Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **438,** 2474–2482 (2014)
Advance Access publication 2014 January 16

doi:10.1093/mnras/stt2366

The cross-power spectrum between 21 cm emission and galaxies in hierarchical galaxy formation models

Jaehong Park,[1⋆] Han-Seek Kim,[1⋆] J. Stuart B. Wyithe[1,2] and C. G. Lacey[3]

[1]*School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia*
[2]*ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), Redfern, NSW 2016, Australia*
[3]*Department of Physics, Institute for Computational Cosmology, University of Durham, South Road, Durham DH1 3LE, UK*

Accepted 2013 December 5. Received 2013 November 26; in original form 2013 September 6

ABSTRACT

The correlation between 21 cm fluctuations and galaxies is sensitive to the astrophysical properties of the galaxies that drove reionization. Thus, detailed measurements of the cross-power spectrum and its evolution could provide a powerful measurement of both the properties of early galaxies and the process of reionization. In this paper, we study the evolution of the cross-power spectrum between 21 cm emission and galaxies using a model which combines the hierarchical galaxy formation model GALFORM implemented within the Millennium-II dark matter simulation, with a semi-numerical scheme to describe the resulting ionization structure. We find that inclusion of different feedback processes changes the cross-power spectrum shape and amplitude. In particular, the feature in the cross-power spectrum corresponding to the size of ionized regions is significantly affected by supernovae feedback. We calculate predicted observational uncertainties of the cross-correlation coefficient based on specifications of the Murchison Widefield Array (MWA) combined with galaxy surveys of varying area and depth. We find that the cross-power spectrum could be detected over several square degrees of galaxy survey with galaxy redshift errors $\sigma_z \lesssim 0.1$.

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

1 INTRODUCTION

The prospect of measuring the 21-cm power spectrum from the epoch of reionization is a focus of modern theoretical cosmology (e.g. Morales & Wyithe 2010). A very successful technique has been to employ an *N*-body code to generate a distribution of haloes, and then apply radiative transfer methods in post-processing to model the generation of ionized structure on large scales using various models for the ionizing sources (e.g. Ciardi, Stoehr & White 2003; Sokasian et al. 2003; Iliev et al. 2007, 2008; Zahn et al. 2007; Trac & Cen 2007; Shin, Trac & Cen 2008; Trac, Cen & Loeb 2008). However, when constructing models to assign ionizing luminosities to dark matter haloes, most studies have used a constant mass-to-luminosity relation. On the other hand, the degree to which the important astrophysics governing formation and evolution of high-redshift galaxies will influence observations of the 21-cm power spectrum is not well known. To improve on the source modelling for calculation of the ionizing photon budget in reionization simulations, several studies (Benson et al. 2006; Raičević, Theuns & Lacey 2011; Lacey et al. 2011) have used GALFORM (Cole et al. 2000; Baugh et al. 2005; Bower et al. 2006) combined with Monte Carlo merger trees. However, these studies calculated only the global evolution of reionization, and are not able to address the reionization structure. Most recently, Kim et al. (2013) have combined GALFORM implemented within the Millennium-II dark matter simulation (Boylan-Kolchin et al. 2009), with a semi-numerical scheme to describe the resulting ionization structure. Kim et al. (2013) demonstrated the sensitivity of the ionization structure to the astrophysics of galaxy formation, and found that the strength of supernovae (SNe) feedback is the most important quantity.

In addition to the 21-cm power spectrum, several studies have previously analysed the cross-power spectrum (correlation) between redshifted 21-cm observations and galaxy surveys (Furlanetto & Lidz 2007; Lidz et al. 2009, 2011; Wiersma et al. 2013). These models showed that the cross-power spectrum should be observable, but do not provide a self-consistent link between the astrophysics of galaxy properties and the reionization structure. For example, Furlanetto & Lidz (2007) and Lidz et al. (2009, 2011) used a simple one-to-one relation between luminosity and dark matter halo mass. Conversely, in Wiersma et al. (2013), the cross-power spectrum was predicted using a semi-numerical code for 21-cm emission based on dark matter overdensity cross-correlated with a semi-analytic model

⋆ E-mail: jaehongp@student.unimelb.edu.au (JP); hansikk@unimelb.edu.au (H-SK)

for galaxies. As a result, the calculation did not include the direct relation between galaxies and ionization structure. In this paper our aim is to determine whether the cross-power spectrum can be used to infer the properties of high-redshift galaxy formation. We present predictions for the cross-power spectrum between 21-cm emission and galaxies using the model of Kim et al. (2013) who directly combined detailed models of high-redshift galaxy formation using GALFORM with a semi-numerical description, and predict the resulting redshifted 21-cm power spectrum of different reionization histories. This model provides self-consistent results because the ionizing sources and observed galaxies are the same. These galaxies include both the observed luminous galaxies and the low-mass ($\sim 10^8\,\mathrm{M}_\odot$) galaxies thought to drive reionization.

We begin in Sections 2 and 3 by describing the implementation of GALFORM, our method for modelling the ionization structure, the cross-power spectrum and cross-correlation function, and the cross-correlation coefficient. The cross-power spectra from our method, and the effect of feedback processes on the cross-power spectra, are presented in Section 4. In Section 5 we describe the observational uncertainty. We finish with some conclusions in Section 6.

2 THE GALFORM GALAXY FORMATION MODEL

In this section we summarize the theoretical galaxy formation modelling based on Kim et al. (2013) that is used in our analysis in order to describe the new features for this paper.

We implement the GALFORM (Cole et al. 2000) model, within the Millennium-II dark matter simulation (Boylan-Kolchin et al. 2009). In this study, we specifically use the Lagos implementation of GALFORM (Lagos et al. 2012) model described in Kim et al. (2013). The simulation has a cosmology including fractional mass and dark energy densities with values of $\Omega_\mathrm{m} = 0.25$, $\Omega_\mathrm{b} = 0.045$ and $\Omega_\Lambda = 0.75$, a dimensionless Hubble constant of $h = 0.73$, and a power spectrum normalization of $\sigma_8 = 0.9$. The particle mass of the simulation is $6.89 \times 10^6\,h^{-1}\,\mathrm{M}_\odot$ and we detect haloes down to 20 particles (the minimum halo mass corresponds to $\sim 1.4 \times 10^8\,h^{-1}\,\mathrm{M}_\odot$) in the simulation box of side length $L = 100\,h^{-1}$ Mpc.

Fig. 1 shows the relation between the UV magnitude (the rest-frame 1500 Å AB magnitude) including the effects of dust extinction of galaxies and the host halo mass (top), and between the total Lyman continuum luminosity ($\dot{N}_\mathrm{Lyc}$) of each galaxies and the host halo mass (bottom) from the GALFORM model. Of particular note is that the luminosity of an ionizing source is not simply proportional to the host halo mass as is often assumed in reionization models (Lidz et al. 2009, 2011; Iliev et al. 2011). In part this is because of the distribution of satellite galaxies. The broad scatter of the relation indicates that physically motivated modelling for ionizing sources during the reionization should be included to understand the epoch of reionization. We note that this magnitude is not the same as ionizing luminosity. However, as shown in Fig. 1, the UV magnitude (the rest-frame 1500 Å AB magnitude) is closely related to the ionizing luminosity. Furthermore, Fig. 1 shows that the predicted ionizing luminosity to mass ratio from the model is not a simple one-to-one relation between luminosity and dark matter halo mass.

3 THE IONIZATION MODEL

In this section we summarize the calculation of the ionized structure (Section 3.1) and describe calculation of cross-power spectrum and cross-correlation function (Sections 3.2 and 3.3).

