

Clustering redshifts with the 21cm-galaxy cross-bispectrum

Caroline Guandalin ¹★, Isabella P. Carucci,^{2,3} David Alonso⁴ and Kavilan Moodley^{5,6}

¹*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, Rua do Matão 1371, CEP 05508-090, São Paulo, Brazil*

²*Dipartimento di Fisica, Università degli Studi di Torino, via P. Giuria 1, I-10125 Torino, Italy*

³*INFN – Istituto Nazionale di Fisica Nucleare, Sezione di Torino, via P. Giuria 1, I-10125 Torino, Italy*

⁴*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

⁵*Astrophysics Research Centre, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa*

⁶*School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa*

Accepted 2022 August 16. Received 2022 August 15; in original form 2021 December 17

ABSTRACT

The cross-correlation between 21-cm intensity mapping (IM) experiments and photometric surveys of galaxies (or any other cosmological tracer with a broad radial kernel) is severely degraded by the loss of long-wavelength radial modes due to Galactic foreground contamination. Higher-order correlators are able to restore some of these modes due to the non-linear coupling between them and the local small-scale clustering induced by gravitational collapse. We explore the possibility of recovering information from the bispectrum between a photometric galaxy sample and an IM experiment, in the context of the clustering-redshifts technique. We demonstrate that the bispectrum is able to calibrate the redshift distribution of the photometric sample to the required accuracy of future experiments such as the Rubin Observatory, using future single-dish and interferometric 21-cm observations, in situations where the two-point function is not able to do so due to foreground contamination. We also show how this calibration is affected by the photometric redshift width $\sigma_{z,0}$ and maximum scale k_{max} . We find that it is important to reach scales $k \gtrsim 0.3 h \text{ Mpc}^{-1}$, with the constraints saturating at around $k \sim 1 h \text{ Mpc}^{-1}$ for next-generation experiments.

Key words: methods: statistical – techniques: miscellaneous – techniques: photometric – cosmology: observations – large-scale structure of Universe – radio lines: general.

1 INTRODUCTION

Galaxy redshift surveys are a cornerstone of contemporary cosmology, with the spectroscopic ones providing highly accurate maps of the three-dimensional galaxy distribution (Percival 2013; DESI Collaboration 2016; Alam et al. 2021). When dealing with discrete dark matter tracers, like galaxies, it is imperative to have a considerable number of observations for the n -point signal to overcome its noise. With the spectroscopic technique, acquiring millions of galaxy spectra is a time consuming process and, even though it is desirable to have high-precision redshifts to reach specific science goals – e.g. baryon acoustic oscillations (Seo & Eisenstein 2007), redshift-space distortions (Kaiser 1987), or Doppler lensing (Bacon et al. 2014) –, interesting physical signatures coming from relativistic and primordial non-Gaussian effects require us to observe the largest scales that are only reachable with a large volume coverage (Dalal et al. 2008; Seljak 2009; Camera et al. 2013; Alonso et al. 2015; Alonso & Ferreira 2015; Schmittfull & Seljak 2018).

Photometric surveys appear as a promising, less expensive route to map the 3D distribution of matter inhomogeneities, covering vast regions of the sky while sampling billions of galaxies in a short amount of time. However, this comes at the expense of large redshift uncertainties, resulting in a limited radial resolution. Specifically, the presence of photometric-redshift (photo- z) errors damp small radial scales, acting similarly to redshift-space distortions (Blake & Bridle 2005). While it may be possible to obtain relatively precise photo- z s for galaxy samples featuring strong spectral features (e.g. the 4000 Å break), this is not the case for the overall galaxy population, resulting in a severe loss of high- k_{\parallel} modes.¹ The presence of large redshift uncertainties impact summary statistics used in cosmological inference and, eventually, degrade cosmological parameter constraints (Ross et al. 2011; Chaves-Montero, Angulo & Hernández-Montegudo 2018).

Fortunately, the presence of different experiments tracing the same large-scale structure (LSS) opens up a range of possibilities to explore combinations of multiple tracers that can complement each other to overcome the specific barriers of each survey. In particular, an avenue to diminish uncertainties in the redshift distribution of a photometric sample, and their impact on cosmological analyses, is to cross-correlate the photometric sample with a spectroscopic one, in what is commonly called the ‘clustering redshifts’ technique (Newman 2008; McQuinn & White 2013). The limiting factor in this case is the overlap in area, redshift, and range of scales probed by both surveys.

* E-mail: c.m.guandalin@qmul.ac.uk

¹In what follows we will label radial (angular) scales with a \parallel (\perp) subscript.

For many LSS analyses, resolving individual galaxies to obtain a detailed map at small scales, as is done spectroscopic galaxy surveys (in the optical or in HI), is not necessary. In the intensity mapping (IM) approach, a map of the LSS is built from the integrated 21-cm emission line of many (unresolved) individual sources at a given angular scale, providing a fast way to map large volumes, up to high redshifts, with spectroscopic precision (Battye, Davies & Weller 2004; McQuinn et al. 2006; Chang et al. 2008; Wyithe & Loeb 2008; Peterson et al. 2009; Bagla, Khandai & Datta 2010; Battye et al. 2013; Masui et al. 2013; Switzer et al. 2013; Bull et al. 2015; Wolz et al. 2021). Therefore, it might be possible, in principle, to make use of IM data to calibrate photometric redshifts (Alonso et al. 2017; Cunnington et al. 2019a).

Although the IM technique overcomes the weakness of the neutral hydrogen signal of an individual galaxy, by integrating the emission from many galaxies and, thus, delivering a larger signal-to-noise, it is prone to foreground contamination that surpasses the strength of the cosmological signal by several orders of magnitude (Santos, Cooray & Knox 2005; Alonso, Ferreira & Santos 2014; Wolz et al. 2014; Cunnington et al. 2019b; Carucci, Irfan & Bobin 2020; Spinelli et al. 2022). Foreground contamination is dominant on the largest radial scales, and its removal renders only the smaller (high k_{\parallel}) modes of the cosmological signal accessible for analysis. This has a severe impact on the cross-correlation between IM and photometric survey data and, therefore, on the possibility of applying the standard clustering redshifts method.

In this paper we will explore a method to overcome this difficulty, by employing higher-order correlators encoding the non-linear coupling between long- and short-wavelength modes induced by gravity. The idea is based on the principle that we can access the lost radial modes in the IM experiment by exploring the impact of the long-wavelength fluctuations on the local HI inhomogeneities. In particular, we will focus on the ability of the bispectrum involving one photometric galaxy sample, which contains the long-wavelength information, and two IM modes to reconstruct the redshift distributions of the photometric sample. In spirit, this is similar to the idea of tidal-mode reconstruction applied to the cross-correlation of 21-cm data with CMB lensing and the projected kinematic Sunyaev-Zel'dovich effect (Li, Zhu & Pen 2019; Hotinli & Johnson 2022).

This paper is structured as follows. The rationale behind this method is presented in Section 2, where we derive expressions for the uncertainty on photo- z parameters achievable through clustering redshifts with the power spectrum and the bispectrum. The survey details and specifications, along with the LSS modelling used, are described in Section 3. Our main results are presented and discussed in Section 4, where we quantify the regime where the bispectrum constraints overcome the power spectrum, as well as the photo- z constraints achievable by different Stage-IV surveys. Finally, conclusions are drawn in Section 5. The relevant mathematical derivations are presented in Appendix A, and some variations of the results presented in Section 4.1 are shown and discussed in Appendix B.

2 CLUSTERING REDSHIFTS WITH 2- AND 3-POINT FUNCTIONS

In the absence of precise redshift information, LSS analyses are carried out tomographically, projecting the data from different galaxy samples (usually lying on a more or less compact range of redshifts) on to the two-dimensional celestial sphere. In the flat-sky approximation (which we use throughout this paper), let $\Delta_g(\mathbf{r}_{\perp})$ be the projected galaxy overdensity in a given redshift bin, related to the three-dimensional overdensity $\delta_g(r_{\parallel}, \mathbf{r}_{\perp})$ via

$$\Delta_g(\mathbf{r}_{\perp}) = \int dr_{\parallel} \phi(r_{\parallel}) \delta_g(r_{\parallel}, \mathbf{r}_{\perp}), \quad (1)$$

where $\phi(r_{\parallel})$ is the radial selection function, which is unknown a priori. The object of the ‘clustering redshifts’ or ‘cross-correlation redshifts’ method is to constrain $\phi(r_{\parallel})$ through the use of a second tracer of the same underlying LSS, but with a well-understood radial distribution. This second tracer is commonly a spectroscopic galaxy sample, although other probes, such as CMB lensing (Alonso et al. 2021; Robertson et al. 2021), may be used to recover certain properties of $\phi(r_{\parallel})$.

The standard implementation of the clustering redshifts approach involves the use of the two-point cross-correlation between the projected overdensity and ‘spectroscopic’ tracer. Treating $\phi(r_{\parallel})$ as a histogram (see equation A2), the cross-correlation of $\Delta_g(\mathbf{r}_{\perp})$ with spectroscopic sources within one of the histogram bins a will be proportional to the number of photometric galaxies in that bin (and zero if there are none). Thus, if the linear biases of the spectroscopic and photometric samples can be measured or calibrated, all histogram bin heights ϕ_a characterizing $\phi(r_{\parallel})$ can be determined from these cross-correlations.

This simple picture becomes murkier when we use 21-cm IM observations as the ‘spectroscopic’ tracer. In this case, although the 21-cm data are able to resolve the smaller radial scales, long-wavelength radial modes are lost due to foreground contamination. Thus, since the projected galaxy overdensity Δ_g is only sensitive to radial scales larger than the support of $\phi(r_{\parallel})$, the range of scales over which both tracers have a significant cross-correlation is severely reduced, degrading significantly the performance of the standard clustering redshifts approach (Alonso et al. 2017; Cunnington et al. 2019a, b; Witzemann et al. 2019). The use of higher-order correlators, exploiting the coupling between small-scale and large-scale modes caused by non-linear gravitational collapse, is a potential workaround to this problem.

