

# Predictions for the 21cm-galaxy cross-power spectrum observable with SKA and future galaxy surveys

Dijana Vrbanc<sup>1,\*</sup>, Benedetta Ciardi<sup>1</sup>, Vibor Jelić<sup>2</sup>, Hannes Jensen<sup>3</sup>, Ilian T. Iliev<sup>4</sup>, Garrelt Mellema<sup>3</sup>, Saleem Zaroubi<sup>5,6,7</sup>

<sup>1</sup>Max Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85748 Garching bei München, Germany

<sup>2</sup>Ruder Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

<sup>3</sup>Department of Astronomy and Oskar Klein Centre, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden

<sup>4</sup>Astronomy Centre, Department of Physics and Astronomy, Pevensey II Building, University of Sussex, Falmer, Brighton BN1 9QH, UK

<sup>5</sup>Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

<sup>6</sup>Department of Natural Sciences, Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel

<sup>7</sup>Department of Physics, The Technion, Haifa 32000, Israel

Accepted - Received -; in original form -

## ABSTRACT

In this paper we use radiative transfer + N-body simulations to explore the feasibility of measurements of cross-correlations between the 21 cm field observed by the Square Kilometer Array (SKA) and high- $z$  Lyman Alpha Emitters (LAEs) detected in galaxy surveys with the Subaru Hyper Supreme Cam (HSC), Subaru Prime Focus Spectrograph (PFS) and Wide Field Infrared Survey Telescope (WFIRST). 21cm-LAE cross-correlations are in fact a powerful probe of the epoch of reionization as they are expected to provide precious information on the progress of reionization and the typical scale of ionized regions at different redshifts. The next generation observations with SKA will have a noise level much lower than those with its precursor radio facilities, introducing a significant improvement in the measurement of the cross-correlations. We find that an SKA-HSC/PFS observation will allow to investigate scales below  $\sim 10 h^{-1}$  Mpc and  $\sim 60 h^{-1}$  Mpc at  $z = 7.3$  and  $6.6$ , respectively. WFIRST will allow to access also higher redshifts, as it is expected to observe spectroscopically  $\sim 900$  LAEs per square degree and unit redshift in the range  $7.5 \leq z \leq 8.5$ . Because of the reduction of the shot noise compared to HSC and PFS, observations with WFIRST will result in more precise cross-correlations and increased observable scales.

**Key words:** galaxies: high redshift - cosmology:observations - reionization - intergalactic medium

## 1 INTRODUCTION

The detection of the 21 cm line from high-redshift neutral hydrogen (HI) is one of the last observational frontiers and the goal of several present and upcoming radio facilities such as the Low Frequency ARray<sup>1</sup> (LOFAR), the Murchison Widefield Array<sup>2</sup> (MWA), the Hydrogen Epoch of Reionization Array<sup>3</sup> (HERA), and the Square Kilometer Array<sup>4</sup> (SKA). Such observational campaigns will provide the first constraints on the history of hydrogen reionization

and the evolution of the properties of the high- $z$  intergalactic medium (IGM) (see Pritchard & Loeb 2012 for a review). Tomography of the 21 cm line will offer information on the temperature and ionization state of the IGM, on the topology of reionization and possibly on the properties of the sources of ionizing photons (see e.g. Tozzi et al. 2000; Ciardi & Madau 2003; Zaroubi et al. 2012), while the detection (or even an upper limit) of fluctuations and power spectrum of the differential brightness temperature will provide invaluable insight on the timing of reionization and statistical estimates of its properties (e.g. Madau, Meiksin & Rees 1997; Mellema et al. 2006; Baek et al. 2009; Patil et al. 2014). The amount of HI along lines of sight towards high- $z$  radio loud sources could instead be measured with 21 cm absorption systems (e.g. Carilli, Gnedin & Owen 2002; Furlanetto 2006; Xu et al. 2009; Ciardi et al. 2013; Semelin 2015).

\* E-mail: dvrbanec@mpa-garching.mpg.de

<sup>1</sup> www.lofar.org

<sup>2</sup> http://www.mwatelescope.org

<sup>3</sup> https://reionization.org

<sup>4</sup> https://www.skatelescope.org

It has also been suggested that correlating the 21 cm signal with observations in different frequency bands would reduce systematic effects and confirm the cosmological origin of the signal, in addition to potentially provide additional and independent information on both the reionization process and the sources it is being correlated with. Investigations have been made with respect to correlations with the cosmic microwave background radiation (Salvaterra et al. 2005; Jelić et al. 2010; Ma et al. 2018b), high- $z$  galaxies (Lidz et al. 2009; Wiersma et al. 2013; Vrbanec et al. 2016; Sobacchi, Mesinger & Greig 2016; Hutter et al. 2017; Kubota et al. 2018), NIR (Fernandez et al. 2014) and X-ray (Ma et al. 2018a) background radiation. Particular attention has been devoted to the possibility of cross-correlating the 21 cm signal with Ly $\alpha$  emitters (LAEs) surveys, as the 21 cm field is expected to anti-correlate (in terms of cross-power spectra and cross-correlation coefficients) with the galaxy field on large scales (where most of the HI lies once reionization is well underway but where there is a paucity of galaxies) and become roughly un-correlated on small scales, within the ionized regions. More specifically, Sobacchi, Mesinger & Greig (2016) investigated the effect of morphology and the mass of LAEs hosting halos on the cross-correlations, showing that LOFAR or SKA in combination with the Subaru's Hyper Suprime-Cam (HSC) should be able to distinguish between a half-neutral and a completely ionized universe. Hutter, Trott & Dayal (2018) expanded on that investigation demonstrating that neutral hydrogen fractions of 0.1, 0.25, 0.5 could be distinguished. Kubota et al. (2018) confirmed the detectability of the cross-correlations with MWA, SKA, HSC and the Subaru Prime Focus Spectrograph (PFS)<sup>5</sup>, focusing on the error budget of these instruments and possible LAEs observational strategies.

