Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **536,** 1645–1659 (2025) https://doi.org/10.1093/mnras/stae2684
Advance Access publication 2024 December 04

Reionization relics in the cross-correlation between the Lyα forest and 21 cm intensity mapping in the post-reionization era

Paulo Montero-Camacho [1,2]★ Catalina Morales-Gutiérrez,[3,4] Yao Zhang,[2] Heyang Long [5,6] and Yi Mao [2]

[1]*Department of Mathematics and Theory, Peng Cheng Laboratory, Shenzhen, Guangdong 518066, China*
[2]*Department of Astronomy, Tsinghua University, Beijing 100084, China*
[3]*Department of Physics, University of Costa Rica, 11501 San José, Costa Rica*
[4]*Space Research Center (CINESPA), University of Costa Rica, 11501 San José, Costa Rica*
[5]*Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[6]*Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*

Accepted 2024 November 25. Received 2024 November 18; in original form 2024 September 26

ABSTRACT

The tumultuous effects of ultraviolet photons that source cosmic reionization, the subsequent compression and shock-heating of low-density regions, and the modulation of baryons in shallow potential wells induced by the passage of ionization fronts, collectively introduce perturbations to the evolution of the intergalactic medium in the post-reionization era. These enduring fluctuations persist deep into the post-reionization era, casting a challenge upon precision cosmology endeavours targeting tracers in this cosmic era. Simultaneously, these relics from reionization also present a unique opportunity to glean insights into the astrophysics that govern the epoch of reionization. In this work, we propose a first study of the cross-correlation of Lyα forest and 21 cm intensity mapping, accounting for the repercussions of inhomogeneous reionization in the post-reionization era. We investigate the ability of Square Kilometre Array (SKA) × Dark Energy Spectroscopic Instrument (DESI)-like, SKA × MUltiplexed Survey Telescope (MUST)-like, and Packed Ultra-wideband Mapping Array (PUMA) × MUST-like instrumental set-ups to achieve a high-signal-to-noise ratio (SNR) in the redshift range $3.5 \leq z \leq 4$. Moreover, we assess how alterations in integration time, survey area, and reionization scenarios impact the SNR. Furthermore, we forecast the cross-correlation's potential to constrain cosmological parameters under varying assumptions: considering or disregarding reionization relics, marginalizing over reionization astrophysics, and assuming perfect knowledge of reionization. Notably, our findings underscore the remarkable capability of a futuristic PUMA × MUST-like set-up, with a modest 100-h integration time over a 100 deg^2 survey, to constrain the ionization efficiency error to $\sigma_\zeta = 3.42$.

Key words: intergalactic medium – dark ages, reionization, first stars.

1 INTRODUCTION

Cosmic reionization, the transformative phase during which our Universe shifted from pre-dominantly neutral to highly ionized, is estimated to have occurred approximately around redshift $z \sim 8$ (e.g. Planck Collaboration VI 2020). While the overarching mechanisms driving cosmic reionization are broadly understood (Haiman 2016; Furlanetto 2019), the lack of direct observations introduces significant uncertainties, particularly regarding the timeline of reionization (Keating et al. 2020; Gnedin 2022; Jin et al. 2023; Roth et al. 2024). With the advent of the *JWST* (Gardner et al. 2006) our knowledge of the sources of ultraviolet photons responsible for reionization is likely to increase dramatically. Moreover, reionization unfolds as a markedly inhomogeneous inside–out process (Lee et al. 2008; Choudhury, Haehnelt & Regan 2009) wherein denser regions undergo ionization first since sources of ultraviolet photons are expected to pre-dominantly be situated within these dense environments.

The post-reionization era is a treasure trove of cosmological information. The relatively high redshifts characteristic of this period not only facilitate the exploration of dark matter candidates (Palanque-Delabrouille et al. 2020; Puchwein et al. 2022) but also offer a unique opportunity to study cosmic evolution and structure formation (Visbal, Loeb & Wyithe 2009; Wyithe & Loeb 2009) before the onset of non-linearities imposes significant constraints on the observational landscape.

The Lyman α (Lyα) forest, the absorption features observed in the spectra of background quasars, stands as a pivotal tool for exploring the Universe in the post-reionization era. Its applications extend to the study of H I and He II reionization (Cen et al. 2009; Montero-Camacho & Mao 2020; Upton Sanderbeck & Bird 2020), the investigation of the matter power spectrum on scales beyond the reach of galaxy surveys (Weinberg et al. 2003), constrain the evolution of the Universe (du Mas des Bourboux et al. 2020), and the inference of cosmological parameters (Chabanier et al. 2019).

★ E-mail: pmontero@pcl.ac.cn

Moreover, the forest also offers a unique window into the impact of neutrino masses (Yèche et al. 2017). However, obtaining reliable cosmological measurements of the Lyα forest at high redshifts proves challenging due to the sparse sampling of quasars (Yèche et al. 2020; Montero-Camacho & Mao 2021; Chaussidon et al. 2023), particularly regarding potential measurements of the 3D flux power spectrum.[1]

Intensity mapping (IM) involves a trade-off of sacrificing angular resolution to concentrate on the integrated emission from unresolved sources. Analogous to the Lyα forest, IM utilizing the 21 cm hyperfine transition of hydrogen presents a versatile probe capable of delving into various aspects of cosmology in the post-reionization era. It can probe H I reionization (Long et al. 2023), explore early universe cosmological parameters (e.g. primordial non-Gaussianity), and constrain the evolution of the universe and structure formation (Visbal et al. 2009; Wyithe & Loeb 2009; Bull et al. 2015; Castorina & White 2019). Additionally, it can investigate the nature of dark matter (Carucci et al. 2015) and more (Cosmic Visions 21 cm Collaboration 2018; Square Kilometre Array Cosmology Science Working Group 2020).

Nevertheless, the full potential of this observable is impeded by foregrounds, such as galactic synchrotron emission, which surpass the amplitude of the cosmological signal by several orders of magnitude (Wolz et al. 2015). Despite dedicated efforts to mitigate foreground effects (see e.g. Zuo, Chen & Mao 2023; Diao, Grumitt & Mao 2024), a conservative perspective suggests that confirming the cosmic origin of the signal might necessitate cross-correlation with another probe.[2] This cross-correlation typically considers high-z galaxies (Furlanetto & Lidz 2007; Hutter et al. 2023; La Plante et al. 2023) as the additional tracer.

The intense heating of the intergalactic medium (IGM) during cosmic reionization triggers a substantial increase in the IGM temperature, reaching a few times 10^4 K. As ionization fronts propagate, inducing a rise in temperature, low-density regions experience shocks that both heat and compress the gas. These shocks, which originate in denser regions due to the increase in Jeans mass with temperature, will push gas in minivoids to higher adiabats compared to gas in denser environments (Hirata 2018) – such as minihaloes. This high-entropy mean-density gas is then responsible for the long-lasting impact of reionization in the Lyα forest (Montero-Camacho et al. 2019; Montero-Camacho & Mao 2020), essentially constituting the long-lasting *memory of reionization* at $z \sim 2$ in the forest. Simultaneously, reionization exerts influence over the baryon abundance within a given halo (Long, Givans & Hirata 2022). The thermal kick resulting from the passage of an ionization front expels some baryons from the halo, particularly for haloes with shallow potential wells. This modulation is responsible for long-lasting reionization relics in 21 cm IM, persisting up to $z \sim 3$ (Long et al. 2023).

Therefore, both the Lyα forest and H I 21 cm intensity mapping exhibit broad-spectrum contamination originating from the imprints of cosmic reionization. This contamination is especially pronounced at large scales and high redshifts in the post-reionization era, where the lingering effects of reionization are more pronounced due to less time to dissipate the additional injected energy (Montero-Camacho & Mao 2020). Managing this systematic is not merely a challenge essential for obtaining cosmological information free from bias in the post-reionization era, but it also represents a novel opportunity to gain insights into the intricate processes governing the reionization history of the Universe (Montero-Camacho & Mao 2021).

The observational programmes of these two probes are at different developmental stages. The Lyα forest has achieved commendable signal-to-noise ratio (SNR), e.g. the extended Baryonic Spectroscopic Survey (eBOSS) (Castorina & White 2019; du Mas des Bourboux et al. 2020) and Dark Energy Spectroscopic Instrument (DESI) early (Gordon et al. 2023; Ramírez-Pérez et al. 2023; Ravoux et al. 2023) and year-1 results (DESI Collaboration 2024). In stark contrast, the 2-point functions of the cosmological 21 cm signal at higher redshifts ($z > 2$) are currently constrained only by upper limits (see e.g. Mertens et al. 2020; Abdurashidova et al. 2022; Munshi et al. 2024). Despite this, because of the inherent challenges in measuring the 21 cm autopower spectrum, there is considerable interest in the cross-correlation of these two distinct tracers as a promising avenue for cosmological studies.

The Lyα forest $\times$ H I 21 cm IM cross-correlation was initially proposed in Guha Sarkar et al. (2011) as a robust and independent probe of the post-reionization IGM. Carucci, Villaescusa-Navarro & Viel (2017) underscored the reduced sensitivity of this cross-correlation to foreground contamination, emphasizing its potential to untangle degeneracies inherent in modelling parameters. Furthermore, the Ly$\alpha \times$ 21 cm cross-correlation exhibits promising prospects for constraining dark energy (Dash & Guha Sarkar 2021) and $f(r)$ gravity models, particularly when coupled with cross-correlations involving cosmic microwave background (CMB) lensing (Dash, Sarkar & Sarkar 2023). Moreover, this observable has also emerged as a potential avenue for constraining the nature of dark matter (Sarkar, Pal & Guha Sarkar 2019).

Regarding the observability of this cosmological probe, the Owens Valley Widefield Array (OWFA) (Ali & Bharadwaj 2014; Bharadwaj, Sarkar & Ali 2015) and a spectroscopic instrument such as BOSS/eBOSS were anticipated to achieve a robust 6σ detection at $z = 3.35$ with 200 h of integration (Sarkar, Bharadwaj & Guha Sarkar 2018). Meanwhile, the Square Kilometre Array Phase 1 Mid-frequency (SKA1-Mid) in conjunction with an eBOSS-like survey could reach a peak SNR of 15 at $z = 2.5$ (Guha Sarkar & Datta 2015).

However, the existing studies/forecasts have yet to incorporate the enduring impact of inhomogeneous reionization on the post-reionization IGM. In this work, we delve into the repercussions of these reionization remnants in the Lyα forest $\times$ H I 21 cm IM cross-correlation within the redshift range $3.5 \leq z \leq 4$. We specifically target this range because of the heightened abundance of Lyα spectra compared to higher redshifts and the pronounced strength of reionization relics in the forest during this period (Montero-Camacho & Mao 2020). Nevertheless, it is crucial to note that the impact of reionization imprints will remain significant for this cross-correlation in the broader redshift range of $3 \lesssim z < 6$.

The rest of this work is organized as follows. In Section 2, we outline our model for capturing the impact of reionization in the post-reionization IGM. We present the simulations required to compute the reionization relics in Section 3. The instrumental configurations considered in this study are introduced in Section 4. Section 5 presents our findings regarding the effectiveness of the instrumental set-ups in detecting the Ly$\alpha \times$ 21 cm cross-correlation. Additionally, we consider deviations in survey design and strategy,

[1]Note that with the current generation of available instruments, a significant increase in the number of high-redshift quasars is occurring (Yang et al. 2023); however, despite these developments, achieving the statistical density required to reduce the mean separation between line of sights for a broad-band 3D power spectrum analysis remains a formidable task that may be achieved with the current generation of available instruments – see Karim et al. (2024) for recent efforts in this direction

[2]Notably, the CHIME collaboration recently detected cosmological 21 cm emission at $z \approx 1$ through cross-correlation with eBOSS tracers (Amiri et al. 2023).

providing insights into possible gains. In Section 6, we demonstrate the importance of accounting for this novel effect through three distinct Fisher forecasts. Finally, we summarized our findings in 7.

Throughout this work, we use $h = 0.6774$, $\Omega_b h^2 = 0.0223$, $\Omega_c h^2 = 0.1188$, $A_s = 2.148 \times 10^{-9}$, and $n_s = 0.9667$. In agreement with the '*TT* + *TE* + *EE* + *lowP* + *lensing* + *ext*' cosmology from Planck Collaboration XIII (2016). Furthermore, we use the values of $b_{\rm H\,I}$ and $\Omega_{\rm H\,I}$ reported in Crighton et al. (2015). In contrast, we obtain the bias and redshift space distortion (RSD) parameters for the forest from Arinyo-i-Prats et al. (2015).

