Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **519,** 6246–6256 (2023)
Advance Access publication 2023 January 12

https://doi.org/10.1093/mnras/stad127

Interferometric H I intensity mapping: perturbation theory predictions and foreground removal effects

Alkistis Pourtsidou [1,2,3]★

[1] *Institute for Astronomy, The University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK*
[2] *Higgs Centre for Theoretical Physics, School of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3FD, UK*
[3] *Department of Physics & Astronomy, University of the Western Cape, Cape Town 7535, South Africa*

Accepted 2023 January 8. Received 2022 November 11; in original form 2022 July 1

ABSTRACT

We provide perturbation theory predictions for the H I intensity mapping power spectrum multipoles using the Effective Field Theory of Large Scale Structure, which should allow us to exploit mildly non-linear scales. Assuming survey specifications typical of proposed interferometric H I intensity mapping experiments like Canadian Hydrogen Observatory and Radio transient Detector and PUMA, and realistic ranges of validity for the perturbation theory modelling, we run mock full shape Markov chain Monte Carlo (MCMC) analyses at $z = 0.5$, and compare with Stage-IV optical galaxy surveys. We include the impact of 21cm foreground removal using simulations-based prescriptions, and quantify the effects on the precision and accuracy of the parameter estimation. We vary 11 parameters in total: three cosmological parameters, seven bias and counter terms parameters, and the H I brightness temperature. Amongst them, the four parameters of interest are: the cold dark matter density, ω_c, the Hubble parameter, h, the primordial amplitude of the power spectrum, A_s, and the linear H I bias, b_1. For the best-case scenario, we obtain unbiased constraints on all parameters with < 3 per cent errors at 68 per cent confidence level. When we include the foreground removal effects, the parameter estimation becomes strongly biased for ω_c, h, and b_1, while A_s is less biased ($<2\sigma$). We find that scale cuts $k_{\rm min} \geq 0.03\, h\,{\rm Mpc}^{-1}$ are required to return accurate estimates for ω_c and h, at the price of a decrease in the precision, while b_1 remains strongly biased. We comment on the implications of these results for real data analyses.

Key words: *(cosmology:)* large-scale structure of Universe – cosmology: observations – cosmology: theory – methods: statistical.

1 INTRODUCTION

Over the next few years, observations of the redshifted 21cm line emission from neutral hydrogen gas (H I) with a new generation of radio telescopes can push the boundaries of our understanding of cosmology and galaxy evolution. Remarkably, H I surveys can trace the matter distribution from the present time ($z = 0$) to the Epoch of Reionization ($z \sim 10$) and beyond, mapping a large part of the observable volume of the Universe.

In the meantime, spectroscopic optical galaxy surveys have already proven extremely successful at mapping the low redshift Universe, and providing exquisite constraints on dark energy and gravity (see e.g. Mueller et al. 2018; Alam et al. 2021). These surveys operate by detecting galaxies in three-dimensional (3D), i.e. by measuring the redshift and angular position of each galaxy very precisely. In the radio wavelengths, due to the weakness of the H I signal, being competitive with optical galaxy surveys using the traditional approach of detecting individual galaxies is extremely challenging. This challenge gave rise to an alternative observational technique, dubbed H I intensity mapping (IM), which maps the entire H I flux coming from many galaxies together in 3D voxels (Battye, Davies & Weller 2004; Chang et al. 2008; Peterson et al. 2010; Seo et al. 2010; Ansari et al. 2012). Provided several observational challenges and systematic effects are mitigated or controlled, the H I IM method has the potential to provide detailed maps of the Universe back to ∼1 billion yr after the big bang (Ahmed et al. 2019; Kovetz et al. 2020; Moresco et al. 2022).

A number of H I IM experiments are expected to come online over the coming years, with some of them already taking data with pilot surveys. Examples are the proposed MeerKLASS survey using the South African MeerKAT array (Santos et al. 2017), FAST (Hu et al. 2020), CHIME (Bandura et al. 2014), HIRAX (Newburgh et al. 2016; Crichton et al. 2022), Tianlai (Li et al. 2020; Wu et al. 2021), PUMA (Slosar et al. 2019), and Canadian Hydrogen Observatory and Radio transient Detector (CHORD; Vanderlinde et al. 2019). Pathfinder surveys with the *Green Bank Telescope*, *Parkes*, CHIME, and MeerKAT, have achieved detections of the cosmological 21 cm emission, but have relied on cross-correlation analyses with optical galaxy surveys (Chang et al. 2010; Masui et al. 2013; Anderson et al. 2018; Amiri et al. 2022a; Wolz et al. 2022; Cunnington et al. 2023).

A major challenge for the H I IM method is the presence of strong astrophysical emission: 21cm foregrounds such as galactic synchrotron (Zheng et al. 2017), point sources, and free–free emission, contaminate the maps and can be orders of magnitude stronger than the cosmological H I signal (Oh & Mack 2003). Hence, they

★ E-mail: alkistis.pourtsidou@ed.ac.uk

have to be removed. We can differentiate these dominant foregrounds from the signal taking advantage of their spectral smoothness (Liu & Tegmark 2011; Chapman et al. 2012; Wolz et al. 2014; Alonso et al. 2015; Shaw et al. 2015; Cunnington et al. 2020). As an example, 21cm foreground removal studies using low-redshift H I IM simulations and real data employ blind foreground removal techniques like principal component analysis (PCA, Switzer et al. 2013, 2015; Alonso et al. 2015) or independent component analysis (Hyvärinen 1999; Wolz et al. 2017). The procedure of foreground removal results in H I signal loss, removing power from modes along the parallel and perpendicular to the line-of-sight (LoS) directions, with the parallel to the LoS effect being more severe than the perpendicular one (Witzemann et al. 2019; Cunnington et al. 2020).

The aim of this work is to investigate the systematic biases from 21cm foreground removal assuming state-of-the-art interferometric H I IM experiments. To quantify how these systematic biases propagate on the cosmological parameter estimation, we will model the H I signal using perturbation theory and run full shape MCMC analyses on synthetic data contaminated with 21cm foreground removal effects. We will also benchmark our predictions against a 'Stage-IV' spectroscopic optical galaxy survey like DESI (Aghamousa et al. 2016) or Euclid (Laureijs et al. 2011; Blanchard et al. 2020).

The paper is organized as follows: In Section 2, we present the perturbation theory model we will use. In Section 3, we produce our synthetic mock data and contaminate them with simulated 21cm foreground removal effects. We present the results of our full shape MCMC runs in Section 4. In Section 5, we summarize our findings and conclude.

2 THEORETICAL MODELLING

Our observable is the H I power spectrum multipoles, and we follow the formalism used in optical galaxy surveys analyses. Similarly to optical galaxies, redshift space distortions (RSDs) introduce anisotropies in the observed H I power spectrum. In order to account for this, we consider the 3D power spectrum as a function of redshift z, k, and μ, where k is the amplitude of the wave vector and μ the cosine of the angle between the wave vector and the LoS component. This gives $k_\perp = \sqrt{1-\mu^2}$ and $k_\parallel = k\mu$.

Before, we present the 1-loop perturbation theory model we will use, it is useful to discuss linear theory. We can model RSDs by considering the Kaiser effect (Kaiser 1987), which is a large-scale effect dependent on the growth rate, f. To linear order, the anisotropic H I power spectrum can be written as:

$$P_{\rm HI}(k,\mu) = \left(\overline{T}_{\rm HI} b_1 + \overline{T}_{\rm HI} f\mu^2\right)^2 P_{\rm m}(k) + P_{\rm SN} + P_{\rm N}. \tag{1}$$

Here, $P_{\rm m}(k)$ is the underlying matter power spectrum, b_1 is the (linear) H I bias, and $\overline{T}_{\rm HI}$ is the mean H I brightness temperature. $P_{\rm N}$ is the thermal noise of the telescope and $P_{\rm SN}$ is the shot noise, $P_{\rm SN} = \overline{T}^2_{\rm HI}(1/\overline{n})$, where $\overline{n}$ is the number density of objects. The $P_{\rm SN}$ contribution is expected to be subdominant (smaller than the thermal noise of the telescope) and is usually neglected (Villaescusa-Navarro et al. 2018). The noise power spectrum for a typical interferometer is given by (Zaldarriaga, Furlanetto & Hernquist 2004; Bull et al. 2015):

$$P_{\rm N} = T^2_{\rm sys} r^2 y_\nu \left(\frac{\lambda^4}{A_{\rm e}^2}\right) \frac{1}{2n(u = k_\perp r/2\pi) t_{\rm total}} \left(\frac{S_{\rm area}}{\rm FOV}\right). \tag{2}$$

Here, $A_{\rm e}$ is the effective beam area, FOV $\approx \lambda/(D_{\rm dish})^2$, r is the comoving distance to the observation redshift z, and $y_\nu = c(1+z)^2/(\nu_0 H(z))$ with $\nu_0 = 1420$ MHz. $T_{\rm sys}$ is the system temperature, $S_{\rm area}$ is the survey area, and $t_{\rm total}$ is the total observing time.

