Inferring the properties of the sources of reionization using the morphological spectra of the ionized regions
Gazagnes, Simon; Koopmans, Léon V. E.; Wilkinson, Michael H. F.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stab107

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Gazagnes, S., Koopmans, L. V. E., & Wilkinson, M. H. F. (2021). Inferring the properties of the sources of reionization using the morphological spectra of the ionized regions. Monthly Notices of the Royal Astronomical Society, 502(2), 1816–1842. Advance online publication.
https://doi.org/10.1093/mnras/stab107

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Inferring the properties of the sources of reionization using the morphological spectra of the ionized regions

Simon Gazagnes 1,2,3*, Léon V. E. Koopmans 2 and Michael H. F. Wilkinson 3

1 Energy and Sustainability Research Institute Groningen, University of Groningen, PO Box 1103, NL-9700 CC Groningen, the Netherlands
2 Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands
3 Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, PO Box 407, NL-9700 AK Groningen, the Netherlands

ABSTRACT

High-redshift 21-cm observations will provide crucial insights into the physical processes of the Epoch of Reionization. Next-generation interferometers such as the Square Kilometer Array will have enough sensitivity to directly image the 21-cm fluctuations and trace the evolution of the ionizing sources. In this work, we develop an inferential approach to recover the sources and IGM properties of the process of reionization using the number and, in particular, the morphological pattern spectra of the ionized regions extracted from realistic mock observations. To do so, we extend the Markov Chain Monte Carlo analysis tool 21CMMC by including these 21-cm tomographic statistics and compare this method to only using the power spectrum. We demonstrate that the evolution of the number-count and morphology of the ionized regions as a function of redshift provides independent information to disentangle multiple reionization scenarios because it probes the average ionizing budget per baryon. Although less precise, we find that constraints inferred using 21-cm tomographic statistics are more robust to the presence of contaminants such as foreground residuals. This work highlights that combining power spectrum and tomographic analyses more accurately recovers the astrophysics of reionization.

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

1 INTRODUCTION

The formation of the first stars and galaxies through gravitational instability of small density fluctuations, several hundred million years after recombination, marks the beginning of a key phase transition in the Universe history referred to as the Cosmic Dawn. These first light sources emitted photons with enough energy to ionize the neutral hydrogen, which propagated in the intergalactic medium (IGM) and progressively reionized the entire Universe. This latter era is called the Epoch of Reionization (EoR), where the transition from the Cosmic Dawn is rather ill defined (see e.g. Ciardi & Ferrara 2005; Morales & Wyithe 2010; Pritchard & Loeb 2012; Furlanetto 2016; Dayal & Ferrara 2018 for reviews). Many unknowns remain about the nature of these sources, and especially their efficiency in re-ionizing the surrounding neutral-hydrogen environment. Over the past decade, both observational and theoretical studies have suggested that a population of low-mass, very star-forming galaxies with an average escape fraction of ionizing photons of $f_{\text{esc}} = 0.1–0.2$, were the dominant suppliers of ionizing photons during the EoR (Ouchi et al. 2009; Bouwens et al. 2012; Robertson et al. 2013; Dressler et al. 2015; Finkelstein et al. 2019; Mason et al. 2019; Trebitsch et al. 2020). Nevertheless, the contribution of brighter sources is still actively discussed (Fontanot, Cristiani & Vanzella 2012; Fontanot et al. 2014; Robertson et al. 2015; Madau & Haardt 2015; Mitra, Choudhury & Ferrara 2018; Naidu et al. 2020; Dayal et al. 2020).

Further observational studies are required to constrain the escape fraction of these sources during reionization. Directly measuring $f_{\text{esc}}$ at high redshifts is very challenging because of the limited mean free path of these photons. Hence, indirect proxies are needed to estimate the ionizing efficiency of the sources of reionization, and some progress has been recently made in that direction using analogues at lower redshift (e.g. Schaerer et al. 2016; Verhamme et al. 2017; Chisholm et al. 2018, 2020; Henry et al. 2018; Gazagnes et al. 2018, 2020; Wang et al. 2019; Cen 2020; Izotov et al. 2020).

The detection of the 21-cm line resulting from the spin-flip of the neutral Hydrogen electron can provide complementary information to investigate the nature of ionizing sources and the propagation of ionizing fronts. First-generation interferometers, such as the Murchison Widefield Array1 (MWA; Tingay et al. 2013; Bowman et al. 2013), the Low-Frequency Array2 (LOFAR; van Haarlem et al. 2013), the Giant Metrewave Radio Telescope (GMRT)3 (Swarup et al. 1991), or the Precision Array for Probing the Epoch of

* E-mail: s.r.n.gazagnes@rug.nl

1http://www.mwatelescope.org
2http://www.lofar.org
3http://www.ncra.tifr.res.in/ncra/gmrt

© The Author(s) 2021.

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://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Reionization\(^4\) (PAPER; Parsons et al. 2010), aim at detecting the 21-cm fluctuations during the EoR. While no detection has been achieved yet, they yield tremendous improvements in our understanding of interferometric observations. Recently, improved upper limits on a statistical detection of the 21-cm fluctuations (at the 2\(\sigma\) level) using the power spectrum have been published in the literature. Trott et al. (2020) reported \(\Delta_{21}^2 \leq (43 \text{ mK})^2\) at \(k = 0.14 \text{ h Mpc}^{-1}\) and \(z = 6.5\) using 110 h of MWA observations and Mertens et al. (2020) found \(\Delta_{21}^2 \leq (73 \text{ mK})^2\) at \(k = 0.075 \text{ h Mpc}^{-1}\) and \(z = 9.1\) using 141 h of LOFAR observations. The latter results already put some light on the astrophysics of reionization by ruling out exotic scenarios incompatible with this upper limit (Ghara et al. 2020; Mondal et al. 2020; Greig et al. 2021).

The progress with current instruments, but also the development of the second-generation interferometers such as the Hydrogen Epoch of Reionization Array\(^5\) (HERA; DeBoer et al. 2017) and the Square Kilometre Array\(^6\) (SKA; Mellema et al. 2013; Koopmans et al. 2015) will progressively tighten our understanding of the EoR by either yielding deeper limits on the power spectrum of the 21-cm fluctuations, or directly imaging it with redshift. Additional insights can be extracted from the non-Gaussian part of the 21-cm fluctuations, which are not encoded in the power spectrum. Over the past decades, a large number of theoretical studies have introduced new statistical formalism to extract information from the 21-cm signal. Shimabukuro et al. (2016), Majumdar et al. (2018), Watkinson et al. (2019), or Hutter et al. (2020) have used the 21-cm bispectrum and shown that it could provide valuable insights about the ionization topology and the size distribution of the ionized structures. Similarly, Gorce & Pritchard (2019) have shown that comparable insights could be extracted using 3-points correlations functions, Watkinson & Pritchard (2014) and Banet et al. (2020) investigated the skewness and kurtosis of the 21-cm intensity probability distribution function (PDF) and have shown that these non-Gaussian measurements can trace the astrophysics of reionization.

Finally, HERA and SKA should have enough sensitivity to map the 21-cm fluctuations and provide images of the evolution of the ionizing fronts during the EoR. The use of these 21-cm tomographic images has already been shown useful to analyse the size statistics of the ionized or neutral regions (Kakiichi et al. 2017; Giri et al. 2018a, 2019), the morphology of the brightness temperature fluctuations (Chen et al. 2019; Kapahi, Chingangbam & Appleby 2019), or the topology of the ionizing field (Elbers & van de Weygaert 2019; Giri & Mellema 2020). All these studies have emphasized that alternative approaches to the power spectrum could provide very valuable insights to understand the astrophysics of cosmic reionization, and could be used to break the degeneracy between models with similar power spectra (Kakiichi et al. 2017).

However, further analysis is required to assess how these approaches can be incorporated within a 21-cm analysis tool to quantify their robustness in inferring the astrophysical information lying in the 21-cm fluctuations. Recently, several studies explored the use of deep-learning approaches to infer the astrophysics of the process of reionization using 21-cm images (e.g. Gillet et al. 2019; Hassan et al. 2019). In this paper, we investigate how independent statistics extracted from these images can be included in a Bayesian statistical inference framework. To do so, we extend 21CMMC (Greig & Mesinger 2015, 2017), a Markov Chain Monte Carlo (MCMC) analysis tool, originally designed to quantify reionization properties using the power spectrum of the 21-cm fluctuations. 21CMMC has been widely used over the past years to investigate the theoretical constrains set by different 21-cm experiments (Greig & Mesinger 2015, 2017; Park et al. 2019), break the degeneracy between different reionization topologies (Binnie & Pritchard 2019), propose optimal designs for 21-cm instruments (e.g. with the SKA; Greig, Mesinger & Koopmans 2020a), or rule out astrophysical models incompatible with the recent upper limits obtained from LOFAR (Greig et al. 2021) and MWA (Greig et al. 2020) data, respectively. In this work, we adapt it to assess the constraints set by 21-cm tomographic statistics and investigate their robustness compared to theoretical results obtained using the power spectrum. To achieve this, we create realistic mock observations using the point spread function and noise profile of the SKA1-Low telescope configuration, assuming 1000 h of observations at redshifts 10, 9, and 8. We use DISCCOFAN (Gazagnes & Wilkinson 2019) to extract the number-count and the morphological pattern spectra (Maragos 1989; Wilkinson & Westenberg 2001; Urbach, Roerdink & Wilkinson 2007; Westenberg, Roerdink & Wilkinson 2007) encoding the shape characteristics of the individual ionized regions. These statistics provide an intuitive approach to investigate the evolution of the ionizing field at different redshifts, and to infer the sources and IGM properties of the process of reionization.

This work is organized as follow. Section 2 details the implementation of 21CMMCFAST, the semi-numerical code embedded in 21CMMC that simulates the redshifted 21-cm signal during the EoR. Section 3 explains the mathematical description of the statistics extracted from the 21-cm tomographic images, and Section 4 analyses how these statistics vary with redshift, different reionization scenarios, or are impacted by the characteristics of the interferometer. Section 5 details the 21CMMC set-up, and Section 6 shows the results of our MCMC analysis. Section 7 further assesses how 21-cm tomographic statistics robustly trace the underlying astrophysics of the reionization and can combined with power spectrum analyses. Finally, Section 8 summarizes our main conclusions. In this paper, we assume a \(\Lambda\)CDM Universe with cosmological parameters values \(\Omega_{\Lambda} = 0.692, \Omega_m = 0.048, H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}, \) and \(\sigma_8 = 0.81\).

## 2 SIMULATIONS AND INSTRUMENT SENSITIVITY

We use 21CMMC (Greig & Mesinger 2015), a Bayesian parameter estimation tool which combines a Markov Chain Monte Carlo (MCMC) algorithm with 21CMMCFAST (Mesinger, Furlanetto & Cen 2011), to explore how the shape of the ionized regions extracted from 21-cm tomographic images can be used to infer the properties of the sources of reionization. Section 2.1 details the 21CMMCFAST parametrization and Section 2.2 describes the interferometer characteristics adopted in this work.

### 2.1 21CMMCFAST

21CMMCFAST\(^7\) is a publicly available semi-numerical code that simulates the evolution of the 21-cm signal during the Cosmic Dawn and the EoR using an analytical model introduced in Furlanetto, Zaldarriaga & Hernquist (2004b). Its implementation is detailed by Mesinger et al. (2011). 21CMMCFAST first generates realizations of the evolved

---

\(^4\)http://eor.berkeley.edu

\(^5\)https://reionization.org/

\(^6\)https://www.skatelescope.org/

\(^7\)https://github.com/andreimesinger/21cmFAST
baryonic and velocity density field using the Zel’dovich approximation (Zel’Dovich 1970) at different redshifts. The brightness temperature contrast, $\delta T_b$, which is proportional to the 21-cm signal intensity and evaluated relative to the temperature of the Cosmic Microwave Background (CMB), is then computed as (Furlanetto, Oh & Briggs 2006):

$$\delta T_b \approx 27(1-x_{\text{ion}})(1+\delta_b)\left(\frac{H(z)}{dV/dt + H(z)}\right)\left(1 - \frac{T_{\text{CMB}}}{T_S}\right) \times \left(\frac{1+z}{10}\right)^{\frac{1}{2}} \left(\frac{\Omega_m h^2}{0.023}\right) \ln K,$$

(1)

where $x_{\text{ion}}$ is the global ionized fraction of the IGM in the universe, $\delta_b$ is the fractional overdensity in baryons, $H(z)$ is the Hubble parameter, $dV/dt$ is the comoving velocity gradient along the line of sight, $T_S$ is the spin temperature, and $T_{\text{CMB}}$ is the CMB temperature. $\delta T_b$ is evaluated at a redshift $z$, such as $z = v_0/v - 1$ where $v_0$ is the rest frame 21-cm frequency (1420 MHz).

We include the effects of peculiar velocities and assume that the spin temperature is above the CMB temperature ($T_S > T_{\text{CMB}}$) such that, in the absence of foregrounds or instrumental effects, the presence of ionized structures directly relates to the absence of 21-cm signal ($\delta T_b = 0$ mK). This assumption, while considered valid during most of the reionization ($0 < z < 10$, Chen & Miralda-Escudé 2004; Baek et al. 2010), could break down during the early stages, or around under-dense regions where the IGM is not sufficiently heated. We discuss the impact of this hypothesis in Section 7.3.

The ionized regions are recovered using an excursion set formalism that compares the number of ionizing photons per time-step to the density of baryons in regions of decreasing scales $R$ (Mesinger et al. 2011). A given cell, $x$, is identified as ionized if it fulfills the following criterion:

$$\xi f_{\text{coll}}(x, z, R, M_{\text{min}}) > 1,$$

(2)

where $f_{\text{coll}}(x, z, R, M_{\text{min}})$ represents the collapse fraction estimated within a given scale $R$ around the cell $x$ at a redshift $z$, and depends on the minimum mass of the star-forming halo $M_{\text{min}}$ (Press & Schechter 1974). The parameter $\xi$ characterizes the ionizing efficiency of the sources and is further detailed in Section 2.1.1. 21CMFAST additionally includes partially ionized cells by setting their ionizing fraction to the value of $\xi f_{\text{coll}}(x, z, R, M_{\text{min}})$ when $R$ reaches the minimum scale.

The focus of this study is to explore how the morphological properties of the ionized regions can be used to infer the properties of the ionizing sources. For this proof-of-concept work, we restrict ourselves to a reionization parameterization with three parameters: the ionizing efficiency of the reionization sources ($\xi$), the mean free path of ionizing photons ($R_{\text{mfp}}$), and the minimum virial temperature of dark haloes hosting star-forming galaxies ($T_{\text{vir}}^{\text{min}}$). We detail each of them in the following sub-sections.

2.1.1 The ionizing efficiency $\xi$

The ionizing efficiency of the galaxies during the EoR is defined as

$$\xi = 30 \frac{f_{\text{esc}}}{0.3} \frac{f_e}{0.05} \frac{4000}{1 + n_{\text{rec}}} \frac{2}{N_V},$$

(3)

where $f_{\text{esc}}$ is the escape fraction of ionizing photons, $f_e$ is the fraction of galactic gas in stars, $N_V$ is the number of ionizing photons produced per baryon within a star, and $n_{\text{rec}}$ is the typical number of recombinations of a hydrogen atom. Equation (3) assumes a constant ionizing efficiency for all galaxies formed in haloes with sufficient mass ($M > M_{\text{min}}$). We note that this results in a sharp drop of the non-ionizing UV-luminosity function for haloes with mass $M < M_{\text{min}}$ (see fig. 1 in Greig & Mesinger 2015).

