University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Greig, B;Ting, YS;Kaurov, AA

Title:
Exploring the cosmic 21-cm signal from the epoch of reionization using the wavelet scattering transform

Date:
2022-06-01

Citation:
Greig, B., Ting, Y. S. & Kaurov, A. A. (2022). Exploring the cosmic 21-cm signal from the epoch of reionization using the wavelet scattering transform. Monthly Notices of the Royal Astronomical Society, 513 (2), pp.1719-1741. https://doi.org/10.1093/mnras/stac977.

Persistent Link:
https://hdl.handle.net/11343/320453

License:
CC BY

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **513,** 1719–1741 (2022) https://doi.org/10.1093/mnras/stac977
Advance Access publication 2022 April 9

Exploring the cosmic 21-cm signal from the epoch of reionization using the wavelet scattering transform

Bradley Greig [1,2]★ Yuan-Sen Ting (丁源森)[2,3,4] and Alexander A. Kaurov[5]

[1] *School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*
[2] *ARC Centre of Excellence for All-Sky Astrophysics in 3 Dimensions (ASTRO 3D)*
[3] *Research School of Astronomy and Astrophysics, Australian National University, Cotter Road, Weston Creek, ACT 2611, Canberra, Australia*
[4] *Research School of Computer Science, Australian National University, Acton, ACT 2601, Australia*
[5] *Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA*

Accepted 2022 April 1. Received 2022 March 6; in original form 2021 December 14

ABSTRACT

Detecting the cosmic 21-cm signal during the Epoch of Reionization and Cosmic Dawn will reveal insights into the properties of the first galaxies and advance cosmological parameter estimation. Until recently, the primary focus for astrophysical parameter inference from the 21-cm signal centred on the power spectrum (PS). However, the cosmic 21-cm signal is highly non-Gaussian rendering the PS suboptimal for characterizing the cosmic signal. In this work, we introduce a new technique to analyse the non-Gaussian information in images of the 21-cm signal called the Wavelet Scattering Transform (WST). This approach closely mirrors that of convolutional neural networks with the added advantage of not requiring tuning or training of a neural network. Instead, it compresses the 2D spatial information into a set of coefficients making it easier to interpret while also providing a robust statistical description of the non-Gaussian information contained in the cosmic 21-cm signal. First, we explore the application of the WST to mock 21-cm images to gain valuable physical insights by comparing to the known behaviour from the 21-cm PS. Then we quantitatively explore the WST applied to the 21-cm signal by extracting astrophysical parameter constraints using Fisher Matrices from a realistic 1000 h mock observation with the Square Kilometre Array. We find that: (i) the WST applied only to 2D images can outperform the 3D spherically averaged 21-cm PS, (ii) the excision of foreground contaminated modes can degrade the constraining power by a factor of ∼1.5–2 with the WST and (iii) higher cadences between the 21-cm images can further improve the constraining power.

Key words: galaxies: high-redshift – intergalactic medium – cosmology: theory – dark ages, reionization, first stars – diffuse radiation – early Universe.

1 INTRODUCTION

Following recombination, the early Universe is rendered opaque to most forms of radiation owing to the pervasive neutral hydrogen fog. It remains in this state until the ignition of star-formation in the first primordial galaxies begins to locally ionize the neutral hydrogen, referred to as the Cosmic Dawn (CD). Over time, as these galaxies become larger, more abundant and clustered the cumulative ionizing radiation from these sources eventually renders the intergalactic medium (IGM) almost completely ionized, referred to as the Epoch of Reionization (EoR).

The most promising tool at our disposal for probing the EoR and CD is the 21-cm hyperfine spin-flip transition of neutral hydrogen. This signal, measured relative to a background radiation source, such as the Cosmic Microwave Background (see e.g. Gnedin & Ostriker 1997; Madau, Meiksin & Rees 1997; Shaver et al. 1999; Tozzi et al. 2000; Gnedin & Shaver 2004; Furlanetto, Oh & Briggs 2006; Morales & Wyithe 2010; Pritchard & Loeb 2012), can be seen in either emission or absorption depending on both the thermal and ionization states of the IGM. By measuring the time-evolution (frequency dependence) of the 21-cm radiation along with its spatial distribution in the IGM, we are able to build up a 3D movie of the IGM in the early Universe. With this, we can infer the typical properties of the galaxies responsible for driving the EoR and CD.

Detecting this 3D 21-cm signal requires large-scale radio interferometer experiments, which are specifically designed to be sensitive to the spatial fluctuations. These naturally operate in the Fourier domain, resulting in the power spectrum (PS) being the preferred statistic to characterize the 21-cm signal. However, the 21-cm signal is non-Gaussian owing to the complex nature of the ionization morphology. Thus, the PS is suboptimal as it completely disregards the important phase information in the signal (i.e. the locations of the ionized regions). For the first generation of interferometer experiments, such as the Low-Frequency Array (LOFAR; van Haarlem et al. 2013), the Murchison Wide Field Array (MWA; Tingay et al. 2013; Wayth et al. 2018), the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010), the Owens Valley Radio Observatory Long Wavelength Array (OVRO-LWA; Eastwood et al. 2019), and the upgraded Giant Metrewave Radio Telescope (uGMRT; Gupta et al. 2017) focusing on the PS is not a concern as these are likely sensitive to only a low signal-

★ E-mail: greigb@unimelb.edu.au

© 2022 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommonsorg/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

to-noise detection of the 21-cm signal. However, next-generation experiments such as the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al. 2017), NenuFAR (New extension in Nançay Upgrading loFAR; Zarka et al. 2012), and the Square Kilometre Array (SKA; Mellema et al. 2013; Koopmans et al. 2015) should yield sufficiently high sensitivity to be able to measure the 21-cm signal beyond the PS. Further, the SKA has been specifically designed to be able to provide the first 3D tomographic images of the 21-cm signal completely bypassing the need to analyse the signal in the Fourier domain.

Recognizing this, in recent years several alternative non-Gaussian probes of the 21-cm signal have been explored. The natural extension to the PS in the presence of non-Gaussian information is to consider the Bispectrum, the Fourier transform of the three-point correlation function. The sensitivity of the Bispectrum to the 21-cm signal has been extensively studied in the literature (e.g. Yoshiura et al. 2015; Shimabukuro et al. 2016; Majumdar et al. 2018; Watkinson et al. 2019; Hutter et al. 2020; Majumdar et al. 2020; Kamran et al. 2021) along with closely related approaches such as the triangle correlation function (focussing on the phase information only; Gorce & Pritchard 2019) and the position-dependent power spectrum (Giri et al. 2019a). In accessing this non-Gaussian information, the Bispectrum has been shown to improve over that from the PS for astrophysical parameter recovery from a mock 21-cm signal (Shimabukuro et al. 2017; Tiwari et al. 2021; Watkinson, Greig & Mesinger 2021).

Beyond the Fourier domain, numerous alternative approaches to analyse images of the 21-cm signal have been explored. These include the one-point statistics of the brightness temperature (Watkinson & Pritchard 2014; Shimabukuro et al. 2015; Kubota et al. 2016; Banet et al. 2021; Gorce, Hutter & Pritchard 2021), the morphological and/or topological features of the 21-cm signal (e.g. Yoshiura et al. 2017; Bag et al. 2019; Chen et al. 2019; Elbers & van de Weygaert 2019; Kapahtia, Chingangbam & Appleby 2019; Gazagnes, Koopmans & Wilkinson 2021; Giri & Mellema 2021; Kapahtia et al. 2021) and the distribution of the sizes of ionized regions (Kakiichi et al. 2017; Giri et al. 2018a; Giri, Mellema & Ghara 2018b; Giri et al. 2019b; Bianco et al. 2021). Alternatively, convolutional neural networks (CNNs) have been considered which are trained to be able to extract 2D or 3D features from images of the cosmic 21-cm signal and used to extract astrophysical information from mock observations (e.g. Gillet et al. 2019; Hassan et al. 2019; La Plante & Ntampaka 2019; Hassan, Andrianomena & Doughty 2020; Kwon, Hong & Park 2020; Mangena, Hassan & Santos 2020; Prelogović et al. 2021; Zhao et al. 2022).

In this work, we introduce a new technique sensitive to the non-Gaussian information contained in images called the Wavelet Scattering Transform (WST). Specifically for this work, we focus on the WST only in two dimensions. First introduced by Mallat (2012), the WST was used for signal processing of high-dimensional data sets and has since been applied to several different areas of astronomy, for example, for the interstellar medium (Allys et al. 2019; Regaldo-Saint Blancard et al. 2020; Saydjari et al. 2021), weak lensing (Cheng et al. 2020; Cheng & Ménard 2021b) and large-scale structure (Allys et al. 2020; Valogiannis & Dvorkin 2021). The WST applies a family of wavelet filters to an input image, extracting spatial features on different physical scales. Wavelets preserve the locality of the signal both spatially and in frequency (unlike the Fourier transform), ensuring the non-Gaussian information of the 21-cm signal is retained. By performing iterative convolutions of the input image by these filters, we measure the clustering of the spatial features, characterizing the non-Gaussian information. Importantly for this work, in using predefined values for the filters, the WST is capable of characterizing the global statistical properties of the signal by compressing the non-Gaussian information into a manageable set of scattering coefficients for performing robust statistical analyses.

While the bispectrum is also capable of measuring the non-Gaussian information, it suffers from the successive multiplication of the input information. As such, the variance of the statistic increases rapidly and is sensitive to outliers in the data. The WST is capable of circumventing this by taking the modulus after each successive filtering ensuring stability in the statistic (i.e. no divergence in the noise of the statistic). As a result, since the scattering coefficients are binning different Fourier modes, the variance of the scattering coefficients should approximately be Gaussian. The WST also shares features with CNNs, where wavelet filters are used to analyse the signal. In CNNs, the spatial scales and rotations of the wavelets are learnt by the training of the network, with the activation functions and pooling consistent with the modulus and reduction to scattering coefficients. On the other hand, since the WST uses a predefined set of wavelets we do not require any complex neural network training. Further, since they are predefined, the information is easier to interpret. For more in-depth discussions between these approaches we refer the reader to Cheng & Ménard (2021a).

First, to gain valuable insight into the WST and its application to the 21-cm signal, we apply the WST to a series of 21-cm simulations with different characteristics. Next, we progressively add in instrumental effects to mimic 'realistic' images expected from an SKA-like observation to ascertain whether the WST remains sensitive to the cosmological signal. Finally, we perform a Fisher Matrix analysis to quantitatively assess the performance of the WST applied to 21-cm images.

The remainder of this paper is organized as follows. In Section 2 we summarize the 21-cm simulations used in this work along with a description of the various astrophysical models. We then introduce the WST in Section 3 and apply it first to images of the raw, simulated 21-cm before considering it applied to realistic mock 21-cm images with the SKA. In Section 4 we perform Fisher Matrix forecasts for the WST applied to mock 21-cm images from an SKA-like observation. Finally, in Section 5 we provide a brief discussion before finishing with our closing remarks. Unless stated otherwise, all quantities are in comoving units and we adopt the cosmological parameters: $(\Omega_\Lambda, \Omega_M, \Omega_b, n, \sigma_8, H_0) = (0.69, 0.31, 0.048, 0.97, 0.81, 68\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$, consistent with recent results from the Planck mission (Planck Collaboration VI 2020).

2 SIMULATING THE 21-CM SIGNAL

In this work, we simulate the cosmic 21-cm signal using the seminumerical simulation code 21CMFAST[1] (Mesinger & Furlanetto 2007; Mesinger, Furlanetto & Cen 2011). In particular, we use the latest public release v3 (Murray et al. 2020) and the Park et al. (2019) astrophysical parametrization which provides prescriptions for the ultraviolet (UV) and X-ray properties of the galaxies responsible for reionization. Below, we briefly summarize the key ingredients for simulating the cosmic 21-cm signal with 21CMFAST, and defer the interested reader to the aforementioned publications for further details.

[1] https://github.com/21cmfast/21cmFAST

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

2.1 Galaxy UV properties

First, we assume that the stellar mass, M_*, of a typical galaxy is directly related to its host halo mass, $M_{\rm h}$ (e.g. Kuhlen & Faucher-Giguère 2012; Dayal et al. 2014; Behroozi & Silk 2015; Mitra, Choudhury & Ferrara 2015; Mutch et al. 2016; Ocvirk et al. 2016; Sun & Furlanetto 2016; Yue, Ferrara & Xu 2016; Hutter et al. 2020),

$$M_*(M_{\rm h}) = f_* \left(\frac{\Omega_{\rm b}}{\Omega_{\rm m}}\right) M_{\rm h}, \tag{1}$$

and that the fraction of galactic gas in stars, f_*, can also be related to the host halo mass through a power-law relation,[2]

$$f_* = f_{*,10} \left(\frac{M_{\rm h}}{10^{10}\,{\rm M}_\odot}\right)^{\alpha_*}. \tag{2}$$

Here, α_* is the power-law index and $f_{*,\,10}$ is the normalization of this expression evaluated for a dark matter halo of mass 10^{10} M$_\odot$, and both are free parameters in this model.

The star-formation rate (SFR) for these galaxies is then determined by dividing the stellar mass by a characteristic time-scale,

$$\dot{M}_*(M_{\rm h}, z) = \frac{M_*}{t_* H^{-1}(z)}, \tag{3}$$

where $H^{-1}(z)$ is the Hubble time and t_* is a free parameter allowed to vary between zero and unity.

Similar to above, the escape fraction of UV ionizing photons, $f_{\rm esc}$, is assumed to be a power law with halo mass

$$f_{\rm esc} = f_{\rm esc,10} \left(\frac{M_{\rm h}}{10^{10}\,{\rm M}_\odot}\right)^{\alpha_{\rm esc}}, \tag{4}$$

with power-law index, $\alpha_{\rm esc}$ and the normalization $f_{\rm esc,\,10}$ both being free parameters.

Finally, only some fraction of haloes contribute to reionization due to inefficient cooling and/or feedback processes in low mass haloes which prevent the formation of star-forming galaxies. To account for this a duty-cycle is included,

$$f_{\rm duty} = \exp\left(-\frac{M_{\rm turn}}{M_{\rm h}}\right) \tag{5}$$

with the fraction, $(1 - f_{\rm duty})$, accounting for the suppression of star-forming galaxies below some scale $M_{\rm turn}$ (e.g. Giroux, Sutherland & Shull 1994; Shapiro, Giroux & Babul 1994; Hui & Gnedin 1997; Barkana & Loeb 2001; Springel & Hernquist 2003; Mesinger & Dijkstra 2008; Okamoto, Gao & Theuns 2008; Sobacchi & Mesinger 2013a, b) which is also free parameter.

One of the key advantages of this parametrization is that galaxy UV luminosity functions (LFs), following some simple conversions, can be produced allowing the 21CMFAST parametrization to be constrained by observed high-z galaxy LFs.

2.2 Galaxy X-ray properties

It is thought that stellar remnants in the first galaxies emit X-rays capable of escaping the host galaxy and heating the IGM. This X-ray heating of the IGM is modelled by computing the cell-by-cell angle-averaged specific X-ray intensity, $J(\boldsymbol{x}, E, z)$, (in erg s^{-1} keV^{-1} cm^{-2} sr^{-1}),

$$J(\boldsymbol{x}, E, z) = \frac{(1+z)^3}{4\pi} \int_z^\infty {\rm d}z' \frac{c{\rm d}t}{{\rm d}z'} \epsilon_{\rm X} e^{-\tau}. \tag{6}$$

This is calculated by integrating the comoving X-ray specific emissivity, $\epsilon_{\rm X}(\boldsymbol{x}, E_e, z')$ back along the light-cone accounting for attenuation of the X-rays by the IGM given by $e^{-\tau}$. The specific emitted emissivity, $E_{\rm e} = E(1+z')/(1+z)$, is then,

$$\epsilon_{\rm X}(\boldsymbol{x}, E_{\rm e}, z') = \frac{L_{\rm X}}{\rm SFR} \left[(1+\bar{\delta}_{\rm nl}) \int_0^\infty {\rm d}M_{\rm h} \frac{{\rm d}n}{{\rm d}M_{\rm h}} f_{\rm duty} \dot{M}_* \right], \tag{7}$$

where the quantity in square brackets is the SFR density along the light-cone with $\frac{{\rm d}n}{{\rm d}M_{\rm h}}$ corresponding to the halo mass function (HMF)[3] and $\bar{\delta}_{\rm nl}$ is the mean, non-linear density in a shell centred on the simulation cell $(\boldsymbol{x}, z)$.

The emissivity is normalized by the specific X-ray luminosity per unit star formation escaping the host galaxies, $L_{\rm X}$/SFR (erg s^{-1} keV^{-1} $M_\odot^{-1}$ yr) which is assumed to follow a power law with respect to photon energy, $L_{\rm X} \propto E^{-\alpha_X}$. This is then normalization to the integrated soft-band (<2 keV) luminosity per SFR (in erg s^{-1} $M_\odot^{-1}$ yr),

$$L_{\rm X<2\,keV}/{\rm SFR} = \int_{E_0}^{2\,{\rm keV}} {\rm d}E_e \; L_{\rm X}/{\rm SFR}\,, \tag{8}$$

where E_0 corresponds to the energy threshold below which X-ray photons are absorbed by the host galaxy. This amounts to three free parameters for the X-ray properties, $L_{\rm X<2\,keV}$/SFR, E_0, and α_X, where for this work we assume $\alpha_X = 1$ consistent high-mass X-ray binary observations in the local Universe (Mineo, Gilfanov & Sunyaev 2012; Fragos et al. 2013; Pacucci et al. 2014).