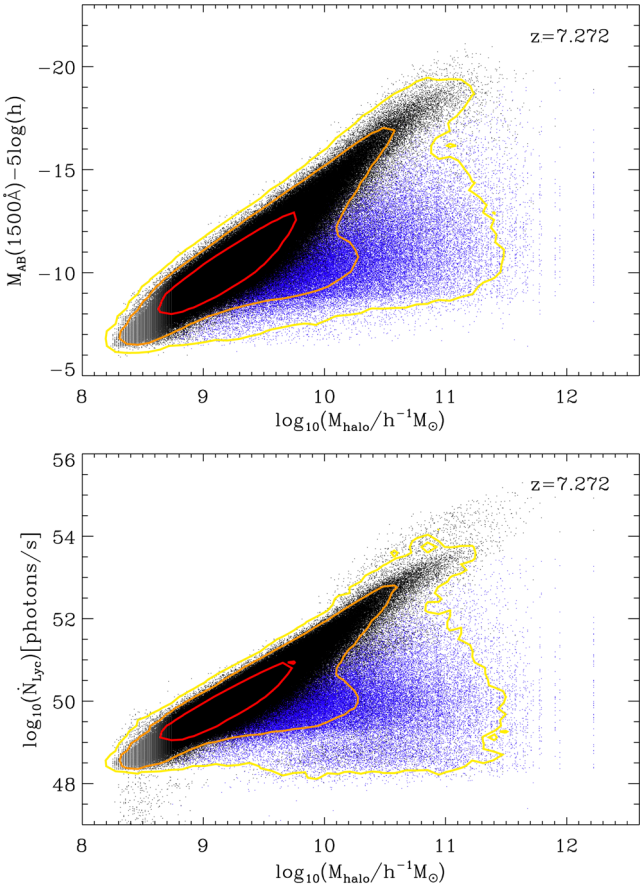

Figure 1. The relation between the UV magnitude (the rest-frame 1500 Å AB magnitude) and the host halo mass (top), and between the total Lyman continuum luminosity ($\dot{N}_\mathrm{Lyc}$) of each galaxies (bottom) for galaxies and the host halo mass at $z = 7.272$ from GALFORM. In each panel, black and blue dots represent central and satellite galaxies. Red, orange and yellow colours represent 1 (68.3 per cent), 2 (95.4 per cent) and 3-sigma (99.7 per cent) levels.

3.1 Semi-numerical scheme to calculate the evolution of ionizationed structure

Mesinger & Furlanetto (2007) introduced an approximate but efficient method for simulating the reionization process, referred to as a semi-numerical technique. In this paper we apply a semi-numerical technique to find the ionization structure resulting from GALFORM galaxies within the Millennium-II dark matter simulation.

The simulation box is divided into cells. We calculate the number of photons produced by galaxies in each cell that enter the IGM and participate in reionization to be

$$N_{\gamma,\mathrm{cell}} = f_\mathrm{esc} \int_0^{t_z} \dot{N}_{\mathrm{Lyc,cell}}(t)\,\mathrm{d}t, \quad (1)$$

where f_esc is the escape fraction of ionizing photons produced by galaxies. Here $\dot{N}_{\mathrm{Lyc,cell}}(t)$ is the total Lyman continuum luminosity of the N_cell galaxies within the cell expressed as the emission rate of ionizing photons (i.e. units of photons s^{-1}).

The ionization fraction within each cell is calculated as

$$Q_\mathrm{cell} = \left[\frac{N_{\gamma,\mathrm{cell}}}{(1+F_\mathrm{c})N_{\mathrm{H\,I,cell}}}\right], \quad (2)$$

where F_c denotes the mean number of recombinations per hydrogen atom and $N_{\rm H\,I,cell}$ is the total number of hydrogen atoms within a cell. The total number of hydrogen atoms is given by

$$N_{\rm H\,I,cell} = n_{\rm H\,I}(\delta_{\rm DM,cell} + 1)V_{\rm cell}, \quad (3)$$

where $n_{\rm H\,I}$ is the mean comoving number density of hydrogen atoms, $\delta_{\rm DM,\,cell}$ is the dark matter over density, which is based on the Millennium-II Simulation density field, and $V_{\rm cell}$ is the comoving volume of the cell. We assume that the overdensity of neutral hydrogen follows the dark matter and self-reionization of a cell occurs when $Q_{\rm cell} \geq 1$. It is complicated to theoretically predict the values of F_c and $f_{\rm esc}$, and the values are not known. In this paper, we use the values of $(1 + F_c)/f_{\rm esc}$ in table 2 of Kim et al. (2013). These parameters provide a reionization history with a mass-averaged ionization fraction of $\langle x_i \rangle = 0.056, 0.16, 0.36, 0.55, 0.75$ and 0.95, corresponding to a redshift of $z = 9.278, 8.550, 7.883, 7.272, 6.712$ and 6.197, respectively. The mass-averaged ionization fraction is calculated as

$$\langle x_i \rangle = \frac{\sum_{j=1}^{\rm N} Q_j N_{{\rm H\,I},j}}{\sum_{j=1}^{\rm N} N_{{\rm H\,I},j}}, \quad (4)$$

where N denotes the number of cells. We divide the Millennium-II simulation box into 256^3 cells, yielding cell side lengths of $0.3906\,h^{-1}$ Mpc and comoving volumes of $0.0596\,h^{-3}$ Mpc^3.

Based on equation (2), individual cells can have $Q_{\rm cell} \geq 1$. On the other hand, cells with $Q_{\rm cell} < 1$ may be ionized by photons produced in a neighbouring cell. In order to find the extent of ionized regions we therefore filter the $Q_{\rm cell}$ field using a sequence of real space top hat filters of radius R (with $0.3906 < R < 100\,h^{-1}$ Mpc), producing one smoothed ionization field Q_R per radius. At each point in the simulation box we find the largest R for which the filtered ionization field is greater than unity (i.e. ionized with $Q_R \geq 1$). All points within the radius R around this point are considered ionized. Ionization cells with $0 < Q_{\rm cell} < 1$ which are not part of an ionized $Q_R \geq 1$ region retain their values.

3.2 The cross-power spectrum

The 21-cm brightness temperature contrast may be written as

$$\tilde{\delta}_{21}(\boldsymbol{r}) = T_0(z)[1 - Q(\boldsymbol{r})](1 + \delta_{\rm DM,cell}), \quad (5)$$

where $T_0(z) = 23.8(\frac{\Omega_{\rm b}h^2}{0.021})[(\frac{0.15}{\Omega_{\rm m}h^2})(\frac{1+z}{10})]^{\frac{1}{2}}$ mK (Zaldarriaga, Furlanetto & Hernquist 2004). For convenience, we define $\delta_{21}(\boldsymbol{r}) \equiv \tilde{\delta}_{21}(\boldsymbol{r})/T_0(z)$, so that $\delta_{21}(\boldsymbol{r})$ is a dimensionless quantity. Galaxy overdensity is given by

$$\delta_{\rm gal}(\boldsymbol{r}) = \frac{\rho_{\rm gal}(\boldsymbol{r}) - \bar{\rho}_{\rm gal}}{\bar{\rho}_{\rm gal}}, \quad (6)$$

where $\rho_{\rm gal}(\boldsymbol{r})$ is a galaxy density field and $\bar{\rho}_{\rm gal}$ is mean density. Defining $\hat{\delta}_{21}(k)$ to be the Fourier transform of $\delta_{21}(k)$, the cross-power spectrum is given by

$$\left\langle \hat{\delta}_{21}(\boldsymbol{k}_1)\hat{\delta}_{\rm gal}(\boldsymbol{k}_2) \right\rangle \equiv (2\pi)^3 \delta_{\rm D}(\boldsymbol{k}_1 + \boldsymbol{k}_2) P_{21,\rm gal}(\boldsymbol{k}_1), \quad (7)$$

where $\delta_{\rm D}(k)$ is the Dirac delta function. The dimensionless cross-power spectrum is

$$\Delta^2_{21,\rm gal}(k) = \frac{k^3}{(2\pi^2)} P_{21,\rm gal}(k). \quad (8)$$

3.3 The cross-correlation function

The cross-correlation function is defined as

$$\xi_{1,2}(\boldsymbol{r}) = \langle \delta_1(\boldsymbol{x})\delta_2(\boldsymbol{x} + \boldsymbol{r}) \rangle\,. \quad (9)$$