2 MODELLING THE CROSS-CORRELATION OF LYα AND 21 CM IM WITH REIONIZATION RELICS

In this work, we consider the inclusion – at first-order – of the memory of reionization, that is the long-lasting impact of inhomogeneous reionization, in both Lyα flux fluctuations and H I 21 cm fluctuations. We examine the repercussions of this inclusion when embedded into the *traditional* cross-correlation of the Lyα forest and 21 cm intensity mapping. Mathematically,

$$P_{21,F}(\boldsymbol{k}, z) = P^{\rm Fid.}_{21,F}(\boldsymbol{k}, z) + P^{\rm Mem.}_{F}(\boldsymbol{k}, z) + P^{\rm Mem.}_{21}(\boldsymbol{k}, z)\,, \tag{1}$$

where the first term on the right-hand side corresponds to the conventional cross-correlation of the Lyα forest and 21cm IM without reionization relics. The second term accounts for the memory of reionization present in the Lyα forest, originating in the underdense regions. The final term represents the memory of reionization in H I 21 cm IM, i.e. the memory of reionization sourced by overdense regions. The three terms are given by (Carucci et al. 2017; Montero-Camacho et al. 2019; Long et al. 2023)

$$P^{\rm Fid.}_{21,F} = b_F(z) b_{\rm H\,I}(z)[1 + \beta_F(z)\mu^2][1 + \beta_{\rm H\,I}(z)\mu^2] P_m(k, z)\,, \tag{2}$$

$$P^{\rm Mem.}_{F} = b_{\rm H\,I}(z)[1 + \beta_{\rm H\,I}(z)\mu^2] b_\Gamma(z) P_{m,\psi}(k, z)\,, \tag{3}$$

$$P^{\rm Mem.}_{21} = b_F(z)[1 + \beta_F(z)\mu^2] P_{m,\Xi}(k, z)\,, \tag{4}$$

where b_F and β_F are the usual flux bias and redshift space distortion parameter, respectively – and similarly for 21 cm quantities. Moreover, μ is the angle with respect to the line of sight, b_Γ is the radiation bias defined in Arinyo-i-Prats et al. (2015) and Hirata (2018), and P_m is the linear matter power spectrum.

Furthermore, $P_{m,\psi}$ and $P_{m,\Xi}$ are the cross-power spectrum of matter and change of transparency of the IGM due to the impact of reionization, and the cross-power spectrum of matter and change of neutral hydrogen density induced by the passage of reionization fronts, respectively. These cross-power spectra are defined as follows (Montero-Camacho et al. 2019; Long et al. 2023)

$$P_{m,\psi}(k, z_{\rm obs}) = -\int {\rm d}z \frac{\partial \psi(z, z_{\rm obs})}{\partial z} P_{m,x_{\rm H\,I}}(k, z) \frac{D(z_{\rm obs})}{D(z)}\,, \tag{5}$$

$$P_{m,\Xi}(k, z_{\rm obs}) = -\int {\rm d}z \frac{\partial \Xi(z, z_{\rm obs})}{\partial z} P_{m,x_{\rm H\,I}}(k, z) \frac{D(z_{\rm obs})}{D(z)}\,, \tag{6}$$

where the integration covers the epoch of reionization, $P_{m,x_{\rm H\,I}}$ is our proxy for the correlation of matter and ionized bubble spatial distribution, which accounts for the patchy nature of reionization. D is the growth factor, while ψ denotes the transparency of the IGM, and Ξ represents the modulation of the neutral hydrogen due to the passage of ionization fronts. They are given by

$$\psi(z_{\rm re}, z_{\rm obs} \,|\, \overline{z}_{\rm re}) = \Delta \ln \tau_1 = \ln \left[\frac{\tau_1(z_{\rm re}, z_{\rm obs})}{\tau_1(\overline{z}_{\rm re}, z_{\rm obs})} \right]\,, \tag{7}$$

$$\Xi(z_{\rm re}, z_{\rm obs} \,|\, \overline{z}_{\rm re}) = \Delta \ln \rho_{\rm H\,I} = \ln \left[\frac{\rho_{\rm H\,I}(z_{\rm re}, z_{\rm obs})}{\rho_{\rm H\,I}(\overline{z}_{\rm re}, z_{\rm obs})} \right]\,, \tag{8}$$

where $3.5 \leq z_{\rm obs} \leq 4$, $z_{\rm re}$ is the *local* redshift of reionization, and $\overline{z}_{\rm re}$ serves as a reference redshift of reionization. The parameter τ_1 is the optical depth that must be assigned in simulations to a patch of gas with mean density and temperature $T = 10^4$ K in order for the mean transmitted flux to match observations, i.e. τ_1 is a normalization factor that guarantees that an optical depth cube return sensible results by matching to the observations of Kim et al. (2007). The transparency is defined as a relative measure, benchmarked against a fiducial scenario with $\overline{z}_{\rm re} = 8$. Conversely, Ξ characterizes the response of haloes with shallow potential wells to the passage of ionization fronts, modelling the perturbations in neutral hydrogen density induced by the response relative to those that occur in a fiducial local reionization scenario. We plot some of the ingredients of the reionization relics in Fig. B1 for reference.

Similarly, the Lyα flux and the 21 cm power spectra are respectively given by

$$P^{\rm 3D}_F(\boldsymbol{k}, z) = b_F^2(1 + \beta_F\mu^2)^2 P_m + 2 b_F b_\Gamma (1 + \beta_F \mu^2) P_{m,\psi} \tag{9}$$

$$P_{21}(\boldsymbol{k}, z) = b_{\rm H\,I}^2(1 + \beta_{\rm H\,I}\mu^2)^2 P_m + 2 b_{\rm H\,I}(1 + \beta_{\rm H\,I}\mu^2) P_{m,\Xi} \tag{10}$$

where we have neglected terms with higher order in Ξ and ψ. Likewise, we disregard non-linear corrections, justified by our focus on large scales.

Even though we extract b_F and β_F from Arinyo-i-Prats et al. (2015), the maximum redshift in their tables is 3. Hence, for our $z > 3$ calculations, we include a redshift evolution factor $[(1 + z)/(1 + z_p)]^{3.55}$ (Palanque-Delabrouille et al. 2013b) with pivot $z_p = 3$ to account for the evolution of the flux bias and redshift space distortion parameter (see equation 10 and surrounding text of Montero-Camacho, Zhang & Mao 2024b).

While both equations (5) and (6) emerge from the IGM's response to inhomogeneous reionization, they represent distinct response mechanisms. Equation (7) covers the response to the local reionization process by underdense gas (Hirata 2018). Physically, it originates in minivoids, as underdense gas undergoes ultraviolet heating, shock heating, and compression during the reionization process. In contrast, equation (8) captures the response to local reionization by denser regions, pre-dominantly arising from minihaloes. Essentially, this is the modulating effect on the number of baryons allowed inside a given halo following the *thermal kick* delivered by an ionization front (Long et al. 2022). Note that both mechanisms differ in their origin on the small scales but are influenced equivalently by the patchy nature of reionization – parametrized here as $P_{m,x_{\rm H\,I}}$ and present in both equations (5) and (6).

To facilitate comparison between the different components of equation (1), we have chosen to utilize the spherically averaged power spectrum,[3] i.e.

$$P^{\rm Sph.}_{21,F}(k, z) = \frac{1}{2\pi} \int_0^{2\pi} {\rm d}\phi \int_0^1 {\rm d}\mu\, P_{21,F}(\boldsymbol{k}, z)\,, \tag{11}$$

where the normalization factor already accounts for the extra factor of 2 due to the μ-symmetry.

We plot equation (11) in Fig. 1 at a few selected redshifts. The impact of reionization is more pronounced at higher redshifts,

[3]This choice was also used by Villaescusa-Navarro et al. (2015) and accounts for the difficulty of having enough angular resolution to have a reliable measurement at different μ-bins.

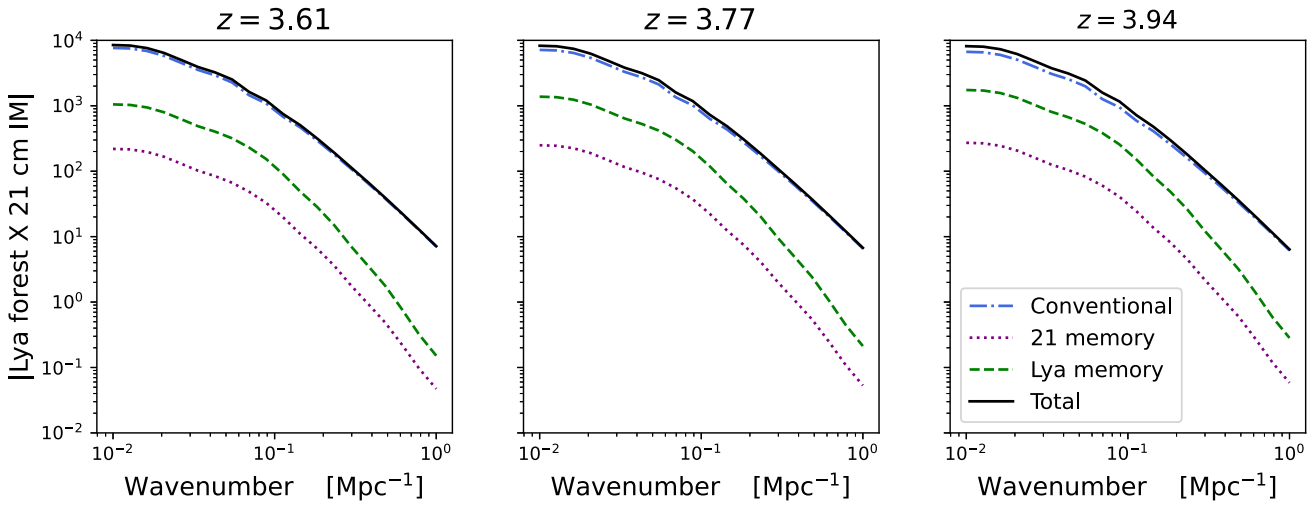

Figure 1. The magnitude of the spherically averaged cross-power spectrum for the Lyα forest $\times$ H I 21 cm IM cross-correlation is illustrated, encompassing its components, including the imprint of reionization in both probes. The blue dash–dotted line corresponds to the *conventional* signal without any relics from cosmic reionization. The purple dotted and green dashed curves represent the memory of reionization in the Lyα forest and 21 cm power, respectively. We highlight that the memory of reionization in the 21 cm signal is negative, leading to a competition in which the Lyα forest, which is positive, dominates. The influence of reionization becomes more pronounced at higher redshifts and larger scales.

introducing a competition between the 21 cm effect, arising from the modulation of baryons due to reionization, and the Lyα forest effect, originating in underdense regions. The former aims to diminish the signal's amplitude while the latter enhances it. Furthermore, the more substantial (in absolute value) impact of the Lyα sourced term compared to the H I 21 cm term is anticipated due to its greater influence on voids, evident in the bimodal temperature-density relation. See fig. 3 to 5 in Montero-Camacho & Mao 2020 versus baryon modulation displayed in fig. 9 of Long et al. 2023. The modulation effect of reionization on the baryon content of haloes contributes up to approximately 25 per cent of the overall 21 cm intensity mapping signal at $z = 5.5$, while the impact of inhomogeneous reionization on the Lyα forest accounts for roughly 60 per cent of the Lyα power spectrum at $z = 4$. While our focus is on the $3.5 < z < 4.$ redshift range, we expect this qualitative trend – a more pronounced reionization imprint on the Lyα forest – to persist up to higher redshifts ($4 \leq z \lesssim 6$) especially given that the impact of self-shielding in minihaloes will also enhance the Lyα power spectrum at large scales (Park et al. 2024) because the surviving neutral hydrogen protected in those dense gas regions will lead to a strong absorption at high redshifts, boosting the power spectrum at large scales, which aligns with the effect of inhomogeneous reionization. Regarding the wavenumber trend, the larger deviation at large scales compared to the conventional term is explained by the coupling to the reionization bubble scale in the cross-correlation of the matter and neutral hydrogen fraction field, $P_{m,x_{\rm H\,I}}$.

3 SIMULATIONS

To accurately model the reionization process, it is imperative to probe under the Jeans length prior to the passage of an ionization front ($\sim$ 100 kpc). Failure to do so would result in the loss of the ability to track the response by the small-scale structure due to the significant wiping out of these structures caused by the increasing Jeans scale throughout reionization (Hirata 2018). Besides, reionization occurs in an inhomogeneous way. Thus, to secure enough statistical power, there is a need to simulate large comoving volumes (a few hundred Mpc) with substantial ionized bubbles each with a radius of a few Mpc. However, the dynamical range required for these considerations is too large to accomplish with reasonable computational resources using a single simulation.