The antennae distribution function $n(u)$ can be calculated using a fitting formula (Ansari et al. 2018). For a square array with $N_{\rm s}^2$ receivers, the number of baselines as a function of physical distance of antennas is given by

$$n_{\rm b}^{\rm phys}(l) = n_0 \frac{a + b(l/L)}{1 + c(l/L)^d} e^{-(l/L)^e}, \tag{3}$$

where $n_0 = (N_{\rm s}/D_{\rm dish})^2$, $L = N_{\rm s} D_{\rm dish}$, and the uv-plane density is

$$n(u) = \lambda^2 n_{\rm b}^{\rm phys}(l = u\lambda). \tag{4}$$

The H I abundance and clustering properties have been studied using simulations and semi-analytical modelling (see e.g. Padmanabhan, Refregier & Amara 2017; Villaescusa-Navarro et al. 2018; Spinelli et al. 2020). The clustering of H I should be accurately described by perturbative methods at mildly non-linear scales (Sarkar, Bharadwaj & Anathpindika 2016; McQuinn & D'Aloisio 2018; Castorina & White 2019; Sarkar & Bharadwaj 2019; Sailer et al. 2021; Karagiannis, Maartens & Randrianjanahary 2022; Qin et al. 2022). Modelling non-linear scales is necessary in order to get precise and accurate cosmological constraints with instruments like HIRAX, CHIME, CHORD, and PUMA, and it also helps break degeneracies, e.g. between b_1 and the primordial power spectrum amplitude, $A_{\rm s}$. Similar degeneracies exist for $\overline{T}_{\rm HI}$, which is proportional to the H I mean density, $\Omega_{\rm HI}$. Accurate (< 5 per cent) measurements of $\Omega_{\rm HI}(z)$ are available at low redshifts (Crighton et al. 2015), and it can also be constrained by joint analyses of different probes (Obuljen et al. 2018; Chen et al. 2019) or by exploiting very small scales that can be described using bespoke H I halo models (Chen et al. 2021).

In this work, we will use the 'EFTofLSS' formalism that has been developed to model the power spectrum multipoles of biased tracers in redshift space. The main difference between this model and the standard 1-loop Standard Perturbation Theory formalism (Bernardeau et al. 2002) is that the EFTofLSS approach accounts for the impact of non-linearities on mildly non-linear scales by introducing effective stresses in the equations of motion. This results in the addition of counter terms to the 1-loop power spectrum, which represents the effects of short distance physics at long distances.

The EFTofLSS model we will employ is described in various papers (see e.g. Perko et al. 2016 and references therein), and we refer to D'Amico et al. (2020) for its application to the DR12 BOSS data. Main assumptions are that we live in a spatially expanding, homogeneous, and isotropic background space–time, and that we work on subhorizon scales with $\delta, \theta \ll 1$ (where δ and θ are the density and velocity perturbations, respectively).

The 1-loop redshift space galaxy power spectrum then reads (Perko et al. 2016; D'Amico et al. 2020)

$$\begin{aligned} P_g(k,\mu) =\ & Z_1(\mu)^2 P_{11}(k) \\ & + 2\int \frac{d^3q}{(2\pi)^3} Z_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}, \mu)^2 P_{11}(|\boldsymbol{k}-\boldsymbol{q}|) P_{11}(q) \\ & + 6 Z_1(\mu) P_{11}(k) \int \frac{d^3q}{(2\pi)^3} Z_3(\boldsymbol{q}, -\boldsymbol{q}, \boldsymbol{k}, \mu) P_{11}(q) \\ & + 2 Z_1(\mu) P_{11}(k) \left(c_{\rm ct}\frac{k^2}{k_{\rm M}^2} + c_{{\rm r},1}\mu^2\frac{k^2}{k_{\rm M}^2} + c_{{\rm r},2}\mu^4\frac{k^2}{k_{\rm M}^2}\right) \\ & + \frac{1}{\bar{n}_g}\left(c_{\epsilon,1} + c_{\epsilon,2}\frac{k^2}{k_{\rm M}^2} + c_{\epsilon,3} f\mu^2\frac{k^2}{k_{\rm M}^2}\right), \end{aligned} \tag{5}$$

where $k_{\rm M} = 0.7h\,{\rm Mpc}^{-1}$ and $\bar{n}_{\rm g}$ is the mean galaxy density. The various terms are summarized nicely in Nishimichi et al. (2020), and

we follow this description here: the $c_{\rm ct}$ term represents a combination of a higher derivative bias and the speed of sound of dark matter; the $c_{{\rm r},i}$ terms represent the redshift-space counter terms, while the $c_{\epsilon,i}$ terms represent the stochastic counter terms. The kernels Z_1, Z_2, and Z_3 are the redshift-space galaxy density kernels appearing in the 1-loop power spectra. They are expressed in terms of the galaxy density and velocity kernels and four bias parameters: $\{b_1, b_2, b_3,$ and $b_4\}$. For flat Lambda cold dark matter, which we will assume in this work, the logarithmic growth rate f is calculated by solving for the linear growth factor D (with a the scale factor), and yields: $f(a) = \frac{(5a-3D(a))\Omega_{\rm m}}{2D(a)\left(\Omega_{\rm m}+a^3(1-\Omega_{\rm m})\right)}$.

The model of equation (5) has recently been implemented in a publicly available PYTHON code, PYBIRD (D'Amico, Senatore & Zhang 2021). In principle, the model can describe any biased tracer of matter, so we can straightforwardly apply it to H I. Following the literature we perform the following changes of variables: $b_2 = \frac{1}{\sqrt{2}}(c_2 + c_4)$, $b_4 = \frac{1}{\sqrt{2}}(c_2 - c_4)$, $c_{\epsilon,\rm mono} = c_{\epsilon,1} + \frac{1}{3}c_{\epsilon,2}$, and $c_{\epsilon,\rm quad} = \frac{2}{3}c_{\epsilon,2}$, and also fix $c_4 = c_{\rm r,2} = c_{\epsilon,\rm mono} = 0$ so that our final set of nuisance parameters is: $\{b_1, c_2, b_3, c_{\rm ct}, c_{\rm r,1}, c_{\epsilon,1}, c_{\epsilon,\rm quad}\}$. We will comment on these choices when we construct our mock data in Section 3.

The cosmological parameters that the code takes as input are: the cold dark matter density $\omega_{\rm c} = \Omega_{\rm c}h^2$, the baryonic matter density $\omega_{\rm b} = \Omega_{\rm b}h^2$, the Hubble parameter h, the amplitude of the primordial power spectrum, $A_{\rm s}$, and the scalar spectral index, $n_{\rm s}$. We will describe the code and other software we used to speed-up the parameter inference in more detail in Section 4.

3 MOCK DATA

For our analysis we produce synthetic H I monopole and quadrupole data running PYBIRD for a central redshift $z = 0.5$. We will not use the hexadecapole as it is not expected to add significant cosmological information, and it is more affected by non-linear uncertainties. In addition, as shown in Cunnington et al. (2020) and Soares et al. (2021), the H I IM hexadecapole (as well as higher order multipoles) can be used for identifying the effects of foreground removal and other systematics. Not using the hexadecapole allows us to set $c_{\rm r,2} = 0$. The choice $c_4 = c_{\epsilon,\rm mono} = 0$ is motivated by the assumption that the functions multiplying c_4 and $c_{\epsilon,\rm mono}$ are too small to affect the results. These assumptions follow the BOSS data analyses choices (D'Amico et al. 2020), but they will need to be reaffirmed with bespoke H I simulations and real data. The fiducial cosmological parameters are (Aghanim et al. 2020):

$$\{\omega_{\rm c}, h, A_{\rm s}, \omega_{\rm b}, n_{\rm s}\} = \{0.1193, 0.677, 3.047, 0.0224, 0.967\}.$$

For setting the fiducial values of the nuisance parameters, we perform fits to H I IM simulations. These are described in Section A, and we find

$$\{b_1, c_2, b_3, c_{\rm ct}, c_{\rm r,1}, c_{\epsilon,1}, c_{\epsilon,\rm quad}\} = \{1.1, 0.6, 0.1, 0.1, -10, 0, -0.8\}.$$

We remark that the value of the linear H I bias b_1 is in very good agreement with values found at similar redshifts in other works (Sarkar et al. 2016; Villaescusa-Navarro et al. 2018). We also note that in all our MCMC forecasts we will marginalize over the nuisance parameters.

The model in equation (5) has to be rescaled by the square of the H I brightness temperature, $\bar{T}_{\rm HI}(z)$, which in turn depends on the H I abundance, $\Omega_{\rm HI}(z)$. Using the fitting function from SKA Cosmology Science Working Group et al. (2020) we set $\bar{T}_{\rm HI}(z = 0.5) = 0.168$ mK as our fiducial value for this parameter.