Throughout, we consider a range of ionizing efficiencies from 1 to 250. As discussed in Greig & Mesinger (2017), this range explores a large variety of EoR models, with sources having both very low and large ionizing efficiency. The choice of $\xi$ strongly affects the evolution of the EoR, such that increasing $\xi$ will lead to a faster reionization process. Nevertheless, its impact also depends on the minimal virial temperature $T_{\text{vir}}^{\text{min}}$, since the combination of both these parameters sets the average ionizing budget during reionization.

2.1.2 The minimum virial temperature $T_{\text{vir}}^{\text{min}}$

The collapse fraction $f_{\text{coll}}$ is the fraction of mass in a given volume enclosed in haloes of individual mass larger or equal to $M_{\text{min}}$. In 21CMFAST, $M_{\text{min}}$ is defined through the minimum virial temperature of star-forming haloes, $T_{\text{vir}}^{\text{min}}$, and expressed as (Barkana & Loeb 2001):

$$M_{\text{min}} = 10^8 h^{-1} \left(\frac{\mu}{0.6}\right)^{-3/2} \left(\frac{\Omega_m \Delta_c}{\Omega_m^{18} \sigma^2}\right)^{1/2} \left(\frac{T_{\text{vir}}^{\text{min}}}{1.98 \times 10^4 K}\right)^{3/2} \left(\frac{1+z}{10}\right)^{-3/2} M_\odot,$$

(4)

with $\mu$ the mean molecular weight, $\Omega_m^{18}$ the evolved $\Omega_m$ at redshift $z$, and $\Delta_c = 18\pi^2 + 82d - 39d$ with $d = \Omega_m^{18} - 1$. As mentioned above, the choice of $T_{\text{vir}}^{\text{min}}$ impacts the position of the sharp drop in the UV luminosity function ($\xi = 0$ if $T_{\text{vir}} < T_{\text{vir}}^{\text{min}}$). We adopt the same prior as used in Greig & Mesinger (2017), such that $T_{\text{vir}}^{\text{min}}$ varies from $10^4$ to $10^5$ K. The lower limit of $T_{\text{vir}}^{\text{min}}$ corresponds to the minimum temperature to trigger efficient atomic cooling (Barkana & Loeb 2001; Kimm et al. 2017), while the upper limit is chosen to be consistent with the host halo mass of high redshift Lyman break galaxies observations (Kuhlen & Faucher-Giguère 2012; Barone-Nugent et al. 2014).

2.1.3 The mean free path of photons $R_{\text{mfp}}$

The growth of the ionized regions is a dynamic process, which mainly depends on the recombination rate of the hydrogen atoms and the propagation distance of the ionizing photons. The balance between both processes sets the physical size of the ionized bubbles.

The current 21CMMC version explicitly computes inhomogeneous recombinations. However, we use a simplification that assumes a maximum horizon for the ionizing photons in the IGM. This parameter is denoted by $R_{\text{mfp}}$ and fixes the maximum smoothing scale used in equation (2). Its impact is most significant when the ionized regions size becomes closer to $R_{\text{mfp}}$ (Greig & Mesinger 2017). Small values of the mean horizon of the ionizing photons (e.g. <10 Mpc) can delay the late stage of reionization. On the other hand, Greig & Mesinger (2017) note that values larger than 15 Mpc have little impact on the power spectrum of the 21-cm fluctuations because the fusion of the ionized regions is then the dominant growing process.

We adopt a flat prior of $R_{\text{mfp}} \in [5, 25]$ Mpc, similar to Greig & Mesinger (2015) and Greig & Mesinger (2017).

2.2 Interferometer characteristics

We create mock observations using the point spread function (PSF) and the expected noise from 1000 h of observations with
Table 1. Characteristics of the interferometer SKA1-Low and observation parameters used in this work.

| Parameter          | Values         |
|--------------------|---------------|
| $N_{\text{ant}}$   | 512           |
| $T_{\text{sys}}$   | 60(1163)K     |
| $A_{\text{eff}}$   | 969 m$^2$     |
| $\Delta t$         | 10 s          |
| $I_{\text{sky}}$   | 6 h           |
| $t_{\text{obs}}$   | 1000 h        |

SKA1-Low. The current configuration of SKA1-Low consists of 512 stations, where 224 of them are randomly distributed in a central circular core of 500 m in radius. The remaining 288 stations are placed in 36 clusters located in three spiral arms out to a radius of 50 km from the central core. The noise maps are computed using the python package TOOLS21cm (Giri, Mellema & Jensen 2020) which uses the formalism detailed in Ghara et al. (2017) and Giri et al. (2018a). The system noise per visibility is a Gaussian random variable with mean zero, and variance $\sigma^2$ defined as

$$\sigma^2 = \frac{\sqrt{2k_B T_{\text{sys}}}}{A_{\text{eff}} \sqrt{\Delta t / \Delta \nu}}^2,$$

where $k_B$ is the Boltzmann constant, $T_{\text{sys}}$ is the telescope temperature that accounts for the receiver and sky temperature ($T_{\text{sky}} \Delta \nu^{-2.55}$), $A_{\text{eff}}$ is the effective collecting area of the individual receivers, $\Delta \nu$ is the frequency resolution of the data, and $\Delta t$ is the integration time. $T_{\text{sys}}$ and $A_{\text{eff}}$ are fixed by the SKA1-Low technical design. Additionally, we adopt $\Delta t = 10$ s per visibility, 6 h of observing per day ($t_{\text{obs}}^{\text{skymap}}$), and a total observation time ($t_{\text{obs}}^{\text{tot}}$) of 1000 h. The noise maps are obtained by generating the uv maps $U(u,v)$ using the baseline distribution in the gridded uv plane. Then, a noise cube is derived in the Fourier domain using Gaussian distributed values with mean zero and variance $\sigma^2$ for both the real and imaginary parts. The cells in the uv-plane that are not sampled are set to 0, and we divide the noise values in the remaining bins by $\sqrt{2k_B T_{\text{sys}}/A_{\text{eff}} \Delta \nu}$. Finally, the noise values are further scaled down by a factor $I_{\text{sky}}^{\text{obs}} I_{\text{sky}}^{\text{source}}$. In reality, the baseline distribution varies in frequency such that a distinct uv map should be generated for each channel. However, in practice, these differences are small when the frequency bandwidth of the observation is relatively narrow. Consequently, we generate noise cubes using a single uv distribution which corresponds to the central frequency of the observation. Table 1 summarizes the interferometer characteristics and observation parameters for this work.

### 3 THE MORPHOLOGY OF THE IONIZED REGIONS

The power spectrum is a well-defined approach to analyse the excess of the 21-cm signal at different scales and already provides valuable clues about the evolution of the ionization fronts during the EoR. However, it is blind to the non-Gaussianities of the 21-cm fluctuations (Furlanetto, Zaldarriaga & Hernquist 2004a), while the latter can provide important information to understand and constrain the underlying physical processes during reionization (Koopmans et al. 2015). A wide range of novel approaches, using higher-order and topological statistics, have been investigated to extract the non-Gaussian information lying in the 21-cm fluctuations (see Section 1 and Greig 2019). In this work, we define a new approach based on a pure morphological description of the ionized regions. This method provides a simplistic but intuitive description of the shape of the ionized regions, which can be efficiently extracted from noisy and PSF-convolved 21-cm image cubes. Thus, it is well suited for a Bayesian framework analysis that requires comparing thousands of models in a reasonable amount of time. Section 3.1 introduces these 21-cm tomographic statistics, and Section 3.2 describes DISCOFAN, a massively parallelized image processing tool designed to extract the morphology of the structures observed in image and image cubes.

#### 3.1 Morphological attributes

Several studies already highlighted theoretical differences in the morphology of the ionized regions for diverse reionization models (e.g. Kakiichi et al. 2017; Giri et al. 2018a; Kapahiuia et al. 2019; Gorce & Pritchard 2019). In this work, we use the morphological pattern spectra (Maragos 1989) of the ionized regions extracted from 21-cm image cubes. This mathematical formalism is based on scale-invariant morphological attributes derived from the moment-of-inertia matrix of each ionized region. This approach has been previously introduced for detection and extraction of anomalies (aneurysms and stenoses) in blood-vessels (Wilkinson & Westenberg 2001; Urbach et al. 2007; Westenberg et al. 2007). For a discrete case (e.g. quantized images or volumes), the moment of inertia matrix $I_o$ of a three-dimensional (3D) object $O$ is defined by

$$I_o = \begin{bmatrix} I_{oxx} & I_{oyy} & I_{ozz} \\ I_{oyy} & I_{oyy} & I_{ozz} \\ I_{ozz} & I_{ozz} & I_{ozz} \end{bmatrix}$$

with

$$I_{oxx} = \sum_o (x - \bar{x})^2 + V_o/12,$$

$$I_{oyy} = \sum_o (y - \bar{y})^2 + V_o/12,$$

$$I_{ozz} = \sum_o (z - \bar{z})^2 + V_o/12,$$

$$I_{xyx} = \sum_o (x - \bar{x}) \times (y - \bar{y}),$$

$$I_{yz} = \sum_o (x - \bar{x}) \times (z - \bar{z}).$$

where $[x, y, z]$ are the coordinates of the voxels belonging to $O$, $[\bar{x}, \bar{y}, \bar{z}]$ are the coordinates of its centre of mass, and $V_o$ is its volume. This formalism is derived from the continuous case where $I_o^{\text{skymap}} = \iiint x^2 dV$. The factor $V_o/12$ is introduced to account for the moment of inertia of the individual cubic voxels. From $I_o$ and its eigenvalue decomposition, we extract four moments-invariant attributes encoding different morphological properties of a 3D geometrical structure: the elongation $E$, the flatness $F$, the non-compactness $N$, and the sparseness $S$.

**Non-compactness $N$** is defined as

$$N = \frac{\text{Tr}(I_o)}{V_o^{3/5}},$$
where $\text{Tr}(L)$ is the trace of the moment of inertia tensor $L$. The unit of $L$ is consistent with a length $L$ to the power 5, while the volume scales with $L^3$, hence the ratio $\frac{\text{Tr}(L)}{V^{5/3}}$ is scale-invariant. $N$ reaches a minimum value of 0.25 for perfectly spherical objects and increases for complex asymmetric structures.

**Elongation and flatness:** $\mathcal{E}$ and $\mathcal{F}$ are derived using the ratio of the eigenvalues of $L_o$, referred as $\lambda_1$, $\lambda_2$, and $\lambda_3$ with $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$, such that

$$\mathcal{E} = \frac{|\lambda_1|}{|\lambda_2|}$$

and

$$\mathcal{F} = \frac{|\lambda_2|}{|\lambda_3|}. \quad (8)$$

The eigenvalues $\lambda$ represent the variance of the coordinates of an object along its main axes, scaled by its volume. In other words, they probe the spatial expansion of a 3D structure along the directions of maximal growth, such that $|\lambda_1| = |\lambda_2| = |\lambda_3|$ for perfectly spherical objects or $|\lambda_1| > |\lambda_2| = |\lambda_3|$ for elongated shapes. Thus, $\mathcal{E}$ and $\mathcal{F}$ refer to the divergence between the object growth in different directions. Combined, they provide information about the eccentricity of a geometrical shape. From the description above, it comes trivially that spherical objects have $\mathcal{E} = \mathcal{F} = 1$, cylindrical or cigar-shaped objects have $\mathcal{E} > 1$ and $\mathcal{F} \approx 1$, and more complex tri-axial shapes have $\mathcal{E}$ and $\mathcal{F}$ larger than 1. The analysis of $\mathcal{E}$ and $\mathcal{F}$ is useful to reveal a dominant direction in the growth of the ionized regions, and therefore provide information whether the propagation of the ionizing radiation is isotropic.

**Sparseness:** $\mathcal{S}$ is defined as

$$\mathcal{S} = \frac{\pi^{20/3} \prod_{i=1}^{3} |\lambda_i|}{6 V^{10/3}}. \quad (9)$$

It corresponds to the ratio of the expected volume computed using the coordinate variance along the three main axes (product of the three eigenvalues), versus the actual volume, obtained by summing over all the voxels belonging to the object. Broadly speaking, $\mathcal{S}$ relates to the filling factor of an object. It reaches 1 for filled spheres or ellipsoids, 1.127 for solid rectangular blocks, and increases as the filling factor of an object. It reaches 1 for filled spheres or cubes, 0.707 for cubes, and 0.125 for solid spheres. For elongated shapes, it increases as the filling factor of an object.

Overall, these four morphological attributes provide different information about the shape of the ionized regions. $\mathcal{E}$ and $\mathcal{F}$ highlight the structures eccentricity, while $\mathcal{N}$ and $\mathcal{S}$ describe the shape complexity, through its degree of asymmetry and porosity. Throughout this work, we also consider the total number of individual ionized regions observed, $n_{\text{sub}}$, and the volume of the ionized regions, $V$. Both measures are reliable probes of the underlying astrophysics of reionization (e.g. Kakiichi et al. 2017; Giri et al. 2018a). In the next section, we detail the algorithm used to extract $n_{\text{sub}}$, the volume, and four morphological attributes of each ionized region directly from the 3D 21-cm images.

### 3.2 Extracting morphological attributes with DISCCOFAN

To extract the number and the morphological pattern spectra of the individual ionized regions from 21-cm tomographic images, we use DISCCOFAN$^{10}$ (DIStributed Connected COmponent Filtering and ANalysis; Gazagnes & Wilkinson 2019; Gazagnes & Wilkinson submitted), a recent method that combines image processing and mathematical morphology techniques to efficiently find, select, and analyse the connected structures in two- and three-dimensional data sets. Contrary to classical pixel-based approaches, DISCCOFAN works on a region-based representation of the data referred to as a component tree (Salembier, Oliveras & Garrido 1998). During the first step (the flooding), the algorithm goes through all the voxels and group them in connected components. The typical clustering rule to define a connected component is to select all the neighbours’ voxels above a given threshold set. The nested relations between the connected components at different intensity levels are stored in a tree structure, which allows us to perform complex multiscale operations on this optimized hierarchical representation. The use of similar tree-techniques has been successfully used to improve the detection of faint-objects in optical astronomical surveys (Teeninga et al. 2016; Haigh et al. 2021).