We adopt the methodology introduced by Montero-Camacho et al. (2019) and Long et al. (2023) and employ a hybrid approach to compute equations (5) and (6). This approach relies on seminumerical simulations using 21CMFAST (Mesinger, Furlanetto & Cen 2011; Murray et al. 2020), which track the patchy nature of reionization within a 400 Mpc box. In addition, we leverage high-mass resolution small-box ($L = 1152h^{-1}$ ckpc) simulations run using a modified version of GADGET-2 (Springel 2005; Hirata 2018) to monitor the response to ionization fronts in both dense (crucial for H I 21 cm) and underdense (Lyα forest) regions. This hybrid simulation strategy allows us to capture the impact of reionization across a range of scales and redshifts with minimal trade-offs. Naturally, this hybrid methodology will not be sensitive to higher order correlations between the reionization field and the small-scale density field (Montero-Camacho et al. 2019).

The small-box simulations include adiabatic expansion (including Hubble expansion), shock heating, Compton heating, and cooling for neutral gas (accounting for residual ionization). In addition, we incorporate the long-lasting impact of X-ray pre-heating in the post-reionization IGM (Montero-Camacho et al. 2024b). For ionized gas, we also include Compton cooling, He II cooling, recombination cooling, photoionization heating, and free–free cooling. These simulations have been described – and tested – in detail in Hirata (2018) and Long et al. (2022).

To surpass the limitation on the smallest wavenumber imposed by the size of our 21CMFAST boxes, we utilize a simple liner bias to estimate the effect of reionization at $k \lesssim 0.06$ Mpc^{-1} as follows:

$$P_{m,X}(k,z) = \frac{P_{m,X}(k_{\rm cut} \approx 0.06, z)}{P_m(k_{\rm cut}, z)} P_m(k,z)\,, \tag{12}$$

where $X = \{\psi, \Xi\}$. This approximation should be sufficient for larger scales than the ionized bubble scale. Note, that this restriction is due to the box side of our large-box (400 Mpc) simulations and it is a

self-imposed restriction based on available computational resources. Given the size of our simulation box, the largest scale we can probe is $k_{\rm min} \approx 0.02\ {\rm Mpc}^{-1}$. However, the number of modes in the first few k-bins would be too small to provide meaningful statistical power. Thus, we choose $k_{\rm cut} = 0.06\ {\rm Mpc}^{-1}$ as our cut-off.

4 TELESCOPES

To measure the cross-correlation between 21 cm and the Lyα forest, we require a radio interferometer and a spectroscopic telescope with some non-negligible overlap in sky coverage and redshift. Unfortunately, our choice for estimating the SNR is not straightforward due to the timeline of the DESI (DESI Collaboration 2022) versus the timeline of the SKA (Braun et al. 2019). While DESI may potentially extend its operational life as DESI-II, allowing for direct operational overlap with SKA, the exact timeline overlap remains uncertain.

Concurrently, upcoming spectroscopic instruments such as MegaMapper (Schlegel, Kollmeier & Ferraro 2019), the MUltiplexed Survey Telescope (MUST; Zhang et al. 2024), and MaunaKea Spectroscopic Explorer (MSE; Percival et al. 2019), categorized as Stage V spectroscopic instruments, are anticipated to be operational around the time when 21 cm radio interferometers come online. However, we underscore that the detailed designs of these instruments are currently in the developmental phase. Navigating the evolving landscape of these instrument timelines and designs will be vital to optimizing an observational programme capable of measuring the cross-correlation between 21 cm and the Lyα forest.

Given these 'instrumental' constraints, we have opted to explore several distinct instrumentation scenarios.

(i) SKA1-Low × DESI-like spectroscopic telescope (referred to as SKA in relevant figures).
(ii) SKA1-Low × DESI-MUST-like hybrid (referred to as MUST).
(iii) PUMA-like × DESI-MUST-like hybrid (referred to as PUMA).

Across all scenarios, we maintain a redshift coverage of $3.5 \leq z \leq 4$ and consider two sky coverage overlaps – 100 and 1000 deg^2. Likewise, we vary the integration times, considering both 100 and 1000 h, which will mainly affect the radio observations as described below.

Our base reference scenario involves the SKA1-Low × DESI-like pair, with 100 h of integration time and a 100 deg^2 overlap between the instruments. This reference scenario assumes a *Planck*-like reionization timeline, which may be underestimating the impact of reionization since *Planck* uses a hyperbolic tangent as a sigmoid to model reionization that proves a poor fit of astrophysical constraints on the timeline of reionization (see fig. 4 of Montero-Camacho et al. 2024a).

4.1 SKA1-Low

SKA1-Low, our baseline 21 cm instrument, operates as an aperture array radio interferometer. We consider only the dense core of the array, which is most sensitive to the 21 cm power spectrum. Do note that the outer stations in the array are crucial for calibration and foreground removal purposes. The SKA1-Low, upon completion, offers versatility in supporting various observing modes. Previous works have often focused on observing mode 1, characterized by a dense core comprising 224 stations of 40 m in diameter, each containing 256 dipole antennas (Square Kilometre Array Cosmology Science Working Group 2020; Zhao, Mao & Wandelt 2022). These core stations are distributed over a diameter of approximately 1 km.

Nevertheless, as pointed out in SKA1-LOW Configuration – Constraints & Performance Analysis section 5.4.3, power spectrum measurements are better served by 'substations' with 10 m of diameter. Here, we chose to consider only the observing mode 4, as detailed in table 2 of SKA1-LOW Configuration – Constraints & Performance Analysis. This mode, supported by the instrument's correlator, boasts 6 times more correlatable elements (substations) at the expense of a reduced bandwidth (8.4 MHz compared to 300 for mode 1).

To quantify the unique baselines covered by the dense core, we utilize the number density of baselines n_b from Villaescusa-Navarro et al. (2015) and normalize it to recover the total amount of distinct baselines covered by the dense core, as outlined in section 2 of Long et al. (2023).

The system temperature – the sum of the sky temperature, the ground reflections, and the temperature of the receiver – is given by

$$T_{\rm sys}(\nu) = 60 \left(\frac{300\ {\rm MHz}}{\nu}\right)^{2.55} \times 1.1 + 40 \quad [{\rm K}], \tag{13}$$

while the effective collecting area per station A_e is given by (Square Kilometre Array Cosmology Science Working Group 2020)

$$A_e(\nu) = A_{e,{\rm crit}} \times \begin{cases} \left(\frac{\nu_{\rm crit}}{\nu}\right)^2, & \nu > \nu_{\rm crit} \\ 1, & \nu \leq \nu_{\rm crit} \end{cases}, \tag{14}$$

where $\nu_{\rm crit} = 110$ MHz and $A_{e,{\rm crit}}$ is the collecting area in m^2 for the 256 dipole antennas of 3.2 m^2 each for observing mode 1. Thus, we approximate the effective area for observing mode 4 by decreasing equation (14) by a factor of 6.

Moreover, the field of view is given by

$${\rm FoV}(\lambda) = \left(\frac{\lambda}{\sqrt{0.7} D_{\rm phys}}\right)^2, \tag{15}$$

where $D_{\rm phys}$ is the diameter of the correlatable element and λ is the observing wavelength.

4.2 PUMA

The Packed Ultra-wideband Mapping Array design (PUMA; Cosmic Visions 21 cm Collaboration 2018; Slosar et al. 2019) – *currently* – consists of 32 000 antennas distributed in a hexagonal-close packing in a compact circle of roughly 1.5 km diameter. Each antenna has a diameter of 6 m. We take the baseline number density $n_{\rm b}$ directly from appendix D of Cosmic Visions 21 cm Collaboration (2018).

For PUMA, the system temperature is given by

$$T_{\rm sys}(\nu) = 25 \left(\frac{400\ {\rm MHz}}{\nu}\right)^{2.75} + 2.7 + \frac{300}{9} + \frac{50}{0.81} \quad [{\rm K}]. \tag{16}$$

The effective collecting area per antenna is simply the effective area of the antenna dish, i.e. $\pi(\sqrt{0.7} D_{\rm phys}/2)^2$. Simultaneously, the field of view is given by equation (15) but with $D_{\rm phys} = 6$ m. We set the bandwidth to 8.4 MHz (same as SKA1-Low observing mode 4).

We consider PUMA as an optimistic/futuristic 21 cm telescope for our purposes, essentially our best benchmark.

4.3 DESI-like

We assume a DESI-like spectrograph as our baseline Lyα forest instrument, which is likely a conservative estimate given the projected timeline of SKA. To gauge the performance of the final – 5 yr – DESI

data, we use the quasar luminosity function (QLF) to quantify the expected number of Lyα quasars observable with DESI (Palanque-Delabrouille et al. 2013a; Yèche et al. 2020) and the spectrograph performance of DESI – see equation (21). Detailed information about the DESI instrument is available in DESI Collaboration (2022), the quasi-stellar object (QSO) and Lyα QSO target selection can be found in Chaussidon et al. (2023), and the spectroscopic pipeline is described in Guy et al. (2023).

4.4 DESI-MUST-like hybrid

We refine our initially conservative DESI-like scenario by enhancing its performance to align more closely with the standards expected for Stage V spectroscopic instruments. Pragmatically, this improvement involves adjusting the aliasing term in the covariance computation, effectively augmenting the effective density of lines of sight beyond what is anticipated with our original DESI specs. In summary, we implement an optimistic factor of 3 reduction in both the second and third terms of equation (21). Given that MUST will likely be the World's first Stage V spectroscopic instrument,[4] we decided to refer to this scenario as MUST.

5 RESULTS

In the redshift range of interest, the cross-correlation of Ly$\alpha\times$ 21 cm IM presents a pragmatically more accessible measurement compared to the auto-correlation analyses of both the Lyα forest or of the H I 21 cm field. This disparity arises from distinct challenges associated with each observable. For the Lyα forest, the QLF peaks around $z\sim 2$, resulting in a substantial reduction in the number density of quasars at $z\sim 4$. Conversely, for 21 cm intensity mapping, the signal strength is significantly weaker than the foregrounds by several orders of magnitude. Notably, these challenges are uncorrelated, rendering the cross-correlation less susceptible to their impacts (Carucci et al. 2017). Since this is the first study of the effects of the relics from reionization in this observable, we will not consider foreground contamination nor several Lyα forest systematics, like continuum fitting (Sun, Ting & Cai 2023), spectra with broad absorption lines (Filbert et al. 2024), UV clustering (Long & Hirata 2023), spectra with damped Lyα systems (Wang et al. 2022), and several other astrophysical and instrumental systematics. Note that this simpler approach is often used in other studies of this cross-correlation (e.g. Guha Sarkar et al. 2011; Carucci et al. 2017; Sarkar et al. 2019) at other redshifts.

5.1 Impact on SNR

Under these considerations, we can compute the errors on the observables following Villaescusa-Navarro et al. (2015) and Carucci et al. (2017)

$$\sigma^2[P_{21}(\boldsymbol{k},z)] = \left(P_{21}^{\rm T}(\boldsymbol{k},z)\right)^2 , \tag{17}$$

$$\sigma^2[P_F(\boldsymbol{k},z)] = \left(P_F^{\rm T}(\boldsymbol{k},z)\right)^2 , \text{and} \tag{18}$$

$$\sigma^2[P_{21,F}(\boldsymbol{k},z)] = \frac{1}{2}\left(P_{21,F}^2 + \sigma[P_{21}]\sigma[P_F]\right) . \tag{19}$$

For 21 cm, $P^{\rm Tot.}$ has contributions from cosmic variance and thermal noise, i.e.

$$P_{21}^{\rm T}(\boldsymbol{k},z) = P_{\rm H\,I} + T_{\rm sys}^2(z)\chi^2(z)\lambda(z)\frac{1+z}{H(z)}\left(\frac{\lambda^2(z)}{A_e}\right)^2\left(\frac{S_{\rm area}}{{\rm FoV}(z)}\right) \times\left(\frac{1}{N_{\rm pol}t_{\rm int}n_{\rm b}(u=k_\perp\chi(z)/2\pi)}\right) , \tag{20}$$

where χ is the comoving distance, H is the Hubble parameter, and $\lambda = 21\,{\rm cm}/(1+z)$. Moreover, $T_{\rm sys}$ stands for the system temperature of the radio interferometer, $N_{\rm pol}=2$ is the number of polarizations, A_e the effective area of the correlatable unit (dishes or stations), $S_{\rm area}$ is the survey area, FoV is the field of view, $n_{\rm b}$ quantifies the number density of baselines in the uv-plane as a function of wavenumber, and $t_{\rm int}$ corresponds to the integration time. The shot noise is subdominant for the redshift range of interest here (Castorina & Villaescusa-Navarro 2017), thus we do not include it in equation (20).