We calculate the cross-correlation function using the Fourier transform,

$$\xi_{21,\rm gal}(r) = \frac{1}{(2\pi)^3} \int P_{21,\rm gal}(k) \frac{\sin kr}{kr} 4\pi k^2 {\rm d}k. \quad (10)$$

We note that the integration range of equation (10) forced by a finite box size introduces uncertainties for the predicted cross-correlation function (i.e. information about the turn over scale). However, this equation allows us to qualitatively investigate the cross-correlation function. We also calculate the cross-correlation coefficient,

$$r_{21,\rm gal}(k) = \frac{P_{21,\rm gal}(k)}{\sqrt{P_{21}(k)P_{\rm gal}(k)}}. \quad (11)$$

4 THE CORRELATION BETWEEN 21-CM EMISSION AND GALAXIES

In this section we present predictions for the cross-power spectrum, cross-correlation function and cross-correlation coefficient between 21-cm emission and galaxies as a function of redshift, luminosity and host halo mass (Section 4.1). We also discuss the effect of feedback processes on the cross-power spectrum, cross-correlation function and cross-correlation coefficient (Section 4.2).

4.1 Predictions for the correlation between 21-cm emission and galaxies

Fig. 2 shows the redshift evolution of the cross-power spectrum (top-left) and cross-correlation coefficient (bottom-left panel), and of the cross-correlation function (right panel) between redshifted 21-cm emission and galaxies. We show three examples which have UV magnitude limits, $M_{\rm AB}(1500\,\text{Å}) - 5\log(h) < -18$, in the model. This magnitude threshold corresponds to the deepest 'wide' area survey with Wide Field Camera 3/infrared and the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey on *Hubble Space Telescope* (Bouwens et al. 2011; Finkelstein et al. 2012). At each redshift, we calculate a mass-averaged ionization fraction, $\langle x_i \rangle$. From the correlation function, galaxies and 21-cm emission are anti-correlated at small separations while at large separations we find a weak correlation. These regions are separated by a transition wavenumber at which the cross-correlation coefficient and cross-correlation function change from negative to positive. Galaxies are correlated with 21-cm emission on scales larger than the ionized regions, but anti-correlated on smaller scales. The size of ionization regions therefore corresponds to this transition wavenumber. We find that the transition wavenumber from negative to positive cross-correlation coefficient increases as redshift decreases since the size of ionized regions generated by galaxies increases as the Universe evolves. We note that the cross-correlation function at $z = 6.197$ ($\langle x_i \rangle = 0.95$) has a shape that is different at small scales. We interpret this as being due to noise because the 21-cm emission regions are rare. In addition to the limitation of the integration range mentioned in Section 3.3, this noise indicates that a larger simulation box is necessary if one wants to measure the shape of the cross-correlation function at small k. Fig. 3 shows the comparison of cross-power spectra and cross-correlation functions between different host halo mass thresholds at $z = 7.272$ ($\langle x_i \rangle = 0.55$). We find that more massive haloes exhibit stronger anti-correlation as expected (Lidz et al.

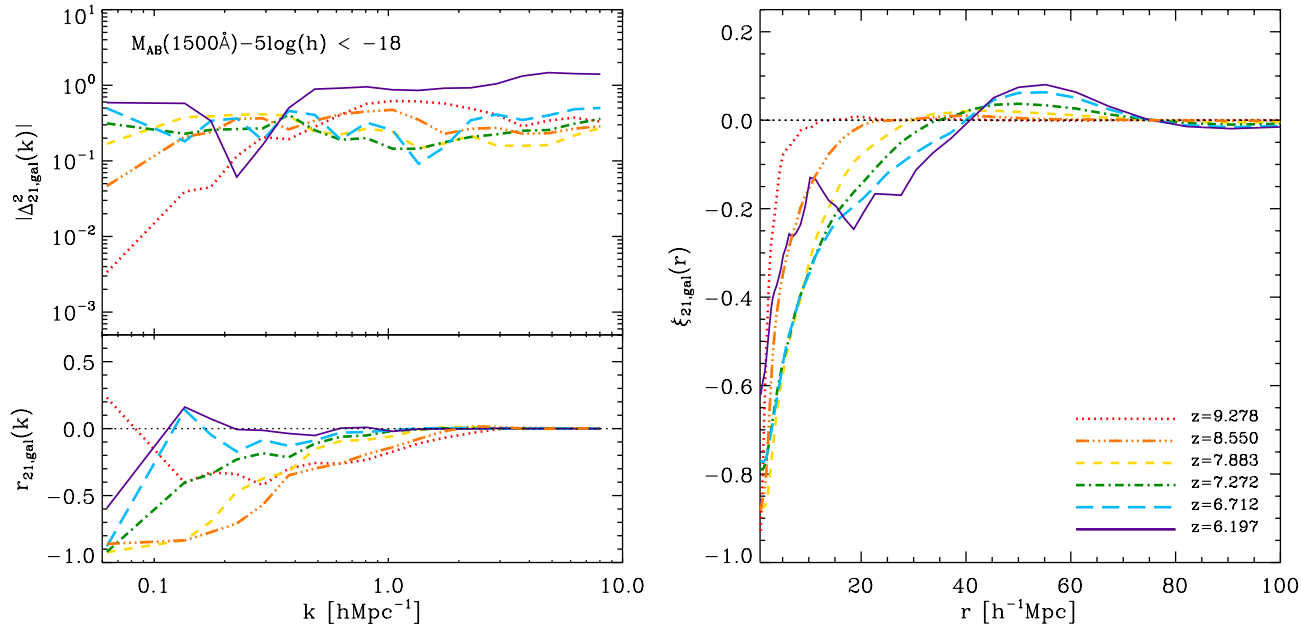

Figure 2. Redshift evolution of the cross-power spectrum and cross-correlation function between 21-cm fluctuations and the galaxies which have the UV magnitude less than −18 in the model. Left panel: the absolute value of the cross-power spectrum (top) and cross-correlation coefficient (bottom). Right panel: the corresponding cross-correlation function. In each panel, dotted (red), dash–three dotted (orange), dashed (yellow), dash–dotted (green), long-dashed (blue), and solid (purple) lines represent results from at z ($\langle x_i \rangle$) = 9.278 (0.056), 8.550 (0.16), 7.883 (0.36), 7.272 (0.55), 6.712 (0.75) and 6.197(0.95), respectively.