2.3 Ionization and thermal state of the IGM

21CMFAST generates the velocity and evolved density fields using second-order Lagrange perturbation theory (e.g. Scoccimarro 1998). The ionization field is then obtained from the evolved density field using an excursion-set approach (e.g. Furlanetto, Zaldarriaga & Hernquist 2004). Here, the cumulative number of ionizing photons, $n_{\rm ion}$, are compared to the total number of neutral hydrogen atoms plus cumulative recombinations, $\bar{n}_{\rm rec}$ (e.g. Sobacchi & Mesinger 2014) in spheres of decreasing radii. A simulation cell is ionized when,

$$n_{\rm ion}(\boldsymbol{x}, z|R, \delta_R) \geq (1+\bar{n}_{\rm rec})(1-\bar{x}_e), \tag{9}$$

where the $(1-\bar{x}_e)$ factor corresponds to ionization by X-rays. The cumulative number of ionizing photons per baryon inside a spherical region of size, R and overdensity, δ_R is,

$$n_{\rm ion} = \bar{\rho}_b^{-1} \int_0^\infty {\rm d}M_{\rm h} \frac{{\rm d}n(M_h, z|R, \delta_R)}{{\rm d}M_{\rm h}} f_{\rm duty} \dot{M}_* f_{\rm esc} N_{\gamma/b}, \tag{10}$$

where $\bar{\rho}_b$ is the mean baryon density and the total number of ionizing photons per stellar baryon is given by $N_{\gamma/b}$[4]

The thermal state of the neutral IGM is then determined by self-consistently calculating the heating and cooling rates from structure formation, Compton scattering off CMB photons, and heating following partial ionizations as well as X-ray heating and ionizations. This sets the IGM spin temperature, $T_{\rm S}$, which is obtained

[2]This power-law behaviour between M_* and $M_{\rm h}$ is consistent with the mean behaviour recovered from both semi-empirical fits to observations (e.g. Harikane et al. 2016; Tacchella et al. 2018; Behroozi et al. 2019) and semi-analytic model predictions (e.g. Mutch et al. 2016; Yung et al. 2019; Hutter et al. 2020).

[3]Here, we adopt the Sheth–Tormen HMF (Sheth & Tormen 1999).

[4]By default 21CMFAST assumes $N_{\gamma/b} = 5000$, consistent with that of a Salpeter initial mass function (Salpeter 1955).

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

by the weighted mean of the gas, $T_{\rm K}$, and CMB, $T_{\rm CMB}$, temperatures,

$$T_{\rm S}^{-1} = \frac{T_{\rm CMB}^{-1} + x_\alpha T_\alpha^{-1} + x_{\rm c} T_{\rm K}^{-1}}{1 + x_\alpha + x_{\rm c}}, \tag{11}$$

and depends on both the local gas density and the intensity of the Lyman-α (Lyα) radiation impinging on the simulation cell. Here, x_α is the Wouthuysen-Field coupling coefficient (Wouthuysen 1952; Field 1958) and $x_{\rm c}$ is the collisional coupling coefficient between the free electrons and CMB photons. The Lyα background is formed by the contribution of X-ray excitations of neutral hydrogen atoms and the direct stellar emission of Lyman band photons by the first sources. For further details, see Mesinger et al. (2011).

2.4 21-cm brightness temperature

The cosmic 21-cm signal is measured as a brightness temperature fluctuation relative to the background CMB (e.g. Furlanetto et al. 2006):

$$\delta T_{\rm b}(\nu) = \frac{T_{\rm S} - T_{\rm CMB}(z)}{1+z}\left(1 - {\rm e}^{-\tau_{\nu_0}}\right)\ {\rm mK}, \tag{12}$$

where τ_{ν_0} is the optical depth of the 21-cm line,

$$\tau_{\nu_0} \propto (1+\delta_{\rm nl})(1+z)^{3/2}\frac{x_{\rm H\,I}}{T_{\rm S}}\left(\frac{H}{{\rm d}v_{\rm r}/{\rm d}r + H}\right). \tag{13}$$

This expression depends on the neutral hydrogen fraction, $x_{\rm H\,I}$, gas overdensity, $\delta_{\rm nl} \equiv \rho/\bar{\rho} - 1$, the Hubble parameter, $H(z)$, and the line-of-sight gradient of the peculiar velocity. Here, everything is evaluated at the redshift $z = \nu_0/\nu - 1$ and we have dropped the spatial dependence for brevity.

2.5 Simulation setup

To gain valuable insights into how the WST extracts astrophysical information from the cosmic 21-cm signal we consider four different reionization scenarios. For each, we simulate the full 21-cm light cone spanning $z = 5.9 - 35$ using 21CMFAST with a transverse comoving length of 250 Mpc and 128 voxels per side length (see Appendix A for discussions on convergence testing). In Table 1 we provide the astrophysical parameters for each of these four models, and provide a brief description of each below:

(1) *Fiducial model*: In this work, we adopt the default model considered in Park et al. (2019) as our fiducial model.[5] This was found to match a range of observational constraints including observed UV LFs at $z = 6 - 10$ and the electron scattering optical depth, $\tau_{\rm e}$.

(2) *Cold reionization*: Here, reionization proceeds as per our fiducial model above except we significantly increase the amplitude of the 21-cm signal by considering a scenario whereby the IGM undergoes very little or no X-ray heating. This class of model is referred to as 'cold' reionization (e.g. Mesinger, Ewall-Wice & Hewitt 2014; Parsons et al. 2014).

(3) *Large haloes*: We increase our characteristic turn-over scale for star-forming galaxies, $M_{\rm turn}$, to ensure reionization is driven by larger, more biased galaxies leading to a different ionization morphology than our fiducial model.

(4) *Extended reionization*: We decrease the ionizing efficiency of our star-forming galaxies to produce a slower, more extended reionization driven by a dominant population of faint star-forming galaxies. This results in a considerably later reionization ($z \sim 5$) while also further reducing the amplitude of the 21-cm signal.

The models above are primarily chosen to be illustrative of different features present in the 21-cm signal, not necessarily to be consistent with existing observational constraints. The resultant reionization history (IGM neutral fraction as a function of redshift) and the mean 21-cm brightness temperature signal (global signal) are shown in the top two panels of Fig. 1.

[5]Though see Qin et al. (2021) for a more up-to-date model using a broader range of recent observations.

3 APPLYING THE WAVELET SCATTERING TRANSFORM TO THE 21-CM SIGNAL

3.1 The wavelet scattering transform

The wavelet scattering transform (WST) is the convolution of a target field (image), $I(\boldsymbol{x})$, by a family of rotated and dilated wavelet filters. The key feature of wavelet filters is that they preserve locality both spatially and in frequency (unlike the Fourier Transform). Following the convolution operation, the modulus of the filtered field is calculated before spatially averaging to compress the information down to a single scattering coefficient. In this work, we specifically focus on applying the WST to 2D images, following closely the approach used by Cheng et al. (2020) which explored the WST for analysing the weak lensing signal in 2D images. In particular, we adopt the Morlet filter as our wavelet family. Note however, that the approach can be generalized to any dimension. It will be interesting to determine how well the WST performs in 3D with the 21-cm signal having a clear redshift dependence along the line of sight. However, initial exploration using wavelets by Trott (2016) highlighted the improvements in handling this redshift dependence relative to the power spectrum.

Given an input 2D image of the 21-cm signal, $I_0(x, y)$, the zeroth order scattering coefficient is simply the mean of the signal,

$$s_0 = \langle I_0(x, y)\rangle, \tag{14}$$

which, for the 21-cm signal, is simply the mean brightness temperature, $\delta\bar{T}_{\rm b}$, that is, the global 21-cm signal.

The family of first-order fields, $I_1(x, y)$, are obtained by the convolution of the original image by a family of wavelets, $\psi^{j_1,l_1}(x, y)$, where j and l correspond to different physical extents and rotations of the filters. Taking the modulus and then averaging we obtain the family of first-order scattering coefficients, $s_1^{j_1,l_1}$,

$$s_1^{j_1,l_1} = \langle\, |I_0 * \psi^{j_1,l_1}|\,\rangle. \tag{15}$$

We can then extend this to second-order by filtering our first-order fields by another set of filters, $\psi^{j_2,l_2}(x, y)$. Taking the modulus and averaging following the convolution of this second family of filters produces the second-order scattering coefficients, $s_2^{j_1,l_1,j_2,l_2}$,

$$s_2^{j_1,l_1,j_2,l_2} = \langle\, |I_1 * \psi^{j_2,l_2}|\,\rangle = \langle\, ||I_0 * \psi^{j_1,l_1}| * \psi^{j_2,l_2}|\,\rangle. \tag{16}$$

The key step above is the application of the modulus operation. Taking the modulus measures the strength of the spatial features prior to the secondary convolution and modulus operation. This measures the strength of the clustering of spatial features, which in effect is the non-Gaussian information. This same procedure can be followed to produce higher order coefficients, however, in this work we only consider the scattering coefficients up to the second order. By design, the information cascades very quickly, thus going to even higher order coefficients usually does not capture much additional information beyond the second order. The largest physical scale possible (denoted J) is set by the requirement that 2^J cannot

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

Table 1. A summary of the astrophysical parameters used for our four different reionization models used for exploring the WST applied to the 21-cm signal. See Section 2.5 for a description of each model.

Model type	$\log_{10}(f_{*,10})$	α_*	$\log_{10}(f_{\rm esc,10})$	$\alpha_{\rm esc}$	t_*	$\log_{10}(M_{\rm turn})$ ($M_\odot$)	$\log_{10}\left(\frac{L_{X<2keV}}{SFR}\right)$ (erg s^{-1} $M_\odot^{-1}$ yr)	E_0 (keV)
Fiducial model	−1.30	0.50	−1.00	−0.50	0.5	8.7	40.50	0.50
Cold reionization	−1.30	0.50	−1.00	−0.50	0.5	8.7	38.00	0.50
Large haloes	−0.70	0.50	−1.00	−0.50	0.5	9.9	40.50	0.50
Extended reionization	−1.65	0.50	−1.00	−0.50	0.5	8.0	40.50	0.50

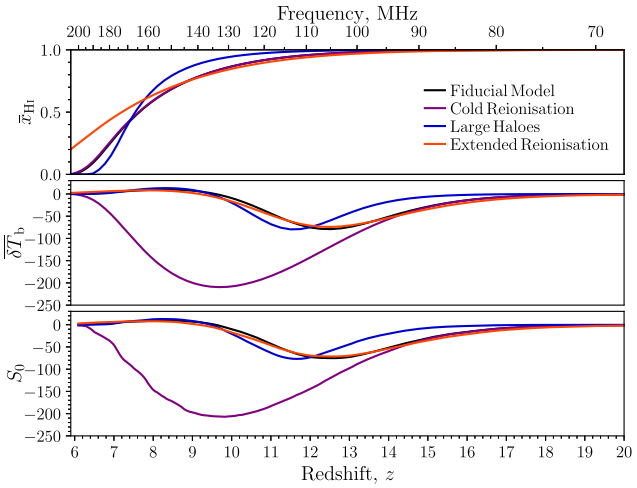


Figure 1. The globally averaged IGM neutral fraction, $\bar{x}_{\rm HI}$ (top panel), and the mean 21-cm brightness temperature, $\overline{\delta T_{\rm b}}$ (middle panel) as a function of redshift calculated from the four different astrophysical models outlined in Section 2.5. For comparison, we provide the evolution in the S_0 scattering coefficient recovered from the wavelet scattering transform in the bottom panel.

exceed the number of pixels in any one dimension of the image. The total number of available scattering coefficients is dictated by the number of filters used with J defining the maximal spatial scale and L, the total number of possible orientations (i.e. rotations) of the filters (each is rotated by π/L). Thus, with a total of $J \times L$ filters used there are $J^n\, L^n$ possible combinations of coefficients, with n being the order of scattering coefficients we are considering (i.e. in our case, $n = 2$).

In this work, we reduce the total number of scattering coefficients by averaging over all possible orientations of the filters. This assumes that the 21-cm signal is isotropic. That is, for the 21-cm signal we measure the following,

$$S_0 = s_0, \tag{17}$$

$$S_1^{j_1} = \left\langle s_1^{j_1,l_1} \right\rangle_{l_1}, \tag{18}$$

$$S_2^{j_1,j_2} = \left\langle s_2^{j_1,l_1,j_2,l_2} \right\rangle_{l_1,l_2}. \tag{19}$$

This reduces the total number of scattering coefficients down to a more manageable number of $1 + J + J^2$. For our simulation setup (128 voxels per side length), this corresponds to $J = 7$ and for convenience we only consider four distinct rotations, $L = 4$. The physical scales correspond to a dyadic sequence (2^j) resulting in seven logarithmically spaced bins ranging from ∼5.2 Mpc ($j = 0$) to ∼333.3 Mpc ($j = 6$). To obtain the scattering coefficients using the WST we use the convenient PYTHON package KYMATIO[6] (Andreux et al. 2018). For the second-order scattering coefficients KYMATIO by default only calculates those with $j_2 > j_1$ as the coefficients with $j_2 \leq j_1$ contain very little cosmological information as the 'high' frequency information (small physical scales) gets discarded first. Conveniently, this further reduces the total number of scattering coefficients down to $1 + 7 + 21 = 29$.

3.2 Filtering a single 21-cm image

As a first demonstration of the WST to the 21-cm signal, we apply the WST to a single 21-cm image from one of our four simulations. In Fig. 2 we show the 21-cm image at $z = 7.5$ ($\bar{x}_{\rm HI} \sim 0.5$) from our fiducial simulation (panel i), and the first and second-order filtered images for $l = 2$ (panels ii and iii, respectively) along with the corresponding scattering coefficients in the bottom left-hand corner of each panel (white text). Each filtered image is normalized to unity to amplify the spatial features rather than the amplitude of the signal (the scattering coefficients are calculated prior to the normalization).

We can see from Fig. 2 that as we increase j we are progressively becoming more sensitive to physically larger scale features in the 21-cm signal. This is evident as for $j < 4$, the original ionized regions (i.e. ionized bubbles) are still clearly present for the first-order filtered images. The second-order filters then pick up the clustering of these ionized bubbles of this same scale. Eventually, for increasing j we begin to consider scales larger than the individual ionized regions, becoming more sensitive to large-scale fluctuations in brightness temperature, driven primarily by thermal effects from X-ray heating. Importantly, this interpretation is only illustrative, as it is difficult to interpret the underlying physics of reionization from a single image. It is more illuminating to follow the evolution of the scattering coefficients as a function of redshift (frequency), which we investigate next.

3.3 WST of the raw, simulated 21-cm light-cone

To truly understand what we are learning about the EoR and CD from the WST, we switch from focusing on a single 21-cm image from one simulation to the full 21-cm light cone from the four different reionization simulations we discussed in Section 2.5. In doing so, we can more readily compare the performance of the WST against our established understanding of how we know the 21-cm signal behaves. For this, we apply the WST to images extracted from our simulated 21-cm light cone with a cadence of 1 MHz between 50–206 MHz (corresponding to $z = 27.4 - 5.9$). The lower end of this range (50 MHz) corresponds to the low frequency edge of the SKA observing band, while the upper end is dictated by the end of

[6] https://www.kymat.io/

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

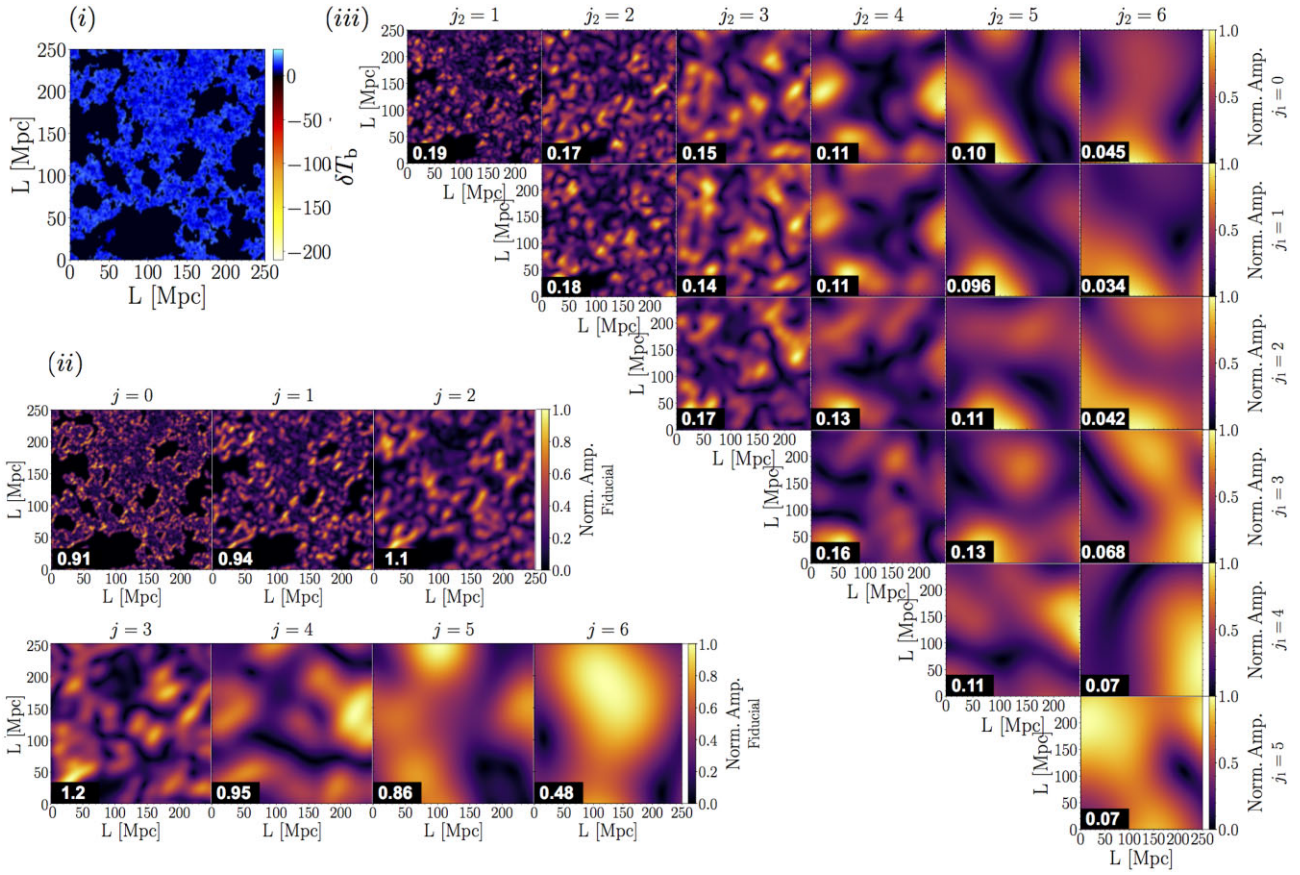


Figure 2. The wavelet scattering transform applied to a single, 21-cm image at $z = 7.5$ ($\bar{x}_{\rm H\,I} \sim 0.5$) from our fiducial model. (i): The simulated 21-cm brightness temperature output directly from 21CMFAST, (ii): The 21-cm signal filtered by the first-order family of wavelets in increasing order of spatial scale, j, (iii): The 21-cm signal filtered at second order, with the two scales represented by j_1 (rows) and j_2 (columns). Each filtered image is normalized to unity, with the white number in the bottom left-hand corner corresponding to the scattering coefficient prior to normalization (e.g. S_1^j for first-order and $S_2^{j_1,j_2}$ for second-order). For all, the filtered images assume $l = 2$ for the orientation.

reionization for our fiducial setup. For each image, we then calculate all relevant scattering coefficients.