In contrast, for the Lyα forest we have (McDonald & Eisenstein 2007; Font-Ribera et al. 2014; Montero-Camacho & Mao 2021)

$$P_F^{\rm T}(\boldsymbol{k},z) = P_F^{\rm 3D}(\boldsymbol{k},z) + P_F^{\rm 1D}(k_\parallel,z)P_w^{\rm 2D}(z) + P_N^{\rm eff}(z) . \tag{21}$$

The first term corresponds to cosmic variance, i.e. it comes from $\langle\delta_F\delta_F^*\rangle$, while the second term represents the aliasing noise due to the sparse sampling of quasars. The third term describes the effective noise due to the spectrograph performance. Note that at high redshifts ($z\gtrsim 3.2$), the aliasing term tends to dominate equation (21) since the limited number of quasars becomes more restrictive (see fig. 6 of Montero-Camacho & Mao 2021). Naturally, this trend reduces the expected SNR of Lyα forest surveys attempting a measurement of the 3D flux power spectrum at high redshifts.

Given equations (17, 18, or 19), the SNR for P_i can be written as

$${\rm SNR}_i^2 = N_k\frac{P_i^2}{\sigma^2[P_i]} = \frac{V_{\rm Survey}k^3\epsilon{\rm d}\mu}{4\pi^2}\frac{P_i^2}{\sigma^2[P_i]} , \tag{22}$$

where $\epsilon = {\rm d}k/k$ and $V_{\rm Survey}$ is the – overlap – volume between the radio interferometer and the spectrograph. We consider three redshift bins centred[5] at $z_{\rm c} = [3.61, 3.77, 3.94]$ with widths defined by the bandwidth of the radio interferometer, i.e. $\Delta\nu = 8$ MHz. Consequently, our survey volume for a redshift bin centred at $z_{\rm c}$ is given by

$$V_{\rm Survey}(z_{\rm c}) = \frac{4\pi}{3}f_{\rm sky}[\chi(z_{\rm max})^3 - \chi(z_{\rm min})^3)] \tag{23}$$

with χ as the comoving distance. Note that $f_{\rm sky}$, the sky coverage, depends on the area of overlap.

In what follows, we focus on the spherical-averaged power spectrum, i.e. equation (11) and analogues, to compute the appropriate SNR for Ly$\alpha\times$ 21 cm, Lyα forest, and 21 cm intensity mapping.

We illustrate the observability of the cross-correlation compared to the autocorrelation of Lyα and 21 cm in Fig. 2 for our base scenario (SKA1-Low × DESI-like, $t_{\rm int} = 100$ h, $A_{\rm sky} = 100\,{\rm deg}^2$ and *Planck*'s reionization timeline). We highlight that all of the curves here include the long-lasting relics from cosmic reionization, i.e. equations (5) and (6). Unsurprisingly, the SNR for the Lyα forest is less than unity because of the drop in available line of sight at these redshifts and because of the survey volume. In principle, it is necessary to observe more quasars to reduce the mean separation between forests to make the 3D flux power spectrum observation

[4]MUST's first light is currently planned for 2029, see https://must.astro.tsinghua.edu.cn/en for more details.

[5]The centre of the bins were chosen to coincide with the last three bins used in Montero-Camacho & Mao (2021)'s forecast.

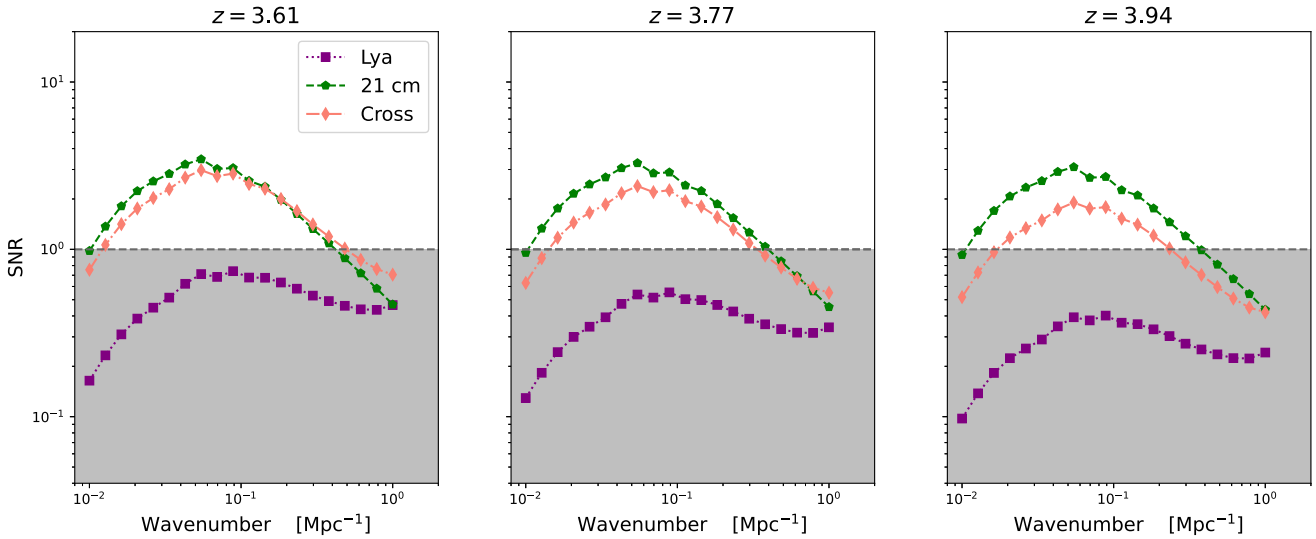

Figure 2. The SNR for the baseline configuration of the Lyα forest (purple squares), 21 cm intensity mapping (green pentagons), and Ly$\alpha\times$ 21 cm (orange diamonds) is presented as a function of wavenumber in our baseline scenario (see the text for details). The three panels correspond to different redshift bins. The cross-correlation is, naturally, the more pragmatic measurement since the 21 cm signal would be severely impacted by foregrounds. Notably, the SNR for the forest is somewhat impeded by the chosen survey volume, although the drop in the number density of quasars leads to difficulties at high redshifts regardless of sensible choices for survey volume. The grey-shaded region corresponds to SNR ≤ 1.

feasible. Although Fig. 2 demonstrates that the 21 cm signal has a larger SNR, this is only true in the absence of foregrounds, which will severely bury the autopower spectrum. In contrast, the cross-correlation is more robust against foreground contamination (Furlanetto & Lidz 2007; Sarkar et al. 2018; Zhou, Tan & Mao 2021) and it is already competitive with the 21 cm autocorrelation with the baseline instrumental set-up.

Fig. 2 provides a compelling rationale for the significance of the Ly$\alpha\times$ 21 cm cross-correlation, particularly in the presence of foregrounds. Detection in the cross-correlation can guarantee the cosmological nature of the 21 cm signal. Besides, the SNR is similar to the autocorrelation at the ionized bubble scale (and for our first redshift bin). This is consistent with the findings of Carucci et al. (2017) at a lower redshift ($z = 2.4$). However, we underscore that in the absence of reionization relics, the cross-correlation will not have a similar SNR to that of the auto 21 cm SNR in the lowest redshift bin considered here. The boost provided by the memory of reionization is crucial to be competitive, yet it is not sufficient at higher redshift bins because of a significant rise in aliasing noise – $P_w^{\rm 2D}$ in equation (21) – at those redshifts.

Furthermore, akin to the findings in Carucci et al. (2017), the cross-correlation can surpass the SNR of the 21 cm autocorrelation at smaller scales. However, this trend is obscured at high redshifts due to the poor sampling of quasars.

Having established the interest in the cross-correlation of the Lyα forest and 21 cm IM at the redshift range of interest ($3.5 \leq z \leq 4$), we now ponder the impact of the survey strategy.

5.2 Importance of survey design

Our main objective here is to underscore the importance of a non-negligible overlap between spectroscopic instruments and radio interferometers. Given that SKA1-Low is under development in the Southern hemisphere and DESI – also likely DESI II – is in the Northern hemisphere, we investigate the potential advantages and insights that could be derived from this cross-correlation with diverse instrumental configurations.

In order to reduce the noise in the radio interferometer, a widely employed strategy is to consider an increase in the integration time since the thermal noise is inversely proportional to $t_{\rm int}$. Hence, we explore the impact of a tenfold increase in $t_{\rm int}$ across various instrumental configurations – specifically SKA1-Low $\times$ DESI-like, SKA1-Low $\times$ DESI-MUST-like hybrid, and PUMA-like $\times$ DESI-MUST-like hybrid.

We assume that the Lyα forest observations will be completed regardless of the integration time; nevertheless, we reward the longer observation time with a factor of 2 reduction in the spectrograph effective noise $P_N^{\rm eff}$. In other words, this implies that the extended observation time allows for more exposures. Note that we do not modify the aliasing error – second term in equation (21) that significantly dominates the error budget – even when considering additional observation time.

Fig. 3 showcases the dependence of the SNR for different integration times and telescope pairs. Unsurprisingly, there is a consistent enhancement in SNR across all redshifts and configurations. In particular, the SKA1-Low configurations benefit the most from this strategy, manifesting improvements of ≈ 2.6 in overall SNR (see Table 1). Intriguingly, the PUMA-like $\times$ DESI-MUST-like hybrid exhibits negligible improvement at large scales.

From equation (20), the thermal noise of the radio interferometer is inversely proportional to the total integration time. The reason Fig. 3 shows a larger impact for the pairs using SKA1-Low is because the radio telescope dominates the error budget for those set-ups, i.e. it is larger than the cosmic variance contribution and $\sigma[P_{\rm 21}] > \sigma[P_F]$. This is also why the change from DESI-like to DESI-MUST-hybrid does not result in a major improvement and largely conserves the increase in SNR that was present in the 100-h baseline scenario. In contrast, for the PUMA $\times$ MUST set-up, the Lyα forest survey now dominates the error, gains are then possible at larger wavenumbers given the large drop at large scales (as can be seen in Fig. 2). We also highlight that the number density of baselines also restricts the shape

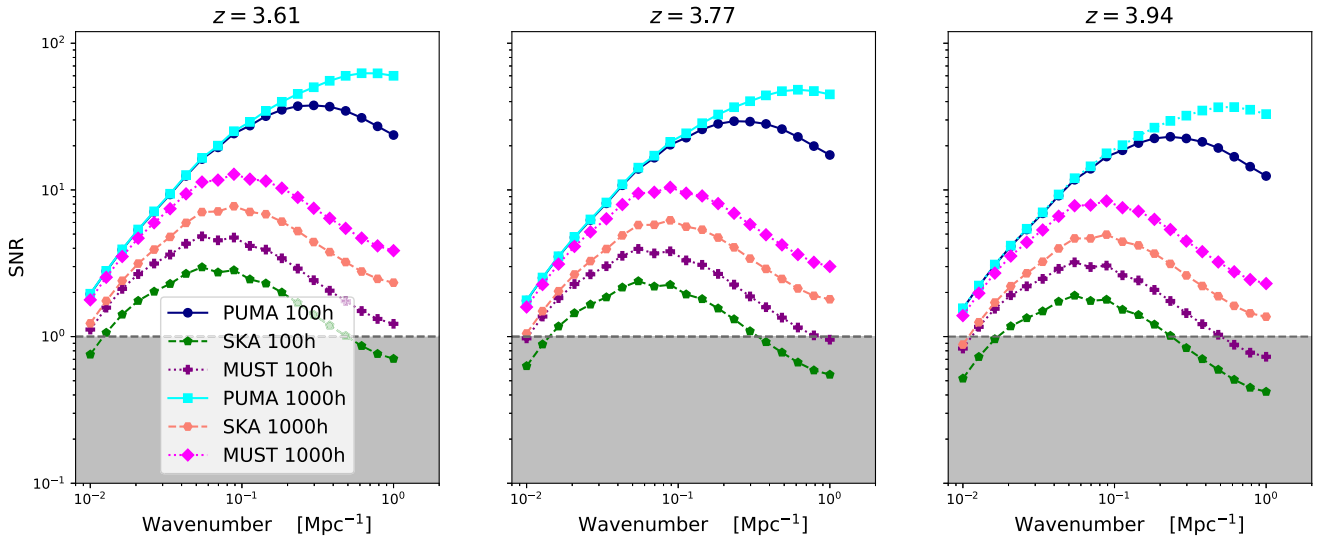

Figure 3. SNR for the Lyα forest $\times$ 21 cm intensity mapping accounting for variations in the integration time of the radio telescopes. We assume that the Lyα forest survey is also completed during that integration time and reward the additional integration time (see the text for details). Shown are the SNR for: PUMA $\times$ MUST 100 h (blue circles) and 1000 h (cyan squares); SKA1-Low $\times$ DESI 100 h (green pentagons) and 1000 h (orange octagons); SKA1-Low $\times$ MUST 100 h (purple crosses) and 1000 h (magenta diamonds). Increasing the total integration time allows for a better measurement in all configurations.

Table 1. Summary of the total SNR at $z = 3.61$ for the different scenarios considered in this work. The fiducial scenario corresponds to 100 h, 100 deg^2 of survey area, and a *Planck*-like reionization scenario. We highlight that increasing the survey area by tenfold results in a marginal increase to the SNR for the baseline instrumental set-up (see Section 5.2).