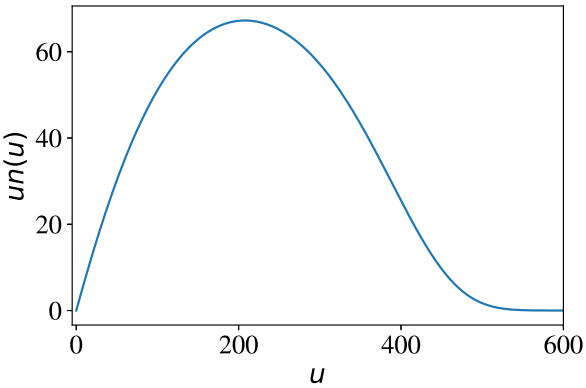

Figure 1. Baseline density for our CHORD-like array, calculated using equation (4).

We also need a data covariance. To calculate this, we will assume an ambitious CHORD-like IM survey (Vanderlinde et al. 2019). CHORD is a successor to CHIME (Amiri et al. 2022b), aiming to perform a very large sky H I IM survey. Its core array consists of 512 dishes, each 6 m in diameter. The bandwidth is large, covering the 300–1500 MHz band, or redshifts up to $z = 6$. We will assume $T_{\rm sys} = 50$ K in our forecasts. Another very ambitious proposal is PUMA, a close-packed interferometer array with 32 000 dishes, covering the frequency range 200–1100 MHz, or redshifts $0.3 < z < 6$ (Slosar et al. 2019). We expect both of these instruments to be able to achieve similar signal-to-noise ratios (S/N), and we will focus on CHORD from now on.

Aiming to establish how CHORD can complement and compete with state-of-the-art optical galaxy surveys, we choose a low redshift bin centred at $z = 0.5$ with width $\Delta z = 0.3$. Our fiducial CHORD-like survey covers 20 000 deg^2 on the sky, resulting in a survey bin volume $V_{\rm sur} = 4246^3\,({\rm Mpc}/h)^3$. We assume a 20 000-h survey and calculate the noise power spectrum using equation (2), with the fitting parameters needed in equation (4) being $a = 0.4847$, $b = -0.3300$, $c = 1.3157$, $d = 1.5974$, and $e = 6.8390$ (Ansari et al. 2018). The corresponding baseline density $n(u)$ for our CHORD-like array is shown in Fig. 1 (see appendix A of Karagiannis et al. 2022 for the case of a HIRAX-like array).

We can now calculate $P_{\rm N}$ for our CHORD-like survey using equation (2). Dividing by $\bar{T}^2_{\rm HI}$ and inverting, we can define an effective mean number density $\bar{n}_{\rm HI}$. For a Stage-IV spectroscopic galaxy survey like DESI (Aghamousa et al. 2016) or Euclid (Laureijs et al. 2011; Blanchard et al. 2020), the shot noise is the inverse of the number density of galaxies, $\bar{n}_{\rm g}$. In Fig. 2, we plot the S/N (squared) for the power spherically averaged power spectrum (the monopole, P_0) for different surveys. Stage-IV corresponds to a spectroscopic optical galaxy survey with $\bar{n}_{\rm g} = 0.0005\,h^3\,{\rm Mpc}^{-3}$. 'CHORD-optimal' corresponds to an idealized case for an interferometric H I IM survey without any systematic effects, while 'CHORD-with-sys' illustrates the case where sensitivity is lost at small k (see also fig. 15 in Ansari et al. 2018). This can be due to foreground removal effects which mainly affect the small $k_\parallel$, and/or inability to probe the small $k_\perp$ due to baseline restrictions. For the case of the CHORD-like survey at $z = 0.5$, this can result to loss of sensitivity in a range $\sim 0.001 \leq k \leq 0.06\,h\,{\rm Mpc}^{-1}$, and we will consider different $k_{\rm min}$ scale cuts in our forecasts to take this into account. In all cases, Fig. 2 demonstrates that a CHORD-like experiment can achieve a higher S/N in the non-linear regime, compared to a Stage-IV optical galaxy survey.

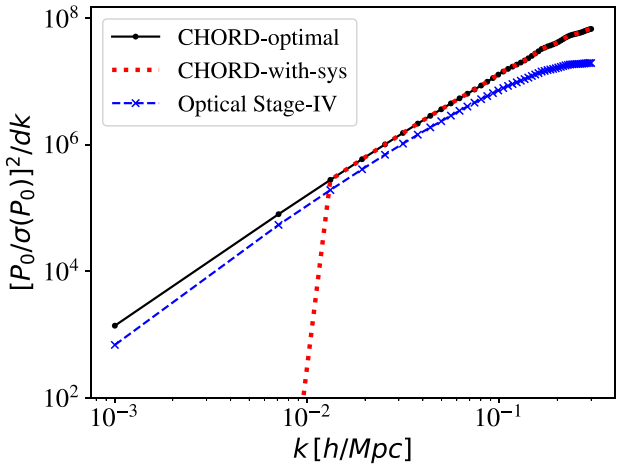

Figure 2. *S/N* (squared) for the spherically averaged power spectrum, P_0, for different surveys. 'Optical Stage-IV' represents a spectroscopic galaxy survey with $\bar{n}_g = 0.0005\,h^3\,\mathrm{Mpc}^{-3}$. 'CHORD-optimal' corresponds to an idealized H I IM survey, while 'CHORD-with-sys' illustrates the case where sensitivity is lost at small k.

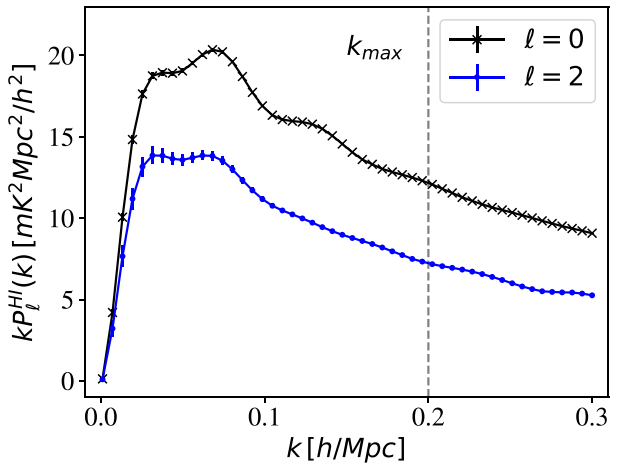

Figure 3. Our fiducial monopole ($\ell = 0$) and quadrupole ($\ell = 2$) data, assuming the CHORD-optimal H I IM survey. The vertical dashed grey line denotes the maximum wavenumber (smallest scale used in the MCMC) $k_{max} = 0.2h\,\mathrm{Mpc}^{-1}$ at our chosen central redshift $z = 0.5$.

We can now proceed to calculate the multipole covariances analytically, using the Gaussian approximation (Taruya, Nishimichi & Saito 2010; Soares et al. 2021). We present the resulting mock data and measurement errors in Fig. 3 for the CHORD-optimal case. We notice that in the case of the monopole the error bars are not large enough to be visible. At this low redshift, and with such high *S/N*, non-linear uncertainties are expected to become important at a relatively low k. Hence, we choose the range of validity for the EFTofLSS modelling to be $k < k_{max} = 0.2\,h$ Mpc (this should be a good assumption, but it has to be validated with tailored H I simulations for these experiments). We are now ready to perform MCMC forecasts.

4 MCMC ANALYSES

To calculate posterior distributions on the parameters we have run MCMCs using the ensemble slice sampling codes `EMCEE` (Foreman-Mackey et al. 2013) and `zeus` (Karamanis, Beutler & Peacock 2021). The latter has been recently used to run mock full shape MCMC analyses assuming galaxy surveys specifications using the `MATRYOSHKA` suite of neural network based emulators (Donald-McCann et al. 2022; Donald-McCann, Koyama & Beutler 2023). Due to the impressive increase in computational speed for the inference (∼3 orders of magnitude improvement with respect to the `PYBIRD` runs), we opted for this set-up to run and present our final MCMC forecasts.[1] We vary three cosmological parameters, [ω_c, h, and $\ln(10^{10}A_s)$], seven bias and counter terms parameters, $(b_1, c_2, b_3, c_{ct}, c_{r,1}, c_{\epsilon,1},$ and $c_{\epsilon,\mathrm{quad}})$, and, in the case of IM, we also vary $\bar{T}_{HI}$. The scalar spectral index n_s is fixed to its true value, and so is the baryon fraction $f_b = \omega_b/(\omega_b + \omega_c)$.

For the three cosmological parameters and b_1, we assume the uniform flat priors shown in Table 1. We do not employ Planck priors on A_s, ω_c, and h because we wish to assess the precision versus accuracy performance of interferometric H I IM independently of CMB experiments. For $\bar{T}_{HI}$ we take a flat prior [0,1]. For the rest of the bias and counter terms parameters, we follow D'Amico et al. (2020) and set

$$c_2 \sim \mathcal{U}(-4,4), \qquad b_3 \sim \mathcal{N}(0,2),$$
$$c_{ct} \sim \mathcal{N}(0,2), \qquad c_{r,1} \sim \mathcal{N}(0,8),$$
$$c_{\epsilon,1}\ /\ n_{HI} \sim \mathcal{N}(0,400), \qquad c_{\mathrm{quad}} \sim \mathcal{N}(0,2).$$

Finally, we assume a Gaussian likelihood given by

$$\ln \mathcal{L}(P^d \mid \theta) = -\frac{1}{2}(P^d - P^m)^T \boldsymbol{C}^{-1}(P^d - P^m)\,, \tag{6}$$

with P^d being the mock data (the power spectrum monopole, P_0, and quadrupole, P_2), P^m being the EFTofLSS model predictions for a given set of parameters, θ, and $\boldsymbol{C}$ being the covariance matrix.