The next two subsections present a Fisher matrix-based approach to calculate the uncertainties on $\phi(r_{\parallel})$ achievable through the use of 2-point and 3-point correlators involving a photometric galaxy sample and a set of HI maps. The details of this calculation are described in Appendix A, and we only provide the final results here. For simplicity, and since it will allow us to connect the final constraints with the specific requirements for photo- z calibration in Stage-IV experiments, we will assume that the radial selection function is described by a Gaussian

function centred at distance r_{\parallel}^* with standard deviation σ_{\parallel} :

$$\phi(r_{\parallel}) \propto \exp\left[-\frac{(r_{\parallel} - r_{\parallel}^*)^2}{2\sigma_{\parallel}^2}\right]. \quad (2)$$

2.1 Clustering redshifts with two-point functions

The Fisher matrix for a data vector made up of a set of power spectra between two-dimensional fields $\mathcal{P}^{XY}(k_{\perp})$ is given by:

$$F_{\alpha\beta}^P = \sum_{XX'} \frac{A}{4\pi} \int_0^{\infty} dk_{\perp} k_{\perp} \partial_{\alpha} \mathcal{P}^{XY}(k_{\perp}) \partial_{\beta} \mathcal{P}^{X'Y'}(k_{\perp}) \mathcal{I}^{XX'}(k_{\perp}) \mathcal{I}^{Y'Y'}(k_{\perp}), \quad (3)$$

where A is the comoving survey area, \mathcal{I}^{XY} is the XY element of the inverse matrix of \mathcal{P}^{XY} , $\mathbf{X} = \{X, Y\}$ runs over all the maps used, and ∂_{α} denotes the derivative with respect to a given parameter θ_{α} . In our case, this will correspond to any parameters describing the radial selection function, such as a set of histogram heights, Fourier coefficients or the Gaussian parameters (r_{\parallel}^* , σ_{\parallel}). Once $F_{\alpha\beta}^P$ has been calculated, the covariance of all parameters is simply given by its inverse.

Following the derivation in Appendix A, the variances of r_{\parallel}^* and σ_{\parallel} can be approximated as:

$$\text{Var}_P^{-1}(r_{\parallel}^*) \simeq \frac{A}{2\pi} \int dk_{\perp} k_{\perp} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dk_{\parallel}}{2\pi} k_{\parallel}^2 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})}, \quad (4)$$

$$\text{Var}_P^{-1}(\sigma_{\parallel}) \simeq \frac{A}{2\pi} \int dk_{\perp} k_{\perp} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dk_{\parallel}}{2\pi} \sigma_{\parallel}^2 k_{\parallel}^4 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})}, \quad (5)$$

and their covariance is $\text{Cov}(r_{\parallel}^*, \sigma_{\parallel}) = 0$. Here P_{gg} , P_{gh} , and P_{hh} are the auto- and cross-spectra of the galaxy (g) and HI (h) fields in three dimensions. \mathcal{I}^{TT} is the element of \mathcal{I}^{XY} corresponding to the autocorrelation of the projected photometric sample, and is given by

$$\mathcal{I}^{TT}(\mathbf{k}_{\perp}) = \left\{ \int \frac{dk_{\parallel}}{2\pi} e^{-k_{\parallel}^2 \sigma_{\parallel}^2} P_{gg}(k) \left[1 - \frac{P_{gh}^2(k)}{P_{gg}(k)P_{hh}(k)} \right] \right\}^{-1}, \quad (6)$$

where we use k as a shorthand for $(k_{\parallel}, \mathbf{k}_{\perp})$. Note that, implicitly, all autocorrelations (P_{gg} and P_{hh}) include noise contributions, as described in Section 3.

A number of approximations have been made to reach this result:

(i) We have only considered the Gaussian contribution to the power-spectrum covariance (4-point function). This likely not a bad approximation on the scales used here, and in the presence of instrumental noise.

(ii) We have disregarded all information from the autocorrelation of the projected galaxy overdensity field $\Delta_g(\mathbf{k}_{\perp})$. As discussed in Appendix A2, this is appropriate for r_{\parallel}^* , but the discarded information is, at face value, important for σ_{\parallel} . This is because the amplitude of the projected galaxy autocorrelation is highly sensitive to the width of the redshift distribution (Alonso et al. 2021). In a realistic scenario this sensitivity would be degenerate with the unknown galaxy bias (which we do not marginalise over here). Furthermore, for the purposes of this exercise, we are only interested in the contribution from the 21-cm cross-correlation.

(iii) We have assumed the galaxy shot noise to be described by a Poisson distribution, and that its stochasticity is uncorrelated with possible small-scale fluctuations for the HI samples. Furthermore, because the IM observations are not built from discrete tracers, but from the integrated emission of neutral hydrogen inside galaxies, a Poissonian noise should be negligible if compared with the HI autocorrelation signal. Nonetheless, if the number of HI-emitting galaxies is low, such that the brightness temperature is not enough to deliver a high signal-to-noise, then corrections should be accounted for (Wolz et al. 2016, 2019).

(iv) We have only kept the leading-order terms in $\epsilon = P_{gh}/P_{hh}$. This is exact in the limit of zero mode overlap between the HI maps and the projected galaxy distribution, and would lead to at most $\mathcal{O}(1)$ corrections² in the final result (similar to the effect of other approximations used here). As mentioned in Appendix A2.3 and item (ii) above, the effects of this assumption on the final constraints on r_{\parallel}^* are negligible.

Equation (4) is illustrative of the problem posed by foregrounds to the clustering redshifts method. Focusing on the inner integral over k_{\parallel} , we see that, on the one hand, the exponential prefactor dampens the contribution of radial modes smaller than the photo- z width σ_{\parallel} . On the other hand, the ratio P_{gh}/P_{hh} is only significant on scales smaller than those dominated by foreground contamination. If σ_{\parallel} is larger than the largest foreground-clean scale, both factors conspire to suppress the integrand over the whole k_{\parallel} range. Radio foregrounds are expected to affect 21-cm observations at least on scales $k_{\parallel} \lesssim k_{\parallel}^{\text{FG}} \sim 0.02 h \text{ Mpc}^{-1}$ (Shaw et al. 2015), which coincides with the physical scale corresponding to a photo- z width $\sigma_z \simeq 0.05$ at redshift $z \sim 0.8$. Thus, foreground contamination can present a significant challenge to the clustering-redshifts method using two-point correlators involving 21 cm.

²By $\mathcal{O}(1)$ we mean corrections that could change our results by a factor of ~ 2 . Similarly, the inclusion of off-diagonal contributions in the covariance matrix, at least for the coupling of squeezed bispectrum configurations, have also been shown to degrade primordial non-Gaussian constraints by a factor of 2 (Biagetti et al. 2021).

2.2 Clustering redshifts with three-point functions

The non-linear nature of gravitational collapse leads to a statistical coupling between long- and short-wavelength modes of the density field. This can be understood intuitively: the presence of a long-wavelength fluctuation on a small patch is equivalent to shifting the mean background density, which affects the amplitude and growth of inhomogeneities. Thus the variations in the local power spectrum are statistically coupled to the density field on large scales, leading to a non-vanishing squeezed-limit bispectrum (Bernardeau et al. 2002; Chiang et al. 2014, 2017; Chiang & Slosar 2018a; de Putter 2018). This is the key fact behind the idea of large-scale density reconstruction: by examining the statistical correlations of the small-scale density fluctuations, we can build an estimator of the large-scale overdensity. In the simplest formalism, using only information encoded in the bispectrum, this takes the form of a quadratic estimator in which the long-wavelength overdensity is reconstructed from a quadratic combination of small-scale modes (Jasche & Wandelt 2013; Leclercq et al. 2013; Chiang et al. 2014; Zhu et al. 2016, 2018; Li, Dodelson & Croft 2020). More optimally, this can be done via forward modelling (Schmittfull, Baldauf & Zaldarriaga 2017; Modi et al. 2019, 2021), making use of virtually all N -point correlators.

We could thus reconstruct the long-wavelength fluctuations lost to foreground removal in the HI field by using quadratic combinations of the small-scale HI fluctuations. The resulting large-scale modes could then be correlated with the projected galaxy overdensity, in what would effectively amount to a three-point correlator between two short-scale HI modes and one long-wavelength galaxy overdensity mode, to constrain the properties of the galaxy redshift distribution. Since the constraints imposed by photo- z uncertainties and foreground contamination affect only the radial k_{\parallel} modes, without restricting the transverse modes to any particular range, rather than phrasing our forecasts in terms of this quadratic reconstruction, we will quantify the total amount of information enclosed in the bispectrum.

The equivalent of equation (3) for the bispectra is

$$F_{\alpha\beta}^B = \sum_{XX'} \frac{A}{4\pi} \frac{1}{6} \int_0^\infty dk_{\perp} dq_{\perp} dp_{\perp} \frac{k_{\perp} q_{\perp} p_{\perp}}{\pi^2 A_T} \partial_{\alpha} \mathcal{B}^{XYZ}(k_{\perp}, p_{\perp}, q_{\perp}) \partial_{\beta} \mathcal{B}^{X'Y'Z'}(k_{\perp}, p_{\perp}, q_{\perp}) \mathcal{I}^{XX'}(k_{\perp}) \mathcal{I}^{YY'}(q_{\perp}) \mathcal{I}^{ZZ'}(p_{\perp}), \quad (7)$$

where \mathcal{B}^{XYZ} is the bispectrum between three 2D fields, $A_T = \frac{1}{2} \sqrt{2k_{\perp}^2 q_{\perp}^2 + 2q_{\perp}^2 p_{\perp}^2 + 2p_{\perp}^2 k_{\perp}^2 - k_{\perp}^4 - q_{\perp}^4 - p_{\perp}^4}$ is the area of a triangle with sides $(k_{\perp}, p_{\perp}, q_{\perp})$, and $X = \{X, Y, Z\}$.