	Total SNR		
	SKA $\times$ DESI	SKA $\times$ MUST	PUMA $\times$ MUST
Fid.	8.50	13.9	110
	Increase integration time		
1000 h	21.9	35.9	167
	Increase survey area		
1000 deg^2	8.89	14.6	187
	Reionization		
None	7.76	12.8	103
Early	8.10	13.3	106
Late	8.75	14.3	115

of the SNR at these scales, which is why the SKA1-Low $\times$ MUST can be unexpectedly competitive with PUMA $\times$ MUST's SNR at some wavenumbers for a 1000 h of integration time.

It is unclear whether major projects like SKA or a Stage V spectroscopic instrument could allocate 1000 h/many exposures to observe the same overlapping area, given the promising science programmes they aim to address. Nevertheless, it is evident that with such a time investment the community could expedite the progress in measurements of the cross-correlation by a decade, achieving observations competitive with what may be available with Stage II 21 cm radio telescopes – at least at large scales. We acknowledge, however, that this is a significant time commitment.

Naively, one might expect that expanding the overlapping area would perhaps be a more straightforward enterprise, especially for spectroscopic instruments located in the Southern hemisphere like the 4-metre Multi-Object Spectroscopic Telescope (4MOST; Richard et al. 2019). However, it is important to note that increasing the survey area produces a larger thermal noise – see equation (20). Conversely, the number of modes N_k is directly proportional to the survey area. Thus, as long as the thermal noise is manageable, an increase in the survey area will improve SNR due to the additional modes. For the PUMA set-up, where Lyα dominates the error budget at large scales, one could anticipate that the increase in modes may give rise to an enhancement in SNR up to a large k value. In contrast, the impact of this choice is expected to generate limited gains for the SKA and MUST configurations due to the dominance of the 21 cm error, particularly with a tenfold increase in thermal noise.

We confirm these expectations in Fig. 4 by augmenting the area of overlap by a factor of 10 from the baseline scenario (from 100 to 1000 deg^2).

The interpretation of this finding is convoluted. Initially, it may be more appealing to allocate 100 h of a telescope's time to an observation spanning different pointings. Unfortunately, this approach does not lead to a significant enhancement in observations for the late 2020s configurations. In terms of the overall SNR, aggregated across the k-bins, there is only a very modest increase of approximately 1.05 compared to the baseline scenario.

Moreover, if radio telescopes happen to be located in the global south and spectroscopic instruments keep the trend of being located in the Northern hemisphere, then the overlap may be small to begin with and hence preference would be given to securing a small overlap in the survey footprints. To our knowledge the possible location of PUMA is not decided yet, consequently, our results advocate for the consideration of overlap with Stage V spectroscopic instruments. Furthermore, our results also showcase that a larger overlap can produce significant gains for a PUMA cross-correlation at large scales. For PUMA, the overall SNR augments by a factor of 1.7 when enlarging the survey area by a factor of 10. Besides, this enhancement is limited by the assumption of Stage V instrument performance. In principle, PUMA may instead share the skies with potential Stage VI spectroscopic instruments.

While we have only scratched the surface regarding the implications of survey strategy for the Ly$\alpha \times$ H I 21 cm cross-correlation at $3.5 \leq z \leq 4$, we deem further exploration of the instrumental set-up beyond the scope of this work. This study focuses on demonstrating the impact of cosmic reionization in the cross-correlation during

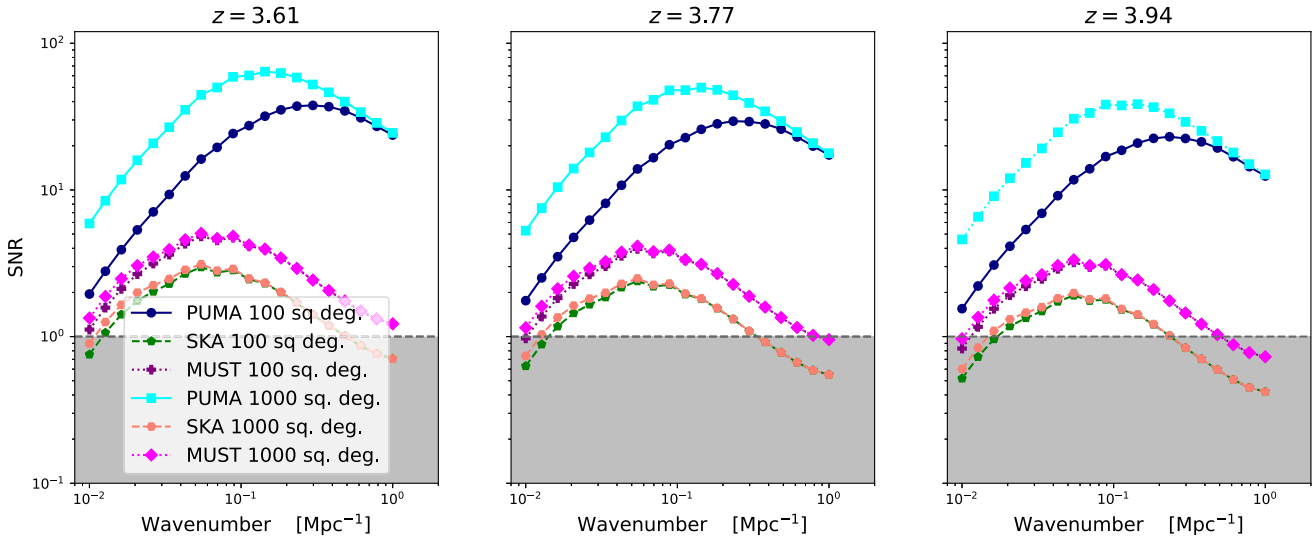

Figure 4. Similar to Fig. 3, but modifying the survey area. Notably, increasing the survey area does little to improve the SNR for the configurations with SKA1-Low since the error is dominated by the interferometric instrument and the thermal noise increases with survey area – see equation (20).

the post-reionization era. Therefore, it would be remiss of us not to address the dependence of the signal on the uncertain timeline of reionization (Montero-Camacho & Mao 2021; Jin et al. 2023).

5.3 Dependence on reionization history

The timeline of reionization remains uncertain, although abundant direct detection of the ionizing sources is becoming possible thanks to the *JWST* (Gardner et al. 2006). In fact, *JWST* has already started to revolutionize our understanding of galaxy formation at the epoch of reionization (EoR) redshifts (e.g. Bradley et al. 2023; Adams et al. 2023; Atek et al. 2023; Donnan et al. 2023). Nevertheless, even though perfect knowledge of the reionization timeline would help to model the EoR, it would still be non-trivial to translate these constraints into a single mapping of reionization astrophysics.

Consequently, it is intriguing to ponder the impact of different reionization scenarios in our SNR. We limit our study of reionization scenarios to three distinct reionization timelines. Our fiducial case aligns with *Planck*'s timeline of reionization $z_{\rm mid} = z(\overline{x}_{\rm HI} = 0.5) \approx 7.7$, besides its duration, defined as $\Delta z = z(\overline{x}_{\rm HI} = 0.10) - z(\overline{x}_{\rm HI} = 0.90)$ is 3.83. Furthermore, we complement this model with scenarios representing later and earlier reionization scenarios, having mid-points at $z_{\rm mid} = 6.88$ and 8.41, and durations of $\Delta z = 4.15$ and 3.59, respectively. Note that these additional models are roughly consistent with the 1σ error reported by Planck Collaboration VI (2020) but we caution that the best fit from *Planck* is not able to reproduce astrophysical constraints in a satisfactory way – see fig. 4 of Montero-Camacho et al. (2024a) – due to the use of the hyperbolic tangent reionization model (Lewis 2008). A more adequate choice would be to consider the range of reionization profiles allowed in the Gompertz model of reionization. Its best fit to CMB data, in conjunction with astrophysical constraints, indicates a midpoint of reionization at ≈ 7 (Montero-Camacho et al. 2024a), more closely aligned with our late reionization scenario.

We obtain our three reionization models by modifying the ionization efficiency ζ in 21CMFAST. This parameter governs the overall timeline of reionization, although there are degeneracies (Pober et al. 2014; Park et al. 2019; Montero-Camacho & Mao 2021). Physically, it quantifies the ability of photons to escape their parent galaxies and reach the intergalactic medium (Dayal & Ferrara 2018).

In Fig. 5, we illustrate the effect of including the memory of reionization in $P_{21,F}$ compared to neglecting its existence. Naturally, the influence of the remnants of reionization is more pronounced in the highest redshift bin and for the delayed reionization scenario. As time progresses, gas dissipates the additional injected energy during reionization, resulting in a lesser effect in the Lyα forest. Meanwhile, the modulation of the baryons in shallow potential wells becomes more subdued due to the growth of affected galaxies, as a result, the significance of the memory of reionization in H I 21 cm gradually diminishes as well.

As expected, the primary gains in SNR occur at wavenumbers associated with the reionization-bubble scales, and they are minor at small scales. Note that the maximum importance overlaps with the best window in terms of SNR, as shown in Fig. 2. Hence, the cross-correlation of Lyα and 21 cm intensity mapping holds significant promise for shedding light on the astrophysics of reionization and it could be used to indirectly constrain the timeline of reionization through its impact in the post-reionization intergalactic medium. Future work will investigate what could be gained from such efforts (see Section 6 for a preliminary forecast).

Furthermore, Fig. 5 demonstrates that the impact of inhomogeneous reionization on the post-reionization IGM will increase the SNR for all timelines of reionization and redshifts considered here. In fact, the enhancement can reach 30 per cent at its peak. However, we underscore that there is a fundamental competition in $P_{21,F}$ between the response of underdense regions (forest) and that of denser regions (intensity mapping) to the reionization process. It is plausible that the overall effect would be a reduction of the signal-to-noise at redshifts close to the tail end of reionization. Fortunately, further scrutiny of this hypothesis will not require high-mass resolution simulations like the ones used throughout this work since the role of the small-scale structure should be subdued at these redshifts. Nevertheless, the feasibility of such a measurement is at best uncertain given the current capabilities of spectroscopic instruments and the sparse density of quasars at such high redshifts.

We summarize our findings regarding observation strategy and dependence with reionization history in terms of SNR in Table 1.

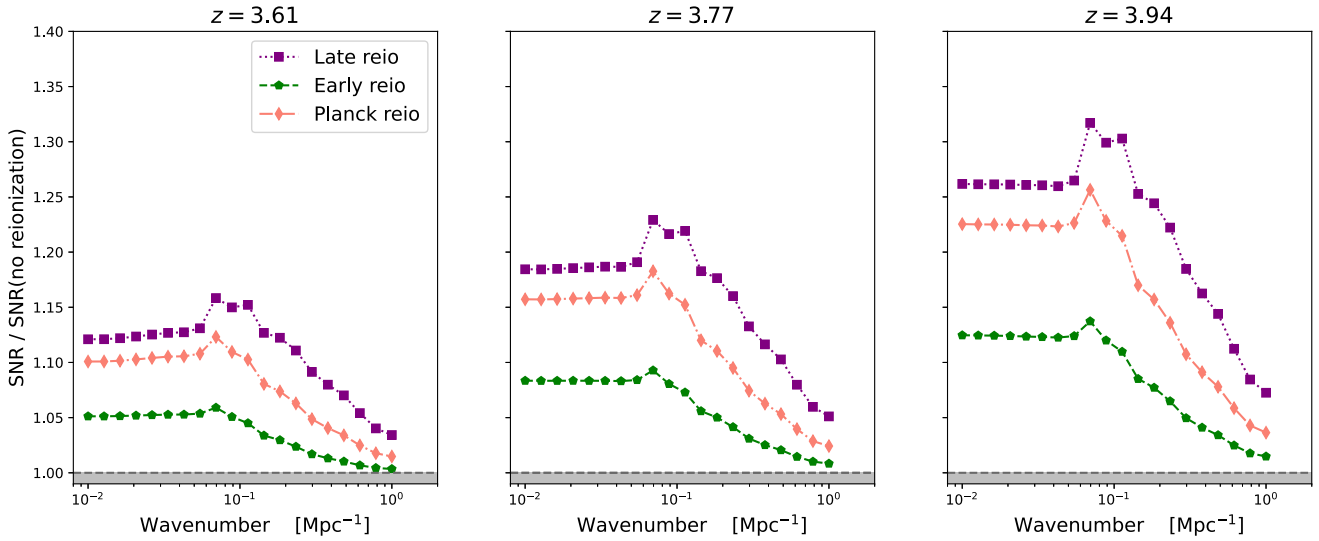

Figure 5. Impact of reionization history in the SNR of the cross-correlation assuming the SKA1-Low × DESI baseline configuration, i.e. $t_{\rm int} = 100$ h and $S_{\rm survey} = 100$ deg^2. Shown are the SNR for a late reionization (purple squares), a *Planck*-like reionization (orange diamonds), and an early reionization scenario (green pentagons). The behavior at small k is driven by equation (12).