In Vrbanec et al. (2016, hereafter V2016) we had investigated the feasibility of cross-correlations at  $z = 6.6$  and  $7.3$  with LOFAR and HSC, as an agreement was in place among the two teams to observe the common field ELAIS-N1. As the LOFAR peak performance turned out to be at redshifts higher than initially envisioned (i.e.  $z \sim 8.5 - 10.5$ , Patil et al. 2017), this has compromised the ability of a cross-correlation exercise, discussed in V2016, to yield significant detection. Therefore, we extend the investigation performed in V2016 to the SKA with the HSC, the PFS and the Wide Field Infrared Survey Telescope (WFIRST)<sup>6</sup>.

The paper is structured as follows. In section 2 we present the simulations and methodology adopted, in section 3 the results, followed by a discussion in section 4. Our conclusions are summarized in section 5.

## 2 METHODOLOGY

In our cross-correlation analysis we follow closely what was done in V2016. We thus refer the reader to that paper for more details.

To compute the cross-correlations we have used the simulation of reionization described in Iliev et al. (2014). This has been run in a box of length  $425h^{-1}$  cMpc (equivalent

to  $\sim 4$  deg at  $z=7$ ) with 165 billion particles distributed on a  $10,976^3$  cells grid ( $3.9h^{-1}$ kpc gravity force resolution). The N-body simulation has been initialized at  $z = 300$  using the Zel'dovich approximation together with a power spectrum of the linear fluctuations generated with the code CAMB (Lewis, Challinor & Lasenby 2000). The outputs of the simulation have been mapped onto a grid of  $504^3$  cells and post-processed with the radiative transfer code C<sup>2</sup>-RAY (Mellema et al. 2006a) to model IGM reionization. Finally, the results have been used as input to a Ly $\alpha$  radiative transfer method to study high- $z$  LAEs and their observability (Jensen et al. 2013, 2014). The LAEs so modelled reproduce the observed luminosity functions. The reader should refer to the original papers for more details.

As in V2016, we evaluate the differential brightness temperature associated to the neutral hydrogen distribution obtained from the simulations as (e.g. Field 1959; Madau, Meiksin & Rees 1997; Furlanetto, Oh & Briggs 2006):

$$\delta T_b = 28.5 \text{ mK} (1 + \delta) x_{HI} \left( \frac{\Omega_b}{0.042} \frac{h}{0.73} \right) \times \left[ \left( \frac{1+z}{10} \right) \left( \frac{0.24}{\Omega_m} \right) \right]^{1/2}, \quad (1)$$

where  $x_{HI}(1 + \delta) = n_{HI}/\langle n_H \rangle$  is the average density of neutral hydrogen in units of the average density of hydrogen at redshift  $z$ , and we assume the spin temperature to be much larger than the temperature of the CMB.  $\delta T_b$  is then used to produce mock observations with SKA (see section 3). Subaru and WFIRST mock observations are instead produced using the Ly $\alpha$  intrinsic and transmitted luminosities extracted from the same simulations (see section 3).

Following Lidz et al. (2009), we decompose the 21cm-galaxy cross-power spectrum at wave number  $k = |k|$ ,  $\Delta_{21,gal}^2(k)$ , into three terms:

$$\begin{aligned} \Delta_{21,gal}^2(k) &= \tilde{\Delta}_{21,gal}^2(k) / \delta T_{b0} \\ &= \langle x_{HI} \rangle [\Delta_{x_{HI},gal}^2(k) + \Delta_{\rho,gal}^2(k) + \Delta_{x_{HI}\rho,gal}^2(k)], \quad (2) \end{aligned}$$

where  $\Delta_{x_{HI},gal}^2$ ,  $\Delta_{\rho,gal}^2$  and  $\Delta_{x_{HI}\rho,gal}^2$  are, respectively, the neutral fraction-galaxy power spectrum, the density-galaxy power spectrum and neutral density-galaxy power spectrum.  $\delta T_{b0}$  is the 21 cm brightness temperature relative to the CMB for neutral gas at the mean density, and  $\langle x_{HI} \rangle$  is the volume-averaged HI fraction. Given the fields  $a$  and  $b$ ,  $\Delta_{a,b}^2$  is their dimensionless cross-power spectrum, and it is written as  $\Delta_{a,b}^2(k) = k^3 P_{a,b}(k) / 2\pi^2$  in 3D, and  $\Delta_{a,b}^2(k) = k^2 P_{a,b}(k) / \pi$  in 2D. Here  $P_{a,b}$  is the dimensional cross-power spectrum of the two fields, which are represented in terms of their fluctuations at any given location  $r$ , i.e.  $\delta_a(r) = (a(r) - \langle a \rangle) / \langle a \rangle$ , and the same for  $b$ . We refer the reader to Lidz et al. (2009) and V2016 for a more detailed discussion of the various terms.

<sup>7</sup> The theoretical cross-power spectrum is evaluated as  $\langle a \rangle = (\sum_{i=1}^N a_i) / N$ , with  $N$  number of pixels contained in the part of the simulation considered. With the exception of the galaxy field in mock observations, we calculate all quantities in this way. The galaxy field is evaluated as  $\langle N_{gal} \rangle = N_{gal} / V$ , with  $V$  volume of the survey and  $N_{gal}$  number of galaxies in the mock observation. This choice was made because of an easier comparison to the shot noise of the power spectrum,  $P_{shot}(k) = 1/n_{gal}$ , where  $n_{gal}$  is the average number of galaxies in the volume covered by the survey.

<sup>5</sup> <http://pfs.ipmu.jp/>

<sup>6</sup> <https://wfIRST.gsfc.nasa.gov/>

In addition, we have evaluated the cross-correlation function, which characterizes the changes in real space of the correlation between two fields,  $a$  and  $b$ . We write their cross-correlation function as  $\xi_{a,b}(\mathbf{r}) = \langle \delta_a(\mathbf{x})\delta_b(\mathbf{x} + \mathbf{r}) \rangle$ , where  $\delta_a(\mathbf{x})$  and  $\delta_b(\mathbf{x})$  are the fractional fluctuation of field  $a$  and  $b$  at location  $\mathbf{x}$ , respectively. The observed cross-correlation function is evaluated as:

$$\xi_{21,LAE}(r) = \frac{\sum_{\mathbf{x}} \delta_{LAE}(\mathbf{x})\delta_{21}(\mathbf{x} + \mathbf{r})}{N_{pair}(r)}, \quad (3)$$

where  $N_{pair}(r)$  is the number of 21cm-LAE pairs at a separation  $r$ , while  $\delta_{LAE}$  and  $\delta_{21}$  are the fractional fluctuations of the LAE and 21cm fields, respectively.