3.3.1 S_0 coefficients

In Fig. 1 we compare the evolution of the IGM neutral fraction, $\bar{x}_{\rm H\,I}$, and the mean 21-cm brightness temperature, $\overline{\delta T_{\rm b}}$ against the S_0 coefficient for all four of our astrophysical models. Comparing the second and third panels, it is clear that the S_0 coefficient is equivalent to the mean 21-cm brightness temperature, which was to be expected from Section 3.1. Note, while the evolution of the S_0 coefficient appears to be less smooth than the mean brightness temperature, this is because they are computed differently. For $\overline{\delta T_{\rm b}}$ it is measured from each of the full 3D co-eval (single redshift) cubes used to construct the light cone. Whereas for S_0 we are determining it from a single 2D image. Thus for S_0 we are calculating the mean signal from considerably less information resulting in considerably more variation between images. Importantly for this work, a 21-cm interferometer experiment only measures the statistical fluctuations in the 21-cm signal, not the mean signal. Thus in practice, S_0 extracted from real 21-cm images will return a mean zero signal. This mean signal is what is observed by global signal experiments, such as the Experiment to Detect the Global EoR Signature (EDGES; Bowman & Rogers 2010).

3.3.2 S_1 coefficients

As discussed in-depth by Cheng et al. (2020), the S_1 coefficients can be closely related to the power spectrum. For example, for the 21-cm signal, the power spectrum, $P_{21}(\boldsymbol{k})$ is calculated by taking the square of the Fourier transformed fluctuations in the 21-cm brightness temperature, $P_{21}(\boldsymbol{k}) \propto \langle |\delta_{21}(\boldsymbol{k})|^2 \rangle$, where $\delta_{21}(\boldsymbol{k}) \equiv \delta T_{\rm b}(\boldsymbol{x})/\overline{\delta T_{\rm b}}(\boldsymbol{x}) - 1$. If we instead substitute the Morlet wavelet with a plane-wave (i.e. Fourier) filter, $\psi = {\rm e}^{-i\boldsymbol{k}.\boldsymbol{x}}$, and take the square of the resultant filtered image we obtain,

$$P(\boldsymbol{k}) \propto \langle\, |I_1|^2 \rangle \propto \langle\, |I_0 * {\rm e}^{-i\boldsymbol{k}.\boldsymbol{x}}|^2 \rangle. \tag{20}$$

Remembering that the first-order scattering coefficients are given by $s_1^{j,l} = \langle\, |I_0 * \psi^{j,l}| \,\rangle$, it is straightforward to see the similarities between the scattering coefficients and the Fourier power spectrum. Both measure the strength of the spatial fluctuations as a function of scale. Therefore, each of the S_1 scattering coefficients can be interpreted as coarsely binned power spectra. The key advantage of the WST is that by convolving the signal with a Morlet filter, the filtered signal retains both the structural and phase information in both real/redshift space and in wavelet space, as is evident in Fig. 2. In contrast, the Fourier transform de-localizes the signal. Subsequently, by convolving the wavelet filtered images iteratively with the same family of filters then enables the exploration of the clustering of

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

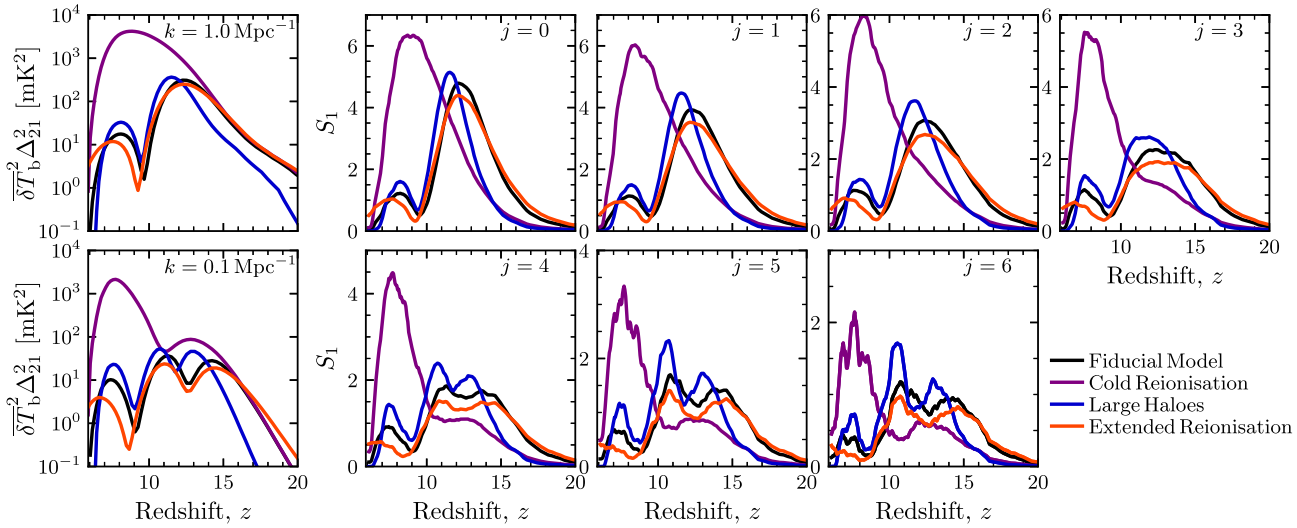


Figure 3. A direct comparison between the evolution of the 21-cm power spectrum and the recovered S_1 scattering coefficients. Left most column is the evolution in the 21-cm power ($k = 1.0$ and $0.1\ \mathrm{Mpc}^{-1}$, respectively) for the four different astrophysical models. In the remaining panels, we present the evolution of all possible S_1 scattering coefficients. Note that for the cold reionization model, we reduce the amplitude of these coefficients by a factor of 3 to more easily present the evolution of all four models.

the phase information (i.e. the second-order coefficients, discussed further in the next section).

It is well established that the evolution of the 21-cm power as a function of redshift shows three distinct peaks at large scales (i.e. $k = 0.1\ \mathrm{Mpc}^{-1}$, see e.g. Pritchard & Furlanetto 2007; Baek et al. 2010). These correspond to three important epochs; the EoR, the Epoch of Heating (EoH), and Wouthuysen-Field (WF) coupling (Wouthuysen 1952; Field 1958) in increasing order of redshift. In the left most panel of Fig. 3, we show the 21-cm power at two different spatial scales ($k = 1.0$ and $0.1\ \mathrm{Mpc}^{-1}$, respectively) for the four astrophysical models outlined in Section 2.5. Across these four models, the location of these peaks differ in both amplitude and redshift, dependent on the specifics of the astrophysics in each model. For example, the cold reionization model is $\sim$2 orders of magnitude larger than the other models, highlighting the impact that limited IGM heating can have on the amplitude of the 21-cm brightness temperature signal. Further, owing to the limited X-ray heating there is no EoH and thus there are only two peaks.

In the remaining panels of Fig. 3 we present the evolution of the S_1 scattering coefficients for the four different astrophysical models, in increasing order of physical scale (j) for the wavelet filters. As discussed in Section 3.2 the different j wavelet filters access different spatial scales, with low j corresponding to small physical scales and high j corresponding to large physical scales. Thus, the evolution of the S_1 scattering coefficients mirror the equivalent evolution of the 21-cm power at different Fourier scales. Low j scales are equivalent to large k Fourier modes, which is clearly evident as the evolution of the $k = 1.0\ \mathrm{Mpc}^{-1}$ mode (dashed curves) closely matches the low j scattering coefficient evolution. Equally the large j coefficients more closely resemble that of the smaller scale $k = 0.1\ \mathrm{Mpc}^{-1}$ Fourier mode, showing the distinctive three peaked structure for the cosmic 21-cm signal. Thus, the first-order scattering coefficients comparable information to the equivalent Fourier power spectrum.

It is important to note that for Fig. 3 and for the remainder of this work, we perform a simple box-car filtering to smooth out noise features for the evolution of the scattering coefficients. In effect, we perform a running average of each scattering coefficient over a 10 MHz range (we apply the WST on images at a 1 MHz cadence). In Appendix B we discuss this in more detail, highlighting the smoothed coefficients relative to the raw output (at 1 MHz cadence) in Fig. B1. Note, these noise like features could be reduced by considering images of larger physical extent.

3.3.3 S_2 *coefficients*

In the previous section, we highlighted the similarities between the S_1 coefficients and the Fourier power spectrum. Additionally, we highlighted the key advantage of the WST of the Fourier transform in that it always preserves both the structural and phase information. Following the WST properties outlined in Section 3.1, the S_2 coefficients are obtained from the subsequent convolution of the first-order family of filtered images after taking the modulus. If the S_1 coefficients are related to the Fourier power spectrum (i.e. the spatial information of the original input field), the S_2 coefficients are then related to the power spectrum of the first-order filtered images of the input field. Thus, the S_2 coefficients measure the spatial information (clustering) of the spatial information of the original field (see Cheng et al. 2020 for further details). In effect, these S_2 coefficients contain information up to the 4-point function of the input signal field. Thus, in analogy to the S_1 coefficients, S_2 can be thought of as containing information from coarsely binned bispectra and trispectra (i.e. the Fourier transform of the three and four point correlation functions). The advantage of the WST is that this binning of the higher order coefficients is considerably less noisy than that of the higher order polyspectra,[7] the bispectrum, and trispectrum (Cheng & Ménard 2021a). However, unlike for the power spectrum above, we do not demonstrate the common features to these higher order functions. Primarily, the difficulty arises from determining the equivalence between the WST binning and the appropriate closure configuration for the bispectrum or trispectrum. Such a study is beyond the scope of this work.

[7]Secondly, the WST is considerably cheaper in terms of total compute time than any higher order polyspectra as it does not require binning of the statistical information (only convolutions and modulus operations).

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

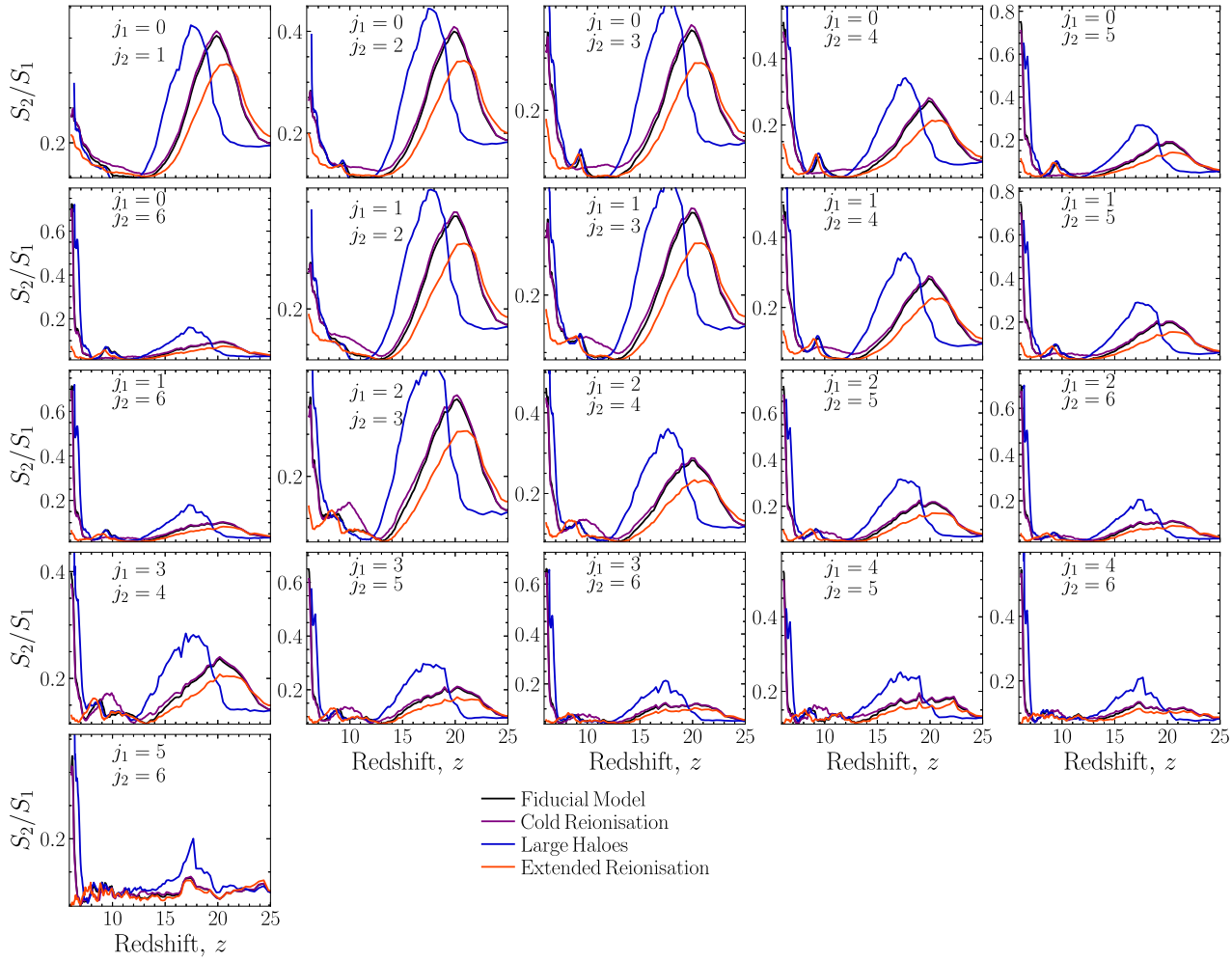


Figure 4. A comparison of the evolution for all combinations of the de-correlated S_2 coefficients (dividing out the dependence from the S_1 coefficients) for the four astrophysical models outlined in Section 2.5.

Equally, via the same properties of the WST, the S_2 coefficients are strongly correlated to the S_1 coefficients (since they are the filtered version following the modulus of S_1). Thus, if we present the evolution of the S_2 coefficients (as we have done for the S_1 coefficients above) these will be difficult to visually distinguish from the S_1 coefficients. Instead, we can produce de-correlated S_2 coefficients by simply dividing our $S_2^{j_1,j_2}$ coefficients by the corresponding first-order coefficient, $S_1^{j_1}$ (see e.g. Bruna et al. 2015; Cheng et al. 2020).

In Fig. 4 we present the evolution of the de-correlated S_2 coefficients as a function of redshift for the four different astrophysical models. These de-correlated coefficients contain notably different structural information to those from the S_1 coefficients as they are picking up information from the higher order moments, making it difficult to perform a similar qualitative analysis as we performed previously. Note, here we explicitly provide all possible combinations to highlight the available information. However, in practice it is quite clear many of these combinations display very similar behaviour. As such, for the remainder of this work we will only present a select subset of all possible combinations for the de-correlated S_2 coefficients.

Primarily, these de-correlated S_2 coefficients obtain the largest amplitudes towards the tail-end of reionization (lowest z), where the 21-cm brightness temperature is expected to be highly non-Gaussian (due to the complex morphology in the ionization field). The peaks at the highest redshifts ($z \sim 15 - 20$) are driven by the first appearance of the ionizing sources (i.e. the rarest, most clustered peaks in the density field) and their imprint on the IGM brightness temperature through the WF-coupling effect. The signal is strongest on the smallest physical scales (low j_1 and/or low j_2) with the peak location shifting to lower redshift from the extended reionization model down to the large haloes model. The extended reionization model has the smallest $M_{\rm turn}$ (reionization driven by smaller, fainter galaxies) meaning galaxies form first in this model, with the large haloes model (highest $M_{\rm turn}$) forming their first galaxies last. Thus, the additional information contained in the S_2 coefficients provides additional insight into the physical properties of the EoR and CD.

It is important to re-emphasize that here we are presenting the de-correlated S_2 coefficients, not the S_2 coefficients. This choice is motivated by wanting to highlight the additional non-Gaussian cosmological information contained in these coefficients beyond that held in the S_1 coefficients. Finally, if the S_2 coefficients truly contain the non-Gaussian information describing the 21-cm signal, we would expect that the features presented here in Fig. 4 should disappear once we randomize the phases in our filtered first-order images of the 21-cm signal. In Appendix C we verify that this is indeed the case.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

3.4 The WST applied to realistic mock images of the 21-cm signal

Previously, we have explored the WST applied directly to the raw simulated 21-cm signal. However, in practice from an interferometer experiment we will observe a corrupted 21-cm signal, which contains instrumental effects such as thermal noise and finite instrumental resolution. Thus, we now shift our attention to the WST applied to more realistic mock 21-cm images containing these effects to see if we are still sensitive to the 21-cm signal.

3.4.1 Instrumental effects

In this work, we adopt the SKA1–low as our fiducial 21-cm interferometer experiment. This choice is motivated by the expectation that the SKA will be able to provide 3D tomographic images during the EoR. We model the SKA antenna configuration using the design outlined in the SKA System Baseline Design document.[8] Here, 512 37.5 m antennae stations are distributed within a 500 m core radius and the system temperature, $T_{\rm sys}$, is modelled as $T_{\rm sys} = 1.1T_{\rm sky} + 40$ K. In this work, we adopt $T_{\rm sky} = 60\left(\frac{\nu}{300\ \rm MHz}\right)^{-2.55}$ K (Thompson, Moran & Swenson 2007). Images of the 21-cm signal are achieved by phase-tracking the sky, for which we assume a single 6 h track per night for a total observing time of 1000 h.

To construct our mock images of the 21-cm signal using SKA1–low, we modify the publicly available PYTHON module 21CMSENSE[9] (Pober et al. 2013, 2014). Nominally, 21CMSENSE uses the gridded *uv*-visibilities of any instrument design to yield an estimate of the thermal noise power spectrum, $P_{\rm N}(k)$, given an input 21-cm power spectrum measurement according to the following expression,

$$P_{\rm N}(k) \approx X^2Y\frac{\Omega^{'}}{2t}T_{\rm sys}^2. \tag{21}$$

Here, X^2Y is the cosmological conversion between observing bandwidth, frequency, and comoving distance, $\Omega^{'}$ is a beam-dependent factor derived in Parsons et al. (2014), t is the total time spent by all baselines within a specific Fourier mode and $T_{\rm sys}$ is the system temperature.

