



# Prospects for 21 cm Galaxy Cross-correlations with HERA and the Roman High-latitude Survey

Paul La Plante<sup>1,2,3,4</sup>, Jordan Mirocha<sup>5</sup>, Adélie Gorce<sup>5</sup>, Adam Lidz<sup>6</sup>, and Aaron Parsons<sup>1</sup><sup>1</sup>Department of Astronomy, University of California, Berkeley, CA 94720, USA; [paul.laplante@unlv.edu](mailto:paul.laplante@unlv.edu)<sup>2</sup>Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA<sup>3</sup>Department of Computer Science, University of Nevada, Las Vegas, NV 89154, USA<sup>4</sup>Nevada Center for Astrophysics, University of Nevada, Las Vegas, NV 89154, USA<sup>5</sup>Department of Physics and McGill Space Science Institute, McGill University, Montreal, Canada<sup>6</sup>Center for Particle Cosmology, Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

Received 2022 May 25; revised 2022 December 21; accepted 2022 December 22; published 2023 February 13

## Abstract

The cross-correlation between the 21 cm field and the galaxy distribution is a potential probe of the Epoch of Reionization (EoR). The 21 cm signal traces neutral gas in the intergalactic medium and, on large spatial scales, this should be anticorrelated with the high-redshift galaxy distribution, which partly sources and tracks the ionized gas. In the near future, interferometers such as the Hydrogen Epoch of Reionization Array (HERA) are projected to provide extremely sensitive measurements of the 21 cm power spectrum. At the same time, the Nancy Grace Roman Space Telescope (Roman) will produce the most extensive catalog to date of bright galaxies from the EoR. Using seminumeric simulations of reionization, we explore the prospects for measuring the cross-power spectrum between the 21 cm and galaxy fields during the EoR. We forecast a  $12\sigma$  detection between HERA and Roman, assuming an overlapping survey area of  $500 \text{ deg}^2$ , redshift uncertainties of  $\sigma_z = 0.01$  (as expected for the high-latitude spectroscopic survey of Ly $\alpha$ -emitting galaxies), and an effective Ly $\alpha$  emitter duty cycle of  $f_{\text{LAE}} = 0.1$ . Thus the HERA–Roman cross-power spectrum may be used to help verify 21 cm detections from HERA. We find that the shot-noise in the galaxy distribution is a limiting factor for detection, and so supplemental observations using Roman should prioritize deeper observations, rather than covering a wider field of view. We have made a public GitHub repository containing key parts of the calculation, which accompanies this paper: [https://github.com/plaplant/21cm\\_gal\\_cross\\_correlation](https://github.com/plaplant/21cm_gal_cross_correlation).

*Unified Astronomy Thesaurus concepts:* Cosmology (343); Large-scale structure of the universe (902); Reionization (1383)

## 1. Introduction

The Epoch of Reionization (EoR) is one of the last remaining frontiers in observational cosmology. The EoR is the time period when the first luminous sources in the universe formed and gradually photoionized neutral hydrogen in the surrounding intergalactic medium (IGM). This is thought to occur roughly 0.5–1 Gyr after the Big Bang. This large-scale transition of the universe has yet to be fully observed: in particular, the redshift evolution of the average ionization fraction,  $x_i(z)$ , and the overall topology of the reionization process remain highly uncertain. Measurements of the ionization history and its spatial fluctuations would help pin-down the properties of early star-forming galaxies, as well as cosmological parameters such as the optical depth of cosmic microwave background (CMB) photons to electron scattering,  $\tau$ . As such, providing measurements of the EoR has important ramifications for multiple fields in astrophysics and cosmology.

Observationally, radio interferometers are attempting to measure the 21 cm signal of neutral hydrogen (Madau et al. 1997). The hyperfine transition of neutral hydrogen atoms emits and absorbs radiation with a rest wavelength of 21 cm, which can imprint an excess or deficit in brightness relative to the CMB backlight. During the EoR, the 21 cm signal is

expected to have significant fluctuations due to the presence of large ionized regions, because once ionized, the hydrogen in the IGM no longer emits a 21 cm signal. While many experiments seek the 21 cm power spectrum (e.g., the Low Frequency Array; van Haarlem et al. 2013; the Murchison Widefield Array, MWA; Bowman et al. 2013; and the Owens Valley Long Wavelength Array; Hallinan et al. 2015), in this work we focus on the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al. 2017), currently under construction in the Karoo desert of South Africa. Once completed, HERA will be the most sensitive instrument to date measuring the 21 cm signal, and is expected to provide the first statistically significant detection of the 21 cm power spectrum (DeBoer et al. 2017).

At the same time, wide-field infrared telescopes are planning to observe a large ensemble of high-redshift galaxies directly, e.g., Euclid<sup>7</sup> and the Nancy Grace Roman Space Telescope (Roman; Spergel et al. 2015). In this work, we focus in particular on Roman, expected to launch in 2026, as its wide-field surveys are deeper than the nominal Euclid survey. Among other scientific goals, Roman is slated to observe  $2200 \text{ deg}^2$  of the sky with both imaging and spectroscopy, the so-called High Latitude Survey (HLS), with imaging and spectroscopic campaigns referred to as the HLIS and HLSS, respectively. Although this survey is designed to detect galaxies with redshift  $1 \lesssim z \lesssim 3$ , it can also detect galaxies at



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

<sup>7</sup> <https://www.cosmos.esa.int/web/euclid/euclid-survey>

redshifts of  $z \gtrsim 5$  via the Lyman-break technique (Steidel et al. 1996). These high-redshift galaxies include some of the sources thought to be responsible for ionizing the universe during the EoR. In principle, the galaxy field measured in such a wide-field survey and the 21 cm signal should be statistically anticorrelated on large spatial scales during the EoR, at least in the expected case that large-scale overdensities are ionized first, i.e., in an “inside-out” reionization scenario. This anticorrelation is due to the galaxy field tracing highly biased regions of cosmic structure, whereas the 21 cm signal comes predominantly from lower-density, neutral regions. As such, the 21 cm and galaxy fields should be anticorrelated, at least on spatial scales that are large relative to the typical size of ionized bubbles.

The cross-power spectrum of the 21 cm and galaxy fields is also interesting in principle as a means to validate measurements made of the 21 cm auto-power spectrum from HERA. Detecting the cross-power spectrum between the 21 cm and an independent tracer will provide important evidence that should confirm inferences made from HERA measurements alone, such as information about the size distribution of ionized bubbles and their evolution with redshift. Furthermore, at low redshift the only 21 cm fluctuation measurements made to date have been in cross-correlation with galaxy and quasar catalogs, rather than as a 21 cm auto-power spectrum. For example, the recent results from the Canadian Hydrogen Intensity Mapping Experiment (CHIME<sup>8</sup>) report a detection of the cross-power spectrum between the 21 cm signal and galaxies and quasars from eBOSS (CHIME Collaboration et al. 2022). Additionally, HI intensity maps from  $z \sim 1$  made by the Green Bank Telescope have been used in cross-correlation with galaxy surveys to measure the hydrogen abundance and bias parameters (Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013; Wolz et al. 2022). As such, modeling and measuring the 21 cm galaxy cross-power spectrum at high redshift is useful, independent of the 21 cm auto-power spectrum.

Previous studies (Furlanetto & Lidz 2007; Wyithe & Loeb 2007; Lidz et al. 2009; Vrbancic et al. 2020) have looked at similar prospects for cross-correlation between 21 cm experiments and wide-field galaxy surveys. We move past this previous work by making cross-spectrum forecasts for the upcoming HERA and Roman data for the first time. Importantly, we also explicitly account for the “foreground wedge” (Datta et al. 2010; Pober et al. 2013; Parsons et al. 2014). That is, previous studies accounted only for foregrounds preventing measurements of Fourier modes with long-wavelength line-of-sight components, i.e., low- $k_{\parallel}$  modes. In fact, the frequency dependence of the instrumental response of an interferometer leads to mode-mixing, and this corrupts some high  $k_{\parallel}$  modes as well. Nevertheless, this contamination is expected mostly to occupy a wedge-shaped region in the  $k_{\parallel}$ - $k_{\perp}$  plane. This further degrades the prospects for cross-power spectrum measurements, as discussed in this work. Furthermore, we include a detailed treatment of the Roman observations. Specifically, we account for the nominal magnitude and flux limits for the HLIS and HLSS, a complete handling of redshift uncertainties, and projections for the joint overlap area between HERA and Roman, now that such information is available (Spergel et al. 2015; Doré et al. 2018;

Wang et al. 2022). Finally, we model how the cross-correlation signal varies as a function of ionization history and galaxy properties. Although these conclusions may be partly tied to the particular seminumeric simulation of reionization employed, these model variations nevertheless have important implications for the detectability and interpretation of the cross-spectrum signal.

We organize the rest of this paper as follows. In Section 2, we described the seminumeric simulations used for this study. In Section 3, we present the primary results of our work. In Section 4, we discuss the feasibility of detection for upcoming 21 cm surveys. In Section 5, we explore ways that observations can improve on the fiducial measurement strategies. Finally, in Section 6, we provide a summary and avenues for future research. Throughout this work, we assume a  $\Lambda$  cold dark matter cosmology with parameters consistent with the Planck 2018 results (Planck Collaboration et al. 2020).

## 2. Methods

In this section, we describe the simulations used to explore the EoR and to model the cross-correlation power spectrum between the 21 cm and galaxy fields. We expect the 21 cm and galaxy fields to be well-correlated on relatively large scales ( $\gtrsim 1 h^{-1}$  Mpc) based on previous work (Lidz et al. 2009). Thus we opt to simulate a large volume with moderate resolution, as this helps capture features of the field on the scales probed by upcoming observations. We begin by describing the seminumeric reionization method that we use, followed by our galaxy modeling. We also include a discussion of the various observational systematics present for both the 21 cm measurements as well as the galaxy field.

### 2.1. 21 cm Modeling

Accurately simulating the EoR is a computationally difficult problem. The complex interplay between dark matter, baryons, and photons necessitates employing  $N$ -body methods, hydrodynamics, and radiative transfer to capture the variety of physical effects self-consistently. Furthermore, the formation of luminous objects on subkiloparsec scales has implications for the ionization state of the IGM on scales of tens of megaparsecs, which requires a tremendous amount of dynamic range in the simulation. For now, even state-of-the-art simulation packages are incapable of handling such a tremendous workload. Furthermore, for the present study, we are most interested in the behavior of the IGM on relatively large scales, and so many of the small-scale details inside of individual galaxies are not relevant. Thus, we opt to use a seminumeric scheme for simulating reionization, which captures the main features in a reliable fashion while remaining relatively cheap computationally to allow for an exploration of the parameter space of various ionization histories. Specifically, we make use of the *zreion* seminumeric reionization code (Battaglia et al. 2013), which has previously been applied to modeling the 21 cm signal in La Plante et al. (2014), La Plante & Ntampaka (2019), and La Plante et al. (2020).

The central *Ansatz* of *zreion* is that the matter density field  $\delta_m(\mathbf{r})$  and the redshift at which a particular portion of the IGM is reionized  $z_{\text{re}}(\mathbf{r})$  are correlated on large scales. We begin by

<sup>8</sup> <https://chime-experiment.ca/en>

defining the matter overdensity field  $\delta_m(\mathbf{r})$  as:

$$\delta_m(\mathbf{r}) \equiv \frac{\rho_m(\mathbf{r}) - \bar{\rho}_m}{\bar{\rho}_m}, \quad (1)$$

where  $\bar{\rho}_m$  is the mean matter density, and the equivalent “overdensity” field for the redshift of reionization  $\delta_z(\mathbf{r})$  is

$$\delta_z(\mathbf{r}) \equiv \frac{[z_{\text{re}}(\mathbf{r}) + 1] - [\bar{z} + 1]}{\bar{z} + 1}, \quad (2)$$

where  $\bar{z}$  is the mean value for the  $z_{\text{re}}(\mathbf{r})$  field. The `zreion` method expresses the correlation between  $\delta_m(\mathbf{r})$  and  $\delta_z(\mathbf{r})$  as a scale-dependent bias factor  $b_{zm}(k)$ , which is defined as:

$$b_{zm}^2(k) \equiv \frac{\langle \delta_z^* \delta_z \rangle_k}{\langle \delta_m^* \delta_m \rangle_k} = \frac{P_{zz}(k)}{P_{mm}(k)}. \quad (3)$$

We parameterize this bias factor using two parameters,  $k_0$  and  $\alpha$ , and express the bias as a function of Fourier wavenumber  $k$ :

$$b_{zm}(k) = \frac{b_0}{\left(1 + \frac{k}{k_0}\right)^\alpha}. \quad (4)$$

We use the value of  $b_0 = 1/\delta_c = 0.593$ , where  $\delta_c$  is the critical overdensity in spherical collapse halo models. Given a cosmological density field, `zreion` relies only on the value of the parameters  $\{\bar{z}, k_0, \alpha\}$  to determine the redshift of reionization  $z_{\text{re}}(\mathbf{x})$ . The midpoint of reionization—defined as the time at which the universe is 50% ionized by volume—is largely determined by the mean value  $\bar{z}$  (though it is not identically equal to the parameter), and the duration is controlled by  $k_0$  and  $\alpha$ .

For the current work, we run a series of dark-matter-only simulations that contain  $1024^3$  particles in a cubic volume with length  $L = 2 h^{-1} \text{ Gpc}$  on a side, which corresponds to an angular extent of  $\theta \approx 18 \text{ deg}$  at  $z = 8$ . This is about twice the extent of the instantaneous field of view (FOV) of HERA (DeBoer et al. 2017). We first generate a set of initial conditions based off of transfer functions obtained from CAMB (Lewis et al. 2000). Given these initial conditions, we use second-order Lagrangian perturbation theory (2LPT) to evolve the particle positions as a function of redshift. Although 2LPT does not capture the nonlinear evolution of particles to the same extent as  $N$ -body methods, the results are sufficiently accurate on the scales of interest (Scoccimarro 1998). To generate the ionization field, we use 2LPT to evolve the particles to the midpoint of reionization  $\bar{z}$ . To compute the matter density field  $\delta_m(\mathbf{x})$ , we use triangular shaped clouds to deposit the particles in the simulation onto a regular cubic lattice. With the density field in hand, we apply a fast Fourier transform (FFT) and apply the bias from Equation (4) to  $\delta_m(\mathbf{k})$  to compute  $\delta_z(\mathbf{k})$ . After applying an inverse FFT, we use Equation (2) to compute  $z_{\text{re}}(\mathbf{x})$ . To convert from this quantity at a particular redshift  $z_0$  into an ionization field  $x_i(\mathbf{x}, z_0)$ , we set all values of the ionization field to have a value of 1 where  $z_{\text{re}}(\mathbf{x})$  is greater than  $z_0$  (meaning that portion of the volume was reionized at an earlier time), and 0 otherwise.

Once we have the ionization and density fields, we can convert to the 21 cm brightness temperature by using (Madau et al. 1997):

$$\delta T_b(\mathbf{r}, z) = T_0(z)[1 + \delta_m(\mathbf{r})][1 - x_i(\mathbf{r})], \quad (5)$$

where the quantity  $T_0(z)$  is

$$T_0(z) = 26 \left( \frac{T_S - T_\gamma}{T_S} \right) \left( \frac{\Omega_b h^2}{0.022} \right) \times \left[ \left( \frac{0.143}{\Omega_m h^2} \right) \left( \frac{1+z}{10} \right) \right]^{\frac{1}{2}}. \quad (6)$$

In this equation,  $T_S$  is the spin temperature of neutral hydrogen and  $T_\gamma$  is the temperature of the CMB. It is common in the literature to assume the spin temperature to be coupled to the temperature of the gas  $T_{\text{gas}}$ . This is thought to occur via X-ray photoheating, perhaps driven by X-ray binaries in early galaxies, such that the gas becomes substantially hotter than  $T_\gamma$  before it is significantly ionized (Furlanetto et al. 2006). For contrast, it is also useful to consider a “cold reionization” scenario in which there is no such heating. In this case, we suppose that  $T_{\text{gas}}$  cools adiabatically between when the gas thermally decouples from the CMB (near  $z \sim 200$ ) and reionization. This makes  $T_{\text{gas}}$  smaller than  $T_\gamma$ , so the brightness temperature in Equation (5) is negative and has a larger amplitude than in our fiducial scenario. We consider this case in Section 4.1.