For the Gaussian selection function assumed here, as shown in Section A, this reduces to the following expression for the variance of r_{\parallel}^* and σ_{\parallel}^* :

$$\text{Var}_B^{-1}(r_{\parallel}^*) = \frac{1}{6} \frac{A}{4\pi} \int dk_{\perp} dp_{\perp} dq_{\perp} \frac{k_{\perp} p_{\perp} q_{\perp}}{\pi^2 A_T} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dp_{\parallel}}{2\pi} \frac{dk_{\parallel}}{2\pi} k_{\parallel}^2 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{B_{ghh}^2(k_{\parallel}, p_{\parallel}, -k_{\parallel} - p_{\parallel}; k_{\perp}, p_{\perp}, q_{\perp})}{P_{hh}(p_{\parallel}, q_{\perp}) P_{hh}(-k_{\parallel} - p_{\parallel}, p_{\perp})}, \quad (8)$$

$$\text{Var}_B^{-1}(\sigma_{\parallel}^*) = \frac{1}{6} \frac{A}{4\pi} \int dk_{\perp} dp_{\perp} dq_{\perp} \frac{k_{\perp} p_{\perp} q_{\perp}}{\pi^2 A_T} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dp_{\parallel}}{2\pi} \frac{dk_{\parallel}}{2\pi} \sigma_{\parallel}^2 k_{\parallel}^4 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{B_{ghh}^2(k_{\parallel}, p_{\parallel}, -k_{\parallel} - p_{\parallel}; k_{\perp}, p_{\perp}, q_{\perp})}{P_{hh}(p_{\parallel}, q_{\perp}) P_{hh}(-k_{\parallel} - p_{\parallel}, p_{\perp})}, \quad (9)$$

assuming $\text{Cov}(r_{\parallel}^*, \sigma_{\parallel}^*) = 0$. Here $B_{ghh}(k_{\parallel}, q_{\parallel}, p_{\parallel}; k_{\perp}, q_{\perp}, p_{\perp})$ is the bispectrum between the 3D galaxy overdensity $\delta_g(\mathbf{k})$ and the HI field at two different wavenumbers \mathbf{q} and \mathbf{p} .

These equations demonstrate the ability of higher-order correlators to bypass the loss of radial modes to foregrounds affecting the 2-point function. Although, as in the case of the power spectrum, photometric uncertainties limit the range of usable small-scale radial modes of the galaxy overdensity through the exponential damping term, this does not limit the range of small-scale radial HI modes: p_{\parallel} (and therefore $|-k_{\parallel} - p_{\parallel}|$) can take arbitrarily large values, deep into the region unaffected by foreground contamination. In other words, power spectra require k_{\parallel} to be the same for the HI and galaxy overdensity fields, and the lack of overlap between them due to photometric redshift uncertainties and foreground mode loss suppresses the cross-correlation significantly. In turn, bispectra only force the radial mode of the galaxy field to be $k_{\parallel}^g = -k_{\parallel}^{h1} - k_{\parallel}^{h2}$, where $k_{\parallel}^{h1,2}$ are the radial wavenumbers for the two HI legs. While only small k_{\parallel}^g survive due to photo- z uncertainties, the HI modes can be arbitrarily large (thus avoiding the foreground-dominated regime) as long as they satisfy the previous equality.

3 SURVEY MODELLING

3.1 Power spectra and bispectra of biased tracers

For generic biased tracers x , y and z of the dark matter field, the three-dimensional power spectrum and bispectrum are given by

$$P_{xy}(k) = \mathcal{K}_{xy} P_{\text{nl}}(k) + N_{xy} \delta_{xy}, \quad (10)$$

and

$$B_{xyz}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \mathcal{K}_{xyz} B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \mathcal{J}_{xyz}^{112} P(k_1) P(k_2) + \mathcal{J}_{xyz}^{121} P(k_1) P(k_3) + \mathcal{J}_{xyz}^{211} P(k_2) P(k_3), \quad (11)$$

respectively. In the above equations we are omitting the redshift dependence, which is implied in every quantity appearing on both correlators. We introduced the kernels $\mathcal{K}_{xy} \equiv \mathcal{K}_x \mathcal{K}_y$ and $\mathcal{K}_{xyz} \equiv \mathcal{K}_x \mathcal{K}_y \mathcal{K}_z$ to include particularities of each tracer, for example: $\mathcal{K}_g = b_{g,1}$ gives the linear bias for a particular galaxy sample, and $\mathcal{K}_h = b_{h,1} \bar{T}_h$ includes the mean temperature of the 21-cm signal. In equation (10), N_{xy} is the power-spectrum noise which we neglect in the case $x \neq y$, as already mentioned in Section 2.1.

One may argue that, because the HI is tracing galaxies, there is the possibility of overlap between the photometric sample and the galaxies traced by the IM, which could lead to a correlated shot noise. However, even if this is the case, the Fisher matrix formalism that we base

ourselves in is free from such contributions, since the noise in the signal will not depend on the parameters under consideration (i.e. it is independent of the photometric redshifts). Finally, although the \mathcal{I}^{TT} contribution in the inverse covariance weighting of equations (3) and (7) might depend on some correlated noise coming through the cross-correlation P_{gh} [see equation (6)], only scales $k \lesssim 0.1 h \text{ Mpc}^{-1}$ will be affected, where the impact of stochastic noise should be unimportant. At these scales, the most relevant effect comes from foreground suppression, which overcomes other possible sources of noise.

In equation (11), $B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ is the matter bispectrum induced by non-linear gravitational dynamics of the dark matter field. We do not include any noise component in B_{ghh} (although shot noise and instrumental noise are fully taken into account in the Fisher matrix calculation) for two reasons: firstly, any noise contribution would not depend on the Fisher matrix parameters, and thus drops out when taking derivatives with respect to them. Moreover, the autocorrelation of galaxies in the bispectrum will not be considered, and thus shot-noise contributions from galaxies are irrelevant. The 21-cm instrumental noise is also irrelevant at the 3-point function level (Cunnington, Watkinson & Pourtsidou 2021). Finally, $\mathcal{J}_{xyz}^{ijk} \equiv \mathcal{J}_x^i \mathcal{J}_y^j \mathcal{J}_z^k$ accounts for the second-order bias $b_{x,2}$ contributions induced by the non-linear clustering. As an example: $\mathcal{J}_{ghh}^{112} \equiv b_{g,1} b_{h,1} b_{h,2} \bar{T}_h^2$.

In the remaining of this work, we will adopt the fitting formula proposed in Gil-Marín et al. (2012), based on Λ CDM dark-matter-only N -body simulations:

$$B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2^{\text{eff}}(\mathbf{k}_1, \mathbf{k}_2)P_{\text{nl}}(k_1)P_{\text{nl}}(k_2) + \text{perms}, \quad (12)$$

with

$$F_2^{\text{eff}}(\mathbf{k}_i, \mathbf{k}_j) = \frac{5}{7} \tilde{a}(k_i) \tilde{a}(k_j) + \frac{1}{2} \cos(\theta_{ij}) \left(\frac{k_i}{k_j} + \frac{k_j}{k_i} \right) \tilde{b}(k_i) \tilde{b}(k_j) + \frac{2}{7} \cos^2(\theta_{ij}) \tilde{c}(k_i) \tilde{c}(k_j), \quad (13)$$

recovering the tree-level result in the limit of $\tilde{a}, \tilde{b}, \tilde{c} \rightarrow 1$ (valid for large scales)³ and replacing the non-linear matter power spectrum $P_{\text{nl}}(\mathbf{k})$ by the linear one. We also assume a flat Λ CDM model with $h = 0.674$, $n_s = 0.965$, $\sigma_8 = 0.811$, $\Omega_m = 0.315$, and $\Omega_b = 0.049$ (Aghanim et al. 2020). Finally, non-linearities in the bias will be modelled via the phenomenological fit proposed by Lazeyras et al. (2016), which relates b_2 with the linear bias b_1 for dark matter haloes in a flat Λ CDM universe as

$$b_2(b_1) = 0.412 - 2.143 b_1 + 0.929 b_1^2 + 0.008 b_1^3. \quad (14)$$

Note that, although we shall assume that equation (14) holds exactly for both galaxies and neutral hydrogen, its applicability in these cases is uncertain. This is particularly relevant when dealing with galaxies whose formation depends upon several astrophysical processes that may induce a large scatter in the bias relation (Zennaro et al. 2022). This, however, allows us to roughly quantify the impact of non-linear bias on our results.

3.2 Galaxy surveys: LSST

Galaxy surveys are characterized by, among other things, the galaxy counts per unit redshift bin dz and solid angle $d\Omega$, $n_g(z) = dN/dz/d\Omega$. For this, we will assume the common parametrization (Ma, Hu & Huterer 2006):

$$n_g(z) = A z^\alpha \exp \left[- \left(\frac{z}{z_0} \right)^\beta \right]. \quad (15)$$

In our case, the redshift distribution sets the mean number density of galaxies at a given redshift, which determines the shot-noise level in equation (10),

$$N_{gg} = \frac{1}{\bar{n}_g}, \quad (16)$$

after converting equation (15) to the appropriate units.

Our discussion will focus on the Vera Rubin Observatory's Legacy Survey of Space and Time (LSST). In this case, the redshift distribution is well approximated by the parameters $(z_0, \alpha, \beta) = (0.28, 2, 0.9)$. The linear bias will be modelled as (The LSST Dark Energy Science Collaboration 2021; Ballardini, Matthewson & Maartens 2019; Nicola et al. 2020)

$$b_{g,1}(z) = 0.95 D^{-1}(z), \quad (17)$$

where $D(z)$ is the growth factor normalized to $D(z=0) = 1$. The second-order bias is then determined through equation (14).