We model the effects of finite instrument resolution and thermal noise from our modified version of 21CMSENSE by performing the following steps:

(i) we first 2D Fourier transform the input (simulated) image of the 21-cm signal

(ii) we then filter our image using the output gridded *uv*-visibilities for the SKA1–low baseline configuration from 21CMSENSE. Cells with finite *uv*-coverage are multiplied by unity, all others are set to zero

(iii) in each cell we then calculate the thermal noise power, $P_{\rm N}(k_x, k_y)$, using equation (21), where k_x and k_y are the two transverse (on sky) directions and add a randomly sampled value from this measured power spectrum to each cell to mimic the effect of thermal noise

(iv) finally, we inverse Fourier transform back to image space to obtain our mock 21-cm image.

[8] http://astronomers.skatelescope.org/wp-content/uploads/2016/09/SKA-TEL-SKO-0000422_02_SKA1_LowConfigurationCoordinates-1.pdf

[9] https://github.com/jpober/21cmSense

3.4.2 Wedge mode removal

Unfortunately, observations of the 21-cm signal are hit by another source of error. The visibilities (*uv* coverage) of an interferometer experiment are frequency dependent. This means line-of-sight (frequency dependent) power can be thrown into the transverse (frequency independent) Fourier modes. This leakage of power results in a well-defined contaminated 'wedge' in cylindrical 2D Fourier space (Datta, Bowman & Carilli 2010; Morales et al. 2012; Parsons et al. 2012; Trott, Wayth & Tingay 2012; Vedantham, Shankar & Subrahmanyan 2012; Thyagarajan et al. 2013; Liu, Parsons & Trott 2014a, b; Thyagarajan et al. 2015a, b; Pober et al. 2016; Murray & Trott 2018). While it is possible to mitigate or 'clean' these contaminated modes (see e.g. Chapman & Jelić 2019 for a comprehensive review, or by using machine learning Gagnon-Hartman et al. 2021) in this work we take the conservative approach of cutting all contaminated 'wedge' modes from our 21-cm light cone.

The location of this 'wedge' in 2D cylindrical space is determined by,

$$k_{\parallel} = mk_{\perp} + b, \tag{22}$$

where $k_{\parallel}$ and $k_{\perp}$ are the line-of-sight and transverse Fourier modes, b is a additive buffer of $\Delta k_{\parallel} = 0.1\,h$ Mpc^{-1} extending beyond the horizon limit and,

$$m = \frac{D_{\rm C}H_0E(z)\sin(\theta)}{c(1+z)}. \tag{23}$$

Here, $D_{\rm C}$ is the comoving distance, H_0 is the Hubble constant, $E(z) = \sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_{\Lambda}}$, and $\sin(\theta)$ accounts for the viewing angle of the telescope, for which we conservatively take as $\theta = \pi/2$ (i.e. a zenith pointing observation).

To mimic the removal of these 'wedge' contaminated modes, prior to the steps outlined in Section 3.4.1 we first:

(i) Construct a 3D volume of the 21-cm signal centred on our current observing frequency.

(ii) Fourier transform, apply the instrument effects, and thermal noise as per the steps outlined in Section 3.4.1.

(iii) Excise all modes below the foreground contaminated 'wedge' as defined by equation (22).

(iv) Inverse Fourier transform back.

3.4.3 Realistic mock 21-cm image

In Fig. 5 we provide a demonstration of this pipeline applied to a simulated 21-cm signal at 150 MHz ($z \sim 8.5$) from our fiducial model. The left-hand panel shows a 2D image of the raw, simulated 21-cm signal, the middle panel contains an idealized mock image following the steps outline in Section 3.4.1 (e.g. perfect foreground removal) and in the right-hand panel we show a realistic mock signal with the foreground 'wedge' modes removed following Sections 3.4.1 and 3.4.2.

Importantly, interferometer experiments do not measure the mean signal, instead they only measure the amplitude of the spatial fluctuations. In effect, we observe an image with the mean signal removed. Applying our pipelines above takes this into account (i.e. after Fourier transforming we remove the DC mode), thus, the mock images shown in the middle and left of Fig. 5 only show the amplitude of the fluctuations around the mean. This has important consequences for the S_0 coefficients (i.e. the mean signal) since these will not be accessible from images of the 21-cm signal obtained from

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

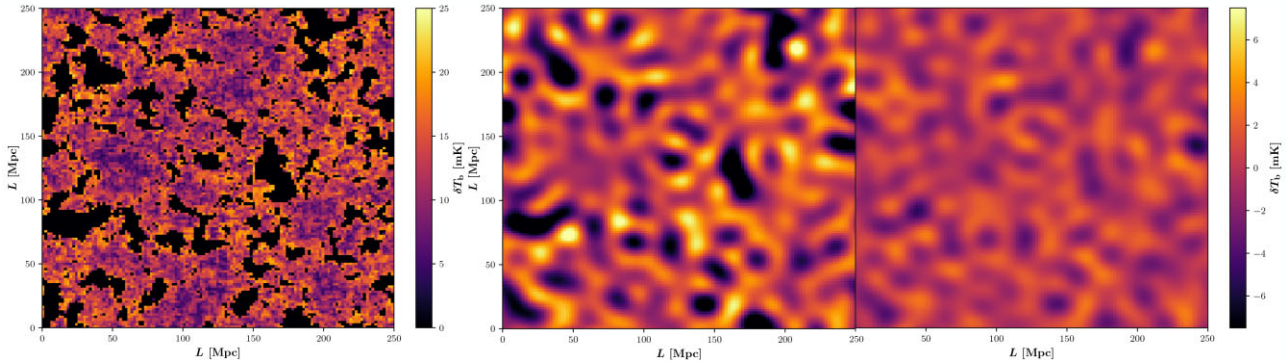


Figure 5. Schematic highlighting the incremental steps towards obtaining 'realistic' mock 21-cm images from a 21-cm interferometer experiment such as SKA1–low, assuming a 1000 h observation. Left-hand panel: the 21-cm signal extracted directly from our simulations, middle panel: the 21-cm signal including thermal noise and finite resolution by applying the SKA *uv*-coverage (see Section 3.4.1) and in the right-hand panel: the 21-cm signal from the middle panel but with contaminated foreground 'wedge' modes removed (see Section 3.4.2).

radio interferometers.[10] The effective removal of the mean 21-cm signal from the image also impacts on the higher order scattering coefficients. That is, their relative amplitudes reduce following mean removal.

The impact of finite resolution and noise from a realistic 21-cm experiment is clearly shown in the middle panel of Fig. 5. All small-scale features below our instrumental resolution are lost, leaving only the moderate to large-scale information. Importantly though, in terms of reionization studies, for the most part the ionization structure (ionized regions) are still identifiable. This implies that we would expect low j scattering coefficients to be heavily affected and/or dominated by the instrumental effects, with the higher j scattering coefficients being more sensitive to the 21-cm signal. However, this assumes perfect foreground removal (i.e. no loss of structural information). In the final panel, we see the impact of removing these contaminated foreground 'wedge' modes which notably further reduces the amplitude and structural information in the resultant mock image.

3.5 Evolution of WST scattering coefficients under instrumental effects

Previously, in Section 3.3 we explored the evolution of the WST scattering coefficients as a function of redshift to ascertain the utility of the WST applied to the 21-cm signal. Here, we return to this, now investigating the evolution of the scattering coefficients from realistic mock 21-cm images from an experiment such as the SKA1–low. Unlike previously, where we explored four different astrophysical models, here we only focus on our fiducial model, as we are specifically interested in how the instrumental effects impact our ability to infer information on reionization. Further, we apply the WST to 30 different realizations of our 21-cm signal from our fiducial model to obtain the mean evolution in the scattering coefficients. This choice is motivated by: (i) taking into account the sample variance, which will be necessary for our astrophysical forecasting in the next section and (ii) to better estimate the mean thermal noise obtained from various thermal noise realizations used in our mock images.

In Fig. 6, we compare the mean evolution of all S_1 scattering coefficients after applying the WST to the 21-cm signal. The black curves correspond to the WST applied directly to the raw, simulated 21-cm signal (i.e. no instrumental effects). The red and blue curves include both thermal noise and the expected SKA1–low *uv*-coverage, with the latter also including the removal of foreground contaminated 'wedge' modes. Finally, to understand the noise floor from thermal noise, the purple curve corresponds to the mean evolution in the S_1 scattering coefficients including only thermal noise (i.e. no 21-cm signal). This effectively sets the lower bound for the amplitude of the scattering coefficients, with the cosmological information needing to be above this level to be differentiable from the noise.

As established previously, by including thermal noise and the expected SKA1–low *uv*-coverage, the amplitude of the S_1 coefficients for the low j scattering coefficients are considerably impacted. We lose most of the signal on spatial scales below the SKA resolution (as seen by the precipitous drop in amplitude for $j \leq 2$). As we increase the spatial scales of our wavelet filters (increase j), the impact of the instrumental effects are minimized. For $j \gtrsim 3$, the scattering coefficients including the instrumental effects closely match those from the raw simulation output, and are still sensitive to the three peaks denoting the various epochs of the cosmic 21-cm evolution (EoR, EoH, and WF-coupling). Thus, the WST is relatively unaffected by instrumental effects on moderate to large spatial scales (as is the case for the 21-cm PS, see e.g. Greig & Mesinger 2015) for realistic images of the 21-cm signal. It is these scales that are important for astrophysical parameter inference.

Once we remove the foreground 'wedge' modes, the amplitude of the S_1 scattering coefficients is drastically reduced across all spatial (j) scales. Further, we can no longer differentiate between the two peaks in the cosmic 21-cm signal arising from the EoH and WF-coupling eras. Nevertheless, despite the amplitude of the S_1 coefficients dropping by $\sim$2–3 relative to the red curves, the amplitude remains above that set by the thermal noise floor. Though, it does reach the noise floor earlier and for a wider range of redshifts. Importantly, this implies that cosmological information is still accessible from the S_1 scattering coefficients applied to realistic mock images of the 21-cm signal.

In Fig. 7, we perform the same comparison, this time for the de-correlated S_2 coefficients. Here, we specifically limit ourselves to a small subset of S_2 coefficients noting that many exhibit similar behaviours as seen in Fig. 4. Similar to above, we find that for any S_2 coefficients on small spatial scales (i.e. low j_1 or j_2) the instrumental effects dominate over the cosmological signal. Similarly, on large spatial scales, we find that the S_2 coefficients,

[10] We verified this by applying the WST to these mock images and recover $S_0 = 0$ for all redshifts.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

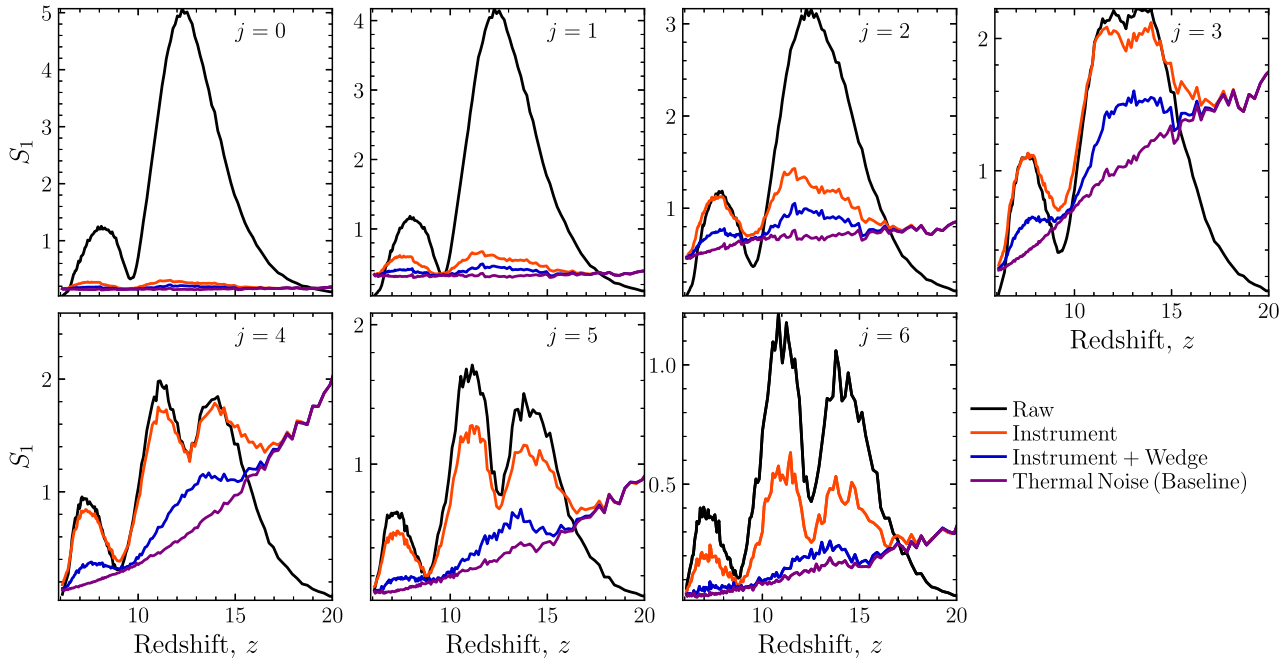


Figure 6. The evolution in the S_1 scattering coefficients in the presence of instrumental effects as a function of redshift for our fiducial model averaged over 30 different simulated 21-cm realizations. We assume a 1000 h observation with SKA1–low. The black curve corresponds to raw, simulated 21-cm images (no instrumental effects), the red curve includes thermal noise and the SKA *uv*-coverage, the blue curve includes thermal noise, the SKA *uv*-coverage and the removal of foreground contaminated ('wedge') modes and finally the purple curves are pure thermal noise realizations only (including SKA1–low resolution).

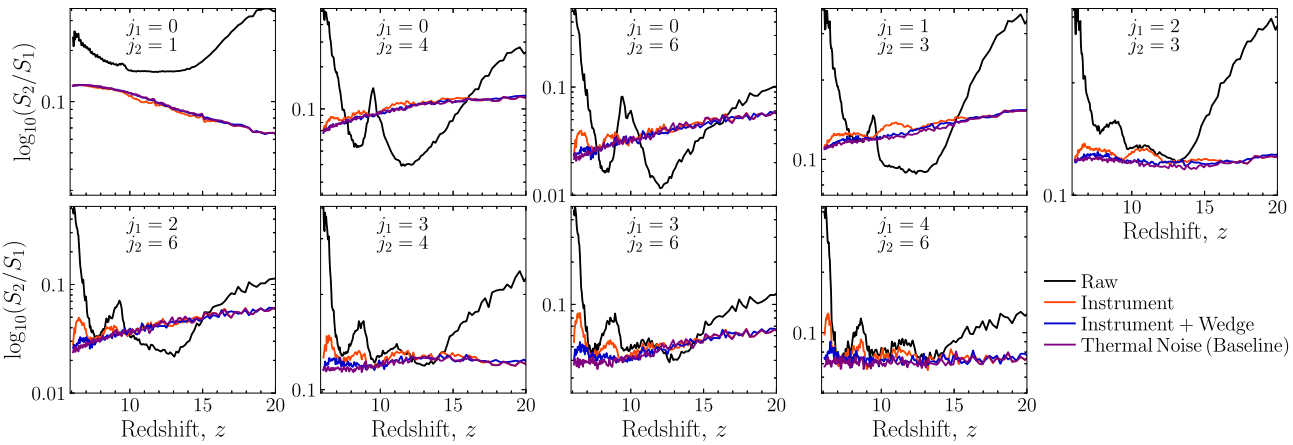


Figure 7. The same as Fig. 6, except for the all the de-correlated S_2 coefficients. Here, we only present a subset of all possible scattering coefficients.

even in the presence of all sources of instrumental effects, remain above the thermal noise level implying the possibility of being able to detect the non-Gaussian information from the 21-cm signal in amongst the noise. However, this appears to only be the case at $z < 10$, and predominately at $z \lesssim 8$. This implies that most of the additional non-Gaussian information that we can obtain from realistic 21-cm images using the WST will come during the EoR, where the non-Gaussianities are strongest due to the complex ionization morphology. Finally, we re-emphasize that here we are presenting the de-correlated S_2 coefficients to highlight the non-Gaussian information beyond that held by the S_1 coefficients. As discussed previously, the actual S_2 coefficients are larger in amplitude as they retain the spatial information held by the S_1 coefficients.

4 ASTROPHYSICAL PARAMETER FORECASTS

Thus far, we have only qualitatively explored the WST applied to the 21-cm signal, gaining physical insight and intuition by comparing the evolution of the scattering coefficients as a function of redshift. In this section, we perform a quantitative analysis of the WST in the context of astrophysical parameter inference, determining how well we can constrain our input astrophysical model using our realistic mock images of the 21-cm signal from the SKA1–low. For the remainder of this work, we focus on the fiducial model outlined in Section 2.5.

While it is plausible to explore the WST applied directly to the 21-cm signal from 3D simulations in a fully Bayesian framework using 21CMMC (Greig & Mesinger 2015, 2017; Greig & Mesinger 2018; Park et al. 2019), for simplicity we instead perform a Fisher Matrix

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

analysis. Primarily, this choice is motivated by several uncertainties in the construction of the likelihood function for the WST, which can be more trivially accounted for with a Fisher Matrix. For example, determining the cosmic (sample) variance (and associated covariance matrix) for each of the scattering coefficients requires numerical estimation from many realizations. While this is tractable for the Fisher Matrix, as we only need to estimate it for our mock observation, for direct MCMC we would need to determine this for each set of sampled astrophysical parameters. In the future, we will return to this and perform a more robust exploration of the WST with 21CMMC.

4.1 Fisher matrices

The Fisher Information Matrix (Fisher 1935) estimates the amount of information contained about a model parameter, θ_i, given some observational measurement. In our case, we can write the Fisher Matrix as,

$$\mathbf{F}_{ij} \equiv -\left\langle \frac{\partial^2 \ln(\mathcal{L})}{\partial\theta_i \partial\theta_j} \right\rangle = \sum_z \frac{\partial^2}{\partial\theta_i \partial\theta_j} \mathbf{S}^{\mathbf{T}} \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{S}, \tag{24}$$

where we sum over the redshifts, z, of each individual 2D image, $\mathcal{L}$ is the likelihood function (the probability density of the model given our input parameters), $\mathbf{\Sigma}$ is the covariance matrix containing the cosmic variance and thermal noise errors on each of the scattering transform coefficients and $\mathbf{S}$ contains all possible scattering coefficients (i.e. first and second order). In defining the Fisher Matrix in this way, we assume that each observational measurement of the scattering transform coefficients are statistically independent in redshift (frequency). Thus, we sum the information over all possible redshifts for which we extract a 21-cm image.