To evaluate the cross-correlations, we used simulation boxes at  $z = 6.68, 7.3$ , and  $8.06$ , when the volume (mass) averaged ionized fractions are  $\langle x \rangle = 0.93$  (0.95),  $0.48$  (0.58), and  $0.21$  (0.30), respectively. These redshifts were chosen because HSC has two narrow-band filters observing at  $z = 6.6$  and  $7.3$ , while WFIRST will observe in the redshift range  $7.5 < z < 8.5$ .

To compute a 3D cross-power spectrum we merged bins so that  $\Delta k > 0.02 h \text{ Mpc}^{-1}$ , which corresponds to the smallest mode resolved by a field of view (FoV) of  $16 \text{ deg}^2$ , i.e. equivalent to the one covered by our simulations. We also avoided correlations in power induced by the window function<sup>8</sup> by having a binning with  $\Delta \log k = 0.02$ . As the FoV of HSC and PFS is  $7 \text{ deg}^2$  and  $1.7 \text{ deg}^2$  at  $z = 6.6$  and  $7.3$ , respectively (see section 3), when evaluating cross-power spectra for these instruments we used  $\Delta \log k = 0.03$  (0.05) and  $\Delta k > 0.04$  (0.07)  $h \text{ Mpc}^{-1}$  for  $z = 6.6$  (7.3).

### 3 CORRELATIONS OBSERVABLE WITH SKA AND FUTURE GALAXY SURVEYS

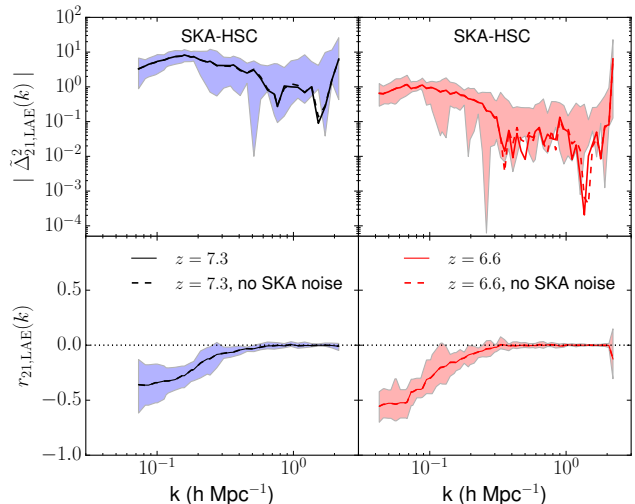
While the reader can find in V2016 a discussion on the theoretical correlations (including the effect of 2D projection), in this section we will analyze the observability of 21cm-galaxy cross-correlations with SKA and galaxy surveys planned with the Subaru HSC and PFS, and with WFIRST.

Once the differential brightness temperature is evaluated within a simulation box using eq. 1, the theoretical 3D 21cm-galaxy cross-power spectrum can be calculated with eq. 2 (see section 3.1 of V2016 for an extensive discussion). Finally, to produce mock observations, the various instrument characteristics need to be taken into account.

To build mock observations with SKA, we model the instrument characteristics using the OSKAR simulator<sup>9</sup> for the preliminary configuration of SKA1-LOW presented in Chapman et al. (2015). Full correlation between all 866 core stations (with a maximum baseline length of 5.29 km) is assumed to produce the images of the point spread function (PSF), which have been generated with the Common

<sup>8</sup> The sphere that is used to evaluate a spherically averaged  $P(k)$  for a simulation box of length  $425 h^{-1} \text{ cMpc}$ , should have an equivalent volume, i.e. a radius  $R = 264 h^{-1} \text{ cMpc}$ . The first zero of a window function for a spherical top-hat is at  $dk \cdot R \sim 4.5$ , meaning that  $k$ -values at distances shorter than  $4.5/R = 0.02 h \text{ Mpc}^{-1}$  are correlated (Feldman, Kaiser & Peacock 1994; Furlanetto & Lidz 2007; Lidz et al. 2009; Wiersma et al. 2013).

<sup>9</sup> <https://bitbucket.org/account/user/oscaremr/projects/OSCAR>



**Figure 1.** *Top panels:* 2D unnormalized by  $\delta T_{b0}$ , circularly averaged 21cm-LAE cross-power spectra between SKA and HSC at  $z = 7.3$  (left panel) and  $6.6$  (right). *Bottom panels:* 21cm-LAE cross-correlation coefficient,  $r_{21,LAE}$ , corresponding to  $\Delta^2_{21,LAE}$ . The field of view is  $1.7 \text{ deg}^2$  and  $7 \text{ deg}^2$  at  $z = 7.3$  and  $z = 6.6$ , respectively. Solid (dashed) lines refer to the cross-power spectrum with (without) SKA noise. Shaded areas indicate scatter from 10 mock observations.

Astronomy Software Applications package (CASA, Brogan et al. 2007) across a  $5 \text{ deg}$  FoV. The coordinates of the baselines have been generated for a 12-hour synthesis observation. The noise is then normalized assuming a 1000 hour integration time following the prescription described in e.g. Thompson, Moran & Swenson (2001).

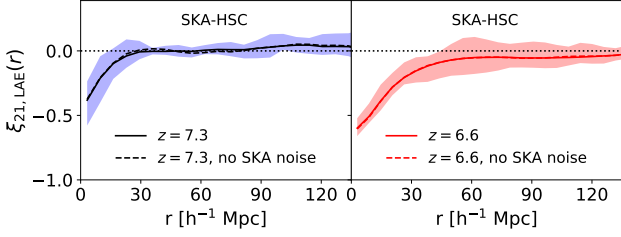
In the following we will investigate the cross-correlation between SKA mock observations thus obtained and various galaxy surveys.