### 2.1.1. Changing the Ionization History

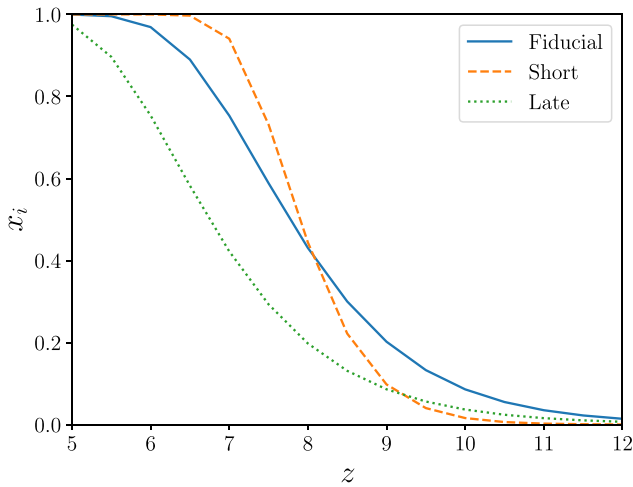
To explore the extent to which these results depend on the precise ionization history, we vary the parameters of our seminumeric reionization model described in Section 2.1 to produce alternate plausible histories. In addition to our “fiducial” scenario, we run a series of simulations that have a shorter reionization history with a comparable midpoint (the “short” scenario). We also run a simulation that has a slightly later midpoint of reionization (the “late” scenario),<sup>9</sup> which can improve our understanding of how the results are impacted by the midpoint of reionization occurring at redshift values that are not covered by the Roman grism (as mentioned above in Section 2.4.2).

Figure 1 shows the ionization fraction as a function of redshift for the different ionization histories described. In all cases, we use the same values of the galaxy bias parameters described below in Section 2.2. As discussed more there, we do not include any explicit connection between the ionization field and the galaxy field besides the fact that both are treated as biased tracers of the matter density field. However, given that we are primarily interested in relatively bright and biased galaxies in the HLS observations on large scales, we do not expect this simplification to be a significant source of error.

In the results below in Sections 3 and 4, we present results for all three of the ionization histories. In general, the amplitude of the cross-spectrum increases as the duration of reionization decreases. This is due to a larger amplitude of the 21 cm fluctuations on large scales for these shorter scenarios, which is a general feature of the seminumeric model used (and explored more in La Plante et al. 2014).

<sup>9</sup> The `zreion` parameters for the fiducial scenario are:  $\bar{z} = 8$ ,  $\alpha = 0.2$ , and  $k_0 = 0.9 h \text{ Mpc}^{-1}$ . The short scenario parameters are:  $\bar{z} = 8$ ,  $\alpha = 0.564$ , and  $k_0 = 0.185 h \text{ Mpc}^{-1}$ . The late scenario parameters are the same as the fiducial scenario, but with  $\bar{z} = 7$ .





**Figure 1.** The global ionization fraction  $x_i$  as a function of redshift  $z$ . We show this quantity for the three scenarios explored in this work. As mentioned in Section 2.1.1, these different histories help show the extent to which our forecasts are affected by details specific to one particular history.

## 2.2. Galaxy Modeling

Like reionization, galaxy formation is a rich and complicated physical process that involves the interactions between dark matter, baryons, and radiation over many decades in length scale. As a result, self-consistent simulations of galaxy formation and reionization—especially in the large volumes relevant to cross-correlations—remain exceedingly challenging (Iliev et al. 2006; Trac & Cen 2007; O’Shea et al. 2015; Ocvirk et al. 2016; Lewis et al. 2022). Given that we are most interested in the large-scale correlation between the 21 cm and galaxy fields, we use a linear bias model to rapidly generate a galaxy abundance field rather than simulating it from first principles, though the linear bias values we adopt are themselves based on semianalytic and hydrodynamical models.

We first describe our model for star formation in high- $z$  galaxies in Section 2.2.1 and the resulting predictions for the galaxy bias and abundance. Then, we describe our approach to Ly $\alpha$  emission in Section 2.2.2, with an emphasis on the likelihood that galaxies detected in the Roman high-latitude imaging survey (HLIS) are also detected in the high-latitude spectroscopic survey (HLSS).

### 2.2.1. Galaxy Properties

The bias of galaxies—one of two key inputs to our model—can be readily computed from simulations using an expression analogous to Equation (3), but using the galaxy overdensity field in place of the reionization redshift field. This is precisely what has been done for the BLUETIDES simulations (Feng et al. 2014; Waters et al. 2016). In this work, we will employ both the BLUETIDES predictions for the galaxy bias, as well as an efficient semiempirical model implemented in ARES.<sup>10</sup> Both models have been designed to match the high-redshift rest-UV luminosity functions (UVLFs), and so are in good agreement with current data sets by construction. ARES allows us to explore potential extensions to the HLS, for which galaxy bias predictions from simulations are not readily available. We summarize the most pertinent aspects of this model here briefly,

and refer the interested reader to Mirocha et al. (2017) and Mirocha et al. (2020) for more details.

Our basic approach is similar to many semiempirical models put forth in recent years. We assume that the star formation rate (SFR) in galaxies is driven by halo growth,  $\dot{M}_* = f_*(M_h)\dot{M}_h$ , where  $f_*$  is a halo mass-dependent star formation efficiency. We derive halo mass accretion rates (MARs) under the assumption that halos evolve at fixed number density (see Furlanetto et al. 2017), which provides a good match to MARs derived from numerical simulations (Mirocha et al. 2021). We assume  $f_*$  is a double power law, and calibrate its free parameters by jointly fitting to high- $z$  UVLFs and UV colors ( $\beta$ ) from Bouwens et al. (2015) and Bouwens et al. (2014). Two key assumptions remain. First, we adopt the BPASS version 1 single-star models (Eldridge & Stanway 2009) with a stellar metallicity of  $Z = 0.004$ .<sup>11</sup> Second, whereas many models in the literature neglect dust or adopt empirical relationships between dust attenuation and UV color, we self-consistently forward model the full rest-UV spectrum of each model galaxy, with additional free parameters governing the dust production efficiency and scale length allowed to vary as well (see Mirocha et al. 2020). This results in slightly different relationships between, e.g.,  $M_{UV}$  and the UV extinction  $A_{UV}$  than are predicted from the relationship between infrared excess and  $\beta$  at lower redshifts, such as that of Meurer et al. (1999).

For semiempirical models like this, for which there is no simulation box, one must derive the galaxy bias as a weighted integral over the halo mass function,  $dn/dm$ , and halo bias,  $b_h$ ,

$$b_g(\geq m_{\min}; z) = \frac{\int_{m_{\min}}^{\infty} b_h(m; z) \frac{dn}{dm} dm}{\int_{m_{\min}}^{\infty} \frac{dn}{dm} dm} \quad (7)$$

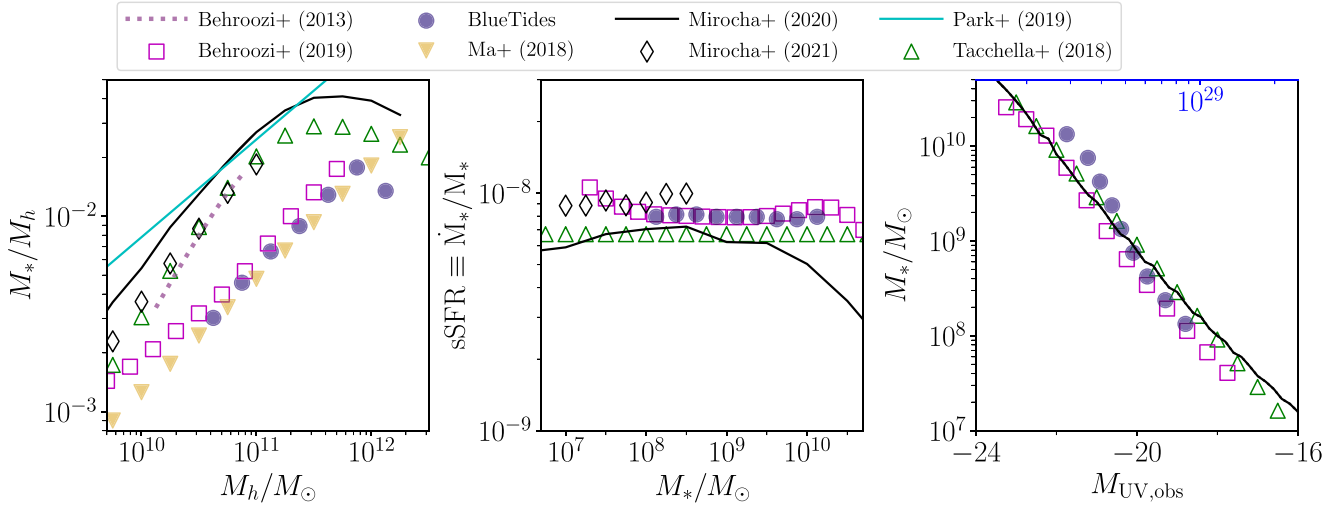
where the minimum mass is related to the limiting magnitude of the survey,  $m_{\min} = m_{\min}(m_{AB, \text{lim}})$ . We use a Tinker et al. (2008) mass function throughout, and adopt their fitting formula for the halo bias as well (Tinker et al. 2010).

In Figure 2, we compare the basic properties of galaxies in the ARES semiempirical model (solid black) to many others from the literature.<sup>12</sup> The most striking differences among this set of models occur in the stellar mass–halo mass (SMHM) relation (left), while the specific star formation rate (sSFR) of galaxies (center) and stellar mass–UV magnitude relation (right) are much more similar. At a glance, it is the semiempirical models that generally predict higher stellar masses at fixed halo mass (Behroozi et al. 2013a, 2013b; Tacchella et al. 2018; Park et al. 2019; Mirocha et al. 2020), as well as more not shown here (e.g., Sun & Furlanetto 2016), while simulation-based estimates like BLUETIDES (Feng et al. 2015, 2016) and, e.g., FIRE-2 (Ma et al. 2018), exhibit lower  $M_*/M_h$  ratios at fixed  $M_h$ , more in line with

<sup>11</sup> Note that our results are insensitive to this choice. Because the model is calibrated empirically, any change in  $Z$  must be met with a commensurate change in  $f_*$ , which keeps the rest-UV emission roughly constant. See, e.g., Section 3.4 in Mirocha et al. (2017) for additional discussion.

<sup>12</sup> Note that we show two versions of the ARES model: a default approach using the mean halo MAR computed following Furlanetto et al. (2017), and another in which halos histories are extracted from  $N$ -body simulations (Mirocha et al. 2021). The latter agrees more with Tacchella et al. (2018), who also anchor their semiempirical model to  $N$ -body simulations. The reduced SMHM is due to the steeper growth histories of simulated halos, thus making them brighter and bluer, allowing one to reduce the overall normalization of the star formation efficiency (Mirocha et al. 2021).

<sup>10</sup> <https://github.com/mirochaj/ares>

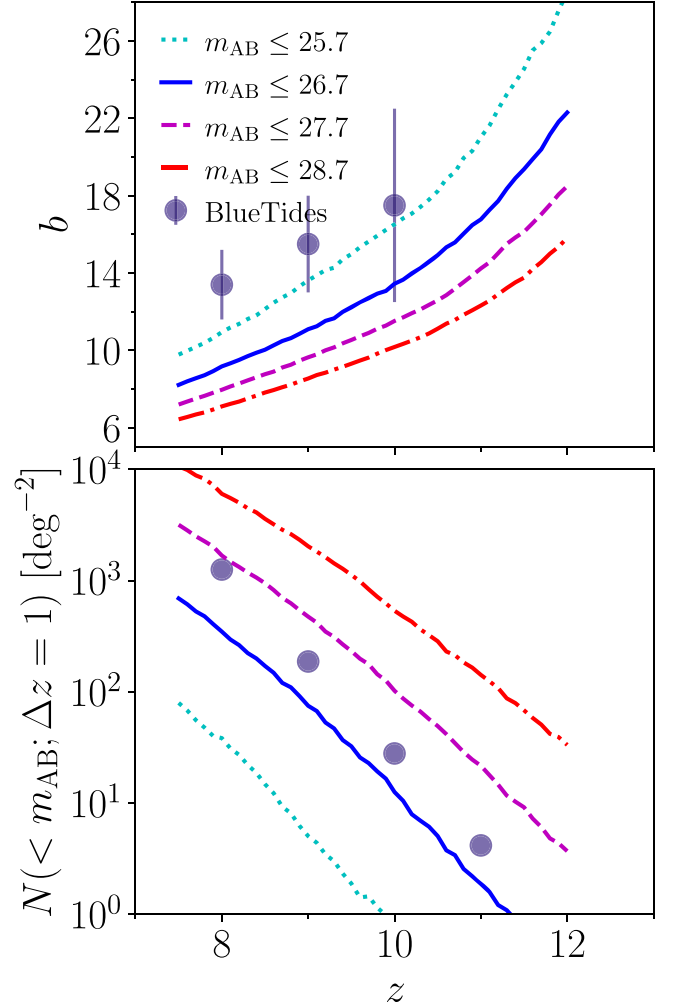


**Figure 2.** Predictions for galaxy properties at  $z \sim 8$  from several models. Left: stellar mass—halo mass (SMHM) relation. Center: specific star formation rate (sSFR) vs. stellar mass. Right: relationship between stellar mass and observed UV magnitude (upper axis is  $L_{1600}/[\text{erg s}^{-1} \text{Hz}^{-1}]$  for BLUETIDES). Note that not all data sets are shown in every panel.

UNIVERSEMACHINE predictions (Behroozi et al. 2019) and other similar semiempirical models built on  $N$ -body simulations (e.g., Moster et al. 2018). However, one can also find examples in the literature of very detailed ab initio simulations that predict higher SMHM ratios than BLUETIDES or FIRE (e.g., *Renaissance*, SERRA; Xu et al. 2016; Pallottini et al. 2022), so the source of the differences here remains (as far as we know) unclear.

To our knowledge, only Waters et al. (2016) explicitly provided predictions for the bias of Roman HLS sources. From Figure 2, we expect the bias of sources in BLUETIDES to be higher than the ARES models described earlier in this section: because there is good agreement between predictions for the sSFR as a function of stellar mass among all models (middle panel), and reasonably good agreement also in predictions for the  $M_{\text{UV}}-M_*$  relation (right), we conclude that at fixed stellar mass, the models predict very similar intrinsic and dust-reddened luminosities. As a result, the SMHM relation will dominate differences in galaxy bias predictions. Indeed, this is the case, as we show in Figure 3 (top), along with predictions for the abundance of high- $z$  galaxies (bottom). Here, we show the BLUETIDES predictions as well as ARES models with four different magnitude cuts: the nominal HLS limiting magnitude of 26.7 (solid blue), as well as scenarios 1 magnitude shallower (dotted cyan) and 1 and 2 magnitudes deeper (dashed and dashed-dotted curves, respectively). As expected from Figure 2, these models differ nontrivially in their predictions: while BLUETIDES predicts  $b(z=8) \simeq 13.5$ , the Mirocha et al. (2020) models predict  $b(z=8) \simeq 9$ , with more rapid evolution at high redshifts than a linear extrapolation of BLUETIDES would suggest. In the bottom panel, we compare predictions for the surface density of galaxies as a function of redshift. Once again, there are noticeable differences, at the  $\sim 2$ – $3\times$  level.

The differences between model predictions for the SMHM relation (and thus galaxy bias) are certainly interesting and warrant attention (see, e.g., Section 4.1 in Tacchella et al. 2018, for additional discussion). In this work, however, we will remain agnostic about which of these models (if any) are correct, and instead use their differences to motivate a plausible range of possibilities to explore in our cross-correlation forecast. Because BLUETIDES has provided predictions for the



**Figure 3.** Predictions for the galaxy bias and abundance at high redshifts. Top: galaxy bias as a function of apparent UV magnitude cut,  $m_{\text{AB}}$ , computed with ARES (lines) compared to BLUETIDES (points), which adopted the nominal HLS value of  $m_{\text{AB}} = 26.7$ . Bottom: galaxy surface densities over the same redshift interval.

bias of HLS sources explicitly, we will explore this as one possible scenario, and use the ARES models as a contrasting case, for which we can efficiently generate alternative scenarios for different survey parameters or galaxy properties. For additional comparisons of the BLUETIDES and ARES predictions, see Appendix A.