We obtain parameter forecasts for our individual astrophysical parameters by calculating the covariance matrix, $\mathbf{C}_{ij}$, which is simply given by the inverse of the Fisher Matrix. The resultant confidence intervals obtained through this method are considered as lower bounds on the true precision (i.e. best possible constraints) given the observational setup.

4.2 Comparing the WST to the 21-cm power spectrum

First, in order to be able to provide a baseline from which to compare the performance of the WST against, we compute the Fisher Matrix using the 3D spherically averaged 21-cm PS. While we could theoretically compare the 21-cm PS extracted from a 2D 21-cm image to more directly compare to the 2D WST, we refrain from doing so as it is not what will be explored in practice. Instead, in this work, we want to compare the WST to the established literature, for which the 3D spherically averaged 21-cm PS is the benchmark as it is the most well explored and understood. Alternatively, we could instead apply the WST in 3D, however, interpreting the scattering coefficients will be less trivial owing to the redshift dependence along the line of sight. However, in future, we will return to this.

It is important to note that in making this choice we cannot perform a straight-forward like-for-like comparison. As we will be comparing the 2D WST to the 3D spherically averaged PS, inherently there will be more statistical information available with the 3D PS as it will average over more structural information also leading to a smaller sample variance error due to the extra information. Further, by measuring the 21-cm signal across a bandwidth with the 3D PS, the instrumental errors (i.e. noise) will also reduce down for the PS compared to that of a single 2D image. Nevertheless, the comparison serves to highlight our understanding that astrophysical inference from the first-order scattering coefficients behaves in a similar way to the 21-cm PS. Further, from this, we can then determine how well things improve when considering the additional non-Gaussian information from the S_2 coefficients in Section 4.3.

4.2.1 *21-cm power spectrum Fisher matrix*

For the mock observation using the 21-cm PS, we closely follow the approach of Park et al. (2019). We break the observed 21-cm light cone along the line-of-sight (redshift) direction into equal comoving volumes equivalent in size to our 250 Mpc simulation. For each chunk of the 21-cm light-cone we then compute the 3D spherically averaged 21-cm PS resulting in 12 independent measurements of the 21-cm PS[11] between $z \sim 5.9 - 27.4$. Using 21CMSENSE, we determine the observational errors (sample variance and thermal noise) on our mock observation of the 21-cm PS assuming a 1000 h observation with the SKA1–low. Combining these sources of error in quadrature, we then compute our Fisher matrix for the 21-cm PS using a similar format to that expressed in equation (24). Following Park et al. (2019), we restrict our Fisher matrix computation from the 21-cm PS to eight linear[12] spaced Fourier bins between $k = 0.1 - 1.0$ Mpc^{-1}, with the lower limit set by astrophysical foreground contamination and the upper limit being shot-noise within the 21-cm simulations.

Importantly for this comparison, unlike Park et al. (2019) we do not consider the modelling uncertainty term. This approximately quantifies how much the simulated 21-cm PS obtained from a seminumerical simulation differs to that from a full radiative transfer simulation using the same initial conditions. In order to be able to perform a fair comparison between the WST and the 21-cm PS we must ignore this source of error as we do not have an equivalent estimate for the modelling uncertainty for the WST. In the future, we will return to this. Additionally, this implies we cannot perform a like-for-like comparison between our Fisher Matrix estimates for the 21-cm PS and the direct MCMC performed in Park et al. (2019) (their 21-cm only case). However, while not shown here, we verified that when including this modelling uncertainty, our Fisher matrix framework recovered comparable constraints to those presented in Park et al. (2019). Equally, we recovered similarly good agreement with the 21-cm Fisher matrix code, 21cmFish (Mason et al., in preparation).

4.2.2 S_1 *scattering coefficients Fisher matrix*

In Section 3.3.2, we highlighted that the first-order scattering coefficients are analogous to the 21-cm PS. Here, we quantify this by estimating a Fisher matrix using just the S_1 coefficients. In order to be able to perform this comparison, we limit the number of 21-cm images used for the WST to ensure we have a comparable number of independent measurements. Specifically, we extract the scattering coefficients from 21-cm images extracted at the central redshift (frequency) of the 12 equal comoving cubes of the 21-cm

[11]The central redshift for each of these individual pieces of the 21-cm light-cone correspond to: $z = 6.3$, 7.0, 7.7, 8.6, 9.6, 10.8, 12.2, 13.8, 15.8, 18.1, 21.0, and 24.7.

[12]This choice to linearly sample the bins in Fourier space is to increase the sampling towards moderate to large spatial scales (i.e. small k), where there is increased sensitivity to the astrophysical information. However, over the range in which we compute this, the differences between this and logarithmically sampling are minimal.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

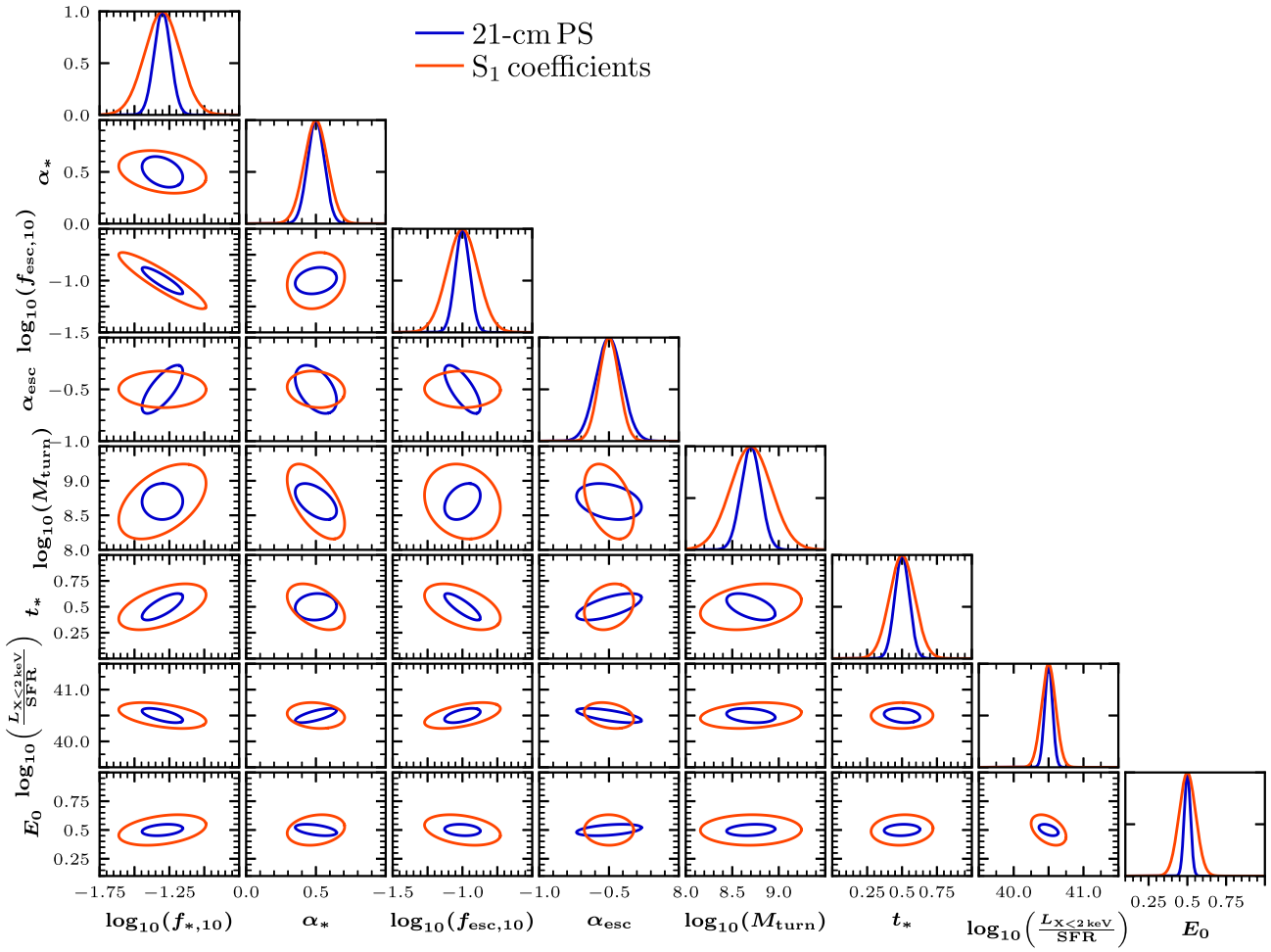


Figure 8. The recovered confidence intervals for our astrophysical parameters from the fiducial model estimated using Fisher Matrices assuming a 1000 h observation with SKA1–low. The 1D probability distribution functions for each parameter are shown along the diagonal, whereas the 2D dimensional confidence intervals are shown below. Here, these contours correspond to the 95 percentiles. The blue curves correspond to the constraints using the 21-cm power spectrum while the red curves correspond to the S_1 scattering coefficients obtained from the 2D WST.

light cone see previous section. Further, for our simulation setup, we have seven S_1 coefficients, compared to the eight 21-cm PS bins per redshift. Thus, we have a comparable number of data points across the two approaches.

For this comparison, we only consider the impact of the SKA1–low instrument resolution and thermal noise on our mock 21-cm images (i.e. following Section 3.4.1) leaving an exploration of the impact of excising the foreground contaminated 'wedge' modes until Section 4.4. To compute the Fisher matrix for the S_1 coefficients (and subsequently all WST Fisher matrices), we first determine the mean scattering coefficients averaged across our 30 independent realizations of the 21-cm signal. We then take the numerical derivatives of these mean coefficients with respect to our astrophysical parameters. This is chosen to minimize any potential numerical noise artefacts in the scattering coefficients owing to our chosen image size (i.e. 250 Mpc). In averaging across these 30 realization, we additionally determine our error covariance (sample variance) for the scattering coefficients. We then add this in quadrature with the thermal noise error for the scattering coefficients obtained from 21-cm images to obtain an estimate of the total noise, σ from equation (24). This thermal noise error is determined by averaging different noise realizations (i.e. no 21-cm signal) and applying the same SKA1–low *uv*-coverage, amounting to the mean uncertainty on the scattering coefficients from thermal noise only (no 21-cm signal).

4.2.3 *Comparing astrophysical parameter constraints*

In Fig. 8 we compare the resultant constraints on our astrophysical parameters from our fiducial model from the 21-cm PS (blue) and the S_1 scattering coefficients (red). Along the diagonal are the marginalized 1D probability distribution functions (PDFs), while the remaining panels are the recovered 2D PDFs (95 percentiles). We summarize the resultant 1σ confidence intervals on our individual astrophysical parameters in the top two rows of Table 2.

Overall, we find that the 21-cm PS recovers our astrophysical parameters at $\sim$1.5–2 times higher precision than the 2D WST. However, this was to be expected given that we are comparing the 3D spherically averaged 21-cm PS to the 2D WST applied to 21-cm images. Thus, while in this scenario we have a similar number of independent data points (redshifts and filtering scales), the 3D PS contains additional redshift evolution information by averaging information along the line-of-sight direction, which is crucial for constraining our astrophysical parameters. Further, the 2D images of the 21-cm signal used in the WST analysis contain larger sources

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

Table 2. The recovered 1σ confidence intervals for our astrophysical parameters from the 21-cm power spectrum, various combinations of the scattering transform coefficients (S_1, S_2 and $S_1 + S_2$) and different observational effects. In all, we assume a 1000 h observation with the SKA1-low including its finite *uv*-coverage (instrumental resolution) and thermal noise. We additionally consider the case where we also remove contaminated foreground 'wedge' modes before analysing our 21-cm images (wedge removed). Finally, we consider altering the frequency cadence for the sampling of the 21-cm images used in our mock observation (2, 5, and 10 MHz). See the text for further details.

	$\log_{10}(f_{*,10})$	α_*	$\log_{10}(f_{\rm esc,10})$	$\alpha_{\rm esc}$	$\log_{10}(M_{\rm turn})$ ($M_\odot$)	t_*	$\log_{10}\left(\frac{L_{X<2keV}}{SFR}\right)$ $\left(\rm erg\,s^{-1}\,M_\odot^{-1}\,yr\right)$	E_0 (keV)
Mock Obs.	− 1.30	0.50	− 1.00	− 0.50	8.7	0.5	40.50	0.50
21-cm PS	±0.06	±0.06	±0.05	±0.09	±0.10	±0.05	±0.06	±0.02
Instrumental effects								
S_1	±0.13	±0.08	±0.11	±0.07	±0.22	±0.09	±0.10	±0.05
S_2 (de-corr.)	±0.18	±0.12	±0.15	±0.14	±0.69	±0.15	±0.30	±0.13
$S_1 + S_2$	±0.07	±0.05	±0.07	±0.05	±0.13	±0.05	±0.06	±0.04
Wedge removed								
S_1	±0.20	±0.13	±0.22	±0.12	±0.54	±0.11	±0.24	±0.16
$S_1 + S_2$	±0.12	±0.07	±0.13	±0.07	±0.31	±0.07	±0.14	±0.07
Wedge removed (different cadence)								
$S_1 + S_2$ (2 MHz)	±0.05	±0.03	±0.04	±0.03	±0.13	±0.03	±0.06	±0.03
$S_1 + S_2$ (5 MHz)	±0.07	±0.04	±0.07	±0.04	±0.20	±0.04	±0.09	±0.05
$S_1 + S_2$ (10 MHz)	±0.10	±0.05	±0.09	±0.06	±0.26	±0.06	±0.12	±0.06

of error, namely larger sample variance (as we are only averaging over a 2D image rather than a 3D cube for the 21-cm PS) and higher thermal noise as we do not average across the bandwidth (line-of-sight direction). Theoretically, for this comparison, we could reduce either of these sources of noise, which would notably improve the performance of the WST relative to the 21-cm PS, however, the actual performance is less relevant, instead it is about gaining intuition about the WST. Taking into the account caveats, quite clearly this demonstrates that the information contained in the S_1 coefficients is equivalent to the 21-cm PS.

Importantly, from Fig. 8 we mostly recover similar astrophysical degeneracies for the two statistical methods. This confirms our intuition that the S_1 scattering coefficients behave in a similar way to the 21-cm PS, and follows from Cheng et al. (2020), who noted the similarities between the cosmological constraints obtained from weak lensing from the PS and the S_1 coefficients. However, we do observe some notable changes in the orientations for some ellipses potentially implying different sensitivities to the underlying astrophysics. These likely arise from the fact that the 21-cm PS is the square of the modulus of the convolved field whereas the S_1 coefficients are simply the modulus, allowing some differences to occur. Note, in principle it is also possible to combine both statistics to further improve the constraining power, however, in this work we refrain from doing so focussing more on the comparison between the two approaches.

4.3 Full WST

The true value of the WST for astrophysical parameter inference comes from the inclusion of the additional non-Gaussian information obtained in the second-order coefficients. Therefore using the same setup as in the previous section we now highlight the available improvement when including this non-Gaussian information. In Fig. 9 we compare the 1D and 2D marginalized PDFs for just the S_1 coefficients (red curve), the de-correlated S_2 coefficients (blue curve) and then combining all coefficients (i.e. $S_1 + S_2$; yellow curve). Additionally, we provide the recovered 68th percentiles for each of the astrophysical parameters in Table 2.

With the addition of the non-Gaussian information following the inclusion of the S_2 coefficients, we find $\sim$1.5–2 $\times$ improvements for the astrophysical parameter constraints relative to those from just the S_1 coefficients. To highlight how the second order scattering coefficients adds constraining power, we focus on the de-correlated S_2 coefficients, which isolate out the non-Gaussian information by averaging out the dependence of the first-order information. These are highlighted by the blue curves in Fig. 9. The astrophysical constraints from just the de-correlated S_2 coefficients are comparable, though marginally worse than those from S_1 for all expect $M_{\rm turn}$, $\frac{L_{X<2keV}}{SFR}$ and E_0. Instead for these, the constraints are $\sim$2–2.5 times worse than the S_1 coefficients. However, this is consistent with the picture from Watkinson et al. (2021) and Tiwari et al. (2021), where for a much simpler three parameter reionization model the constraints from just the isosceles bispectrum (a single configuration of the bispectrum) alone are larger than from the 21-cm PS.[13] For $\frac{L_{X<2keV}}{SFR}$ and E_0 the reduction in constraining power arises owing to the fact that the 21-cm signal is most sensitive to the X-ray parameters during the Epoch of Heating (i.e. $z \gtrsim 10$). During this time (see Fig. 4) the S_2 scattering coefficients are considerably less sensitive, thus there is little constraining power on the X-ray parameters from the non-Gaussian information.