6 A SIMPLE FORECAST

Having explored the potential gains in SNR due to different strategies and instrumental set-ups, we now turn our attention to quantifying the potential cosmological gains using the Fisher Matrix formalism (Heavens 2009).

Since no other work has looked at the cross-correlation of Ly$\alpha\times$ 21 cm at this redshift range, we will forecast based on three different scenarios: (i) conventional signal, i.e. no memory of reionization, (ii) memory of reionization with perfect reionization knowledge, and (iii) marginalizing over the memory of reionization due to uncertainty regarding its timeline/modelling. Hence, we can illustrate the potential advantage gained by the reionization relics but simultaneously demonstrate the dangers of ignoring this effect. Furthermore, note that our strategies do not include any sort of priors from other data sets,[6] and hence future forecasts are likely to improve in these projections.

The Fisher matrix is given by

$$F_{\alpha\beta} = \sum_{i}^{z-\rm bins}\sum_{j}^{k-\rm bins} \sigma^{-2}_{z_i k_j}[P_{21,F}]\frac{\partial P_{21,F}}{\partial\theta_\alpha}(z_i,k_j)\frac{\partial P_{21,F}}{\partial\theta_\beta}(z_i,k_j)\,, \qquad (24)$$

where we have assumed uncorrelated errors between the different bins and that the posterior distribution can be reasonably well-described by a Gaussian. For simplicity, we consider only two cosmological parameters to forecast $\theta_i = \{\sigma_8, n_s\}$ in (i) and (ii). For (iii), the conservative point of view on the progress of the field, we add the ionization efficiency ζ to $\boldsymbol{\theta}$.

The uncertainty in equation (24) is obtained from equation (19). Besides, we use the same three redshift bins centred at 3.61, 3.77, and 3.94 used throughout the rest of this paper. For the wavenumbers, we consider 20 logarithmic bins from 0.01 to 1 Mpc^{-1}.

[6]One could, for instance, follow section 2.4 of Long et al. (2023) to incorporate CMB priors; however, this would not address the degeneracy between σ_8/n_s and ζ. Instead, one could consider using the recent results from Montero-Camacho et al. (2024a) to simultaneously break that degeneracy.

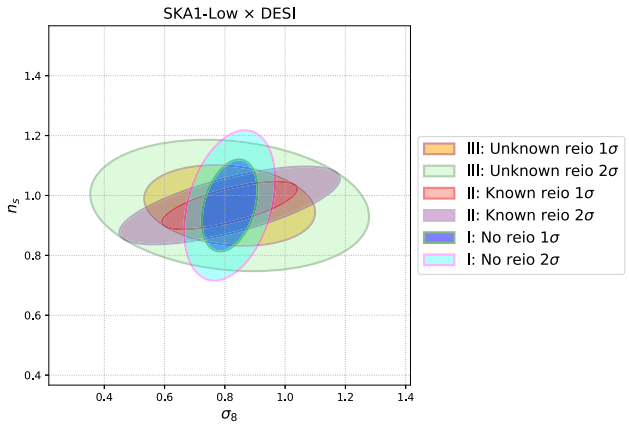

Figure 6. Forecast for the ability of SKA1-Low × DESI-like to constrain σ_8 and n_s using 100 deg^2 and 100 h of integration time. Shown are the confidence ellipses – 1σ and 2σ – for different strategies where we consider the case of no reionization (I), i.e. the conventional signal; impact of reionization with known reionization (II), and impact of reionization with uncertain reionization (III). Notably this configuration, which uses only three redshift bins, is not very powerful compared to other cosmological programmes expected to be operational by the end of the 2020s; however, it still showcases how reionization relics will jeopardize our ability to constrain cosmology from this observable.

To compute the derivatives in equation (24), we run additional simulations that cover deviations of 3 per cent in σ_8 and n_s around their fiducial values of 0.8159 and 0.9667, respectively. Meanwhile, for ζ we also consider 3 per cent variation around $\zeta = 30$ but we check the convergence of our choice for the ionization efficiency in Appendix A. Furthermore, we illustrate some of the tendencies for different properties of this simulation suite in Appendix B.

We plot the results of the SKA × DESI-like (PUMA × MUST-like) forecast in Fig. 6 (Fig. 7) for the three different strategies considered in this work. The confidence ellipses correspond to the 1 and 2σ

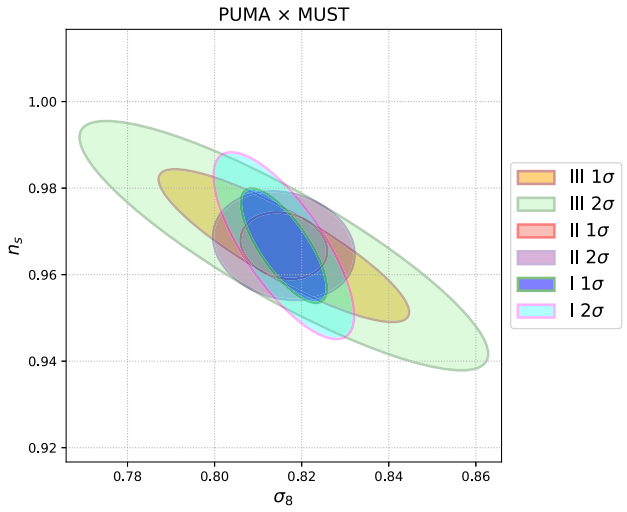

Figure 7. Similar to Fig. 6, but for PUMA × MUST-like. In contrast to Fig. 6, this set-up places competitive constraints on the cosmological parameters. Therefore, our results emphasize the importance of considering the impact of reionization to prevent miscalculations of errors and the potential introduction of biases. Note that the combination with other data sets and the use of more than three redshift bins will significantly enhance the real constraining power expected from these instruments.

Table 2. Projected errors for the Fisher forecast of Section 6. Shown are the 1σ errors obtained for two instrumental configurations (SKA1-Low × DESI-like and PUMA × MUST-like) and for three different scenarios that correspond to no reionization (I), perfect knowledge of reionization/no marginalization over EoR astrophysics (II), and marginalization over ionization efficiency (III), respectively. Both instrumental set-ups have 100 h of integration time and 100 deg^2 of sky coverage.

Forecasted errors						
	SKA × DESI			PUMA × MUST		
Strategy	σ_{σ_8}	σ_{n_s}	σ_ζ	σ_{σ_8}	σ_{n_s}	σ_ζ
I	0.0608	0.1011	–	0.0065	0.0087	–
II	0.1474	0.0524	–	0.0065	0.0051	–
III	0.1864	0.0884	30.39	0.0189	0.0116	3.42

contours for the $n_s \times \sigma_8$ plane. In addition, we tabulate the projected errors in Table 2.

Both instrumental configurations exhibit a trend of greater ability to constrain the tilt of the primordial power spectrum when the impact of reionization is accounted for and known but when marginalization over the ionization efficiency is required the constraint power for the tilt degrades for the PUMA × MUST-like configuration. This trend of better n_s constraints was also present in the forecast for Lyα forest done by Montero-Camacho & Mao (2021). As seen in their figs 9 and 10, which include the memory of reionization in the 3D flux power spectrum from the Lyα forest, the tilt plays a more important role than the amplitude. We attribute this to the increase (or decrease) of faint galaxies that would happen by modifying the value of n_s while a similar increase in σ_8 will affect the environment more uniformly, hence influencing the ionizing sources to a lesser degree than the tilt. Note that reionization is driven by faint galaxies in our 21CMFAST simulations. The larger significance of the impact of the tilt is also shown in Fig. B1 where the impact of changing n_s and σ_8 in the reionization history is illustrated.

In the absence of reionization relics, the cross-correlation – without any external information from other cosmological probes – seems to constrain σ_8 more than n_s. Interestingly, this is the opposite behaviour of the results for H I 21 cm intensity mapping auto power spectrum albeit with A_s instead (see the 1σ error in table 4 of Long et al. 2023). Disregarding the difference of σ_8 with A_s, this trend could be due to the use of *Planck* priors in the autocorrelation, the additional parameters (like Ω's that will likely be degenerate with A_s and σ_8), or perhaps due to the anticorrelated nature of the cross-correlation since in the low-density regions the forest is sensitive to, an increase in σ_8 would result in enhanced structure formation, and consequently a lesser Lyα flux due to the increased absorption. This extra absorption or lack of could play a significant role in the range where the SNR is not dominated by the 21 cm error, i.e. where the Lyα forest dominates the error budget, as seen on large scales for the PUMA × MUST configuration. In these regions, the enhanced sensitivity to the Lyα forest increases the constraining power on σ_8, which could explain the trend of stronger constraint on σ_8 for the cross-correlation relative to the trend in the autocorrelation.

The Lyα× 21 cm cross-correlation will eventually become a competitive probe of the astrophysics that governs cosmic reionization. However, the constraining power for 100 h of integration time is too weak to offer real insights into the reionization process. This research direction is likely to become promising with a more ambitious cross-correlation programme, say 1000 h in a SKA × MUST-like set-up. Unexpectedly, even a conservative 100-h cross-correlation survey will be highly competitive once PUMA (Cosmic Visions 21 cm Collaboration 2018) starts taking data. For reference, 100 h of integration time would greatly improve on the projected error for a full DESI (5 yr) constraint using the 3D flux power spectrum ($\sigma_\zeta = 11.6$; Montero-Camacho & Mao 2021) by roughly a factor of 3. Its constraining power would be of similar strength to that obtained by demoting the optical depth to reionization to derived-parameter using symbolic regression and CMB data in conjunction with astrophysical data (Montero-Camacho et al. 2024a).

6.1 Constraint on the timeline of reionization

Here, we use the results of our PUMA × MUST-like Strategy III forecast – marginalize over reionization astrophysics – to constrain the timeline of reionization, assuming our fiducial model correctly represents the Universe's neutral hydrogen fraction. In essence, we assume $\zeta = 30$ and we construct the range of allowed $\overline{x}_{\rm H\,I}$ based on our forecasted error σ_ζ.

For clarity, we decided to only consider the PUMA configuration, which is significantly stronger than the SKA1-Low × DESI-like constraint (the error is smaller by almost a factor of 10). This stronger constraining power also justifies our choice of only considering ζ due to the very small errors for σ_8 and n_s expected from this instrumental configuration. Additionally, note that our choice of $\zeta = 30$ physically aligns with a scenario where a significant population of faint galaxies is responsible for driving the reionization process – see the *faint galaxies* model introduced in Greig & Mesinger (2018).

We plot our findings in Fig. 8 along with the current state of direct and indirect observations (i.e. optical depth constraint from the CMB). For our fiducial model, with an ionization efficiency of $\zeta = 30$, the ionization of neutral hydrogen begins gradually around $z \sim 12$ as denser regions start forming sources of ultraviolet photons capable of ionizing hydrogen. As time progresses, more H II bubbles form, expand, and merge, leading to an asymmetric reionization process in which the later stages proceed more rapidly. This contrasts with the symmetric sigmoid model, based on a hyperbolic tangent

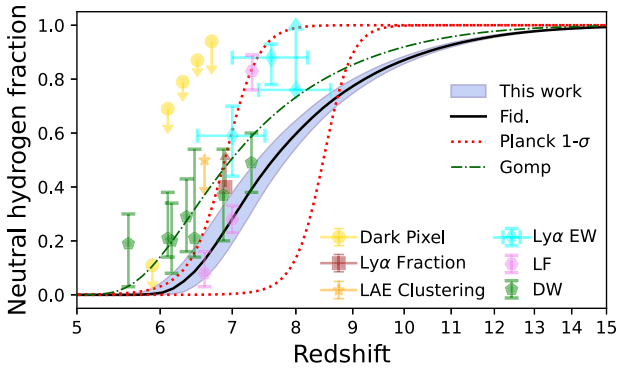

Figure 8. Inferred constraint on the reionization history by our PUMA × MUST set-up using 100 h of integration and assuming a 100 deg^2 of overlap (solid line). Also shown are current observational constraints on the timeline of reionization including dark pixel fraction (McGreer, Mesinger & D'Odorico 2015; Jin et al. 2023), high-redshift galaxies through their clustering, luminosity evolution (LF) and equivalent width (EW) (Ouchi et al. 2010; Mesinger et al. 2015; Sobacchi & Mesinger 2015; Mason et al. 2018, 2019; Hoag et al. 2019; Morales et al. 2021), and high-z quasars damping wings (DW) (Greig et al. 2022, 2024; Ďurovčíková et al. 2024; Spina et al. 2024). We also include the *Planck* indirect constraints on the timeline of reionization (dotted line, Planck Collaboration VI 2020), which relies on a hyperbolic tangent to parametrize the reionization process. In addition, we include the best fit using Gompertz reionization (dash-dotted line, Montero-Camacho, Li & Cranmer 2024a), which uses *Planck* data but does not rely on the hyperbolic tangent parametrization and provides a better fit to astrophysical observations. Note that we have anchored the constraint around the fiducial model, thus the key feature is the width of the constraint, rather than its exact location.

function, used in e.g. Planck Collaboration VI (2020) constraint to model reionization.

