimated noise+sample variance error, but since we are using 21CMFASTv3 for generating our mock observations, we set the modelling error factor to $A = 0.0$. We show the bispectrum of these two random seeds in Fig. 3, we also plot in thin lines the full range of bispectrum produced in the repeat sampling we used to estimate the 1σ sample-variance errors (which are the error bars on each of our random seed bispectra). The plots from top to bottom correspond to $k_1 = [0.12, 0.3, 0.7, 0.98]$ and $z = [9.0, 8.0, 7.0, 6.3]$. As can be seen from this plot, seed 6937 is our "extreme" seed and seed 54 321 is our "standard" seed. For the interested reader, we have included the equivalent plots for the power spectrum in Appendix A.

The top plot of Fig. 4 shows the resulting credible intervals when we use the standard seed and assume the parameters of our fiducial model for our mock observed data. The forward model and mock observed data used for the analysis behind this plot both include instrumental effects (i.e. *uv* sampling and noise). As before, the largest grey contour shows the bispectrum-only case, the red contours the power-spectrum only case, and the blue contours the bispectrum + power spectrum case. For both models the true parameters values (marked with the black dashed lines) lie within the 95 per cent cred-

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

ible intervals for all three combinations of statistic, however for the fiducial model the power spectrum posterior is bimodal. Furthermore, there is more probability density in the mode that is centred around different parameter values to the truth, leading to biased marginal statistics (this can be seen from the marginalized statistics for this case, which we show in the bottom plot of Fig. 4). The posterior for the bispectrum-only case has its probability-density focused around the true parameter values for $T_{\rm vir}$ and ζ, but as with the power spectrum exhibits bias towards larger $R_{\rm max}$. All is saved by combining the power spectrum with the bispectrum, the marginal statistics of which do not suffer from bias on the inferred parameter values.

We can attribute some blame for this bias to sample variance. That is, numerical fluctuations in the statistics measured from the mock observation compared to the statistics measured from the realizations sampled within the MCMC, which use different random seeds to the mock. Further, it is well known that the $R_{\rm mfp}$ parameter causes a 'knee' feature in the power spectrum, meaning it is sensitive to one or a small number of k-bins in the power spectrum (as well as how wide or narrow the bins are in Fourier space). Numerical fluctuations due to different seeds can cause the amplitude of the 'knee' to differ or appear in neighbouring k-bins. These differences will be measurable by our likelihood function resulting in biases in the inferred parameters. Once we combine the statistics, in some sense we average out over this variance resulting in an improved ability to infer our fiducial values. We could additionally mitigate this by removing the $R_{\rm mfp}$ parameter in favour of a more self-consistent treatment in 21CMFASTv3using inhomogeneous recombinations (e.g. Sobacchi & Mesinger 2014). However, we defer this to future work.

If we now consider the results when we use the 'extreme' seed for generating our mock observed data sets, then we see that the 95 per cent credible intervals for all combinations of summary statistic still contain the true model parameters for all parameters. However, they are in a lower probability region of the posterior than they were for the case of the more standard seed. This can be seen in Fig. 5, where the top plot shows the corner plot for the fiducial model with extreme seed. We see that for the case of the 'extreme' seed, the weight is more evenly spread across the two posterior modes resulting in marginal statistics (which are summarized in the bottom plot of Fig. 5) that are less biased than one might imagine from examining the credible intervals by eye. The bias of the marginal posterior's mean is even reduced for the pspec-only case relative to the results using the more standard seed for the mock observed data set. Combining the bispectrum still improves the robustness of the results; however, the bias on the marginal statistics of $R_{\rm max}$ is not as reduced when the bispectrum and power spectrum are combined as it is for the more standard seed.

As can be seen in the corner plot of Fig. 6 (top), there is much less of an issue with bi-modality in the posterior for mock observed data generated with the parameters of our late reionization model; clearly this region of parameter space is less generic (i.e. the model statistics are very distinct from those of other models). The marginal statistics for this model are summarized in the bottom plot of 6. We see that for our late reionization model, using the bispectrum in combination with the power spectrum still overall reduces bias on the marginal statistics (although at the cost of introducing a small bias on the marginal statistics of ζ) and shrinks the credible intervals relative to those of either statistic alone. We found that even in test runs where we fixed the modelled initial conditions, using the standard seed and the extreme seeds for the data that the results for our late reionization model were still robust with no serious issue with biased results.