The procedure to derive the ionized region statistics from noisy, low-resolution 21-cm image cubes remains the same for all the tests performed in this work. We first extract the ionized regions directly from the 21-cm mock observations, using an independent thresholding technique (see more details in Section 4.3). DISCCOFAN is then applied on the resulting segmented image cubes encoding the recovered ionized regions such that voxels with a value of 0 and 1 are defined as neutral and ionized, respectively. DISCCOFAN finds and extracts all the individual connected components that have spatially connected voxels with a value equal to 1, assuming that each voxel has 26 neighbours in three dimensions. The total number of connected components found corresponds to the number of ionized regions $n_{\text{sub}}$. The volume of each region is derived by summing over all the voxels that belong to it, and by multiplying this number with the volume of the voxel. Finally, DISCCOFAN simultaneously returns the morphological attributes $\mathcal{E}$, $\mathcal{F}$, $\mathcal{N}$, and $\mathcal{S}$ using the equations presented in Section 3. As applied on binary image cubes, the approach of DISCCOFAN is very similar to a classical friend-of-friend algorithm. However, this tool provides a powerful framework to analyse the full range of 21-cm intensities, and we plan to explore this aspect in future works.

DISCCOFAN can process very large data sets (>100 Gigapixels; Gazagnes & Wilkinson submitted). It is massively parallelized using shared and distributed memory techniques, thus is fast and efficient to analyse the properties of connected structures in any data set. Using a single CPU process, it takes around 10 s to create a model on a 128$^3$ box with 21CMFAST, while it takes less than a second to extract the morphological pattern spectra with DISCCOFAN. This asset makes this tool well suited to be incorporated in a Bayesian inferential framework, where thousands of models need to be processed in a reasonable time.

### 4 MORPHOLOGY OF THE IONIZED REGIONS

In this section, we investigate the evolution of the morphological pattern spectra of the ionized regions as a function of redshift (Section 4.1), different reionization scenarios (Section 4.2), and the impact of the expected noise profile and PSF of SKA1-Low on the recovered statistics (Section 4.3).

#### 4.1 Evolution during reionization

We first investigate the evolution of the morphological attributes ($V$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{N}$, and $\mathcal{S}$) during reionization using simulated 21-cm maps. We generate co-eval boxes of 512$^3$ cells with 2 Mpc resolution using 21CMFAST with the set of parameters $\xi = 30$, $T_{\text{vir}}^{\text{min}} = 5.10^4 K$, and $R_{\text{mp}} = 15$ Mpc (referred as the FAINT SOURCES

---

10https://github.com/sgazagnes/disccofan
Inference using the reionization morphology

Figure 1. Evolution of the morphological attributes of the ionized regions for the FAINT SOURCES model simulations ($512^3$ Mpc$^3$), for three snapshots corresponding to an ionized fraction $x_{\text{ion}}$ of 0.1, 0.3 and 0.7, respectively. The diagonal and lower panels show, respectively, the 1D and 2D marginalized PDFs. The contours enclose successively 68, 95, and 99.7 per cent of the ionized region distributions. The diamonds show the morphological attribute values of the largest ionized region in each cube.

This model characterizes a cosmic reionization dominated by a population of several sources with a relatively low ionizing efficiency (Mesinger, Greig & Sobacchi 2016; Greig & Mesinger 2017). We consider three different redshifts corresponding to snapshots, where the ionizing fraction ($x_{\text{ion}}$) in the box is successively 0.1, 0.3, and 0.7. This choice enables us to investigate the evolution of the ionized regions morphology for three different reionization stages, where the universe is mostly dominated by (1) individual ionized regions; (2) ionized filaments resulting from the fusion of the ionized structures; and (3) neutral islands spanning through a main ionized percolated region that fills most of the universe (Chen et al. 2019). For this noise-free exploratory case, we extract the ionized regions from the synthetic 21-cm maps using $\delta T_b = 0$ mK as a threshold, and we apply DISCCOFAN on the resulting 3D binary images.

Fig. 1 shows the resulting distributions of the five morphological attributes for the three snapshots as a corner plot. The diagonal and lower panels show, respectively, the one-dimensional (1D) and 2D marginalized probability density functions (PDFs). The contours in the lower panels enclose successively 68, 95, and 99.7 per cent of the ionized regions distributions. Additionally, we highlight the morphological properties of the largest ionized region in each box using diamonds. The largest scales are typically more sensitive to sample variance. However, resolving the properties of individual ionized regions with SKA is expected to be simpler as their size grows larger (Ghara & Choudhury 2020). Hence, it should be possible to extract valuable information about the astrophysics of the reionization process using the morphological properties of these large structures.

We observe that the typical size of the ionized regions shrinks as reionization progresses, while the largest objects grow larger. The probability to observe several large ionized structures decreases because more and more of these regions fuse into a larger percolated cluster, spanning the whole box, with boundaries that are artificially infinite. Fig. 1 shows that a large ionized region of $\approx 10^7$ Mpc$^3$ is already present at $x_{\text{ion}} = 0.1$, suggesting that a percolated object already formed at this early stage. This is consistent with Furlanetto & Oh (2016) who found that the formation of a percolated cluster typically happens when the universe is still only mildly ionized ($x_{\text{ion}}$ between 0.1 and 0.2).

Fig. 11 shows the resulting distributions of the five morphological attributes for the three snapshots as a corner plot. The diagonal and lower panels show, respectively, the one-dimensional (1D) and 2D marginalized probability density functions (PDFs). The contours in the lower panels enclose successively 68, 95, and 99.7 per cent of the ionized regions distributions. Additionally, we highlight the morphological properties of the largest ionized region in each box using diamonds. The largest scales are typically more sensitive to sample variance. However, resolving the properties of individual ionized regions with SKA is expected to be simpler as their size grows larger (Ghara & Choudhury 2020). Hence, it should be possible to extract valuable information about the astrophysics of the reionization process using the morphological properties of these large structures.

We observe that the typical size of the ionized regions shrinks as reionization progresses, while the largest objects grow larger. The probability to observe several large ionized structures decreases because more and more of these regions fuse into a larger percolated cluster, spanning the whole box, with boundaries that are artificially infinite. Fig. 1 shows that a large ionized region of $\approx 10^7$ Mpc$^3$ is already present at $x_{\text{ion}} = 0.1$, suggesting that a percolated object already formed at this early stage. This is consistent with Furlanetto & Oh (2016) who found that the formation of a percolated cluster typically happens when the universe is still only mildly ionized ($x_{\text{ion}}$ between 0.1 and 0.2).

Corner plots from Figs 1, 2, and 4 have been made using the PYTHON package PAPPY: https://github.com/drphilmarshall/pappy.
The distributions of $E$ and $F$ also provide some interesting insights about the growth of the ionized regions. The elongation and flatness are large when the ionized regions are small, but decrease as their volume gets larger, with $E$ and $F$ tending closer to 1. This suggests that the growth mechanisms of the small structures are likely anisotropic, but the fusion of the individual regions equalizes the spatial extension of the larger merged structures. However, the $E$ and $F$ values derived in the largest ionized region should be taken with caution. Indeed, the box has a finite length, such that the maximal spatial extension of an object is limited by the size of the simulation. Hence, the percolated cluster will always have $E \approx F \approx 1$, because it spans over the whole box.

Interestingly, the 2D PDF of $F$ versus $E$ highlights that only few objects have $E$ and $F = 1$ that supports that the shape of the ionized structures strongly diverges from the idea of spherical bubbles. On the other hand, objects with the largest values of $E$ ($F$) have $F$ ($E$) closer to 1, suggesting that these regions are more extended along a specific direction, and resemble cylindrical or filamentary structures. This outcome is somewhat expected from inside–out reionization, as implemented in 21CMFAST (Furlanetto, Hernquist, along a specific direction, and resemble cylindrical or filamentary ($E$, $F$)) closer to 1, suggesting that these regions are more extended

This can be seen as the 99.7 per cent contours enclose a narrower range of $S$ and $N$, for ionized regions with the same volume at any redshift. This effect is also particularly noticeable on the properties of the percolated cluster in each box. As reionization progresses, these objects become more compact and ‘filled’ because the remaining neutral patches are progressively ionized. Hence, even when the universe is predominantly ionized, comparing the morphology of these percolated clusters might still provide valuable insights about the properties of the ionizing sources. This is further discussed in the next section.

Overall, Fig. 1 emphasizes that the evolution of the morphology of the ionized regions probes the reionization history. Mainly, early reionization stages form complex, eccentric, and overall porous ionized structures due to the fusion of the individual ionized regions along the high-density filaments of the underlying density field. As reionization progresses, these structures continue to grow as their ionizing fronts are propagating. However, their maximal size becomes limited by the merging mechanisms, such that they are more likely to fuse early with the percolated cluster. In the next section, we compare these statistics for two different reionization scenarios.

### 4.2 Morphological spectra for different reionization scenarios

Understanding how the neutral hydrogen was ionized during to reionization requires one to constrain the properties of the sources that dominantly contributed to the ionizing budget of the EoR. Studies generally investigate the impact of different populations of reionization sources, labelled through their ionizing efficiency, the typical mass of the host haloes, or the hardness of their X-ray spectral energy distribution (SED). We use the FAINT SOURCES and BRIGHT SOURCES models, previously introduced in Mesinger et al. (2016) and Greig & Mesinger (2017), and defined by the following sets of parameters:

(i) FAINT SOURCES: $\xi = 30$, $T_{\text{vir}}^{\text{min}} = 5.0 \times 10^4 K$, and $R_{\text{mid}} = 15$ Mpc.

(ii) BRIGHT SOURCES: $\xi = 200$, $T_{\text{vir}}^{\text{min}} = 3.0 \times 10^5 K$, and $R_{\text{mid}} = 15$ Mpc.

The FAINT SOURCES model, already introduced in Section 4.1, characterizes a reionization history dominated by a large population of galaxies with a relatively low ionizing efficiency. On the other hand, the BRIGHT SOURCES model has larger $\xi$ and $T_{\text{vir}}^{\text{min}}$, and is characterized by fewer sources with larger ionizing efficiency. By construction, these models have fairly similar reionization histories (see fig. 2 in Greig & Mesinger 2017) and match both the inferred evolution of the cosmic star formation rate (SFR) density using the extrapolated observed luminosity functions of Bouwens et al. (2015) and the electron scattering optical depth, $\tau_e = 0.078 \pm 0.012$ from Planck Collaboration et al. (2016). These two scenarios have strong implications for the patchiness of reionization and the morphology of the observed ionized regions, and provide a first glance at what differences would we observe for an EoR driven by AGNs or by fainter galaxies (Robertson et al. 2015; Finkelstein et al. 2019). We note that these models are not unique to explain reionization, as recent studies have shown that ‘intermediate’ scenarios, driven by brighter galaxies, could very well match the current observational constraints (e.g. Ly $\alpha$ damping wings of quasars and galaxies at $z > 7$) (Naidu et al. 2020). Nevertheless, they are interesting for astrophysical parameter forecasting frameworks to explore which information is needed to favour or rule out specific scenarios. Similar to Section 4.1, in this noise-free case, we extract the ionized regions from the synthetic 21-cm maps using $\delta T_\text{th} = 0$ mK as the threshold value.

In Fig. 2, we investigate the distribution of the morphological attributes of the ionized regions for these two models using a snapshot with $x_{\text{ion}} \approx 0.2$. This ionizing fraction corresponds to redshift 9 and 8.5 for the FAINT SOURCES and BRIGHT SOURCES, respectively. Similarly to Fig. 1, the contours plotted in the lower panels include successively 68, 95, and 99.7 per cent of the distributions and we highlight the properties of the largest (percolated) ionized region in the box with diamonds.

We note several modest differences between the observed distributions for both models. The size distribution of the ionized regions suggests that larger ionized structures have formed when reionization is dominated by brighter sources. Additionally, the distributions of $S$ and $N$, show that the FAINT SOURCES model produces more sparse and asymmetric ionized regions. We note that these differences are more significant in the tails of these distributions, suggesting that outliers are important to discriminate between these different reionization scenarios.

The distributions of $E$ and $F$ stay roughly similar for both models. As mentioned above, this might suggest that this property is independent of the reionization scenario as it probes the correlation...
between the ionization field and the underlying density field. This property should prove useful to investigate the global topology of reionization or more complex reionization models, as discussed in Section 4.1.

The morphology of the percolated object in each model provides similar insights. It appears more porous and less compact in the FAINT SOURCES model. This is expected since a population of fainter sources form many individual ionized regions, whose fusions result in more porous and larger structures compared to a reionization scenario driven by fewer brighter sources. Interestingly, both percolated clusters have similar size, suggesting that their morphological properties can provide important additional insights on the nature of the underlying sources.

We note that, in this section and Section 4.1, we used a single set of initial conditions to perform this analysis. In theory, the tails of the different distributions presented in Figs 1 and 2 might change for different set of initial conditions. This cosmic variance effect can have significant consequences when comparing the observed ionized regions number-count and morphology in various astrophysical models. Nevertheless, we found that, when using large box sizes ($512^3$ cells with 2 Mpc resolution), the impact of cosmic variance is negligible compared to the difference between these two models. Additionally, in Section 6.3, we further assess this effect on our inferential framework using smaller mock observations of $128^3$ cells with a 250 Mpc box size. We show that using different sets of initial conditions does not significantly impact the recovered parameter intervals, which suggests that the variations due to cosmic variance is typically smaller than the variation between astrophysical models.

Overall, this section highlights that the most significant differences between these two reionization scenarios are related to the size, sparseness, and compactness of the ionized regions. Additionally, while we did not discuss it in this section, the total number of observed ionized regions should also provide additional information to compare different models. Several studies showed that we should observe fewer ionized regions for scenarios with brighter sources, or with a higher minimum mass of the star-forming haloes (Kakiichi et al. 2017; Giri et al. 2018a). Nevertheless, these differences are more complicated to extract from realistic observations, as instrumental effects tend to smooth the contours of the ionized regions and suppress the information lying in the smaller scales.

4.3 Impact of the point spread function and noise

In this section, we investigate the robustness of the morphological properties of the ionized regions extracted from realistic 21-cm tomographic observations with SKA1-Low. To do so, we simulate
To investigate the effects of noise and resolution on the ionized region morphology, we compare the results obtained in Section 4.1 for the FAINT SOURCES model using a snapshot where $x_{\text{ion}} = 0.2$. We display on Fig. 3 the impact of noise on the simulated 3D image cubes, by showing from left to right: The simulated brightness temperature intensity map with an intrinsic resolution of 2 Mpc (0.730 arcmin) resolution; the 21-cm signal observed with noise corresponding to the maximum baseline of 10 km with SKA1-Low; the 21-cm signal observed after applying a smoothing kernel in space and frequency corresponding to a maximum baseline of 2 km ($\approx$2.9 arcmin or 7.9 Mpc resolution); the extracted ionized regions using the Triangle thresholding approach (see details in Section 4.3).