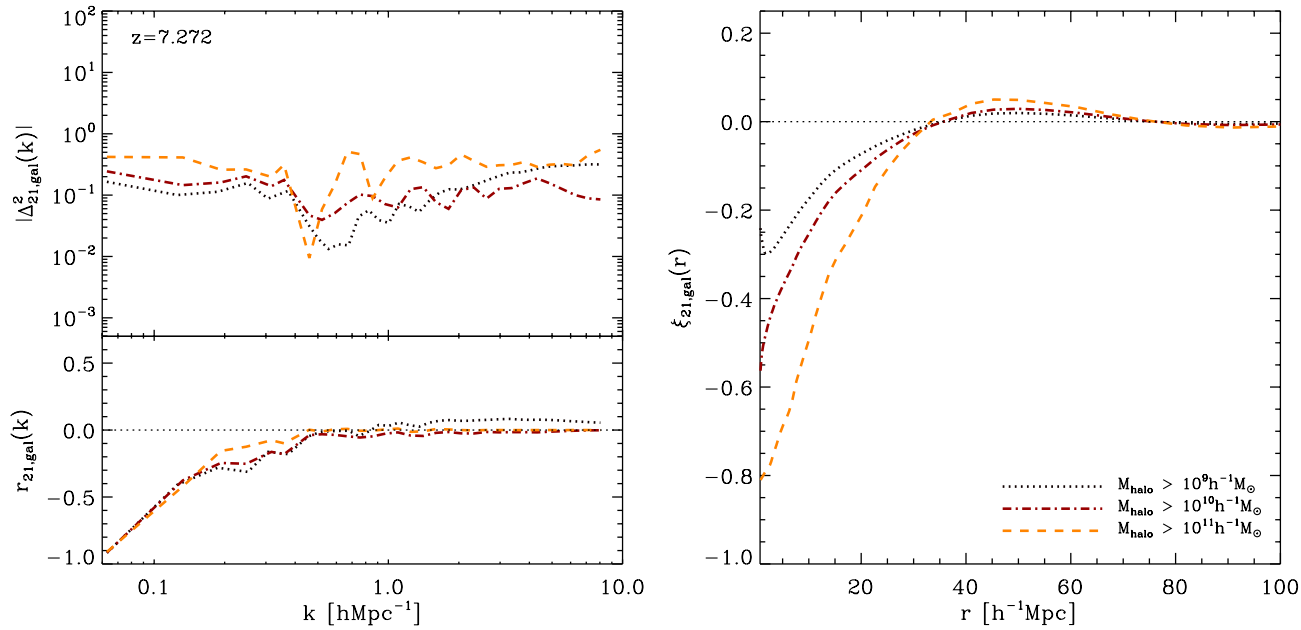

Figure 3. Comparison of the cross-power spectrum and cross-correlation function for different host halo mass thresholds at $z = 7.272$ ($\langle x_i \rangle = 0.55$). Left panel: the absolute value of the cross-power spectrum (top) and cross-correlation coefficient (bottom). Right panel: the corresponding cross-correlation function. In each panel, the dotted (dark brown), dot–dashed (brown), dashed (orange), long-dashed (yellow) lines show the cross-correlation using galaxies which are included in 10^9, 10^{10}, 10^{11} and 10^{12} h^{-1} $M_\odot$, respectively.

2009; Wiersma et al. 2013). The same trend is also shown in Fig. 4 where we compare the results from calculations with different UV magnitude [$M_{AB}(1500\,\text{Å}) - 5\log(h)$] thresholds. Fig. 4 shows that the transition wavenumber is similar for galaxy samples selected at different luminosity thresholds, since this scale is primarily set by the size of H II regions.

4.2 The effect of feedback processes

In order to investigate the effect on the power spectrum of different feedback processes in galaxy formation, we follow a similar method to Kim et al. (2013). We use the Lagos et al. (2012) galaxy formation model as our fiducial case, and then consider two variants of this (hereafter called NOSN models) which have SNe

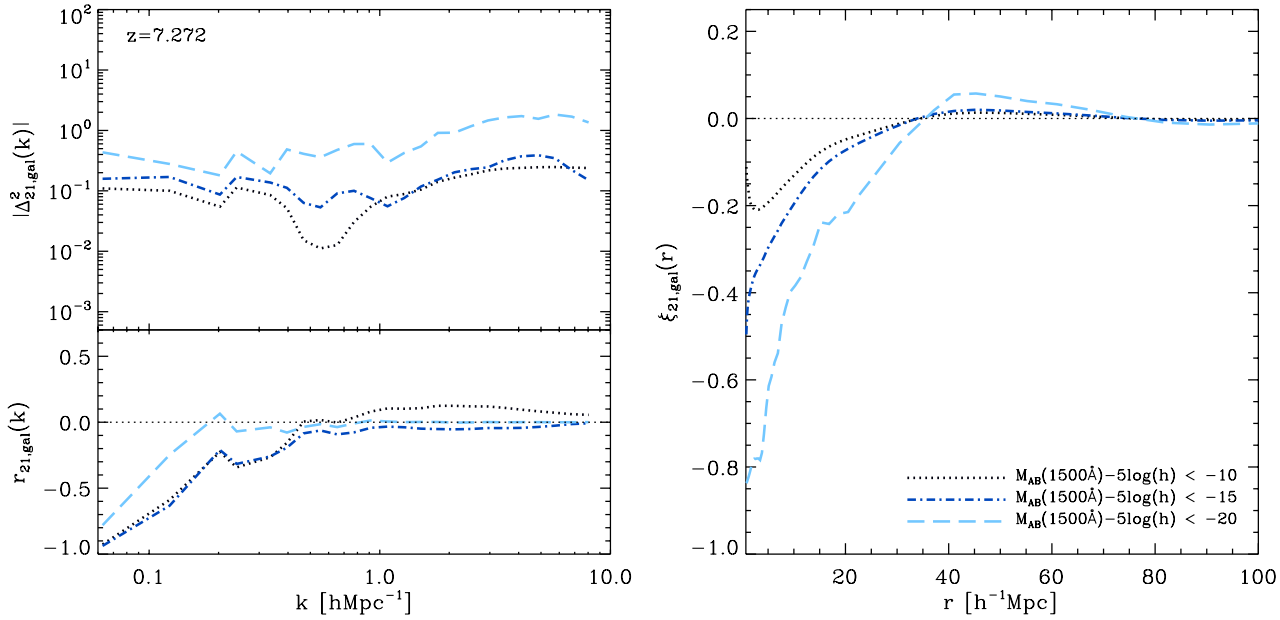

Figure 4. The same as Fig. 3 but results are computed based on different UV magnitude thresholds. In each panel, the dotted (black), dot–dashed (blue) and long dashed (sky-blue) lines show the cross-correlation using galaxy samples which are, respectively, more luminous than magnitude limits of −10, −15 and −20.

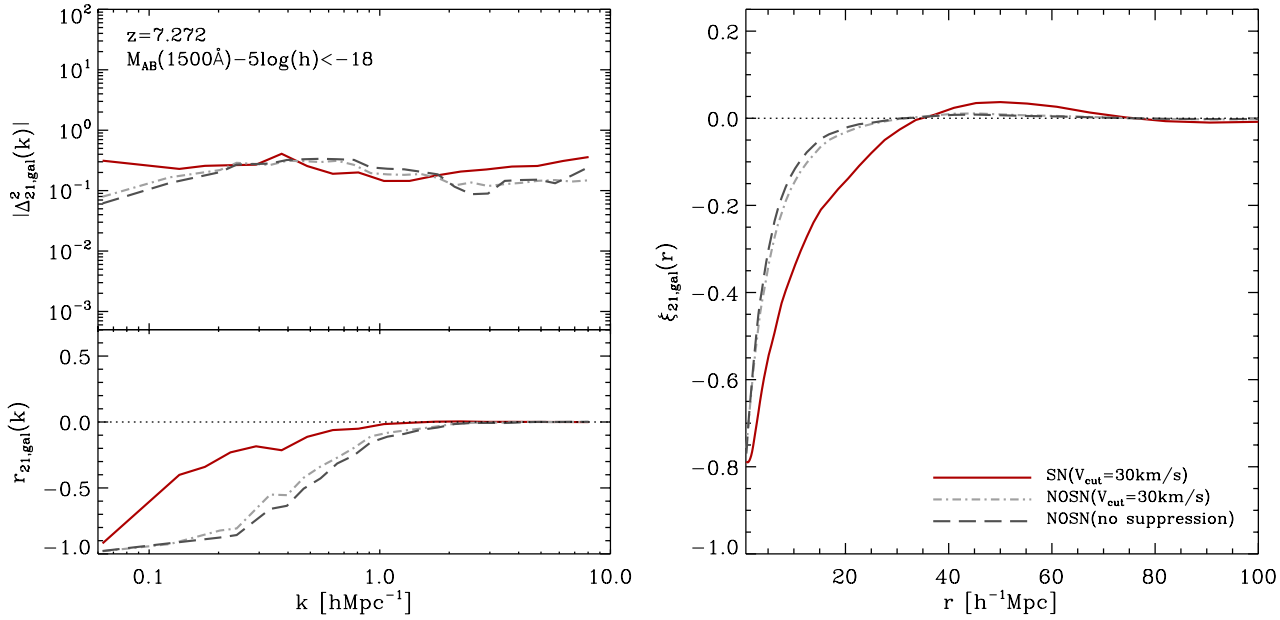

Figure 5. The same as Fig. 4 but results are computed based on different feedback processes. In each panel, solid (red), dot–dashed (light grey) and long dashed (dark grey) lines represent our model, NOSN($V_{\rm cut} = 30\,{\rm km\,s^{-1}}$), and NOSN(no suppression) models, respectively.