4.1 The systematics-free, Stage-IV survey scenario

We start by comparing the performance of a Stage-IV spectroscopic galaxy survey and an analogous IM survey, assuming both of them are free of systematic effects. The volumes of the surveys are taken to be exactly the same,[2] but the effective mean number densities (i.e. the noise components) are different as we have described in detail in Section 3 (see e.g. Fig. 2). Following up on the discussion in the previous section, we emphasize again that in the case of H I IM there is an additional overall amplitude parameter, $\bar{T}_{HI}^2 \propto \Omega_{HI}^2$, which we vary. This means that a total of 10 (11) parameters are varied in the MCMC for the optical (IM) surveys under consideration.

The MCMC contours for the idealized case are shown in Fig. 4. We are able to recover the true values in an unbiased manner, which also confirms the accuracy of the `MATRYOSHKA` emulator (see Donald-McCann et al. 2023 for a suite of validation tests). The Stage-IV spectroscopic optical galaxy survey and the CHORD-like H I IM survey have similar constraining power when the same survey volume is assumed. An exception is the primordial amplitude A_s. In the CHORD-like IM case, additional degeneracies due to varying $\bar{T}_{HI}$ increase the uncertainty in A_s compared to the optical case.

[1] An alternative approach to speed-up the inference is to use a fast and accurate linear matter power spectrum emulator such as `bacco` (Aricò, Angulo & Zennaro 2021) or `CosmoPower` (Spurio Mancini et al. 2022) as input in a perturbation theory code, instead of running a Boltzmann solver like `CAMB` (Lewis, Challinor & Lasenby 2000) or `CLASS` (Blas, Lesgourgues & Tram 2011; Lesgourgues 2011).

[2] H I IM surveys can cover a much wider redshift range compared to spectroscopic optical galaxy surveys; however, our goal here is to compare their performance on a given redshift bin.

Table 1. Fiducial values, prior ranges, and marginalized constraints for the four parameters of interest (68 per cent confidence level). We vary 10 parameters in total for optical galaxy surveys, and 11 for IM (adding the $\bar{T}_{\rm HI}$ parameter). For the IM surveys, we consider different $k_{\rm min}$ limits to mitigate 21cm foreground removal effects.

Parameters of interest		$\omega_{\rm c}$	h	$\ln(10^{10}A_{\rm s})$	b_1
Fiducial values		0.1193	0.677	3.047	1.1
Priors		[0.101,0.140]	[0.575,0.748]	[2.78,3.32]	[0,4]
Case	k-range ($h\,{\rm Mpc}^{-1}$)				
Optical galaxy survey (Fig. 4)	$0.001 < k < 0.2$	$0.119^{+0.003}_{-0.003}$	$0.676^{+0.008}_{-0.008}$	$3.05^{+0.04}_{-0.04}$	$1.1^{+0.03}_{-0.02}$
IM-noFG (Fig. 4)	$0.001 < k < 0.2$	$0.119^{+0.003}_{-0.003}$	$0.676^{+0.007}_{-0.007}$	$3.00^{+0.08}_{-0.08}$	$1.1^{+0.02}_{-0.02}$
IM-subFG, $k_{\rm min} = 0.01\,h\,{\rm Mpc}^{-1}$ (Fig. 7)	$0.01 < k < 0.2$	$0.1282^{+0.0011}_{-0.0005}$	$0.697^{+0.003}_{-0.002}$	$3.12^{+0.06}_{-0.05}$	$0.95^{+0.02}_{-0.02}$
IM-subFG, $k_{\rm min} = 0.03\,h\,{\rm Mpc}^{-1}$ (Figs 7 and 8)	$0.03 < k < 0.2$	$0.125^{+0.004}_{-0.002}$	$0.689^{+0.010}_{-0.005}$	$3.12^{+0.07}_{-0.07}$	$0.96^{+0.02}_{-0.02}$
IM-subFG, $k_{\rm min} = 0.05\,h\,{\rm Mpc}^{-1}$ (Figs 7 and 8)	$0.05 < k < 0.2$	$0.121^{+0.007}_{-0.004}$	$0.679^{+0.015}_{-0.010}$	$3.09^{+0.09}_{-0.07}$	$0.98^{+0.03}_{-0.03}$

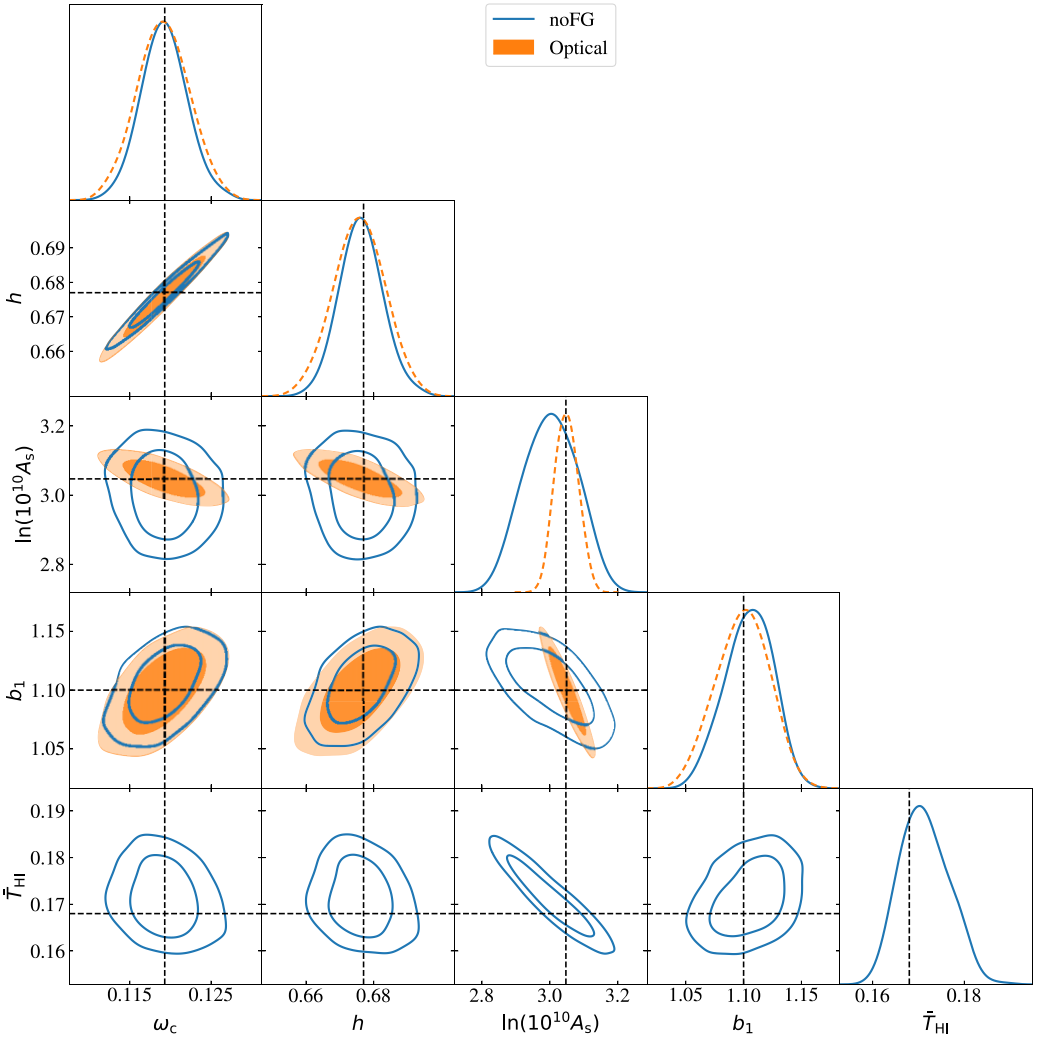

Figure 4. Marginalized 1D and 2D posterior distributions, and 1σ and 2σ contours, for a Stage-IV spectroscopic optical galaxy survey and a CHORD-like IM survey. The black dashed line shows the fiducial (true) parameters. We show the four parameters of interest and $\bar{T}_{\rm HI}$, but we remind the reader that a total of 10 parameters are included in the MCMC for the optical survey, and 11 for the IM survey (see Appendix B, Fig. B1).

From Table 1 we see that, in the absence of systematics, the CHORD-like H I IM survey determines $\omega_{\rm c}$ with < 3 per cent error, h with 1 per cent error, $\ln(10^{10}A_{\rm s})$ with < 3 per cent error, and b_1 (the linear H I bias) with < 2 per cent error. The survey can also constrain $\bar{T}_{\rm HI} = 0.171^{+0.005}_{-0.006}$, demonstrating how exploiting mildly non-linear scales can break degeneracies.

