

UC Berkeley

UC Berkeley Previously Published Works

Title

Optical depth to reionization from perturbative 21 cm clustering

Permalink

<https://escholarship.org/uc/item/9cc3w789>

Journal

Journal of Cosmology and Astroparticle Physics, 2022(10)

ISSN

1475-7516

Authors

Sailer, Noah

Chen, Shi-Fan

White, Martin

Publication Date

2022-10-01

DOI

10.1088/1475-7516/2022/10/007

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Optical depth to reionization from perturbative 21cm clustering

Noah Sailer,^{a,b,c} Shi-Fan Chen^{a,b} and Martin White^{a,b,c}

^aDepartment of Physics, University of California, Berkeley, CA 94720, USA

^bBerkeley Center for Cosmological Physics, UC Berkeley, CA 94720, USA

^cLawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

E-mail: nsailer@berkeley.edu, shifan.chen@berkeley.edu, mwhite@berkeley.edu

Abstract. The optical depth τ is the least well determined parameter in the standard model of cosmology, and one whose precise value is important for both understanding reionization and for inferring fundamental physics from cosmological measurements. We forecast how well future epoch of reionization experiments could constraint τ using a symmetries-based bias expansion that highlights the special role played by anisotropies in the power spectrum on large scales. Given a parametric model for the ionization evolution inspired by the physical behavior of more detailed reionization simulations, we find that future 21cm experiments could place tight constraints on the timing and duration of reionization and hence constraints on τ that are competitive with proposed, space-based CMB missions provided they can measure $k \approx 0.1 h \text{ Mpc}^{-1}$ with a clean foreground wedge across redshifts spanning the most active periods of reionization, corresponding to ionization fractions $0.2 \lesssim x \lesssim 0.8$. Significantly improving upon existing CMB-based measurements with next-generation 21cm surveys would require substantially longer observations (~ 5 years) than standard $\mathcal{O}(1000 \text{ hour})$ integration times. Precise measurements of smaller scales will not improve constraints on τ until a better understanding of the astrophysics of reionization is achieved. In the presence of noise and foregrounds even future 21cm experiments will struggle to constrain τ if the ionization evolution deviates significantly from simple parametric forms.

Contents

1	Introduction	1
2	Extracting the optical depth from 21cm clustering	4
2.1	The optical depth as a derived parameter	4
2.2	Ionization fraction from 21cm clustering	5
2.3	Parametric reionization history	6
3	The 21cm clustering signal in Lagrangian Perturbation Theory	6
4	Forecasts	9
4.1	Model and fiducial scenario	9
4.2	Idealistic survey	10
4.3	More realistic scenarios	12
5	Discussion and conclusions	15

1 Introduction

Within the Standard (Λ CDM) Model of Cosmology the optical depth to reionization τ is the least well-determined of the standard parameters [1]. Though it is an almost purely astrophysical parameter its uncertainty has a large impact on inferences about fundamental physics because, when fitting cosmic microwave background (CMB) data, it is strongly degenerate with parameters like the amplitude of primordial fluctuations (A_s , σ_8) and the sum of the neutrino masses M_ν . In addition to hindering higher-precision measurements of these parameters, τ derives from the Thomson scattering of free electrons and is thus a probe of the history of cosmic reionization [2–4]. Beyond suppressing the amplitude of CMB temperature anisotropies this Thomson scattering results in a range of physical phenomena in the CMB including large-scale (E-mode) polarization [5] and the kinetic Sunyaev-Zeldovich (kSZ) effect [6–9] that can be used to measure τ . In addition, probes of the Epoch of Reionization (EoR) have the potential to directly constrain the global ionization history and therefore the abundance of free electrons that source τ [10–15]. Our goal in this paper is to investigate this final method using recent advances combining cosmological perturbation theory with techniques based on effective field theories.

There have been numerous approaches to modeling the 21cm signal from the EoR, ranging from detailed radiative transfer simulations coupled to hydrodynamics [16, 17] to semi-analytic [18] and hybrid [19] models. The majority are based on N-body simulations. By contrast, our approach to modeling 21cm observations will be based upon cosmological perturbation theory with a symmetries-based bias expansion. This approach is now quite commonplace in large-scale structure studies, but differs in some key respects from the standard approaches within the reionization field (see references above for examples of analytic or semi-numerical models and refs. [20–22] for notable examples similar in spirit to our approach). No matter how complex the physical processes leading to ionized regions, as long as they obey certain symmetries (e.g. rotational invariance and the equivalence principle) and are quasi-local we can write δ_{21} as an expansion in powers of the density field, tidal field and

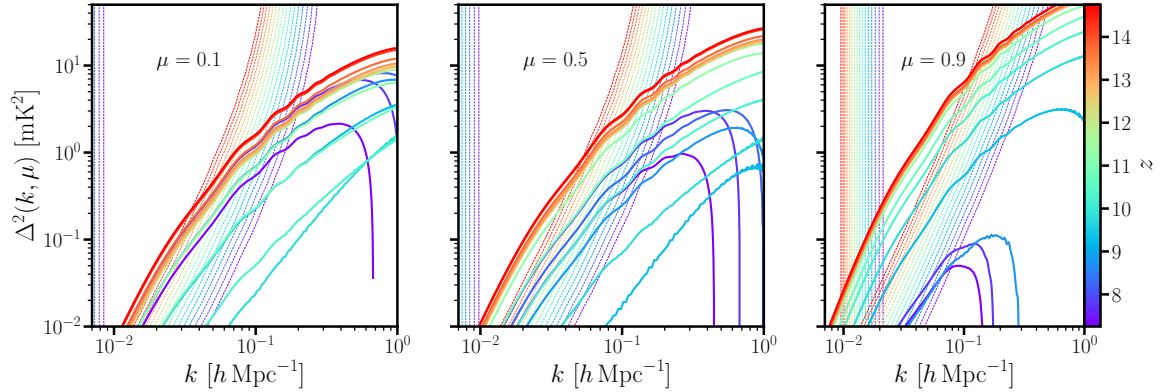


Figure 1. The dimensionless 21cm power spectrum ($\Delta^2 = k^3 P / 2\pi^2$) as a function of redshift for our fiducial model (solid lines, see §4.1) and the thermal noise per mode (steeply rising, dashed lines) for the case of 1000 14m dishes with a 10 year integration time (first column of Table 1). The step in thermal noise (vertical lines) at low- k is due to a absence of baselines capable of measuring modes with $k_{\perp} \chi / 2\pi < D / \lambda_{21}^{\text{obs}}$, where D is the diameter of the dishes. The panels show $\mu = 0.1$ (left), $\mu = 0.5$ (middle) and $\mu = 0.9$ (right) with colors indicating redshift. Note that the spectrum is quite anisotropic (the amplitudes vary significantly with μ , particularly at lower redshifts), with the shape encoding valuable information about the cosmology and reionization history.

low-order derivatives [23, 24]. This expansion will be valid on large scales, in our case on scales larger than the typical size of an ionized bubble [21, 22]. The impact of small-scale physics not explicitly included in the model is then handled by correction terms (i.e. counterterms), which are also restricted by symmetry considerations. By marginalizing over the bias parameters and counterterms we can focus on the cosmological information contained in large scale modes while efficiently removing sensitivity to small-scale physics. Since our expansion is based entirely on symmetry, it makes minimal assumptions about that physics. Conversely, information about the sources and morphology of reionization is encapsulated in the small-scale modes and bias coefficients.

Given that perturbation theory and the bias expansion marginalize away much of the “astrophysics” of the 21cm signal it is worth asking whether it is still possible to recover useful information about reionization within such a framework, or whether the study of reionization necessarily requires detailed modeling of the small-scale physics of the EoR (e.g. the composition of ionizing sources). For example, while the overall amplitude of the observed 21cm power spectrum is proportional to the brightness temperature \bar{T}_b and can therefore be used to predict the electron fraction (§2.2), even at linear order in perturbation theory the power spectrum will also be proportional to the square of the linear bias b^2 whose value cannot be determined absent detailed knowledge of small-scale behavior of neutral hydrogen, the host halos of ionizing sources etc., leading to an exact degeneracy with the electron fraction.

However, while many potential measurements of the global reionization history, e.g. the overall amplitude of the 21cm signal, are polluted by small-scale physics, the presence of fundamental symmetries implies that others are protected. For example, the equivalence principle dictates that all matter behaves identically under the force of gravity, such that the peculiar velocity of neutral hydrogen on large scales must be the same¹ as that of matter

¹The relative velocity between baryons and dark matter is discussed in §5.

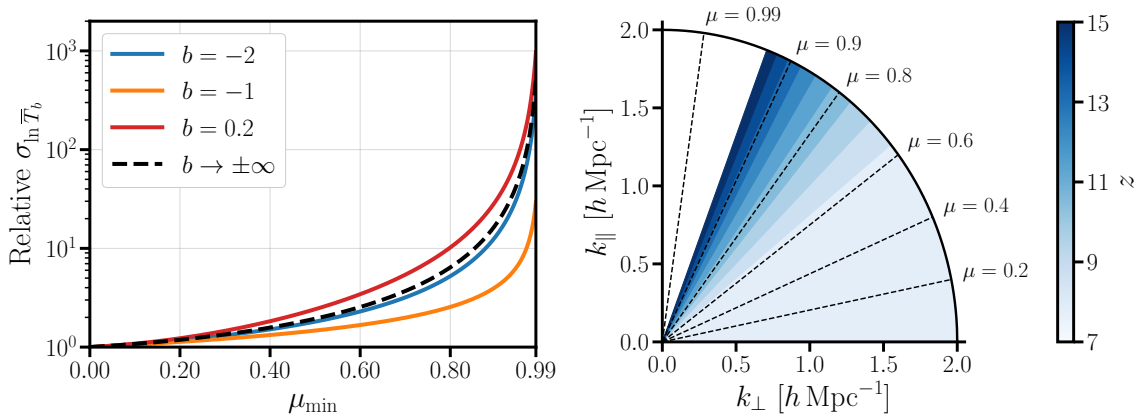


Figure 2. *Left:* Error on $\ln \bar{T}_b$ (normalized to $\mu_{\min} = 0$) after marginalizing over the linear bias, as a function of μ_{\min} , following the simplified scenario outlined in §1. The colors denote the assumed fiducial value of the linear bias. *Right:* Foreground wedge for our fiducial scenario ($N_w = 3$; Eq. 4.5) as a function of redshift, assuming 14m diameter dishes (see §4.3 for a more detailed discussion of the foreground wedge).