Finally, once we have the linear bias in hand, we construct the galaxy field from the matter density field  $\delta_m$  through the relation:

$$\delta_g(\mathbf{k}, z) = b_g(z)\delta_m(\mathbf{k}, z). \quad (8)$$

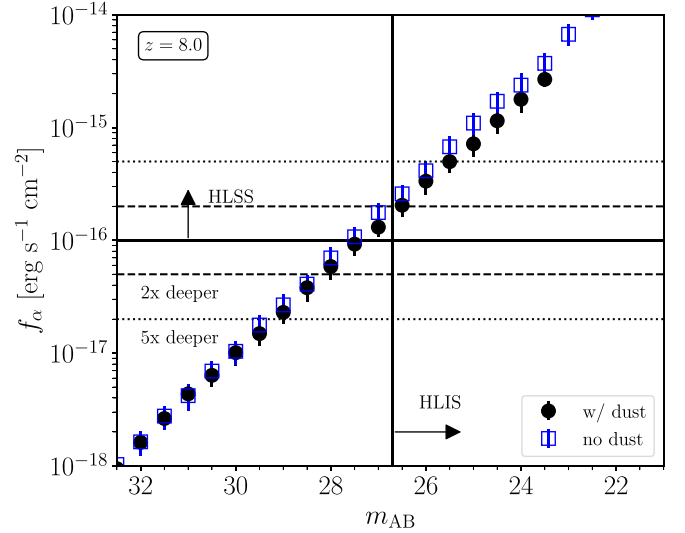
Note that given the relatively large values for the bias  $b_g$ , as seen in Figure 3, the simulated galaxy distribution does sometimes reach nonphysical values in the simulation, where  $\delta_g < -1$ . Although this is a limitation of our current treatment, we do not expect these values to cause significant inaccuracies in the resulting predictions given that most of the cross-spectrum sensitivity comes from large spatial scales where a linear biasing model is a good approximation.

### 2.2.2. Nebular Emission

So far, we have only considered the bias and abundance of galaxies detected in at least one band in the HLIS. However, a key question is whether or not the galaxies detected in the imaging survey will also be detected spectroscopically, since redshift uncertainties  $\sigma_z \gtrsim 0.1$  (e.g., photometric redshifts alone) are expected to prevent a cross-correlation detection (Furlanetto & Lidz 2007). Specifically, if the redshift uncertainties in the galaxy survey are too large, this prevents measuring higher  $k_{\parallel}$  modes in the galaxy distribution, while 21 cm surveys generally lose the low  $k_{\parallel}$  modes to foreground contamination. For 21 cm galaxy cross-power spectrum measurements, it is thus crucial to obtain good spectroscopic redshifts in the galaxy survey, as this will help ensure that each survey measures some common Fourier modes. We discuss the impact of measurement uncertainties in more detail in Section 2.4.2, and here focus on whether galaxies bright enough to be detected in imaging ought to also be bright enough in their Ly $\alpha$  emission to be detected in the spectroscopic survey.

We take a simple approach and assume photoionization equilibrium in the HII regions of galaxies. This allows us to relate the luminosity of Ly $\alpha$  to the recombination rate ( $\sim 2/3$  of recombinations result in Ly $\alpha$  photons), which is in turn related to the intrinsic ionizing photon production rate, one of the main predictions of the galaxy model. The implementation in ARES is described in more detail in Section 2.2.2 of Sun et al. (2021).

In Figure 4, we show our predictions for the relationship between intrinsic Ly $\alpha$  line luminosity and apparent UV magnitude at  $z=8$ , with nominal sensitivity limits for the HLIS and HLSS overlaid for reference (vertical and horizontal lines, respectively). For the flux limits, we adopt the nominal sensitivity quoted in Wang et al. (2022), and include factors of 2x and 5x deeper/fainter surveys for reference as well. Objects detected in the upper-right quadrant of this plot are detected in both imaging and spectroscopy, while objects in the lower-right or upper-left quadrants are only detected in imaging or spectroscopy, respectively. We can see clearly that an object detected in imaging should also be detectable spectroscopically via its Ly $\alpha$  emission. This remains the case regardless of whether one includes dust or not (circles versus squares), since

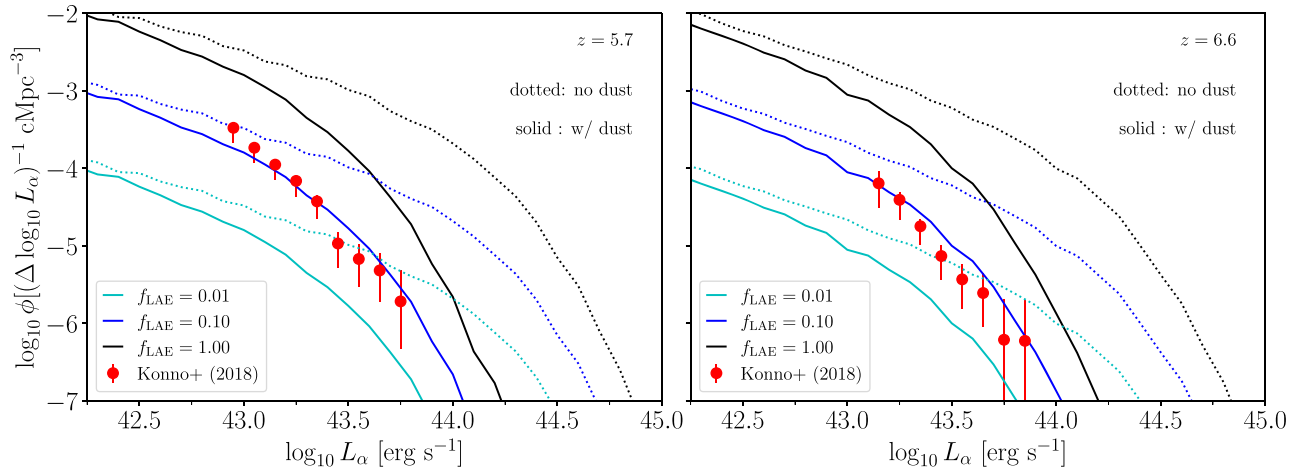


**Figure 4.** Relationship between intrinsic Ly $\alpha$  line luminosity and galaxy UV magnitude. Solid lines indicate the nominal imaging (vertical) and spectroscopic (horizontal) survey depths. The dashed and dotted curves indicate factors of 2 and 5 deviations from the nominal spectroscopic survey, respectively. All HLIS sources detected in the rest-UV continuum should be detected in the HLSS with the nominal sensitivity, neglecting IGM transmission effects.

dust here attenuates the UV continuum and Ly $\alpha$  similarly. Setting the dust contents to zero by hand will boost the overall number of galaxies, as well as the magnitude of the brightest galaxy, but not the relationship between Ly $\alpha$  flux and  $m_{AB}$ . Additionally, each of the galaxies in the HLSS will be observed from multiple roll angles, and so to reduce the incidence of false positives, galaxies will be required to be observed from at least three different rolls.

Figure 4 also tells us that galaxies that are slightly too faint to be detected in imaging may still be detectable in the HLSS. However, given that Roman uses a grism, the expectation is to only extract spectra for known sources identified in imaging. The result shown in Figure 4 is not entirely a surprise. The nominal HLS limiting magnitude and flux limits are chosen to optimize low- $z$  science, e.g., the ability to detect  $1 \lesssim z \lesssim 3$  galaxies both in imaging of the rest-optical continuum (with some contamination from lines) and the rest-optical emission lines themselves, like H $\alpha$  and [O III]. Because the continuum of star-forming galaxies is relatively flat from the UV through the optical, and recombination results in  $\sim 1$  Ly $\alpha$  photon for every H $\alpha$  photon, it makes sense that sensitivities set for low- $z$  science goals are about right for detecting high- $z$  galaxies in Ly $\alpha$  and UV continuum as well.

Before moving on, note that so far we have neglected IGM transmission effects, effectively assuming that every HLS galaxy resides in a very large fully ionized bubble. In all that follows, we adopt an Ly $\alpha$  emitter (LAE) duty cycle or fraction,  $f_{LAE}$ , which reduces the number of spectroscopically detected galaxies but not the luminosity of any individual object. In practice, the LAEs observable by Roman will have some spatial and luminosity dependence primarily owing to absorption from neutral regions in the IGM. However, for the sake of



**Figure 5.** The LAE galaxy luminosity function predicted by ARES and measured from SILVERRUSH (Konno et al. 2018). As can be seen, an effective LAE duty cycle of  $f_{\text{LAE}} = 0.1$  with dust reddening agrees reasonably well with the data at both  $z \sim 5.7$  and  $z \sim 6.6$ . We use this value as our fiducial value for the rest of the analysis, but include predictions for  $f_{\text{LAE}} = 1$  and  $f_{\text{LAE}} = 0.01$  as additional points of comparison.

simplicity in this initial study, we instead use an overall factor and defer a more realistic treatment to future work.<sup>13</sup>

Figure 5 shows a comparison between the LAE galaxy luminosity function predicted from ARES and the measurements from the SILVERRUSH survey (Konno et al. 2018). We include our galaxy models both with and without dust reddening, and plot different factors of  $f_{\text{LAE}}$ . As can be seen from the figure, a value of  $f_{\text{LAE}} = 0.1$  and including dust reddening agrees reasonably well with the available measurements at  $z \sim 5.7$  and  $z \sim 6.6$ . Note that our overall factor  $f_{\text{LAE}}$  is not inferred the same way as  $f_{\text{duty}}$ , which sometimes appears in the interpretation of these measurements. Constraints on  $f_{\text{duty}}$  are typically based on clustering measurements: one infers a large-scale bias, which can be related to a minimum halo mass and thus a prediction for the total number density of halos. The value for  $f_{\text{duty}}$  is then computed as the number of LAEs divided by the number of expected halos. In our case, the number of LAEs in any  $L_\alpha$  bin is reduced both by  $f_{\text{LAE}}$  and dust reddening. As a result, we do not need extreme values of  $f_{\text{LAE}} = 0.01$  unless dust is completely negligible.

In the following analysis, we choose  $f_{\text{LAE}} = 0.1$  as our default value, but also include predictions for  $f_{\text{LAE}} = 1$  and  $f_{\text{LAE}} = 0.01$  for the sake of comparison and understanding how this values impacts our final results. For example,  $f_{\text{LAE}} = 0.01$  may be a more realistic value at higher redshift, when the IGM becomes more neutral and the opacity to Ly $\alpha$  photons increases due to there being smaller neutral regions (and hence less time for photons to redshift out of the Ly $\alpha$  transition window).

### 2.3. The 21 cm Galaxy Cross-spectrum

Once we have the 21 cm brightness temperature field defined in Equation (5) and the galaxy field defined in Equation (8), we are able to compute the cross-spectrum  $P_{21 \times \text{gal}}$ . For the sake of comparing the results with observations, we compute this quantity as a function of  $k_\perp$  and  $k_\parallel$ , i.e., the Fourier modes that

lie in the plane of the sky and along the line of sight, respectively. Given the relatively small angular extent of our simulations, we use the flat-sky approximation when computing power spectra. We average over all modes that contribute to a given cylindrical wavenumber bin, described by  $(k_\perp, k_\parallel)$ :

$$P_{21 \times \text{gal}}(k_\perp, k_\parallel) = \langle \delta T_b^* \delta_g \rangle_{k_\perp, k_\parallel}. \quad (9)$$

We first construct a pair of light cones—one each for the 21 cm field and the galaxy field—in a self-consistent fashion using our 2LPT simulations described above. Once we have these light cones, we apply FFTs and then compute the cross-power spectrum as described above.

Figure 6 shows a visualization of the 21 cm field and the galaxy fields as generated by the seminumeric models described above in Sections 2.1 and 2.2. These fields are generated from a realization of our fiducial ionization history by averaging over the respective light cones between  $7.5 \leq z \leq 8.5$ . Note that in practice when computing the cross-spectrum defined above in Equation (9), we compute the FFT in three dimensions and do not average prior to computing this cross-spectrum. This approach ensures that we retain the full Fourier space information and can accurately model the effect of avoiding modes that are contaminated by the 21 cm foregrounds described below in Section 2.4.1.

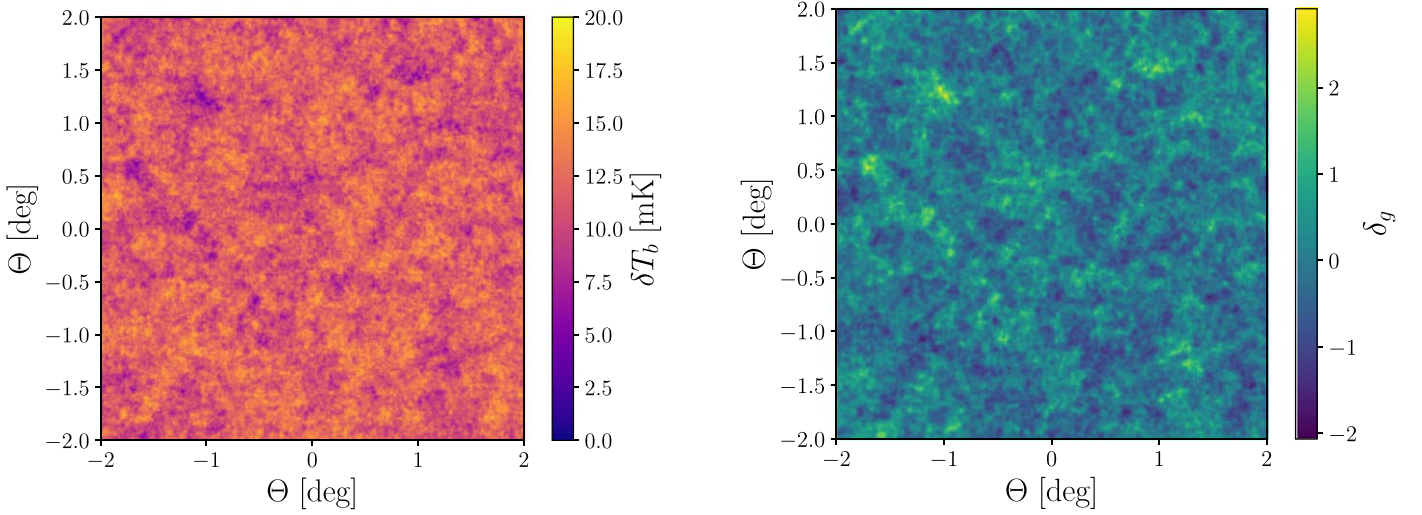
## 2.4. Observational Effects

### 2.4.1. 21 cm Observations

In general, 21 cm observations from interferometers are subject to foreground contamination that are many orders of magnitude larger than the cosmological signal from reionization (Morales & Wyithe 2010; Pober et al. 2013). The dominant foreground is due to galactic synchrotron radiation from the Milky Way, which roughly follows a power law in frequency and hence is smooth in Fourier space. Naïvely, these bright foregrounds should be restricted to small  $k_\parallel$  modes in Fourier space. However, the chromatic response of the interferometer causes this foreground signal to scatter into high  $k_\parallel$  modes, creating a “wedge” in  $(k_\perp, k_\parallel)$  Fourier space (Parsons et al. 2012).

<sup>13</sup> Because the optical depth to Ly $\alpha$  is so large close to the line center, even an LAE in an ionized bubble will generally be attenuated. For instance, if the intrinsic line were a simple Gaussian of some width, the blue half of the line will typically be lost, even in a large ionized bubble. In future work we plan to account for attenuation of LAEs due to the IGM explicitly, but for now use the overall factor  $f_{\text{LAE}}$  to include these various physical effects.





**Figure 6.** A visualization of our simulation volumes. Left: the 21 cm field generated from our simulations described in Section 2.1. This  $2^\circ \times 2^\circ$  field represents about 1% of the simulated sky area from one of our simulations. Right: a self-consistently generated galaxy field described in Section 2.2. These two-dimensional slices are averaged over the redshift window  $7.5 \leq z \leq 8.5$ , which spans the midpoint of reionization for this realization. Although we show the two-dimensional field for illustration purposes, we compute the cross-spectrum in three dimensions to avoid catastrophic cancellation of Fourier modes contaminated by 21 cm foregrounds.

Mathematically, the slope of the wedge  $m(z)$  is implicitly a function of baseline length and cosmology, and can be expressed as (Thyagarajan et al. 2015):

$$m(z) \equiv \frac{k_{\parallel}}{k_{\perp}} = \frac{\lambda(z)D_c(z)f_{21}H(z)}{c^2(1+z)^2}, \quad (10)$$

where  $\lambda(z)$  is the wavelength of the 21 cm signal at a given redshift  $z$ ,  $D_c(z)$  is the comoving distance to that redshift,  $f_{21}$  is the rest-frame frequency of the 21 cm signal, and  $H(z)$  is the Hubble parameter. Note that this form of the foreground wedge in Equation (10) accounts for maximal data contamination: physically, the bright foreground contamination extends down to the horizon of the interferometer beam. Improved calibration of interferometers may make it possible to work “inside the wedge” at points in Fourier space where the contamination from foregrounds may be less severe, but for the purposes of this work, we assume this maximal amount of contamination. To include this effect, we explicitly track which portions of Fourier space for each redshift are subject to this foreground contamination. For the redshift values relevant to the EoR,  $m \sim 3$ , which leads to a significant amount of Fourier space being subject to this contamination.