feedback turned off. We use two variants of the NOSN model. First, we consider the inclusion of photoionization feedback using $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$, where $V_{\rm cut}$ is a threshold value of the host halo's circular velocity (Kim et al. 2013). Secondly, we removed both SNe feedback and photoionization feedback by setting $V_{\rm cut} = 0\,{\rm km\,s^{-1}}$. We refer to this second model as NOSN (no suppression) in this paper. Since turning off SNe feedback in the Lagos et al. (2012) model changes the bright end of the UV luminosity function, we have changed some other parameters so that the NOSN models still match the observed UV luminosity functions at $z = 7.272$. Specifically, we introduce a stellar initial mass function dominated by brown dwarfs, with $\Upsilon = 4$, and also reduce the star formation time-scale in bursts by setting $f_{\rm dyn} = 2$ and $\tau_{*\rm burst,\,min} = 0.005$ Gyr [see Cole et al. (2000) and Lacey et al. (2011) for more details of these parameters]. In Fig. 5 we show the resulting comparison of cross-power spectra and cross-correlation functions at $z = 7.272$ ($\langle x_i \rangle = 0.55$). The locations of transition wavenumbers between the Lagos et al. (2012) model and the two NOSN models are significantly different (see also ionization structure for these models in Kim et al. 2013). In particular, the Lagos et al. (2012) model has a larger transition scale. On small scales, the cross-correlation function of the Lagos et al. (2012) model shows stronger anti-correlation than the two NOSN models between 21-cm emissions and galaxies.

We note that for the three different models, we have used the same mass-averaged ionization fraction, $\langle x_i \rangle$, at each redshift (listed in Section 3.1). While these models are forced to have the same ionization history, Fig. 5 still shows different cross-power spectrum shapes. This is because of the effect of feedback processes. SN feedback suppresses the formation of galaxies within small dark matter haloes, and consequently this process enhances the galaxy bias of the ionizing emission. Photoionization feedback also suppresses the formation of low-luminosity galaxies, but the effect is not significant compared to the effect of SNe feedback (see also Kim et al. 2013). This weak effect of photoionization feedback is revealed by small difference between two NOSN models.

5 DETECTABILITY

In this section we describe the error estimation of the cross-correlation coefficient (Section 5.1) and discuss observational requirements for future galaxy surveys (Section 5.2). Our examples are based on the MWA-like observations of the 21-cm signal combined with various hypothetical galaxy redshift surveys.

5.1 Error estimate in the cross-correlation coefficient

In order to estimate the sensitivity of future surveys, we calculate the error on the cross-correlation coefficient (Furlanetto & Lidz 2007; Lidz et al. 2009). For convenience we use the notation of Lidz et al. (2009) for the cross-correlation coefficient,

$$r_{21,\mathrm{gal}}(k) = \frac{P_{21,\mathrm{gal}}(k)}{\sqrt{P_{21}(k)P_{\mathrm{gal}}(k)}} \equiv \frac{A(k)}{\sqrt{B(k)C(k)}}. \quad (12)$$

The error on the cross-correlation coefficient can be written as

$$\frac{\sigma_r^2}{r^2}(k) = \frac{\sigma_A^2}{A^2}(k) + \frac{\sigma_B^2}{4B^2}(k) + \frac{\sigma_C^2}{4C^2}(k) - \frac{\sigma_{AB}^2}{AB}(k) - \frac{\sigma_{AC}^2}{AC}(k) + \frac{\sigma_{BC}^2}{2BC}(k). \quad (13)$$

This equation has variances of the cross-power spectrum between 21-cm and galaxy, and the auto-power spectra of both the 21-cm emission and galaxies. It also has the covariance between different pairs of power spectra. The components of equation (13) are given by

$$\sigma_A^2(k,\mu) = \mathrm{var}[P_{21,\mathrm{gal}}(k,\mu)] = \frac{1}{2}\left[P_{21,\mathrm{gal}}(k,\mu) + \sigma_B(k,\mu)\sigma_C(k,\mu)\right], \quad (14)$$

$$\sigma_B^2(k,\mu) = \mathrm{var}[P_{21}(k,\mu)] = \left[P_{21}(k,\mu) + \frac{T_{\mathrm{sys}}^2}{T_0^2}\frac{1}{Bt_{\mathrm{int}}}\frac{D^2\Delta D}{n(k_\perp)}\left(\frac{\lambda^2}{A_e}\right)^2\right]^2, \quad (15)$$

$$\sigma_C^2(k,\mu) = \mathrm{var}[P_{\mathrm{gal}}(k,\mu)] = \left[P_{\mathrm{gal}}(k,\mu) + n_{\mathrm{gal}}^{-1}e^{k_\parallel^2\sigma_\chi^2}\right]^2, \quad (16)$$

$$\sigma_{AB}^2(k,\mu) = \mathrm{cov}[P_{21,\mathrm{gal}}(k,\mu), P_{21}(k,\mu)] = \left[P_{21,\mathrm{gal}}(k,\mu)P_{21}(k,\mu)\right], \quad (17)$$

$$\sigma_{AC}^2(k,\mu) = \mathrm{cov}[P_{21,\mathrm{gal}}(k,\mu), P_{\mathrm{gal}}(k,\mu)] = \left[P_{21,\mathrm{gal}}(k,\mu)P_{\mathrm{gal}}(k,\mu)\right], \quad (18)$$

and

$$\sigma_{BC}^2(k,\mu) = \mathrm{cov}[P_{21}(k,\mu), P_{\mathrm{gal}}(k,\mu)] = \left[P_{21}(k,\mu)P_{\mathrm{gal}}(k,\mu)\right], \quad (19)$$

where μ is the cosine of the angle between $\boldsymbol{k}$ and the line of sight. To introduce large-scale redshift space distortions we use the relation $P(k,\mu) = (1+\beta\mu^2)^2P(k)$, where $\beta = \Omega_{\mathrm{m}}^{0.6}(z)/b$ and b is a bias factor, between the redshift space power spectrum and the real space (Kaiser 1987). We use $b_{\mathrm{gal}}^2(k) = P_{\mathrm{gal}}(k)/P_{\mathrm{DM}}(k)$ which is scale dependent and assume $b_{21} = 1$ for 21-cm power spectrum.