and other tracers like galaxies. Since 21cm surveys infer line-of-sight distances in terms of observed frequencies these peculiar velocities are then encoded in anisotropies in the observed clustering due to redshift-space distortions (RSD), where an object at \mathbf{x} appears to be at redshift-space position $\mathbf{s} = \mathbf{x} + \mathbf{u}$, with $\mathbf{u} = \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v}_{\text{pec}})/\mathcal{H}$. Within linear theory the redshift-space power spectrum is given by [25]

$$P_{\text{obs}}(k) = \bar{T}_b^2 (b + f\mu^2)^2 P_{\text{lin}}(k), \quad \mu = \hat{\mathbf{n}} \cdot \hat{\mathbf{k}} \quad (1.1)$$

where $\hat{\mathbf{n}}$ is the line of sight unit vector and $f(z) = d \ln D / d \ln a$ is the growth rate, with $f \simeq 1$ to sub-percent precision at the matter dominated redshifts where reionization occurs (see Fig. 1). As the growth rate and P_{lin} are purely a function of cosmology, the size of the anisotropy (μ^2) on large scales is then uncontaminated by small-scale physics and for any given cosmology can be used to infer the brightness temperature \bar{T}_b , unlike the isotropic piece proportional to b^2 [10, 11]. In addition to the linear bias, \bar{T}_b is perfectly degenerate in linear theory with the amplitude of matter fluctuations σ_8 within a single redshift bin. This degeneracy can be broken² using the different time-evolution of $\sigma_8(z) \propto (1+z)^{-1}$ and $\bar{T}_b(z) \propto (1+z)^{1/2} [1-x(z)]$, or by combing 21cm observations with primary CMB fluctuations (which constrain $\sigma_8 e^{-\tau}$).

The left panel of Figure 2 shows our relative ability to measure the brightness temperature as a function of the minimum angular scale μ_{\min} in an idealized setup³. Clearly, since

²At lower redshifts non-linear evolution of large-scale structure can be used to break this degeneracy [26, 27], but at the early times relevant for EoR such effects are very small on the scales that can be reliably modeled.

³Specifically, we consider a scenario where the survey is noiseless and the signal is given by linear theory, and the treat the linear power spectrum P_{lin} as fixed. Our ability to constraint \bar{T}_b is then independent of the linear power spectrum shape and the Fisher matrix is given by

$$F_{ij} \propto \int_{\mu_{\min}}^1 d\mu \left[\partial_i (b + \mu^2)^2 \bar{T}_b^2 \right] \left[\partial_j (b + \mu^2)^2 \bar{T}_b^2 \right] / (b + \mu^2)^4 \bar{T}_b^4,$$

where $i \in \{\bar{T}_b, b\}$.

our ability to constrain \bar{T}_b depends on resolving the anisotropy due to RSD, reducing the angular range over which we fit results in dramatic reductions in constraining power (see also ref. [28]). This is particularly relevant for 21cm interferometers wherein chromaticity in instrument leads to foreground leakage into a “wedge” below some μ_{\min} that renders those modes unusable (§4.3). The wedge angle corresponding to our fiducial setup is shown in the right panel of Figure 2.

Our goal in this paper will be to investigate the feasibility of using the EoR 21cm signal in redshift space to constrain the global reionization history within a perturbation theory framework. The remainder of this paper is organized as follows: Section 2 discusses the relationship between the optical depth and reionization as probed by 21cm measurements. This section describes our parameterizations of the ionization fraction evolution. Section 3 gives a quick overview of Lagrangian perturbation theory and the symmetries-based bias expansion that we employ throughout this work. Section 4 discusses the Fisher matrix formalism that we employ in our forecasts, and the fiducial models and scale cuts that we investigate. We conclude in Section 5 with a discussion of implications from our forecasts.

2 Extracting the optical depth from 21cm clustering

In this section we describe our method for constraining τ from a measurement of 21cm clustering during the reionization era. We start by identifying the physical quantities that we need to constrain in order to infer τ and then describe how these are related to the 21cm power spectrum before discussing how we choose to parameterize the evolving ionization fraction.

2.1 The optical depth as a derived parameter

We begin by reviewing how the optical depth can be computed given a cosmological background and a reionization history (i.e. the evolution of the ionization fraction) [12]. The mean CMB optical depth is by definition $\tau = \sigma_T \int_0^{z_{\text{rec}}} dz \bar{n}_e(z) dr/dz$, where σ_T is the Thompson cross-section, $\bar{n}_e(z)$ is the volume-averaged free electron (proper) number density, $z_{\text{rec}} \simeq 1100$ is redshift of last scattering, and dr/dz is the line-of-sight (proper) distance per unit redshift. Throughout we assume that the Universe is flat, in which case $dr/dz = c/(1+z)H(z)$, and use overlines to denote *volume* averages. We assume that free electrons are sourced solely from hydrogen and (singly ionized) helium, neglecting the free electrons from doubly-ionized helium⁴ and additional isotopes of H and He, as well as heavier elements. We further assume that hydrogen and helium have identical ionization fractions: $x_{\text{HII}} \equiv x_{\text{HeII}}$. This is a standard approximation in the literature [12, 29, 30], due to the similar ionization potentials (13.6 vs 24.6 eV) of the two elements and the fact that neutral H efficiently shields the He from ionizing radiation. Given these assumptions, the free electron number density can be rewritten as $n_e(\mathbf{x}) = x_{\text{HII}}(\mathbf{x})\rho_b(\mathbf{x})/\mu m_{\text{H}}$, where ρ_b is the baryon density, $\mu = 1 + (m_{\text{He}}/m_{\text{H}} - 1)Y_p/4$ is the mean molecular weight and $Y_p = 4n_{\text{He}}/n_b$ is the primordial helium abundance. It is standard to rewrite the baryon density as $\rho_b(\mathbf{x}) \equiv \bar{\rho}_b(z)[1 + \delta_b(\mathbf{x})]$, where $\bar{\rho}_b(z) = 3H_0^2\Omega_b(1+z)^3/8\pi G$ is the mean baryon density, and δ_b is the density contrast. With these definitions and as-

⁴Ref. [12] performs a back-of-the-envelope calculation to show that doubly-ionized helium contributes to $\sim 1.4\%$ of the total optical depth. Assuming that the best fit $\tau = 0.0561$ from Planck is the truth, this amounts to $\tau_{\text{HeIII}} \sim 0.0008$.

sumptions in hand, the optical depth can be rewritten as [12]

$$\begin{aligned}\tau &= \frac{3H_0^2\Omega_b\sigma_{TC}}{8\pi G\mu m_H} \int_0^{z_{\text{rec}}} dz \frac{(1+z)^2}{H(z)} x_{\text{HII}}(1+\delta_b)(z) \\ &= 0.0691 \frac{\omega_b}{h\mu} \int_0^{z_{\text{rec}}} dz \frac{(1+z)^2}{E(z)} x(z),\end{aligned}\tag{2.1}$$

where $\omega_b = \Omega_b h^2$, $H_0 \equiv 100 h \text{ km/s/Mpc}$, $E(z) = H(z)/H_0$, and we've defined the *mass-weighted* average ionization fraction $x(z) \equiv x_{\text{HII}}(1+\delta_b)(z)$. In going from the first to second line of Eq. (2.1) we have substituted the values of quantities which are known to well beyond the precision of our τ measurement. In particular we see that τ may be derived from the baryon fraction ω_b , the expansion history $E(z)$, the primordial helium abundance $\mu = 1 + 0.7428 Y_p$, and the ‘‘reionization history’’ $x(z)$.

2.2 Ionization fraction from 21cm clustering

The 21cm brightness temperature, defined as the difference between the observed 21cm brightness and that of the CMB, at redshift-space position \mathbf{s} is given by [31, 32]:

$$\begin{aligned}T_b(\mathbf{s}, z) &= \frac{3}{16} \frac{\hbar c^3 A_{10}}{k_B \nu_{21}^2} \frac{[1 - x_{\text{HII}}(\mathbf{x}, z)] n_{\text{H}}(\mathbf{x}, z)}{(1+z)(dv_{\parallel}/dr_{\parallel})} \left[1 - \frac{T_{\text{CMB}}(z)}{T_s(\mathbf{x}, z)} \right] \\ &\equiv \bar{T}_b(z) [1 + \delta_{21}(\mathbf{s}, z)]\end{aligned}\tag{2.2}$$

where \mathbf{x} is the corresponding position in real space, $\langle \delta_{21} \rangle = 0$, $A_{10} = 2.85 \times 10^{-15} \text{ Hz}$ is the spontaneous emission coefficient of the 21cm transition, $\nu_{21} = 1.42 \text{ GHz}$ is the rest frame 21cm transition frequency and T_s is the spin temperature, defined so that the ratio of the number of excited (spin aligned) hydrogen atoms to the number of hydrogen atoms in the ground state (spins anti-aligned) is given by $3 \exp[-2\pi\hbar\nu_{21}/k_B T_s]$ ⁵. The derivative $dv_{\parallel}/dr_{\parallel}$ is the (proper) gradient of the (proper) velocity along the line of sight and sources the redshift space anisotropy in the observed signal; dividing a real-space density by this factor yields the redshift-space density. We assume that the spin temperature dominates over the CMB ($T_s \gg T_{\text{CMB}}$), and discuss the limitations of this approximation in §4.3. Under these assumptions, the mean 21cm brightness temperature takes the form:

$$\bar{T}_b(z) = 189.1 [1 - x(z)] \frac{(1+z)^2}{E(z)} \frac{\omega_b}{h\mu} \left(1 - \frac{Y_p}{4} \right) \text{ mK},\tag{2.4}$$

where we have substituted the values of known quantities.

A 21cm interferometer directly measures the Fourier modes $T_b(\mathbf{k})$, from which one measures the power spectrum:

$$P_{\text{obs}}(\mathbf{k}, z) = \bar{T}_b^2(z) P_{21}(\mathbf{k}, z) + P_N(\mathbf{k}, z),\tag{2.5}$$

⁵This expression follows from the more general form

$$T_b(\mathbf{s}, z) = \frac{T_s - T_{\text{CMB}}}{1+z} (1 - e^{-\tau_{21}(\mathbf{s})})\tag{2.3}$$

where τ_{21} is the 21cm optical depth, which in the small τ_{21} limit gives $1 - e^{-\tau_{21}} \approx \tau_{21}$. Since the 21cm signal comes from the diffuse, optically-thin IGM this approximation is expected to be a good one, but it is worth noting that in the opposite, optically thick limit the nonlinearity of the exponential mapping tends to contaminate the linear redshift-space signal critical to obtaining τ constraints. This is due to the fact that the exponential mapping breaks the number conservation symmetry that protects the Kaiser form of RSD [33].

where $\langle \delta_{21}(\mathbf{k})\delta_{21}(\mathbf{k}') \rangle = (2\pi)^3 \delta^D(\mathbf{k} + \mathbf{k}') P_{21}(\mathbf{k})$ and $P_N(\mathbf{k}, z)$ is the thermal noise of the instrument. In particular, we note that the mean 21cm brightness temperature controls the overall amplitude of the observed (redshift space) power spectrum of the 21cm signal. This mean temperature is in turn proportional to the neutral fraction $1 - x$. Thus, by measuring the amplitude of the 21cm power spectrum one can in turn infer a measurement of the ionization fraction x .