To improve the noise resolution, the percolated clusters were smoothed using a Gaussian and tophat kernel, which could accentuate their extension and shape. Interestingly, the percolated cluster has a smaller volume, but most of SKA’s collecting area is at baselines smaller than 2 km. This is because the PSF of the instrument, and the instrumental effects complicate the extraction of the ionized regions from 21-cm tomographic observations. To optimally identify the ionized regions, Giri et al. (2018b) recently proposed a new approach using a super-pixel thresholding method (SLIC; Achanta et al. 2011). The authors show that this method performs better than typical segmentation approaches when the ionization fraction of the box is larger than 0.1. Nevertheless, this method is relatively time-consuming, making it less attractive when including it within a Bayesian inference framework such as 21CMMC. We choose a different thresholding approach, referred to as the Triangle thresholding (Zack, Rogers & Latt 1977). This strategy is based on the PDF of the pixel intensities in the image. It finds the optimal threshold value by constructing a line between the histogram peak and the farthest measurement in the tail of the histogram. The threshold is derived by taking the point of maximum distance between this line and the histogram level. This technique is also referred to as the maximum deviation method in Giri et al. (2018b), and is found to be particularly effective for bi-modal distributions. Hence, it is well suited to extract the ionized regions from noisy 21-cm observations because the presence of these structures should imprint a clear peak on the PDF of the 21-cm intensities for observations well within the EoR. Nevertheless, this might be more complex for the very early stages where the ionization fraction is too low to robustly identify these regions (Kakikuchi et al. 2017; Giri et al. 2018b). Overall, we found that this method was giving good performance up to $x_{\text{ion}} > 0.05$ and that it gives similar performance as the SLIC approach (see Appendix A). The choice of our particular approach is motivated by its better computational efficiency. Using a single CPU process and a mock 21-cm image cube of $128^3$ cells, it takes less than 1 s to derive the segmented image using the Triangle thresholding approach, while SLIC takes more than 10 s.

To optimally identify the ionized regions, we developed a new approach using a super-pixel thresholding method (SLIC; Achanta et al. 2011). The authors show that this method performs better than typical segmentation approaches when the ionization fraction of the box is larger than 0.1. Nevertheless, this method is relatively time-consuming, making it less attractive when including it within a Bayesian inference framework such as 21CMMC. We choose a different thresholding approach, referred to as the Triangle thresholding (Zack, Rogers & Latt 1977). This strategy is based on the PDF of the pixel intensities in the image. It finds the optimal threshold value by constructing a line between the histogram peak and the farthest measurement in the tail of the histogram. The threshold is derived by taking the point of maximum distance between this line and the histogram level. This technique is also referred to as the maximum deviation method in Giri et al. (2018b), and is found to be particularly effective for bi-modal distributions. Hence, it is well suited to extract the ionized regions from noisy 21-cm observations because the presence of these structures should imprint a clear peak on the PDF of the 21-cm intensities for observations well within the EoR. Nevertheless, this might be more complex for the very early stages where the ionization fraction is too low to robustly identify these regions (Kakikuchi et al. 2017; Giri et al. 2018b). Overall, we found that this method was giving good performance up to $x_{\text{ion}} > 0.05$ and that it gives similar performance as the SLIC approach (see Appendix A). The choice of our particular approach is motivated by its better computational efficiency. Using a single CPU process and a mock 21-cm image cube of $128^3$ cells, it takes less than 1 s to derive the segmented image using the Triangle thresholding approach, while SLIC takes more than 10 s.
Inference using the reionization morphology

Figure 4. Comparison of the morphological attributes of the ionized regions recovered on the synthetic simulation (green) and on the realistic mock observation assuming 1000 h observations with SKA1-Low (orange), using a snapshot of the FAINT SOURCES model with $x_{\text{ion}} = 0.20$. The contours enclose successively 68, 95, and 99.7 per cent of the ionized regions distributions. The diamonds show the morphological attribute values of the largest ionized region in each cube.

stage, and on the contrast of the 21-cm signal fluctuations that can vary for different reionization scenarios. Consequently, accurately analysing the impact of the instrumental effects of the recovered ionized morphology should be done as a function of multiple reionization stages and different reionization scenarios. However, the performance of the inference framework will indirectly reveal whether the extracted statistics are robust enough against these effects.

5 21CMMC SET-UP

21CMMC$^{12}$ is an MCMC sampler designed to explore the astrophysical parameter space of the Cosmic Dawn (also sometimes called the Epoch of Heating; EoH) and EoR. It includes a modified version of the PYTHON module COSMOHAMMER (Akeret et al. 2013) which is built on the top of the EMCEE python module (Foreman-Mackey et al. 2013), using an affine invariant ensemble sampler (Goodman & Weare 2010). 21CMMC employs a streamlined version of 21CMFAST (see Section 2.1) to efficiently simulate the 21-cm brightness temperature fluctuations for different sets of astrophysical parameters. It has been designed to prepare forthcoming observations and already used in many studies to quantify the constraints and degeneracies among the reionization model astrophysical parameters (e.g. Greig & Mesinger 2015, 2017, 2018; Park et al. 2019). In this work, we use it to explore the use of 21-cm tomographic statistics as an alternative to the power spectrum to recover the EoR astrophysical parameters. In Section 5.1, we summarize the original implementation of 21CMMC based on the power spectrum of the 21-cm observations. We detail, in Section 5.2, the new likelihood function implemented in 21CMMC, which uses the ionized regions statistics extracted from 21-cm images. Finally, in Section 5.3, we present the mock observations and general set-up used in this work.

5.1 21CMMC using the power spectrum

In 21CMMC, the likelihood function for the 21-cm power spectrum ($\Delta$) is defined as a $\chi^2$ statistics. It is expressed as

$$L(\Delta) = \prod_{i=1}^{N_k} \exp \left( \frac{(\Delta_{\text{obs}}(k_i)^2 - \Delta_{\text{mod}}(k_i)^2)}{2\sigma_{\Delta}(k_i)^2} \right),$$

where $\Delta_{\text{obs}}$ and $\Delta_{\text{mod}}$ are the power spectrum of the observation and the model, respectively, $k$ is the Fourier mode, $N_k$ the number of independent Fourier modes included (8 in this work), and $\sigma_{\Delta}$ is the

---

$^{12}$https://github.com/BradGreig/21CMMC
21-cm power spectrum uncertainty defined as
\[ \sigma_A(k) = \sigma_{\text{thermal}}(k)^2 + \sigma_{\text{obs}}(k)^2 + \sigma_{\text{mod}}(k)^2, \]
where \( \sigma_{\text{thermal}} \) is the thermal noise, and \( \sigma_{\text{obs}} \) and \( \sigma_{\text{mod}} \) are the variance of the observation and model, respectively. In this work, \( \sigma_{\text{thermal}} \) is estimated by deriving the uv coverage for a 1000 h observation with SKA1-Low (see parameters in Table 1), simulating the thermal noise using a SEFD of 2500 Jy at the central frequency of the observation, and generating the corresponding thermal noise power spectrum and uncertainty.\(^1\) We note that, in Greig & Mesinger (2015), Greig & Mesinger (2017), the authors included an additional modelling uncertainty, typically fixed to 20 per cent of the mock power spectrum value, to account for the semi-numerical approximations compared to fully numerical codes (Zahn et al. 2011; Hutter 2018), and differences in the radiative transfer equations implementation (Iliev et al. 2006). Nevertheless, in this work, we aim to investigate and compare the performance of a parameter inferential approach using the ionized regions morphological pattern spectra and using the 21-cm power spectrum. Hence, including such error term for the power spectrum would require an equivalent for the 21-cm tomographic statistics, which is not trivial to define. Consequently, we choose to exclude this modelling uncertainty for both cases. Additionally, the eight independent \( k \) bins are sampled in logarithmic space in the interval [0.1, 1] Mpc\(^{-1}\). The lower bound of this interval corresponds to the foreground corruption limit, while the upper bound is fixed by the thermal-noise limit.

5.2 21CMMC using tomographic statistics

Defining a likelihood function to compare the distribution of the morphological attributes of the ionized regions is not trivial. We need to compare two distributions of \( n_{\text{bub}} \), five-dimensional (5D) vectors \( \vec{v} \), such that \( \vec{v} = [V, E, F, S, N] \), and \( n_{\text{bub}} \) is the number of ionized regions extracted in the 21-cm images.\(^2\) Comparing 5D distributions can be computationally costly and penalize the MCMC run if the time taken to compute the likelihood value is too large with respect to the time required to simulate the models. Additionally, the problem is complex because \( n_{\text{bub}} \) significantly fluctuates for different reionization scenarios. Nevertheless, this latter aspect, if handled properly, can also provide additional constraints during the inference, because the number of ionized regions should be closely connected to the parameters of the ionizing sources (Kakiichi et al. 2017; Giri et al. 2018a).

Hence, we choose to follow the approach from Vegetti & Koopmans (2009) to define our likelihood function. The authors used a strategy that divides the likelihood function into two independent factors: the first one accounting for the likelihood of observing a given number of objects using a Poissonian distribution, and the second to compare the properties of these objects.

We define a morphological likelihood function, \( \mathcal{L}(\text{morphology}) \), using two separate factors: a Poissonian factor that compares the number of ionized regions between the model and the observation, and a distance factor that compares the two distributions of 5D vectors, independently of the number of objects in the distributions. In Appendix C, we provide a general expression to define a likelihood function suitable to compare \( n \)-by-\( d \) distributions, where \( d \) is the dimension of the statistics used, and \( n \) the number of objects that might vary for different sets of parameters. For our case \((d = 5)\), the final likelihood expression is

\[
\mathcal{L}(\text{morphology}) = \exp \left( -n_{\text{bub}} \right) \times \left( n_{\text{bub}} \right)^{n_{\text{obs}} - 1} \frac{n_{\text{obs}}}{n_{\text{bub}}} \times \exp \left( -\lambda \times \frac{n_{\text{bub}}}{n_{\text{obs}}} \sum_{i} \arg \min_{j \in n_{\text{mod}}} d_{\text{maha}}(\vec{v}_{\text{obs},i}, \vec{v}_{\text{mod},j})^2 \right),
\]

where \( n_{\text{bub}} \) and \( n_{\text{obs}} \) are the number of ionized bubbles in the observation and in the model, respectively, \( \lambda \) is a regularization parameter, \( d_{\text{maha}} \) is the Mahalanobis distance (Mahalanobis 1936), and \( \vec{v}_{\text{obs}} \) and \( \vec{v}_{\text{mod}} \) are the 5D vectors carrying the five morphological attributes for each individual ionized region observed in the observation and model. A complete description of this likelihood function can be found in Appendix C, and we only briefly describe the general idea here. As mentioned above, the factor \( \exp \left( -n_{\text{bub}} \right) \times \left( n_{\text{bub}} \right)^{n_{\text{obs}} - 1} \frac{n_{\text{obs}}}{n_{\text{bub}}} \) assumes that the fluctuation of the number of ionized regions observed follows a Poissonian distribution. The second factor provides a way to compare the distribution of 5D vectors, by finding, for each ionized region in the observation, the ionized structure in the sampled model with similar morphological attributes. This is done by extracting the pair of vectors \( \vec{v}_{\text{mod}} \) and \( \vec{v}_{\text{obs}} \) such that the Mahalanobis distance between them is minimal. The choice of the Mahalanobis distance is motivated by the fact that classical distance metrics, such as the Euclidean distance, are not suited for high-dimensional applications (Aggarwal, Hinneburg & Keim 2001). Additionally, the Mahalanobis distance accounts for the covariance of the 5D distribution of morphological attributes observed (the distributions are normalized to unit variance for each parameter), such that it provides an unbiased metric that is suitable to compare the structures in different distributions of morphological attributes.

The minimum Mahalanobis distances are computed, elevated to the power five (i.e. number of dimensions), and summed for all the ionized regions in the observation. In other words, this can be understood as minimizing the total volume enclosed between the ionized region morphological distribution in the observation and model in a 5D space. This approach should favour models whose ionized regions reproduce the observations closest. The factor \( \exp \left( -\lambda \times \frac{n_{\text{bub}}}{n_{\text{obs}}} \sum_{i} \arg \min_{j \in n_{\text{mod}}} d_{\text{maha}}(\vec{v}_{\text{obs},i}, \vec{v}_{\text{mod},j})^2 \right) \) ensures that this second term is normalized to the number of ionized regions in the observations (see details in Appendix C). Finally, the parameter \( \lambda \) regulates the weight given either to the Poisson factor or to the normalized distance metric that compares the 5D distributions. Optimally, \( \lambda \) should be included as an additional parameter during the MCMC inference, such that its best value is determined during the sampling process. However, adding a parameter would further increase the complexity and computational time of the 21CMMC run. Hence, in this work, we choose to fix it a priori by performing several test-cases to find an arbitrary optimal value. We find that fixing \( \lambda \) to 0.001 provides the best inference results, and we use this value for all results shown in this paper.

We note that the function defined in equation (12) is not a likelihood function, but rather an approximate penalty function, which is typically suited for regularized maximum likelihood estimation.
problems. This is because the second factor is not a ‘proper’ statistics, contrary to the Poisson factor. Hence, it comes with some caveats, such as the absence of well-defined error terms related to the variance of the 21-cm tomographic statistics with respect to noise or sample variance. Nevertheless, by construction, $\lambda$ can be considered as the inverse variance of the distribution, and therefore, act as a proxy for the error. Additionally, this approach still provide valuable insights to understand whether and how 21-cm tomographic statistics can be included in a Bayesian inference framework, and combined to power spectrum analyses to optimize future observational studies. We will refine this framework in future studies to account for the variance of the 21-cm tomographic statistics. In Section 6, we discuss the inference performance of (1) using only the ionized regions number count (the Poissonian factor), (2) using only the ionized regions morphology (the distance factor), or (3) using both. Increasing or decreasing $\lambda$ simply shifts the result of (3) towards (2) or (1), respectively.

5.3 Mock observations

Extracting and analysing the ionized regions from noisy 21-cm images is typically more computationally expensive than deriving the power spectrum. Hence, we choose to focus on relatively small boxes of $128^3$ cells ($250^3$ Mpc$^3$). We use two sets of mock observations, corresponding to the FAINT SOURCES and the BRIGHT SOURCES models (Mesinger et al. 2016), to compare the constraining power of the ionized regions morphological pattern spectra and power spectrum. Additionally, we combine observations at three different redshifts, 10, 9, and 8. The mock power spectra are extracted using the simulated Fourier visibilities and combined with the total theoretical uncertainty computed using the procedure detailed in Section 5.1. We then create realistic 21-cm mock images including the SKA1-Low PSF, a random noise realization, and the additional 3D smoothing.

The models are sampled using a different set of initial conditions (i.e. different realization of the density field) on the same grid size as the mock observations. Studies using the power spectrum typically use a larger box size to create the mock observations, while the models are produced on a smaller grid (but keeping the same cell resolution) (e.g. Greig & Mesinger 2015, 2017). Nevertheless, 21-cm tomographic statistics are not independent of the box size used (e.g. the number of ionized structures will increase for larger boxes); thus, we keep the same grid for both the models and the mock observations. We note that our experiment is likely more affected by sample variance due to this relatively small box size, and we further discuss this point in Section 6.3.