The first term in equation (15) comes from a sample variance within the finite volume of the survey and the second term comes from the thermal noise of the 21-cm telescope. We have assumed specifications of the MWA for the calculation of thermal noise. In the thermal noise term, $T_{\mathrm{sys}} \sim 250[(1+z)/7]^{2.6}$K denotes the system temperature of the telescope; $B = 8$ MHz is the survey bandpass; t_{int} is the integration observing time. We use 1000 h total observing time in this calculation; D and ΔD are the comoving distance to the survey volume and the comoving survey depth, $\Delta D = 1.7(\frac{B}{0.1\mathrm{MHz}})\sqrt{\frac{1+z}{10}}(\frac{\Omega_{\mathrm{m}}h^2}{0.15})^{-1/2}$ (Furlanetto, Oh & Briggs 2006), respectively; $n(k_\perp)$ denotes the number density of baselines in observing the transverse component of the wave vector, where $k_\perp = \sqrt{1-\mu^2}k$. Observing the signal of $k_\perp$ in each Fourier cell is related to the length of baseline and the antenna configuration. Here, we follow the method of Morales (2005), Bowman, Morales & Hewitt (2006) and Datta, Bharadwaj & Choudhury (2007) for calculation of $n(k_\perp)$. The maximum value of the transverse component of the wave vector is $k_{\perp,\,\mathrm{max}} = 2\pi L_{\mathrm{max}}/(D\lambda)$, where $L_{\mathrm{max}} = 750$ m is the maximum baseline distance in the antenna array. This limit is due to the maximum angular resolution of the telescope related to L_{max}. On the other hand, the minimum line-of-sight wavenumber is set by the bandpass $k_{\mathrm{min}} = 2\pi/\Delta D$; The observed wavelength is $\lambda = 0.21\,\mathrm{m} \times (1+z)$, and A_{e} is the effective collecting area of each antenna. We use $A_{\mathrm{e}} \sim N_{\mathrm{dip}}\lambda^2/4$ (Bowman et al. 2006), where $N_{\mathrm{dip}} = 16$ is the number of dipoles. We have assumed 500 antenna elements.[1]

From equations (14)–(21), we compute the errors of the power spectra averaged over a spherical shell of the logarithmic width $\epsilon = d\ln k$ for individual k modes. For example, the error of the cross-power spectrum is given by

$$\frac{1}{\sigma_A^2(k)} = \sum_\mu \frac{\epsilon k^3 V_{\mathrm{survey}}}{4\pi^2}\frac{\Delta\mu}{\sigma_A^2(k,\mu)}, \quad (20)$$

where V_{survey} is the effective survey volume for a radio telescope, $V_{\mathrm{survey}} = D^2\Delta D(\lambda^2/A_{\mathrm{e}})$. The value of λ^2/A_{e} corresponds to the solid angle of the survey, which for the MWA corresponds to $\sim$800 deg^2. Note that if the galaxy survey volume is less than the 21-cm survey volume, then the variance is increased by a factor of $V_{\mathrm{survey,\,21}}/V_{\mathrm{survey,\,gal}}$. The MWA is designed to operate at frequencies between 80 and 300 MHz in order to observe the 21-cm signal at $6 < z < 30$. When the 21-cm signal is observed at $z \sim 7$, the wavelength is $\sim$1.7 m corresponding to $\sim$200 MHz.

Fig. 6 shows the 21-cm power spectrum with errors estimated based on equation (15) for cases including different feedback processes. The 21-cm power spectra show obvious differences between the models for SNe feedback, especially at large scales. Fig. 6

[1] The down scoped MWA has been constructed with 128 antennas. We use 500 here, corresponding to an upgraded array.

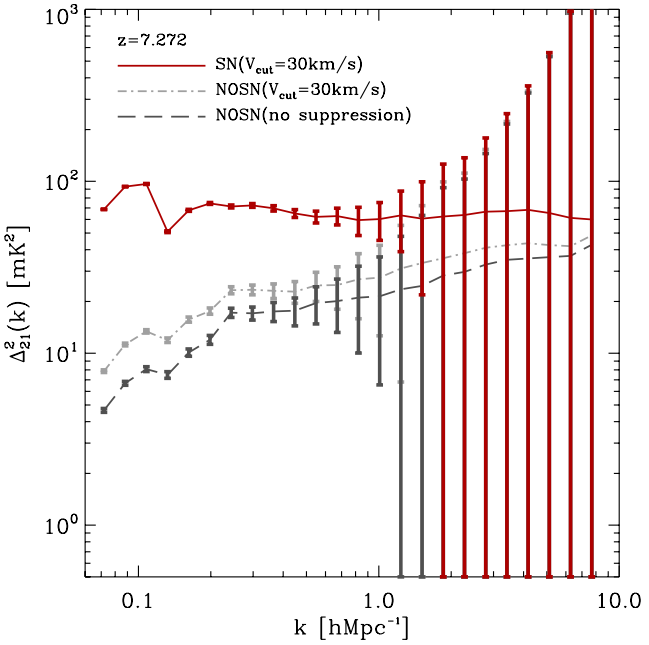

Figure 6. The 21-cm power spectrum with estimated errors, based on an 800 deg^2 survey area, at $z = 7.272$. We assume 1000 h total observing time, and, based on the assumption of 8 MHz bandwidth, the survey depth is about 0.2 redshift units. Red represents the power spectrum from our model including SNe feedback with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$. The light grey and dark grey lines represent power spectrum from the NOSN models with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$ and no suppression, respectively.

reinforces the importance of detailed modelling of galaxy formation during reionization (Kim et al. 2013).

The error on the galaxy power spectrum is expressed in equation (16). The galaxy shot-noise is dependent on the number density of galaxies observable ($n_{\rm gal}$), $k_\parallel = \mu k$, and $\sigma_\chi = c\sigma_z/H(z)$, where σ_z is the galaxy redshift error. Here, we assume a Gaussian distribution of redshift errors.

5.2 Observational requirements for future galaxy surveys

Following Lidz et al. (2009), we begin by considering a galaxy number density of $n_{\rm gal} = 1.6 \times 10^{-4}\,h^3\,{\rm Mpc^{-3}}$ for a survey in combination with 21-cm observations from the MWA. To match this number density in our galaxy catalogue we use a magnitude threshold, with a value of -19.4 at $z = 7.272$ and -19.8 at $z = 6.712$, in UV magnitude $[M_{\rm AB}(1500\,{\rm Å}) - 5\,\log(h)]$. We also match the number density of NOSN models in the same way. To find the general requirements for detection of the cross-correlation, in Fig. 7 we show the signal-to-noise ratio (S/N) for the cross-correlation coefficient as a function of survey area ($A_{\rm survey}$) and redshift error (σ_z). In our calculations, we assume 1000 h total observing time for the MWA. The left panel of Fig. 7 shows the total S/N, which is calculated by summing up the S/N in each k bin,

$$({\rm S/N})^2_{\rm total} = \sum_i^{N_{\rm bin}} \left(\frac{\Delta k}{\epsilon k_i}\right) ({\rm S/N})^2_i, \qquad (21)$$

where i represent ith bin and Δk is the bin size. We assume a redshift error (σ_z) of 0.05 as an example value from narrow-band survey for Lyman α emitters (Ouchi et al. 2008, 2010). The S/N is calculated with a survey area of 800 deg^2, and then scaled the S/N with the relation ${\rm S/N} \propto 1/\sqrt{A_{\rm survey}}$. Our default model that includes SNe feedback with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$ shows increased S/N compared with the results of the NOSN models. The default model predicts a 3σ detection of cross-correlation with a 2 deg^2 survey area. Survey areas greater than 10 deg^2 will provide detailed high S/N measurements.

As a specific example, we also calculate the total S/N as a function of σ_z by assuming the survey area of $A_{\rm survey} = 5\,{\rm deg^2}$. The total S/N in the central panel of Fig. 7 shows that measurements will require redshift uncertainties less than 0.1. The NOSN models show a similar shape to the default model, but have lower S/N. Lower accuracy redshifts ($\sigma_z > 0.1$) wash out the cross-correlation signal. An error of $\sigma_z \sim 0.1$ provides measurement only on larger scales ($k[h\,{\rm Mpc^{-1}}] < 0.2$) (right-hand panel in Fig. 7). To measure the cross-correlation over a broad range of k, redshift uncertainties, σ_z, less than 0.01 will be required.