2.3 Parametric reionization history

The presence of a series of acoustic peaks in the angular power spectra of the CMB indicates that the Universe was dense and ionized at early times and then underwent a rapid transition to being (largely) neutral at $z \simeq 1100$. This neutral period lasted for a significant time before the first astrophysical sources (in currently popular models, massive stars in relatively low mass early galaxies) reionized the Universe. Lower limits on the redshift of reionization come from the lack of complete absorption of UV photons from high redshift QSOs [34]. While significant questions remain, there are numerous observational constraints on the process of reionization [2–4]. Current observations point towards a “late and fast” reionization period, though with considerable uncertainty [35].

Rather than use a simulation-based model to infer the evolution of the ionization fraction from a set of physical parameters (star formation rates, escape fractions, baryon content, etc.), we will instead take a more astrophysics-agnostic approach and adopt an analytic parameterization of $x(z)$. Ref. [19] found that the Weibull function, which has three free parameters:

$$x(z) = \text{Exp} \left[-\max \left(\frac{z - a}{b}, 0 \right)^c \right], \quad (2.6)$$

is capable of accurately fitting a wide range of reionization scenarios. We will adopt this Weibull form as the default model in our forecasts. A comparison of the best-fit Weibull function with our fiducial reionization history is shown in the left panel of Fig. 3.

In §4 we will consider how freeing up the functional form of $x(z)$ impacts our constraints. To examine this we will model $x(z)$ with Lagrange interpolating functions, as proposed by ref. [36]. The parameters of the model are a set of mass-weighted ionization fractions $\{x(z_i)\}$ for a chosen (fixed) set of redshifts $\{z_i\}$. These points are smoothly connected via:

$$x(z) = \min \left[\exp \left(\sum_{i=1}^N p_i(z) \right), 1 \right] \quad \text{where} \quad p_i(z) \equiv \ln(x(z_i)) \prod_{j \neq i} \frac{\ln(1+z) - \ln(1+z_j)}{\ln(1+z_i) - \ln(1+z_j)}. \quad (2.7)$$

We consider how our constraints vary as a function of the number of free parameters in §4.2.

3 The 21cm clustering signal in Lagrangian Perturbation Theory

To robustly account for scale-dependent bias and structure formation in the quasi-linear regime we will adopt the formalism of Lagrangian Perturbation Theory (LPT). The literature on perturbative modeling of large-scale structure is vast and we will not try to give a comprehensive overview of the subject or rederive results; rather, our goal is to give a reasonably self-contained summary and highlight important physical points distinguishing the 21cm signal from the more standard treatment of galaxies. A review of perturbation theory can be found in refs. [37, 38], and of bias in ref. [24].

Within LPT cosmological structure formation is modeled as the gravitational dynamics of an effective fluid whose short-wavelength (UV) physics are integrated out, with corresponding effects on large-scale structure parametrized by effective-theory counterterms and stochastic contributions [39, 40]. Unlike in Eulerian effective theories of large-scale structure (see e.g. [41, 42]), which model the resulting fluid via its densities and velocity potentials, within the Lagrangian picture of fluid dynamics one traces the trajectories of infinitesimal fluid elements labeled by their initial (Lagrangian) positions \mathbf{q} . The dynamics are encoded in a displacement vector $\Psi(\mathbf{q}, \eta)$ defined such that the Eulerian (comoving) positions \mathbf{x} of the fluid elements at a conformal time η is $\mathbf{x}(\mathbf{q}, \eta) = \mathbf{q} + \Psi(\mathbf{q}, \eta)$, and where the displacement vector is sourced by the gravitational potential to follow $\ddot{\Psi}(\mathbf{q}) + \mathcal{H}\dot{\Psi}(\mathbf{q}) = -\nabla_{\mathbf{x}}\Phi(\mathbf{x})$, with \mathcal{H} defined as the conformal Hubble parameter. The gravitational potential in turn follows Poisson’s equation, with the matter (over)density obtained via number conservation:

$$1 + \delta_{\text{m}}(\mathbf{x}, \eta) = \int d^3\mathbf{q} \delta^D(\mathbf{x} - \mathbf{q} - \Psi(\mathbf{q}, \eta)). \quad (3.1)$$

The velocity field is given straightforwardly as the time derivative of the displacement.

In general, cosmological observables like the 21cm signal do not directly probe the matter field described by the displacements Ψ but rather are determined by complex small-scale astrophysical phenomena and their response to the large-scale matter distribution over cosmic time [24]. For example, whether a galaxy exists at a spacetime point (\mathbf{x}, η) will generically depend on the matter density, velocity gradients and tidal fields in a neighborhood about its trajectory $\mathbf{q} + \Psi(\mathbf{q}, \eta)$ approximately the size of its Lagrangian halo radius R_h . While the size of the nonlocality in the galaxy bias expansion is governed primarily by R_h (since the dominant physical effect is the nonlinear gravitational collapse of the Lagrangian patch into a halo), the 21cm signal during reionization is sensitive to other, stronger nonlocalities such as the mean free path λ_{mfp} of ionizing photons, which unlike nonrelativistic matter can travel significant ($\sim H^{-1}$) distances in a Hubble time in the absence of absorption. In this case the size of nonlocalities is roughly set by the size of the ionized bubbles during reionization, with smaller bubbles earlier on in the process corresponding to our ability to perturbatively model the signal over a wider range of scales [21].

Within the (Lagrangian) effective-theory framework [23, 43–46] we can write the quasi-local dependence of the 21cm signal at \mathbf{x} on neighboring points along its trajectory — that is, nearby in \mathbf{q} — as an expansion order-by-order in the initial conditions at \mathbf{q}

$$F(\mathbf{q}) = 1 + b_1^{\text{L}}\delta(\mathbf{q}) + \frac{1}{2}b_2^{\text{L}}(\delta^2(\mathbf{q}) - \langle\delta^2\rangle) + b_s^{\text{L}}(s^2(\mathbf{q}) - \langle s^2\rangle) + \dots \\ + R_*^2\nabla^2\delta(\mathbf{q}) + \dots + \epsilon(\mathbf{q}). \quad (3.2)$$

This functional is then dynamically advected to the real-space position \mathbf{x} [47]:

$$1 + \delta_{21}(\mathbf{x}) = \int d^3\mathbf{q} F(\mathbf{q}) \delta^D(\mathbf{x} - \mathbf{q} - \Psi(\mathbf{q})) \quad (3.3)$$

in the same way as the matter field (Eq. 3.1).⁶ In this way the 21cm clustering signal is captured in LPT via a combination of dynamics (Ψ) and dependence on the initial conditions $F(\mathbf{q})$.

⁶Fluctuations in the 21cm signal can be sourced by fluctuations in either the baryon density δ_b or the electron fraction δ_x and treating these separately is common in 21cm studies. However, since it can equally be thought of as fluctuations in the density of neutral hydrogen, in this work we directly model this product using a single bias expansion instead of separately. This avoids needing to model correlated fluctuations in δ_b and δ_x when evaluating $\delta_b\delta_x$, which in perturbation theory language is known as a “composite operator”.

The form of the functional $F(\mathbf{q})$ is restricted by symmetries and the equivalence principle to depend only on scalar combinations of the density, velocity gradients and the tidal field. Within perturbation theory these quantities can all be written in terms of the initial density field, i.e. to linear order scalars formed from $\partial_i v_j$ and $\partial_i \partial_j \Phi \sim (k_i k_j / k^2) \delta$. To second order, contributions to the bias expansion can be summarized by its dependence on δ^2 and $s^2 = s_{ij} s_{ij}$, where $s_{ij} \equiv (\partial_i \partial_j / \partial^2 - \delta_{ij} / 3) \delta$ is the shear field. Beyond the local value of these initial conditions along the trajectory starting at \mathbf{q} , the existence of nonlocalities due to e.g. ionized bubbles can be handled by including derivative corrections such as $\nabla^2 \delta$, i.e. performing a Taylor series expansion of δ around \mathbf{q} . These corrections become increasingly prominent as we probe scales closer to the bubble size. Depending on how small in scale we need to fit to it may be necessary to include terms higher order in density and derivatives, though it is important to stress that our perturbative treatment of $F(\mathbf{q})$ as a quasi-local functional will necessarily break down on scales smaller than the bubble size where the physics of reionization are nonlocal.

The last term in Eq. 3.2 is the residual, $\epsilon(\mathbf{q})$, which represents stochastic small-scale contributions to the clustering of the 21cm signal uncorrelated with the large-scale modes we will be interested in. To lowest order these modes only correlate in close proximity $\langle \epsilon(\mathbf{q}) \epsilon(\mathbf{q}') \rangle \sim \delta^D(\mathbf{q} - \mathbf{q}')$ or, equivalently, have a flat spectrum in Fourier space: $P_\epsilon(k) = \text{const}$.