Fig. 7 shows the results for our late reionization model when the extreme seed is used to generate the mock observed data. As with

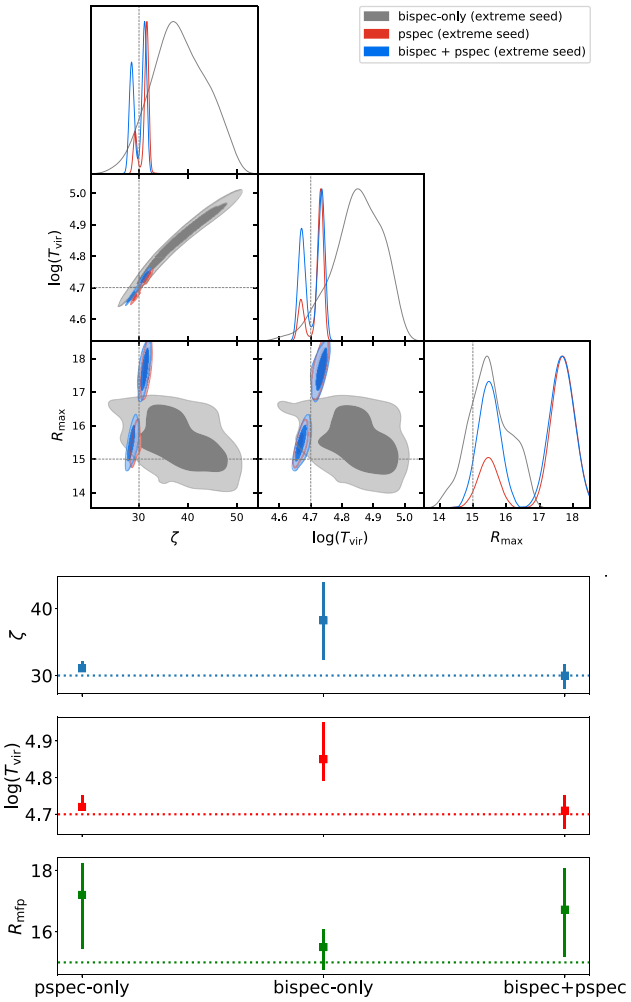

Figure 5. Corner plot (top) of credible intervals when using mock observed data generated using the fiducial model and the 'extreme' seed for the bispec-only (grey contours), pspec-only (red contours), and bipsec+pspec (blue contours) as summary statistics in the likelihood. The blacked dashed lines indicate the parameter values used to generate the mock observed data sets for each model. The bottom plot shows the mean ±68 per cent credible intervals for each parameter. All include SKA-LOW (phase 1) instrumental effects (assuming negligible primary beam effects), as well as sample variance, which we model the associated standard deviation using MC methods and using the parameters of the fiducial model. All cases contain the truth within their 95 per cent credible intervals, albeit in a lower probability region of the posterior.

the standard seed, the results for our late reionization model are less biased than they are for the fiducial model with the 95 per cent credible intervals of all combinations of statistic containing the true model and with the combining of the bispectrum and power spectrum improving the quality of the constraints. As we will discuss further in the following paragraph, this is because this model is at a much early stage of the reionization process for which differences between seeds are suppressed relative to that of the fiducial model.

What is potentially important about the results of the our late reionization model analysis is that we have used the standard deviation due to sample variance as calculated for the fiducial model, rather than calculating it for the our late reionization model parameters, i.e. we have seen no serious negative impact from assuming sample variance is the same in both regions of model parameter space, despite