In addition to the foreground contamination, we include the effects of thermal noise present in measurements from HERA. For power spectrum measurements, the primary analysis pipeline for HERA uses the so-called “delay transform” of radio interferometer visibilities to estimate the power spectrum. Under the flat-sky approximation, the thermal noise contribution to the variance of the power spectrum can be written as a function of the observational frequency  $\nu$  and wavenumber  $u$  as (Parsons et al. 2014):

$$P_{21}^{\text{noise}}(u, \nu) \approx \frac{T_{\text{sys}}^2 \Omega_p^2(\nu) X^2(\nu) Y(\nu)}{\Omega_{pp}(\nu) t_{\text{int}} N_{\text{pol}} N_{\text{bl}}(u)}, \quad (11)$$

where  $T_{\text{sys}}$  is the system temperature of the interferometer,  $\Omega_p = \int d[2] |A(\mathbf{l})|^2$  is the integral of the primary beam of the antenna  $A(\mathbf{l})$ , and  $\Omega_{pp} = \int d[2] |A(\mathbf{l})|^4$  is the integral of the square

of the primary beam.  $X(\nu)$  and  $Y(\nu)$  are factors that convert the “observed units” of the interferometer into cosmological units (Furlanetto et al. 2006).  $X(\nu)$  accounts for the conversion in the plane of the sky:

$$X(\nu) \equiv \frac{dr_{\perp}}{dl} = D_M(\nu), \quad (12)$$

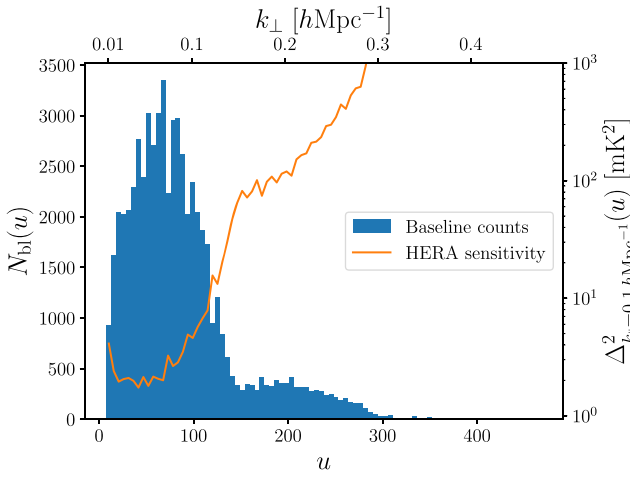
where  $D_M(\nu)$  is the transverse comoving distance ( $D_M = D_c$  for a flat universe); and  $Y(\nu)$  accounts for the conversion along the line of sight:

$$Y(\nu) \equiv \frac{dr_{\parallel}}{d\nu} = \frac{c(1+z)}{H(z)\nu}. \quad (13)$$

Note that  $X(\nu)$  has units of comoving megaparsecs per radian, and  $Y(\nu)$  has units of comoving megaparsecs per hertz. Finally,  $t_{\text{int}}$  is the integration time of the measurement,  $N_{\text{pol}}$  is the number of instrumental polarizations used to estimate the power spectrum, and  $N_{\text{bl}}(u)$  is the number of baseline pairs at a given  $u$ , where  $u$  denotes the physical separation of HERA dishes in units of observed wavelength. Each baseline of length  $u$  samples wavenumbers with transverse components,  $k_{\perp}$ , according to  $u(\nu) = D_M(\nu)k_{\perp}/2\pi$ . For HERA, we assume a system temperature of  $T_{\text{sys}} = 400$  K, an observation time of  $t_{\text{int}} = 200$  hr, and  $N_{\text{pol}} = 2$  for two independent linear polarizations. We assume an overall observing season of 1000 hr, which is distributed among five nonoverlapping fields. Such an approach was taken in the recent reanalysis of Phase I HERA data (The HERA Collaboration et al. 2022). We discuss combining observations from different regions of the sky more below in Section 4.3.

The number of baselines that observe a wavenumber  $u$  depends on the configuration of the array. We show the baseline distribution for HERA in Figure 7. HERA features many short baselines, so most of the baselines probe modes for which  $u \lesssim 300$ . We also show the noise floor for an observation, given by Equation (11). For the current work, we limit ourselves to considering only the projected sensitivity for HERA. Next-generation radio telescopes, such as the





**Figure 7.** The baseline distribution and sensitivity of HERA-350. We plot the distribution of baselines for HERA as a function of wavenumber  $u$  for the fully constructed 350 element array at  $z=8$ . The noise power spectrum in Equation (11) is inversely proportional to the number of baselines that observe a particular  $u$  mode on the sky, and is shown in orange. We plot the dimensionless noise power spectrum  $\Delta^2(k) = k^3 P(k)/2\pi^2$  with a fixed value of  $k_{\parallel} = 0.1 \text{ h Mpc}^{-1}$ , for 200 hr of observation (i.e., a single field).

Square Kilometre Array (SKA), are projected to provide data with even greater sensitivity. However, as we will see below in Section 4.1, the 21 cm signal variance is dominated by cosmic variance of the modes used to make the measurements, rather than the instrumental uncertainty. As such, the additional sensitivity from the SKA may not significantly improve the measured significance of the 21 cm auto-power spectrum. Nevertheless, the SKA may offer a significant improvement over HERA by being able to use more of the observable Fourier space, parameterized by the foreground wedge  $m$  in Equation (10). The SKA may also be able to observe more sky area that overlaps with the Roman HLS, yielding more joint sky coverage. It is also worth noting that given the improved sensitivity offered by the SKA, it may be possible to do a “stacking” type analysis in real (map) space, rather than Fourier space as assumed in this manuscript. It may be worthwhile to revisit some of these types of forecasts in the future to investigate the prospects for the SKA.

#### 2.4.2. Galaxy Surveys

Upcoming galaxy surveys will provide us the deepest and most comprehensive catalog of high-redshift objects to date. Nevertheless, these objects are still relatively faint and rare, and as such are only sparsely sampled. Furthermore, the redshifts of the galaxies in these samples are only approximately known. Assuming Poisson statistics and accounting for redshift uncertainties, the noise power spectrum for the galaxy field is:

$$P_{\text{gal}}^{\text{noise}}(k_{\parallel}) = \frac{e^{k_{\parallel}^2 \sigma_z^2}}{n_{\text{gal}}}, \quad (14)$$

where  $\sigma_z = c\sigma_z/H(z)$  is the comoving distance uncertainty along the line of sight, given a  $1\sigma$  redshift precision of  $\sigma_z$ , and  $n_{\text{gal}}$  is the comoving number density of galaxies. The HLS in Roman is expected to have a total survey area of  $2200 \text{ deg}^2$  (Spergel et al. 2015). However, the HLS footprint is not expected to fully overlap with the regions of the sky surveyed by HERA. For the purposes of this analysis, we assume an

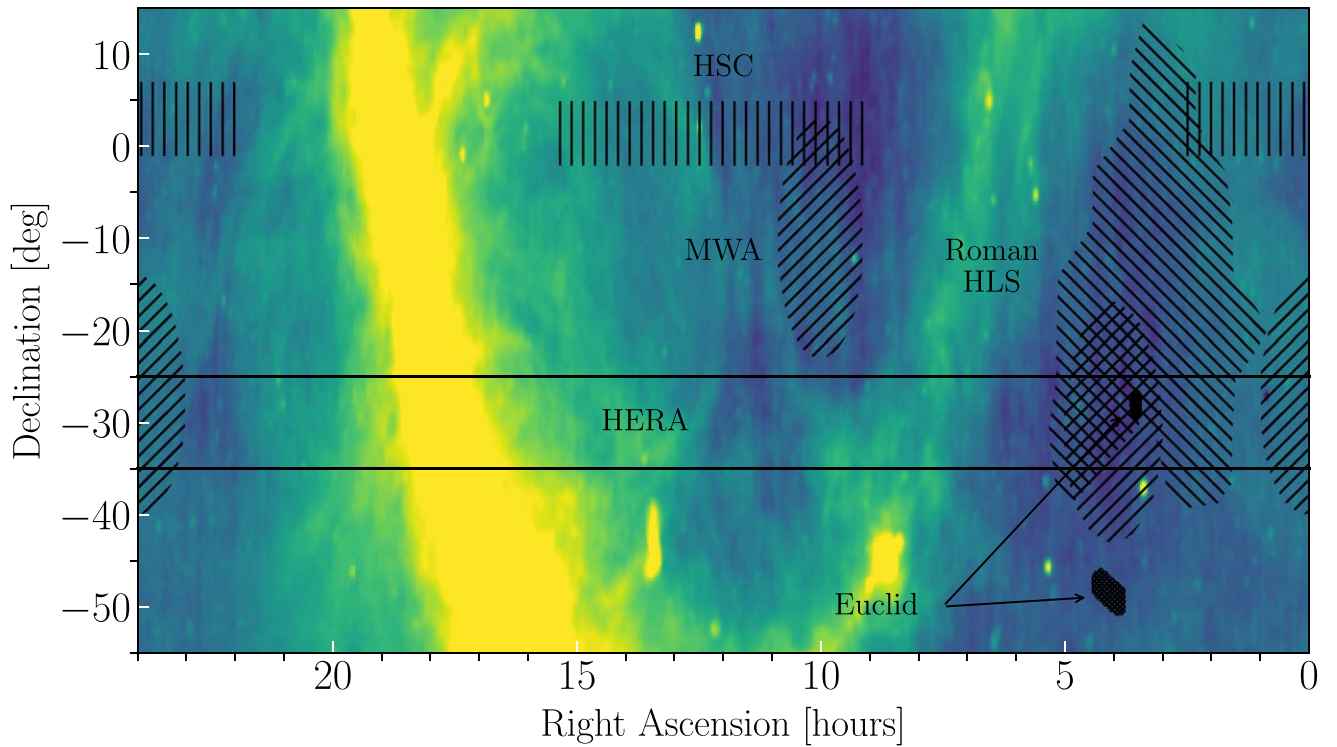
overlapping sky region of  $500 \text{ deg}^2$ , as is expected given the nominal HLS footprint (see Figure 8). Note that the planned Euclid deep fields are comparable to the HLS in depth. Unfortunately, only one lies in the HERA stripe, and is relatively narrow in the sky area it covers. We discuss trade-offs between depth and area further in Section 5.

Also vital to the success of any cross-correlation effort are precise redshift measurements. Photometric redshift estimates are expected to yield uncertainties of  $\sigma_z \sim 0.5$  using the Lyman-break technique (Bouwens et al. 2015; Finkelstein 2016), while we expect uncertainties  $\sigma_z \gtrsim 0.1$  to preclude a cross-correlation detection (Lidz et al. 2009). Given the high redshifts of interest, a successful HERA–Roman cross-correlation requires spectroscopically determined redshifts. Fortunately, the reference spectroscopic survey (Wang et al. 2022) expects  $\sigma_z \simeq 0.001(1+z)$ , and so redshift measurements in principle should not be a limiting factor. We explore the effect of redshift uncertainty explicitly in Section 4.1 below.

Given the sparsity of strong rest-frame UV lines in star-forming galaxies,  $\text{Ly}\alpha$  is likely the only line to be detected in Roman spectroscopy at the redshifts of interest. Furthermore, grism spectroscopy will only be extracted at the locations of sources detected in imaging. As a result, we have two requirements: (i) that galaxies be detected via the Lyman-break technique, and (ii) that  $\text{Ly}\alpha$  is brighter than the sensitivity of the spectroscopic survey. Though there are potential systematic observing issues, we defer a discussion of these to Section 4.4.1.

The first requirement is explored in Figure 9, where we show the spectral coverage of the Roman photometry and spectroscopy. The bottom panel shows the transmission curves for the filters that will be used to find dropouts, while in the top panel, we show the redshift ranges corresponding to dropouts in each filter. We follow Drakos et al. (2022), and for simplicity assign the dropout filter as the bluest filter that contains  $\text{Ly}\alpha$ . All galaxy magnitudes are reported in the filter just redward of the dropout filter, in order to avoid contamination from the break and the  $\text{Ly}\alpha$  line itself. Though the dropout technique can be used to identify galaxies at redshifts at  $z \lesssim 4$  with Roman, the grism covers only  $\lambda \in [1, 1.93] \mu\text{m}$ . As a result, only photometrically selected galaxies at  $z \gtrsim 7.2$  have a chance at an  $\text{Ly}\alpha$  detection.

The second requirement for a galaxy to be included in cross-correlation analyses—that  $\text{Ly}\alpha$  is bright enough to be detected—is more difficult to assess. As discussed in Section 2.2 and shown in Figure 4, given the nominal HLSS line flux limits (Wang et al. 2022), any galaxy detected photometrically should also have detectable  $\text{Ly}\alpha$  emission assuming negligible absorption from the IGM. This is of course an optimistic assumption. In general, the occurrence rate of LAEs will be reduced by intergalactic absorption, particularly sources residing in small bubbles. Though the intrinsic LAE fraction at high redshifts is unknown, simulations suggest that  $\text{Ly}\alpha$  escape could be nontrivial at the redshifts of interest here (Garel et al. 2021; Smith et al. 2022). Furthermore, the HLS preferentially picks out bright sources that very well may live in large bubbles and so be subject to little attenuation from the IGM. We explore  $f_{\text{LAE}} = 0.1$  as our fiducial case, but also explore contrasting scenarios in which only 1% or 100% of galaxies detected in the HLIS have detectable  $\text{Ly}\alpha$  emission in the HLSS. We will denote these cases via  $f_{\text{LAE}} = 0.01$  and  $f_{\text{LAE}} = 1$ , respectively. The pessimistic scenario may apply if



**Figure 8.** Overlap of HERA and upcoming wide-field galaxy surveys. HERA observes a stripe of constant decl.,  $\delta = 30^\circ \pm 5^\circ$ , which partially overlaps with the nominal Roman HLS footprint (cross-hatching between  $1 \lesssim \text{R. A./hr} \lesssim 5$ ). Two Euclid deep fields are planned in the South that have sensitivity comparable to the HLS—the Deep Field South and Deep Field Fornax—but only the latter is in the HERA stripe. One of the three MWA EoR fields overlaps with Roman as well (circular cross-hatching), and the most northerly MWA field overlaps with LAE campaigns being conducted with Subaru Hyper-Suprime Cam (vertical hatching; see, e.g., Ouchi et al. 2018; Trott et al. 2021). The background color-scale is the global sky model at 150 MHz (De Oliveira-Costa et al. 2008).

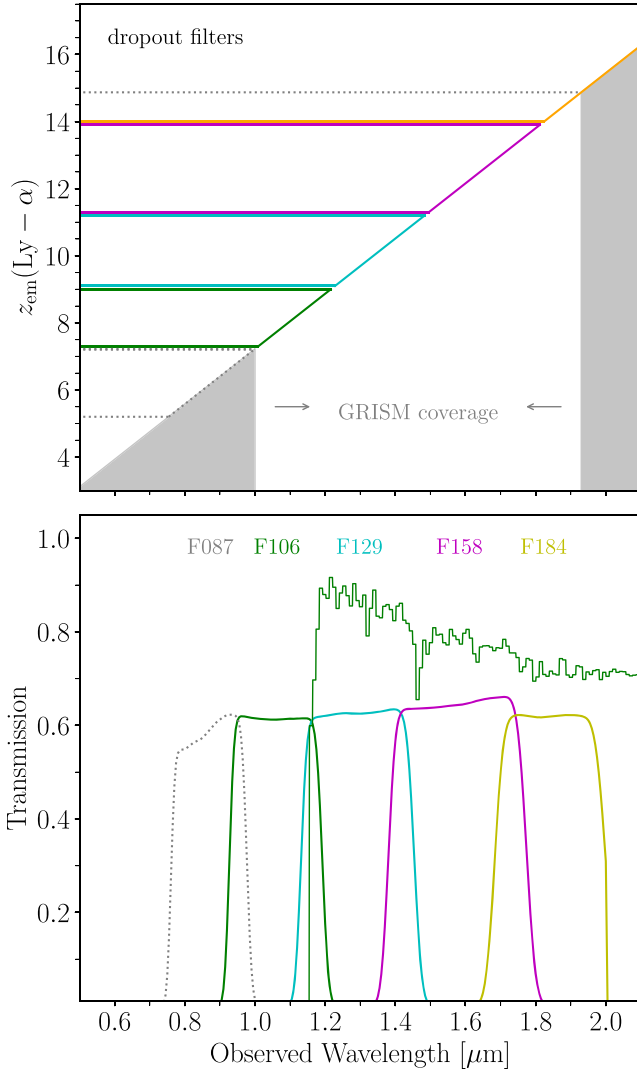
the LAE populations uncovered by the Subaru Hyper-Suprime Cam (e.g., Ouchi et al. 2018) are representative of all reionization era LAEs, while the  $f_{\text{LAE}} = 1$  case is included as a maximally optimistic limit.