Finally, we need to account for contributions to the *observed* densities due to peculiar velocities in the form of redshift-space distortions (RSD). Specifically, since line-of-sight distances are inferred from redshifts they incur an additional Doppler contribution from the peculiar velocity $\mathbf{u} = \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v}_{\text{pec}}) / \mathcal{H}$, such that the 21cm signal in redshift-space is given by [48–50]

$$1 + \delta_{21}(\mathbf{s}) = \int d^3 \mathbf{q} F(\mathbf{q}) \delta^D(\mathbf{s} - \mathbf{q} - \Psi_s(\mathbf{q})) \quad ; \quad \Psi_s(\mathbf{q}) = \Psi + \mathbf{u}. \quad (3.4)$$

In the limit where the real-to-redshift-space mapping is unique ($\mathbf{s} \leftrightarrow \mathbf{x} \leftrightarrow \mathbf{q}$) we can integrate over the delta function to recover the familiar expression $\delta(\mathbf{s}) = aH / (dv_{\parallel} / dr) \delta(\mathbf{x})$, where v_{\parallel} is the line of sight velocity including Hubble flow [32]. Within the Einstein-de Sitter approximation, which is an excellent approximation at the redshifts where reionization occurs, the m^{th} order displacement obeys the simple relation

$$\Psi_s^{(m)}(\mathbf{q}, \eta) = \Psi^{(m)} + mf \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \Psi^{(m)} \right). \quad (3.5)$$

with $f \simeq 1$, though it is important to note that, like the real-space densities, the velocities too require effective-theory corrections to properly account for their dependence on small-scale physics [46]. To lowest order, the above calculation recovers the familiar Kaiser [25] form:

$$1 + \delta_{21} = \frac{1 + b_1^L \delta(\mathbf{q})}{1 + (aH)^{-1} dv_{\parallel}^{\text{pec}} / dr}$$

and, using that $v_{\text{pec},i} / (aH) = -f \nabla_i^{-1} \delta$ we have in Fourier space

$$\delta_{21}(\mathbf{k}) = ([1 + b_1^L] + f\mu^2) \delta(\mathbf{k}), \quad \mu = \hat{\mathbf{n}} \cdot \hat{\mathbf{k}}$$

such that the linear theory power spectrum is given by $(b + f\mu^2)^2 P_{\text{lin}}(k)$ [25], where $b = 1 + b_1^L$. This is the familiar result used in the introduction.

4 Forecasts

Now we turn to our forecasts of how well a (future) 21cm interferometer could determine τ by measuring the evolving 21cm power spectrum. Our forecasts use the Fisher matrix formalism, with an extensive set of nuisance terms to handle the complexities of the reionization process. These forecasts allow us to determine the key observations that would allow a robust extraction of τ and highlight critical assumptions. We begin by defining our fiducial cosmology, reionization history and 21cm clustering model before describing the range of scales and redshifts used in our fits.

4.1 Model and fiducial scenario

In this work we use CLASS [51] to compute the linear CDM+baryon power spectrum, and velocileptors⁷ [46] to model non-linearities and redshift-space distortions (see §3) to one-loop order within LPT:

$$P_{21}(\mathbf{k}) = P_{21}^{1\text{-loop}}(\mathbf{k}) + (\alpha_0 + \alpha_2\mu^2 + \alpha_4\mu^4)k^2 P_{cb}^{\text{Zel}}(\mathbf{k}) + N_0 + N_2(\mu k)^2 + N_4(\mu k)^4, \quad (4.1)$$

where N_2 and N_4 encode the effects of small-scale velocities, N_0 is the shot noise, P_{cb}^{Zel} is the CDM+baryon power spectrum in the Zel'dovich approximation, and $\mu = \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}$ is the cosine of the angle between the wavevector and the line of sight. This model includes a bias expansion to quadratic order in the linear CDM+baryon density contrast (i.e. the first line of Eq. 3.2). For further details on the calculations see refs. [46, 50].

We use the FishLSS⁸ [27] code to compute the Fisher information for the full-shape 21cm power spectrum $P_{\text{obs}}(\mathbf{k})$ in the basis:

$$(A_s, n_s, h, \omega_c, \omega_b, M_\nu) \cup \{\ln \bar{T}_b, b, b_2, b_s, \alpha_0, \alpha_2, \alpha_4, N_0, N_2, N_4\}_z, \quad (4.2)$$

where M_ν is the sum of the neutrino masses, which we hold fixed unless otherwise stated, and the biases here (b, b_2, b_s) are in the Eulerian frame. We treat all time-dependent nuisance terms (the right-hand set in Eq. 4.2; the z subscript indicates time-dependence) in different redshift bins as separate, independent variables which are individually varied. We direct the reader to ref. [27] for additional details regarding the numerical computation of derivatives and evaluation of the full shape Fisher matrix, as well as information regarding the CMB Fisher matrices used below. Unless otherwise stated we will include a prior on $\Lambda\text{CDM}+M_\nu$ from Planck primary CMB measurements in all of our forecasts.

As discussed in §2.3 we will parameterize the reionization history with set of parameters $\{x_i\}$. Using this parametrization we can trade the free parameters $\{\ln \bar{T}_b\}_z$ for the set $\{x_i\}$ via the chain rule: $\partial_{x_i} P_{\text{obs}}(z_j) = (\partial_{x_i} \ln \bar{T}_b) \partial_{\ln \bar{T}_b} P_{\text{obs}}(z_j)$, since the value of the ionization fraction has little impact on the evolution of large-scale matter fluctuations and primarily impacts P_{obs} through an amplitude change. In our fiducial scenario we use the Weibull form ($\{x_i\} = \{a, b, c\}$), in which case our fiducial full set of free parameters is given by

$$\underbrace{(A_s, n_s, h, \omega_c, \omega_b, a, b, c)}_{\equiv \theta} \cup \{b, b_2, b_s, \alpha_0, \alpha_2, \alpha_4, N_0, N_2, N_4\}_z. \quad (4.3)$$

In this scenario our total number of free parameters is $8+9 \times (\text{number of redshift bins})$. Once we have the Fisher matrix in the basis Eq. (4.3), we then average over the nuisance parameters

⁷<https://github.com/sfschen/velocileptors>

⁸<https://github.com/NoahSailer/FishLSS>

$\{b, b_2, b_s, \alpha_0, \alpha_2, \alpha_4, N_0, N_2, N_4\}_z$ to obtain the nuisance-averaged 8×8 Fisher matrix \mathbf{F} , and use Eq. (2.1) to propagate the errors on $\boldsymbol{\theta}$ to the optical depth: $\sigma_\tau^2 = \partial_{\boldsymbol{\theta}}\tau \cdot \mathbf{F}^{-1} \cdot \partial_{\boldsymbol{\theta}}\tau$.

We base our fiducial reionization history off of the SCORCH [52] simulations, which generalize the radiation-hydrodynamic code Radhydro [53, 54] to include varying radiation escape fractions: $f_{\text{esc}}(z) = f_8[(1+z)/9]^{a_8}$. We choose to adopt the redshift-independent scenario ($a_8 = 0$) as our fiducial model for $x(z)$, which is shown in the left panel of Fig. 3. This scenario predicts $\tau = 0.0639$ when all of the cosmological parameters are fixed at their best-fit Planck values [5], which is within 1.1σ of the Planck measurement $\tau = 0.0561 \pm 0.0071$. In Fig. 3 we also show the best fit Weibull function ($a = 6.62, b = 1.85, c = 1.14$), which agrees with the simulation output to within $|\delta x| \leq 0.02$, resulting in a very small systematic shift in the optical depth: $|\delta\tau| = 0.0005$. Additionally Fig. 3 shows the cumulative contributions to τ from redshifts $< z$ (normalized to the total τ). Note that ≈ 0.5 of the total τ comes from redshifts below 5 where the Universe is fully ionized (and hence has no uncertainty from ignorance of x). Most of the rest comes from $z < 11$.

For the clustering we adopt the evolution of b measured from simulations⁹ in ref. [21] and set our fiducial counterterm $\alpha_0(z)$ from the effective bubble radius evolution $R_{\text{eff}}(z)$ in the same work (Fig. 3). These biases are consistent with those measured by ref. [22], however other theoretical calculations have found different values [55]. The effective bubble radius¹⁰ is defined such that $\delta_{21} \supset b(1 - R_{\text{eff}}^2 k^2/3)\delta$, so that the power spectrum $P_{21} \supset -2b^2 R_{\text{eff}}^2 k^2 P/3 + \dots$, and so we identify $\alpha_0 = -2b^2 R_{\text{eff}}^2/3$. For the higher order biases, b_2 and b_s , we assume vanishing Lagrangian bias corresponding to the coevolution values for Eulerian biases and we further take $N_0 = N_2 = N_4 = \alpha_2 = \alpha_4 = 0$. Since these terms are relevant primarily at small scales and in the absence of receiver noise, the precise values are not important. We have explicitly checked that changing the fiducial value of b_2 from its coevolution value to $b_2(z) = -2$, a value that roughly agrees with that measured by ref. [21] from simulations, impacts our forecasts (for the idealized scenario of §4.2) by less than 7%.

In Fig. 3 we note that the linear bias starts off positive and eventually becomes negative while R_{eff} grows towards the end of reionization. The growth of the bubbles is self-explanatory. Ref. [21] gives a simple explanation for the evolution of the bias assuming that the ionized bubbles trace the locations of the sources in proportion to their ionizing luminosity and that recombinations and light travel delays can be neglected. In this model a large region containing more sources will have more ionization so $\delta_{\text{HII}} = \delta_S + \delta$, where δ_S is the source overdensity field. Writing $\delta = (1-x)\delta_{\text{HI}} + x\delta_{\text{HII}}$ we note that the first term on the rhs is δ_{21} while the last is $x(\delta + \delta_S)$ and thus $\delta_{21} = (1-x)\delta - x\delta_S = (1-x[1+b_S])\delta$. Assuming the sources of reionization are highly clustered we expect b to be negative once reionization gets underway.

4.2 Idealistic survey

We first examine an idealized noiseless survey which covers half the sky over the redshift range $7 < z < 15$, which we will split into 16 redshift bins¹¹ with $\Delta z = 0.5$. To investigate

⁹These values are taken from the left-hand panel of Fig. 7. We fit a Weibull function to the reionization history of ref. [21], finding $(a, b, c) = (6.73, 1.0, 0.92)$, and use this to appropriately translate these curves to our fiducial reionization history (assuming that $b(x)$ and $R_{\text{eff}}^2(x)$ are fixed).

¹⁰The measured values of R_{eff}^2 are poorly constrained at early times [21], so the negative values of R_{eff}^2 may just be a numerical artifact. This has a negligible impact on our constraints since in this regime the hard k_{max}^h cut dominates over the bubble size such that $\alpha_0 k^2 \ll 1$.

¹¹We have checked that doubling the number of redshifts bins (so that $\Delta z = 0.25$) does not significantly impact our forecasts, suggesting that $\Delta z = 0.5$ is sufficient for convergence.