#### 3.1 Subaru Hyper Supreme Cam

In this section we will discuss the observability of the cross-correlation between SKA and HSC.

HSC will observe four fields of  $7 \text{ deg}^2$  at  $z = 6.6$  as part of the deep layer, and four fields of  $1.7 \text{ deg}^2$  (two at  $z = 6.6$  and two at  $z = 7.3$ ) as part of the ultradeep layer. Observations are made with narrow-band filters of  $\Delta z = 0.1$ , corresponding to about one tenth of our simulation length, and thus the galaxies will appear as if lying on a single plane. In this case then the observed cross-power spectrum will be a circularly averaged 2D one. Following the procedure outlined and tested in V2016, to match the size of the HSC's FoV the box dimension has been reduced by removing its external cells. We then select 1375 (20) galaxies at  $z = 6.6$  (7.3) with Ly $\alpha$  equivalent width  $> 20 \text{ \AA}$  and luminosity  $> 2.5 \times 10^{42} \text{ erg s}^{-1}$  to mimic the HSC sensitivity and expected number of observed LAEs per field<sup>10</sup>. Finally, to match the width of the narrow-band filters we have divided the boxes into 10 sub-boxes of 50 slices each, and every sub-box has

<sup>10</sup> Note that, because of the small number of galaxies expected to be observed at  $z = 7.3$ , the shot noise is dominant and the resulting correlations cannot be considered statistically significant.

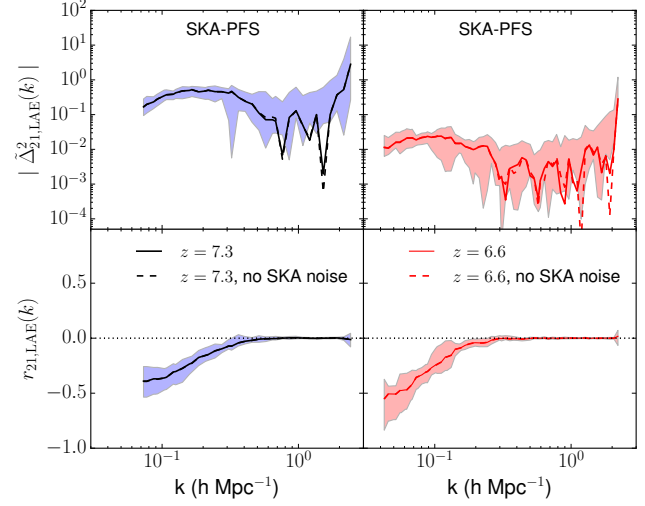


**Figure 2.** 2D 21cm-LAE cross-correlation function between SKA and HSC at  $z = 7.3$  (left panel) and  $6.6$  (right). Solid (dashed) lines refer to the cross-correlation function with (without) SKA noise. The black dotted lines indicate zero correlation and shaded areas indicate scatter from 10 mock observations.

been collapsed onto a single plane to mimic the fact that HSC observations will provide 2D galaxy maps. Each of the 10 HSC mock observation is correlated to the corresponding SKA mock observation, and the resulting 21cm-LAEs cross-power spectra or cross-correlation functions are averaged to get a sample independent result.

The resulting 2D, unnormalized by  $\delta T_{b,0}$ , circularly averaged 21cm-LAE cross-power spectra are shown in Fig. 1, together with the corresponding cross-correlation coefficients, defined as  $r_{21,gal}(k) = P_{21,gal}(k)/[P_{21}(k)P_{gal}(k)]^{1/2}$ . The expected dependence of the intensity of the cross-correlation on redshift is maintained, as well as a clear anti-correlation at the largest scales. Comparing the resulting cross-power spectra with the ones between LOFAR and HSC (Fig. 4 in Vrbanec et al. 2016) it is easy to notice that the spectra observed by SKA and HSC are smoother, more similar to the theoretical ones, and with smaller scatter, so that larger scales (up to  $k < 0.3 h \text{ Mpc}^{-1}$ ) are now accessible and an anti-correlation can be detected also at  $z = 7.3$ . The 21cm-LAE cross-power spectra with and without the SKA noise show very little difference, and the resulting cross-correlation coefficients have an anti-correlation stronger (i.e. closer to the theoretical value) than the one observed with LOFAR and HSC ( $\sim -0.4$  at  $z = 7.3$  and  $\sim -0.6$  at  $z = 6.6$ ). As discussed in V2016, also with SKA as a reference instrument for 21 cm observations, the anti-correlation at  $z = 6.6$  is stronger than at  $z = 7.3$  (differently from the theoretical expectations), as here observations are dominated by shot-noise due to the small number of observed galaxies. Additionally, the smaller field of view at  $z = 7.3$  covers less of the large-scale modes that would boost the anti-correlation strength at large-scales. It should also be noted that the cross-correlation coefficient is more affected by noise than the power spectrum because it contains the contribution from various power spectra.

In Fig. 2 we present the 2D 21cm-LAE cross-correlation functions, which again show a clear anti-correlation on small scales both at  $z = 6.6$  and  $7.3$ . Also in this case, a clear improvement is seen in comparison to results expected from LOFAR (see upper panel of Fig. 8 in V2016), for which the scatter estimated from 10 mock observations is large enough to prevent a clear detection at any redshift and scale.