As detailed in Section 4.3, both the noise and the PSF impact the number and morphological attributes of the ionized regions. Consequently, these instrumental effects must also be included when sampling the parameter space during the inference process, to consistently compare models with observations. During the MCMC sampling process, the ionized regions number-count and morphological spectra of each model sampled are extracted on the fly by (1) using the Triangle Thresholding approach to segment the ionized regions from the mock noisy, low-resolution 21-cm tomographic image cubes, and (2) applying DISCOFAN on the resulting segmented volumes (similar to Section 4.3). We note that the location of the ionized regions is always extracted directly from the mock 21-cm observations, such that, in principle, the same approach can be applied on real observational data sets. To accelerate the convergence of the MCMC analysis, we keep the noise realization fixed for each model sampled (but chosen different than for the mock observations). We performed several test-cases to ensure that the choice of a particular noise realization has little impact on the final inference results. The stopping criterion in 21CMMC is defined relative to a certain number of sample iterations, rather than to a convergence criterion. Therefore, we test for the convergence of the Markov chains using two tests. The first one is the Gelman–Rubin diagnostic (Gelman & Rubin 1992), which computes the sample mean and variance from multiple chains, and check whether they are similar enough to indicate approximate convergence. Additionally, we also use the Geweke diagnostic Geweke (1991), which compares the similarity of the mean and variance of segments from the beginning and end of a single chain. We note that these two tests estimate the convergence of the chains but are no proof of actual convergence. They are usually used to prove a failure to converge, rather than guaranteeing that the chains converged to a global minimum.

Overall, running 21CMMC using the 21-cm tomographic statistics, 3000 iterations and 50 walkers takes around 8 d on a machine with 80 CPUs. Using only the power spectrum with the same set-up takes approximately a factor of 2 less time. Assessing the performance of this approach on larger observations will be crucial to understand the full potential of 21-cm tomographic statistics. To do so, using emulators of the tomographic 21-cm image cubes, such as in Chardin et al. (2019) and List & Lewis (2020), might prove useful to extend this work.

6 RESULTS

This section presents the inference results obtained with 21CMMC. Section 6.1 details the recovered parameter intervals using the 21-cm tomographic statistics for both the FAINT SOURCES and BRIGHT SOURCES mock observations. Section 6.2 compares these results to the constraints inferred using the power spectrum of the 21-cm fluctuations. Finally, Section 6.3 explores the impact of using different initial conditions, and Section 6.4 investigates the consequences of including simulated foreground residuals, as one of the dominant systematic errors, on these results.

6.1 Inference using the ionized regions number-count and morphology

We first assess the performance of the 21-cm tomographic inference approach to recover the parameters of the sources of cosmic reionization. In Figs 5 and 6, we compare the outputs of 21CMMC for the FAINT SOURCES and BRIGHT SOURCES models, respectively, considering three different cases based on the likelihood function defined in equation (12): (1) using the number of ionized regions observed (i.e. the Poissonian factor); (2) using the morphology of the ionized regions (i.e. the distance factor); and (3) combining (1) and (2). The results are shown as corner plots, with the diagonal panels providing the normalized marginalized 1D PDFs of the three parameters, and the lower panels representing the 2D likelihood contours, at the 1 and 2 $\sigma$ level (thick and thin lines, respectively). Additionally, Table 2 shows the recovered median values and the associated 16th and 84th percentile errors, assuming that our framework, and in particular the factor of $1/\lambda$, decently traces the variance for the attribute values (see discussion in Section 5.2).

15Corner plots from Figs 5, 7, 6, 8, and 12 have been made using the python package CORNER (Foreman-Mackey 2016, https://github.com/dfm/corner.py/blob/main/docs/index.rst).
Figure 5. The 1D marginalized PDFs (diagonal panels) and 2D likelihood contours (lower panels) of the three astrophysical parameters $\zeta$, $\log_{10}(T_{\text{min}}^\text{vir})$, and $R_{\text{mfp}}$ for the FAINT SOURCES mock observations assuming 1000 h observations with SKA at redshifts 10, 9, and 8. The dashed lines show the fiducial parameters ($\zeta = 30$, $\log_{10}(T_{\text{min}}^\text{vir}) = 4.7$ and $R_{\text{mfp}} = 15$ Mpc). We compare three MCMC cases, using the number of ionized regions (green); the morphology of the ionized regions (orange); and combining both (purple).

Figure 6. Same as Fig. 5 for the BRIGHT SOURCES mock observations ($\zeta = 200$, $\log_{10}(T_{\text{min}}^\text{vir}) = 5.48$, and $R_{\text{mfp}} = 15$ Mpc).
In the following sub-sections, we detail the results of each inference case.

6.1.1 Comparing only the number of ionized regions

Using the number of ionized regions information already provides remarkably good constraints for the recovered ionizing efficiency and the minimal virial temperature for both reionization scenarios. We note that the median of the posterior distributions are slightly offset compared to the fiducial parameters (around +15 for \(\zeta\) and +0.2 for \(\log_{10}(T_{\text{vir}}^{\text{min}})\) for the FAINT SOURCES, and around −25 for \(\zeta\) for the BRIGHT SOURCES model), but these fairly accurate intervals already support that this information is useful to infer information about the ionizing sources. This is not unexpected since both the values of \(\zeta\) and \(T_{\text{vir}}^{\text{min}}\) drive the number of ionized regions. As discussed in Section 4.2, models with lower \(T_{\text{vir}}^{\text{min}}\) have fewer ionized regions because the ionizing sources can only exist within the densest haloes. Similarly, a larger \(\zeta\) will boost the number of ionizing photons produced by the sources, accelerating the pace at which ionized bubbles grow, and thus increasing the merging rate during the early stages. In both cases, the observed number of individual ionized regions should decrease. Consequently, this information can already rule out a large number of models that have diverging \(n_{\text{phot}}^{\text{mod}}\) and \(n_{\text{phot}}^{\text{obs}}\). The 2D likelihood contours in Fig. 5 show the degeneracy between the ionizing efficiency and the minimum mass of the star-forming haloes. In the FAINT SOURCES model, the recovered number of ionized regions does not significantly vary for models with larger \(\zeta\) and \(\log_{10}(T_{\text{vir}}^{\text{min}})\), while for the BRIGHT SOURCES model, large fluctuations of the ionizing efficiency of the sources still produce a similar number of ionized regions, but small variations of \(T_{\text{vir}}^{\text{min}}\) have a significant impact on \(n_{\text{phot}}^{\text{obs}}\). The degeneracy between \(\zeta\) and \(T_{\text{vir}}^{\text{min}}\) is expected because they both impact the number of ionized regions in each model, which depends on the criterion used in equation (2) (\(\zeta \times f_{\text{coll}}\)) that sets the ionized fraction in each cell. In Section 7.1, we further demonstrate that the 2D likelihood contours of \(\zeta\) and \(T_{\text{vir}}^{\text{min}}\) almost exactly follow the isocountours of the average \(\zeta \times f_{\text{coll}}\) for a given reionization scenario.

We note that the number of ionized regions does not provide enough information to accurately infer the mean free path of the ionizing photons for both the FAINT SOURCES and BRIGHT SOURCES models. While the \(R_{\text{mfp}}\) fiducial value is within the recovered interval, the large fractional errors show that this parameter is not tightly constrained. In theory, \(R_{\text{mfp}}\) regulates the maximum scale to which an ionized region can grow around the ionizing sources, and becomes important only when the ionized regions grow larger than this value. Greig & Mesinger (2017) emphasized that its impact on the power spectrum is less significant when its value is larger than \(\sim 15\) Mpc because the merging of the ionized bubbles becomes the most dominant growth mechanism. Our result suggests that only models with values lower than 8 Mpc seem to be robustly ruled out. Overall, the impact of \(R_{\text{mfp}}\) fluctuations is likely too complex to recover for our experimental set-up, given the limited box size and low-resolution aspect of the data sets (see discussion in Section 6.3).

We note that changing the box size or its resolution will impact the number of ionized regions observed and could affect these results. Typically, when using larger box sizes, we expect to rule out more efficiently astrophysical models that do not closely reproduce the number of ionized regions in the mock observations. This is because increasing \(R_{\text{mfp}}\) typically causes larger variations in the Poisson likelihood factor, which significantly increase the differences in the log-likelihood values. However, to accurately quantify these effects, we need to test our inference framework on larger boxes. This is currently computationally not feasible on machines available to us, but will be investigated in future works.

6.1.2 Comparing only the ionized regions morphology

Comparing the morphological properties of the ionized regions also provides valuable information to constrain \(\zeta\) and \(T_{\text{vir}}^{\text{min}}\) for both reionization models. For the FAINT SOURCES model, the recovered ionizing efficiency of the sources is slightly offset (around +27), while the minimum virial temperature is more robustly inferred. On the other hand, \(\zeta\) is better constrained for the BRIGHT SOURCES case, but \(T_{\text{vir}}^{\text{min}}\) interval is shifted by around −0.14 (logarithmic). The recovered values are still at \(\pm 3\sigma\), however, given the particular definition of the likelihood function implemented, the robustness of these uncertainties should be taken with caution. Overall, these results suggest that the morphological properties of the ionized structures also provide independent information to recover the properties of the underlying reionization sources. The accurate inference of \(T_{\text{vir}}^{\text{min}}\) is slightly surprising, since the virial temperature regulates the number of ionizing sources formed by selecting the mass threshold of the haloes that can form such sources. Therefore, it is unclear

### Table 2.

The median values, and the associated 16th and 84th percentile errors of the recovered astrophysical parameters, \(\zeta\), \(\log_{10}(T_{\text{vir}}^{\text{min}})\), and \(R_{\text{mfp}}\) for the two different reionization scenarios. We assumed 1000 h of observations with SKA1-Low for three different epochs at \(z = 10, 9,\) and 8. We compare the inference results using only ionized regions number count; only the ionized regions morphology; both the ionized regions number-count and morphology; and the power spectrum.

| Case / method                      | \(\zeta\)  | \(\log_{10}(T_{\text{vir}}^{\text{min}})\) | \(R_{\text{mfp}}\) (Mpc) |
|-----------------------------------|------------|------------------------------------------|--------------------------|
| **FAINT SOURCES**                 |            |                                          |                          |
| Ionized regions number count      | 30.00      | 4.70                                     | 15.00                    |
| Ionized regions morphology        | 45.19\textsuperscript{\pm} 19.57 | 4.93\textsuperscript{\pm} 0.15          | 14.66\textsuperscript{\pm} 5.85 |
| Ionized regions number count and morphology | 57.52\textsuperscript{\pm} 12.37 | 4.85\textsuperscript{\pm} 0.08          | 17.32\textsuperscript{\pm} 5.13 |
| Power spectrum                    | 35.02\textsuperscript{\pm} 18.36 | 4.78\textsuperscript{\pm} 0.19          | 12.54\textsuperscript{\pm} 5.26 |
| **BRIGHT SOURCES**               |            |                                          |                          |
| Ionized regions number count      | 200.00     | 5.48                                     | 15.00                    |
| Ionized regions morphology        | 176.85\textsuperscript{\pm} 38.81 | 5.44\textsuperscript{\pm} 0.11          | 15.57\textsuperscript{\pm} 5.17 |
| Ionized regions number count and morphology | 208.22\textsuperscript{\pm} 30.85 | 5.34\textsuperscript{\pm} 0.06          | 11.53\textsuperscript{\pm} 0.89 |
| Power spectrum                    | 211.38\textsuperscript{\pm} 23.34 | 5.51\textsuperscript{\pm} 0.02          | 11.55\textsuperscript{\pm} 0.07 |

Inference using the reionization morphology 1829
how this parameter is connected to the morphological properties of the sources. Nevertheless, the distance factor in equation (12) ensures that the models with the best likelihood values have similar morphological pattern spectra than in the observation. Hence, models with different $T_{\text{vir}}^{\text{min}}$ values might be efficiently ruled out because they do not produce ionized regions with the same morphological spectra. In Section 7.1, we actually show that the ionized regions morphology is indirectly pre-determined by the average ionizing budget per baryon, which depends on both the values $\zeta$ and $T_{\text{vir}}^{\text{min}}$.

We note in Fig. 5 that the morphology of the ionized regions does not help to place a robust constraint on $R_{\text{mfp}}$, but can only be used to rule out models with $R_{\text{mfp}} < 8$ Mpc. In theory, fluctuations in $R_{\text{mfp}}$ impact the volume of the ionized structures, such that using the size distribution of the ionized regions should provide information to infer this IGM property. Nevertheless, as mentioned already in Section 6.1.1, these differences might be too complex to accurately measure given our experimental set-up.

### 6.1.3 Combining the number-count and morphology of the ionized regions

Finally, combining the number count and the morphological attributes of the ionized regions provides the best inference results for the joint set of values of $\zeta$ and $T_{\text{vir}}^{\text{min}}$ for both the FAINT SOURCES and BRIGHT SOURCES models. The posterior distributions peak closer to the fiducial values, and have smaller 16th and 84th percentile errors. On the other hand, it does not improve the inferred interval of $R_{\text{mfp}}$, where the 2D likelihood contours are typically extended over the whole $R_{\text{mfp}}$ prior range. In Section 6.3, we highlight that the recovered interval is strongly affected by the set of initial conditions. Nevertheless, we expect this outcome to be different for experiments where both the observation and the models are sampled on a larger grid because fluctuations in the $R_{\text{mfp}}$ values should be more significantly observable based on the morphological properties or the number of the ionized structures.

Overall, Figs 5 and 6 highlight that the combination of the number and morphological attributes of the ionized structures extracted from 21-cm images provides an alternative approach to the power spectrum to disentangle different EoR scenarios, and constrain properties of the ionizing sources. This first test-case paves the way for more complex, and deeper explorations of these alternative approaches using Bayesian inference frameworks.

### 6.2 Comparison with the power spectrum

Several studies (e.g. Greig & Mesinger 2015, 2017, 2018; Binnie & Pritchard 2019; Park et al. 2019) showed that a Bayesian statistical framework using the power spectrum would already provide tight constraints on the EoR parameters or models. However, the power spectrum is only a limited statistic, because it does not encode the phase information from the Fourier modes, nor the information about non-Gaussianities or non-ergodic effects. In theory, 21-cm tomographic images give a more complete description of the evolution of the ionizing fronts and probe additional astrophysical properties through the non-Gaussianities of the 21-cm signal. In this section, we simply explore how the constraints inferred using the 21-cm tomographic statistics compare to the ones set by the power spectrum of the 21-cm observations.

In Figs 7 and 8, we compare the inferred 1D PDFs and 2D likelihood contours when using the power spectrum or the ionized regions number count and morphological spectra for the FAINT SOURCES and BRIGHT SOURCES models, respectively. Additionally, the corresponding medians and 16th and 84th percentile errors can be found in Table 2. Overall, the power spectrum provides tighter constraints on $\zeta$ and $T_{\text{vir}}^{\text{min}}$ compared to the 21-cm tomographic statistics. These results are not unexpected since, as shown in Appendix B, the theoretical uncertainties on the mock power spectrum used in this work are relatively small at most of the $k$ modes sampled, which strongly rules out all models that do not closely reproduce its shape. Additionally, the power spectrum is extracted directly from the simulated brightness temperature fluctuations in the Fourier space, while the ionized morphological attributes are extracted from noisy, lower resolutions, image cubes. Hence, these tomographic statistics are more sensitive to the quality of the 21-cm images, which depends on the additional processing steps, such as the choice of the smoothing kernel, the uv weighting scheme, or the robustness of the ionized regions extraction method.