Importantly, in several cases the recovered contours from the de-correlated S_2 coefficients are (or are close to) orthogonal to those from the S_1 coefficients. This orthogonality indicates that the S_2 coefficients are sensitive to additional information which is capable of breaking degeneracies amongst the astrophysical

[13]Tiwari et al. (2021) also consider the scenario of all possible unique bispectrum configurations, showing an improvement over the 21-cm PS. In doing so, this uses more information than available from the S_2 coefficients as these are coarsely binned over a few j spatial configurations only. However, computing the likelihood for the bispectrum is considerably more challenging owing to the noisiness of the statistic.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

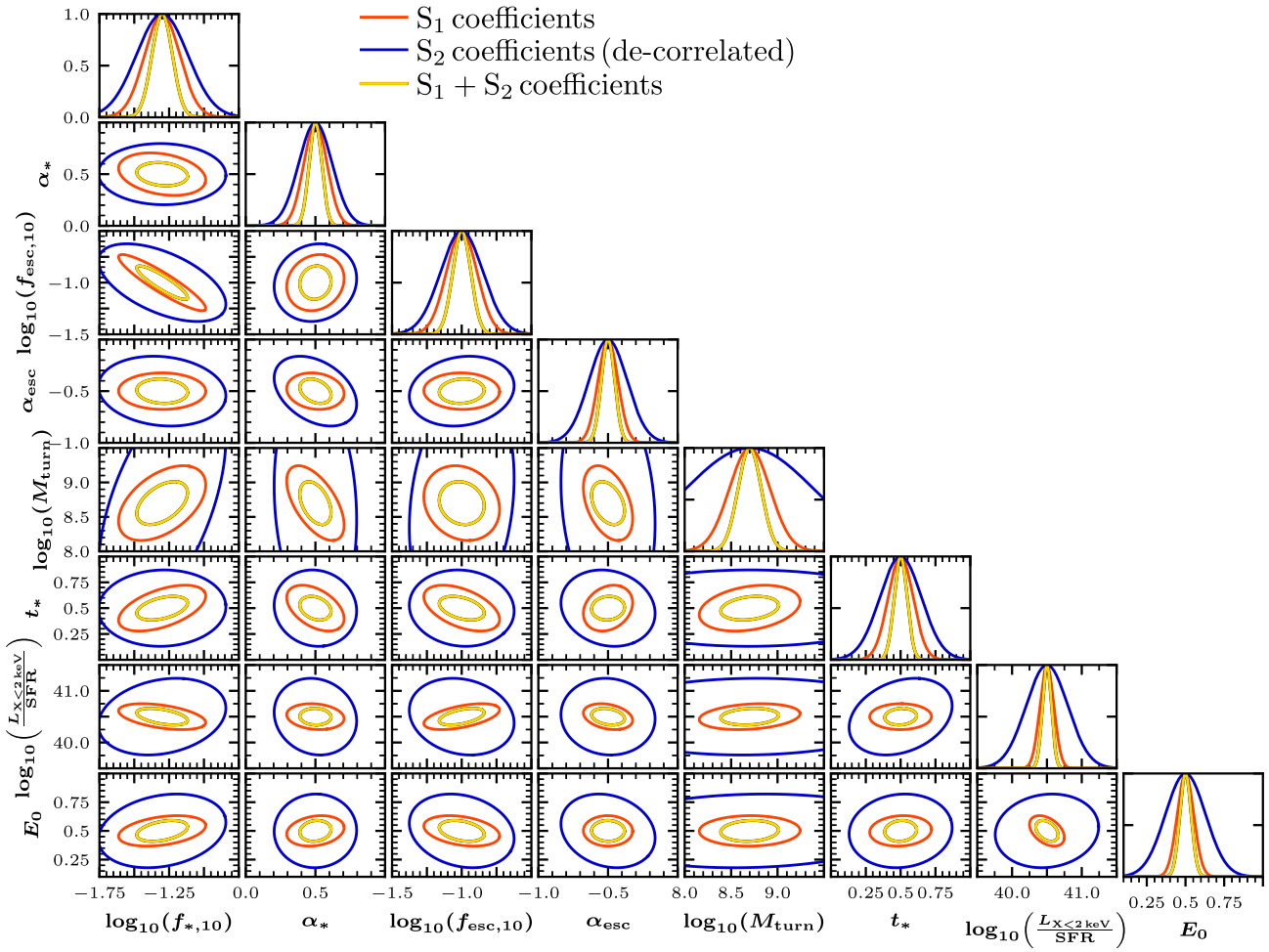


Figure 9. Similar to Fig. 8, except we compare the astrophysical parameter constraints obtained from just the S_1 coefficients (red), the de-correlated S_2 coefficients (blue), and the full combination of $S_1 + S_2$ coefficients (yellow).

parameter constraints from just the S_1 coefficients alone. This is the value of including the non-Gaussian information from the WST when characterizing the 21-cm signal, where we are now additionally measuring the clustering of the sizes of the spatial fluctuations in the 21-cm signal rather than just the sizes of the spatial fluctuations with Gaussian statistics such as the 21-cm PS. Even for parameter combinations where the de-correlated S_2 coefficients are not orthogonal to the S_1 constraints, we still see improvements when combining the S_1 and S_2 coefficients.

To fully appreciate the improvement in the constraining power with the 2D WST when using the non-Gaussian information, in Fig. 10, we compare the marginalized PDFs for the full 2D WST (red curves) and the 3D 21-cm PS (blue curves). From this we clearly see that the full 2D WST is now comparable to that from the 21-cm PS, even marginally improving on the constraining power for α_* and $\alpha_{\rm esc}$. This highlights the true value of the WST that despite the fact we are using considerably less information by restricting ourselves to 2D images we are able to recover comparable constraints to statistics utilizing the full 3D information. This demonstrates that the full WST is superior to the 21-cm PS, and is the main result of this work.

In theory, we could construct the Fisher matrix for the bispectrum and compare more directly to the WST. However, in this work we refrain from doing so as the bispectrum is a considerably more complex statistic than the 21-cm PS. While analogies can be trivially drawn between the binning of the PS and the information in the S_1 coefficients (see Fig. 3) the same is not true for the bispectrum. A detailed exploration of all possible closure configurations for the bispectrum will be required to ascertain the analogies to the information contained in S_2 coefficients. Thus, we defer such a detailed exploration to future work. Nevertheless, this highlights another advantage of the WST. That is, the simplicity to which we are able to access and utilize the non-Gaussian information from the higher order scattering coefficients relative to other statistical methods.

4.4 Full WST with foreground contaminated modes removed

In Fig. 11 we next compare the 1D and 2D marginalized PDFs from combining the $S_1 + S_2$ coefficients in the presence of contamination by astrophysical foregrounds. The red curves correspond to no contamination (i.e. perfect removal of foreground contamination), whereas the blue curves correspond to the removal of all Fourier modes from within the contaminated foreground 'wedge' from our interferometer experiment. We also provide the recovered 68th percentiles for each of the astrophysical parameters in Table 2.

Following the removal of these contaminated foreground 'wedge' modes (see Fig. 5 for a visualization) we find at most a factor of ∼1.5–2 reduction in the constraining power of our astrophysical

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

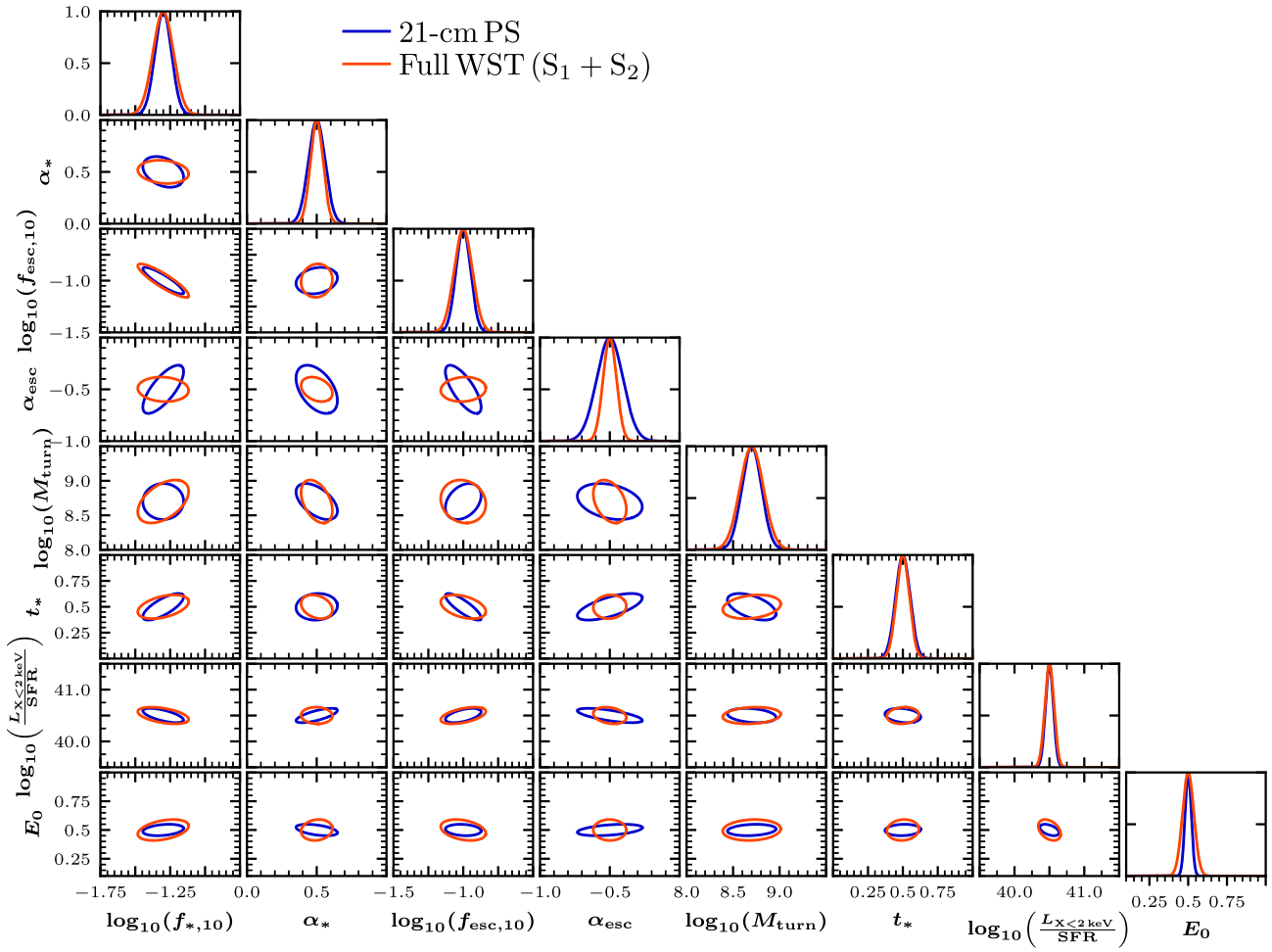


Figure 10. Similar to Fig. 8, except we compare the astrophysical parameter constraints obtained from full 2D WST (i.e. the $S_1 + S_2$ coefficients; red curves) relative to those obtained from the 3D 21-cm PS (blue curve).

parameters. Nevertheless, this is still a factor of ∼2 improvement over that achievable with just the S_1 coefficients (i.e. our proxy for the 21-cm PS, see Table 2). This loss in constraining power following the removal of the wedge modes is to be expected, given the loss of spatial information in the 21-cm signal. While excising these contaminated foreground modes reduces our constraining power from 21-cm images, there are promising approaches in the literature using machine learning to attempt to recover the lost cosmic signal from these foreground contaminated modes (see e.g. Gagnon-Hartman et al. 2021). Thus our ability to constrain astrophysical parameters from images of the 21-cm signal could be further improved by combining both methods. We leave such an investigation to future work. It is also important to note that thus far we have only considered a fairly limited sampling in redshift (frequency) of 2D images of the 21-cm signal. In principle, we can mitigate some of these losses in constraining power simply by increasing the number of independent images used in our WST analysis. We explore this in the next section.

4.5 Observing cadence

Previously, to aid comparisons between the 21-cm PS and the WST for astrophysical parameter recovery from a mock observation of the 21-cm signal with the SKA1–low we tailored our analysis for the 21-cm PS. That is, we analysed 21-cm images extracted from the central frequency of equal comoving volumes taken from our simulated 21-cm light-cone (to obtain a similar number of measurements). In effect, this resulted in 12 images which were not equally spaced in frequency, with the largest separation of ∼18 MHz during the EoR, down to ∼10 MHz at $z \sim 20$. In practice, as we observe the full 21-cm light-cone spanning the observing frequency of our interferometer experiment, we can instead extract images at a fixed cadence in frequency. In particular, for our mock observation with the SKA1–low, we consider the 21-cm light-cone spanning $z \sim 5.9 - 27.4$, corresponding to a frequency coverage of ∼206–50 MHz. Thus, in this section we determine the astrophysical parameter constraints from the WST exploring different cadences in frequency between the 21-cm images extracted from the 21-cm light-cone.

Specifically, we consider three possible cadences in frequency for our 21-cm images: (i) 2 MHz, (ii) 5 MHz, and (iii) 10 MHz. That is, we apply the WST to a 2D image of the 21-cm signal extracted from the full light-cone every 2, 5, or 10 MHz. These choices are arbitrary, but serve to demonstrate the possible constraining power available for different cadences. It remains to be seen how many independent images will be achievable with an experiment such as the SKA1–low, but theoretically even higher cadences are plausible. In total, these observational setups result in 79, 32, and 16 independent images of the 21-cm signal extracted from the full observed light cone, respectively. Note however that beyond about $z \sim 15$, the

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

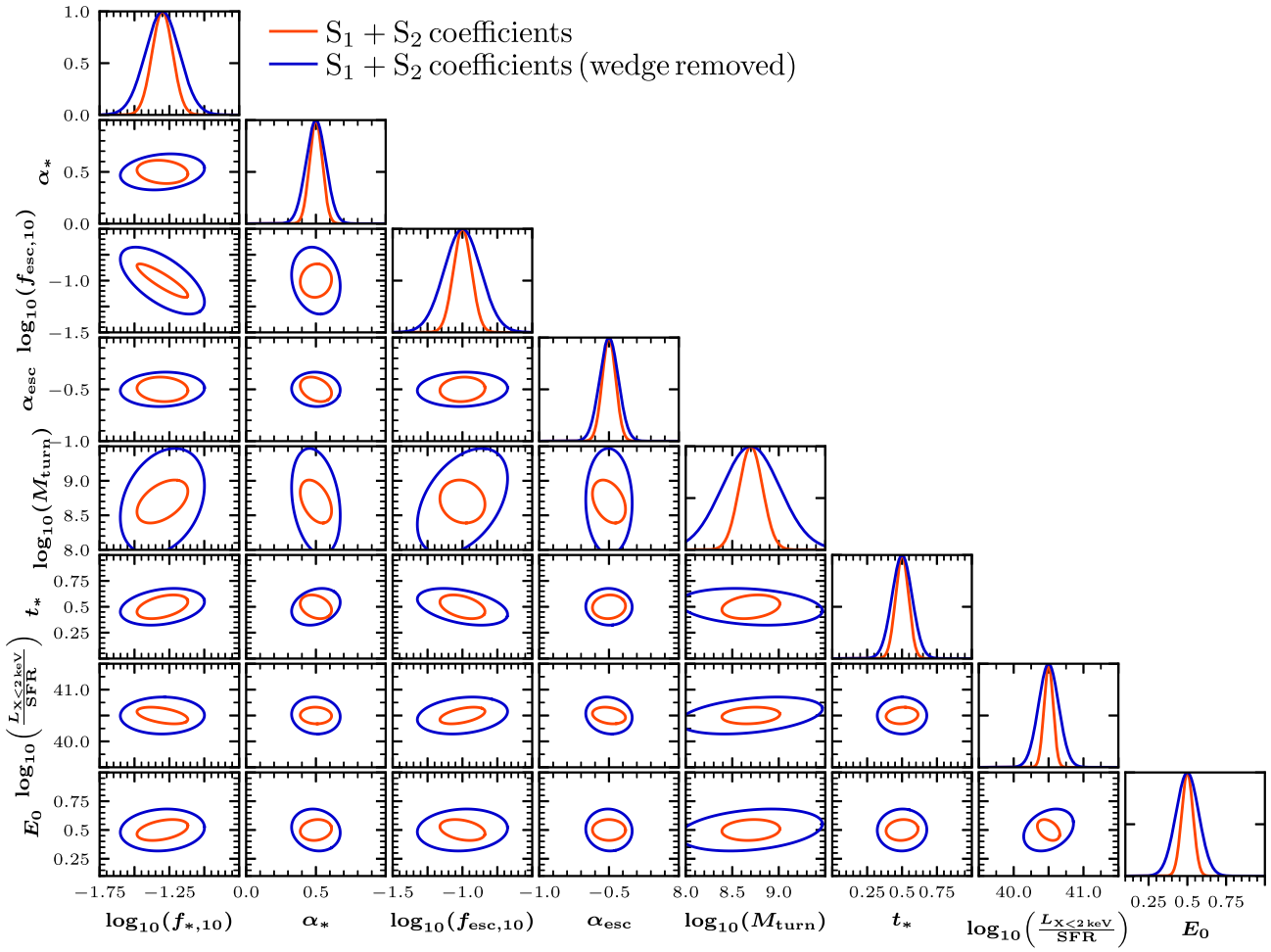


Figure 11. Similar to Fig. 8, except we explore the impact of removing the contaminated foreground 'wedge' modes from our mock 21-cm images using the $S_1 + S_2$ coefficients. The blue curves have these contaminated modes removed, whereas the red curves assumes no foreground contamination (i.e. perfect removal).

images will be swamped by thermal noise (see Fig. 6 thus not all images add additional constraining power. In all scenarios, we have additional information relative to the previous sections where we only had 12 images corresponding to the central redshift for the 3D 21-cm PS. Further, by now evenly sampling in frequency space, we additionally obtain more measurements during the EoR and EoH unlike previously where we had a cadence of ∼18 MHz during these epochs. For all, we consider all observational sources of error with our mock SKA1 – low observation, instrumental resolution, thermal noise, and the removal of foreground contaminated 'wedge' modes.

In Fig. 12 we present the 1D and 2D marginalized PDFs for our astrophysical parameters for these three different cadences in 21-cm images. The red, blue, and yellow curves correspond to the 2, 5, and 10 MHz scenarios, respectively. Further, we summarize the 68 percentile constraints on our astrophysical parameters in the final three rows of Table 2. As expected, following the addition of more independent information by increasing the cadence of the sampled 21-cm images we observe improvements for all astrophysical parameters relative to the cases explored previously. In the case of the 2 MHz sampling (∼6.5 × more 21-cm images) we recover improvements in our astrophysical parameter constraints by a factor of ∼2–3. For 5 MHz sampling (∼3 × more 21-cm images) we recover improvements of ∼1.5–2. Even for the 10 MHz scenario, where the number of images was similar to our previous cases (16 compared to 12) we still note improvements of ∼10–30 per cent for most of the astrophysical parameters. These improvements can be attributed to the inclusion of additional 21-cm images (information) during the EoR and EoH relative to previously where the cadence varied.