Fig. 7 illustrates the conditions required for measurement of the 21-cm-galaxy cross-correlation. Before concluding we discuss these requirements with respect to real galaxy surveys. In this study, we have used UV magnitude cuts to select galaxy samples which relate to observed Lyman-break galaxies. However, Lyman-break galaxies, which are photometrically selected, have $\sigma_z \gtrsim 0.5$ at $z \sim 6.5$ (Beckwith et al. 2006) much longer than the $\sigma_z < 0.1$ requirement. As a result, Lyman-break surveys will not be sufficient to detect the cross-correlation (Wiersma et al. 2013). On the other hand, Lyα emitters selected from narrow-band surveys have $\sigma_z \sim 0.05$–0.1 (Ouchi et al. 2008, 2010). Thus, a detection could be made based on the precision and volume of current Lyα surveys. Our semi-numerical model does not predict Lyα luminosity (see Orsi et al. 2008). However, the difference between simulated populations of Lyα emitters and Lyman-break galaxies is found not to be significant (Dayal & Ferrara 2012). For the purpose of our calculation, we therefore use star-forming galaxies with UV magnitudes $[M_{\rm AB}(1500\,{\rm Å}) - 5\,\log(h)]$ corresponding to the number density of $n_{\rm gal} = 1.6 \times 10^{-4}\,h^3\,{\rm Mpc^{-3}}$, which is seen in the Subaru Deep Field at $z \sim 6.6$ (Kashikawa et al. 2006).

The largest Lyα survey (Ouchi et al. 2010) covered 1 deg^2 at $z \sim 6.6$, and used a narrow-band filter with a central wavelength of 9196 Å and a full width at half-maximum of 132 Å. These values correspond to a survey depth of $\Delta z \sim 0.11$ at $z = 6.6$. This is smaller than, but comparable to, the survey depth of MWA observation which is $\Delta z \sim 0.3$ corresponding to the bandwidth of 8 MHz assumed for this paper. For the survey at $z \sim 7.3$, Shibuya et al. (2012) have a survey depth of $\Delta z \sim 0.18$, which used the central wavelength of 10052 Å and a full width at half-maximum of 214 Å. This is also smaller than, but comparable to, the MWA observation of the survey depth of $\Delta z \sim 0.38$.

While 1 deg^2 represents the largest high-redshift survey at the current time, future surveys will be larger. For example, in the next 5 years Hyper Suprime-Cam on the Subaru telescope will observe 10^5 galaxies at $z \sim 5.7$ and 6.5 in a survey area of $\sim$30 deg^2, and 100 s of galaxies at $z \sim 7$ in a survey area of 3 deg^2 (M. Ouchi, private communication). As shown in Fig. 7, this increased survey area will improve the S/N, so that the cross-power spectra signal could be detected with high significance.[2]

Based on the requirement from Section 5.2, we assume a 5 deg^2 galaxy survey field and a redshift error of 0.05 for a future galaxy survey. We also assume 1000 h total observing time. Fig. 8 shows the predicted errors for the cross-correlation coefficient within spherical

[2] The Subaru Deep Field is not accessible to the MWA. See Wiersma et al. (2013) for a calculation of LOFAR sensitivity.

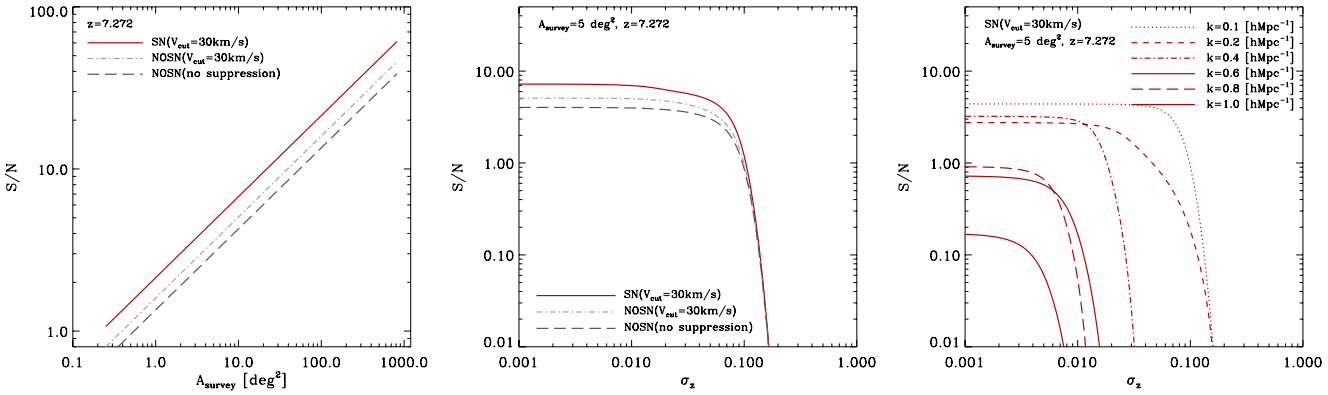

Figure 7. The S/N for the cross-correlation coefficient as a function of survey area and relative redshift error at $z = 7.272$. Left panel: plots of S/N as a function of survey area ($A_{\rm survey}$) for different models. We assume $\sigma_z = 0.05$. Central panel: a plot of S/N as a function of redshift error, σ_z, with $A_{\rm survey} = 5\,{\rm deg}^2$ for the default model. In the left and central panels, solid (brown), long-dashed (light grey) and dotted (dark grey) lines represent results from our model including SNe feedback with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$ (the default model), the NOSN models with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$ and no suppression, respectively. Right panel: plots of S/N as a function of σ_z at different wavenumbers for default model. Dotted, dashed, dot–dashed, three dot–dashed, long dashed and solid lines represent $k = 0.1$, 0.2, 0.4, 0.6, 0.8 and 1.0 $h\,{\rm Mpc^{-1}}$, respectively. In each panel, we assume 1000 h total observing time.