As stated in Section 2 (equation 2), we will assume Gaussian redshift bins characterized by a mean redshift \bar{z} and standard deviation σ_z , which can be related to the mean comoving distance and its scatter via $r_{\parallel}^* = \chi(\bar{z})$, $\sigma_{\parallel} = c\sigma_z/H(z)$, where $H(z)$ is the expansion rate and $\chi(z)$ is the radial comoving distance to redshift z . We will characterize the evolution of the standard deviation as $\sigma_z = \sigma_{z,0}(1+z)$, and we will consider values for $\sigma_{z,0}$ in the range $[0.02, 0.05]$, with a fiducial value $\sigma_{z,0} = 0.03$. We will explore redshifts $z < 2.5$, and we will fix the survey area to $\Omega_{\text{sky}} = 13\,800 \text{ deg}^2$.

³For more details on these functions, we refer the reader to Gil-Marín et al. (2012).

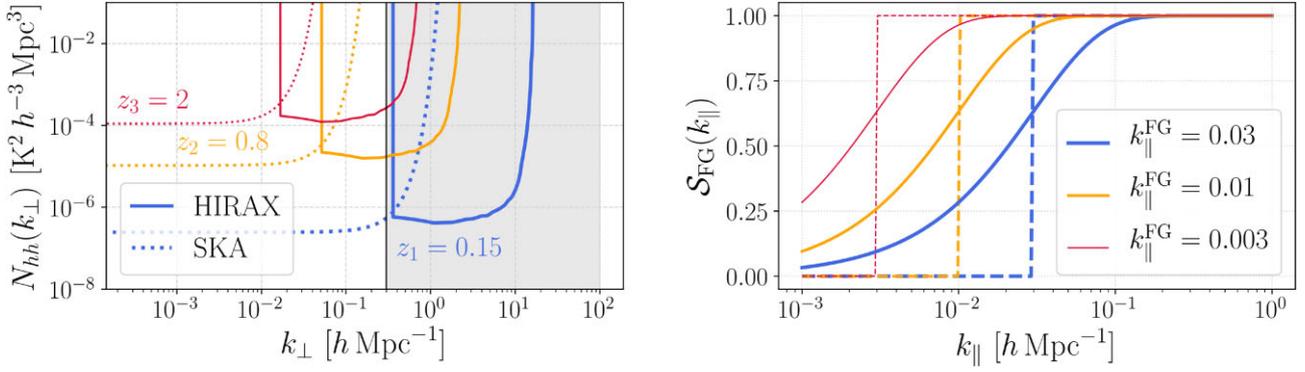


Figure 1. (Left): HIRAX noise power spectrum $N_{hh}(k_{\perp})$ (solid), as described by equation (31), and the SKA noise (equation 29) including beam smoothing [i.e. $N_{hh} S_b^{-2}(k_{\perp})$] (dotted). The beam depicted is for SKA’s single-dish mode, with fixed dish diameter $D_{\text{dish}} = 15$ m. The shaded region corresponds to the modes eliminated by our fiducial $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$ cutoff, but we stress that the $z_1 = 0.15$ will not be covered by HIRAX. Different colours correspond to different redshifts. (Right): Foreground damping function, equation (22). The functions $S_{\text{FG}}^{\text{HS}}$ (dashed) and $S_{\text{FG}}^{\text{SM}}$ (solid) are shown for different foreground damping scales $k_{\parallel}^{\text{FG}}$, with $k_{\parallel}^{\text{min}} = k_{\parallel}^{\text{FG}}$ in the Heaviside model.

3.3 21-cm intensity mapping

3.3.1 The cosmological HI signal

IM of the HI signal maps the fluctuations δT_h of the brightness temperature T_h around its homogeneous component \bar{T}_h . Assuming that HI is a biased tracer of the dark matter field, the temperature fluctuations can be written as (Abdalla & Rawlings 2005):

$$\delta T_h(\mathbf{x}) \equiv \bar{T}_h(z) \delta_h(\mathbf{x}) \equiv \bar{T}_h(z) \left[b_{h,1}(z) \delta(\mathbf{x}) + \frac{b_{h,2}(z)}{2} \delta^2(\mathbf{x}) \right]. \quad (18)$$

The linear biasing scheme $b_{h,1}(z) \equiv b_h(z)$ and the mean brightness temperature \bar{T}_h are taken from Santos et al. (2018), Ballardini et al. (2019):

$$b_h(z) = 1.307 (0.66655 + 0.17765 z + 0.050223 z^2), \quad (19)$$

and

$$\bar{T}_h(z) = (0.055919 + 0.23242 z - 0.024136 z^2) \text{ [mK]}. \quad (20)$$

The second-order bias $b_{h,2}$ is obtained from equation (14) after fixing $b_{h,1}$ to equation (19).

3.3.2 Foreground contamination

Foregrounds originated from astrophysical emissions within and outside our Galaxy have a major impact on the observed temperature fluctuations δT_h^{obs} . The cleaning of contaminated modes removes the cosmological signal on large radial scales, where it is confused with the spectrally smooth foregrounds (Petrovic & Oh 2011; Alonso et al. 2014; Wolz et al. 2014; Cunnington et al. 2019b; Spinelli et al. 2022). This effect can be modelled by introducing a damping function to the observed temperature fluctuation field:

$$\delta T_h^{\text{obs}}(\mathbf{k}) = S_{\text{FG}}(k_{\parallel}) \delta T_h(\mathbf{k}). \quad (21)$$

We will consider two cases based on the functions proposed in Cunnington, Camera & Pourtsidou (2020), Cunnington et al. (2021): (1) a conservative Heaviside (HS) model, in which all modes below a given scale are completely suppressed, and (2) a smooth (SM) model where large radial modes are exponentially suppressed:

$$S_{\text{FG}}^{\text{HS}}(k_{\parallel}) = \begin{cases} 0, & k_{\parallel} < k_{\parallel}^{\text{FG}} \\ 1, & k_{\parallel} > k_{\parallel}^{\text{FG}} \end{cases}, \quad \text{and} \quad S_{\text{FG}}^{\text{SM}}(k_{\parallel}) = 1 - \exp\left[-\frac{k_{\parallel}}{k_{\parallel}^{\text{FG}}}\right]. \quad (22)$$

Above, $k_{\parallel}^{\text{FG}}$ is the characteristic damping scale below which modes are suppressed after foreground removal. $k_{\parallel}^{\text{FG}} = 0$ represents the limit of perfect cleaning. The HS model is the most conservative case, assuming that no information can be retrieved below $k_{\parallel}^{\text{FG}}$. In practice, some the cosmological signal will not be absolutely suppressed on large scales, and the SM model aims to represent this. However, the level of suppression will depend on multiple factors, including foreground complexity, instrumental effects (e.g. polarization leakage or frequency-dependent beams), and choice of cleaning method. It is not obvious that the corresponding transfer function can be sufficiently well characterized in the foreground-dominated regime, and therefore we will use the HS model as the fiducial scenario in our analysis. Both choices are shown in the right-hand panel of Fig. 1 for various values of $k_{\parallel}^{\text{FG}}$.

Another constraint on the modes accessible to the IM survey related to foregrounds comes in the form of the so-called ‘wedge effect’. This is an issue associated with the frequency dependence of the transverse modes induced by the chromatic response of the experiment (Seo &

Hirata 2016). For the case of interferometry, we consider the case where all modes (see Alonso et al. 2017, and references therein)

$$k_{\parallel} < k_{\parallel}^{\text{hor}} \equiv \frac{\chi(z) H(z)}{c(1+z)} k_{\perp}, \quad (23)$$

become inaccessible by the survey, defining the ‘horizon wedge’. It appears from a phase-delay modulation $e^{2\pi i \mathbf{u} \cdot \boldsymbol{\theta}}$ induced by the fringe pattern of the telescope. Here, $\mathbf{u} = \mathbf{d} v/c$ is the baseline vector, with \mathbf{d} the ground distribution vector of antennas, and $\boldsymbol{\theta}$ is the 2-dimensional source location (e.g. see Choudhuri, Bull & Garsden 2021). Because $2\pi \mathbf{u}$ sets the angular resolution, associated with the transverse modes \mathbf{k}_{\perp} , when we cross-correlate the signal received by different antennas, the frequency structure gets shifted and a foreground, in principle smooth in frequency, acquires a chromatic response. Therefore, we can use the model shown in equation (23) to represent the radial modes $k_{\parallel}^{\text{hor}}$ we lose due to this modulation, which depends on the perpendicular scales \mathbf{k}_{\perp} . Nevertheless, note that various systematics could lead to a loss of modes beyond the wedge as defined in equation (23), and a thorough study of these will be necessary to fully quantify the feasibility of the method proposed here in practice.

The second modulation is due to the telescope’s primary beam, affecting both the interferometer and single-dish observations: since the telescope’s beam has a frequency dependence, if we observe a foreground at a fixed sky position $\boldsymbol{\theta}$, its spectrum can be more smoothed if it falls near the side lobes of the beam, or more enhanced by falling near the central lobe’s peak. This effect adds chromaticity to a smooth foreground. For this ‘primary-beam wedge’, the excluded modes will be modelled as

$$k_{\parallel} < k_{\parallel}^{\text{pb}} \equiv \sin\left(\frac{\theta_{\text{FWHM}}}{2}\right) k_{\parallel}^{\text{hor}}, \quad (24)$$

where θ_{FWHM} is the beamwidth, in radians, defined at its full-width-half-maximum (FWHM): $\theta_{\text{FWHM}}(z) = 0.21(1+z) D_{\text{dish}}^{-1}$. The primary beam wedge is therefore a less restrictive version of the horizon wedge.