4.2 Contaminating the data vector with 21cm foreground removal effects

In order to contaminate our synthetic data vector (i.e. the H I power spectrum multipoles P_0 and P_2 in Fig. 3) with 21cm foreground removal effects, we will use the simulations-based prescription by

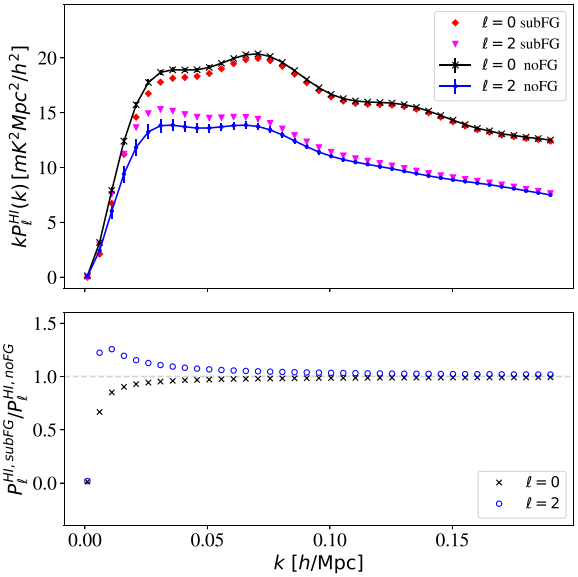

Figure 5. Top panel: Fiducial monopole ($\ell = 0$) and quadrupole ($\ell = 2$) data, with and without foreground removal effects, assuming a CHORD-like IM survey. Bottom panel: The ratios of the power spectrum multipoles illustrate the amplitude and scale dependence of the foreground removal effects.

Soares et al. (2021). This prescription can fit H I IM simulations with foreground removal effects, assuming that PCA or FastICA with $N_{\rm IC} = 4$ independent components was used for the foreground cleaning. The choice $N_{\rm IC} = 4$ corresponds to an excellent calibration scenario (which will hopefully be the case by the time CHORD and PUMA come online) and no polarization leakage (Wolz et al. 2014; Alonso et al. 2015; Liu & Shaw 2020; Cunnington et al. 2021). For existing H I IM surveys, we know that much more aggressive foreground removal (much higher $N_{\rm IC} \sim 20$–30) is employed to deal with more complicated foregrounds, noise, and unknown systematics (see e.g. Masui et al. 2013; Switzer et al. 2013; Wolz et al. 2022; Cunnington et al. 2023).

We present the result of contaminating our data vector with 21cm foreground removal effects in Fig. 5. This is our mock data vector for the remainder of the paper. As we can see, foreground cleaning results in the damping of power across a wide range of scales in the spherically averaged power spectrum (the monopole, P_0). This is a well-known effect, which has also been identified in the context of high-redshift 21cm surveys of the Epoch of Reionization (Petrovic & Oh 2011). Higher order multipoles were first studied extensively in Blake (2019), Cunnington et al. (2020), and Soares et al. (2021), focusing on post-reionization H I. In the case of the quadrupole (P_2), where $P(k, \mu)$ is weighted as a function of μ, we see an enhancement of power on large scales. It is also important to note that, both for P_0 and P_2, while the largest effect is clearly on small k, there is still an effect along larger k. Given that the error bars of our chosen surveys are extremely small, the large k effect may introduce a significant systematic bias as well. We can only verify and quantify this at the level of the parameter inference, and we will do so in the following sections.

We remark that we assume no prior knowledge of the 21cm foreground removal fitting function from Soares et al. (2021). That is, we will not attempt to include a model (and associated nuisance parameters we need to vary) for the 21cm foreground removal effects in our theory vector. We will instead follow a more conservative, 'data-driven' approach, imposing scale cuts on the contaminated

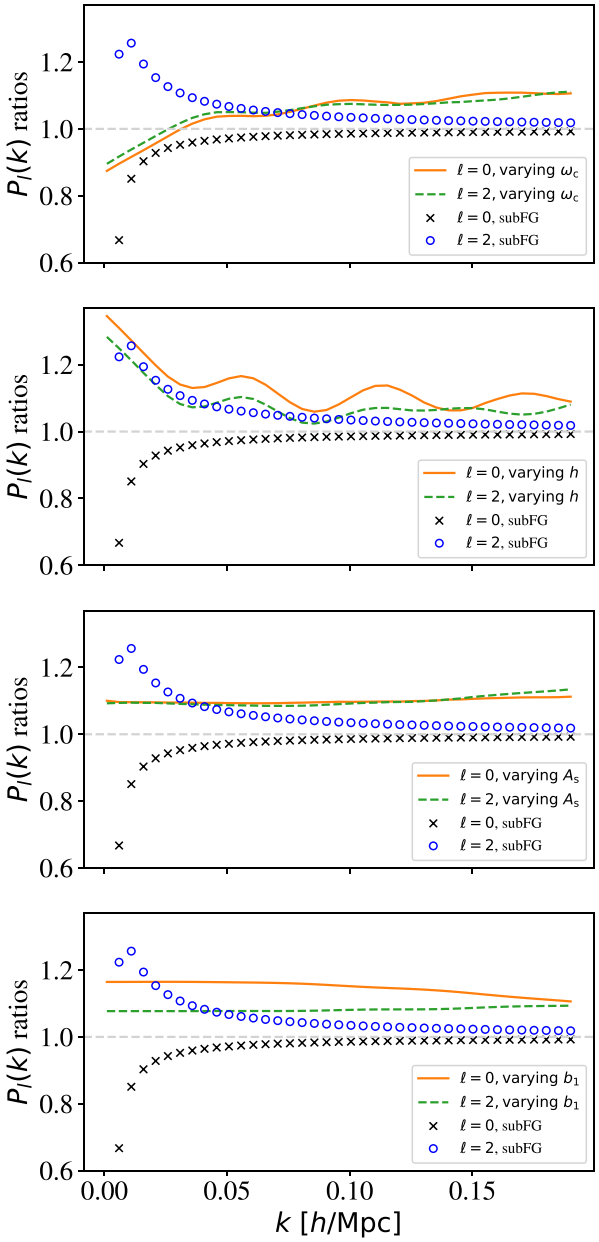

Figure 6. The effect of $\sim$10 per cent variations of the four parameters of interest on the EFTofLSS model predictions. For comparison, we also plot the simulations-based effect of 21cm foreground removal.

data vector. However, the former method has been shown to be very promising in the context of single-dish experiments, and it would be important to re-evaluate its performance when realistic, end-to-end simulations are available (see e.g. the discussion in Spinelli et al. 2022).

Finally, before presenting our MCMC analyses using the contaminated data vector, it is instructive to see how varying our four parameters of interest affects the predictions of the EFTofLSS model of equation (5), and compare with the features of the 21cm foreground removal effects. This comparison is shown in Fig. 6. It is well known that different cosmological and nuisance parameters affect the power spectrum amplitude and shape in distinct ways. Comparing these with our simulated 21cm foreground removal effects suggests that some parameters, e.g. h, have features on the mildly non-linear scales

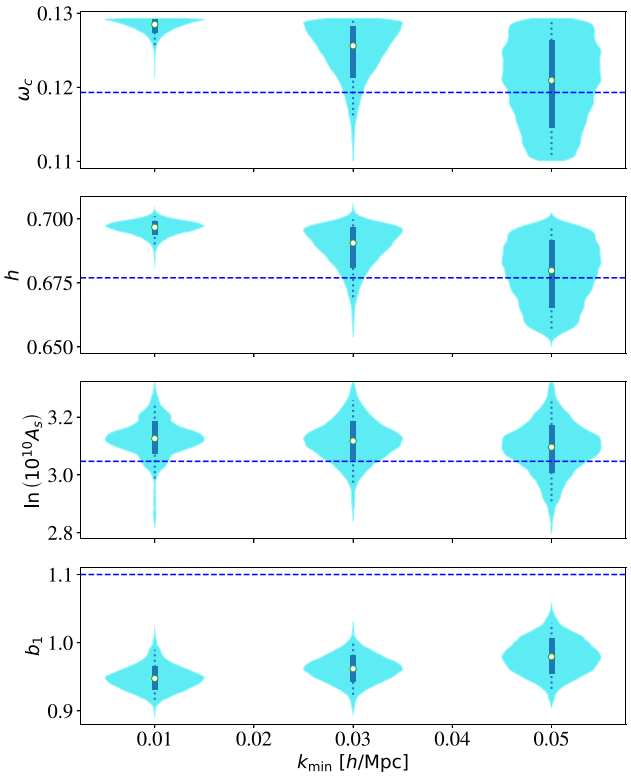

Figure 7. Violin plots showing the marginalized 1D posteriors on the four parameters of interest for the different $k_{\rm min}$ limits we have considered for the CHORD-like IM survey contaminated by foreground removal effects. The dashed blue lines show the true (fiducial) values. The white points show the median values. The thick solid (thin dotted) blue lines show the 1σ (2σ) regions. The shaded cyan regions show the density of the marginalized posteriors. We remind the reader that a total of 11 parameters are included in the MCMC.

that might make them easier to constrain in an unbiased way (i.e. to disentangle them from the foreground removal effects) than others, e.g. $A_{\rm s}$ and b_1.