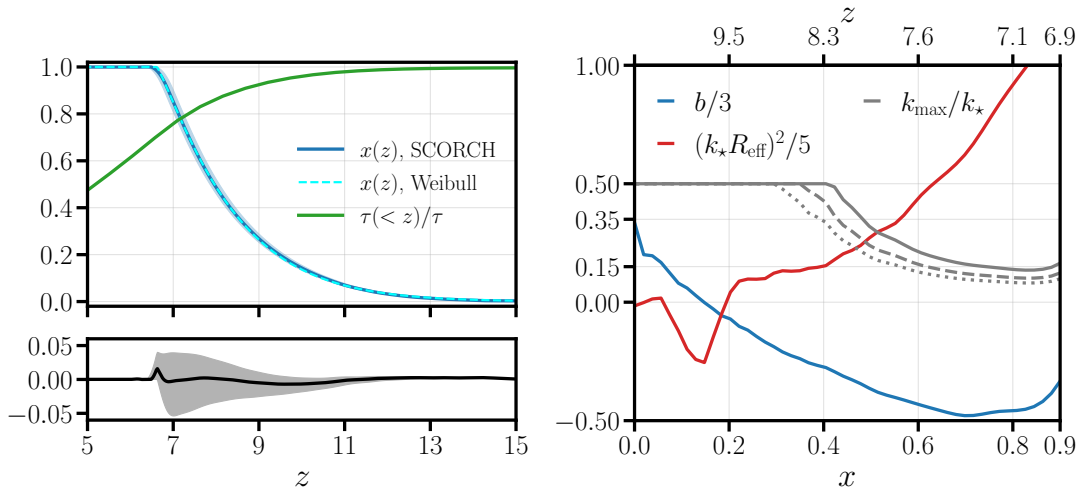


Figure 3. *Left:* In the top panel we show our fiducial reionization history (blue) and the best fit Weibull function (cyan). The bottom panel shows the residual $|\delta x| < 0.02$. This residual results in a systematic error of $|\delta\tau| = 0.0005$. The edges of shaded bands represent the δz by which one can shift the reionization history and change the optical depth by $|\delta\tau| = 0.001$. In green we show the (relative) cumulative contribution to the optical depth for redshifts $< z$. *Right:* The fiducial linear Eulerian bias b , the effective bubble size R_{eff} and a representative evolution of k_{max} (see Eq. (4.4)), where $k_{\star} = 1 h \text{Mpc}^{-1}$. The grey solid, dashed and dotted lines correspond to $X_{\alpha} = 3, 4, 5$ respectively. Such limits ensure that for $k < k_{\text{max}}$ the corrections are perturbative and the higher-order counterterms are negligible.

the sensitivity of these measurements in the $\mathbf{k} - z$ parameter space, we will first consider redshift-independent scale cuts, with our fiducial cuts being $k_{\parallel}^{\text{min}} = 0.05 h \text{Mpc}^{-1}$, $k_{\text{max}} = 0.5 h \text{Mpc}^{-1}$ and $\mu_{\text{min}} = 0.7$. A more realistic treatment which accounts for the redshift dependence of the foreground wedge and k_{max} will be considered in §4.3. For this idealized scenario, and using the best-fit Weibull function to parameterize the reionization history, we find a fiducial error on the optical depth $\sigma_{\tau} = 0.00016$ with a Planck prior on ΛCDM , and 0.00026 without (21cm clustering only). This illustrates that if reionization proceeds smoothly in redshift and can be parameterized by a few parameters then there is more than enough signal in the clustering to infer τ with very high accuracy.

Exploring this further, Fig. 4 shows how the errors depend upon the range of scales and redshifts probed. As Fig. 4 makes clear, the constraints are more sensitive to an experiment’s ability to resolve large scales (i.e. low k modes) and obtain sufficient dynamic range in μ than to the range of redshifts it probes or its ability to probe very small scales. Among other things this suggests that foreground subtraction and instrument calibration will matter as much for measuring τ as raw instrumental sensitivity. Quantitatively we see that the constraints on τ degrade significantly if the experiment cannot measure $k < 0.2 h \text{Mpc}^{-1}$, but it is only necessary to measure well modes up to $k \simeq 0.2 h \text{Mpc}^{-1}$ to get almost all of the constraining power in the data. This is because the smaller-scale modes mostly serve to constrain astrophysical parameters. All of these modes are linear, from the perspective of dynamics.

Much of the integral responsible for τ comes from ‘intermediate’ redshifts, so the survey does not need to probe very high or low redshifts in order to constrain τ very well. This

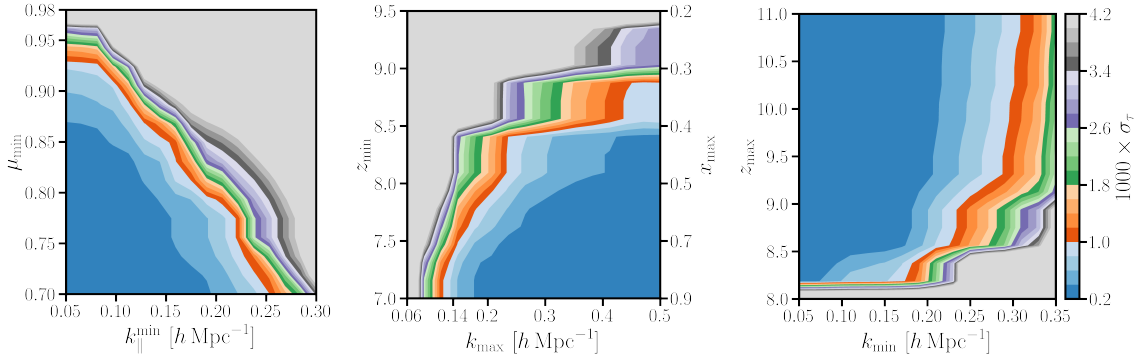


Figure 4. Errors on the optical depth (times a thousand) for a noiseless survey covering $7 < z < 15$ over half the sky. The survey has been split into 16 redshift bins with $\Delta z = 0.5$. The fiducial scale cuts are $k_{\max} = 0.5 \text{ h Mpc}^{-1}$, $k_{\parallel}^{\min} = 0.05 \text{ h Mpc}^{-1}$ and $\mu_{\min} = 0.7$, and we include a prior on ΛCDM from Planck primary CMB measurements.

is also illustrated in Fig. 4 where we see that the errors on τ remain small for $z_{\min} < 8$ and $z_{\max} > 9.5$ (in our fiducial reionization history) corresponding to coverage of ionization fractions $0.2 \lesssim x(z) \lesssim 0.5$. Only once the survey is unable to probe those epochs does the inferred error on τ rise dramatically. Varying the Weibull parameter b , which affects the duration of reionization, in our fiducial model within upper limits on the duration of reionization from the CMB ($b < 2.8$) changes our forecasts for σ_{τ} by $< 25\%$. We expect that $T_S \gg T_{\text{cmb}}$ for $z \lesssim 10$ [2–4], justifying our neglect of spin temperature fluctuations in our fiducial model. If reionization is particularly extended or has a non-trivial high z component this assumption would need to be revisited. Such scenarios are disfavored, but if the spin temperature can fluctuate while x is significantly different from zero then the connection between 21cm clustering and τ would be much more complex.

Finally, to investigate how sensitive our constraints are to the number of free variables in our parameterization of the reionization history, we consider swapping the Weibull function for a Lagrangian interpolating function (see §2.3) with N free parameters, which we choose to be the ionization fractions $\{x(z_i)\}$ at redshifts $z_i = 7, 7 + 6/N, \dots, 13$. For $N = 3, 5, 7, 9$ we find $1000 \times \sigma_{\tau} = 0.20, 0.25, 1.95, 4.47$ for our set of fiducial scale and temporal cuts. The scaling of the error with N gets significantly worse with more realistic scale and temporal cuts: for example, after enforcing $z_{\max} = 10$ and adding a redshift-dependent foreground wedge ($N_w = 3$, see §4.3) the error increases from 0.4 to 6.1 for $N = 3, 5$. It is worth noting that not all of the loss of constraining power with increasing N is necessarily physical: astrophysical processes contributing to reionization typically evolve on long (fraction of Hubble) time scales, leading to constraints on the variation of $x(z)$ between interpolating points. Introducing physically motivated smoothness priors to parametrically fit e.g. the expansion history have recently been studied in ref. [56], and we leave it to future work to develop these techniques in the context of reionization.

4.3 More realistic scenarios

Having determined what range of scales and redshifts are most important for constraining τ we now consider more realistic scenarios in which we add instrument noise, foregrounds and redshift evolution. Our purpose is not to model any particular 21cm survey in detail, but rather to explore the impact that various observational choices have on the inferred

constraints on τ . The current generation of 21cm interferometers does not have sufficient sensitivity to improve upon existing measurements of τ , so we shall consider hypothetical next-generation instruments with integration times on the order of years, consistent with proposed post-reionization 21cm surveys [57], as opposed to the more standard 1000 hours¹².

Our fiducial instrument will be a hexagonally packed array of 10^3 , non-tracking, 14 m dishes. This is a scaled up version of the Hydrogen Epoch of Reionization Array (HERA; [58]). We compute the baseline distribution using 21cmSense¹³ [59, 60]. The thermal noise power is based on Appendix D of ref. [61]. At high redshifts we are sky noise dominated and since most of our constraining power on τ comes from intermediate k the baseline distribution is of secondary importance and we are mostly sensitive to the total collecting area (and the integration time). HERA frequencies range from 50 to 250 MHz, which translates to $4.8 < z < 27.5$ and we shall assume a future array would cover a comparable frequency range. As we have seen, such a wide redshift range is more than sufficient to provide tight constraints on τ .

The range of wavenumbers transverse to the line of sight that can be probed by an interferometer depends upon the distribution of baseline separations [62]. Since the most constraining modes for our purposes are the low k modes we are primarily interested in the number of short baselines. If we take a minimum dish separation of 14 m for our fiducial cosmology $k_{\perp}^{\min} = 6 \times 10^{-3} h \text{ Mpc}^{-1}$ at $z = 10$ and $4 \times 10^{-3} h \text{ Mpc}^{-1}$ at $z = 15$. The baselines that probe $k_{\perp} \sim 0.1 h \text{ Mpc}^{-1}$ are 240 m at $z = 10$ and 380 m at $z = 15$. All of these baselines would be well probed by next-generation experiments.