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

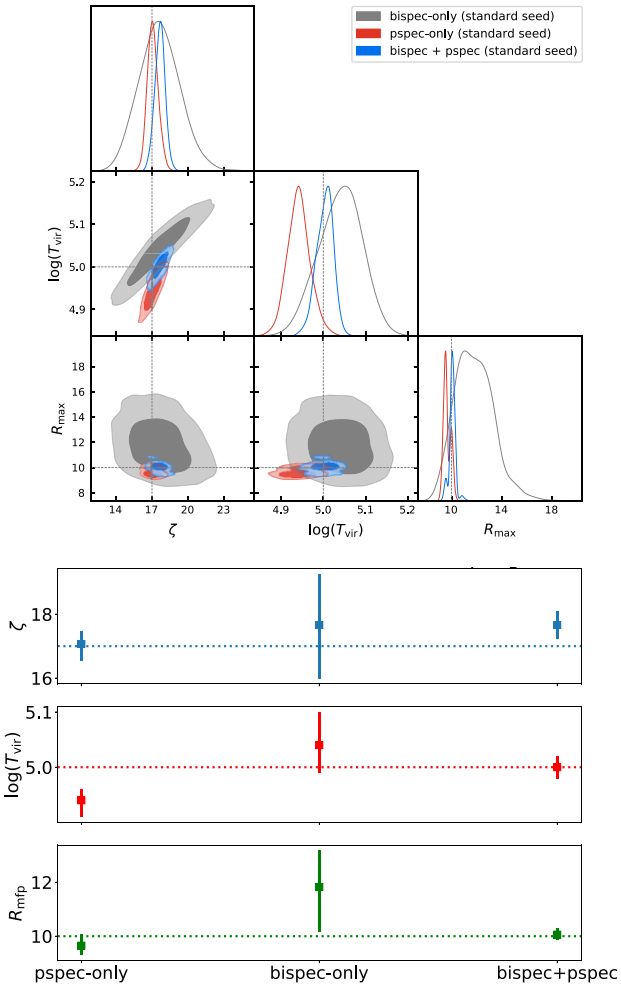

Figure 6. Corner plot (top) of credible intervals when using mock observed data generated using our late reionization model and the standard seed for the bispec-only (grey contours), the pspec-only (red contours), and bipsec+pspec (blue contours) as summary statistics in the likelihood. The blacked dashed lines indicate the parameter values used to generate the mock observed data sets for each model. The bottom plot shows the mean ±68 per cent credible intervals for each parameter. All include SKA-LOW (phase 1) instrumental effects (assuming negligible primary beam effects), as well as sample variance, which we model the associated standard deviation using MC methods and using the parameters of the fiducial model. While all cases contain the truth within their 95 per cent credible intervals, the posterior probability mass for the pspec-only case is concentrated in a different region of model parameter space, resulting in biased marginal statistics.