Importantly, with cadences of 2 or 5 MHz, the recovered astrophysical parameter constraints are comparable to those from the original 21-cm PS analysis. In fact, for a 2 MHz cadence, the 2D WST outperforms the 3D 21-cm PS for almost all astrophysical parameters by up to a factor of ∼2 (for α_* and $\alpha_{\rm esc}$). This, while still only using 2D images which individually contain considerably larger sample variance and thermal noise errors. In principle, we could further reduce the thermal noise in our 2D images by averaging over multiple frequency channels within a fixed bandwidth. Alternatively, we could alter the physical size of our wavelet filters to extract more spatial information. As discussed earlier, we could also instead consider the 3D WST to fully access the available 3D information. Thus, there are many potential avenues to further improve the performance of the WST on astrophysical parameter recovery from the 21-cm signal, some of which we will explore in future. This highlights the unique potential of the WST.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

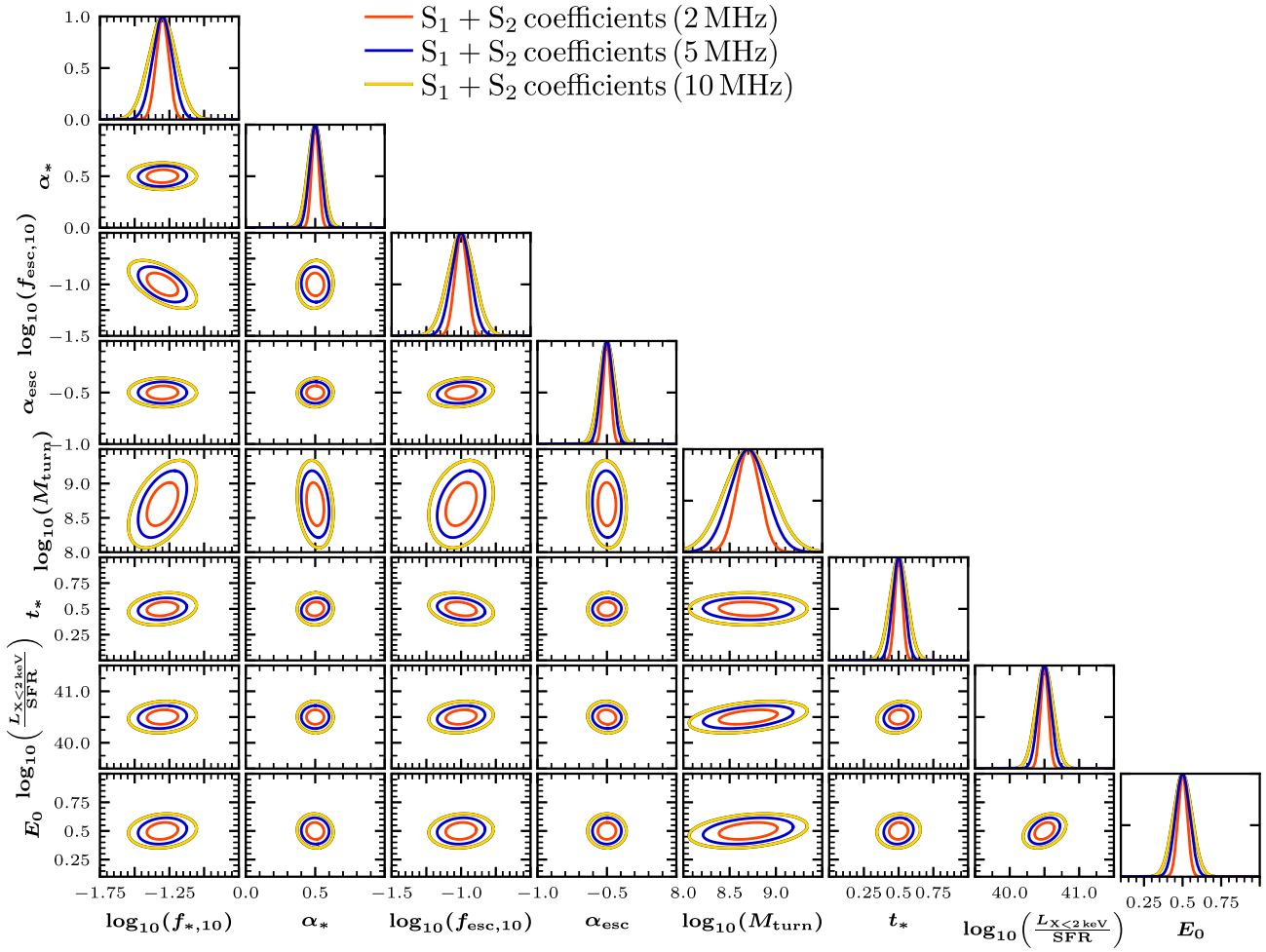


Figure 12. Similar to Fig. 8, except we compare the astrophysical constraints from combining the $S_1 + S_2$ coefficients for different cadences in frequency between the 21-cm images extracted from the full 21-cm light-cone observed with SKA1–low. We consider cadences of 2 MHz (red), 5 MHz (blue), and 10 MHz (yellow). Here, we consider all instrumental effects including the removal of contaminated foreground 'wedge' modes.

5 CONCLUSION

We have introduced the WST and applied it to the 21-cm signal for the first time. The WST (Mallat 2012) convolves a family of wavelet filters of different physical extents and rotations to an input image. The key advantage of the WST comes from its usage of wavelet filters, which preserve locality of the signal both spatially and in frequency. Thus, after successive convolutions and modulus operations applied to the images we are able to extract non-Gaussian information from the 21-cm signal. Importantly, the WST provides a clearly defined series of scattering coefficients, which enables them to be used in a robust statistical analysis.

We performed a qualitative exploration of the WST applied to simulated 21-cm images extracted from the light cone of four different astrophysical models to gain physical insight. These 21-cm simulations were generated using v3 of 21CMFAST (Mesinger & Furlanetto 2007; Mesinger et al. 2011; Murray et al. 2020), specifically the Park et al. (2019) flexible galaxy parametrization. These four astrophysical models were chosen to highlight different features in the evolution of the 21-cm signal. We considered a fiducial model, which matches all existing observational constraints, a 'cold' reionization scenario which amplifies the 21-cm signal in the absence of little to no X-ray heating, a model driven by large haloes (resulting in large ionized regions) and an extended model with reionization driven by faint, low mass haloes (small ionized regions).

Next, we performed a quantitative analysis of the WST by recovering the precision to which we can recover our input astrophysical model parameters from a mock 1000 h SKA1–low observation. Specifically, we used Fisher matrices and performed several different explorations for which we summarize below:

(i) A comparison between the 3D 21-cm PS and the S_1 scattering coefficients (i.e. Gaussian information) obtained from 2D images of the 21-cm signal. Qualitatively, we find the S_1 scattering coefficients recover similar astrophysical degeneracies to the 21-cm PS confirming our expectation that the S_1 scattering coefficients behave similarly to the 21-cm PS. Quantitatively, we found the 21-cm PS outperformed the S_1 scattering coefficients, however, these differences can be attributed to the fact that we only considered 2D images, where the statistical and instrumental noise is increased relative to the 3D 21-cm PS.

(ii) The full potential of the WST was highlighted by considering the constraining power when combining the S_1 and S_2 scattering coefficients. We found that despite only using 2D information, we could recover comparable constraints to the 21-cm PS, even improving the constraints in some instances. These improvements are comparable to those achievable when combining the 21-cm PS

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

with the bispectrum (see e.g. Tiwari et al. 2021; Watkinson et al. 2021). The inclusion of the non-Gaussian information provided by the S_2 coefficients provides additional orthogonal information about the astrophysical parameters breaking degeneracies present when considering only the 21-cm PS. This is the main result of this work, highlighting that the WST is superior to the 21-cm PS for extracting astrophysical information. Secondly, that it is trivial to measure, interpret, and utilize the non-Gaussian information with the WST compared to other statistical measures.

(iii) We explored the relative loss of constraining power following the removal of foreground contaminated 'wedge' modes from our 21-cm images. With the WST ($S_1 + S_2$) we observed at most a factor of $\sim$2 reduction in the constraining power of our astrophysical parameters following the excision of these contaminated modes.

(iv) A variety of observationally motivated cadences in frequency between successive images extracted from the 21-cm light cone were considered (arbitrarily chosen to be 2, 5, and 10 MHz). Even with the foreground 'wedge' modes removed, we could further notably improve on the recovered precision of our astrophysical parameters by analysing 21-cm images every 2–5 MHz. In doing so, we found we could improve our constraining power on all model parameters by a factor of $\sim$2 compared to our previous scenarios using the 2D WST.

While we have demonstrated that the 2D WST has the potential to improve astrophysical parameter recovery from the 21-cm signal relative to the 3D 21-cm PS, thus far we have only explored it in the context of Fisher matrices, which are known to overestimate the constraining power owing to several simplifying assumptions (e.g. Gaussian errors). These are suitable for an initial exploration, however, in future we will explore the WST in a more robust Bayesian statistical framework using direct MCMC sampling of 21-cm images with 21CMMC to more accurately account for the inherent modelling and observational uncertainties.

Finally, in this work we have focused on astrophysical parameter inference with the WST. However, there are many other possible applications of the WST to the reionization epoch. For example, is the WST sensitive enough to detect the non-Gaussian information when the 21-cm signal is below the thermal noise? Doing so could provide the first detection of the 21-cm signal. Alternatively, can the WST be used to reconstruct information that would otherwise be contaminated by astrophysical foregrounds or other instrumental effects. Further, can the WST be used for cross-correlation studies between observed galaxies and the 21-cm signal (e.g. Sobacchi, Mesinger & Greig 2016; Hutter, Trott & Dayal 2018; Vrbanec et al. 2020). We will explore these applications of the WST in future work.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their comments. We also thank Sihao Cheng for early discussions on this work, Steven Murray for discussions related to applying interferometer noise effects and Charlotte Mason for discussions regarding Fisher Matrices. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. YST acknowledges financial support from the Australian Research Council through DECRA Fellowship DE220101520. Parts of this work were performed on the OzSTAR national facility at Swinburne University of Technology. OzSTAR is funded by Swinburne University of Technology.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Allys E., Levrier F., Zhang S., Colling C., Regaldo-Saint Blancard B., Boulanger F., Hennebelle P., Mallat S., 2019, A&A, 629, A115
Allys E., Marchand T., Cardoso J. F., Villaescusa-Navarro F., Ho S., Mallat S., 2020, Phys. Rev. D, 102, 103506
Andreux M. et al., 2018, preprint (arXiv:1812.11214)
Baek S., Semelin B., Matteo P. D., Revaz Y., Combes F., 2010, A&A, 523, A4
Bag S., Mondal R., Sarkar P., Bharadwaj S., Choudhury T. R., Sahni V., 2019, MNRAS, 485, 2235
Banet A., Barkana R., Fialkov A., Guttman O., 2021, MNRAS, 503, 1221
Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
Behroozi P. S., Silk J., 2015, ApJ, 799, 32
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Bianco M., Giri S. K., Iliev I. T., Mellema G., 2021, MNRAS, 505, 3982
Bowman J. D., Rogers A. E. E., 2010, Nature, 468, 796
Bruna J., Mallat S., Bacry E., Muzy J.-F., 2015, Ann. Stat., 43, 323
Chapman E., Jelić V., 2019, preprint (arXiv:1909.12369)
Chen Z., Xu Y., Wang Y., Chen X., 2019, ApJ, 885, 23
Cheng S., Ménard B., 2021a, preprint (arXiv:2112.01288)
Cheng S., Ménard B., 2021b, MNRAS, 507, 1012
Cheng S., Ting Y.-S., Ménard B., Bruna J., 2020, MNRAS, 499, 5902
Datta A., Bowman J. D., Carilli C. L., 2010, ApJ, 724, 526
Dayal P., Ferrara A., Dunlop J. S., Pacucci F., 2014, MNRAS, 445, 2545
DeBoer D. R. et al., 2017, PASP, 129, 045001
Eastwood M. W. et al., 2019, AJ, 158, 84
Elbers W., van de Weygaert R., 2019, MNRAS, 486, 1523
Field G. B., 1958, Proc. Inst. Radio Eng., 46, 240
Fisher R. A., 1935, J. R. Stat. Soc., 98, 39
Fragos T. et al., 2013, ApJ, 764, 41
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1
Gagnon-Hartman S., Cui Y., Liu A., Ravanbakhsh S., 2021, MNRAS, 504, 4716
Gazagnes S., Koopmans L. V. E., Wilkinson M. H. F., 2021, MNRAS, 502, 1816
Gillet N., Mesinger A., Greig B., Liu A., Ucci G., 2019, MNRAS, 484, 282
Giri S. K., D'Aloisio A., Mellema G., Komatsu E., Ghara R., Majumdar S., 2019b, J. Cosmol. Astropart. Phys., 2019, 058
Giri S. K., Mellema G., 2021, MNRAS, 505, 1863
Giri S. K., Mellema G., Aldheimer T., Dixon K. L., Iliev I. T., 2019a, MNRAS, 489, 1590
Giri S. K., Mellema G., Dixon K. L., Iliev I. T., 2018a, MNRAS, 473, 2949
Giri S. K., Mellema G., Ghara R., 2018b, MNRAS, 479, 5596
Giroux M. L., Sutherland R. S., Shull J. M., 1994, ApJ, 435, L97
Gnedin N. Y., Ostriker J. P., 1997, ApJ, 486, 581
Gnedin N. Y., Shaver P. A., 2004, ApJ, 608, 611
Gorce A., Hutter A., Pritchard J. R., 2021, A&A, 653, A58
Gorce A., Pritchard J. R., 2019, MNRAS, 489, 1321
Greig B., Mesinger A., 2015, MNRAS, 449, 4246
Greig B., Mesinger A., 2017, MNRAS, 472, 2651
Greig B., Mesinger A., 2018, MNRAS, 477, 3217
Gupta Y. et al., 2017, Curr. Sci., 113, 707
Harikane Y. et al., 2016, ApJ, 821, 123
Hassan S., Andrianomena S., Doughty C., 2020, MNRAS, 494, 5761
Hassan S., Liu A., Kohn S., La Plante P., 2019, MNRAS, 483, 2524
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Hutter A., Dayal P., Yepes G., Gottlöber S., Legrand L., Ucci G., 2021, MNRAS, 503, 3698
Hutter A., Trott C. M., Dayal P., 2018, MNRAS, 479, L129
Kakiichi K. et al., 2017, MNRAS, 471, 1936

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

Kamran M., Ghara R., Majumdar S., Mondal R., Mellema G., Bharadwaj S., Pritchard J. R., Iliev I. T., 2021, MNRAS, 502, 3800
Kapahtia A., Chingangbam P., Appleby S., 2019, J. Cosmol. Astropart. Phys., 2019, 053
Kapahtia A., Chingangbam P., Ghara R., Appleby S., Choudhury T. R., 2021, J. Cosmol. Astropart. Phys., 2021, 026
Koopmans L. et al., 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14). preprint (arXiv:1505.07568)
Kubota K., Yoshiura S., Shimabukuro H., Takahashi K., 2016, PASJ, 68, 61
Kuhlen M., Faucher-Giguère C.-A., 2012, MNRAS, 423, 862
Kwon Y., Hong S. E., Park I., 2020, J. Korean Phys. Soc., 77, 49
La Plante P., Ntampaka M., 2019, ApJ, 880, 110
Liu A., Parsons A. R., Trott C. M., 2014a, Phys. Rev. D, 90, 023018
Liu A., Parsons A. R., Trott C. M., 2014b, Phys. Rev. D, 90, 023019
Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429
Majumdar S., Kamran M., Pritchard J. R., Mondal R., Mazumdar A., Bharadwaj S., Mellema G., 2020, MNRAS, 499, 5090
Majumdar S., Pritchard J. R., Mondal R., Watkinson C. A., Bharadwaj S., Mellema G., 2018, MNRAS, 476, 4007
Mallat S., 2012, Commun. Pure Appl. Math., 65, 1331
Mangena T., Hassan S., Santos M. G., 2020, MNRAS, 494, 600
Mellema G. et al., 2013, Exp. Astron., 36, 235
Mesinger A., Dijkstra M., 2008, MNRAS, 390, 1071
Mesinger A., Ewall-Wice A., Hewitt J., 2014, MNRAS, 439, 3262
Mesinger A., Furlanetto S., 2007, ApJ, 669, 663
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Mineo S., Gilfanov M., Sunyaev R., 2012, MNRAS, 419, 2095
Mitra S., Choudhury T. R., Ferrara A., 2015, MNRAS, 454, L76
Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, ApJ, 752, 137
Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
Murray S. G., Trott C. M., 2018, ApJ, 869, 25
Murray S., Greig B., Mesinger A., Muñoz J., Qin Y., Park J., Watkinson C., 2020, J. Open Source Softw., 5, 2582
Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 462, 250
Ocvirk P. et al., 2016, MNRAS, 463, 1462
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, MNRAS, 443, 678
Park J., Mesinger A., Greig B., Gillet N., 2019, MNRAS, 484, 933
Parsons A. R. et al., 2010, AJ, 139, 1468
Parsons A. R. et al., 2014, ApJ, 788, 106
Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, ApJ, 756, 165
Planck Collaboration VI, 2020, A&A, 641, A6
Pober J. C. et al., 2013, AJ, 145, 65
Pober J. C. et al., 2014, ApJ, 782, 66
Pober J. C. et al., 2016, ApJ, 819, 8
Prelogović D., Mesinger A., Murray S., Fiameni G., Gillet N., 2022, MNRAS, 509, 3852
Pritchard J. R., Furlanetto S. R., 2007, MNRAS, 376, 1680
Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901
Qin Y., Mesinger A., Bosman S. E. I., Viel M., 2021, MNRAS, 506, 2390
Regaldo-Saint Blancard B., Levrier F., Allys E., Bellomi E., Boulanger F., 2020, A&A, 642, A217
Salpeter E. E., 1955, ApJ, 121, 161
Saydjari A. K., Portillo S. K. N., Slepian Z., Kahraman S., Burkhart B., Finkbeiner D. P., 2021, ApJ, 910, 122
Scoccimarro R., 1998, MNRAS, 299, 1097
Shapiro P. R., Giroux M. L., Babul A., 1994, ApJ, 427, 25
Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Shimabukuro H., Yoshiura S., Takahashi K., Yokoyama S., Ichiki K., 2015, MNRAS, 451, 467
Shimabukuro H., Yoshiura S., Takahashi K., Yokoyama S., Ichiki K., 2016, MNRAS, 458, 3003
Shimabukuro H., Yoshiura S., Takahashi K., Yokoyama S., Ichiki K., 2017, MNRAS, 468, 1542
Sobacchi E., Mesinger A., 2013a, MNRAS, 432, L51
Sobacchi E., Mesinger A., 2013b, MNRAS, 432, 3340
Sobacchi E., Mesinger A., 2014, MNRAS, 440, 1662
Sobacchi E., Mesinger A., Greig B., 2016, MNRAS, 459, 2741
Springel V., Hernquist L., 2003, MNRAS, 339, 312
Sun G., Furlanetto S. R., 2016, MNRAS, 460, 417
Tacchella S., Bose S., Conroy C., Eisenstein D. J., Johnson B. D., 2018, ApJ, 868, 92
Thompson A. R., Moran J. M., Swenson G. W., 2007, Interferometry and Synthesis in Radio Astronomy. Wiley, New York
Thyagarajan N. et al., 2013, ApJ, 776, 6
Thyagarajan N. et al., 2015a, ApJ, 804, 14
Thyagarajan N. et al., 2015b, ApJL, 807, L28
Tingay S. J. et al., 2013, Publ. Astron. Soc. Aust., 30, e007
Tiwari H., Shaw A. K., Majumdar S., Kamran M., Choudhury M., 2021, preprint (arXiv:2108.07279)
Tozzi P., Madau P., Meiksin A., Rees M. J., 2000, ApJ, 528, 597
Trott C. M., 2016, MNRAS, 461, 126
Trott C. M., Wayth R. B., Tingay S. J., 2012, ApJ, 757, 101
Valogiannis G., Dvorkin C., 2021, preprint (arXiv:2108.07821)
van Haarlem M. P. et al., 2013, A&A, 556, A2
Vedantham H., Shankar N. U., Subrahmanyan R., 2012, ApJ, 745, 176
Vrbanec D., Ciardi B., Jelić V., Jensen H., Iliev I. T., Mellema G., Zaroubi S., 2020, MNRAS, 492, 4952
Watkinson C. A., Giri S. K., Ross H. E., Dixon K. L., Iliev I. T., Mellema G., Pritchard J. R., 2019, MNRAS, 482, 2653
Watkinson C. A., Greig B., Mesinger A., 2022, MNRAS, 510, 3838
Watkinson C. A., Pritchard J. R., 2014, MNRAS, 443, 3090
Wayth R. B. et al., 2018, Publ. Astron. Soc. Aust., 35, e033
Wouthuysen S. A., 1952, AJ, 57, 31
Yoshiura S., Shimabukuro H., Takahashi K., Matsubara T., 2017, MNRAS, 465, 394
Yoshiura S., Shimabukuro H., Takahashi K., Momose R., Nakanishi H., Imai H., 2015, MNRAS, 451, 266
Yue B., Ferrara A., Xu Y., 2016, MNRAS, 463, 1968
Yung L. Y. A., Somerville R. S., Popping G., Finkelstein S. L., Ferguson H. C., Davé R., 2019, MNRAS, 490, 2855
Zarka P., Girard J. N., Tagger M., Denis L., 2012, in Boissier S., de Laverny P., Nardetto N., Samadi R., Valls-Gabaud D., Wozniak H., eds, SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. French Society of Astronomy and Astrophysics, Paris, p. 687
Zhao X., Mao Y., Cheng C., Wandelt B. D., 2022, ApJ, 926, 151

APPENDIX A: CONVERGENCE TESTS

Throughout this work, we consider 21-cm simulations with a transverse comoving length of 250 Mpc and 128 voxels per side length. To ensure our results are consistent across different volumes and simulation resolutions, here we perform some rudimentary convergence tests. Throughout we only show the convergence tests for the S_1 scattering coefficients, noting we find similar performance for the S_2 coefficients.