### 3. Results

We first show results for the cross-correlation signal  $P_{21\text{cm} \times \text{gal}}$  as a function of  $k_\perp$  and  $k_\parallel$ . This is the raw signal that we are attempting to measure. When computing the cross-spectrum, we construct a light cone of observations that span the full redshift range  $6 \leq z \leq 15$ . In practice, the Roman grism frequency coverage means that we are not sensitive to galaxies at  $z < 7.2$ . Furthermore, the most detectable contributions to the cross-spectrum will come from relatively low redshift values ( $z \lesssim 9$ ). Nevertheless, we include both higher and lower redshift ranges when constructing the light cones and then restrict ourselves to  $7.5 \leq z \leq 12$  for analysis. Once this has been done, we compute the Fourier transform of nonoverlapping windows along the line of sight that each cover 6 MHz of bandwidth. This is representative of the typical bandwidth used in previous HERA data analyses (Abdurashidova et al. 2022a, 2022b). This bandwidth also translates to comoving distances of  $50 \lesssim \chi \lesssim 100 h^{-1} \text{ Mpc}$  depending on the redshift. Although these cubes are not strictly comoving, the evolution induced by the light cone is sufficiently small so as not to induce spurious signal (La Plante et al. 2014). To further decrease the amount of light cone evolution signal included, we apodize in the line-of-sight direction using a four-term Blackman-Harris window. We run 30 independent realizations of these simulations where we change the initial conditions but keep the cosmological and astrophysical parameters fixed, and

average together the resulting power spectra computed in the described fashion. This averaging helps ensure that the spectra are smooth, especially on the scales relevant to upcoming observations.

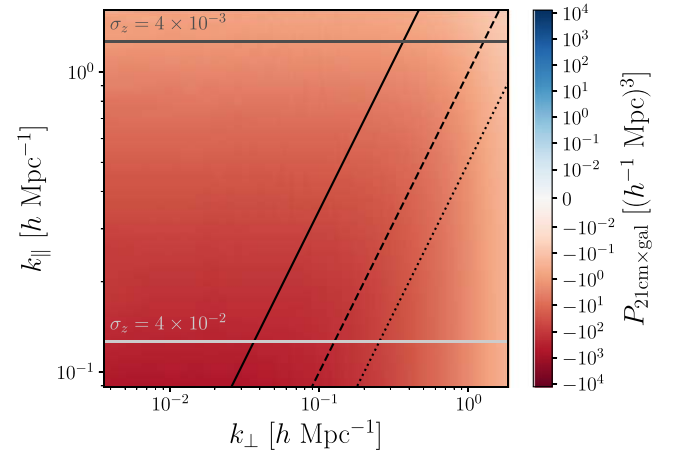
Figure 10 shows the cross-spectrum  $P_{21\text{cm} \times \text{gal}}$  near the midpoint of reionization at  $z \sim 8.01$ , where the amplitude of the cross-spectrum is greatest. In general, the signal is negative, which indicates an anticorrelation. This anticorrelation is expected, as the brightness temperature  $\delta T_b$  in Equation (5) is sourced by regions of neutral hydrogen, which tend to be low-density regions far from galaxies in an inside-out reionization scenario. We also note that the largest-amplitude signal comes from large scales (small values of  $k_\perp$  and  $k_\parallel$ ). We also show as a solid line the value of the slope of the “horizon wedge” defined in Equation (10). For comparison, we also show values of  $m = 1$  (dashed) and  $m = 0.5$  (dotted) lines. These represent varying levels of foreground contamination: the horizon wedge reflects data that are maximally corrupted, whereas less-severe cuts would be possible if improved calibration methods make it possible to recover these data. It is also worth noting that there is a much larger dynamic range in  $k_\perp$  than in  $k_\parallel$ : given the fact that we perform a Fourier transform on a slab with a relatively short axis along the line of sight, there are far fewer  $k_\parallel$  modes available for a given observation. In order to provide a more significant cumulative measurement of the signal, we combine measurements from nonoverlapping redshift windows along the line of sight. Although precise characterization of the  $k$ - and  $z$ -dependence of the signal would be ideal, for the near-term forecast most relevant to HERA, we are focused on a detection. We also ignore potential covariance between different  $k$ - and  $z$ -bins on the assumption that they are relatively free of



**Figure 9.** Roman spectral coverage. Bottom: we show transmission curves for five filters, starting with the  $z$  band (F087), which is not used in the nominal HLS, followed by the  $Y$ ,  $J$ ,  $H$ , and F184 band filters, from left to right, which will all have HLS coverage. An example spectrum of a  $z = 8.5$  dust-free galaxy from our model is shown for reference (with nebular emission neglected for clarity). Top: we show the corresponding redshift range probed by each filter, where the “dropout” filter is the bluest filter containing the  $\text{Ly}\alpha$  line. The fact that grism coverage is restricted to  $\lambda \gtrsim 1 \mu\text{m}$  limits spectroscopic redshifts to galaxy candidates at  $z \gtrsim 7.2$ .

systematic errors. We return to these assumptions in Section 4.4.2.

We also show as horizontal lines several values of  $k_{\parallel}$  corresponding to  $\sigma_z$  galaxy uncertainties defined in Equation (14). We plot values where  $k_{\parallel} = 1/\sigma_z$ , above which the galaxy uncertainty exponentially increases. Given the exponential nature of the noise contributed from this uncertainty, values of  $k_{\parallel}$  greater than these lines face significant contamination. That is, modes with larger  $k_{\parallel}$  are poorly measured in the galaxy survey and do not contribute significantly to a cross-spectrum detection (see further discussion in Section 4). As mentioned above in Section 2.4.2, this implies that photometric surveys are insufficient for present purposes, since the redshift uncertainties in high-redshift Lyman-break surveys are expected to be on the order of  $\sigma_z \simeq 0.5$ .



**Figure 10.** The cross-spectrum  $P_{21\text{cm} \times \text{gal}}$  as a function of  $k_{\perp}$  and  $k_{\parallel}$ . The spectrum has been computed for a window centered on  $z = 8.01$ , which is near the midpoint of reionization for our fiducial scenario. The quantity  $T_0(z)$  from Equation (6) has been divided out. We also plot the slope of the “wedge” to demonstrate the extent of foreground contamination. The solid line is the horizon wedge defined by Equation (10), and more optimistic values of  $m = 1$  (dashed) and  $m = 0.5$  (dotted). Horizontal lines correspond to  $k_{\parallel}$  values associated with uncertainties in redshift determination  $\sigma_z$  defined in Equation (14).

#### 4. Detectability

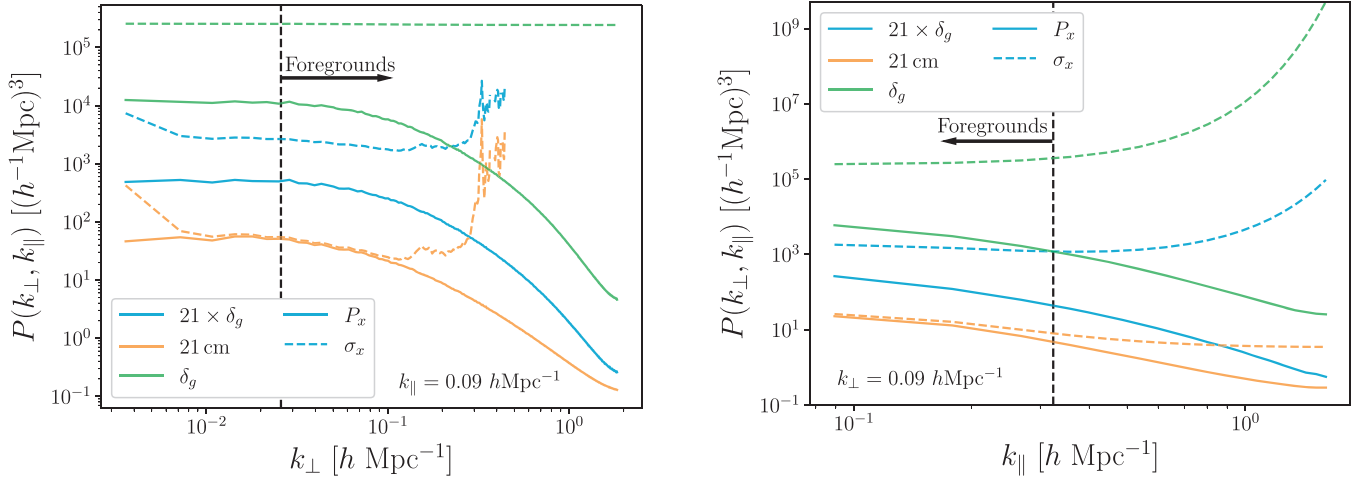
We now turn to forecast the expected signal-to-noise ratio (S/N) for future cross-spectrum measurements. In this analysis, we primarily concern ourselves with the statistical uncertainty of these measurements, and follow related calculations in Furlanetto & Lidz (2007) and Lidz et al. (2009). For 21 cm experiments, we examine the projected sensitivity for HERA. For galaxy surveys, we concern ourselves primarily with the Roman HLS and HLS-like surveys. We do not consider the effect of systematic errors for two reasons: (i) in regions of Fourier space that are not dominated by foreground contamination of the wedge, the Fourier modes should be noise dominated, and (ii) in general, cross-correlation measurements are unbiased (at the cost of higher statistical variance) if there are no sources of joint systematic error between the two measurements. We begin with a more detailed description of the observational approaches of Roman and HERA in Section 4.2, go through a treatment of the S/N in our fiducial reionization scenarios in Section 4.1, explore building up cumulative S/N by combining measurements from different modes in Section 4.3, and then turn to possible systematic errors in Section 4.4.

##### 4.1. S/N Calculations

We are interested in the uncertainty of the cross-spectrum  $\sigma_{21, \text{gal}}$  as a function of  $k_{\perp}$  and  $k_{\parallel}$ , as this defines our expected sensitivity in Fourier space. In this case, the uncertainty can be expressed as (Lidz et al. 2009):

$$\begin{aligned} \sigma_{21, \text{gal}}^2 &= \text{var} \left[ \frac{1}{T_0(z)} P_{21 \times \text{gal}}(k_{\perp}, k_{\parallel}) \right] \\ &= \frac{1}{2} [P_{21 \times \text{gal}}^2(k_{\perp}, k_{\parallel}) \\ &\quad + \sigma_{21}(k_{\perp}, k_{\parallel}) \sigma_{\text{gal}}(k_{\perp}, k_{\parallel})]. \end{aligned} \quad (15)$$





**Figure 11.** The signal and noise quantities of the various components of Equation (15). Solid lines show the signal component, and dashed lines show the corresponding noise. Note that the cross-spectrum signal  $P_{21 \times \delta_g}$  is negative, so we have plotted the absolute value. Left: components as a function of  $k_{\perp}$  at a fixed value of  $k_{\parallel}$ . Right: components as a function of  $k_{\parallel}$  at a fixed value of  $k_{\perp}$ . Vertical dashed lines show the location of the horizon wedge given the fixed values of  $k_{\parallel}$  (left) and  $k_{\perp}$  (right), where the arrows point toward regions that are dominated by foreground contamination.

Similarly, the uncertainty of the 21 cm power spectrum  $\sigma_{21}$  can be written as:

$$\sigma_{21}^2 = \text{var}[P_{21}(k_{\perp}, k_{\parallel})] = \left[ \frac{1}{T_0(z)^2} P_{21}(k_{\perp}, k_{\parallel}) + \frac{1}{T_0(z)^2} P_{21}^{\text{noise}}(k_{\perp}, k_{\parallel}) \right]^2, \quad (16)$$

where  $T_0$  is defined in Equation (6) and  $P_{21}^{\text{noise}}$  is defined in Equation (11). Finally, the uncertainty of the galaxy power spectrum  $\sigma_{\text{gal}}$  can be written as:

$$\sigma_{\text{gal}}^2 = \text{var}[P_{\text{gal}}(k_{\perp}, k_{\parallel})] = [P_{\text{gal}}(k_{\perp}, k_{\parallel}) + P_{\text{gal}}^{\text{noise}}(k_{\parallel})]^2, \quad (17)$$

where  $P_{\text{gal}}^{\text{noise}}$  is defined in Equation (14). These expressions capture the total variance of the observed power spectra, which is due both to cosmic variance and statistical instrumental uncertainties.

Figure 11 shows the signal and noise quantities defined in Equations (15)–(17). The solid lines correspond to the signal, and the dashed lines correspond to the noise. To build intuition for how these quantities evolve with Fourier modes  $k_{\perp}$  and  $k_{\parallel}$ , we hold one mode fixed and plot the signal and noise curves as a function of the other one. We show results for  $f_{\text{LAE}} = 0.1$  and  $\sigma_z = 0.01$  with a redshift window centered at  $z = 8.01$ , near the midpoint of reionization. The panel on the left shows the behavior for fixed  $k_{\parallel} = 0.09 \text{ h Mpc}^{-1}$ , and varies  $k_{\perp}$ . Also shown as a vertical dashed line is the location of the horizon wedge given the value of  $k_{\parallel}$ . As denoted in the figure, the modes to the right of this line are nominally contaminated by foregrounds with a horizon-level wedge defined in Equation (10). In this plot, the effect of the  $uv$  coverage of HERA can be seen in the 21 cm noise curves: at very low values ( $k_{\perp} \lesssim 0.01 \text{ h Mpc}^{-1}$ ) and high values ( $k_{\perp} \gtrsim 0.3 \text{ h Mpc}^{-1}$ ), the experimental uncertainty increases dramatically because of the finite baseline size and the lack of very long baselines, respectively. As seen in Figure 7, many of HERA’s baselines are relatively short, meaning there is good coverage of intermediate  $k_{\perp}$  modes at the expense of large ones. On the right panel, we show the behavior for fixed  $k_{\perp} = 0.09 \text{ h Mpc}^{-1}$

and vary  $k_{\parallel}$ . At relatively large values of  $k_{\parallel}$ , the noise dominates significantly for the galaxy signal. It is also worth noting that in this panel, the foreground-contaminated modes lie to the left of the vertical dashed line.

It is worth discussing some of the large-scale behavior of the trends in this figure to understand how the overall S/N ratio is affected by various observational and experimental considerations. First, it is worth noting that for many values of  $k_{\perp}$ , the 21 cm signal is cosmic-variance dominated. Accordingly, for modes that are well sampled by HERA, a significant detection of the auto-spectrum should be possible. Conversely, the galaxy power spectrum as measured by Roman is dominated by experimental uncertainties, especially at large values of  $k_{\perp}$  and  $k_{\parallel}$ . For large values of  $k_{\perp}$  there is a drop-off in the amplitude of the signal at these modes, whereas the low sensitivity for large values of  $k_{\parallel}$  is driven primarily by the exponential scaling of the redshift-space uncertainty in Equation (14). Note that the variance decreases with the number of modes in a survey, as given by Equation (18). Thus it is possible to decrease the instrument uncertainty by using a sufficiently large volume. We return to prospects for measuring the galaxy auto-spectrum below in Section 4.3. Given the uncertainty of the cross-spectrum is essentially the geometric mean between the individual uncertainties, the S/N of the cross-spectrum is not cosmic-variance dominated, but it also is not as noise dominated as the galaxy spectrum.

#### 4.2. Observational Strategies

In addition to the per-mode uncertainty described above, another important ingredient in the ultimate sensitivity calculations is the number of Fourier modes that may be observed given our survey volume  $V_{\text{survey}}$ . In practice, we consider cross-spectrum estimates in bins in  $k_{\perp}$  and  $k_{\parallel}$ . A given bin will generally receive contributions from many independent Fourier modes, with the Fourier-space resolution set by the survey dimensions, and averaging over these modes will decrease the uncertainties in the bin-averaged measurements.

Consider a cylindrical shell/bin of fixed logarithmic extent  $d \ln k_{\perp}$  and  $d \ln k_{\parallel}$ , where we have constructed an annulus in the plane of the sky with radius  $k_{\perp}$  and at coordinate  $k_{\parallel}$  along the line-of-sight dimension. The number of Fourier modes in our

survey that lie within this cylindrical bin is given by:

$$dN(k_{\perp}, k_{\parallel}) = \frac{k_{\perp}^2 k_{\parallel} V_{\text{survey}}}{(2\pi)^2} d \ln k_{\perp} d \ln k_{\parallel}. \quad (18)$$

Given that the underlying fields are real, there is a Hermitian symmetry of the Fourier transform applied when constructing the power spectrum. As such, we restrict ourselves to estimating the variance only from the upper half-plane so as not to double-count these modes when estimating the variance. We have also taken the absolute value of  $k_{\parallel}$  assuming that positive and negative values are statistically similar.<sup>14</sup>

As discussed above in Section 2.4 and shown in Figure 8, we expect that the overlap between Roman and HERA is about 500 deg<sup>2</sup>. Given that HERA is a drift-scan instrument (rather than a tracking/pointing array), we can treat a joint measurement as a series of single-field patches that we sum together incoherently. This approach is similar to the one advocated in Liu & Shaw (2020), where the number of such patches is parameterized as  $N_{\text{patch}}$ . The instantaneous FOV of HERA is about 10° (DeBoer et al. 2017), so we take the number of such patches to be  $N_{\text{patch}} = 5$ . We divide by the square root of this number when estimating the uncertainties in different wave-number bins.