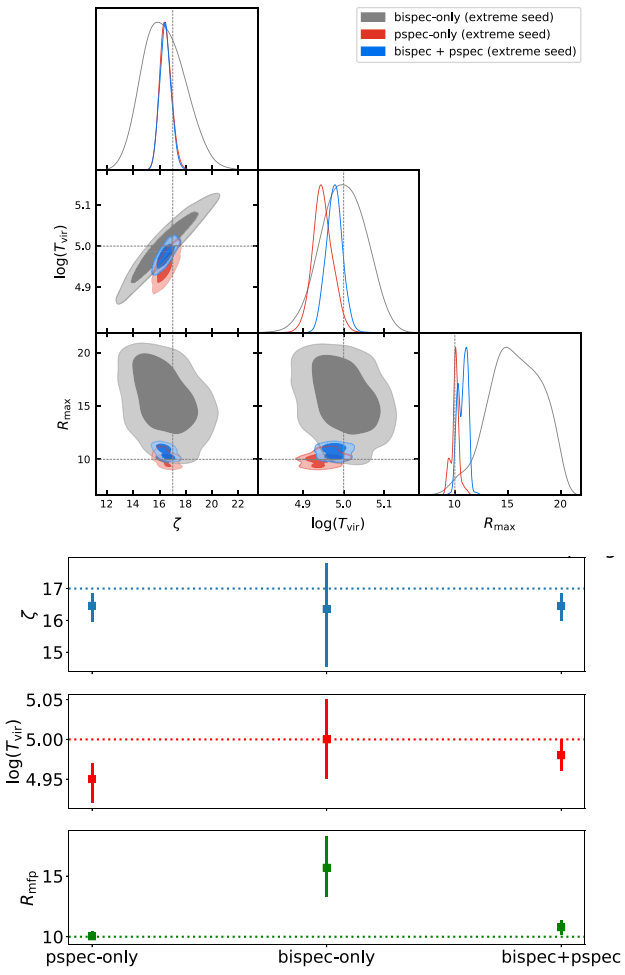

Figure 7. Corner plot (top) of credible intervals when using mock observed data generated using our late reionization model and the 'extreme' seed for the bispec-only (grey contours), pspec-only (red contours), and bipsec+pspec (blue contours) as summary statistics in the likelihood. The blacked dashed lines indicate the parameter values used to generate the mock observed data sets for each model. The bottom plot shows the mean ±68 per cent credible intervals for each parameter. All include SKA-LOW (phase 1) instrumental effects (assuming negligible primary beam effects), as well as sample variance, which we model the associated standard deviation using MC methods and using the parameters of the fiducial model. Here, the combining the bispectrum with the power spectrum still helps less in relieving bias in the credible intervals, although the marginal statistics are seen to return reasonable predictions of the true parameters.

them being very different models. This is likely because the sample-variance error for the later-reionization model is smaller or similar to that of the fiducial model because as the process of reionization is less advanced (the late-reionizaton's ionized fraction is only 0.7 at the lowest redshift we consider as opposed to 0.3 in the fiducial model). In the later stages of reionization (in the regime of sparse neutral islands), the amplitude varies more between realizations, as can be seen by the trend further away from the theoretical sample-variance with decreasing redshift in Fig. 2. This result implies that one could use the sample-variance from a single well-chosen model for all regions in parameter space. A better, and still tractable, option would be to sparsely sample the sample-variance error in parameter space and use some form of interpolation to approximate the sample-variance error in other regions of parameter space. Whether or not, this would be a sufficient approximation, and whether this finding extends to the full covariance matrix should be addressed in future work.

It is clear that using a diagonal covariance matrix and assuming independence between statistical bins are not disastrous assumptions in that the true parameters are constrained by the resulting parameter estimation analysis, even if we consider outlier data. However, it will give stronger and more robust results to not make such assumptions and to use a fully multivariate Gaussian likelihood that includes all correlations between the statistical bins, statistics, and redshifts. We have discussed the difficulty of accurately accounting for sample variance errors, it will equally be challenging to capture correlations between redshifts, which can be seen in Fig. 8, where we plot the

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

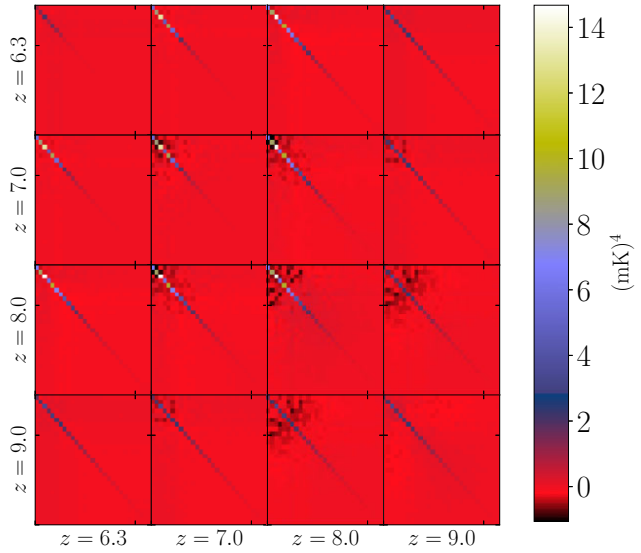

Figure 8. Covariance matrix for the power spectrum for all bins and redshifts considered here. We see that there are correlations between statistical bins in different redshifts, most notably between $z = 8$ and $z = 9$.

covariance matrix for the power spectrum.[7] To avoid correlations across redshift due to each independent realization sampling from the same initial conditions, we split our 2000 realizations into four distinct samples, each containing 500 realizations. This ensures that no two distinct redshift samples contain power spectra obtained from the same initial conditions. These correlations would likely be less severe if we were working with chunks of light cones, which is the more correct thing to consider; however, it is unlikely that there would be no correlations whatsoever. It is also likely that as the complexity of our forward model increases (necessary if we are to fully characterize the instrumental effects, foreground residuals, ionospheric effects, unresolved RFI, and polarization leakage) the assumption of a multivariate Gaussian form for the likelihood will be insufficient.