The minimum wavenumber that can be probed in the line of sight direction depends upon the ability of the experiment to subtract the foregrounds that are many orders of magnitude brighter than the signal, but spectrally smooth. Methods for foreground subtraction are rapidly improving, so we shall simply explore how our constraint on τ depends upon k_{\parallel}^{\min} , which inherently limits the accessible μ -range. For our fiducial scenario we will take $k_{\parallel}^{\min} = 0.1 h \text{ Mpc}^{-1}$. The maximum observationally accessible wavenumber is normally set by thermal noise and the lack of many widely separated baselines. However since we are attempting to focus on perturbative scales this will not be the only limitation of interest to us. Rather our k_{\max} will (also) be set by the scale at which our bias expansion becomes inapplicable, which is mostly set by the evolution of the “effective” bubble size. We shall take k_{\max} to be the smaller of $k_{\max}^h = 0.5 h \text{ Mpc}^{-1}$ and the scale at which our lowest-order counterterm is X_{α}^{-2} of the linear power spectrum:

$$k_{\max}(z) = \min \left[\frac{1}{X_{\alpha} \sqrt{|\alpha_0(z)|}}, k_{\max}^h \right] \quad \text{where} \quad \alpha_0(z) = -\frac{2}{3} b^2(z) R_{\text{eff}}^2(z) \quad (4.4)$$

and consider $X_{\alpha} = 3$ and 4. In principle the validity of the bias expansion in real space is set strictly by the nonlocality scale R_{eff} but we have adopted $1/\sqrt{\alpha_0}$ to serve as a catch-all, including both the validity of the bias expansion and the significance of the linear RSD signal from which most of our constraint derives; we have checked that this choice does not significantly impact our σ_{τ} forecast.

¹²Reducing the integration time for our fiducial SKA-like survey (see the SKA-like column in Table 1) from 5 years to 1000 hours increases the 1σ error on τ from 0.0009 to 0.0067, which is a marginal improvement over Planck alone ($\sigma_{\tau} = 0.0071$). For the 1000 dish scenario, reducing the integration time to 1000 hours results in no improvement over Planck in all of our scenarios.

¹³<https://github.com/steven-murray/21cmSense>

Finally, contamination from systematic errors in the inherently chromatic response of an interferometer, interacting with astrophysical sources that are many orders of magnitude brighter than the signals of interest, leads to contamination within a “foreground wedge” [63, 64]. A particularly lucid and pedagogical introduction to the physical origin of the wedge can be found in ref. [65]. Given the rapidly evolving state of 21cm analysis, we shall take a simple model of the wedge and explore how variations in performance will affect our inference. In particular we follow ref. [57] in modeling the foreground wedge as allowing only modes with

$$k_{\parallel} > \max \left[k_{\perp} \frac{\chi(z)H(z)}{c(1+z)} \sin(\theta_w(z)), k_{\parallel}^{\min} \right] \quad \text{where} \quad \theta_w(z) = N_w \frac{1.22}{2\sqrt{0.7}} \frac{\lambda_{\text{obs}}}{D_{\text{phys}}} \quad (4.5)$$

where $\lambda_{\text{obs}} = (1+z) 21 \text{ cm}$ is the observed wavelength, D_{phys} is the physical diameter of the dish, $k_{\parallel} = k\mu$ and $k_{\perp} = k\sqrt{1-\mu^2}$. For our fiducial scenario we will take $N_w = 3$.

Fig. 2 shows the foreground wedge as a function of redshift for $N_w = 3$ and $D_{\text{phys}} = 14 \text{ m}$. The minimum μ uncontaminated by foregrounds rises from $\simeq 0.5$ at $z = 6$ to $\simeq 0.9$ at $z = 15$. For comparison the minimum μ kept in the current analysis of HERA data is $\simeq 0.98$ [22, 63], which would render measurement of RSD ineffective. By contrast SKA predicts an observable window of $\mu > 0.67$. Clearly considerable effort remains to be done on instrument calibration and foreground removal to enable our forecasts.

Scenario	1000 dishes								SKA-like	Planck	PSD
N_w	3	3	3	3	3	3	3	6	3	–	–
X_{α}	3	3	3	3	3	3	4	3	3	–	–
k_{max}^h [$h \text{ Mpc}^{-1}$]	0.5	0.5	0.5	0.5	0.5	0.3	0.5	0.5	0.5	–	–
k_{\parallel}^{\min} [$h \text{ Mpc}^{-1}$]	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	–	–
z_{max}	12	12	12	10	12	12	12	12	10	–	–
z_{min}	7	7	9	7	7	7	7	7	7	–	–
t_{int} [years]	10	5	5	5	5	5	5	5	5	–	–
$1000 \times \sigma_{\tau}$	1.9	3.0	7.0	3.8	6.9	3.8	3.9	6.3	0.9	7.1	5.8
$\sigma_{M_{\nu}}$ [meV]	30	33	39	33	39	33	34	38	26	83	40

Table 1. Forecasts for the uncertainty on the optical depth and the sum of the neutrino masses for different experimental configurations. All forecasts on the optical depth include a prior from Planck and assume $f_{\text{sky}} = 0.5$. The neutrino mass constraints all include Planck, SO and the emission line galaxy sample of DESI (PSD) following ref. [27]. The baseline constraints for these are shown in the last two columns. Columns 2-9 show the impact of varying each parameter one at a time from our fiducial to a more pessimistic scenario. As a more futuristic example we include an SKA-like forecast, in which the collecting area is increased to 1 km^2 , which increases the sensitivity significantly over our fiducial scenario.

Our main results are summarized in Table 1, which presents forecasts for the uncertainty on the optical depth and the sum of the neutrino masses for different experimental configurations. The 21cm-based forecasts are compared to the existing measurement precision of Planck [1, 5] and to a forecast for how well τ and M_{ν} could be constrained by a combination of future CMB and galaxy redshift surveys (Planck, Simons Observatory and the emission line galaxy sample of DESI, indicated as PSD [27]). Columns 2-9 show the impact of varying experimental configurations, one element at a time, at fixed array size (10^3 dishes) while the column labeled SKA-like shows the improvement afforded by a much larger array (with 1 km^2

collecting area), which reduces the thermal noise by roughly a factor of 5. We see that z_{\min} , k_{\parallel}^{\min} and N_w have the largest impact on σ_{τ} , as expected from our earlier discussion. Once $z_{\min} > 9$ we are no longer measuring the range of redshifts (or x) that contribute the most to τ . For $k_{\parallel}^{\min} > 0.1 h \text{ Mpc}^{-1}$ we become sensitive primarily to the astrophysics controlling bubble sizes which our method explicitly marginalizes over. When $k_{\min} \simeq 0.2 h \text{ Mpc}^{-1}$ our $k_{\max} \leq k_{\min}$ for $x > 0.5$ and we are unable to probe contributions to τ from $z < 8$. For larger N_w the range in μ is too restricted to be able to separate the modes along the line of sight from those transverse to it, and thus we cannot break the degeneracy with bias. By contrast we are relatively insensitive to the smallest scales we model – changes to X_{α} and k_{\max}^h change σ_{τ} only modestly from our baseline configuration. Column 2 and the SKA-like column shows the constraints improve non-negligibly for lower thermal noise levels, indicating that our fiducial array is far from sample-variance limited. The constraints from the 10 yr survey with our fiducial array would have the same precision as the forecasts for σ_{τ} from the proposed LiteBIRD satellite mission [66]. In contrast to the results for a noiseless survey (§4.2, or SKA), here we find that a noisy survey is moderately sensitive to the maximum redshift: decreasing z_{\max} from 12 to 10 to avoid the impact of spin-temperature fluctuations can increase the uncertainty on the optical depth by $\sim 25\%$.

5 Discussion and conclusions

The optical depth to reionization τ is one of the key parameters in the 6-parameter Λ CDM model of cosmology [5]. Holding the amplitude of CMB fluctuation fixed, τ controls the amplitude of fluctuations in the Universe. Currently the tightest constraints on this parameter come from observations of CMB polarization, but in the future 21cm experiments have the potential to dramatically improve upon our constraints. We discuss an approach to constraining τ built around ideas very familiar in large-scale structure studies but less often used in reionization. We emphasize the role of large scale modes, redshift-space distortions and a symmetries-based bias model that allows a principled but efficient marginalization over astrophysical parameters. The split encoded in this formalism means we can focus on the cosmological information contained in large scale modes while efficiently removing sensitivity to small-scale physics (and since our expansion is based entirely on symmetry, it makes minimal assumptions about that physics). Conversely, information about the sources and morphology of reionization is encapsulated in the small-scale modes and bias coefficients, with the large-scale modes being set largely by the timing of reionization. This approach to analyzing 21cm interferometer data has implications for the range of scales and redshifts that contain most of the cosmological signal, and hence on the most impactful observational settings.

Given sufficient dynamic range in scale, and the ability to measure the anisotropy in the clustering that is imprinted by peculiar velocities, future 21cm experiments should be able to place constraints on τ that are much tighter than constraints from existing CMB experiments or even future galaxy redshift surveys and competitive with proposed, future space-based CMB anisotropy missions. In order to make these inferences robust to uncertainties in astrophysical models, a premium should be placed on measuring large scales and on cleaning the foreground wedge as much as possible. By contrast very high angular resolution is less helpful. We find that much of the constraining power to τ is lost once k_{\parallel}^{\min} rises to $\approx 0.2 h \text{ Mpc}^{-1}$ or if the foreground wedge extends past $\approx 3\times$ the primary beam. The experiment

must also measure well the redshifts for which reionization is most active, i.e. those with $0.2 < x < 0.8$.

We have performed our forecasts within the context of Lagrangian perturbation theory, since this has a natural connection to the excursion set formalism and the theory of peaks and easily handles redshift-space distortions. At low redshifts the Lagrangian bias expansion has been shown to perform better for halos and galaxies than the equivalent Eulerian formalism [67], however which is preferable for reionization is currently unknown. The Lagrangian and Eulerian approaches to perturbation theory both predict that the dynamics on the scales and redshifts relevant for determining τ are very close to linear, with small corrections that can be very well described by either formalism. The primary limitation is not the non-linear scale but the scale of non-locality (due to ionized bubbles) in the bias expansion. We have treated this as a fundamental limitation in our forecasts and treated the bias parameters as completely free; however, future information about the process of reionization from a combination of simulations and observations could allow us to place priors on the bias coefficients. Such priors would allow information to be extracted from smaller scales than we have considered. In this regard we encourage simulators (and observers) to provide information on the anisotropy of the power spectrum when publishing their models.

Our forecasts assume that reionization proceeds smoothly and in a manner that can be described as “late and fast”. This is consistent with the current observations, but our understanding of the process of reionization is highly incomplete. Since the low redshift Universe is known to be highly ionized [68], the main cause for concern would be multiple or partial ionizations at high z . Such episodes would leave a clear signal in P_{21} that would be at odds with our default scenario, so would not lead to inference of a highly precise but inaccurate τ unless $x(z)$ and the spin temperature evolution somehow conspired. For this reason we have not considered such scenarios.