3.3.3 Single-dish experiments: SKA

We consider SKA1-MID in the single-dish mode.⁴ This experiment consists of $N_{\text{dishes}} = 133$ dishes with $D_{\text{dish}} = 15$ m in diameter. While it is required Bands 1 and 2 to fully observe the redshift range $0 < z < 3$ (Santos et al. 2015), for simplicity we shall consider Band 1 specifications only: it has a sky coverage of $\Omega_{\text{tot}} = 20\,000$ deg², frequency resolution of $\Delta\nu = 15.2$ kHz, and integration time of $t_{\text{tot}} \approx 10\,000$ h (Bacon et al. 2020).

The finite resolution of the telescope’s primary beam removes structure in the temperature fluctuations on small angular scales. The impact of this smoothing can be taken into account by modifying equation (21) as:

$$\delta T_h^{\text{obs}}(k_{\parallel}, \mathbf{k}_{\perp}) = S_b(k_{\perp}) S_{\text{FG}}(k_{\parallel}) \delta T_h(k_{\parallel}, \mathbf{k}_{\perp}), \quad (25)$$

where $S_b(k_{\perp})$ is the beam window function. We will assume a Gaussian beam of the form $S_b(k_{\perp}) = \exp(-(k_{\perp} \sigma_{\perp})^2/2)$, where σ_{\perp} is related to θ_{FWHM} via

$$\sigma_{\perp} = \chi(z) \frac{\theta_{\text{FWHM}}}{2\sqrt{2 \log 2}}. \quad (26)$$

For the IM, shot noise is not a relevant contribution. The noise in equation (10) comes from thermal fluctuations in the instrument, acting as a source of white noise. For the single-dish mode, it is described by a Gaussian field with variance (Pourtsidou, Bacon & Crittenden 2017)

$$\sigma_{\text{pix}}^2 = T_{\text{sys}}^2 \frac{1}{\Delta\nu t_{\text{tot}}} \frac{\Omega_{\text{tot}}}{\Omega_{\text{pix}}} \frac{1}{N_{\text{dishes}} N_{\text{beams}}}, \quad (27)$$

where $\Omega_{\text{pix}} = 1.133 \theta_{\text{FWHM}}^2$ for a Gaussian beam, $T_{\text{sys}} = T_{\text{rx}} + T_{\text{spill}} + T_{\text{CMB}} + T_{\text{gal}}$ encodes the temperature contributions from the ground (spill-over $T_{\text{spill}} \approx 3$ K), from the CMB ($T_{\text{CMB}} \approx 2.7$ K), and from our Galaxy (T_{gal}) and the system’s noise (T_{rx}) (Bacon et al. 2020):

$$T_{\text{gal}} = 25 \left(\frac{408 \text{ MHz}}{\nu} \right)^{2.75} \text{ K}, \quad \text{and} \quad T_{\text{rx}} = 15 \text{ K} + 30 \text{ K} \left(\frac{\nu}{\text{GHz}} - 0.75 \right)^2. \quad (28)$$

In the remaining of this work, we consider $N_{\text{beams}} = 1$.

Therefore, the noise in the autocorrelation of HI maps is

$$N_{hh} = \sigma_{\text{pix}}^2 V_{\text{pix}}, \quad (29)$$

where

$$V_{\text{pix}} = \Omega_{\text{pix}} \int_{z_{\text{min}}}^{z_{\text{max}}} dz \frac{c \chi^2(z)}{H(z)} \quad (30)$$

⁴Note that it is also possible to carry out IM observations with SKA in interferometer mode. Unfortunately, the SKA baseline distribution is not optimised for cosmology, and most scales used here ($k_{\perp} < 0.3 h \text{ Mpc}^{-1}$) are lost in this case. This could be revisited if the approach described here turned out to be applicable with sufficient robustness on small scales.

is the comoving volume of a ‘pixel’. The dotted curves in the left-hand panel of Fig. 1 shows the SKA noise, equation (29), and the beam smoothing effect, i.e. $N_{hh}(k_{\perp}) \equiv N_{hh} \mathcal{S}_b(k_{\perp})$.

3.3.4 Interferometers: HIRAX

The Hydrogen IM and Real-time Analysis eXperiment (HIRAX, Newburgh et al. 2016; Crichton et al. 2022) is a close-packed array of radio antennas in a square 32×32 configuration. Each of the 6-m dishes will take measurements of the radio sky over $15\,000 \text{ deg}^2$ in the frequency band 400–800 MHz, corresponding to the 21-cm line in the range of redshifts $0.8 \lesssim z \lesssim 2.5$. Like SKA, HIRAX is being deployed in the South African Karoo site, which provides an ideal location in terms of RFI contamination, as well as area overlap with LSST. Besides the standard 21-cm cosmological science cases enabled by its large collecting area and field-of-view, the resulting mapping speed will allow HIRAX to produce large catalogues of localised Fast Radio Bursts (Weltman & Walters 2019). Here, we explore its ability to provide calibrated constraints on the redshift distribution of photometric galaxy samples through the clustering redshifts approach using the power spectrum and bispectrum in the presence of realistic foreground contamination.

We model the noise properties of HIRAX following the model described in appendix C of Alonso et al. (2017):

$$N_{hh}(k_{\perp}) = \frac{4\pi f_{\text{sky}} \chi^2(z) (1+z) T_{\text{sys}}^2 \theta_{\text{FWHM}}^2}{H(z) t_{\text{tot}} \lambda_{21}(z) N_d(\mathbf{d} = \mathbf{k}_{\perp} \chi(z) \lambda_{21}(z)/2\pi)}, \quad (31)$$

where $N_d(\mathbf{d})$ is the number density of baselines (i.e. separations between pairs of antennas) with distance \mathbf{d} , $\lambda_{21}(z)$ is the observed 21-cm line wavelength and we assume $T_{\text{sys}} = 50 \text{ K} + 60 (\nu_{21}(z)/300 \text{ MHz})^{-2.5} \text{ K}$. For simplicity we assume an isotropic density of baselines computed from the distribution corresponding to a 32×32 tightly-packed interferometer (i.e. assuming a 6-m separation between antennas). The solid curves in the left-hand panel of Fig. 1 shows the HIRAX noise, equation (31), as a function of the perpendicular modes.

The advantage of HIRAX over SKA is its ability to cover significantly smaller angular scales, in the range of interest for cosmological studies. Whereas the single-dish SKA beam smooths out comoving scales above $k_{\perp}^{\text{SKA}} \simeq 0.045 h \text{ Mpc}^{-1}$ at $z = 1$ ($\theta_{\text{FWHM}} = 1.6^\circ$), HIRAX can recover the 21-cm fluctuations out to $k_{\perp}^{\text{HIRAX}} \simeq 1 h \text{ Mpc}^{-1}$ at the same redshift. The drawback is that interferometers are not able to resolve scales larger than that determined by the minimum available baseline (corresponding to $k_{\perp} \sim 0.037 h \text{ Mpc}^{-1}$ at $z = 1$ for HIRAX). However, since the number of available modes grows steeply with the maximum k_{\perp} , as does the amplitude of the bispectrum as we approach non-linear scales, we can expect a significant improvement over SKA when using HIRAX, particularly at high redshifts.

4 RESULTS

4.1 Mode-loss and mode reconstruction

In order to determine the regime in which long-wavelength reconstruction via the bispectrum is able to outperform the standard cross-correlation redshifts approach, we start by considering an idealised galaxy sample centred at mean redshift $z = 0.8$, for which equation (17) corresponds to a constant linear bias $b \approx 1.4$, and characterized by a Gaussian window function. We will study the cross-correlation of this sample with a HIRAX-like IM experiment (in terms of angular resolution and noise), varying the two characteristic scales that determine the overlap in k space between both data sets: the width of the Gaussian kernel $\sigma_{z,0}$, the foreground cutoff scale $k_{\parallel}^{\text{FG}}$, and the smallest scale used in the analysis k_{max} (for both radial and perpendicular directions). For completeness, in Appendix B we present the same results, but considering an SKA-like survey.

Fig. 2 shows the number of modes available to the cross-correlation between both samples, given by

$$N_{\text{modes}} = \frac{V}{2\pi^2} \int_0^{k_{\text{max}}} dk_{\parallel} \int_0^{k_{\text{max}}} dk_{\perp} k_{\perp} e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{P_{gh}^2(k_{\parallel}, \mathbf{k}_{\perp})}{P_{gg}(k_{\parallel}, \mathbf{k}_{\perp}) P_{hh}(k_{\parallel}, \mathbf{k}_{\perp})}, \quad (32)$$

where V is the survey volume and we use $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. Both P_{gh} and P_{hh} incorporate the foreground window function, and the angular beam, and the figure shows N_{modes} as a function of $\sigma_{z,0}$ and $k_{\parallel}^{\text{FG}}$ (for the Heaviside cutoff) and as a fraction of the total number of available modes (when $k_{\parallel}^{\text{FG}} = 0$, $\sigma_{z,0} = 0$). For realistic values of these characteristic scales ($k_{\parallel}^{\text{FG}} \sim 0.01 h \text{ Mpc}^{-1}$, $\sigma_{z,0} \gtrsim 0.01$) there is a sharp decrease in the fractional number of modes that can be probed through the galaxy-HI cross-power spectrum, reaching ~ 0.5 per cent for the fiducial case used in this paper: $\sigma_{z,0} = 0.03$ (an optimistic estimate of the photometric redshift accuracy achievable by LSST), and $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$ (Shaw et al. 2015). This fiducial case is shown as a black star in the figure.