A method to bypass all these issues would be to use likelihood-free inference, which bypasses the need to ever pre-calculate a covariance matrix since the likelihood (or posterior depending on the type of likelihood-free inference) is estimated using forward simulations during the inference process. It also means one does not ever need to explicitly write down a likelihood function. This approach will also be able to deal with cross-correlations of the cosmological signal with the noise and foregrounds biasing parameter-inference results, as seen in Nasirudin et al. (2020) who perform far more accurate and detailed forward-modelling than that attempted here (they also use a fully multivariate Gaussian likelihood). We will discuss the application of likelihood-free methods as applied to 21-cm observations in Watkinson, Alsing, Greig & Mesinger (in prep).

5 CONCLUSION

In this work, we have added an isosceles bispectrum likelihood module to the established 21CMMC code that assumes independence between all statistical bins and redshifts. We are able to make this assumption by using a normalized version of the bispectrum in which the contribution of the power spectrum to the bispectrum has been removed. In order to perform our analysis, we use two new publicly available codes BIFFT (to measure the bispectrum with sufficient speed) and PYOBS21 (to simulate *uv* sampling and random instrumental noise for coeval cubes).

We generate various mock observations by varying astrophysical parameters, as well as the random seed for initial conditions. We consider two sets of astrophysical parameters, which result in different reionization histories: a fiducial model and a late reionization model. We also consider two different random seeds – one chosen to produce relatively standard bispectra (in terms of its χ^2 compared to the mean) and another to produce more extreme outlier bispectra data.

Various approaches for handling the bispectrum sample-variance error term have also been considered. We find that the bispectrum sample-variance error cannot be effectively described by propagating the power spectrum sample-variance error on to the bispectrum under the assumption of Gaussianity. We find that using the 1σ error generated using MC methods for a simple 1D Gaussian likelihood is sufficient to constrain the parameters of the three parameter model of reionization considered here. We also find that using the sample-variance error generated under our fiducial model while assuming simulated data from a late-reionization model has no serious negative impact on our results. This is important as it implies that we may be able to get away with a sparse sampling of the bispectrum sample-variance error as a function of parameter space combined with some form of interpolation to estimate the error term at the unsampled points of parameter space.

We find that combining the power spectra and the bispectrum in the likelihood can significantly reduce the bias away from the input reionization parameters, for all of the mock observations and models considered here (see also Gazagnes, Koopmans & Wilkinson 2021). For the late-reionization model, we also see a reduction in the credible limits. These findings hold true even if we consider outlier mock observations.

Further work is needed to establish the improvements from using the bispectrum in more complex models for reionization, such as the mass-dependent paramtrization including spin temperature fluctuations 21CMFASTV3 model. It will also be important for future works that consider the issue of modelling the bispectrum sample variance, to better understand its dependence on simulation resolution and dimensions. It will also be necessary to get a better understanding of how these results will be impacted by the inclusion of more levels of observational realism, as there has already been an indication that foreground residuals and observational effects will be more problematic for the bispectrum (Watkinson et al. 2021c).

ACKNOWLEDGEMENTS

CAW's research is currently supported by a UK Research and Innovation Future Leaders Fellowship, grant number MR/S016066/1 (PI Alkistis Pourtsidou). CAW also thanks Jonathan Pritchard for financial support during the early stages of this project via the European Research Council (ERC) grant number 638743-FIRSTDAWN, as well as the ARC Centre for Excellence in All-Sky Astrophysics in 3D Visitor Fund. This research utilized Queen Mary's Apocrita HPC facility, supported by QMUL ITS Research. http://doi.org/10.5281/zenodo.438045. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through

[7]We do not plot the correlations between the power spectrum and bispectrum, because the amplitude contribution has been normalized out of our bispectrum analysis; we therefore expect correlations between the two statistics to be negligible.

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

project number CE170100013. AM acknowledges support from the ERC under the European Union's Horizon 2020 research and innovation programmes (AIDA – #638809). The results presented here reflect the authors' views; the ERC is not responsible for their use.