Our forecasts suggest that a scaled up version of HERA has the potential to measure the optical depth with statistical uncertainty $\sigma_\tau = 0.002 - 0.004$, while a SKA-like survey could achieve $\sigma_\tau = 0.0009$ (assuming a 5 – 10 year integration time, we find no improvement over Planck with only ~ 1000 hours). We note that this improvement in sensitivity is more modest than models which claim to accurately predict and derive τ -constraints from small-scale 21cm fluctuations (e.g. refs. [12, 14, 15], which use 21cmFAST as the theory model). However, these more ambitious models run an increased risk of returning biased results should their predictions differ significantly from the true small-scale 21cm signal. By contrast, our approach of modeling the 21cm intensity field using a symmetries-based bias expansion has been chosen to mitigate astrophysical uncertainties, however, several notable systematics could remain. First, any error in the best-fit Weibull function to the true reionization history propagates as a systematic error to the optical depth measurement. This systematic error can be reduced by freeing up the functional form of the $x(z)$ parametrization, however, doing so (§4.2) comes with a significant noise penalty when a realistic foreground wedge and z_{\max} are accounted for. Second, we have neglected the contributions from doubly reionized helium, which is expected to contribute $\tau_{\text{He}} \sim 0.0008$ to the optical depth. Even for an SKA-like survey the second ionization of helium would need to be uncertain by $\mathcal{O}(1)$ to give significant errors. Additionally, we have neglected spin temperature fluctuations and uncertainties in the primordial helium abundance. We have shown in §4 how our constraints vary with z_{\max} , which can be lowered to remove any spin-temperature induced systematics, while within our formalism it would be straightforward to marginalize over any uncertainty in Y_p ¹⁴.

¹⁴Given a prior $\sigma_{Y_p} = 0.2437 \pm 0.008$ from Planck [5] and direct measurements of helium abundance,

Finally, in our discussion of perturbation theory we have made the implicit assumption that all matter (dark matter, baryons and neutrinos) in the Universe can be treated as a single fluid. In the case of neutrinos it is known [69, 70] that galaxies primarily trace the cold dark matter and baryons; since the bias coefficients for the 21cm signal primarily reflect those of neutral hydrogen and ionizing sources the same approximation likely holds in this context as well. Residual differences post-recombination in the baryon (b) and dark matter (c) fields can lead to additional bias contributions proportional to relative densities and velocities [71–74]. These contributions are either stationary or decay with time, and thus tend to be extremely subdominant at low redshifts, but could conceivably be a non-negligible effect at the EoR and, at k beyond the acoustic scale, degenerate with the linear bias. However, we note that linear theory contributions to the observed clustering due to the relative velocity v_{bc} should be significantly below 1% of the total matter velocity contributions even at high redshifts and these terms should contribute negligibly to redshift-space distortions. Since our reionization measurement depends mostly on the observed anisotropy it should be quite immune to these additional biases.

Acknowledgments

We would like to thank Avery Meiksin for helpful discussions on Helium reionization, and Joshua Dillon for helpful discussion on thermal noise calculation. S.C. thanks Nickolas Kokron for useful discussions on the form of the 21cm signal. N.S. is supported by the Physics Division of Lawrence Berkeley National Laboratory. S.C. and M.W. are supported by the DOE. This research has made use of NASA’s Astrophysics Data System and the arXiv preprint server. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

Note added. While this paper was being finalized, ref. [22] appeared which takes a highly related approach to modeling 21cm perturbations using a large-scale-structure-inspired bias modeling. Our Lagrangian formalism can be exactly mapped into their Eulerian one [46] so the approaches are highly synoptic.

References

- [1] Planck Collaboration, N. Aghanim, Y. Akrami, et al., *Planck 2018 results. VI. Cosmological parameters*, *arXiv e-prints* (July, 2018) arXiv:1807.06209, [arXiv:1807.06209].
- [2] G. D. Becker, J. S. Bolton, and A. Lidz, *Reionisation and High-Redshift Galaxies: The View from Quasar Absorption Lines*, *Publ. Astron. Soc. Australia* **32** (Dec., 2015) e045, [arXiv:1510.03368].
- [3] A. Mesinger, *Understanding the Epoch of Cosmic Reionization*, *Astrophysics and Space Science Library* **423** (Jan., 2016).
- [4] M. McQuinn, *The Evolution of the Intergalactic Medium*, *Ann. Rev. Astron. & Astrophys.* **54** (Sept., 2016) 313–362, [arXiv:1512.00086].
- [5] Planck Collaboration, Y. Akrami, F. Arroja, et al., *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, *arXiv e-prints* (July, 2018) arXiv:1807.06205, [arXiv:1807.06205].

uncertainties in Y_p would result in $\sigma_\tau \sim 0.0003$. We therefore do not expect varying Y_p to significantly impact our results when realistic noise is included.

- [6] R. A. Sunyaev and Y. B. Zeldovich, *The velocity of clusters of galaxies relative to the microwave background - The possibility of its measurement.*, *Mon. Not. R. Astron. Soc.* **190** (Feb., 1980) 413–420.
- [7] C. Dvorkin and K. M. Smith, *Reconstructing Patchy Reionization from the Cosmic Microwave Background*, *Phys. Rev. D* **79** (2009) 043003, [[arXiv:0812.1566](#)].
- [8] S. Ferraro and K. M. Smith, *Characterizing the epoch of reionization with the small-scale CMB: Constraints on the optical depth and duration*, *Phys. Rev. D* **98** (2018), no. 12 123519, [[arXiv:1803.07036](#)].
- [9] M. A. Alvarez, S. Ferraro, J. C. Hill, et al., *Mitigating the optical depth degeneracy using the kinematic Sunyaev-Zel'dovich effect with CMB-S4*, *Phys. Rev. D* **103** (2021), no. 6 063518, [[arXiv:2006.06594](#)].
- [10] R. Barkana and A. Loeb, *A Method for Separating the Physics from the Astrophysics of High-Redshift 21 Centimeter Fluctuations*, *Astrophys. J. Lett.* **624** (May, 2005) L65–L68, [[astro-ph/0409572](#)].
- [11] M. McQuinn, O. Zahn, M. Zaldarriaga, et al., *Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization*, *Astrophys. J.* **653** (Dec., 2006) 815–834, [[astro-ph/0512263](#)].
- [12] A. Liu, J. R. Pritchard, R. Allison, et al., *Eliminating the optical depth nuisance from the CMB with 21 cm cosmology*, *Phys. Rev. D* **93** (2016), no. 4 043013, [[arXiv:1509.08463](#)].
- [13] T. S. Billings, P. L. Plante, and J. E. Aguirre, *Extracting the optical depth to reionization τ from 21 cm data using machine learning techniques*, 2021.
- [14] B. Greig, C. M. Trott, N. Barry, et al., *Exploring reionization and high- z galaxy observables with recent multiredshift MWA upper limits on the 21-cm signal*, *Mon. Not. Roy. Astron. Soc.* **500** (2020), no. 4 5322–5335, [[arXiv:2008.02639](#)].
- [15] B. Greig et al., *Interpreting LOFAR 21-cm signal upper limits at $z \approx 9.1$ in the context of high- z galaxy and reionization observations*, *Mon. Not. Roy. Astron. Soc.* **501** (2021), no. 1 1–13, [[arXiv:2006.03203](#)].
- [16] A. H. Pawlik, A. Rahmati, J. Schaye, et al., *The Aurora radiation-hydrodynamical simulations of reionization: calibration and first results*, *Mon. Not. R. Astron. Soc.* **466** (Apr., 2017) 960–973, [[arXiv:1603.00034](#)].
- [17] R. Kannan, E. Garaldi, A. Smith, et al., *Introducing the THESAN project: radiation-magnetohydrodynamic simulations of the epoch of reionization*, *Mon. Not. R. Astron. Soc.* **511** (Apr., 2022) 4005–4030, [[arXiv:2110.00584](#)].
- [18] J. B. Muñoz, Y. Qin, A. Mesinger, et al., *The Impact of the First Galaxies on Cosmic Dawn and Reionization*, *arXiv e-prints* (Oct., 2021) arXiv:2110.13919, [[arXiv:2110.13919](#)].
- [19] H. Trac, N. Chen, I. Holst, et al., *AMBER: A Semi-numerical Abundance Matching Box for the Epoch of Reionization*, *Astrophys. J.* **927** (2022), no. 2 186, [[arXiv:2109.10375](#)].
- [20] K. Hoffmann, Y. Mao, J. Xu, et al., *Signatures of cosmic reionization on the 21-cm two- and three-point correlation function I: quadratic bias modelling*, *Mon. Not. R. Astron. Soc.* **487** (Aug., 2019) 3050–3068, [[arXiv:1802.02578](#)].
- [21] M. McQuinn and A. D’Aloisio, *The observable 21cm signal from reionization may be perturbative*, *JCAP* **10** (2018) 016, [[arXiv:1806.08372](#)].
- [22] W. Qin, K. Schutz, A. Smith, et al., *An Effective Bias Expansion for 21 cm Cosmology in Redshift Space*, [[arXiv:2205.06270](#)].
- [23] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming*