4.3 Imposing $k_{\rm min}$ cuts: precision versus accuracy

In this section we perform an MCMC analysis for the CHORD-like survey under consideration, using the contaminated H I IM data vector with different $k_{\rm min}$ limits. Our results for the four parameters of interest are summarized in Table 1 and Fig. 7. The different scale cuts we consider are the following:

(i) *Case* I: We start by imposing a scale cut $k_{\rm min} = 0.01\ h\,{\rm Mpc}^{-1}$ in order to exclude the largest scales where foreground subtraction has the most impact. This is not enough: The parameter estimation becomes strongly biased for $\omega_{\rm c}$, h, and b_1. The primordial amplitude is unbiased within $<2\sigma$ because of the marginalization of $\bar{T}_{\rm HI}$ (i.e. if $\bar{T}_{\rm HI}$ was kept fixed, $A_{\rm s}$ would also be strongly biased). For the H I brightness temperature we get $\bar{T}_{\rm HI} = 0.173^{+0.004}_{-0.004}$.

(ii) *Case* II: Imposing a stricter scale cut $k_{\rm min} = 0.03\ h\,{\rm Mpc}^{-1}$. In this case the $\omega_{\rm c}$, h, and $A_{\rm s}$ parameters are unbiased within 2σ, while b_1 remains biased. For the H I brightness temperature we get $\bar{T}_{\rm HI} = 0.174^{+0.004}_{-0.005}$.

(iii) *Case* III: Imposing a scale cut $k_{\rm min} = 0.05\ h\,{\rm Mpc}^{-1}$. In this case the $\omega_{\rm c}$, h, and $A_{\rm s}$ parameters are unbiased within 1σ, while b_1 remains biased. For the H I brightness temperature we get $\bar{T}_{\rm HI} = 0.176^{+0.004}_{-0.006}$.

In Fig. 8, we compare the idealized case (noFG) and the cases with $k_{\rm min} \geq 0.03\,h\,{\rm Mpc}^{-1}$ limits that lead to unbiased (within 2σ) constraints on $\omega_{\rm c}$, h, and $A_{\rm s}$. Looking back at Fig. 6, we deduce that for $k_{\rm min} \geq 0.03\ h\,{\rm Mpc}^{-1}$ the scale-dependent features of varying $\omega_{\rm c}$ and h are sufficient to constrain them accurately, in contrast to b_1, which gets significantly biased due to the amplitude change from the 21cm foreground removal. The primordial amplitude $A_{\rm s}$ is less affected because of the marginalization of $\bar{T}_{\rm HI}$ (we have checked that when $\bar{T}_{\rm HI}$ is kept fixed, $A_{\rm s}$ becomes strongly biased).

From Table 1 and Figs 7 and 8, we also see that the price to pay for the unbiased estimates of $\omega_{\rm c}$ and h parameters is a decrease in the precision with respect to the idealized case, as expected.

5 CONCLUSIONS

We have provided perturbation theory predictions for H I IM, and performed full shape MCMC analyses including the impact of 21cm foreground removal. Albeit our framework was developed in the context of low-redshift, interferometric H I IM surveys like CHIME, HIRAX, CHORD, and PUMA, our results are also relevant for single-dish surveys as well as Epoch of Reionization surveys. The main conclusions we draw from this work are:

(i) In the idealized case without any systematic biases in the data, interferometric H I IM surveys with instruments like CHORD and PUMA are competitive with Stage-IV optical galaxy surveys. Our results, summarized in Table 1, assume a single-redshift bin centred at $z = 0.5$. In the case of redshift-independent parameters like $\omega_{\rm c}$, $A_{\rm s}$, and h, we naively expect the parameter estimation uncertainties to be reduced by roughly $1/\sqrt{N}$, where N is the number of redshift bins. This is an advantage for H I IM surveys with respect to optical surveys, since the former are not shot-noise limited and can rapidly cover very large redshift ranges. However, the thermal noise of an interferometer can increase rapidly with redshift. Nevertheless, forecasts for experiments similar to what we consider here have shown that interferometric H I IM should be advantageous at high redshift $2 < z < 6$. The cosmological volume spanned at this range is three times higher than the typical optical galaxy surveys range, $0 < z < 2$, with an increased $k_{\rm max}$, i.e. an increased number of (easier to model) linear and mildly non-linear modes. Forecasts for a dedicated 'Stage-II' H I IM experiment showed that $S/N > 1$ is achievable for all modes with $k \leq 0.4h\,{\rm Mpc}^{-1}$ at $z \leq 6$ (Ansari et al. 2018).

(ii) Including 21cm foreground removal effects based on simulations-based prescriptions, the parameter estimation becomes strongly biased for $\omega_{\rm c}$, h, and b_1, while $A_{\rm s}$ is less biased ($<2\sigma$).

(iii) Adopting the scale cuts approach to try and mitigate the biases, we find that scale cuts $k_{\rm min} \geq 0.03\ h\,{\rm Mpc}^{-1}$ are required to return accurate estimates for $\omega_{\rm c}$ and h, at the price of a decrease in the precision, while b_1 remains biased.

In future work, it would be interesting to investigate the possible HI and cosmology dependence of the foreground transfer function. This is a method to correct for signal loss from foreground removal, which has been used in all the H I-galaxy cross-correlation detections so far. Due to the low S/N of current experiments, the cosmology is kept fixed to the Planck best-fitting values and the only parameter we can constrain is $\Omega_{\rm HI} b_1 r$, with r the H I-galaxy cross-correlation coefficient (Masui et al. 2013; Anderson et al. 2018; Wolz et al. 2022; Cunnington et al. 2023). The foreground transfer function is constructed using mock simulations with fixed H I and cosmological parameters. In light of our results, we believe it is important to study how robust the transfer function construction (and the resulting H I

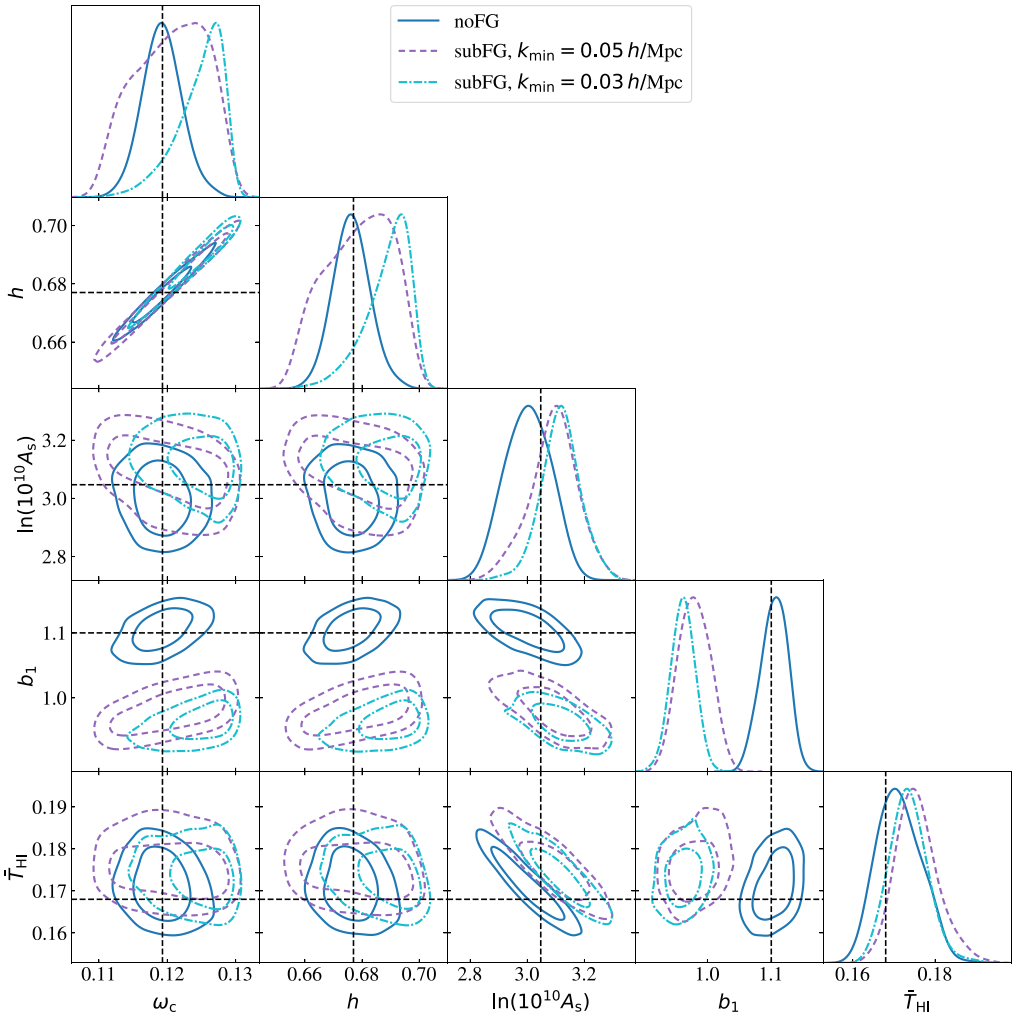

Figure 8. Marginalized 1D and 2D posterior distributions, and 1σ and 2σ contours for a CHORD-like H I IM survey. We compare the idealized case and the cases with $k_{\rm min} \geq 0.03\,h\,{\rm Mpc}^{-1}$ limits that lead to unbiased (within 2σ) constraints on $\omega_{\rm c}$, h, and $A_{\rm s}$. The black dashed line shows the fiducial (true) parameters. We show the four parameters of interest and $\bar{T}_{\rm HI}$, but we remind the reader that a total of 11 parameters are included in the MCMC (see Appendix B, Fig. B1).

signal loss correction) is with respect to varying the parameters in the mock simulations.