To quantify to what extent the ghh bispectrum is able to overcome this loss of modes, we study the dependence of the errors on the photo- z parameters (r_{\parallel}^* , σ_{\parallel}) as a function of the characteristic scales ($\sigma_{z,0}$, $k_{\parallel}^{\text{FG}}$, k_{max}). Fig. 3 shows the standard deviation of r_{\parallel}^* (thick and red) and σ_{\parallel} (thin and blue) for the power spectrum (dotted curves) and bispectrum (solid curves) in four different cases. In each case, the scales not varied are fixed to the fiducial values $\sigma_{z,0} = 0.03$, $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$, and $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. The top-right-hand panel shows the dependence on the photo- z width. The errors on the photo- z parameters from the power spectrum grow rapidly, diverging around $\sigma_{z,0} \sim 0.02$, corresponding to a comoving scale $r \sim 68 h^{-1} \text{ Mpc}$ at $z = 0.8$. In turn, the bispectrum errors show a significantly milder dependence on the photo- z scale, outperforming the power spectrum beyond $\sigma_{z,0} \sim 0.02$.

The top-left-hand panel of Fig. 3 shows the dependence of the errors in the photo- z parameters on $k_{\parallel}^{\text{FG}}$. Although the errors derived from the cross-power spectrum diverge for $k_{\parallel}^{\text{FG}} \sim 0.02 h \text{ Mpc}^{-1}$, roughly corresponding with the physical scale of the photo- z scatter, the bispectrum

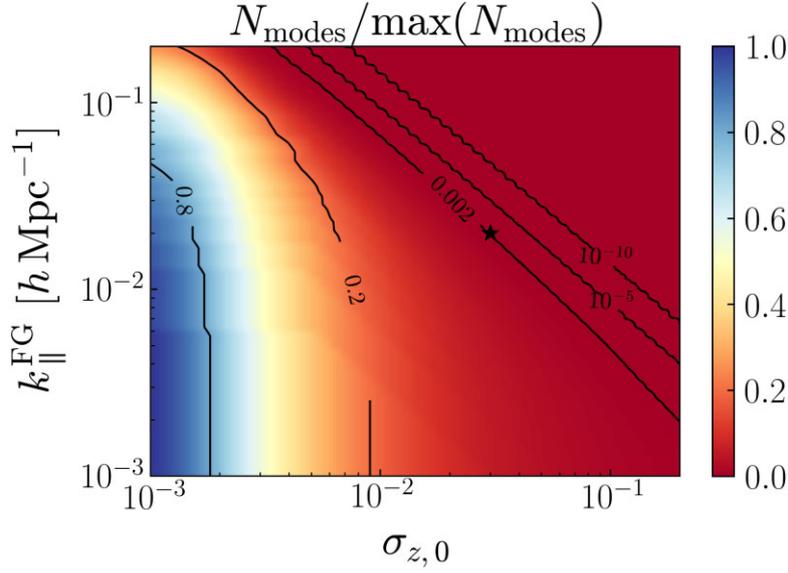


Figure 2. Fractional number of modes available to the galaxy-HI cross-correlation for a HIRAX-like experiment as a function of the photometric redshift error (x -axis) and the foreground cutoff scale (y -axis). The black star marks the fiducial case explored here: $(\sigma_{z,0}, k_{\parallel}^{\text{FG}}) = (0.03, 0.02 \text{ h Mpc}^{-1})$.

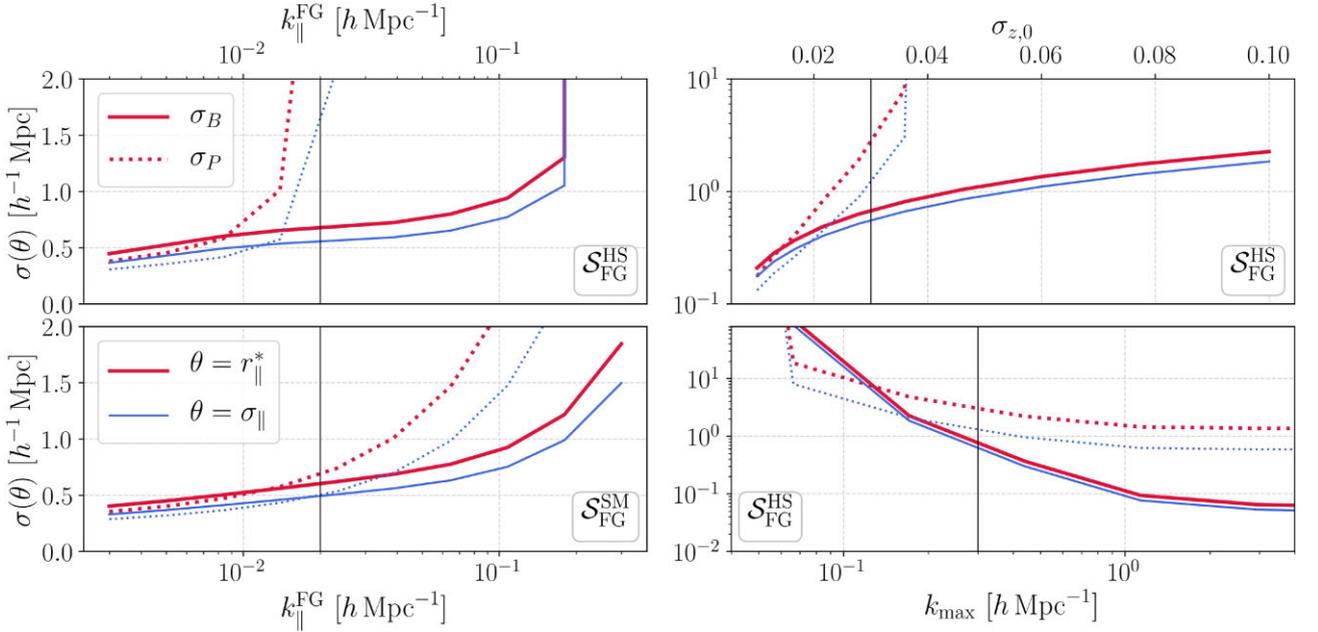


Figure 3. Power-spectrum (*dotted*) and bispectrum (*solid*) constraints obtained by cross-correlating a HIRAX-like (interferometer) survey and a LSST-like, on both Gaussian photo- z parameters, $\theta = r_{\parallel}^*$ (*thick red*), and $\theta = \sigma_{\parallel}$ (*thin blue*), as a function of foreground cutoff/damping scale $k_{\parallel}^{\text{FG}}$ (*left, top and bottom*), photo- z scatter $\sigma_{z,0}$ (*right, top*), and maximum wavenumber k_{max} (*right, bottom*). In the top-left-hand panel, foregrounds are treated via the Heaviside model, $\mathcal{S}_{\text{FG}}^{\text{HS}}(k_{\parallel})$ in equation (22), while the bottom panel shows the same constraints for the smooth exponential damping, $\mathcal{S}_{\text{FG}}^{\text{SM}}(k_{\parallel})$. The plots on the right consider only the $\mathcal{S}_{\text{FG}}^{\text{HS}}(k_{\parallel})$ modelling. A single redshift bin centred at $z = 0.8$ is assumed. Linear bias and number density for the photometric galaxies are, respectively, $b \approx 1.4$ and $\bar{n} \approx 0.004$. We neglect second-order bias contributions and wedge effects in this plot. The vertical lines in each panel show the fiducial values of these scales used in our analysis.

constraints show a much milder growth down to the maximum scale $k_{\text{max}} = 0.3 \text{ h Mpc}^{-1}$. While this panel assumes a conservative binary foreground cutoff (i.e. foregrounds completely eliminate the signal below $k_{\parallel}^{\text{FG}}$), the bottom-left-hand panel of the figure shows the same result for the smooth exponential tapering off. In this case, the bispectrum constraints do not overtake the power spectrum until $k_{\parallel}^{\text{FG}} \sim 0.02 \text{ h Mpc}^{-1}$. This is due to the low noise of both HIRAX and LSST, which always allows us to make use of a sufficiently large number of modes, even if the signal has been suppressed by foreground removal. This, however, assumes that it is possible to model the post-cleaning signal accurately enough on scales $k_{\perp} \ll k_{\parallel}^{\text{FG}}$, which is likely not the case in a realistic scenario. For this reason, we will assume the more conservative binary

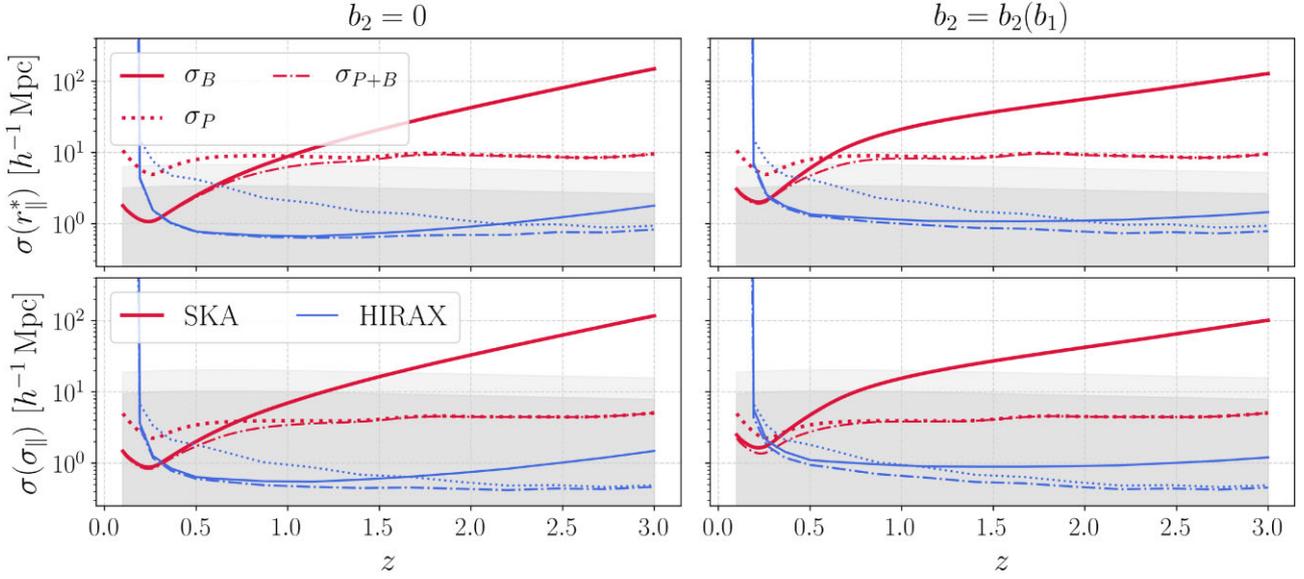


Figure 4. Constraints for the power spectrum (*dotted*), bispectrum (*solid*), and the combination of both statistics (*dot dashed*) obtained with the HIRAX (*thin blue*) and SKA (*thick red*) IM surveys, in the absence of wedge effects. For the latter, we either neglect the second-order bias (*left*), or include it as a function of the linear bias, following equation (14) (*right*). Light-shaded grey regions correspond to the year 1 (Y1) LSST requirements for weak lensing and LSS (3×2 -point) analyses, which are the most stringent ones, whereas darker regions correspond to these requirements for LSST Y10. Characteristic scales are fixed at the fiducial values $\sigma_{z,0} = 0.03$, $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$, and $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. Note that, although we show results for the whole redshift range, HIRAX will only take measurements in the range $0.8 < z < 2.5$.