DATA AVAILABILITY

The codes used to produce this work are all publicly available with links to access provided in the main text. The data underlying this article will be shared on reasonable request to the corresponding author

REFERENCES

Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R., 2001, Phys. Rep., 367, 1
Brillinger D., Rosenblatt M., 1967, in Harris B., ed., Spectr. Anal. time Ser. Wiley, New York, p. 189
Cramér H., 1946, Mathematical Methods of Statistics (PMS-9). Princeton Univ. Press, Princeton
Dewdney P., 2016, Technical report, SKA1 System Baseline Design V2. SKA Office, Manchester, UK
Fisher R. A., 1935, J. R. Stat. Soc., 98, 39
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Furlanetto S. R., Oh S. P., 2005, MNRAS, 363, 1031
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1
Gazagnes S., Koopmans L. V., Wilkinson M. H., 2021, MNRAS, 502, 1816
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Gorce A., Pritchard J. R., 2019, MNRAS, 489, 1321
Greig B., Mesinger A., 2015a, MNRAS, 449, 4246
Greig B., Mesinger A., 2017b, MNRAS, 472, 2651
Hinich M. J., Clay C. S., 1968, Rev. Geophys., 6, 347
Hinich M., Messer H., 1995, IEEE Trans. Signal Process., 43, 2130
Hinich M. J., Wolinsky M., 2005, J. Stat. Plan. Inference, 130, 405
Hutter A., Watkinson C. A., Seiler J., Dayal P., Sinha M., Croton D. J., 2019, MNRAS, 492, 653
Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725
Kaur H. D., Gillet N., Mesinger A., 2020, MNRAS, 495, 2354
Kim Y. C., Powers E. J., 1978, Phys. Fluids, 21, 1452
Koopmans L. et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14). SKA Organisation, Manchester, p. 1
Lewis A., 2011, J. Cosmol. Astropart. Phys., 10, 1475
Liguori M., Sefusatti E., Fergusson J. R., Shellard E. P. S., 2010, Adv. Astron., 2010, 64
Majumdar S., Pritchard J. R., Mondal R., Watkinson C. A., Bharadwaj S., Mellema G., 2017, MNRAS, 476, 4007
Mellema G. et al., 2013, Exp. Astron., 36, 235
Mesinger A., Furlanetto S. R., 2007, ApJ, 669, 663
Mesinger A., Furlanetto S., Cen R., 2010, MNRAS, 411, 955
Mondal R., Bharadwaj S., Majumdar S., 2016, MNRAS, 464, 2992
Murray S., Greig B., Mesinger A., Muñoz J., Qin Y., Park J., Watkinson C., 2020, J. Open Source Softw., 5, 2582
Nasirudin A., Murray S. G., Trott C. M., Greig B., Joseph R. C., Power C., 2020, ApJ, 893, 118
Park J., Mesinger A., Greig B., Gillet N., 2018, MNRAS, 484, 933
Pober J. C. et al., 2013a, AJ, 145, 65
Pober J. C. et al., 2014b, AJ, 782, 66
Qin Y., Mesinger A., Park J., Greig B., Muñoz J. B., 2020, MNRAS, 495, 123
Rao C., 1945, Bull. Calcutta Math. Soc., 37, 81
Scoccimarro R., 2015, Phys. Rev. D, 92, 083532
Scoccimarro R., Colombi S., Fry J. N., Frieman J. A., Hivon E., Melott A., 1998b, ApJ, 496, 586
Scoccimarro R., Sefusatti E., Zaldarriaga M., 2004a, Phys. Rev. D, 69, 1550
Sefusatti E., Crocce M., Scoccimarro R., Couchman H., 2016, MNRAS, 460, 3624
Shaw A. K., Bharadwaj S., Mondal R., 2019b, MNRAS, 487, 4951
Shaw A. K., Bharadwaj S., Mondal R., 2020, MNRAS, 498, 1480
Shimabukuro H., Yoshiura S., Takahashi K., Yokoyama S., Ichiki K., 2016, MNRAS, 468, 1542
Sobacchi E., Mesinger A., 2014, MNRAS, 440, 1662
Tegmark M., Taylor A., Heavens A., 1997, ApJ, 480, 22
Trott C. M. et al., 2019, Publ. Astron. Soc. Aust., 36, e023
Watkinson C. A., Majumdar S., Pritchard J. R., 2017a, MNRAS, 472, 2436
Watkinson C. A. et al., 2019b, MNRAS, 482, 2653
Watkinson C. A., Trott C. M., Hothi I., 2021c, MNRAS, 501, 367
Yoshiura S., Shimabukuro H., Takahashi K., Momose R., Nakanishi H., Imai H., 2015, MNRAS, 451, 266