- era of precision LSS, *Journal of Cosmology and Astro-Particle Physics* **8** (Aug., 2009) 020, [[arXiv:0902.0991](#)].
- [24] V. Desjacques, D. Jeong, and F. Schmidt, *Large-Scale Galaxy Bias*, *ArXiv e-prints* (Nov., 2016) [[arXiv:1611.09787](#)].
- [25] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. R. Astron. Soc.* **227** (July, 1987) 1–21.
- [26] E. Castorina and M. White, *Measuring the growth of structure with intensity mapping surveys*, *Journal of Cosmology and Astro-Particle Physics* **2019** (June, 2019) 025, [[arXiv:1902.07147](#)].
- [27] N. Sailer, E. Castorina, S. Ferraro, and M. White, *Cosmology at high redshift — a probe of fundamental physics*, *JCAP* **12** (2021), no. 12 049, [[arXiv:2106.09713](#)].
- [28] J. C. Pober, *The impact of foregrounds on redshift space distortion measurements with the highly redshifted 21-cm line*, *Mon. Not. R. Astron. Soc.* **447** (Feb., 2015) 1705–1712, [[arXiv:1411.2050](#)].
- [29] P. Madau, A. Meiksin, and M. J. Rees, *21 Centimeter Tomography of the Intergalactic Medium at High Redshift*, *Astrophys. J.* **475** (Feb., 1997) 429–444, [[astro-ph/9608010](#)].
- [30] S. Clesse, L. Lopez-Honorez, C. Ringeval, et al., *Background reionization history from omniscopes*, *Phys. Rev. D* **86** (2012) 123506, [[arXiv:1208.4277](#)].
- [31] Y. Mao, P. R. Shapiro, G. Mellema, et al., *Redshift-space distortion of the 21-cm background from the epoch of reionization - I. Methodology re-examined*, *Mon. Not. R. Astron. Soc.* **422** (May, 2012) 926–954, [[arXiv:1104.2094](#)].
- [32] S. Furlanetto, S. P. Oh, and F. Briggs, *Cosmology at Low Frequencies: The 21 cm Transition and the High-Redshift Universe*, *Phys. Rept.* **433** (2006) 181–301, [[astro-ph/0608032](#)].
- [33] S.-F. Chen, Z. Vlah, and M. White, *The Ly α forest flux correlation function: a perturbation theory perspective*, *Journal of Cosmology and Astro-Particle Physics* **2021** (May, 2021) 053, [[arXiv:2103.13498](#)].
- [34] J. E. Gunn and B. A. Peterson, *On the Density of Neutral Hydrogen in Intergalactic Space.*, *Astrophys. J.* **142** (Nov., 1965) 1633–1636.
- [35] Y. Zhu et al., *Long Dark Gaps in the Ly β Forest at $z < 6$: Evidence of Ultra Late Reionization from XQR-30 Spectra*, [arXiv:2205.04569](#).
- [36] H. Trac, *Parametrizing the Reionization History with the Redshift Midpoint, Duration, and Asymmetry*, *Astrophys. J. Lett.* **858** (2018), no. 2 L11, [[arXiv:1804.00672](#)].
- [37] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, *Large-scale structure of the Universe and cosmological perturbation theory*, *Phys. Rep.* **367** (Sept., 2002) 1–248, [[astro-ph/0112551](#)].
- [38] S. Dodelson and F. Schmidt, *Modern Cosmology*. Elsevier Science, 2020.
- [39] Z. Vlah, M. White, and A. Aviles, *A Lagrangian effective field theory*, *Journal of Cosmology and Astro-Particle Physics* **9** (Sept., 2015) 014, [[arXiv:1506.05264](#)].
- [40] R. A. Porto, L. Senatore, and M. Zaldarriaga, *The Lagrangian-space Effective Field Theory of large scale structures*, *Journal of Cosmology and Astro-Particle Physics* **5** (May, 2014) 022, [[arXiv:1311.2168](#)].
- [41] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, *Cosmological non-linearities as an effective fluid*, *Journal of Cosmology and Astro-Particle Physics* **7** (July, 2012) 51, [[arXiv:1004.2488](#)].
- [42] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, *The effective field theory of cosmological large scale structures*, *Journal of High Energy Physics* **9** (Sept., 2012) 82, [[arXiv:1206.2926](#)].

- [43] L. Senatore, *Bias in the effective field theory of large scale structures*, *Journal of Cosmology and Astro-Particle Physics* **11** (Nov., 2015) 007, [[arXiv:1406.7843](#)].
- [44] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, *Renormalized halo bias*, *Journal of Cosmology and Astro-Particle Physics* **8** (Aug., 2014) 056, [[arXiv:1402.5916](#)].
- [45] Z. Vlah, E. Castorina, and M. White, *The Gaussian streaming model and convolution Lagrangian effective field theory*, *Journal of Cosmology and Astro-Particle Physics* **12** (Dec., 2016) 007, [[arXiv:1609.02908](#)].
- [46] S.-F. Chen, Z. Vlah, and M. White, *Consistent Modeling of Velocity Statistics and Redshift-Space Distortions in One-Loop Perturbation Theory*, *JCAP* **07** (2020) 062, [[arXiv:2005.00523](#)].
- [47] T. Matsubara, *Nonlinear perturbation theory with halo bias and redshift-space distortions via the Lagrangian picture*, *Phys. Rev. D* **78** (Oct., 2008) 083519, [[arXiv:0807.1733](#)].
- [48] T. Matsubara, *Resumming cosmological perturbations via the Lagrangian picture: One-loop results in real space and in redshift space*, *Phys. Rev. D* **77** (Mar., 2008) 063530, [[arXiv:0711.2521](#)].
- [49] Z. Vlah and M. White, *Exploring redshift-space distortions in large-scale structure*, *Journal of Cosmology and Astro-Particle Physics* **2019** (Mar, 2019) 007, [[arXiv:1812.02775](#)].
- [50] S.-F. Chen, Z. Vlah, E. Castorina, and M. White, *Redshift-Space Distortions in Lagrangian Perturbation Theory*, *arXiv e-prints* (Dec., 2020) arXiv:2012.04636, [[arXiv:2012.04636](#)].
- [51] D. Blas, J. Lesgourgues, and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes*, *Journal of Cosmology and Astro-Particle Physics* **7** (July, 2011) 034, [[arXiv:1104.2933](#)].
- [52] A. Doussot, H. Trac, and R. Cen, *SCORCH. II. Radiation-hydrodynamic Simulations of Reionization with Varying Radiation Escape Fractions*, *Astrophys. J.* **870** (2019), no. 1 18, [[arXiv:1712.04464](#)].
- [53] H. Trac and U.-L. Pen, *Out - of - core hydrodynamic simulations for cosmological applications*, *New Astron.* **11** (2006) 273–286, [[astro-ph/0402444](#)].
- [54] H. Trac and R. Cen, *Radiative transfer simulations of cosmic reionization. 1. Methodology and initial results*, *Astrophys. J.* **671** (2007) 1, [[astro-ph/0612406](#)].
- [55] W. Xu, Y. Xu, B. Yue, et al., *The HI bias during the Epoch of Reionization*, *Mon. Not. Roy. Astron. Soc.* **490** (2019), no. 4 5739–5748, [[arXiv:1907.05098](#)].
- [56] M. Raveri, L. Pogosian, K. Koyama, et al., *A joint reconstruction of dark energy and modified growth evolution*, *arXiv e-prints* (July, 2021) arXiv:2107.12990, [[arXiv:2107.12990](#)].
- [57] **Cosmic Visions 21 cm** Collaboration, R. Ansari et al., *Inflation and Early Dark Energy with a Stage II Hydrogen Intensity Mapping experiment*, [[arXiv:1810.09572](#)].
- [58] D. R. DeBoer et al., *Hydrogen Epoch of Reionization Array (HERA)*, *Publ. Astron. Soc. Pac.* **129** (2017), no. 974 045001, [[arXiv:1606.07473](#)].
- [59] J. C. Pober, A. R. Parsons, D. R. DeBoer, et al., *The Baryon Acoustic Oscillation Broadband and Broad-beam Array: Design Overview and Sensitivity Forecasts*, *AJ* **145** (Mar., 2013) 65, [[arXiv:1210.2413](#)].
- [60] J. C. Pober et al., *What Next-Generation 21 cm Power Spectrum Measurements Can Teach Us About the Epoch of Reionization*, *Astrophys. J.* **782** (2014) 66, [[arXiv:1310.7031](#)].
- [61] A. Slosar, Z. Ahmed, D. Alonso, et al., *Packed Ultra-wideband Mapping Array (PUMA): A Radio Telescope for Cosmology and Transients*, in *Bull. Am. Astron. Soc.*, vol. 51, p. 53, Sep, 2019. [[arXiv:1907.12559](#)].

- [62] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy, 3rd Edition*. 2017.
- [63] Z. Abdurashidova, J. E. Aguirre, P. Alexander, et al., *First Results from HERA Phase I: Upper Limits on the Epoch of Reionization 21 cm Power Spectrum*, *Astrophys. J.* **925** (Feb., 2022) 221, [[arXiv:2108.02263](#)].
- [64] H. Gan, L. V. E Koopmans, F. G. Mertens, et al., *Statistical analysis of the causes of excess variance in the 21 cm signal power spectra obtained with the Low-Frequency Array*, *arXiv e-prints* (Mar., 2022) arXiv:2203.02345, [[arXiv:2203.02345](#)].
- [65] M. F. Morales, B. Hazelton, I. Sullivan, and A. Beardsley, *Four Fundamental Foreground Power Spectrum Shapes for 21 cm Cosmology Observations*, *Astrophys. J.* **752** (June, 2012) 137, [[arXiv:1202.3830](#)].
- [66] H. Sugai, P. A. R. Ade, Y. Akiba, et al., *Updated design of the cmb polarization experiment satellite litebird*, *Journal of Low Temperature Physics* **199** (Jan, 2020) 1107–1117.
- [67] C. Modi, M. White, A. Slosar, and E. Castorina, *Reconstructing large-scale structure with neutral hydrogen surveys*, *Journal of Cosmology and Astro-Particle Physics* **2019** (Nov., 2019) 023, [[arXiv:1907.02330](#)].
- [68] X. Fan, M. A. Strauss, R. H. Becker, et al., *Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \sim 6$ Quasars. II. A Sample of 19 Quasars*, *AJ* **132** (July, 2006) 117–136, [[astro-ph/0512082](#)].
- [69] E. Castorina, C. Carbone, J. Bel, et al., *DEMNUni: The clustering of large-scale structures in the presence of massive neutrinos*, *JCAP* **07** (2015) 043, [[arXiv:1505.07148](#)].
- [70] A. E. Bayer, A. Banerjee, and U. Seljak, *Beware of Fake ν s: The Effect of Massive Neutrinos on the Non-Linear Evolution of Cosmic Structure*, *arXiv e-prints* (Aug., 2021) arXiv:2108.04215, [[arXiv:2108.04215](#)].
- [71] J. Yoo, N. Dalal, and U. Seljak, *Supersonic relative velocity effect on the baryonic acoustic oscillation measurements*, *Journal of Cosmology and Astro-Particle Physics* **2011** (July, 2011) 018, [[arXiv:1105.3732](#)].
- [72] D. Tseliakhovich and C. Hirata, *Relative velocity of dark matter and baryonic fluids and the formation of the first structures*, *Phys. Rev. D* **82** (Oct., 2010) 083520, [[arXiv:1005.2416](#)].
- [73] F. Schmidt, *Effect of relative velocity and density perturbations between baryons and dark matter on the clustering of galaxies*, *Phys. Rev. D* **94** (Sept., 2016) 063508, [[arXiv:1602.09059](#)].
- [74] S.-F. Chen, E. Castorina, and M. White, *Biased tracers of two fluids in the Lagrangian picture*, *Journal of Cosmology and Astro-Particle Physics* **2019** (June, 2019) 006, [[arXiv:1903.00437](#)].