foreground cutoff model in the rest of the paper. The figure, however, shows that the quantitative results presented here are very sensitive to the assumptions made about the efficiency of the foreground removal procedure.

Finally, the bottom-right-hand panel of Fig. 3 shows the dependence of the photo- z parameter errors on the maximum wavenumber used in the analysis. We observe that, while the power spectrum constraints do not improve significantly beyond $k_{\text{max}} \sim 0.3 h \text{ Mpc}^{-1}$, the bispectrum errors improve steadily out to $k_{\text{max}} \sim 1 h \text{ Mpc}^{-1}$, where they saturate. This is mainly because, as before, nothing limits the smallest radial scale on which the HI fluctuations can be used in the bispectrum, other than the scale at which the 21-cm noise dominates. Besides that, although the cross-power spectrum is only significant out to the smallest transverse scale mapped by the 21-cm experiment, the triangular inequality allows us to use angular scales that are up to twice as small in the case of the bispectrum.

4.2 Forecasts for Stage-IV surveys

Having gained some intuition about the performance of the clustering redshift method using power-spectra and bispectra as a function of photo- z uncertainty and foreground mode-loss, we now proceed to present forecasts for the errors $\sigma(r_{\parallel}^*)$ and $\sigma(\sigma_{\parallel})$ in the case of Stage-IV experiments. In particular, we will consider the cross-correlation of LSST data with SKA (single-dish mode) and HIRAX (interferometer), as a function of redshift. The radial distribution assumed for the LSST galaxies and their linear bias are described in Section 3.2. For the IM surveys, the foreground cutoff scale is fixed at $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$ (Shaw et al. 2015), and we always assume $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$.

To assess the efficiency of different experiments to constrain the redshift distribution parameters, σ_{\parallel} and r_{\parallel}^* , in Fig. 4 we show the forecast results for both the SKA (thick and red) and HIRAX (thin and blue) surveys, in the wedge-free scenario, fixing the photo- z scatter to $\sigma_{z,0} = 0.03$. The dark and light grey bands show the requirement for LSST in Year-1 (Y1) and Year-10 (Y10). No wedge effects were included for this figure. The top and bottom panels show the constraints on the bin's mean r_{\parallel}^* and width σ_{\parallel} . The left and right-hand panels show constraints considering only a linear bias, and including second-order bias terms respectively.

Let us focus on the r_{\parallel}^* constraints (top panels), for which the LSST requirements are significantly tighter. As shown in the figure, due to the loss of radial modes SKA would not be able to achieve the LSST photo- z calibration requirements using two-point correlations. However, utilizing the information available in the 3-point function it would be possible to outperform the power-spectrum results out to $z \sim 0.5$. At higher redshifts two factors conspire to degrade the bispectrum constraints achievable by SKA: the angular resolution decreases significantly, and the amplitude of the matter bispectrum decreases as the density field becomes more linear (see Appendix C for a quantitative discussion). In Fig. 4, we also showcase the constraining power of a joint 2 + 3-point analysis (dash-dotted curves), where the gains of each correlator are combined.

In the case of HIRAX, with access to smaller angular scales, the bispectrum outperforms the power spectrum out to $z \sim 1.5$ –2. The divergence around $z \sim 0.15$ for HIRAX is expected: we ignore all information on scales $k_{\perp} \gtrsim k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$, which at this redshift fall

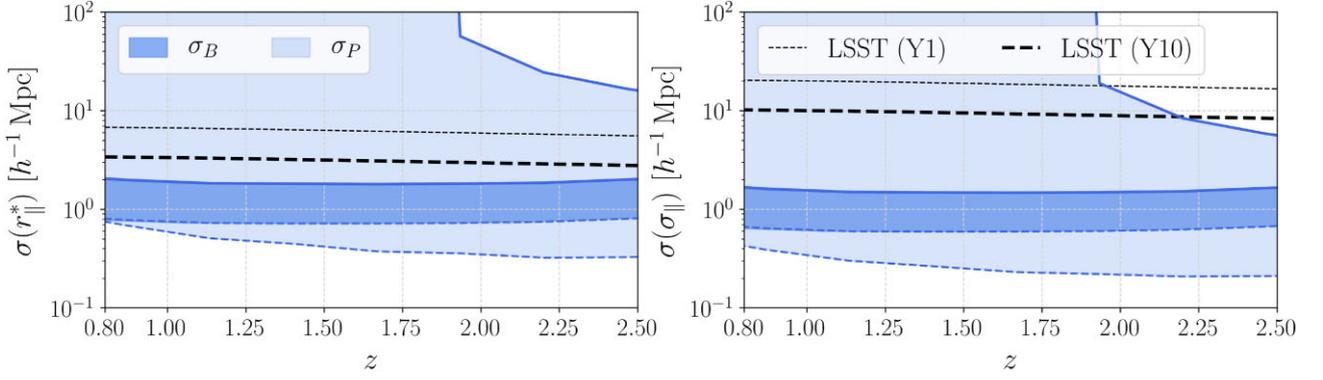


Figure 5. Constraints for the power spectrum (*light*) and bispectrum (*dark*) obtained with a HIRAX-like survey, in the absence of wedge effects, including the second-order bias according to equation (14). We show the regions of constraints extending from an optimistic photo- z scatter ($\sigma_{z,0} = 0.02$, lower limit, *blue-dashed* curves) up to a pessimistic scatter ($\sigma_{z,0} = 0.05$, upper limit, *blue solid*), while keeping the other scales fixed at the fiducial values: $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$, and $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. Black-dashed curves indicate the Y1 (*thin*) and Y10 (*thick*) requirements for the LSST’s 3×2 -point analyses.

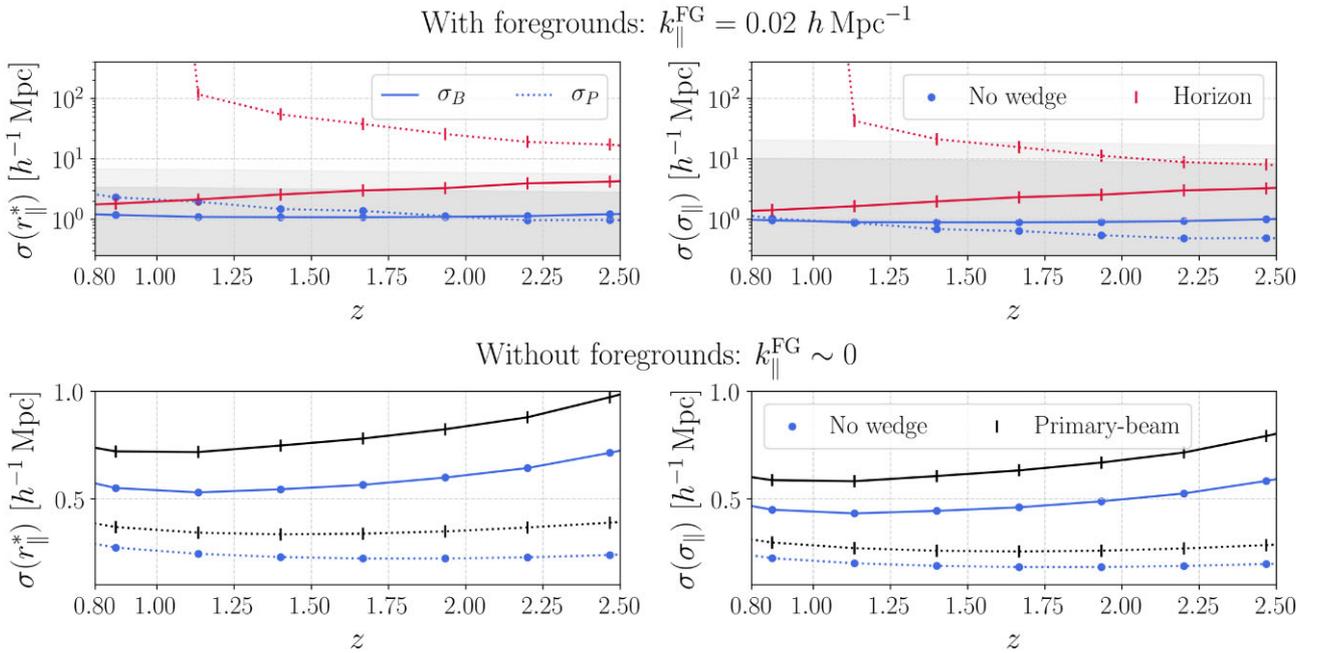


Figure 6. Constraints obtained from the cross-correlations with HIRAX for the power spectrum (*dotted*) and bispectrum (*solid*). The second-order bias correction is included as a function of the linear bias (equation 14). (*Top*): In these plots we explore the effect of horizon wedge (*red bars*) and the wedge-free case (*circles*). Light-shaded regions correspond to the Y1 LSST requirements for the 3×2 -point analyses, whereas darker regions correspond to Y10. Characteristic scales are fixed at the fiducial values $\sigma_{z,0} = 0.03$, $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$, and $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. (*Bottom*): Constraints from HIRAX in the absence of foreground mode loss ($k_{\parallel}^{\text{FG}} \sim 0$), and for $\sigma_{z,0} = 0.03$ and $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. We took the primary-beam wedge (*black bars*) as the only source of radial-mode loss. The wedge-free case (*circles*) is shown for comparison.

below the minimum baseline of the experiment.⁵ Note that the forecasts seem to indicate that, even if better constraints could be achieved with the bispectrum, HIRAX would be able to satisfy the LSST requirements even with the power spectrum. This statement, however, is heavily dependent on the assumed photo- z width.