We note that using the power spectrum only provides a fairly accurate interval for $R_{\text{mfp}}$ for the BRIGHT SOURCES case, but not for the FAINT SOURCES case where the median recovered value is at more than $3\sigma$ than the fiducial value. This is likely because our experiment is significantly impacted by sample variance due to the small box sizes used. In Fig. 9, we show the spherically averaged power spectrum extracted at redshift $z = 8$ from the mock observation of the FAINT SOURCES set of parameters, the power spectrum from the sampled model with the same fiducial set of parameters (but different initial conditions), and the power spectrum from the model with the fiducial $\zeta$ and $T_{\text{vir}}^{\text{min}}$ but $R_{\text{mfp}} = 22$. It highlights that the power spectrum of the model with the same set of fiducial parameters significantly differs at the low $k$ modes (which probe the large scales) compared to the mock observation, such that models with a larger $R_{\text{mfp}}$ (22 Mpc) better reproduce the power spectrum of the mock observation. Typically, the $R_{\text{mfp}}$ fluctuations dominantly affect the large scales (as shown in fig. 2 of Greig & Mesinger 2015), which can be strongly affected by statistical variance in boxes with limited size (Kaur, Gillet & Mesinger 2020). This caveat is further investigated in Section 6.3, where we compare the outputs of 10 MCMC runs with different sets of initial conditions.

Overall, these results suggest that using the power spectrum provides tighter constraints on the ionizing efficiency and the minimum mass of the star-forming haloes compared to statistics derived on noisy, lower resolution 21-cm images. While this might question the use of these statistics to infer the astrophysics of reionization, we show in Section 6.4 that 21-cm tomographic statistics are more robust to the presence of residual artefacts in the observation. Additionally, we note that we compared, but never combined both statistics within the same inference analysis. This is because these approaches are not entirely independent from each other since they both encode for example the typical scale of the ionized structures. We discuss in Section 7.2 how alternative 21-cm tomographic or higher order statistics could be further combined to a power-spectrum analysis to improve the inference results.

### 6.3 Impact of different initial conditions

In 21CMFAST, the set of initial conditions determines the underlying density field, which is then used to compute the brightness temperature fluctuations at different redshifts. In this work, we fixed the initial conditions of the sampled models a priori, but choosing it different than the ones used to create the mock observations. A more robust inference framework would consist in varying the initial conditions as we sample the models during the inference; nevertheless, this is computationally costly as it requires to
Figure 7. The 1D marginalized PDFs (diagonal panels) and 2D likelihood contours (lower panels) of the three astrophysical parameters $\zeta$, $\log_{10}(T_{\text{vir}}^{\text{min}})$, and $R_{\text{mfp}}$ for the FAINT SOURCES model assuming 1000 h observations with SKA at redshifts, 10, 9, and 8. The dashed lines show the fiducial parameters ($\zeta = 30$, $\log_{10}(T_{\text{vir}}^{\text{min}}) = 4.7$ and $R_{\text{mfp}} = 15$ Mpc). We compare the outcomes of 21CMMC using the power spectrum (orange) or the number and morphology of the ionized regions (purple).

Figure 8. Same as Fig. 7 for the BRIGHT SOURCES model (fiducial parameters $\zeta = 200$, $\log_{10}(T_{\text{vir}}^{\text{min}}) = 5.48$, and $R_{\text{mfp}} = 15$ Mpc).
re-compute the evolved density field for each different set of parameters. As already mentioned in the previous section, the choice of the initial conditions can lead to significant differences when the box size is limited. Recently, Kaur et al. (2020) quantified the error using a mock power spectrum extracted from a box of length 1.1 Gpc and showed that boxes larger or equal to 250 Mpc are needed to achieve convergence within 1σ level of the noise. For our experiment, both the mock observation and the models are extracted on boxes with this limited physical size, such that we might be more sensitive to this statistical variance. Hence, in this section, we aim to quantify the impact of sample variance on the previous results by running ten different 21CMMC runs using ten different sets of initial conditions for the sampled models. Because running these tests is computationally expensive (around seven days per run), we perform this test only for the FAINT SOURCES model, and use a lower number of iterations for the MCMC runs.

The top and bottom panels of Fig. 10 show the results of our experiment using the ionized region number count and morphological spectra, and using the power spectrum, respectively. For clarity, we only show the 1D posterior PDFs, since the 2D likelihood contours do not provide useful information. The black line in each figure represents the average of the ten 1D PDFs. We note that these averaged PDFs are not physically motivated, as they do not represent the combination of ten independent results (for example by observing 10 different fields), but simply summarize the typical impact of varying initial conditions. Additionally, we report in Appendix D the median and 16th and 84th percentiles for all 10 cases.

Overall, the top panels in Fig. 10 support the findings from Section 6.1. The ionized morphological and number count information provide fairly accurate constraints for ζ and $T_{\text{vir}}^\text{min}$, such that the interval of median values recovered from the posterior distributions in each case is between 28.28 and 49.38 for the former, and between 5.0 $\times$ 10^4 and 9.0 $\times$ 10^4 K for the latter. Interestingly, these values are all within 1.5σ of the results presented in Section 6.1, suggesting that systematic errors due to sample variance are consistently contained within the width of the recovered parameter interval from independent runs. We also note the degeneracy between the ionizing efficiency and the minimum virial temperature, already highlighted in Section 6.1, such that a larger recovered ζ is typically paired with a larger $T_{\text{vir}}^\text{min}$. The origin of this degeneracy is further discussed in Section 7.1.

Additionally, the top panels in Fig. 10 emphasize that the width and the peak of the posterior distribution of $R_{\text{mfp}}$ significantly vary depending on the set of initial conditions when using the likelihood function based on the number and morphological spectra of the ionized regions. On average, models with very low or very large $R_{\text{mfp}}$ seem slightly more disfavoured, but the large percentile errors prevent us from setting a robust constraint to this IGM property in all individual cases.

When considering the power spectrum, Fig. 10 emphasizes that varying the initial conditions has little impact on the recovered ionizing efficiency and minimum virial temperature. The interval of the lower and higher recovered median values is within 29 and 31 for ζ and 4.7 $\times$ 10^4 and 5.5 $\times$ 10^4 for $T_{\text{vir}}^\text{min}$, and all results are within 1.5σ from the interval obtained in Section 6.1. This suggests that constraints on both parameters are robust against the effect of statistical variance. Nevertheless, this is not the case for $R_{\text{mfp}}$. The 10 posterior distributions have significant differences depending on the set of initial conditions used, such that the average interval of accepted values is within ~14 and 25 Mpc. This is consistent with Greig & Mesinger (2015) who show that this parameter has a stronger impact on the growth of the ionized regions for small values, while it leads to little difference when larger than ~15 Mpc. The results of the 10 runs suggest that this parameter is adjusted to match the realization of the large scale structure from the mock observation.

As discussed before, this caveat can be solved by using a larger grid for both the mock observations and the models sampled, to reduce the statistical variance impacting the large scales. Additionally, it shows that 21-cm tomographic statistics, albeit extracted from noisy and lower resolution images, are still fairly robust against sample variance effects. However, the constraints inferred using the power spectrum seems more precise than using the number and morphological properties of the ionized regions. Nevertheless, we show in the next section that the power spectrum is likely more sensitive to the presence of foreground residuals compared to statistics extracted from 21-cm images, and thus could be less accurate.

### 6.4 Impact of foreground residuals

In Sections 6.2 and 6.3, we showed that 21-cm tomographic statistics provide an alternative approach to the power spectrum to disentangle different reionization scenarios. However, using the power spectrum sets tighter constraints on the parameters of the reionization sources. As we already discussed, this is likely because the 21-cm tomographic statistics only account for the properties of the ionized regions, and are extracted on lower resolution data sets that require additional processing steps. However, 21-cm signal images provide a direct measure of the ionized structures and their properties, while the power spectrum is an averaged statistics. Consequently, the presence of diffuse structures locally have a limited impact on the shape of the ionized regions, while it will affect the power spectrum on large scales. In this section, we investigate this effect using simulations of 3D Gaussian random fields (GRF) having a power spectrum consistent with a power-law $k^\alpha$ with $\alpha = -2.0$ (consistent with observations of large scale diffuse Galactic emission using the Amsterdam-ASTRON Radio Transient Facility and Analysis Centre (AARTFAAC); Gehlot 2019) and a brightness temperate variance of 28.9 mK. We add these GRF to the three mock observations of the FAINT SOURCES model at redshifts 10, 9, and 8. These GRFs can be seen as diffuse structures that have not been correctly removed during the foreground subtraction. Fig. 11 shows a 2D slice of one
Inference using the reionization morphology

Figure 10. Top: The 1D marginalized PDFs for the three astrophysical parameters $\zeta$, $\log_{10}(T_{\text{vir}}^{\text{min}})$ and $R_{\text{mfp}}$ using the likelihood function based on the number count and morphology of the ionized regions and the FAINT SOURCES mock observations. Bottom: Same but using the likelihood function based on the power spectrum. We compare 10 MCMC runs corresponding to 10 different sets of initial conditions (thin colour lines). The black thick line corresponds to the average of the 10 runs.

Figure 11. From left to right: A 2D slice of the Gaussian Random Field (GRF), mimicking foreground residuals, added to the observation at redshift 10; the corresponding power spectrum of the original mock model and of the observation + GRF at redshift 10, 9, and 8.

of the 3D GRF realization, and the mock power spectrum extracted from the modified observations. The contaminated power spectra are shifted towards larger values, and this effect is stronger for larger $k$ bins. This is expected since adding a GRF will enhance the variance if the field is uncorrelated with the 21-cm field. We note that our simulation of the GRF assumes that it fluctuates similarly in the frequency and spatial domain. The latter is not very realistic since these diffuse residuals are typically synchrotron spectra, which are frequency coherent. Nevertheless, it still provides a decent test-case to investigate the impact of foreground contamination.

Fig. 12 shows the result of the MCMC runs using either the number of ionized structures observed and their morphology or the power spectrum of the 21-cm observations. Additionally, we report in Table 3 the median values and associated 16th and 84th percentile errors. Fig. 12 shows that the recovered $\zeta$ and $T_{\text{vir}}^{\text{min}}$ are shifted towards larger values for both cases. This suggests that the presence of diffuse structures likely suppress the information lying in the small-scale structures, such that we recover fewer but brighter sources. Overall, this bias is smaller when using the statistics extracted on 21-cm images. The recovered intervals for $\zeta$ and $\log_{10}(T_{\text{vir}}^{\text{min}})$ are offset by about +10 and +0.1 compared to the constraints obtained in Section 6.1, and the fiducial values are still within 1.5$\sigma$ of the posterior distributions suggesting that the inferred astrophysics is still consistent with the mock parameters. On the other hand, the intervals recovered using the power spectrum are shifted by more than 3$\sigma$ (40 and 0.47 for $\zeta$ and $\log_{10}(T_{\text{vir}}^{\text{min}})$, respectively), which could lead to an incorrect astrophysical interpretation. This larger bias can be explained by the fact that the power spectrum is a spherically averaged metric, hence globally sensitive to contamination affecting the contrast of the 21-cm intensities. Typically, adding a GRF to the observation will always bias the power spectrum upwards if the field is uncorrelated to the 21-cm field. This is consistent with Nasirudin et al. (2020) who report that including realistic foreground models can bias the parameter constraints up to 5$\sigma$, due to the cross-power
Figure 12. The 1D marginalized PDFs (diagonal panels) and 2D likelihood contours (lower panels) of the three astrophysical parameters $\zeta$, $\log_{10}(T^{\text{min}}_\text{vir}[K])$, and $R_{\text{mfp}}$ (Mpc) for the FAINT SOURCES model assuming 1000 h observations with SKA at redshifts, 10, 9, and 8, and including the contaminating foreground structures shown in Figure 11. The dashed lines show the fiducial parameters ($\zeta = 30$, $\log_{10}(T^{\text{min}}_\text{vir}[K]) = 4.7$ and $R_{\text{mfp}} = 15$). We compare the outcomes of 21CMMC using the power spectrum or the number and morphological spectra of the ionized regions.

Table 3. The median values, and the associated 16th and 84th percentile errors for the three astrophysical parameters, $\zeta$, $\log_{10}(T^{\text{min}}_\text{vir}[K])$, and $R_{\text{mfp}}$ (Mpc) for the modified FAINT SOURCES model ($\zeta = 30$, $\log_{10}(T^{\text{min}}_\text{vir}[K]) = 4.7$ and $R_{\text{mfp}} = 15$ Mpc) perturbed with a 3D Gaussian Random Field (see Fig. 11).

| Method        | $\zeta$ (16th) | $\log_{10}(T^{\text{min}}_\text{vir}[K])$ (16th) | $R_{\text{mfp}}$ (Mpc) (16th) |
|---------------|----------------|---------------------------------|------------------------------|
| Tomography    | 46.88$^{+7.15}_{-9.14}$ | 4.90$^{+0.06}_{-0.12}$ | 13.89$^{+3.98}_{-3.82}$ |
| Power spectrum| 71.16$^{+4.57}_{-4.21}$ | 5.17$^{+0.03}_{-0.03}$ | 23.56$^{+1.91}_{-1.65}$ |

between the foreground, thermal noise, and EoR signal. On the other hand, a GRF acts like a correlated noise in the tomographic images, with both up and down fluctuations with zero mean, and thus do not necessarily bias the shape of the PDF of the 21-cm intensities, such that the number and morphological attributes of the ionized regions can still be robustly extracted using the same approach.

We note that we must remain careful when interpreting the derived uncertainties from the 16th and 84th percentile errors. As mentioned in Section 5.2, the likelihood function for the 21-cm tomographic statistics does not include additional variance terms arising from the expected fluctuations of the morphological attributes due to sample variance or thermal noise. Similarly, including an additional 20 per cent modelling uncertainty for the power spectrum, as it was done in Greig & Mesinger (2015), might result in larger recover parameter intervals, such that the fiducial values are within 3$\sigma$. Nevertheless, this would not affect the observed biases of the recovered intervals, which itself supports that 21-cm tomographic statistics are likely more robust to the presence of foreground residuals.

We briefly note that, similar to the results shown in Sections 6.1–6.3, $R_{\text{mfp}}$ is still poorly constrained in both cases. This is expected because of the extra level of complexity of this test.

Overall, the interest of 21-cm tomographic statistics to further constrain the EoR is thus well illustrated by this section. These results highlight that they likely provide more accurate, although less precise, statistics than the power spectrum regarding the presence of potential foreground contamination. This is key information for future observational studies, and we discuss this further in Section 7.2.

7 DISCUSSION

We further discuss here the results presented in Section 6. Section 7.1 investigates the connection between the morphological pattern spectra of the ionized regions and the astrophysics of the reionization. Section 7.2 discusses the potential of 21-cm tomographic statistics for future observational studies. Finally, Section 7.3 considers the limitations of our experiment.

7.1 The ionized regions morphology probes the average ionizing budget per baryon

In this section, we investigate which properties, among the number of ionized regions or their morphological attributes, contain most of the constraining power to infer the astrophysical parameters. To do so, we extract these statistics for a large number of models sampled within the parameter space used in this work, and using a grid interval of 5, 0.1, and 2, for $\zeta$, $\log_{10}(T^{\text{min}}_\text{vir})$, and $R_{\text{mfp}}$, respectively. In total, this
corresponds to 11781 models for which we created realistic mock observations at $z = 10, 9, \text{and} 8$, and extracted the ionized regions number count, volume, sparseness, non-compactness, elongation, and flatness.