In Fig. A1 we explore the redshift evolution of the S_1 scattering coefficients recovered directly from 21-cm simulations (i.e. no instrumental effects) of varying comoving length for fixed voxel resolution. The blue, red, and black curves correspond to transverse lengths of 250 (128), 500 (256), and 1000 Mpc (512 voxels), respectively. Here, we only show j's which are measurable across all three sizes. For all, the recovered evolution in S_1 scattering coefficients are consistent, irrespective of simulation size (at fixed resolution). The variations across each simulation do increase for larger j's, however, these are consistent to within the sample variance error. That is, for increasing j (i.e. larger physical scales) the smaller simulation sizes will have less information on these scales than the larger volumes, thus the smaller

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

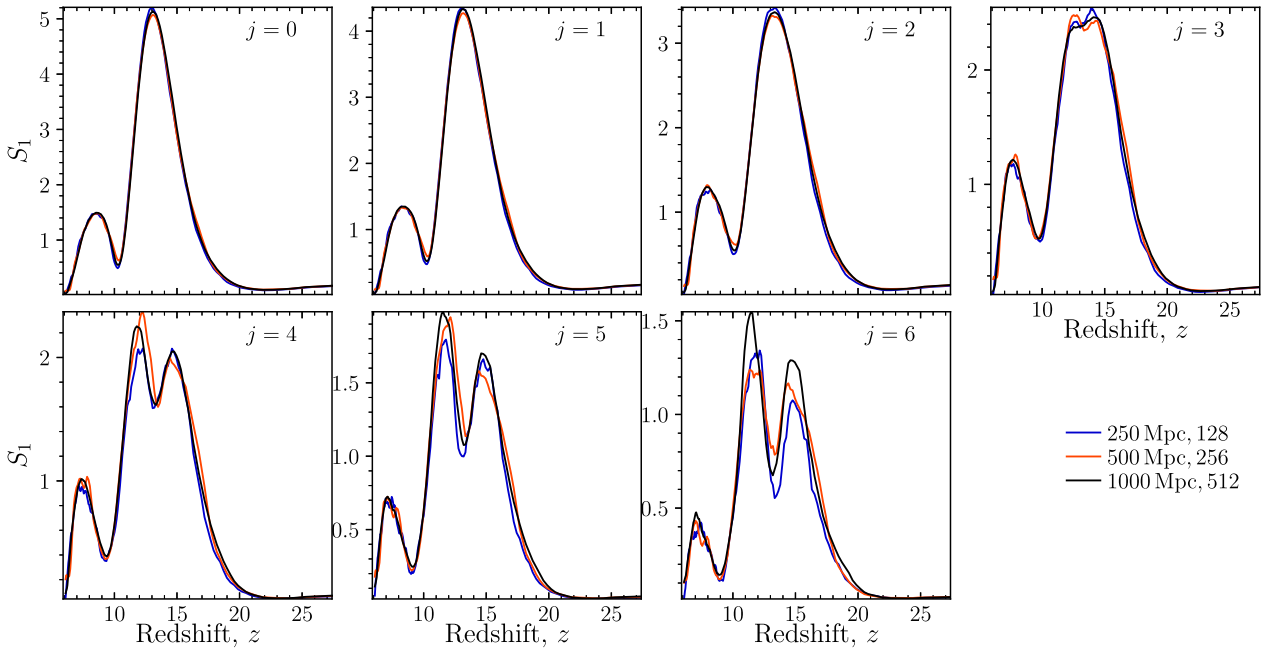


Figure A1. The evolution in the S_1 scattering coefficients for different simulation volumes at fixed resolution. The blue, red, and black curves correspond to a transverse side length of 250 (128), 500 (256), and 1000 Mpc (512 voxels).

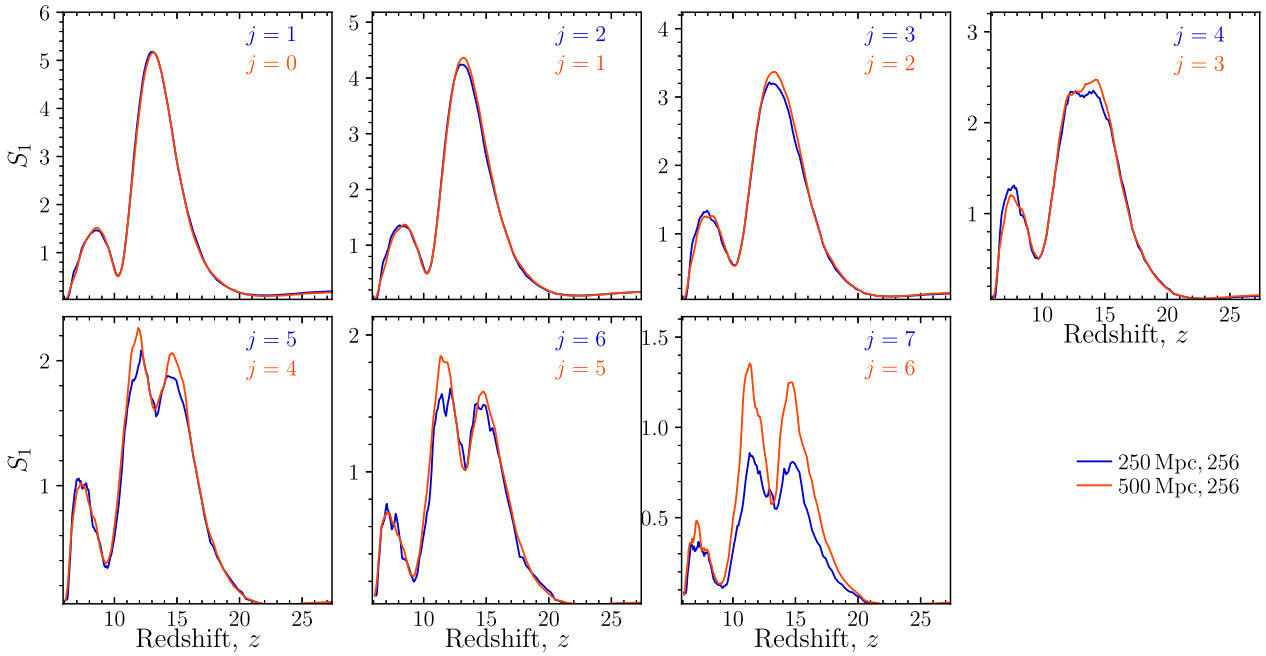


Figure A2. The evolution in the S_1 scattering coefficients for different voxel resolution in our simulations. The blue curves correspond to a simulation of 250 Mpc (256 voxels) whilst the red curve corresponds to 500 Mpc (256 voxels). Note, to accurately compare between resolution we must consider the filters at the same physical extent to ensure the corresponding features are consistent.

sized simulations are impacted more by sample variance than the larger simulations.

In Fig. A2 we instead investigate the redshift evolution of the S_1 scattering coefficients recovered directly from 21-cm simulations (i.e. no instrumental effects) of varying voxel resolution. Here, the blue and red curves correspond to simulations with 250 (256) and 500 Mpc (256 voxels), corresponding to resolutions of $\sim$1 and $\sim$2 Mpc per voxel. Note, to perform this comparison, the j-scales are offset by one to ensure we are comparing to the same physical features in the 21-cm signal (i.e. filters are of the same physical extent). Again, we find the recovered evolution in the S_1 scattering coefficients are consistent with different voxel resolution. Like previously, we again find the variations increase within increasing j-scale, which is due to the increasing sample variance in the higher resolution simulation (i.e. it is at a higher j than the lower resolution simulation).

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

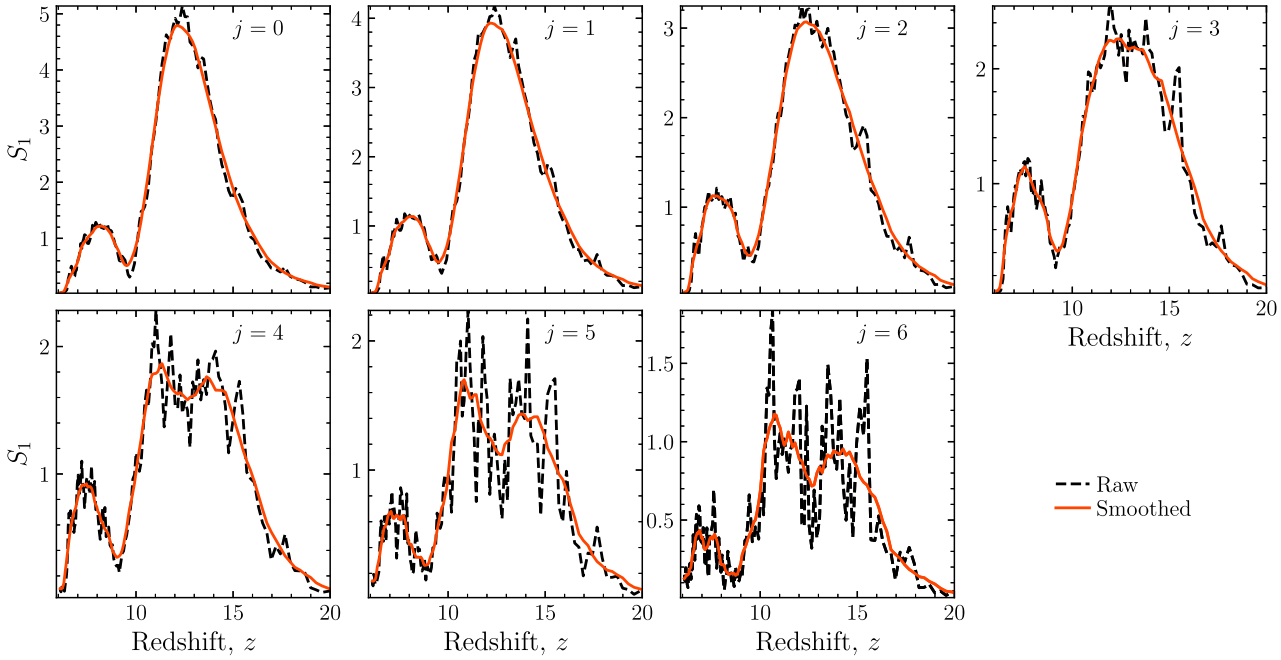


Figure B1. Plotting the evolution of the scattering coefficients from the simulated 21-cm images from a single simulation can be relatively noisy (black dashed curve). Thus, we perform a box-car filtering over 10 MHz to smooth out these features (red solid curve). This more clearly highlights the general trend in the scattering coefficient evolution with redshift, and what is performed throughout this work. Here, we present only the S_1 scattering coefficients, but the same procedure is also applied to the S_2 coefficients.

APPENDIX B: SMOOTHING THE SCATTERING COEFFICIENTS

Throughout Section 3.1 we apply the WST to images of the 21-cm signal extracted from a full 3D light cone from a single simulation to explore the redshift evolution of the scattering coefficients. In particular, we are applying the WST to images that are generated from a single slice in our simulation, corresponding to only $\sim$1.5 comoving Mpc ($\sim$100 kHz). Which is considerably narrower in depth than images that would be obtained from 21-cm interferometer experiments. As such, we experience relatively noisy features in our scattering coefficients when exploring their evolution with redshift, as demonstrated by the black dashed curves in Fig. B1 for our fiducial astrophysical model. To mitigate these noisy features, we instead perform a box-car filtering in frequency, over a 10 MHz bandwidth. That is, for each scattering coefficient we average over the values of the scattering coefficient $\pm$5 MHz from the current frequency. This smoothing is demonstrated by the red curve in Fig. B1, and what is performed throughout this work. In doing so, the general trends for the evolution in the scattering coefficients are more clearly visible.

Importantly, while we could minimize these noisy features by either analysing larger images or increasing the cadence of our light-cone sampling and considering an appropriately narrower box-car filter, the advantage of our current scheme is that it also conveniently mimics the sample variance error. That is, we have only considered one single realization rather than considering a range of different random initial conditions. For example, in Fig. 6, we instead show the mean S_1 scattering coefficients (black solid curve) recovered after averaging across 30 different realizations of our fiducial model, which closely matches the shape and amplitude of the red curve in Fig. B1. Thus our current box-car smoothing approach for illustrating the evolution of the scattering coefficients is sufficient for this work.

APPENDIX C: RANDOMIZING PHASE INFORMATION

In Section 3.3.3, we introduced the S_2 scattering coefficients, highlighting that these contain important non-Gaussian information about the original image. These S_2 coefficients are correlated with their S_1 counterparts, thus in Fig. 4 we explored their evolution with redshift with the de-correlated second-order coefficients (S_2/S_1). These de-correlated second-order coefficients showed additional information that we interpret as being the non-Gaussian signal from the input 21-cm image.

As the S_2 coefficients are obtained by the convolution of the first-order filtered images of the original 21-cm signal, we can verify that the recovered information is truly the non-Gaussian nature of the cosmic signal by randomizing the phase information prior to the second convolution by the wavelet filter. In Fig. C1 we demonstrate this for all combinations of the de-correlated S_2 coefficients for our fiducial model. The red solid curves corresponds to the S_2 coefficients calculated directly from the simulated 21-cm signal, whereas the blue curves correspond to the coefficients after we randomized the phase information in the first-order filtered images before convolution by the second-order family of wavelet filters. Clearly, we see that if we randomize the phase information in the first-order filtered images, the secondary convolution (i.e. S_2) does not yield any information other than a mean signal. Thus, the S_2 coefficients are indeed accessing the non-Gaussian information from the input 21-cm signal.

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022

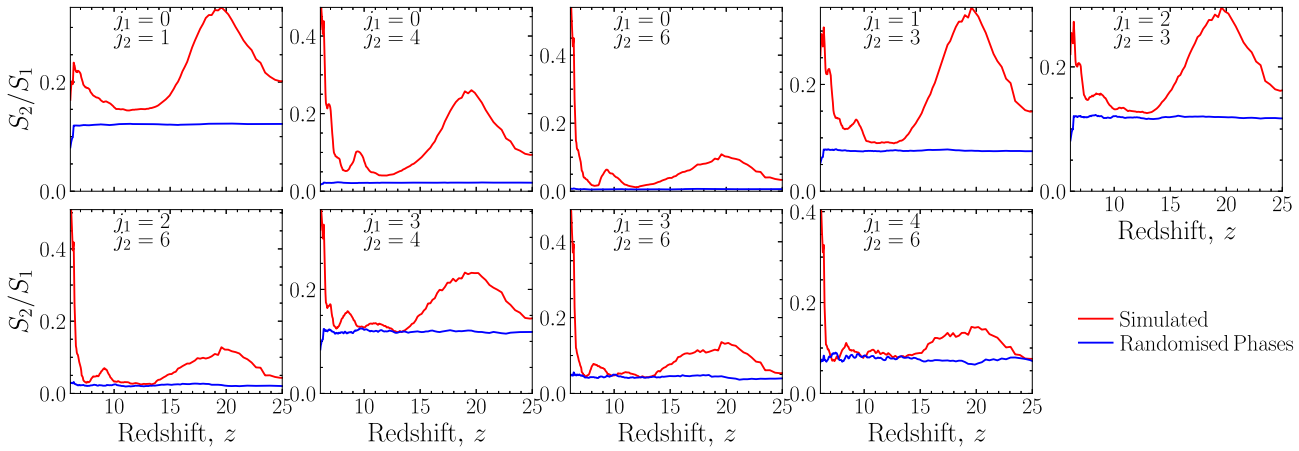


Figure C1. A demonstration of the additional information contained in the de-correlated S_2 coefficients (dividing out the dependence from the S_1 coefficients) compared to the case when the phase information in the first-order filtered images is randomized. Here, we only show the evolution in a select subset of de-correlated S_2 coefficients for our fiducial model.

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

Downloaded from https://academic.oup.com/mnras/article/513/2/1719/6565815 by University of Melbourne Library user on 21 October 2022