We emphasize that the width of our inferred constraint is the true value added by our forecast. Our fiducial value, based on the default in 21CMFASTV3 and aligned with *Planck*'s inferred reionization history (Planck Collaboration VI 2020), provides a reasonable description of the Universe's reionization timeline. However, recent advancements in the more reliable quasar damping wing observations (see e.g. Ďurovčíková et al. 2024; Greig et al. 2024) suggest a slight tension with both our fiducial model and the *Planck* constraint. Notably, these developments also support the Gompertz reionization scenario (Montero-Camacho et al. 2024a), which successfully fits both CMB data and astrophysical constraints on reionization. Future work should consider a fiducial model that better aligns with the Gompertz timeline.

Our forecast indicates that measuring the Ly$\alpha\times$ 21 cm cross-correlation will likely become a powerful new tool to constrain the evolution of the neutral hydrogen fraction. Such a tight constraint as the one in Fig. 8 will impact the standard cosmological model (Lambda cold dark matter), not only by refining the optical depth but also through its dependence on the other cosmological parameters that, in turn, influence the reionization process. Moreover, if observational constraints based on the evolution of the luminosity function of high-z galaxies (represented by pink hexagons in Fig. 8) remain unchanged, the narrowness of our $\overline{x}_{\rm H\,I}(z)$ constraint could suggest a steeper transition at $z > 7$, or perhaps shed light on tensions that may arise between luminosity function and quasar damping wing constraints. Furthermore, we emphasize that reionization is sensitive to alternative cosmological models, such as warm dark matter (e.g. Sitwell et al. 2014), making strong constraints on the reionization timeline a valuable venue to advance our understanding of dark matter – and any other processes that may impact cosmic reionization.

We note that our forecast is conservative due to the exclusion of cosmological information from CMB data. At the same time, our choice to use a single astrophysical parameter (ζ) to model reionization is optimistic. The reionization timeline depends on the ease with which ultraviolet photons escape from their host galaxies, i.e. the ionization efficiency ζ. However, the timeline, especially its speed, is also influenced by the ease of forming the parent galaxies of the ionizing photons, often parametrized by a minimum mass threshold $M_{\rm min}$. Naturally, including an additional free astrophysical parameter in the forecast would widen the error estimates in Table 2, as some of the constraining power would be redirected to account for the extra degree of freedom. As demonstrated in fig. 11 of Montero-Camacho & Mao (2021), the degeneracy between the two astrophysical parameters are significant. Consequently, we expect the error bars in Fig. 8 to increase. While this expected increase in the uncertainty of the $\overline{x}_{\rm H\,I}(z)$ constraint is a concern, we anticipate that the inclusion of CMB data will more than offset this effect.

7 SUMMARY

The cross-correlation between the Lyα forest and 21 cm intensity mapping in the post-reionization era is a promising cosmological probe of the relatively high-redshift intergalactic medium. In particular, it can probe smaller scales before they become fully non-linear (compared to that of traditional galaxy surveys) and it is pragmatically an easier measurement than the autocorrelation of any of those fields (at $z \sim 4$) due to foregrounds or available quasars line of sight. However, just as the Lyα forest (Montero-Camacho et al. 2019) and 21 cm intensity mapping (Long et al. 2023) are sensitive to relics from cosmic reionization, their cross-correlation will also be biased unless appropriate care is taken to handle this broad-band effect.

Regardless of the impact of reionization, our results demonstrate the importance of overlap between radio interferometers and Stage V spectroscopic instruments. Since some of these telescopes are currently in the early planning/design stage, we underscore the significance of guaranteeing a small degree of overlap between the instruments. As shown in Fig. 4, even 100 deg^2 of overlap could lead to a detection and consequently, it would enhance observational programmes aimed at the post-reionzation era.

Furthermore, we found that the gain in signal-to-noise is small for increased survey area in instrumental set-ups that use SKA1-Low. This trend is caused by the increase in the SKA1-Low thermal noise. Nonetheless, the PUMA × MUST set-up does exhibit significant gains, particularly at large scales with increased survey area, since the error is not dominated by thermal noise. In contrast, a longer integration time leads to considerably better measurements across the board, particularly for SKA configurations. Thus, we conclude that our baseline scenario, which we consider our *cheap* option with 100 h of integration time and 100 deg^2 of overlap, is a well-suited observational set-up although the measurements can likely be improved by further adjusting the survey strategy. For instance, a SKA1-Low × Stage V spectroscopic instrument is likely to perform comparatively to a Stage II radio interferometer at large scales with 1000 h of integration time (see Fig. 3). We emphasize that this result highlights the importance of progressing further than Stage V spectroscopic instruments in the 2040s.

The inclusion of reionization relics in the cross-correlation increases the strength of the signal, particularly at high-z and large

scales. Interestingly, late reionization produces an enhancement of up to 30 per cent over the predicted level without reionization imprints at $z \approx 4$. Separating this novel effect from the cosmological information – for instance using physics-inspired templates (Montero-Camacho, Liu & Mao 2023) – will allow for unbiased inference of cosmological parameters and it would unseal an original methodology to investigate the astrophysics of reionization. Future work will focus on mitigation strategies for this observable.

We have demonstrated the expected impact of reionization in the inference of cosmological parameters using a Fisher forecast. Cosmological parameters would be biased if one neglects the reionization relics and there would also be a significant underestimation of the error. However, if our knowledge of the reionization timeline improves significantly, the inclusion of the memory of reionization can result in a stronger constraint (as was the case for our PUMA × MUST forecast in Fig. 7). Furthermore, the cross-correlation of the Lyα forest and H I 21 cm IM will be a promising probe of reionization astrophysics in the next decades.

Our findings should also be interpreted as a cautionary tale – but simultaneously an exciting opportunity – to other post-reionization era high-z ($z \gtrsim 3$) tracers, like Lyα emission (e.g. Croft et al. 2016; Renard et al. 2021, 2024) and the CO rotational transitions (Breysse et al. 2022; Chung et al. 2022). Future work will assess the impact of reionization relics in other cosmological tracers of the post-reionization era.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their insightful suggestions. We are grateful to Chris Hirata and Yifan Zheng for their helpful suggestions and comments. This work was supported by The Major Key Project of PCL. We acknowledge the Tsinghua Astrophysics High-Performance Computing platform at Tsinghua University and PCL's Cloud Brain for providing computational and data storage resources that have contributed to the research results reported within this paper. This work made extensive use of the NASA Astrophysics DataSystem and the following open-source PYTHON libraries/packages: MATPLOTLIB (Hunter 2007), NUMPY (Harris et al. 2020), and SCIPY (Virtanen et al. 2020).

DATA AVAILABILITY

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

REFERENCES

Abdurashidova Z. et al., 2022, ApJ, 925, 221
Adams N. J. et al., 2023, MNRAS, 518, 4755
Ali S. S., Bharadwaj S., 2014, JA&A, 35, 157
Amiri M. et al., 2023, ApJ, 947, 16
Arinyo-i-Prats A., Miralda-Escudé J., Viel M., Cen R., 2015, J. Cosmol. Astropart. Phys., 2015, 017
Atek H. et al., 2023, MNRAS, 519, 1201
Bharadwaj S., Sarkar A. K., Ali S. S., 2015, JA&A, 36, 385
Bradley L. D. et al., 2023, ApJ, 955, 15
Braun R., Bonaldi A., Bourke T., Keane E., Wagg J., 2019, preprint (arXiv:1912.12699)
Breysse P. C., Yang S., Somerville R. S., Pullen A. R., Popping G., Maniyar A. S., 2022, ApJ, 929, 30
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
Carucci I. P., Villaescusa-Navarro F., Viel M., Lapi A., 2015, J. Cosmol. Astropart. Phys., 2015, 047
Carucci I. P., Villaescusa-Navarro F., Viel M., 2017, J. Cosmol. Astropart. Phys., 2017, 001
Castorina E., Villaescusa-Navarro F., 2017, MNRAS, 471, 1788
Castorina E., White M., 2019, J. Cosmol. Astropart. Phys., 2019, 025
Cen R., McDonald P., Trac H., Loeb A., 2009, ApJ, 706, L164
Chabanier S. et al., 2019, J. Cosmol. Astropart. Phys., 2019, 017
Chaussidon E. et al., 2023, ApJ, 944, 107
Choudhury T. R., Haehnelt M. G., Regan J., 2009, MNRAS, 394, 960
Chung D. T. et al., 2022, ApJ, 933, 186
Cosmic Visions 21 cm Collaboration, 2018, preprint (arXiv:1810.09572)
Crighton N. H. M. et al., 2015, MNRAS, 452, 217
Croft R. A. C. et al., 2016, MNRAS, 457, 3541
DESI Collaboration, 2022, AJ, 164, 207
DESI Collaboration et al., 2024, preprint (arXiv:2404.03001)
Dash C. B. V., Guha Sarkar T., 2021, J. Cosmol. Astropart. Phys., 2021, 016
Dash C. B. V., Sarkar T. G., Sarkar A. K., 2023, JA&A, 44, 5
Dayal P., Ferrara A., 2018, Phys. Rep., 780–782, 1
Diao K., Grumitt R. D. P., Mao Y., 2024, preprint (arXiv:2407.11296)
Donnan C. T. et al., 2023, MNRAS, 518, 6011
du Mas des Bourboux H. et al., 2020, ApJ, 901, 153
Ďurovčíková D. et al., 2024, ApJ, 969, 162
Facchinetti G., Lopez-Honorez L., Qin Y., Mesinger A., 2024, J. Cosmol. Astropart. Phys., 2024, 34
Filbert S. et al., 2024, MNRAS, 532, 3669
Font-Ribera A., McDonald P., Mostek N., Reid B. A., Seo H.-J., Slosar A., 2014, J. Cosmol. Astropart. Phys., 2014, 023
Furlanetto S. R., Lidz A., 2007, ApJ, 660, 1030
Gardner J. P. et al., 2006, Space Sci. Rev., 123, 485
Gnedin N. Y., 2022, ApJ, 937, 17
Gordon C. et al., 2023, J. Cosmol. Astropart. Phys., 2023, 045
Greig B., Mesinger A., 2018, in Jelić V., van der Hulst T., eds, IAU Symp. Vol. 333, Peering towards Cosmic Dawn. Astron. Soc. Pac., San Francisco, p. 18
Greig B., Mesinger A., Davies F. B., Wang F., Yang J., Hennawi J. F., 2022, MNRAS, 512, 5390
Greig B. et al., 2024, MNRAS, 530, 3208
Guha Sarkar T., Datta K. K., 2015, J. Cosmol. Astropart. Phys., 2015, 001
Guha Sarkar T., Bharadwaj S., Choudhury T. R., Datta K. K., 2011, MNRAS, 410, 1130
Guy J. et al., 2023, AJ, 165, 144
Harris C. R. et al., 2020, Nature, 585, 357
Heavens A., 2009, preprint (arXiv:0906.0664)
Hirata C. M., 2018, MNRAS, 474, 2173
Hoag A. et al., 2019, ApJ, 878, 12
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Hutter A., Heneka C., Dayal P., Gottlöber S., Mesinger A., Trebitsch M., Yepes G., 2023, MNRAS, 525, 1664
Jin X. et al., 2023, ApJ, 942, 59
Karim M. L. A., Armengaud E., Mention G., Chabanier S., Ravoux C., Lukić Z., 2024, J. Cosmol. Astropart. Phys., 2024, 29
Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2020, MNRAS, 491, 1736
Kim T. S., Bolton J. S., Viel M., Haehnelt M. G., Carswell R. F., 2007, MNRAS, 382, 1657
La Plante P., Mirocha J., Gorce A., Lidz A., Parsons A., 2023, ApJ, 944, 59
Lee K.-G., Cen R., Gott J. Richard I., Trac H., 2008, ApJ, 675, 8
Lewis A., 2008, Phys. Rev. D, 78, 023002
Long H., Hirata C. M., 2023, MNRAS, 520, 948
Long H., Givans J. J., Hirata C. M., 2022, MNRAS, 513, 117
Long H., Morales-Gutiérrez C., Montero-Camacho P., Hirata C. M., 2023, MNRAS, 525, 6036
Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2
Mason C. A. et al., 2019, MNRAS, 485, 3947
McDonald P., Eisenstein D. J., 2007, Phys. Rev. D, 76, 063009
McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499
Mertens F. G. et al., 2020, MNRAS, 493, 1662