As a trade-off to this patch-based approach, we must compute our survey volume  $V_{\text{survey}}$  self-consistently given the angular extent of one of these patches. For each redshift bin  $z_i$ , we assume the FOV of HERA and use this value to compute the comoving distance  $\chi_{\perp}$  using the angular diameter distance  $D_A(z_i)$ . For the line-of-sight component, we compute the comoving distance  $\chi_{\parallel}$  that corresponds to our 6 MHz of bandwidth. The resulting survey volume is thus  $V_{\text{survey}} = \chi_{\perp}^2 \chi_{\parallel}$ . In practice the number of nonoverlapping patches will be dictated by the joint coverage of HERA and Roman. For the forecasts considered here, we assume that it is possible to construct  $N_{\text{patch}} = 5$  fields of observation from which we estimate our S/N. The overall uncertainty is only sensitive to the square root of this quantity, and so having fewer patches does not significantly affect the overall forecast.

### 4.3. Combining Modes

As discussed above in Section 4.2, the number of times an individual mode is measured depends on the details of the survey geometry. To first order, we have approximated this counting factor using Equation (18), which gives the differential number of modes within a volume element of  $(k_{\perp}, k_{\parallel})$  space. By accounting for this factor, we can understand how the expected variance of the measurement is affected by the survey volume and the amount of sky area covered by the two different experiments. We define the bin-averaged S/N  $\hat{s}$  as:

$$\hat{s}(k_{\perp}, k_{\parallel}) \equiv \sqrt{N_{\text{patch}} dN(k_{\perp}, k_{\parallel})} \frac{P_{21 \times \text{gal}}(k_{\perp}, k_{\parallel})}{\sigma_{21 \times \text{gal}}(k_{\perp}, k_{\parallel})}, \quad (19)$$

<sup>14</sup> If one uses linear bins of  $dk_{\perp}$  and  $dk_{\parallel}$  instead of logarithmic ones, the number of modes is given by:

$$dN(k_{\perp}, k_{\parallel}) = \frac{k_{\perp}^2 V_{\text{survey}}}{(2\pi)^2} dk_{\perp} dk_{\parallel},$$

where we have applied similar considerations for mode counting only in the positive half-plane and  $k_{\parallel}$  symmetry.

where  $P_{21 \times \text{gal}}$  is measured empirically from our simulated volumes, and  $\sigma_{21 \times \text{gal}}$  is computed from Equation (15). For the forecasts here, we choose logarithmic bins of fixed extent  $d \ln k_{\perp} = d \ln k_{\parallel} = 0.1$ , though we have verified that the overall results are relatively insensitive to the precise choice of binning scheme.

Figure 12 shows the per-mode S/N ratio as a function of  $k_{\perp}$  and  $k_{\parallel}$  of the cross-spectrum  $P_{21, \text{gal}}$  for our fiducial reionization scenario. The uncertainty  $\sigma_{21 \times \text{gal}}$  is given by Equation (15), and the signal  $P_{21 \times \text{gal}}$  is shown in Figure 10. For the uncertainty calculation, we use the parameters  $f_{\text{LAE}} = 0.1$  and  $\sigma_z = 0.01$ . When the mode-sampling in Equation (18) is applied, we see that the significance of several modes is  $\hat{s} > 1$ . Also consistent with Figure 11, the most significant modes lie in the regions of  $k$ -space that have small values of  $k_{\perp}$  and  $k_{\parallel}$ . We also show the values of wedge corresponding to the horizon (solid),  $m = 1$  (dashed) and  $m = 0.5$  (dotted). In principle there is a significant amount of sensitivity that can be obtained by increasing the number of foreground modes that can be used, though as shown below, a significant detection is possible even if restricted to modes that are expected to be free of foreground contamination.

With these individual uncertainties defined in Equation (19), we compute the cumulative S/N by adding the per-bin sensitivity  $\hat{s}$  for individual  $(k_{\perp}, k_{\parallel})$  bins in quadrature:

$$\left(\frac{S}{N}\right)^2 = \sum_{k_{\perp}, k_{\parallel}, z} \hat{s}^2(k_{\perp}, k_{\parallel}, z). \quad (20)$$

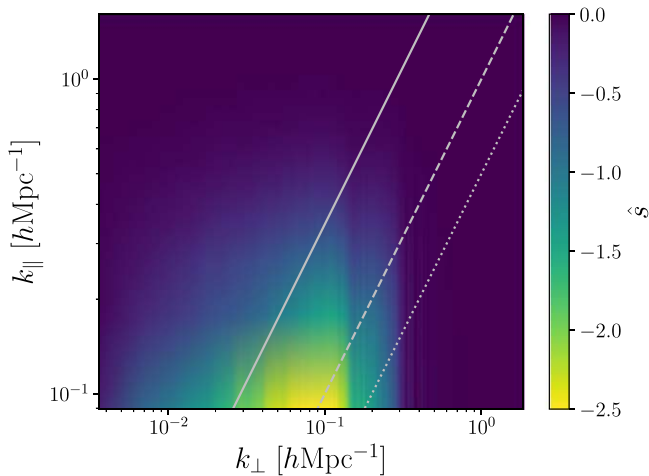
In this way, we can build up cumulative sensitivity by combining measurements from different modes together. We also sum over all nonoverlapping redshift windows between  $7.2 \leq z \leq 12$ . This restriction on the redshift value reflects the inability of the Roman grism to observe and detect LAEs below this threshold. In practice, given the low number density of galaxies expected above  $z \gtrsim 10$ , the sensitivity of these measurements decreases significantly toward high redshift.

To capture how the total uncertainty varies with changes made to individual experimental parameters, we explore a multidimensional parameter space of statistical errors.

Next, we consider how the error bars on the cross-power spectrum vary with changes in both the survey and model parameters. Specifically, we explore how the S/N varies with respect to: (i) the slope of the 21 cm wedge (i.e., the degree to which foregrounds contaminate the signal of interest), (ii) the galaxy redshift uncertainties,  $\sigma_z$ , (iii) the effective duty cycle of sources,  $f_{\text{LAE}}$ , (iv) whether the spin temperature of the IGM is saturated or maximally cold, (v) the ionization history, and (vi) the galaxy bias factor predicted by ARES and BLUETIDES. Exploring these variations will help in formulating optimal observing strategies for upcoming cross-power spectrum measurements.

Table 1 shows the cumulative S/N calculated using Equation (20) for our various combinations of parameters. For our default parameters, we choose a horizon wedge level of foreground contamination,  $f_{\text{LAE}} = 0.1$ ,  $\sigma_z = 0.01$ ,<sup>15</sup> a spin-temperature saturated IGM, our fiducial reionization scenario, and the galaxy bias  $b_g(z)$  as predicted by ARES. We show the resulting changes to the S/N as a function of varying these

<sup>15</sup> This value of  $\sigma_z$  is consistent with the redshift-dependent scaling of  $\sigma_z(z) = 0.001(1+z)$  mentioned above (Wang et al. 2022).



**Figure 12.** The S/N as a function of  $k_{\perp}$  and  $k_{\parallel}$  of the  $P_{21,\text{gal}}$  cross-spectrum, with the uncertainty  $\sigma_{21,\text{gal}}$  given by Equation (15). As in Figure 10, we show the horizon wedge (solid),  $m = 1$  (dashed), and  $m = 0.5$  (dotted) foreground contamination lines. The plotted noise levels assume a photometric redshift uncertainty of  $\sigma_z = 0.01$  and an effective LAE observation fraction of  $f_{\text{LAE}} = 0.1$ . For several modes, the S/N is greater than 1, and combining the individual uncertainties can lead to even more significant detection.

parameters. Encouragingly, given these default parameters, we forecast a  $12\sigma$  detection for the nominal HERA and Roman sensitivities. For the same combination of observational parameters, we forecast a  $282\sigma$  detection of the 21 cm auto-power spectrum from HERA and a  $14\sigma$  detection of the LAE galaxy auto-power spectrum from Roman.<sup>16</sup> We discuss further practical challenges in achieving this cross-power S/N in Section 4.4.

We now briefly discuss some of the general trends on display in Table 1 to contextualize some of the various observational effects at play. Not surprisingly, decreasing the slope of the foreground wedge significantly increases the cumulative S/N ratio. If a value of  $m = 0.5$  is assumed, the cumulative S/N increases by nearly a factor of 3, to a roughly  $32\sigma$  detection. As shown in Figure 12, there are a nontrivial number of modes that are subject to foreground contamination, which are also the ones most strongly detectable. By increasing the number of modes that contribute to the S/N statistic, the prospects for successful detection and characterization are improved dramatically. Furthermore, this improvement relies almost exclusively on improvements to 21 cm calibration and analysis techniques. Although the design of HERA means that reliably accessing these modes may be difficult, it is worth exploring whether there are ways to extract and use some of this information when measuring the cross-correlation spectrum.

For the redshift uncertainty  $\sigma_z$ , we see that there is asymmetric behavior in how the overall uncertainty changes. Increasing the fiducial value to  $\sigma_z = 0.001$  offers a modest improvement to  $14\sigma$ . However, decreasing the uncertainty to  $\sigma_z = 0.1$  leads to a cumulative S/N well below 1. This can be seen in the behavior of the S/N shown in the right panel of Figure 11: as  $k_{\parallel}$  increases, the galaxy noise power (and hence the cross-spectrum variance in the shot-noise dominated limit) increases exponentially. When the uncertainties are such that

$\sigma_z \sim 0.1$  (which is the expected level for photometrically determined redshift values), there are very few  $k_{\parallel}$  values that have an appreciable S/N value. It may be possible to access lower values of  $k_{\parallel}$  by increasing the bandwidth used to observe 21 cm measurements, though the foreground contamination swamps the signal for small values of  $k_{\parallel}$  (Abdurashidova et al. 2022a).

The effective LAE duty cycle  $f_{\text{LAE}}$  also has a significant effect on the projected S/N ratio. As discussed above, changing this quantity only affects the galaxy number density calculation when determining the shot-noise term of the galaxy uncertainty  $P_{\text{gal}}^{\text{noise}}$  in Equation (14). Nevertheless, this particular source of uncertainty is one of the most significant. The left panel of Figure 11 suggests that this shot-noise term dominates the galaxy spectrum uncertainty at all values of  $k_{\perp}$ , which in turn limits the sensitivity of the cross-spectrum measurement. Thus, increasing this quantity to  $f_{\text{LAE}} = 1$  significantly reduces the shot-noise uncertainty and increases the S/N by a factor of about 3. Conversely, using  $f_{\text{LAE}} = 0.01$  decreases the S/N by a similar factor.

Next, we look at the impact of a saturated versus a cold IGM. As mentioned above in Section 2.1, our default model assumes that the spin temperature of hydrogen has been saturated, so that  $T_s \gg T_{\gamma}$ . However, we also investigate the case of a “cold” IGM, in which the gas outside fully ionized regions is not heated. As a result, the 21 cm brightness temperature can be large and negative, which increases the amplitude of the 21 cm auto-power spectrum as well as the cross-spectrum of interest here. Although this assumption significantly increases the amplitude of the 21 cm auto-power spectrum, it only provides a modest increase in the cross-spectrum here (see also Heneka & Mesinger 2020), largely due to the quadratic scaling of the fluctuation amplitude in the 21 cm auto-power spectrum versus the linear scaling here. As discussed above, the 21 cm signal is primarily cosmic-variance limited for the detectable Fourier modes in the cross-power spectrum, and so the higher-amplitude fluctuations in the cold IGM boost both the signal and the noise and do not significantly improve the total S/N.

Furthermore, we investigate the effects of changing the ionization history. As can be seen in Table 1, both the short and the late reionization scenarios lead to slightly smaller projected S/N values compared to the fiducial history chosen here. For the case of the short history, the cross-power spectrum itself has a slightly larger amplitude due to a larger 21 cm signal, noted as a feature in previous applications of this seminumeric model (La Plante et al. 2014). However, this increase in signal amplitude is offset by having fewer redshift windows over which the 21 cm and galaxy signals have an appreciable cross-spectrum amplitude. For the case of the late reionization scenario, imposing the cutoff of  $z \geq 7.2$  means that spectral windows near the midpoint of reionization are thrown out. Despite the smaller amount of redshift information that can be used, there is still enough sensitivity to detect the cross-spectrum at roughly  $9.5\sigma$ .

Finally, we also look at the effect of the linearly biased galaxy model chosen for the galaxy fields. Instead of our models generated from ARES, we use bias values and number densities inferred from the BLUETIDES simulation. As can be seen in Figure 3, the galaxy properties predicted by BLUETIDES feature both a larger galaxy bias  $b_g$  and a greater galaxy number density  $n_g$ . The combination of these effects is a significantly larger predicted value for the cumulative S/N, with a value of

<sup>16</sup> Note that for the HERA and Roman auto-spectra, we have assumed sky coverage areas of  $1000 \text{ deg}^2$  and  $2200 \text{ deg}^2$ , respectively. Also note that the LAE galaxy auto-power spectrum does not include the foreground wedge excision, so the nominal S/N forecast is comparable to the cross-power spectrum.



**Table 1**  
Cumulative S/N Ratio for Different Model Parameters

Bias Model	Ionization History	IGM	$f_{\text{LAE}}$	$\sigma_z$	Wedge Cut	S/N
ARES	fiducial	saturated	0.1	0.01	Horizon	12
ARES	fiducial	saturated	0.1	0.01	1.0	28
ARES	fiducial	saturated	0.1	0.01	0.5	32
ARES	fiducial	saturated	0.1	0.001	Horizon	14
ARES	fiducial	saturated	0.1	0.1	Horizon	0.55
ARES	fiducial	saturated	1	0.01	Horizon	35
ARES	fiducial	saturated	0.01	0.01	Horizon	3.7
ARES	fiducial	cold	0.1	0.01	Horizon	15
ARES	short	saturated	0.1	0.01	Horizon	10.0
ARES	late	saturated	0.1	0.01	Horizon	9.5
BLUETIDES	fiducial	saturated	0.1	0.01	Horizon	29
BLUETIDES	fiducial	saturated	0.1	0.01	1.0	69
BLUETIDES	fiducial	saturated	0.1	0.01	0.5	81

**Note.** The cumulative S/N ratio for different combinations of model parameters and observational assumptions. Our fiducial parameters assume: the reionization history described in Section 2.1.1, a spin-temperature saturated IGM,  $f_{\text{LAE}} = 0.1$ ,  $\sigma_z = 0.01$ , and a horizon wedge level of foreground contamination. We show how the resulting S/N changes when modifying each of these parameters around these fiducial values.

$29\sigma$ . In Table 1, we also show the S/N for different values of the horizon cut. In general, the S/N for BLUETIDES is larger than the S/N for ARES by a factor of 2–3. We find that this trend holds for all of the parameter combinations we explored.

#### 4.4. Sources of Systematic Errors

In addition to the statistical uncertainties discussed above, there are sources of systematic errors both for galaxy and 21 cm uncertainties. We begin by discussing issues related to galaxy observations and potential paths for mitigation, then turn to 21 cm measurements.

##### 4.4.1. Galaxy Survey Errors

For the galaxy surveys, there is the potential for the measured redshift values from Ly $\alpha$  lines to be systematically offset from the true galaxy redshifts (e.g., Shapley et al. 2003; Erb et al. 2014; Shibuya et al. 2014; Mason et al. 2018; Endsley et al. 2022). As we show in the above analysis, accurately determined spectroscopic redshift values are key to yielding a statistically significant detection.

Fortunately, there are two potential solutions to systematic errors in the determination of galaxy redshift values. First, we may use data taken from other instruments as a means of cross-validating and calibrating the measurements from the HLSS. Specifically, the near-infrared spectrograph (Birkmann et al. 2011) aboard the James Webb Space Telescope (JWST) is expected to take high-resolution spectroscopic measurements of many high-redshift galaxies. By comparing several of the higher-quality spectra from JWST with a subset of those taken by Roman, it may be possible to model and remove any systematic uncertainties from the full catalog produced by Roman. Additionally, when endeavoring to measure the cross-correlation signal in practice, we can treat the galaxy redshift values as uncertain quantities, or even a single nuisance parameter quantifying the typical systematic redshift offset (as done in, e.g., CHIME Collaboration et al. 2022), and marginalize over them in Markov Chain Monte Carlo (MCMC)-style analysis. Essentially, the cross-correlation signal is not

significant when the incorrect values for the galaxy sample redshifts are used, but the MCMC technique yields a maximal signal as the accuracy is increased. We can use the measured redshift values to form relatively tight prior distributions for the values used in the analysis, which are made more accurate as the space is sampled.