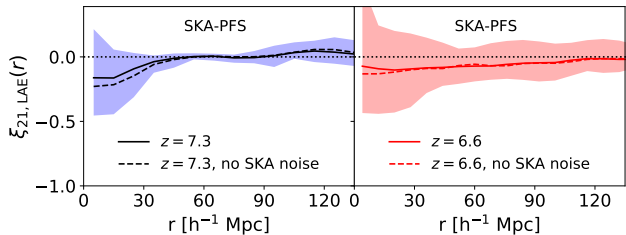
**Figure 3.** *Top panels:* 3D unnormalized by  $\delta T_{b,0}$ , spherically averaged 21cm-LAE cross-power spectra between SKA and PFS at  $z = 7.3$  (left panel) and  $6.6$  (right). *Bottom panels:* 21cm-LAE cross-correlation coefficient,  $r_{21,LAE}$ , corresponding to  $\Delta_{21,LAE}^2$ . The field of view is  $1.7 \text{ deg}^2$  and  $7 \text{ deg}^2$  at  $z = 7.3$  and  $z = 6.6$ , respectively. Solid (dashed) lines refer to the cross-power spectrum with (without) SKA noise. Shaded areas indicate scatter from 10 mock observations.

### 3.2 Subaru Prime Focus Spectrograph

Around 2020 the PFS is expected to be ready for operation. It will focus on the LAEs detected by HSC, but will give the precise position of the objects due to spectroscopic observations. Hence, a 3D instead of 2D cross-correlation can be measured.

As PFS' FoV ( $7 \text{ deg}^2$  at  $z = 6.6$  and  $1.7 \text{ deg}^2$  at  $z = 7.3$ ), depth ( $\Delta z = 0.1$ ), and expected number of observed LAEs per field (1375 at  $z = 6.6$  and 20 at  $z = 7.3$ ) are those of HSC, the 10 sub-boxes obtained from our simulations are the same as those discussed in the previous section, but here the sub-boxes are not collapsed onto a single plane, since PFS will acquire the exact position of the observed LAEs. We thus create 3D mock observations, which are correlated to the corresponding SKA mock observations, resulting in 10 cross-correlations which are then averaged to obtain a sample independent result.

The resulting 3D, unnormalized by  $\delta T_{b,0}$ , spherically averaged 21cm-LAE cross-power spectra are shown in Fig. 3. They look very similar to the ones measured with SKA and HSC (see Fig. 1). However, the strength of the cross-power spectra is now lower due to the spherical average over a 3D volume, where the reduction of strength is higher at  $z = 6.6$  due to the larger observed volume. The cross-power spectra become noisy around the same scale as the cross-power spectra measured with SKA and HSC, since they have similar shot noise (due to the same number of observed LAEs; see section 4). The strength of the anti-correlation at large scales in the cross-correlation coefficient is similar to the one measured with SKA and HSC, i.e.  $\sim -0.4$  at  $z = 7.3$  and  $\sim -0.6$  at  $z = 6.6$ . We should note that the projection effects which enter our calculations when using HSC as reference instrument for LAEs observations do not affect the results. The qualitative behaviour of the various correlations is in



**Figure 4.** 3D 21cm-LAE cross-correlation function between SKA and PFS at  $z = 7.3$  (left panel) and  $6.6$  (right). Solid (dashed) lines refer to the cross-correlation function with (without) SKA noise. The black dotted lines indicate zero correlation and shaded areas indicate scatter from 10 mock observations.

fact the same as in the case of PFS, for which 3D rather than 2D power spectra will be available. However, in 3D the scatter is reduced due to projection effects (see Figs. 2 and 3 of V2016). As for HSC, also in this case, the effect of the SKA noise is minimal.

In Fig. 4 we show the 2D 21cm-LAE cross-correlation functions. Differently from HSC, here the functions are very noisy, due to the higher shot noise (see discussion in Sec. 4) and do not offer any valuable information.

### 3.3 Wide Field Infrared Survey Telescope

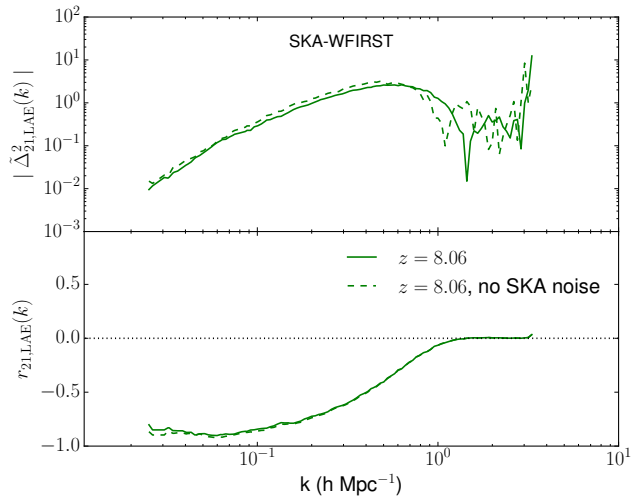
The WFIRST is a space facility expected to observe 900 LAEs per square degree and unit redshift in the range  $7.5 \leq z \leq 8.5$ , with Ly $\alpha$  equivalent width  $>5\text{\AA}$  and luminosity  $> 2.8 \times 10^{42} \text{ erg s}^{-1}$ . Observations will be done spectroscopically over wide areas of the sky. We constructed a mock observation for WFIRST using a FoV corresponding to our full simulation box, i.e.  $16 \text{ deg}^2$ , a depth of  $\Delta z = 1$ , and 14400 observed LAEs, as expected to be detected within this volume. It should be noted that, differently from the HSC and PFS case, here only one mock observation could be designed because the depth of the experiment corresponds to the entire box. The SKA mock observation has been created following the usual procedure, using the same FoV and depth of WFIRST, and it has been correlated to the mock WFIRST observation.

The resulting 3D 21cm-LAE cross-power spectrum and cross-correlation coefficient are shown in Fig. 5. Here the 21cm-LAE cross power spectrum has a shape very similar to the theoretical one (Fig. 2 in V2016), although it becomes noisy on small scales, i.e.  $k > 1 \text{ h Mpc}^{-1}$ . The corresponding cross-correlation coefficient shows strong anti-correlation on large scales ( $r_{21,LAE} \sim -0.9$ ). As in the previous cases, the SKA noise has very little influence on the resulting cross-power spectrum and cross-correlation coefficient.