Haiman Z., 2016, in A. Mesinger, Springer International Publishing Switzerland, Understanding the Epoch of Cosmic Reionization. Astrophysics and Space Science LibraryVol. 423, p.1
S. Furlanetto , 2019, in A. Mesinger , The Cosmic 21-cm Revolution; Charting the first billion years of our universe, IOP Publishing, Bristol, UK, p.1
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Mesinger A., Aykutalp A., Vanzella E., Pentericci L., Ferrara A., Dijkstra M., 2015, MNRAS, 446, 566
Montero-Camacho P., Mao Y., 2020, MNRAS, 499, 1640
Montero-Camacho P., Mao Y., 2021, MNRAS, 508, 1262
Montero-Camacho P., Hirata C. M., Martini P., Honscheid K., 2019, MNRAS, 487, 1047
Montero-Camacho P., Liu Y., Mao Y., 2023, MNRAS, 520, 4853
Montero-Camacho P., Li Y., Cranmer M., 2024a, preprint (arXiv:2405.13680)
Montero-Camacho P., Zhang Y., Mao Y., 2024b, MNRAS, 529, 3666
Morales A. M., Mason C. A., Bruton S., Gronke M., Haardt F., Scarlata C., 2021, ApJ, 919, 120
Munshi S. et al., 2024, A&A, 681, A62
Murray S., Greig B., Mesinger A., Muñoz J., Qin Y., Park J., Watkinson C., 2020, J. Open Source Softw., 5, 2582
Ouchi M. et al., 2010, ApJ, 723, 869
Palanque-Delabrouille N. et al., 2013a, A&A, 551, A29
Palanque-Delabrouille N. et al., 2013b, A&A, 559, A85
Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2020, J. Cosmol. Astropart. Phys., 2020, 038
Park J., Mesinger A., Greig B., Gillet N., 2019, MNRAS, 484, 933
Park H., Lukić Z., Sexton J., Alvarez M. A., Shapiro P. R., 2024, ApJ, 969, 46
Percival W. J. et al., 2019, preprint (arXiv:1903.03158)
Planck Collaboration XIII, 2016, A&A, 594, A13
Planck Collaboration VI, 2020, A&A, 641, A6
Pober J. C. et al., 2014, ApJ, 782, 66
Puchwein E. et al., 2022, MNRAS, 519, 6162
Ramírez-Pérez C. et al., 2023, MNRAS, 528, 6666
Ravoux C. et al., 2023, MNRAS, 526, 5118
Renard P. et al., 2021, MNRAS, 501, 3883
Renard P., Spinoso D., Sun Z., Zou H., Montero-Camacho P., Cai Z., 2024, MNRAS, 535, 826
Richard J. et al., 2019, The Messenger, 175, 50
Roth J. T., D'Aloisio A., Cain C., Wilson B., Zhu Y., Becker G. D., 2024, MNRAS, 530, 5209
Sarkar A. K., Bharadwaj S., Guha Sarkar T., 2018, J. Cosmol. Astropart. Phys., 2018, 051
Sarkar A. K., Pal A. K., Guha Sarkar T., 2019, J. Cosmol. Astropart. Phys., 2019, 058
Schlegel D., Kollmeier J. A., Ferraro S., 2019, BAAS, 51, 229
Sitwell M., Mesinger A., Ma Y.-Z., Sigurdson K., 2014, MNRAS, 438, 2664
Slosar A. et al., 2019, BAAS, 51, 53
Sobacchi E., Mesinger A., 2015, MNRAS, 453, 1843
Spina B., Bosman S. E. I., Davies F. B., Gaikwad P., Zhu Y., 2024, A&A, 688, L26
Springel V., 2005, MNRAS, 364, 1105
Square Kilometre Array Cosmology Science Working Group, 2020, PASA, 37, e007
Sun Z., Ting Y.-S., Cai Z., 2023, ApJS, 269, 4
Upton Sanderbeck P., Bird S., 2020, MNRAS, 496, 4372
Villaescusa-Navarro F., Viel M., Alonso D., Datta K. K., Bull P., Santos M. G., 2015, J. Cosmol. Astropart. Phys., 2015, 034
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Visbal E., Loeb A., Wyithe S., 2009, J. Cosmol. Astropart. Phys., 2009, 030
Wang B. et al., 2022, ApJS, 259, 28
Weinberg D. H., Davé R., Katz N., Kollmeier J. A., 2003, in Holt S. H., Reynolds C. S., eds, AIP Conf. Ser. Vol. 666, The Emergence of Cosmic Structure. Am. Inst. Phys., New York, p. 157
Wolz L. et al., 2015, in T. Bourke, Advancing Astrophysics with the Square Kilometre Array (AASKA14) Proceedings of Science, Giardini Naxos, Italy, p.35
Wyithe J. S. B., Loeb A., 2009, MNRAS, 397, 1926
Yang J. et al., 2023, ApJS, 269, 15
Yèche C., Palanque-Delabrouille N., Baur J., du Mas des Bourboux H., 2017, J. Cosmol. Astropart. Phys., 2017, 047
Yèche C. et al., 2020, RNAAS, 4, 179
Zhang Y. et al., 2024, MNRAS, 530, 1235
Zhao X., Mao Y., Wandelt B. D., 2022, ApJ, 933, 236
Zhou M., Tan J., Mao Y., 2021, ApJ, 909, 51
Zuo S., Chen X., Mao Y., 2023, ApJ, 945, 38

APPENDIX A: DEPENDENCE OF FORECAST ON $\delta\zeta$

The ionization efficiency ζ controls the timing of reionization in 21CMFAST by parametrizing the difficulty for an ionizing photon to escape into the IGM. Consequently, this parameter is not as well constrained or studied as cosmological parameters like σ_8 or n_s. Here, we confirm that the step size of this uncertain parameter does not have a significant impact on our results. We follow Facchinetti et al. (2024) strategy of checking the impact on the forecasted error instead of the performance of the numerical derivative.

Fig. A1 highlights the small difference between the forecasted ζ errors and justifies our choice of 3 per cent variation in ionization efficiency for our Fisher forecast.

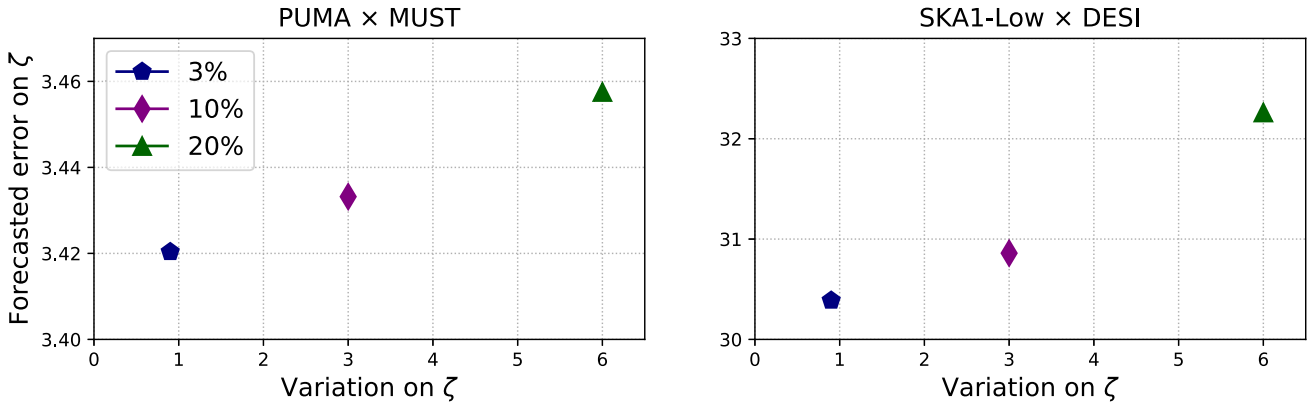

Figure A1. Convergence of the ionization efficiency error with step size. Shown are the forecasted errors for a variation of 3 per cent (blue pentagons), 10 per cent (purple diamonds), and 20 per cent (green triangles) of the fiducial ζ. Given the small difference between the forecasted errors, we opt for the 3 per cent variation in ionization efficiency.

APPENDIX B: TRENDS FOR THE FISHER SIMULATIONS

Here, we show some reionization-related properties of the simulations used in Section 6 to clarify their hierarchy.

For instance, it is clear that n_s will be the most constrained parameter by the inclusion of reionization relics since it dominates most of the different metrics in Fig. B1. In contrast, a weaker constraint in ionization efficiency is also expected, particularly since ionization efficiency does not affect the first term in equation (1).

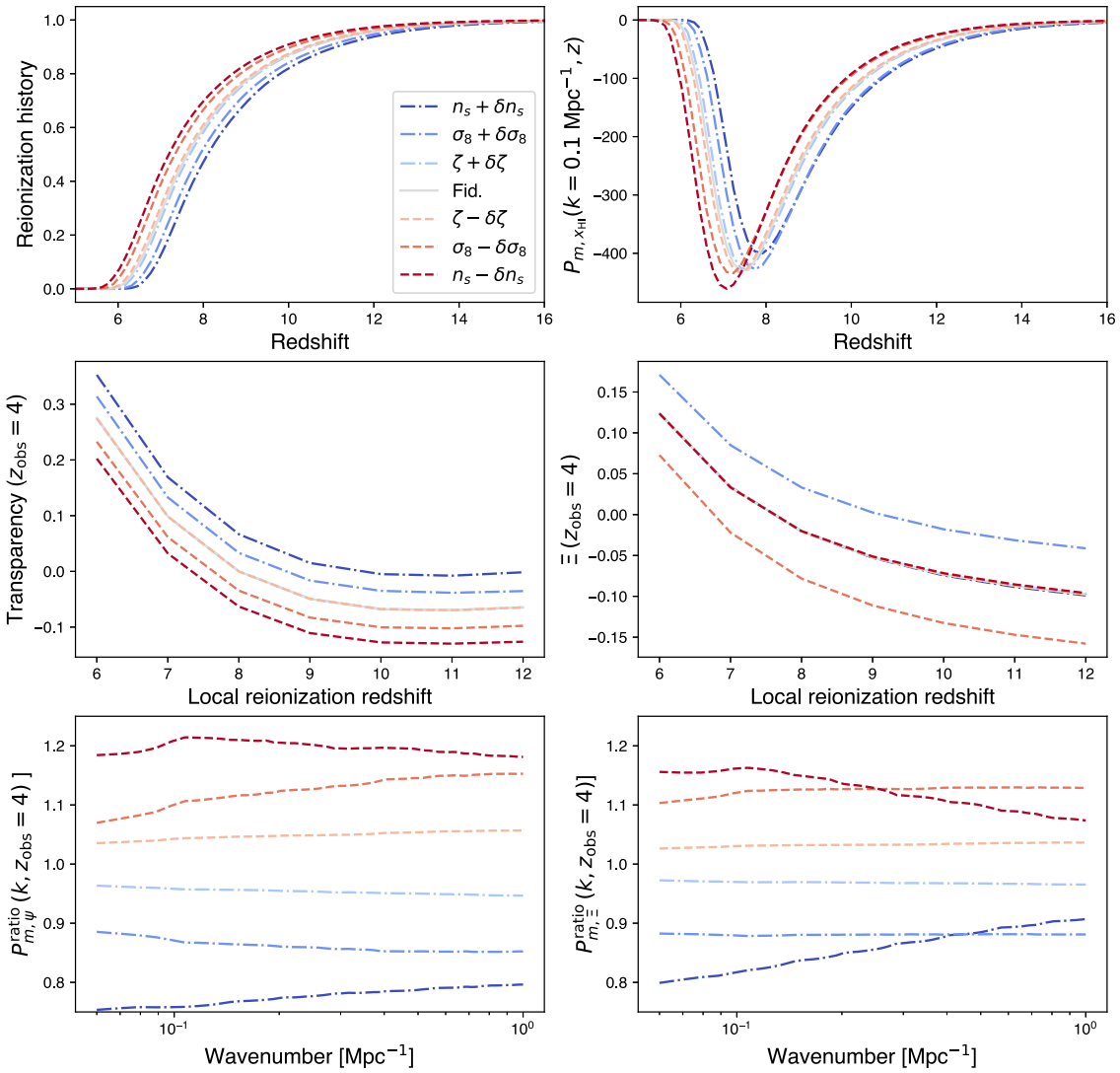

Figure B1. Properties of the simulations used in the Fisher forecast of Section 6. (Top left) Reionization history. (Top right) Cross-correlation of matter and neutral hydrogen field as a function of redshift and evaluated at a k representative of the ionized bubble scale. (Middle left) Relative transparency of the IGM ψ as a function of local reionization redshift, equation (5). (Middle right) Relative neutral hydrogen density Ξ, equation (6). (Bottom left) The cross-power spectrum of matter and transparency of the model divided by that of the fiducial simulation, i.e. equation (7) divided by the reference scenario. (Bottom right) Ratio of the cross-power spectrum of matter and relative neutral hydrogen density, i.e. equation (8) normalized by the reference scenario.

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