Another source of uncertainty is the potential confusion of sources as measured by the grism, given that the instrument covers a relatively wide FOV. These sources can either be other nearby high-redshift galaxies, or interloping low-redshift galaxies or stars. Interlopers that are not high-redshift LAEs will increase the effective shot-noise of the measurement. An additional complication is the confusion of two LAE objects in the Roman survey. As discussed above in Section 2.4.2, we expect that objects observed in the HLSS will also be sufficiently bright in the HLIS so as to be well-localized. In Satpathy et al. (2020), the authors explicitly model and measure the expected number of overlapping galaxy spectra (see their Section 3.1.3). They find that about 7% of galaxies are expected to overlap. Assuming that the overlap of galaxies is not biased in some way, then this only results in an effective decrease in the number density of observed galaxies, which is accounted for in our  $f_{\text{LAE}}$  parameter. As we show, even with a pessimistic value of  $f_{\text{LAE}} = 0.01$ , we expect a significant detection of the cross-spectrum. Thus, this potential overlap should not preclude a successful measurement.

##### 4.4.2. 21 cm Survey Errors

On the 21 cm side, there are also considerations of systematic uncertainties. As discussed at great length throughout the rest of this paper, we know that the wedge and 21 cm foregrounds pose a significant challenge when making measurements and inferences. Although we have chosen a pessimistic value of a horizon cut for our 21 cm foregrounds, the amount of contamination may necessitate increasing this value even further. Previous analysis has shown that a “super-horizon buffer” may be required to fully remove the effect of foreground emission (e.g., Poher et al. 2013), which would lead

to a further decrease in the amount of  $(k_{\perp}, k_{\parallel})$  space that can be used for this cross-correlation. Furthermore, although it may be possible for HERA to produce data from inside the wedge, the primary mode of HERA data analysis thus far has been explicit foreground avoidance. The development and implementation of new techniques may be required to achieve levels of foreground cleaning required to improve the prospects for this measurement.

Separately to the issue of Fourier modes contaminated by foreground contamination per se, we have also ignored contributions of foreground signals to the cross-spectrum variance. Although we expect the foregrounds from the 21 cm data to be uncorrelated with the galaxy survey signal, they nevertheless contribute to the cross-spectrum variance. As explored above, the variance of the 21 cm signal contributes to the overall variance of the cross-power spectrum, so increasing the variance will reduce the forecast S/N values.

Another potential issue for 21 cm measurements is a decrease in the effective number of baselines measuring a given  $u$  mode on the sky. Figure 7 shows the distributions of baseline for the full HERA-350 array. As shown in Figure 11, this full distribution is expected to yield measurements that are cosmic-variance limited, rather than instrument-noise limited. However, if the effective number of baselines is decreased, this may no longer be true, which would increase the effective noise of the measurement. Fortunately, given the highly redundant nature of HERA, this baseline distribution ensures that these low- $k_{\perp}$  modes that yield the highest S/N are very well sampled, and should be cosmic-variance dominated even with a smaller number of antennas.

Another source of systematic errors that may decrease the sensitivity of HERA measurements in practice is the presence of radio frequency interference (RFI), which affects frequency ranges detectable by the instrument. Wider frequency ranges that are relatively free of RFI mean that multiple different redshift windows can be used to increase the cumulative S/N. Additionally, wider frequency ranges mean that broader FFTs can be performed, which probe smaller values of  $k_{\parallel}$ . Given that most of the sensitivity for the cross-correlation statistic comes from low values of  $k_{\parallel}$ , having access to wide frequency ranges free of RFI is appealing. Fortunately, the observed frequency ranges for HERA are relatively wide (Abdurashidova et al. 2022a), so it should be possible to construct several frequency windows for evaluating the cross-correlation S/N to make such a detection possible.

## 5. Discussion

Based on the S/N calculations above in Section 4.1, there are several important findings that have implications for various observation strategies. We now turn our attention to the proposed HLS survey, and ask about how to improve prospects for this cross-correlation measurement. Before discussing these prospects in detail, it is worth making some high-level remarks about this trade-off.

As shown above in Section 4.1 and Figure 11, we expect the galaxy shot-noise term  $1/n_{\text{gal}}$  to be significantly larger than the intrinsic galaxy power spectrum  $P_{\text{gal}}$  on all scales, but specifically for the low- $k$  modes that contain the most sensitivity in the cross-power spectrum. If  $1/n_{\text{gal}} \gg P_{\text{gal}}$  on the scales of interest, then one is shot-noise limited, and it is better to attempt to make deeper observations so that the shot-noise term decreases. Conversely, once the observations are

sufficiently deep such that  $1/n_{\text{gal}} \sim P_{\text{gal}}$ , then one should use wider observations at this depth rather than deeper ones. We now explore some of the quantitative ramifications for the HERA–Roman cross-spectrum.

### 5.1. Increased Sky Coverage

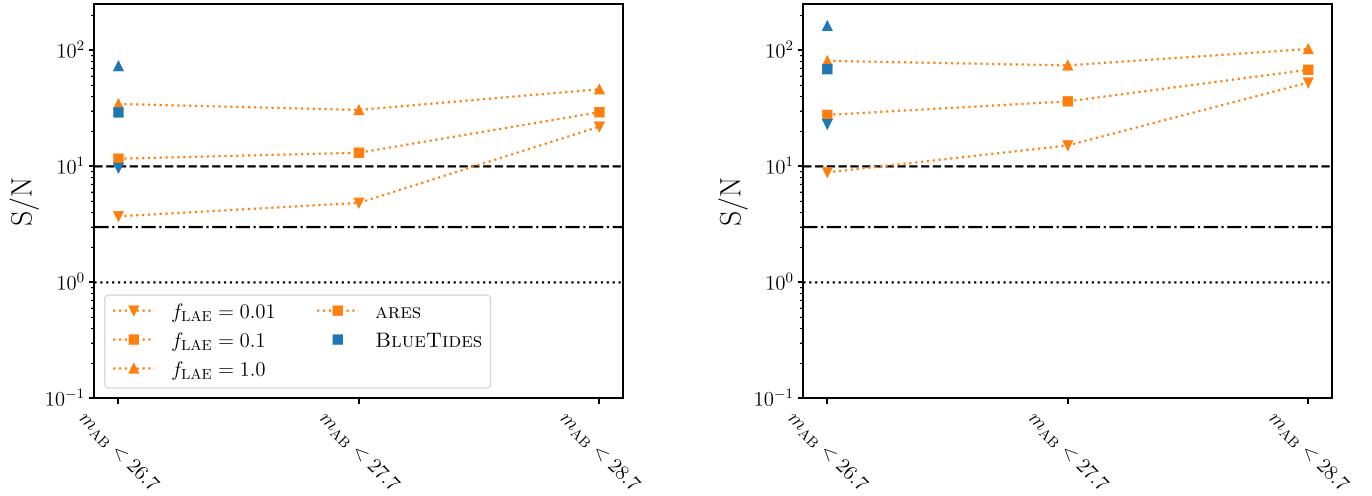
Let us first suppose that we were able to make use of additional time on the instrument, and decided to cover additional sky area (in the HERA stripe) at the same depth. Using the noise formalism outlined above, the effect of increasing the survey area when cross-correlating with HERA is related to the number of nonoverlapping sky patches that are observed  $N_{\text{patch}}$ . Specifically, by assuming that the signals in these patches average together incoherently, we increase the computed S/N value by  $\sqrt{N_{\text{patch}}}$ . In the above analysis, given that the projected overlapping area is expected to be  $500 \text{ deg}^2$  and the primary beam of HERA is about  $10^\circ$ , we have assumed that  $N_{\text{patch}} = 5$ . If we instead assume a 20% increase in sky coverage such that  $N_{\text{patch}} = 6$ , the S/N ratio increases only by about 10%. Although this gain is modest and is straightforward to model, it represents a significant increase in the amount of sky area that is jointly covered by the two instruments. As such, it may not be feasible to expand the overlapping area by such a large degree.

### 5.2. Deeper Observations

Alternatively, given additional observing time, it may instead be used to probe the existing sky coverage with greater depth. A fully self-consistent treatment of this choice is subtle, but we briefly discuss some of the high-level effects. By observing the sky for additional time, the limiting magnitude of objects in the survey increases, yielding more observed galaxies. Given that these galaxies are generally hosted in less-massive dark matter halos, the associated bias of these objects will decrease (though still remain relatively large, as shown in Figure 3), meaning the amplitude of the cross-spectrum will decrease by a small amount. However, given the relatively steep slope of the UVLF for these objects, we expect a significant increase in the number of objects observed.<sup>17</sup> As discussed above in Section 4.1, the cross-spectrum S/N forecasts are limited by shot-noise in the galaxy distribution and the sample variance (i.e., cosmic variance) in the 21 cm measurements. Thus by observing more galaxies, we would be able to reduce this contribution to the total noise budget, and increase the significance of detection.

Figure 13 makes this trade-off more concrete. We show the nominal S/N for our  $500 \text{ deg}^2$  survey as reported in Table 1 and the expected limiting magnitude of the HLIS, where the detected galaxies have  $m_{\text{AB}} < 26.7$ . For our ARES models, we are also able to self-consistently model the change in galaxy bias and number density as a function of this magnitude limit (see Figure 3). We compute the projected S/N for proposed deeper observations that are 1 mag deeper ( $m_{\text{AB}} < 27.7$ ) and 2 mag deeper ( $m_{\text{AB}} < 28.7$ ). To reflect the fact that a smaller sky area is probed, we set  $N_{\text{patch}} = 1$  (i.e., this deeper survey only covers a single patch of  $100 \text{ deg}^2$ ). We can see that going 1 mag deeper leads to a modest increase in projected S/N, from

<sup>17</sup> Note that a much more significant increase in survey depth than we consider here might have quickly diminishing returns, because many newly detected faint galaxies are more likely to be in small bubbles, which impede the transmission of Ly $\alpha$  and as a result, our ability to detect the galaxies in Ly $\alpha$ .



**Figure 13.** The projected S/N as a function of the limiting magnitude  $m_{AB}$  for different combinations of observational parameters. Left: the forecast S/N for a horizon wedge. Right: the S/N for  $m = 1.0$ . We show different values of  $f_{LAE}$  using different symbols, and use color to denote ARES vs. BLUETIDES galaxy models (the latter of which are only shown for  $m_{AB} < 26.7$ ). For our assumed default parameters, 1 magnitude of deeper observations corresponds to a slight increase of the S/N, from  $12\sigma$ – $13\sigma$ . Note that when considering the deeper observations, we use a sky coverage of  $100 \text{ deg}^2$  rather than  $500 \text{ deg}^2$ . Going 2 magnitudes deeper corresponds to a further increase to an S/N of  $29\sigma$ . See Section 5 for more discussion.

$12\sigma$ – $13\sigma$  for the default set of observing parameters, even at the expense of sky coverage. However, going 2 mag deeper leads to a significant increase in the projected S/N, to  $29\sigma$ .

We note also that in principle, these deeper photometric surveys also require deeper grism coverage in the accompanying areas. As shown in Figure 4, going 2 mag deeper in imaging requires roughly five times deeper spectroscopy to observe all of these sources in Ly $\alpha$ . Other studies have considered other possible deep field options (e.g., Drakos et al. 2022), such as covering  $10 \text{ deg}^2$  with relatively deep photometry and spectroscopy. In this case, we project a detection of  $4.0\sigma$  for a limiting magnitude of  $m_{AB} < 27.7$  (i.e., 1 mag deeper than the nominal HLIS), and a  $9.0\sigma$  detection for  $m_{AB} < 28.7$ .

## 6. Conclusion

In this paper, we explore the prospects for the 21 cm galaxy cross-spectrum for upcoming measurements.<sup>18</sup> We show that such a signal is nominally detectable for projected measurements for HERA and the Roman Space Telescope, with a forecasted sensitivity of  $12\sigma$  even under pessimistic assumptions regarding the foreground wedge, i.e., the region of Fourier space unusable owing to foreground contamination. We find that the accuracy of spectroscopically determined redshift values  $\sigma_z$  does not dramatically impact the sensitivity of the measurement, though photometrically determined redshifts will not produce a statistically significant detection. In our fiducial forecast, we assume an effective duty cycle of 10% for the LAE objects detected by the HLS, which has a significant impact on the forecasted sensitivity. If instead the effective duty cycle is 1%, a significant detection is still possible, though this lower significance can be mitigated with a deeper galaxy survey, less 21 cm foreground contamination, or an intrinsically larger galaxy bias consistent with the BLUETIDES predictions. We also explore the impact of varying the reionization history used in our simulations, and find that the

effect is not very significant. The one caveat is for sufficiently late reionization histories (where the midpoint of reionization is  $z \lesssim 7$ ), since the grism on Roman cannot detect Ly $\alpha$  in objects at  $z \lesssim 7.2$ .

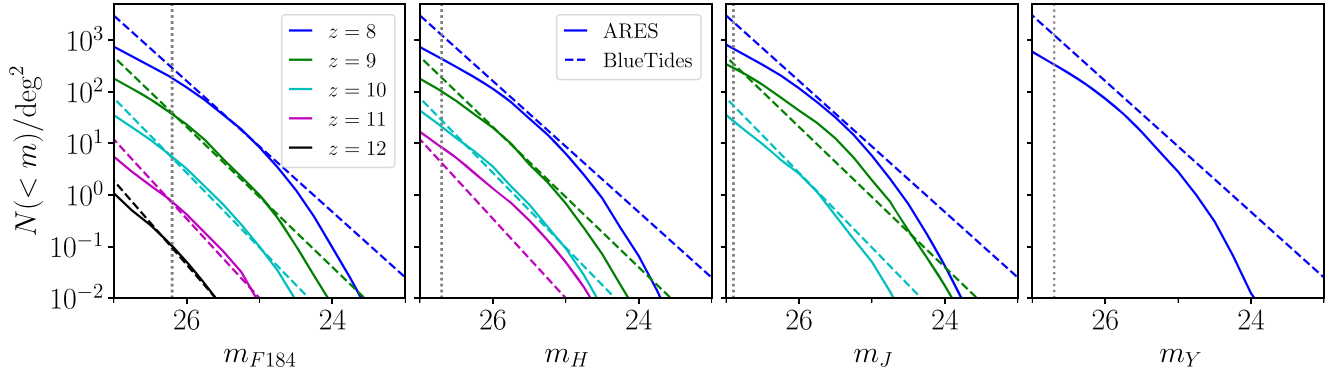
Given the forecast sensitivity for HERA, we expect the 21 cm measurements to be cosmic-variance dominated for the modes that contribute most to the cross-power spectrum. As such, there is not much additional statistical advantage to be gained from next-generation experiments such as the SKA. However, if the SKA allows for the recovery of the signal that is lost to foreground contamination, then the significance of detection may improve dramatically. Indeed, with the projected sensitivity levels, it may be possible to characterize the ionization history, rather than make a simple detection, by dividing the cumulative sensitivity into nonoverlapping redshift windows. Additionally, by using additional  $k$ -bins, it may be possible to constrain the size of ionized bubbles during the EoR. However, we leave a detailed investigation of this possibility for future work.

For the galaxy measurements, we find that the relatively low number of galaxies expected to be measured at high redshift has the largest impact on the overall sensitivity of the measurement. As such, increased observation time of the HLS patch may prove useful for making a more robust measurement. When weighing the potential trade-off between covering more sky area versus performing a deeper survey of the existing footprint, we conclude that a deeper survey is more beneficial. This strategy would lead to a larger number of observed galaxies, thereby decreasing the shot-noise contribution to the overall uncertainty. As this component is the most significant factor, taking steps to decrease it has the largest impact on the overall sensitivity.

We note that our modeling methods in this paper are relatively simple, and do not capture the full interplay between the 21 cm signal and high-redshift galaxies. In particular, there is no explicit link between the galaxies and the resulting ionization field. Additionally, we have modeled LAEs using a constant factor  $f_{LAE}$  to account for the fraction of intrinsic LAEs actually observed by Roman. In reality, the observable LAEs will have some spatial and luminosity dependence

<sup>18</sup> Those interested in understanding more of the implementation details can look at a GitHub repository containing key parts of the calculation: [https://github.com/plaplant/21cm\\_gal\\_cross\\_correlation](https://github.com/plaplant/21cm_gal_cross_correlation).