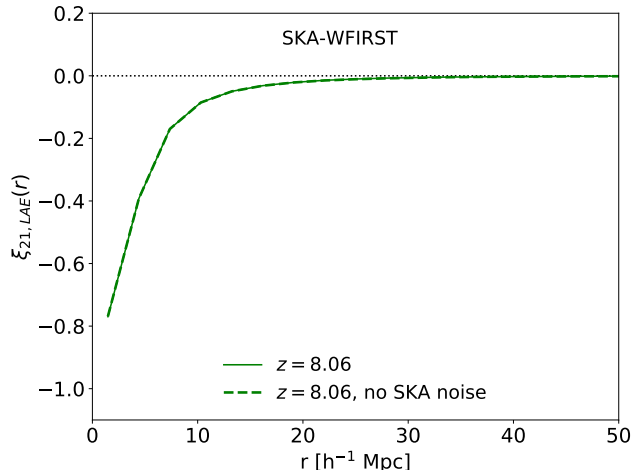
Fig. 6 shows 21cm-LAE cross-correlation function, which suggests that the average scale of ionized bubbles at the redshift in question is  $\sim 30h^{-1} \text{ Mpc}$ . The addition of the SKA noise hardly makes any difference for the cross-correlation function.

## 4 DISCUSSION

The next generation instrument SKA will have much lower noise levels than LOFAR and thus will introduce a signif-

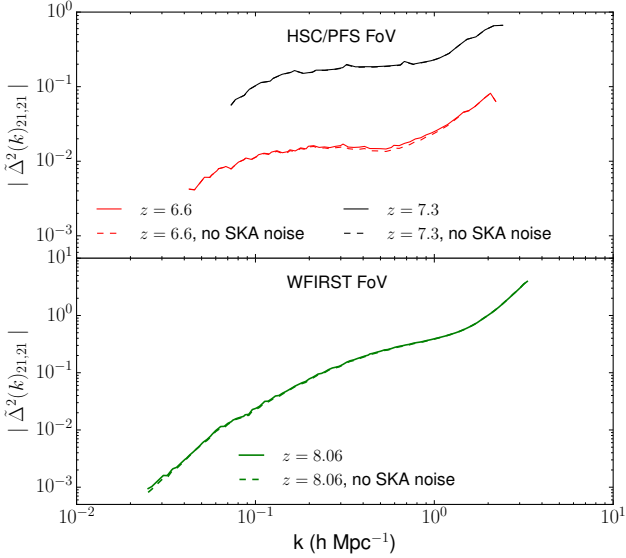


**Figure 5.** *Top panel:* 3D unnormalized by  $\delta T_{b0}$ , spherically averaged 21cm-LAE cross-power spectrum between SKA and WFIRST at  $z = 8.06$ . *Bottom panel:* 21cm-LAE cross-correlation coefficient,  $r_{21,LAE}$ , corresponding to  $\Delta_{21,LAE}^2$ . The field of view is  $16 \text{ deg}^2$ . The solid (dashed) line refers to the cross-power spectrum with (without) SKA noise.



**Figure 6.** 3D 21cm-LAE cross-correlation function between SKA and WFIRST with a FoV of  $16 \text{ deg}^2$  at  $z = 8.06$ . The solid (dashed) line refers to the cross-correlation function with (without) SKA noise. The black dotted line indicates zero correlation.

icant change when used for the observation of 21cm-LAE cross-correlations. The 21cm auto-power spectra with and without SKA noise after 1000 hours of observation in fields of view of  $7 \text{ deg}^2$  at  $z = 6.6$  and  $1.7 \text{ deg}^2$  at  $z = 7.3$  (i.e. equivalent to the ones of HSC and PFS), and in a field of view of  $16 \text{ deg}^2$  at  $z = 8.06$  (i.e. equivalent to the one of WFIRST) are shown in Fig. 7. The difference between results with and without SKA noise is so small it is barely visible, and thus, the main noise contribution to these observations comes from the shot noise due to LAEs observations. The reduction from LOFAR to SKA noise will significantly improve the range of observable scales, from  $k < 0.1 \text{ h Mpc}^{-1}$  to  $k < 0.3 \text{ h Mpc}^{-1}$ , as well as the detectable strength of the

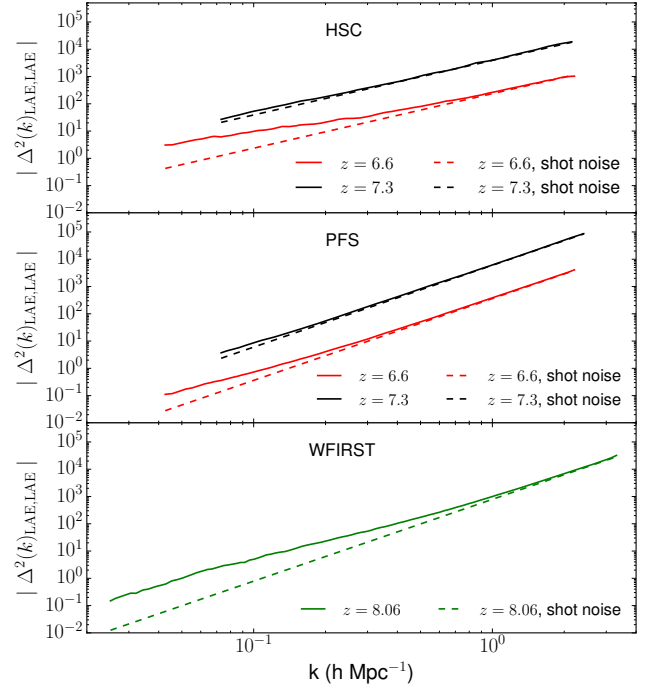


**Figure 7.** *Top panel:* 2D, unnormalised by  $\delta T_{b0}$ , 21cm auto-power spectrum at  $z = 7.3$  (upper set of black lines) and  $z = 6.6$  (lower set of red lines). The field of view is 1.7 deg<sup>2</sup> and 7 deg<sup>2</sup> at  $z = 7.3$  and  $z = 6.6$ , respectively (as HSC and PFS). Solid (dashed) lines refer to the spectrum with (without) SKA noise. *Bottom panel:* 3D, unnormalised by  $\delta T_{b0}$ , 21cm auto-power spectrum with (solid line) and without (dashed) SKA noise at  $z = 8.06$ . The field of view is 16 deg<sup>2</sup> (as WFIRST).

anti-correlation on the large scales of the cross-correlation coefficient, up to  $\sim -0.4$  at  $z = 7.3$  and  $\sim -0.6$  at  $z = 6.6$ .