Fig. 13 shows the colour maps corresponding to these six statistics as a function of $\xi$ and $\log_{10}(T_{\text{vir}}^{\text{min}})$. To do so, we marginalize the values over $R_{\text{amp}}$ by taking the mean over all 11 $R_{\text{amp}}$ realizations for each pair of $\xi$ and $\log_{10}(T_{\text{vir}}^{\text{min}})$. Except for the number of ionized regions, we focus on the average of the morphological attributes over all the ionized structures. We note that this is slightly different from the likelihood framework (equation 12) that compares the distributions in a point-by-point manner. Additionally, we show on the figure the isocontours (dashed black lines) of the average number of ionizing photons emitted in the IGM per baryon, $n_{\text{ion}/b}$, is a time-dependent quantity, which relates both to the ionizing efficiency of the sources (considered independent from time or halo mass in this work), and the minimum mass of the star-forming halo through the equation:

$$n_{\text{ion}/b} = \chi \times \frac{d_{\text{coll}}(z, M_{\text{min}})}{dt}.$$  

(13)

Hence, the average of ionizing photons emitted in the IGM per baryon at a given redshift $z$, referred to as $n_{\text{ion}/b}(z)$, can be approximated using the mean collapse fraction in the box (averaged over all voxels). In 21cmFAST, the collapse fraction is estimated for different excursion sets corresponding to different scales $R$. Nevertheless, by construction, the average $d_{\text{coll}}$ should match the Press–Schechter collapse fraction ($f_{\text{col}}^{\text{PS}}$) defined for a redshift $z$ and a minimum halo mass $M_{\text{min}}$. Thus, we derive $n_{\text{ion}/b}(z)$ at a given redshift using

$$n_{\text{ion}/b}(z) = \chi \times \frac{f_{\text{coll}}^{\text{PS}}(z, M_{\text{min}})}{d_{\text{coll}}}.$$  

(14)

The isocontours in Fig. 13 correspond to the evolution of $n_{\text{ion}/b}(z)$ as a function of $\xi$ and $\log_{10}(T_{\text{vir}}^{\text{min}})$. The white regions in the different panels are the sets of parameters where the ionized fraction in the box is too low to robustly estimate the number of ionized regions (typically $n_{\text{ion}} < 0.05$). These models have large $T_{\text{vir}}^{\text{min}}$ such that too few ionizing sources are formed and $n_{\text{ion}/b}(z)$ is less than 0.05. On the other hand, models with $T_{\text{vir}}^{\text{min}} > 1.5$ have few objects, an average volume larger than $10^7$ Mpc$^3$, and $\Xi$ and $\mathcal{F}$ close to 1, suggesting that the universe is dominated by a single percolated ionized region. The models with the most ‘extreme’ morphological attributes (large $\Xi$, $\mathcal{F}$, $\mathcal{N}$, or $\mathcal{S}$) are located around the isocontour $n_{\text{ion}/b}(z) = 1$. The peculiar morphology of the ionized regions in these models is explained by the fact that this isocontour traces a reionization stage just before percolation. At this point, the ionizing field consist of a few large regions (a few tens in number, with an average volume between $10^7$ and $10^8$ Mpc$^3$), strongly porous and eccentric, which are on the brink to fuse into a single, percolated cluster spanning the whole box.

Overall, Fig. 13 emphasizes that the number of ionized regions and their morphology are strongly correlated with the isocontours of $n_{\text{ion}/b}(z)$. Nevertheless, despite a few exceptions, these statistics remain fairly constant along these isocontours, which they are only sufficient to constrain the value of $n_{\text{ion}/b}(z)$, that contains the correct astrophysics, but might not be enough to disentangle the correct pair of $\xi$ and $T_{\text{vir}}^{\text{min}}$ on this contour. On the other hand, Fig. 13 shows that the shape and location of the isocontours of $n_{\text{ion}/b}(z)$ evolve with redshift. Hence, combining the information about the ionized regions morphology at different redshifts can indirectly trace the time evolution of the ionizing budget per baryon, and thus provides additional information to recover $\xi$ and $T_{\text{vir}}^{\text{min}}$. We test for this hypothesis by deriving the relative difference in $n_{\text{ion}/b}$, noted $n_{\text{ion}/b}^{\text{ref}}$, for each pair of $(\xi, T_{\text{vir}}^{\text{min}})$, relatively to the $n_{\text{ion}/b}$ of the fiducial models. Hence, $n_{\text{ion}/b}^{\text{ref}}$ is defined as

$$n_{\text{ion}/b}^{\text{ref}}(\xi, T_{\text{vir}}^{\text{min}}) = \sum (\frac{n_{\text{mock model}}(\xi, T_{\text{vir}}^{\text{min}}) - n_{\text{ion}/b}(\xi, T_{\text{vir}}^{\text{min}})}{n_{\text{mock model}}(\xi, T_{\text{vir}}^{\text{min}})})^2.$$  

(15)

where $n_{\text{mock model}}(\xi, T_{\text{vir}}^{\text{min}})$ and $n_{\text{ion}/b}(\xi, T_{\text{vir}}^{\text{min}})$ are the average number of ionizing photons emitted in the IGM per baryon at a given redshift for the fiducial model (FAINT SOURCES or BRIGHT SOURCES), and for a given pair of $\xi$ and $T_{\text{vir}}^{\text{min}}$, respectively. Hence, $n_{\text{ion}/b}^{\text{ref}}$ traces the difference in the time evolution of the ionizing budget for all the models sampled relative to the fiducial models considered in this work. We present the evolution of $\ln(n_{\text{ion}/b}^{\text{ref}})$ as a function of $\xi$ and $T_{\text{vir}}^{\text{min}}$ for the FAINT SOURCES or BRIGHT SOURCES cases in Fig. 14, and additionally display the 2D likelihood contours obtained for both models in Section 6.1 (shown in Figs 5 and 6) with black solid lines. We note that the 2D likelihood contours of the inference analysis are consistent with the regions having the lowest $\ln(n_{\text{ion}/b}^{\text{ref}})$ which characterize the parameter space where the difference in the time evolution of the ionizing budget is minimum. Additionally, the degeneracy between these parameters observed in Section 6.1 is directly consistent with the shape of these regions. Therefore, this analysis highlights that the evolution of the ionized regions number count and morphological spectra at different redshifts likely provides the strongest constraints on $\xi$ and $T_{\text{vir}}^{\text{min}}$, because it traces $n_{\text{ion}/b}$. Thus, it supports that these tomographic statistics give robust and independent insights about the physical process of reionization.

### 7.2 Combining power spectrum and tomographic analyses

Koopmans et al. (2015) and Mellema et al. (2015) pointed out the potential of the synergy between the power spectrum and alternative approaches based on 21-cm tomographic images. Recent studies have further shown that the use of higher-order, topological and morphological statistics could provide valuable additional information to disentangle the astrophysics of the EoR (Section 1 and Greig 2019). Several of them emphasized that these approaches would be especially useful to break the degeneracies between several astrophysical models having a similar power spectrum (e.g. Kakiichi et al. 2017; Giri et al. 2018a; Kapahi et al. 2019). In this work, we showed how these 21-cm tomographic statistics could be used in a Bayesian inference framework to disentangle two different reionization scenarios, and that these constraints are theoretically more robust to the presence of residual foregrounds than when using the power spectrum (see Section 6.4). Hence, independently combining both approaches can already provide valuable insights into the presence of contamination in the observation if their outcomes are inconsistent.

This first test-case supports the importance of using both approaches to achieve the most robust scientific results and paves the way for further studies investigating the synergy between power spectra and tomographic analyses. In this work, we focused on an intuitive formalism related to the shape or number of ionized regions and emphasized that these statistics trace the time evolution of the ionizing budget per baryon, thus providing already crucial insights on the astrophysics of the EoR. The use of more advanced methods, encoding, for example, a topological analysis of the ionization field (Elbers & van de Weygaert 2019) could help to extend this analysis further. Additionally, alternative statistics that only encode the non-gaussianities of the 21-cm signal, such as the bi-spectrum (e.g. Watkinson et al. 2019; Hutter et al. 2020) or the skewness and
Figure 13. Evolution of the average values of the morphological attributes as a function of the ionizing efficiency and the minimum virial temperature of the star-forming haloes, at redshifts 10, 9, and 8. From top to bottom: The number of ionized structures observed; their average volume (logarithmic); sparseness; non-compactness; elongation; and flatness. Each individual plot combines 11 871 model regularly sampling the parameter space (see the detail in Section 7.1). The white regions are models where the ionized fraction is too low to extract the ionized regions (typically $x_{\text{ion}} < 0.05$). The black dashed lines corresponds to the isocontours of the average number of ionizing photons emitted per baryon, derived as $\xi \times f_{\text{coll}}^{\text{P}}(z, M_{\text{min}})$. 

MNRAS 502, 1816–1842 (2021)
kurtosis of the 21-cm PDF (Banet et al. 2020), could be more easily combined to the power spectrum within the same MCMC analysis, because they provide independent information. This is less trivial for morphological or topological analysis of the ionizing field, where additional steps would be required to ensure that the information they provide is truly disconnected from the power spectrum.

Hence, further theoretical work is required to investigate the optimal strategy to combine both approaches, to prepare for upcoming observational studies with SKA and HERA.

7.3 Limitations of using tomography

To investigate the use of the 21-cm tomographic statistics to recover the astrophysics of the EoR, we made a few assumptions to either simplify the computation of the 21-cm signal or its analysis. One might thus ask: how dependent are our results from the typical assumptions made? We discuss here the potential caveats arising from these assumptions.

In this work, we ignored the effects of redshift space distortion (RSD). RSD artificially compresses and expands the high- and low-density regions, respectively (Kaiser effect; Kaiser 1987). This effect has been studied in Giri et al. (2018a), where the authors showed that it shifts the ionized bubbles size distribution towards smaller values. This is expected from the typical inside–out reionization topology where the ionized regions grow first on the highest density regions, thus are compressed due to the RSD effects. In Giri et al. (2018a), the authors highlight that accounting for these effects has a limited impact on the final results. However, if the data were affected by RSD, they must be included in the simulations to consistently compare the distribution of the number-count and morphological attributes between the observation and the models.

Additionally, our experimental set-up considers co-eval cubes for both the mock observations and the models. However, true observations consist of lightcones, such that the line of sight axis can cover a large range of redshifts. The impact of lightcone effects have been discussed several times in the literature (e.g. Datta et al. 2012, 2014; Greig & Mesinger 2018). Giri et al. (2018a) investigated its impact on the ionized bubble size distribution, and showed that it shifts the ionized regions size towards larger values. However, Datta et al. (2014) showed that the impact of lightcone effects can be mitigated by using a limited bandwidth observation, less than $\Delta f = 10$ MHz ($\Delta z = 0.5$). Thus, similarly to the RSD effects, we do not expect this issue to be a major obstacle for future 21-cm tomographic analyses.

Another concern is related to the absence of physically motivated foregrounds in our mock observations. Perfect foreground removal is typically assumed in all theoretical papers discussing 21-cm tomographic statistics. This is because, in theory, foregrounds properties are significantly different (e.g. smooth, or statistically independent) from the properties of the 21-cm fluctuations. Nevertheless, in practice, perfectly removing all foreground contamination is a very complex task, and artefacts can remain after applying foreground removal methods. In Section 6.4, we briefly discussed the possible impact of foreground residual but limited our analysis to a specific case using Gaussian random fields with a given power-spectrum shape. This test-case already showed that statistics extracted on 21-cm images are likely more robust to the presence of these artefacts than when using the power spectrum. However, additional studies with more complex physical foregrounds models, are needed to further investigate their impact on power spectra or tomographic analyses. We note however that significant improvements have been made regarding foreground removal algorithms. For example, Chapman et al. (2013) used a theoretical experiment based on the LOFAR sensitivity to show that a Generalized Morphological Component Analysis (GMCA) approach (Bobin et al. 2013) would already perform well to recover the 21-cm images from foreground contaminated observations. In fact, Hothi et al. (2021) show that the use of Gaussian Regression analysis tools (GPR; Mertens, Ghosh & Koopmans 2018) performs better than GMCA when including the PSF and the noise in the mock observations. Hence, the use of already available tools, and the development of futures foreground removal methods, optimal for tomographic applications, will be the key to mitigate the caveats arising from these contaminants.

Finally, the results presented in this work are strongly dependent on the validity of the $T_s \gg T_{\text{CMB}}$ (post-heating) hypothesis. Indeed, the ionized regions in the 21-cm tomographic images can be robustly extracted, even in the presence of noise or additional smoothing, because they imprint a clear peak in the low-end distribution of the 21-cm intensities. However, this is more complex when the spin temperate is not well heated above the CMB temperature because the presence of low-intensity structures also arise from regions where $T_s = T_{\text{CMB}}$. Getting rid of this assumption would allow us to probe higher redshifts, constrain epochs where the EoR and the Epoch of Heating overlap (e.g. as in Greig & Mesinger 2017), and more generally, explore more complex reionization models. Nevertheless, it would require to adopt a different approach to probe the evolution of the morphological information lying in the 21-cm fluctuations, for example, by considering the 21-cm field morphological spectra.
as a function of several threshold sets within the 21-cm intensity range (similar to Kapahi et al. 2019). Such strategy would be simple to implement using the connected component tree framework of DISCCOFAN, but would require to re-define the likelihood function in 21CMMC, which is not a trivial task for this approach. Consequently, dismissing this assumption is likely the more challenging obstacle to overcome to extend this analysis to higher redshifts.

Overall, this work paves the way for further theoretical studies investigating the robustness of 21-cm tomographic statistics, using more complex reionization parametrization and observations including both physically motivated foregrounds and additional instrumental artefacts. The ability of 21-cm images to probe the astrophysics, as shown in Section 6.1, and their apparent better robustness to residual contamination (Section 6.4), support that they are promising for future observational studies.

8 CONCLUSIONS

We have examined the use of statistics extracted from 21-cm tomographic images to constrain the properties of the reionization sources using a Bayesian statistical inference framework. We defined an intuitive formalism based on the number count and morphological pattern spectra of the ionized regions, encoding their eccentricity, porosity, and asymmetry. We implemented a likelihood function within 21CMMC (Greig & Mesinger 2015), a Markov Chain Monte Carlo analysis tool, to quantify the astrophysical constraints set by these 21-cm tomographic statistics. We used a reionization model with three parameters: the ionizing efficiency of the sources ($\zeta$, assumed constant), the mass of the star-forming haloes derived from the minimum virial temperature ($T_{\text{vir}}^{\text{min}}$), and the mean free path of the ionizing photons in the IGM ($R_{\text{mfp}}$). We created mock observations of 128$^3$ cells with 2 Mpc resolution (0.735 arcmin) based on two reionization scenarios dominated by FAINT SOURCES ($\zeta = 30$, $T_{\text{vir}}^{\text{min}} = 5.0 \times 10^4$ K, and $R_{\text{mfp}} = 15$ Mpc), and BRIGHT SOURCES ($\zeta = 200$, $T_{\text{vir}}^{\text{min}} = 3.0 \times 10^5$ K, and $R_{\text{mfp}} = 15$ Mpc), similarly as in Greig & Mesinger (2017). We included the noise profile and point spread function of the SKA1-Low, assuming 1000 h of observation, a maximum baseline of 2 km, and three snapshots at redshift 10, 9, and 8. Our results can be summarized as follow:

(i) The morphological properties of the ionized regions provide valuable information to trace the reionization history (Section 4.1) or differences between different reionization scenarios (Section 4.2). Additionally, these morphological differences are more significant in the largest structures, such as the ionized percolated region that typically dominates the late reionization stages. This supports that crucial astrophysical insights can already be extracted from the morphology of a single, large ionized structure.