**Figure 14.** Comparison of ARES and BLUETIDES predictions for galaxy surface densities in the nominal four HLS filters, from reddest to bluest (left to right). The vertical line in each panel indicates the nominal HLS magnitude limit in that band.

primarily owing to absorption from neutral regions in the IGM. In the future, it would be beneficial to employ a more self-consistent model. This is of course challenging and will be more computationally expensive than the approach we take here, but will be vital for unbiased inference.

Finally, when looking toward future 21 cm experiments, it may be useful to consider how projections for cross-correlation measurements such as the ones presented in this paper may provide guidance for array design and construction. When seeking to confirm or independently verify astrophysical and cosmological parameters measured from 21 cm experiments alone, cross-correlations may provide a key path forward for providing confidence in auto-power spectra.

We thank Steven Furlanetto, Adrian Liu, Andrei Mesinger, and Steven Murray for illuminating discussions regarding this work. We also thank Ronniy Joseph for assistance with the MWA footprints, and Henry Gebhardt for providing the Roman HLS footprint, which was generated by Chris Hirata, projected by Dida Markovic, and is a product of the Roman SIT. This material is based upon work supported by the National Science Foundation under grant Nos. 1636646 and 2206602, the Gordon and Betty Moore Foundation through grant GBMF5215 to the Massachusetts Institute of Technology, and institutional support from the HERA collaboration partners. HERA is hosted by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. A.G.’s work is supported by the McGill Astrophysics Fellowship funded by the Trottier Chair in Astrophysics as well as the Canada 150 Programme and the Canadian Institute for Advanced Research (CIFAR) Azrieli Global Scholars program. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant No. ACI-1548562. Specifically, it used the Bridges-2 system, which is supported by NSF award No. ACI-1928147, at the Pittsburgh Supercomputing Center (PSC; Towns et al. 2014; Brown et al. 2020).

## Appendix A Comparison with BLUETIDES

In this section, we show a more detailed comparison between our fiducial model (solid) and the BLUETIDES (dashed) predictions. First, in Figure 14, we compare predictions for the cumulative surface density of galaxies detected in each

HLS filter. Each panel from left to right indicates a different filter, from the reddest F184 filter to the Y band, which straddles  $\lambda \simeq 1 \mu\text{m}$ . The magnitude limit is slightly different from band to band, as indicated by the vertical lines. Overall, agreement is fairly good, particularly in the F184 and H bands. The greatest discrepancies are at the factor of  $\sim 2$ – $3$  level at  $z = 8$  for  $H$ ,  $J$ , and  $Y$  bands. This level of disagreement is not unexpected given that, e.g., the UVLFs from BLUETIDES are not identical to ours (see Figure 2).

Though galaxy abundances are fairly similar between models, as shown in Figure 3, our predictions for the bias factors are quite different. At  $z = 8$ , BLUETIDES predicts a linear bias of  $b \sim 13$  for HLS galaxies, while our fiducial ARES model predicts lower values of  $b \simeq 8$ . This difference is dominated by the difference in the stellar mass–halo mass relations—at fixed halo mass, BLUETIDES predicts a lower stellar mass. As a consequence, galaxies of a fixed magnitude will be hosted in higher mass halos, meaning they will be more biased.

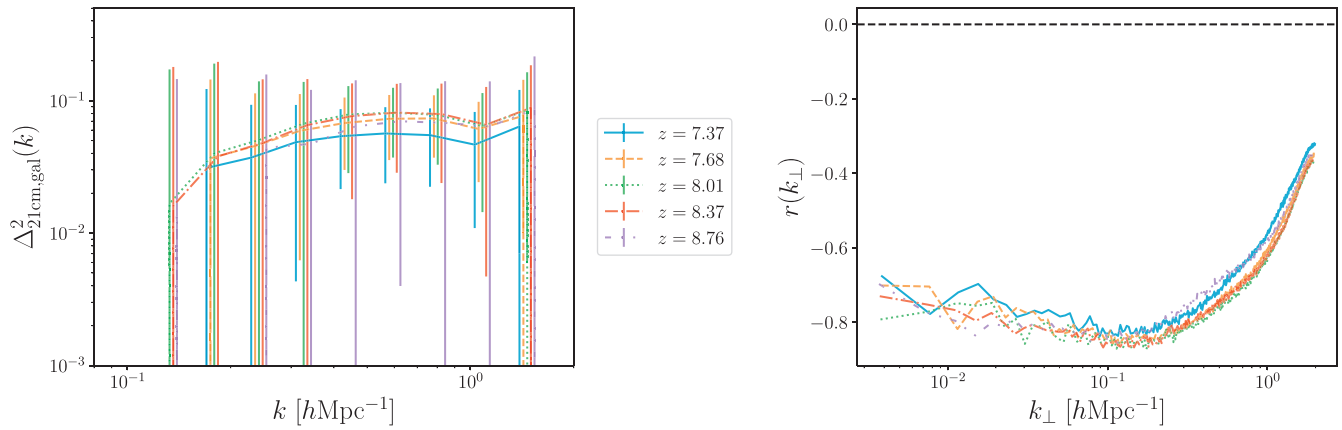
## Appendix B Redshift Evolution of the Signal

To form a more quantitative picture of how the signal of interest evolves with redshift, we include an exploration of two useful metrics. First we explore the spherical power spectrum  $\Delta^2(k) = k^3 P(k) / 2\pi^2$ . This quantity is useful to understand because previous studies (e.g., Lidz et al. 2009) have looked almost exclusively at this quantity rather than the cylindrical power spectra that we consider. We compute the spherical power spectrum  $P(k)$  from our cylindrical power spectra at different redshifts. Given our significantly reduced extent along the  $k_{\parallel}$  dimension, there is a relatively restricted range of  $k$  values that we are in principle sensitive to. Nevertheless, this quantity is useful for building intuition.

As another measure of the behavior of the signal, we compute the cross-correlation coefficient  $r(k_{\perp}, k_{\parallel})$ . As with the power spectrum in the main portion of the paper, we compute this quantity using cylindrical coordinates. Quantitatively,  $r$  can be expressed as:

$$r(k_{\perp}, k_{\parallel}) = \frac{P_{21 \times \delta_g}(k_{\perp}, k_{\parallel})}{\sqrt{P_{21,21}(k_{\perp}, k_{\parallel}) P_{\delta_g, \delta_g}(k_{\perp}, k_{\parallel})}}. \quad (\text{B1})$$

The normalization ensures that  $r \in [-1, 1]$ . A value of  $r = 1$  means the two fields are perfectly correlated, and a value of  $r = -1$  is perfectly anticorrelated. Given the physical picture



**Figure 15.** Redshift evolution of the cross-power spectrum. Left: the absolute value of the spherical power spectrum  $\Delta^2(k)$  for several different redshift bins. The lines are horizontally offset from each other for visual clarity. We show the  $1\sigma$  error bars for each bin. The corresponding ionization fraction values are  $x_i = \{0.63, 0.53, 0.43, 0.33, 0.25\}$ . Right: the cross-correlation coefficient  $r(k_\perp)$  defined in Equation (B1) for the same redshift values. As can be seen, the fields are highly anticorrelated on all scales, though the anticorrelation decreases on small scales ( $k_\perp \gtrsim 1 \text{ h Mpc}^{-1}$ ).

of reionization, we expect these quantities to generally be anticorrelated: the galaxy signal comes from regions of high density, and the 21 cm signal during the bulk of reionization comes from regions of low density.

Figure 15 shows the spherical power spectrum  $\Delta^2(k)$  and the cross-correlation coefficient  $r(k_\perp)$  for a fixed value of  $k_\parallel = 0.1 \text{ h Mpc}^{-1}$ . We also show the  $1\sigma$  error bars, where we have summed in quadrature the uncertainty described by Equation (15) for different  $(k_\perp, k_\parallel)$  modes that correspond to the same  $k$  mode. The cross-correlation signal evolves relatively slowly over this interval. On large scales, the signal peaks near the midpoint of reionization, which is consistent with the behavior of the 21 cm signal (Lidz et al. 2008). Interestingly, the error is smallest at redshifts slightly past the midpoint of reionization. This is a reflection of the fact that the expected number of observed galaxies is larger, meaning the galaxy shot-noise component is smaller. For the cross-correlation coefficient  $r(k_\perp, k_\parallel)$ , the anticorrelation between the fields is generally quite large on large scales, though not quite perfectly anticorrelated. The decrease toward small scales suggests there is no longer a strong correlation between the galaxy and 21 cm fields. This result further confirms the desire to work on relatively large scales, such as those probed by HERA and Roman.

### ORCID iDs

Paul La Plante <https://orcid.org/0000-0002-4693-0102>  
 Jordan Mirocha <https://orcid.org/0000-0002-8802-5581>  
 Adélie Gorcé <https://orcid.org/0000-0002-1712-737X>  
 Adam Lidz <https://orcid.org/0000-0002-3950-9598>  
 Aaron Parsons <https://orcid.org/0000-0002-5400-8097>

### References

- Abdurashidova, Z., Aguirre, J. E., Alexander, P., et al. 2022a, *ApJ*, **925**, 221  
 Abdurashidova, Z., Aguirre, J. E., Alexander, P., et al. 2022b, *ApJ*, **924**, 51  
 Battaglia, N., Trac, H., Cen, R., & Loeb, A. 2013, *ApJ*, **776**, 81  
 Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143  
 Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013a, *ApJ*, **770**, 57  
 Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013b, *ApJL*, **762**, L31  
 Birkmann, S. M., Böker, T., Ferruit, P., et al. 2011, *Proc. SPIE*, **8150**, 81500B  
 Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, *ApJ*, **793**, 115  
 Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, **803**, 34  
 Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, *PASA*, **30**, e031  
 Brown, S. T., Buitrago, P., Hanna, E., et al. 2020, Practice and Experience in Advanced Research Computing, PEARC '21 (New York: Association for Computing Machinery),  
 Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, *Natur*, **466**, 463  
 CHIME Collaboration, Amiri, M., Bandura, K., et al. 2022, arXiv:2202.01242  
 Datta, A., Bowman, J. D., & Carilli, C. L. 2010, *ApJ*, **724**, 526  
 De Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., et al. 2008, *MNRAS*, **388**, 247  
 DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, *PASP*, **129**, 045001  
 Doré, O., Hirata, C., Wang, Y., et al. 2018, arXiv:1804.03628  
 Drakos, N. E., Villaseñor, B., Robertson, B. E., et al. 2022, *ApJ*, **926**, 194  
 Eldridge, J. J., & Stanway, E. R. 2009, *MNRAS*, **400**, 1019  
 Endsley, R., Stark, D. P., Bouwens, R. J., et al. 2022, *MNRAS*, **517**, 5642  
 Erb, D. K., Steidel, C. C., Trainor, R. F., et al. 2014, *ApJ*, **795**, 33  
 Feng, Y., Di Matteo, T., Croft, R., et al. 2015, *ApJL*, **808**, L17  
 Feng, Y., Di Matteo, T., Croft, R., & Khandai, N. 2014, *MNRAS*, **440**, 1865  
 Feng, Y., Di-Matteo, T., Croft, R. A., et al. 2016, *MNRAS*, **455**, 2778  
 Finkelstein, S. L. 2016, *PASA*, **33**, e037  
 Furlanetto, S. R., & Lidz, A. 2007, *ApJ*, **660**, 1030  
 Furlanetto, S. R., Mirocha, J., Mebane, R. H., & Sun, G. 2017, *MNRAS*, **472**, 1576  
 Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *PhR*, **433**, 181  
 Garel, T., Blaizot, J., Rosdahl, J., et al. 2021, *MNRAS*, **504**, 1902  
 Hallinan, G., Bourke, S., Anderson, M., et al. 2015, AAS Meeting Abstracts, **225**, 328.01  
 Heneka, C., & Mesinger, A. 2020, *MNRAS*, **496**, 581  
 Iliev, I. T., Mellema, G., Pen, U.-L., et al. 2006, *MNRAS*, **369**, 1625  
 Konno, A., Ouchi, M., Shibuya, T., et al. 2018, *PASJ*, **70**, S16  
 La Plante, P., Battaglia, N., Natarajan, A., et al. 2014, *ApJ*, **789**, 31  
 La Plante, P., Lidz, A., Aguirre, J., & Kohn, S. 2020, *ApJ*, **899**, 40  
 La Plante, P., & Ntampaka, M. 2019, *ApJ*, **880**, 110  
 Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473  
 Lewis, J. S. W., Ocvirk, P., Sorce, J. G., et al. 2022, *MNRAS*, **516**, 3389  
 Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, *ApJ*, **690**, 252  
 Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., & Hernquist, L. 2008, *ApJ*, **680**, 962  
 Liu, A., & Shaw, J. R. 2020, *PASP*, **132**, 062001  
 Ma, X., Hopkins, P. F., Garrison-Kimmel, S., et al. 2018, *MNRAS*, **478**, 1694  
 Madau, P., Meiksin, A., & Rees, M. J. 1997, *ApJ*, **475**, 429  
 Mason, C. A., Treu, T., de Barros, S., et al. 2018, *ApJL*, **857**, L11  
 Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, *ApJL*, **763**, L20  
 Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *ApJ*, **521**, 64  
 Mirocha, J., Furlanetto, S. R., & Sun, G. 2017, *MNRAS*, **464**, 1365  
 Mirocha, J., La Plante, P., & Liu, A. 2021, *MNRAS*, **507**, 3872  
 Mirocha, J., Mason, C., & Stark, D. P. 2020, *MNRAS*, **498**, 2645  
 Morales, M. F., & Wyithe, J. S. B. 2010, *ARA&A*, **48**, 127  
 Moster, B. P., Naab, T., & White, S. D. M. 2018, *MNRAS*, **477**, 1822  
 Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, *MNRAS*, **463**, 1462  
 O'Shea, B. W., Wise, J. H., Xu, H., & Norman, M. L. 2015, *ApJL*, **807**, L12  
 Ouchi, M., Harikane, Y., Shibuya, T., et al. 2018, *PASJ*, **70**, S13  
 Pallottini, A., Ferrara, A., Gallerani, S., et al. 2022, *MNRAS*, **513**, 5621  
 Park, J., Mesinger, A., Greig, B., & Gillet, N. 2019, *MNRAS*, **484**, 933

- Parsons, A., Poher, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012, *ApJ*, **753**, 81
- Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, *ApJ*, **788**, 106
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *A&A*, **641**, A6
- Poher, J. C., Parsons, A. R., Aguirre, J. E., et al. 2013, *ApJL*, **768**, L36
- Satpathy, S., An, Z., Croft, R. A. C., et al. 2020, *MNRAS*, **498**, 4955
- Scoccimarro, R. 1998, *MNRAS*, **299**, 1097
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, *ApJ*, **588**, 65
- Shibuya, T., Ouchi, M., Nakajima, K., et al. 2014, *ApJ*, **788**, 74
- Smith, A., Kannan, R., Garaldi, E., et al. 2022, *MNRAS*, **512**, 3243
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, *ApJL*, **462**, L17
- Sun, G., & Furlanetto, S. R. 2016, *MNRAS*, **460**, 417
- Sun, G., Mirocha, J., Mebane, R. H., & Furlanetto, S. R. 2021, *MNRAS*, **508**, 1954
- Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, *MNRAS*, **434**, L46
- Tacchella, S., Bose, S., Conroy, C., Eisenstein, D. J., & Johnson, B. D. 2018, *ApJ*, **868**, 92
- The HERA Collaboration, Abdurashidova, Z., Adams, T., et al. 2022, arXiv:2210.04912
- Thyagarajan, N., Jacobs, D. C., Bowman, J. D., et al. 2015, *ApJL*, **807**, L28
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, *ApJ*, **688**, 709
- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, *ApJ*, **724**, 878
- Towns, J., Cockerill, T., Dahan, M., et al. 2014, *CSE*, **16**, 62
- Trac, H., & Cen, R. 2007, *ApJ*, **671**, 1
- Trott, C. M., Jordan, C. H., Line, J. L. B., et al. 2021, *MNRAS*, **507**, 772
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, **556**, A2
- Vrbanec, D., Ciardi, B., Jelić, V., et al. 2020, *MNRAS*, **492**, 4952
- Wang, Y., Zhai, Z., Alavi, A., et al. 2022, *ApJ*, **928**, 1
- Waters, D., Di Matteo, T., Feng, Y., Wilkins, S. M., & Croft, R. A. C. 2016, *MNRAS*, **463**, 3520
- Wolz, L., Pourtsidou, A., Masui, K. W., et al. 2022, *MNRAS*, **510**, 3495
- Wyithe, J. S. B., & Loeb, A. 2007, *MNRAS*, **375**, 1034
- Xu, H., Wise, J. H., Norman, M. L., Ahn, K., & O'Shea, B. W. 2016, *ApJ*, **833**, 84