PFS will introduce significant change compared to HSC by making precise spectroscopic observations of previously detected LAEs, which will enable more precise computations of their positions, thus allowing 3D cross-correlations. However, this important change in LAE information will not result in a significant improvement of the measured cross-correlations, as the same number of LAEs will describe the observed 3D volume poorer than the 2D surface, increasing the shot noise significantly. Thus, the additional information about positions of LAEs and the increase of the shot noise will mostly cancel each other out. As shown in Fig. 8, the shot noise dominates over the LAE auto-power spectrum more significantly in observations with PFS rather than with HSC. At  $z = 7.3$  with PFS the LAE auto-power spectrum starts to dominate over the shot noise at the same scale as with HSC ( $k \sim 0.2$  h Mpc<sup>-1</sup>). However, in both cases this dominance is hardly insignificant. At  $z = 6.6$  the LAE auto-power spectrum starts to be relevant on larger scales ( $k \sim 0.3$  h Mpc<sup>-1</sup> for PFS and  $k \sim 0.7$  h Mpc<sup>-1</sup> for HSC), but the difference with the shot noise in PFS' observations reaches only half an order of magnitude. Thus, even though in this case there are no projection effects, the shape of the cross-power spectra and the strength of the cross-correlation coefficient do not change significantly due to the simultaneous increase of the shot noise.

In the bottom panel of Fig. 8 we show the LAE auto-power spectrum and shot noise power spectrum for WFIRST observations at  $z = 8.06$ . The LAE auto-power spectrum dominates over the shot noise power spectrum on scales larger than  $k \sim 1$  h Mpc<sup>-1</sup>, becoming also an order of magnitude larger. We thus expect that the low noise of both



**Figure 8.** *Top panel:* HSC 2D LAE auto-power spectra at  $z = 7.3$  (black solid line) and  $z = 6.6$  (red solid), and shot noise power spectra at  $z = 7.3$  (black dashed) and  $z = 6.6$  (red dashed). The field of view is 1.7 deg<sup>2</sup> and 7 deg<sup>2</sup> at  $z = 7.3$  and  $z = 6.6$ , respectively. *Middle panel:* PFS 3D LAE auto-power spectra at  $z = 7.3$  (black solid line) and  $z = 6.6$  (red solid), and shot noise power spectra at  $z = 7.3$  (black dashed) and  $z = 6.6$  (red dashed). The field of view is 1.7 deg<sup>2</sup> and 7 deg<sup>2</sup> at  $z = 7.3$  and  $z = 6.6$ , respectively. *Bottom panel:* WFIRST 3D LAE auto-power spectrum (green solid line) and shot noise power spectrum (green dashed) at  $z = 8.06$ . The field of view is 16 deg<sup>2</sup>.

WFIRST and SKA observations should allow a precise measurements on scales larger than  $k \sim 1$  h Mpc<sup>-1</sup>.

While investigations of the 21 cm-LAEs cross-correlation appeared in the literature give results which are qualitatively similar, a quantitative comparison is somewhat misleading, as the methods and models employed are very different. As discussed in previous literature, the quantitative results depend primarily on the spatial distribution and dimension of HII regions, as well as the population of LAEs. As a consequence, different reionization histories will yield different results, with the choice of the minimum mass of star hosting halos and of how halos are populated with LAEs having a major impact. Also the method used to compute the mock observations and the cross-correlations (e.g. by integrating the power spectra rather than directly) are expected to introduce further differences. For these reasons, our results are more in line to those of Hutter et al. (2017) and Kubota et al. (2018), who employ simulations of reionization more similar to ours, compared e.g. to Sobacchi, Mesinger & Greig (2016) who instead use a semi-numeric approach.

## 5 CONCLUSIONS

In this paper we have used radiative transfer + N-body simulations to explore the feasibility of measurements of cross-correlations between the 21 cm field observed by the Square Kilometer Array (SKA) and high- $z$  Lyman Alpha Emitters (LAEs) observed in galaxy surveys with the Subaru Hyper Supreme Cam (HSC), Subaru Prime Focus Spectrograph (PFS) and Wide Field Infrared Survey Telescope (WFIRST). We have investigated both the 21cm-LAE cross-power spectra and cross-correlation functions since the noise affect them differently. For this study we have closely followed a companion paper based on observations with LO-FAR (V2016).

The next generation observations with SKA will have a noise level much lower than those with LOFAR, introducing a significant improvement in the measurement of the cross-correlations. Compared to V2016, we find that an SKA-HSC observation will extend the range of detectable scales from  $k < 0.1 h \text{ Mpc}^{-1}$  to  $k < 0.3 h \text{ Mpc}^{-1}$ , as well as increase the strength of the anti-correlation on large scales up to  $\sim -0.4$  at  $z = 7.3$  and  $\sim -0.6$  at  $z = 6.6$ , compared to  $\sim -0.25$  and  $\sim -0.3$  respectively. The cross-correlation functions are much smoother and have a smaller scatter, so that they can be used to investigate scales above  $\sim 10 h^{-1} \text{ Mpc}$  and  $\sim 60 h^{-1} \text{ Mpc}$  at  $z = 7.3$  and  $6.6$ , respectively.

PFS will introduce an important improvement to Subaru's measurements by making precise spectroscopic observations of LAEs previously detected with HSC. This will enable a more accurate computation of their positions, thus allowing 3D cross-correlations. However, this additional information will not result in a significant improvement of the measured cross-correlations as the same number of LAEs will describe the observed 3D volume poorer than the 2D surface, increasing the shot noise significantly. Thus, the information gained on the position of LAEs and the increase of the shot noise will mostly cancel each other out.