ACKNOWLEDGEMENTS

I acknowledge use of open source software: `PYTHON` (Van Rossum & Drake 1995; Hunter 2007), `NUMPY` (Van Der Walt, Colbert & Varoquaux 2011), `SCIPY` (Virtanen et al. 2020), `TENSORFLOW` (Abadi et al. 2015), `ASTROPY` (Astropy Collaboration et al. 2018), `CORNER` (Foreman-Mackey 2016), and `GETDIST` (Lewis 2019). I thank New Mexico State University (USA) and Instituto de Astrofisica de Andalucia CSIC (Spain) for hosting the Skies & Universes site for cosmological simulation products. My research is funded by a UK Research and Innovation Future Leaders Fellowship (grant MR/S016066/1). For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

DATA AVAILABILITY

The data products will be shared on reasonable request to the corresponding author.

REFERENCES

Abadi M. et al., 2015, preprint (arXiv:1603.04467)
Ade P. A. R. et al., 2016, Astron. Astrophys., 594, A13
Aghamousa A. et al., 2016, preprint (arXiv:1611.00036)
Aghanim N. et al., 2020, Astron. Astrophys., 641, A6
Ahmed Z. et al., 2019, preprint (arXiv:1907.13090)
Alam S. et al., 2021, Phys. Rev. D, 103, 083533
Alonso D., Bull P., Ferreira P. G., Santos M. G., 2015, MNRAS, 447, 400
Amiri M. et al., 2022a, preprint (arXiv:2202.01242)
Amiri M. et al., 2022b, ApJS, 261, 29
Anderson C. et al., 2018, MNRAS, 476, 3382
Ansari R. et al., 2012, Astron. Astrophys., 540, A129
Ansari R. et al., 2018, preprint (arXiv:1810.09572)
Aricò G., Angulo R. E., Zennaro M., 2021, preprint (arXiv:2104.14568)
Astropy Collaboration et al., 2018, AJ, 156, 123
Bandura K. et al., 2014, Proc. SPIE Conf. Ser. Vol. 9145, Canadian Hydrogen Intensity Mapping Experiment (CHIME) pathfinder. SPIE, Bellingham, p. 22
Battye R. A., Davies R. D., Weller J., 2004, MNRAS, 355, 1339
Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Blake C., 2019, MNRAS, 489, 153
Blanchard A. et al., 2020, Astron. Astrophys., 642, A191

Blas D., Lesgourgues J., Tram T., 2011, J. Cosmol. Astropart. Phys., 07, 034
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
Castorina E., White M., 2019, J. Cosmol. Astropart. Phys., 06, 025
Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, Phys. Rev. Lett., 100, 091303
Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, Nature, 466, 463
Chapman E. et al., 2012, MNRAS, 423, 2518
Chen S.-F., Castorina E., White M., Slosar A., 2019, J. Cosmol. Astropart. Phys., 07, 023
Chen Z., Wolz L., Spinelli M., Murray S. G., 2021, MNRAS, 502, 5259
Crichton D. et al., 2022, J. Astron. Telesc. Instrum. Syst., 8, 011019
Crighton N. H. et al., 2015, MNRAS, 452, 217
Croton D. J. et al., 2016, ApJS, 222, 22
Cunnington S., Pourtsidou A., Soares P. S., Blake C., Bacon D., 2020, MNRAS, 496, 415
Cunnington S., Irfan M. O., Carucci I. P., Pourtsidou A., Bobin J., 2021, MNRAS, 504, 208
Cunnington S. et al., 2023, MNRAS, 518, 6262
D'Amico G., Gleyzes J., Kokron N., Markovic K., Senatore L., Zhang P., Beutler F., Gil-Marín H., 2020, J. Cosmol. Astropart. Phys., 05, 005
D'Amico G., Senatore L., Zhang P., 2021, J. Cosmol. Astropart. Phys., 01, 006
Donald-McCann J., Beutler F., Koyama K., Karamanis M., 2022, MNRAS, 511, 3768
Donald-McCann J., Koyama K., Beutler F., 2023, MNRAS, 518, 3106
Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Hu W., Wang X., Wu F., Wang Y., Zhang P., Chen X., 2020, MNRAS, 493, 5854
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Hyvärinen A., 1999, IEEE Trans. Neural Netw., 10, 626
Kaiser N., 1987, MNRAS, 227, 1
Karagiannis D., Maartens R., Randrianjanahary L. F., 2022, J. Cosmol. Astropart. Phys., 11, 003
Karamanis M., Beutler F., Peacock J. A., 2021, MNRAS, 508, 3589
Klypin A., Yepes G., Gottlober S., Prada F., Hess S., 2016, MNRAS, 457, 4340
Knebe A. et al., 2018, MNRAS, 474, 5206
Kovetz E. D. et al., 2020, BAAS, 51, 101
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lesgourgues J., 2011, preprint (arXiv:1104.2932)
Lewis A., 2019, preprint (arXiv:1910.13970)
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Li J. et al., 2020, Sci. China Phys. Mech. Astron., 63, 129862
Liu A., Shaw J. R., 2020, PASP, 132, 062001
Liu A., Tegmark M., 2011, Phys. Rev. D, 83, 103006
McQuinn M., D'Aloisio A., 2018, J. Cosmol. Astropart. Phys., 10, 016
Masui K. W. et al., 2013, ApJ, 763, L20
Moresco M. et al., 2022, Living Rev. Relativ., 25, 6
Mueller E.-M., Percival W., Linder E., Alam S., Zhao G.-B., Sánchez A. G., Beutler F., Brinkmann J., 2018, MNRAS, 475, 2122
Newburgh L. et al., 2016, Proc. SPIE Conf. Ser. Vol. 9906, HIRAX: a probe of dark energy and radio transients. SPIE, Bellingham, p. 99065X
Nishimichi T., D'Amico G., Ivanov M. M., Senatore L., Simonović M., Takada M., Zaldarriaga M., Zhang P., 2020, Phys. Rev. D, 102, 123541
Obuljen A., Castorina E., Villaescusa-Navarro F., Viel M., 2018, J. Cosmol. Astropart. Phys., 05, 004
Oh S. P., Mack K. J., 2003, MNRAS, 346, 871
Padmanabhan H., Refregier A., Amara A., 2017, MNRAS, 469, 2323
Perko A., Senatore L., Jennings E., Wechsler R. H., 2016, preprint (arXiv:1610.09321)
Peterson J. B. et al., 2010, Astro2010: The Astronomy and Astrophysics Decadal Survey. NASA, USA, p. 234
Petrovic N., Oh S. P., 2011, MNRAS, 413, 2103
Qin W., Schutz K., Smith A., Garaldi E., Kannan R., Slatyer T. R., Vogelsberger M., 2022, Phys. Rev. D, 106, 123506
Sailer N., Castorina E., Ferraro S., White M., 2021, J. Cosmol. Astropart. Phys., 12, 049
Santos M. G. et al., 2017, in MeerKAT Science: On the Pathway to the SKA. PoS MeerKAT2016, p. 32
Sarkar D., Bharadwaj S., 2019, MNRAS, 487, 5666
Sarkar D., Bharadwaj S., Anathpindika S., 2016, MNRAS, 460, 4310
Seo H.-J., Dodelson S., Marriner J., Mcginnis D., Stebbins A., Stoughton C., Vallinotto A., 2010, ApJ, 721, 164
Shaw J. R., Sigurdson K., Sitwell M., Stebbins A., Pen U.-L., 2015, Phys. Rev. D, 91, 083514
SKA Cosmology Science Working Group et al., 2020, Publ. Astron. Soc. Aust., 37, e007
Slosar A. et al., 2019, BAAS, 51, 53
Soares P. S., Cunnington S., Pourtsidou A., Blake C., 2021, MNRAS, 502, 2549
Soares P. S., Watkinson C. A., Cunnington S., Pourtsidou A., 2022, MNRAS, 510, 5872
Spinelli M., Zoldan A., De Lucia G., Xie L., Viel M., 2020, MNRAS, 493, 5434
Spinelli M., Carucci I. P., Cunnington S., Harper S. E., Irfan M. O., Fonseca J., Pourtsidou A., Wolz L., 2022, MNRAS, 509, 2048
Spurio Mancini A., Piras D., Alsing J., Joachimi B., Hobson M. P., 2022, MNRAS, 511, 1771
Switzer E. R. et al., 2013, MNRAS, 434, L46
Switzer E. R., Chang T.-C., Masui K. W., Pen U.-L., Voytek T. C., 2015, ApJ, 815, 51
Taruya A., Nishimichi T., Saito S., 2010, Phys. Rev. D, 82, 063522
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Van Rossum G., Drake F. L., Jr, 1995, Python Reference Manual. Centrum voor Wiskunde en Informatica Amsterdam
Vanderlinde K. et al., 2019, Canadian Long Range Plan for Astronomy and Astrophysics White Papers, Vol. 2020. Zenodo, p. 28
Villaescusa-Navarro F. et al., 2018, ApJ, 866, 135
Virtanen P. et al., 2020, Nature Methods, 17, 261
Witzemann A., Alonso D., Fonseca J., Santos M. G., 2019, MNRAS, 485, 5519
Wolz L., Abdalla F. B., Blake C., Shaw J. R., Chapman E., Rawlings S., 2014, MNRAS, 441, 3271
Wolz L. et al., 2017, MNRAS, 464, 4938
Wolz L. et al., 2022, MNRAS, 510, 3495
Wu F. et al., 2021, MNRAS, 506, 3455
Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, ApJ, 608, 622
Zheng H. et al., 2017, MNRAS, 464, 3486