APPENDIX A: THE SAMPLE VARIANCE OF THE POWER SPECTRUM

In Fig. A1 we plot the 2000 power-spectra realizations that we use to calculate our sample-variance errors. We also overplot the two random seeds used for mock observed data in this study. As with the bispectrum the extreme seed (purple solid like) is more than 1σ away from the more standard seed (blue dot–dashed line) for many bins, especially at the later stages of reionization, i.e. $z \leq 7$.

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022

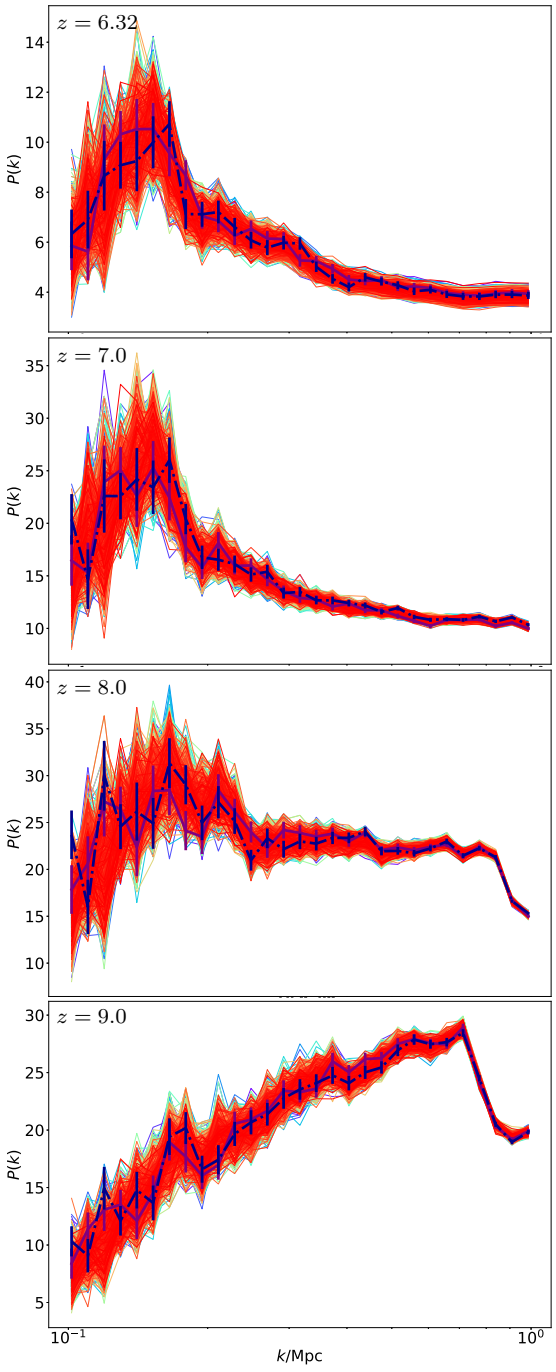

Figure A1. Here, we plot with thin lines all 2000 power spectra used in estimating the error due to sample variance for our simulation dimensions. The plots from top to bottom correspond to $z = [6.3, 7.0, 8.0, \text{and } 9.0]$. We overplot the two random seeds used in our parameter estimation analysis chosen from about 50 trial runs to minimize (54 321) and maximize (6937) the reduced χ^2 between them and the mean of the distribution of the thin lines shown by the thin lines in the plot.

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

Downloaded from https://academic.oup.com/mnras/article/510/3/3838/6498124 by Luisa Ferrini user on 18 March 2022