Unlike Subaru's observations, which will mostly give insight into cross-correlations towards the end of reionization, WFIRST will allow access to also higher redshifts. Specifically, WFIRST is expected to observe spectroscopically about 900 LAEs per square degree and unit redshift in the range  $7.5 \leq z \leq 8.5$ . As a reference, we have computed the 21cm-LAE cross-correlations at  $z = 8.06$ . As the SKA noise at this redshift is still low and the expected detected LAE density is rather large, the resulting cross-correlations are very similar to the theoretical ones, and shot noise becomes a problem only above  $k \sim 0.6 h \text{ Mpc}^{-1}$ . The anti-correlation at large scales reaches a strength of  $\sim -0.8$ , and the cross-correlation function suggests a typical scale for the ionized regions of  $\sim 30 h^{-1} \text{ Mpc}$ . In the future, we are planning to investigate how accurately this estimate reproduces the size of the HII regions in the simulations, and its dependence on the method used to quantify it.

In summary, 21cm-LAE cross-correlations are a powerful probe of the EoR and could provide precious information on the progress of reionization and the typical dimension of ionized regions at various redshifts. As 21 cm observations with SKA will be affected by a very small noise, cross-correlations with future galaxy surveys will be feasible in a range of different redshifts. At the same time, WFIRST is expected to observe a large number density of LAEs, reduc-

ing the shot noise compared to HSC and PFS. Both effects will result in more precise cross-correlations and increasing observable scales, offering information at various redshifts on e.g. the typical scale of ionized regions.

## ACKNOWLEDGEMENTS

The authors would like to thank an anonymous referee for her/his useful comments. The authors acknowledge Paul Shapiro for permission to use the simulations on which this paper was based, described in Iliev et al. (2014). That work was supported in part by grants and allocations of which Shapiro is the P.I., including US NSF grant AST-1009799, NASA grant NNX11AE09G, NASA/JPL grant RSA Nos 1492788 and 1515294, and supercomputer resources from NSF XSEDE grant TG-AST090005 and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. Some of the numerical computations were done on the Apollo cluster at The University of Sussex and the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet and the University of Portsmouth. Part of the computations were performed on the GPC supercomputer at the SciNet HPC Consortium (courtesy Ue-Li Pen). SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund – Research Excellence; and the University of Toronto. The authors thank Kyungjin Ahn for providing the recipe for including subresolution sources in the simulation volume.

## REFERENCES

- Baek S., Di Matteo P., Semelin B., Combes F., Revaz Y., 2009, *A&A*, 495, 389
- Brogan C. L., Fomalont E., McMullin J., Myers S., Radziwill N., Rupen M., 2007, *National Radio Astronomy Observatory Newsletter*, 113, 17
- Carilli C. L., Gnedin N. Y., Owen F., 2002, *ApJ*, 577, 22
- Chapman E. et al., 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 5
- Ciardi B. et al., 2013, *MNRAS*, 428, 1755
- Ciardi B., Madau P., 2003, *ApJ*, 596, 1
- Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23
- Fernandez E. R., Zaroubi S., Iliev I. T., Mellema G., Jelić V., 2014, *MNRAS*, 440, 298
- Field G. B., 1959, *ApJ*, 129, 536
- Furlanetto S. R., 2006, *MNRAS*, 370, 1867
- Furlanetto S. R., Lidz A., 2007, *ApJ*, 660, 1030
- Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *PhysRep*, 433, 181
- Hutter A., Dayal P., Müller V., Trott C. M., 2017, *ApJ*, 836, 176
- Hutter A., Trott C. M., Dayal P., 2018, *MNRAS*, 479, L129
- Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, *MNRAS*, 439, 725
- Jelić V., Zaroubi S., Aghanim N., Douspis M., Koopmans L. V. E., 2010, *MNRAS*, 402, 2279
- Jensen H., Hayes M., Iliev I. T., Laursen P., Mellema G., Zackrisson E., 2014, *MNRAS*, 444, 2114

- Jensen H., Laursen P., Mellema G., Iliev I. T., Sommer-Larsen J., Shapiro P. R., 2013, *MNRAS*, 428, 1366
- Kubota K., Yoshiura S., Takahashi K., Hasegawa K., Yajima H., Ouchi M., Pindor B., Webster R. L., 2018, *MNRAS*, 479, 2754
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Lidz A., Zahn O., Furlanetto S. R., McQuinn M., Hernquist L., Zaldarriaga M., 2009, *ApJ*, 690, 252
- Ma Q., Ciardi B., Eide M. B., Helgason K., 2018a, *MNRAS*, 480, 26
- Ma Q., Helgason K., Komatsu E., Ciardi B., Ferrara A., 2018b, *MNRAS*, 476, 4025
- Madau P., Meiksin A., Rees M. J., 1997, *ApJ*, 475, 429
- Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006a, *New Astronomy*, 11, 374
- Mellema G., Iliev I. T., Pen U., Shapiro P. R., 2006, *MNRAS*, 372, 679
- Patil A. H. et al., 2017, *ApJ*, 838, 65
- Patil A. H. et al., 2014, *ArXiv e-prints*
- Pritchard J. R., Loeb A., 2012, *Reports on Progress in Physics*, 75, 086901
- Salvaterra R., Ciardi B., Ferrara A., Baccigalupi C., 2005, *MNRAS*, 360, 1063
- Semelin B., 2015, *ArXiv e-prints*
- Sobacchi E., Mesinger A., Greig B., 2016, *MNRAS*, 459, 2741
- Thompson A. R., Moran J. M., Swenson, Jr. G. W., 2001, *Interferometry and Synthesis in Radio Astronomy*, 2nd Edition
- Tozzi P., Madau P., Meiksin A., Rees M. J., 2000, *ApJ*, 528, 597
- Vrbanec D. et al., 2016, *MNRAS*, 457, 666
- Wiersma R. P. C., Ciardi B., Thomas R. M., Harker G. J. A., 2013, *MNRAS*, 432, 2615
- Xu Y., Chen X., Fan Z., Trac H., Cen R., 2009, *ApJ*, 704, 1396
- Zaroubi S. et al., 2012, *MNRAS*, 425, 2964

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