APPENDIX A: FITS TO SIMULATIONS

Here, we fit the EFTofLSS model of equation (5) to H I simulations in order to determine fiducial values for our bias and counter terms parameters. The simulations we use have been described in detail in Soares et al. (2022), but we summarize them here for completeness. They are based on the `MULTIDARK-PLANCK` dark matter N-body simulation (Klypin et al. 2016), which follows 3840^3 particles evolving in a box of side $1\,\mathrm{Gpc}\,h^{-1}$. The cosmology is consistent with `PLANCK15` (Ade et al. 2016). From this, the `MULTIDARK-SAGE` catalogue was produced by applying the semi-analytical model `SAGE` (Croton et al. 2016; Knebe et al. 2018). These simulated data products are publicly available in the Skies & Universe[3] web page, and we use the $z = 0.39$ snapshot. The method used for generating H I IM simulations is as follows: Each galaxy in the `MULTIDARK-SAGE` catalogue has an associated cold gas mass, which is used to calculate an H I mass. The H I mass of each galaxy belonging to each voxel is binned together, and then converted into an H I brightness

[3] http://skiesanduniverses.org/

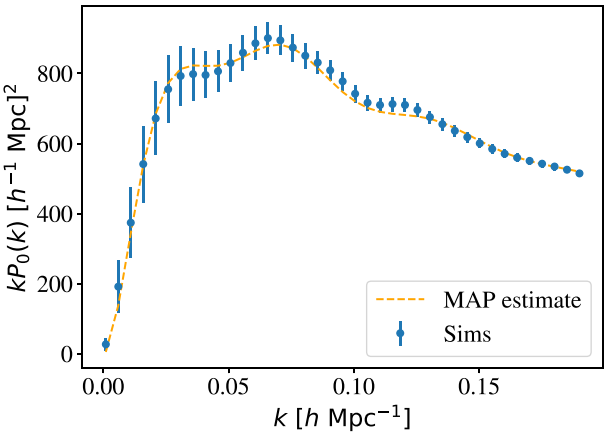

Figure A1. Comparison of the power spectrum monopole (P_0) predictions at the MAP estimate to the simulations measurements.

temperature for that voxel. A crucial limitation comes from the fact that low-mass haloes (lower than $\leq 10^{10} h^{-1}\,\mathrm{M}_\odot$) are not included in the simulation. To account for these missing haloes, which would contribute to the total H I brightness of each voxel, we need to re-scale the mean H I temperature of the simulation to a more realistic value. In the power spectrum measurements, this is an overall amplitude $\bar{T}_{\rm HI}^2$, which we have divided out for the purposes of this fit.

We then proceed to find the maximum a posteriori (MAP) estimate to the power spectrum measurements from these simulations. We plot the result in Fig. A1. The MAP values of the bias and counter terms parameters are as follows:

$$\{b_1, c_2, b_3, c_{\rm ct}, c_{\rm r,1}, c_{\epsilon,1}, c_{\epsilon,\rm quad}\} = \{1.1, 0.6, 0.1, 0.1, -10, 0, -0.8\}.$$

APPENDIX B: FULL POSTERIORS

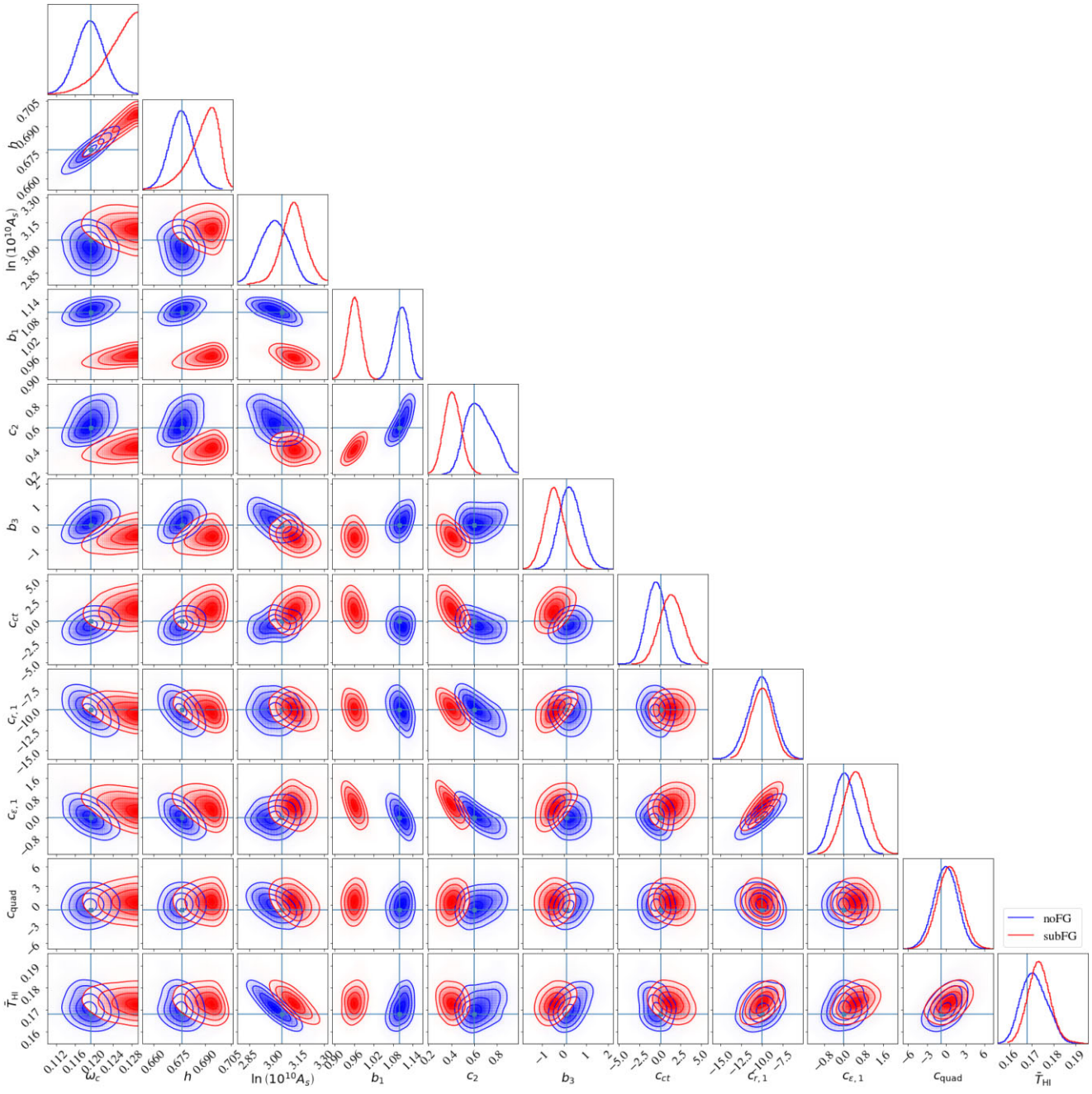

Figure B1. Marginalized 1D and 2D posterior distributions, and 1σ and 2σ contours, for a CHORD-like H I IM survey, including all 11 parameters varied in the MCMC. We show the idealized case (noFG) and a case with 21cm foreground removal effects and a scale cut $k_{\rm min} = 0.03\,h\,{\rm Mpc}^{-1}$ (subFG).

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