(ii) The ionized regions number-count and morphological spectra, incorporated in a Bayesian inference framework, can disentangle reionization scenarios with different ionizing efficiency of the sources of reionization and minimum mass of the star-forming haloes (Section 6.1). This is because these statistics indirectly probe the time evolution of the average number of ionizing photons emitted in the IGM per baryon, which is fixed for a given pair of $\zeta$ and $T_{\text{vir}}^{\text{min}}$ (Section 7.1). This supports the use of 21-cm tomographic statistics as an alternative approach to the power spectrum to recover the astrophysics of the ionizing sources.

(iii) Using the power spectrum of the 21-cm fluctuations provides the tightest constraints for $\zeta$ and $T_{\text{vir}}^{\text{min}}$, and is slightly less sensitive to different sets of initial conditions than the number and morphology of the ionized regions (Sections 6.2 and 6.3). Nevertheless, statistics extracted from realistic mock 21-cm tomographic observations are more robust to the presence of diffuse foreground residuals in the observations than using the power spectrum (Section 6.4). This is shown by the presence of a larger bias in the recovered parameter intervals when using the power spectrum (Fig. 12). Comparing the outcomes of power spectra and tomographic analyses is useful to check the consistency of the recovered astrophysical parameters, and can provide valuable information about the presence of residual structures in the observation (Section 7.2).

(iv) The mean free path of photons, $R_{\text{mfp}}$, cannot be accurately inferred using our experimental set-up, both when using the power spectrum or the tomographic statistics (Section 6.2). This is because the mock observations and the models sampled have a limited box size (250 Mpc), such that the large scales, where fluctuations in $R_{\text{mfp}}$ have a significant impact, are dominated by statistical variance (Fig. 9). This is supported by the results of Section 6.3, where we showed that varying the initial conditions has a significant impact on the recovered $R_{\text{mfp}}$ value, both for the power spectrum and the tomographic statistics.

This work emphasizes that 21-cm tomographic statistics can be successfully incorporated within a Bayesian statistical inference tool, and provide a robust alternative approach to the power spectrum. Additionally, it highlights the importance of combining both approaches in upcoming observational studies with the SKA to achieve the best scientific results. In further exploratory studies, we plan to assess the robustness of these results using more complex reionization parametrizations, by removing the post-heating assumption ($T_S \gg T_{\text{CMB}}$), and generating mock observations including physically motivated foreground contamination. Overall, this paper supports that the use of 21-cm images and, in general, the morphological properties of the ionized structures, provide key information to the physical process of reionization.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous referee for constructive comments that help improve the quality of this manuscript, and the Center for Information Technology of the University of Groningen for their support and for providing access to the Peregrine high performance computing cluster. SG thanks Sambit Giri for several discussions throughout this work, Florent Mertens and Garreld Mellema for helpful suggestions, Andrei Mesinger and Bradley Greig for discussions about 21CMFAST and 21CMMC, and Marco Grzegorczyk for an instructive conversation on the implementation of likelihood functions adapted for high-dimensional statistics. This paper is based on research developed in the DSSC Doctoral Training Programme, co-funded through a Marie Sklodowska-Curie COFUND (DSSC 754315).

Additionally to the various tools referenced in the manuscript, the authors acknowledge the use of the python packages NUMPY (Harris et al. 2020), MATPLOTLIB (Hunter 2007), SCIPY (Virtanen et al. 2020), PANDA (Wes McKinney 2010), and SCIKIT-IMAGE (van der Walt et al. 2014).

DATA AVAILABILITY

The simulation data underlying this article will be shared on request to the corresponding author.
In Section 4.3, we briefly mentioned the technique used in this work to extract ionized regions from noisy observations. We provide here more details about the different thresholding methods. Recently, Giri et al. (2018b) proposed a new approach to optimally identify the ionized regions from realistic observations, using a superpixel thresholding approach, SLIC, introduced in Achanta et al. (2012). While this method shows good performance for their application, the SLIC approach is computationally expensive, which makes it less suitable for Bayesian inference frameworks.

On the contrary, we found the Triangle thresholding approach (Zack et al. 1977) to perform well on the typical noisy image cubes we deal with in this work. This thresholding method works on the histogram of the values in the image. It constructs a line between the peak and farthest measurement in the tail of the histogram, and the threshold value is derived as the value where the distance between this line and the histogram is maximal. This approach is identical to the maximum deviation method as used by Giri et al. (2018b).

In Fig. A1, we compare the performance of four approaches, the SLIC method, the k-means clustering used for example in Kakiichi et al. (2017), and the Triangle thresholding approach. We note that we applied the latter method in two different ways, labelled as Triangle3D and Triangle2D. Triangle3D refers to the case where we apply the triangular thresholding directly on the 3D image. Triangle2D refers to a different approach where we first apply the triangular thresholding to all individual 2D slices and collect the corresponding list of threshold values. We then take the median value of this list and use the latter to segment the entire 3D 21-cm map in a consistent manner. This approach is different from stacking the set of 2D individually thresholded slices and does not bias the recovered spatial extent of the ionized regions in the resulting 3D segmented volume. To compare the performance of these methods, we used several co-evol cubes of $128^3$ cells with $\approx 2$ Mpc resolution, with different neutral fractions in the box. We added a noise cube corresponding to $1000\, h$ observations with SKA1-Low at $z = 10$, and smoothed the final cube assuming a resolution of $2\, Mpc$.

Figure A1. Left: the recovered ionizing fraction in the realistic observations versus its fiducial value. Middle: the correlation coefficient between the recovered binary fields using the different methods, and the ionization field extracted from the fiducial simulation. Right: the computational time of each method.

APPENDIX A: COMPARISON OF METHODS TO EXTRACT THE IONIZED REGIONS FROM MOCK TOMOGRAPHIC OBSERVATIONS
corresponding to maximum baseline of 2 km (the rms of the noise after smoothing is 2.5 mK). The left panel compares the recovered volumetric ionizing fraction $x_{\text{ion}}^{\text{fid}}$ with respect to the fiducial ionizing fraction $x_{\text{ion}}^{\text{fid}}$ in the box. The middle panel shows the evolution of the correlation coefficient ($r_{xy}$, Cramer 1946) between the recovered binary fields using the different methods, and the binary field from the fiducial simulations extracted from the ionization fraction box taking $x_{\text{ion}}^{\text{fid}} > 0.5$ as a threshold to determine if a voxel is ionized or not. The right-hand panel shows the execution time to extract the ionized regions as a function of the fiducial $x_{\text{ion}}^{\text{fid}}$.

Overall, the correlation coefficient shows that the Triangle2D approach performs better than the other methods, while its execution time is lower than 0.1 s. Interestingly, the triangle thresholding approach that finds the best on the 2D slices and taking the median works better than applying it directly on the 3D data. SLIC performs slightly better than Triangle2D, but is two order of magnitude slower. Additionally, the triangle method has a similar maximum deviation method as reported in Giri et al. (2018b), but is two order of magnitude slower. The right-hand panel shows that the triangle thresholding approach that finds the best on the 2D slices and taking the median works better than applying it directly on the 3D data. SLIC performs slightly better than Triangle2D, but is two order of magnitude slower.

**APPENDIX B: POWER SPECTRUM OF THE MOCK OBSERVATIONS**

Fig. B1 shows the 21-cm power spectrum of the FAINT SOURCES ($\zeta = 30$, log$_{10}(v_{\text{vir}}[K]) = 4.7$ and $R_{\text{vir}} = 15$ Mpc) and BRIGHT SOURCES ($\zeta = 200$, log$_{10}(v_{\text{vir}}[K]) = 5.48$, and $R_{\text{vir}} = 15$ Mpc) mock observations (Mesinger et al. 2016) at the three co-eval redshifts used in this work. The total uncertainty on each power spectrum is shown in the data and model. To picture this normalization problem, we consider the factor $d_{\text{maha}}$ as a volume in a $N$-dimensional space. The total volume occupied by the observations in B should be divided by $n_B$, such that the volume per observation in the distribution is proportional to $1/n_B$. Thus, if $n_B$ gets larger than $n_A$, the sum of $d_{\text{maha}}$, which encodes the minimum volume between the observations in A and B, will reach lower values because of $1/n_B$ dependence. To compensate for this effect, we multiply the sum of $d_{\text{maha}}$ by $n_B$ such that the final factor is independent from $n_B$. Similarly, we add a factor $1/n_A$ to ensure that the sum of the distances is also normalized with respect to the number of objects in the distribution A. Hence, this second factor, referred to as the distance factor, is expressed as

$$\exp \left( -\frac{n_B}{n_A} \sum_i^{n_A} \arg \min_j^{n_B} d_{\text{maha}}(\vec{v}_{A,i}, \vec{v}_{B,j}) \right)^{N_B}. \quad (C2)$$

We additionally include a regularization factor $\lambda$ to adjust the relative importance of each term, such that, larger and smaller $\lambda$ values,
respectively, increase and decrease the weight on the factor that compares the \( N \)-dimensional vectors. The final expression of our maximum likelihood estimator function is defined as
\[
\mathcal{L}(\text{morphology}) = \frac{\exp(-n_B) \times n_B^n}{n_A!} \times \exp \left( -\lambda \times \frac{n_B}{n_A} \sum_{j=1}^{n_A} \arg \min d_{\text{mahal}}(\vec{v}_{A_j}, \vec{v}_{B_j}) \right)^N.
\]

We note that this equation is not strictly speaking an exact likelihood function. It suffers a few caveats, as it is not normalized, and does not encode potential systematic relations to the variance of the measurement of the \( N \)-dimensional vectors in \( A \) and \( B \). The presence of the regularization factor \( \lambda \) makes it closer to a regularized function. However, it provides a suitable framework for comparing the distributions of \( n \) \( N \)-dimensional vectors, where \( n \) can significantly fluctuate, and additionally includes the information lying in the distribution outliers, which are typically not accounted for in classical distribution metrics (e.g. mean or standard deviation). This strategy can be less effective when \( N \) is large and the vectors of observations extracted from \( A \) or \( B \) are subject to noise, because distributions with outliers will be penalized by the factor \( d_{\text{mahal}}^N \). However, we also tested different cases using \( d_{\text{mahal}}^K \), with \( K < N \), to reduce this caveat. We found that the best results were achieved when \( K = N \).

### APPENDIX D: SAMPLE VARIANCE TESTS

In this appendix, we detail the individual results of the ten 21CMMC tests with differing initial conditions done in Section 6.3. Table D1 shows the individual median, 16th and 84th percentile errors for the three astrophysical parameters, \( \zeta \), \( \log_{10}(T_{\text{min}}^{\text{max}}) \), and \( R_{\text{filt}} \) considering the FAINT SOURCES model, and two cases using either the power spectrum or the number and morphological properties of the ionized regions.

| Initial conditions set | Case | \( \zeta \) | \( \log_{10}(T_{\text{min}}^{\text{max}}) (K) \) | \( R_{\text{filt}} \) (Mpc) |
|------------------------|------|-------------|-------------------------------|---------------------|
| 1                      | Ionized regions count + morphology | 35.02 \( \pm \) 18.36 | 4.78 \( \pm \) 0.19 | 12.54 \( \pm \) 5.26 |
                      | Power spectrum | 31.29 \( \pm \) 1.24 | 4.74 \( \pm \) 0.03 | 22.11 \( \pm \) 2.11 |
| 2                      | Ionized regions count + morphology | 31.80 \( \pm \) 18.43 | 4.71 \( \pm \) 0.39 | 13.55 \( \pm \) 5.31 |
                      | Power spectrum | 31.11 \( \pm \) 1.29 | 4.72 \( \pm \) 0.03 | 18.24 \( \pm \) 3.42 |
| 3                      | Ionized regions count + morphology | 47.38 \( \pm \) 17.17 | 4.96 \( \pm \) 0.12 | 15.88 \( \pm \) 4.22 |
                      | Power spectrum | 30.09 \( \pm \) 1.48 | 4.71 \( \pm \) 0.03 | 14.83 \( \pm \) 2.56 |
| 4                      | Ionized regions count + morphology | 39.32 \( \pm \) 29.14 | 4.90 \( \pm \) 0.20 | 17.14 \( \pm \) 2.98 |
                      | Power spectrum | 30.90 \( \pm \) 1.28 | 4.70 \( \pm \) 0.03 | 21.65 \( \pm \) 1.99 |
| 5                      | Ionized regions count + morphology | 29.90 \( \pm \) 57.80 | 4.70 \( \pm \) 0.49 | 14.07 \( \pm \) 3.14 |
                      | Power spectrum | 29.65 \( \pm \) 1.13 | 4.66 \( \pm \) 0.03 | 20.12 \( \pm \) 2.55 |
| 6                      | Ionized regions count + morphology | 28.28 \( \pm \) 2.91 | 4.72 \( \pm \) 0.06 | 15.66 \( \pm \) 2.50 |
                      | Power spectrum | 30.05 \( \pm \) 1.23 | 4.70 \( \pm \) 0.03 | 20.34 \( \pm \) 2.24 |
| 7                      | Ionized regions count + morphology | 37.26 \( \pm \) 8.88 | 4.82 \( \pm \) 0.09 | 16.25 \( \pm \) 3.29 |
                      | Power spectrum | 30.64 \( \pm \) 1.26 | 4.73 \( \pm \) 0.03 | 18.61 \( \pm \) 2.23 |
| 8                      | Ionized regions count + morphology | 31.31 \( \pm \) 21.36 | 4.73 \( \pm \) 0.25 | 15.50 \( \pm \) 3.51 |
                      | Power spectrum | 30.48 \( \pm \) 1.32 | 4.70 \( \pm \) 0.03 | 18.01 \( \pm \) 4.41 |
| 9                      | Ionized regions count + morphology | 43.06 \( \pm \) 11.80 | 4.88 \( \pm \) 0.10 | 15.87 \( \pm \) 3.53 |
                      | Power spectrum | 30.37 \( \pm \) 1.23 | 4.71 \( \pm \) 0.03 | 18.57 \( \pm \) 3.42 |
| 10                     | Ionized regions count + morphology | 35.91 \( \pm \) 48.42 | 4.83 \( \pm \) 0.38 | 16.51 \( \pm \) 4.14 |
                      | Power spectrum | 29.87 \( \pm \) 1.16 | 4.69 \( \pm \) 0.03 | 18.58 \( \pm \) 3.18 |

Average Ionized regions count + morphology | 36.92 \( \pm \) 26.73 | 4.82 \( \pm \) 0.24 | 15.53 \( \pm \) 4.16 |

| Power spectrum | 30.24 \( \pm \) 3.34 | 4.70 \( \pm \) 0.05 | 19.29 \( \pm \) 3.35 |

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.