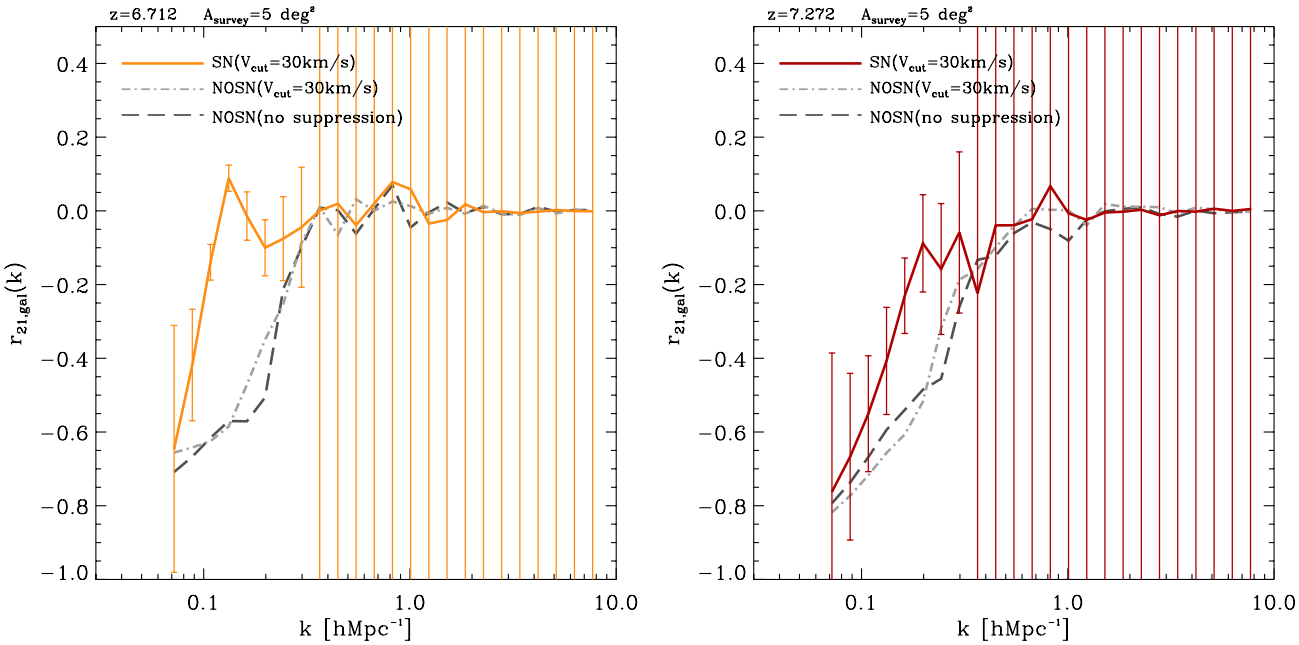

Figure 8. The cross-correlation coefficient at $z = 6.712$ (left panel) and 7.272 (right panel). The error bars are calculated for spherical bins of logarithmic width $\epsilon \equiv d\ln k = 0.5$. We assume a 5 deg^2 galaxy survey field, 1000 h total observing time, the redshift error of 0.05 and galaxy number density of the Subaru Deep Field survey. The solid (orange and red) lines represent the power spectrum from the default model (Lagos et al. 2012) including SNe feedback with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$. The long dashed (light grey) and dot–dashed (dark grey) lines represent power spectrum from the NOSN models with $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$ and no suppression, respectively.

bins of logarithmic width $\epsilon = 0.5$ at $z = 6.712$ and 7.272 for such a galaxy survey combined with the MWA. The estimated errors are exponentially increased near the wavenumber of 1 $h\,{\rm Mpc^{-1}}$, because of the limit of $k_{\perp,\,\rm max}$ for a 21-cm survey. We compare the result with the cross-correlation coefficient from the NOSN models. The result shows that we could observationally distinguish our default model from two different NOSN models.

Lyα observations at $z \gtrsim 7$ over a large area are very challenging. The latest Lyα survey at $z \sim 7.3$ (Shibuya et al. 2012) has a galaxy number density of $\sim 6.7 \times 10^{-6}$ covering a survey area of 0.48 deg^2. This value is smaller than the value we assume for our error estimation. Computing the cross-power spectrum corresponding to this number density is not possible owing to the limited box size of our simulation. However, we have checked that the estimated error would approximately increase by a factor of 2, when using this number density.

6 SUMMARY AND CONCLUSIONS

In this study we have investigated evolution of the cross-power spectrum, cross-correlation function and cross-correlation coefficient between 21-cm emission and galaxies using the model of Kim et al. (2013). This model combines the hierarchical galaxy formation model GALFORM implemented within the Millennium-II dark matter simulation, with a semi-numerical scheme to describe the resulting ionization structure. We find that there is a transition wavenumber,

k, at which the cross-correlation coefficient changes from negative to positive (Lidz et al. 2009). This transition wavenumber is associated with the size of the ionized regions generated by galaxies, and increases with decreasing redshift. We also find the same trend in the cross-power spectrum and cross-correlation function. We calculated the cross-power spectrum as a function of UV luminosity [$M_{\rm AB}(1500\,{\rm \AA}) - 5\log(h)$] and host halo mass. These calculations reveal that bright galaxies and galaxies residing in massive haloes have stronger anti-correlation, but a similar transition wavenumber.

We have studied observational uncertainties in measurement of the cross-correlation coefficient based on the specifications of an upgraded (512 tile) Murchison Widefield Array (MWA) combined with galaxy surveys. The results show that the cross-power spectrum signal could be detected when combined with more than 3 deg^2 of a galaxy survey at the depth of the future galaxy survey having redshift error <0.1. We have also investigated the dependence on the inclusion of feedback processes in the galaxy modelling. We find that the amplitude of the cross-correlation is larger when SNe feedback is considered and that the cross-correlation coefficient has a different shape compared to a model with no SNe feedback. Thus the cross-correlation could be used to determine the importance of SNe feedback in high-redshift galaxies. Our results imply that detailed modelling of reionization processes and galaxy formation is required to predict an accurate cross-correlation between 21-cm emission and galaxies, and to interpret future observational measurements.

ACKNOWLEDGEMENTS

HSK is supported by a Super-Science Fellowship from the Australian Research Council. JSBW acknowledges the support of an Australian Research Council Laureate Fellowship. The Centre for All-sky Astrophysics is an Australian Research Council Centre of Excellence, funded by grant CE110001020. This work was supported in part by the Science and Technology Facilities Council rolling grant ST/I001166/1 to the ICC. The Millennium II Simulation was carried out by the Virgo Consortium at the supercomputer centre of the Max Planck Society in Garching. Calculations for this paper were partly performed on the ICC Cosmology Machine, which is part of the DiRAC Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS and Durham University.

REFERENCES

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Beckwith S. V. W. et al., 2006, AJ, 132, 1729
Benson A. J., Sugiyama N., Nusser A., Lacey C. G., 2006, MNRAS, 369, 1055
Bouwens R. J. et al., 2011, ApJ, 737, 90
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Bowman J. D., Morales M. F., Hewitt J. N., 2006, ApJ, 638, 20
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150
Ciardi B., Stoehr F., White S. D. M., 2003, MNRAS, 343, 1101
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Datta K. K., Bharadwaj S., Choudhury T. R., 2007, MNRAS, 382, 809
Dayal P., Ferrara A., 2012, MNRAS, 421, 2568
Finkelstein S. L. et al., 2012, ApJ, 758, 93
Furlanetto S. R., Lidz A., 2007, ApJ, 660, 1030
Furlanetto S., Oh S. P., Briggs F., 2006, Phys. Rep., 433, 181
Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., 2007, MNRAS, 376, 534
Iliev I. T., Mellema G., Pen U.-L., Bond J. R., Shapiro P. R., 2008, MNRAS, 384, 863
Iliev I. T., Moore B., Gottlöber S., Yepes G., Hoffman Y., Mellema G., 2011, MNRAS, 413, 2093
Kaiser N., 1987, MNRAS, 227, 1
Kashikawa N. et al., 2006, ApJ, 648, 7
Kim H.-S., Wyithe J. S. B., Raskutti S., Lacey C. G., Helly J. C., 2013, MNRAS, 428, 2467
Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2011, MNRAS, 412, 1828
Lagos C. d. P., Bayet E., Baugh C. M., Lacey C. G., Bell T. A., Fanidakis N., Geach J. E., 2012, MNRAS, 426, 2142
Lidz A., Zahn O., Furlanetto S. R., McQuinn M., Hernquist L., Zaldarriaga M., 2009, ApJ, 690, 252
Lidz A., Furlanetto S. R., Oh S. P., Aguirre J., Chang T.-C., Doré O., Pritchard J. R., 2011, ApJ, 741, 70
Mesinger A., Furlanetto S., 2007, ApJ, 669, 663
Morales M. F., 2005, ApJ, 619, 678
Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, MNRAS, 391, 1589
Ouchi M. et al., 2008, ApJS, 176, 301
Ouchi M. et al., 2010, ApJ, 723, 869
Raičević M., Theuns T., Lacey C., 2011, MNRAS, 410, 775
Shibuya T., Kashikawa N., Ota K., Iye M., Ouchi M., Furusawa H., Shimasaku K., Hattori T., 2012, ApJ, 752, 114
Shin M.-S., Trac H., Cen R., 2008, ApJ, 681, 756
Sokasian A., Abel T., Hernquist L., Springel V., 2003, MNRAS, 344, 607
Trac H., Cen R., 2007, ApJ, 671, 1
Trac H., Cen R., Loeb A., 2008, ApJ, 689, L81
Wiersma R. P. C. et al., 2013, MNRAS, 432, 2615
Zahn O., Lidz A., McQuinn M., Dutta S., Hernquist L., Zaldarriaga M., Furlanetto S. R., 2007, ApJ, 654, 12
Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, ApJ, 608, 622

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