Fig. 5 shows the constraints on the photo- z parameters achievable by HIRAX for a photo- z width $\sigma_{z,0}$ in the range [0.02,0.05]. Constraints are shown for the power spectrum (light blue) and bispectrum (dark blue). The power-spectrum constraints have a steep dependence on the photo- z width, and achieving the LSST requirement would be impossible if $\sigma_{z,0} \gtrsim 0.05$. Note that this is a perfectly realistic prospect. On the other hand, the bispectrum constraints show a much milder dependence on $\sigma_{z,0}$, and fall within the requirement for all z and $\sigma_{z,0}$.

Let us now turn to the impact of the foreground wedge. For both the single-dish and interferometer experiments, the primary-beam wedge is irrelevant, since the modes affected by it are already suppressed by our choice of $k_{\parallel}^{\text{FG}} = 0.02 h \text{ Mpc}^{-1}$. The impact of the horizon wedge on HIRAX is shown in the top panels of Fig. 6. The presence of the horizon wedge has a strong impact on the power-spectrum constraints, which

⁵Nevertheless, note that HIRAX will only target the redshift range $0.8 \lesssim z \lesssim 2.5$.

are degraded by more than one order of magnitude, falling outside the LSST requirement at all redshifts. The bispectrum constraints, on the other hand, are significantly less sensitive to the wedge, and the Y10 requirements are still achieved at $z \lesssim 2.2$.

For completeness, the bottom panels of Fig. 6 present the forecast for HIRAX considering primary-beam wedge effects as the only source of foreground mode suppression (i.e. $k_{\parallel}^{\text{FG}} \sim 0$), in which case the primary-beam wedge does have a relevant effect. In these panels, the constraints are all within the LSST's Y10 requirements: $\sigma(r_{\parallel}^*) \sim 3 h^{-1} \text{ Mpc}$ and $\sigma(\sigma_{\parallel}) \sim 10 h^{-1} \text{ Mpc}$. We do not show the horizon wedge for this foreground free scenario since the constraints are the same shown in the top panels of Fig. 6, meaning that, for the HIRAX IM survey, the horizon wedge is the leading cause of mode loss.

Before finishing this section, it is worth turning back to the right-hand panels in Fig. 4, showing the constraints accounting for quadratic bias contributions. Although we have shown that the bispectrum can outperform the two-point function for redshift distribution calibration in the presence of foregrounds, the quantitative value of the achieved constraints depends on the non-linear modelling of the galaxy and HI overdensities. Thus the final numbers presented here should only be interpreted as a rough order-of-magnitude estimation of the achievable constraints. In fact, assuming the optimistic value $\sigma_{z,0} = 0.03$, the power spectrum always provides better constraints on the photo- z parameters for a cautious $k_{\text{max}} \sim 0.15 h \text{ Mpc}^{-1}$, for both SKA and HIRAX (in the absence of a horizon wedge). But, as already discussed, the power spectrum for both surveys is unable to reach the required calibration constraints for $\sigma_{z,0} \sim 0.05$, or when a horizon wedge is present. In the extreme scenario, in which $\sigma_{z,0} \approx 0.05$ and there is a horizon wedge, we are unable to reach the desired constraints – even with the bispectrum – unless we go into the mildly non-linear regime ($k \sim 0.3 h \text{ Mpc}^{-1}$).

5 CONCLUSIONS

The loss of long radial modes to foreground removals hampers the ability of cross-correlations between 21-cm intensity maps and photometric redshift surveys to calibrate the redshift distribution of the latter. In this paper we have explored the possibility of overcoming this issue through the use of higher-order correlations, focusing on the case of the bispectrum. The motivation for doing so is based on the idea of mode reconstruction: the statistical coupling between large- and small-scale modes makes it possible to infer the amplitude of a long-wavelength density fluctuation from the local variance of two small-scale modes.

With this in mind, we have quantified the information content of the bispectrum involving a map of the projected galaxy distribution and two HI maps, and propagated it into forecast constraints on redshift distribution parameters from the cross-correlation of LSST with HIRAX and SKA. In particular, we have used a Fisher matrix approach, focusing on two parameters of a simplified Gaussian radial selection function: its mean comoving distance r_{\parallel}^* and standard deviation σ_{\parallel} .

Any constraints based on the two-point cross-correlation are severely hampered whenever the foreground cutoff scale $k_{\parallel}^{\text{FG}}$ approaches or exceeds the scale associated with the redshift scatter $\sim 1/\sigma_{\parallel}$. For typical photo- z uncertainties ($\sigma_z \sim 0.05$), this happens at $k_{\parallel}^{\text{FG}} \gtrsim 0.02 h \text{ Mpc}^{-1}$, comparable with the performance of usual foreground removal methods (Shaw et al. 2015). In this regime, we have shown that the bispectrum is generally able to outperform power-spectrum constraints, and to calibrate the redshift distribution properties below the requirements of LSST.

This is true at low redshifts ($z \lesssim 0.5$) for SKA in single-dish mode, although the constraints degrade significantly at higher redshifts due to the size of the SKA beam. In contrast, a HIRAX-like experiment in interferometer mode would be able to satisfy the LSST requirements throughout its whole redshift range ($0.8 < z < 2.5$) for redshift widths $\sigma_{z,0} \in [0.02, 0.05]$ using the bispectrum, whereas this would be impossible with the two-point function. This result also holds in the presence of a horizon foreground wedge. Whereas the wedge has a devastating effect on the two-point function constraints, the corresponding mode loss is less severe for the bispectrum, which is still able to achieve the LSST requirements. The impact of the primary beam wedge is insignificant in all cases studied except when assuming a very optimistic $k_{\parallel}^{\text{FG}}$.

Our fiducial forecasts used only scales $k < 0.3 h \text{ Mpc}^{-1}$, where it should be possible to characterize the two- and three-point correlations with existing tools. These constraints can therefore be improved by including smaller scales, as long as the impact of non-linearities can be quantified. Nevertheless, we find that, given the expected sensitivity of SKA and HIRAX, the constraining power saturates at $k_{\text{max}} \sim 1 h \text{ Mpc}^{-1}$. As we showed, it is important to use scales slightly inside the non-linear regime ($k_{\text{max}} \sim 0.3 h \text{ Mpc}^{-1}$), as the constraints can be severely degraded if only modes $k \sim 0.15 h \text{ Mpc}^{-1}$ are used. Hence, it is worth exploring the impact of higher-order biases and the non-linear effects, which plague these scales, on the clustering-redshifts method. We point out, however, that a conservative stand that includes only linear scales ($k_{\text{max}} \sim 0.15 h \text{ Mpc}^{-1}$) is not strictly necessary, and in fact traditional applications of the clustering-redshifts approach using 2-point functions have made use of non-linear scales, e.g. down to separations of $\sim 0.5 \text{ Mpc}$ (Chiang, Ménard & Schiminovich 2019), showing robust constraints on the tomographically reconstructed quantities. Therefore, a target of $k_{\text{max}} \sim 0.3 h \text{ Mpc}^{-1}$ is reasonable for clustering-redshifts calibration.

Based on these findings, we therefore recommend the use of the 3-point function to calibrate redshift distributions for future photometric data sets. A combination of interferometer and single-dish observations would likely be necessary in order to cover the whole redshift range out to $z \sim 3$.

These conclusions are subject to a number of caveats. These are summarized as follows:

- (i) We have neglected all non-Gaussian contributions to the bispectrum covariance. However, the off-diagonal correlations induced by the non-Gaussian terms is known to degrade the bispectrum constraints on primordial non-Gaussianity (Biagetti et al. 2021).
- (ii) We have only considered contributions from 2- and 3-point functions involving a single galaxy overdensity map, ignoring the information contained in the galaxy auto-spectrum and any other bispectrum combinations.

- (iii) We have assumed that the galaxy and HI biases are known exactly, including their evolution in redshift. While this may not be true in practice, this problem is inherent in the clustering-redshifts approach, and not specific to 21-cm data.
- (iv) We extrapolated the empirical second-order bias relation obtained for haloes, equation (14), to galaxies and neutral hydrogen, even though we should not expect this relation to hold for luminous tracers. Furthermore, we have not included any tidal, non-local, or higher-order bias terms.⁶
- (v) We have made specific choices when modelling the impact of foreground contamination and photometric redshift uncertainties. Although these have been guided by achieved or forecast performance, the validity of the models used will depend strongly on e.g. complex instrumental effects, or the availability of the photo- z training method used.
- (vi) Our forecasts have made use of the plane-parallel approximation.

In summary, these conclusions must be treated as a proof-of-concept study, and an order-of-magnitude estimation of the information content of the 21cm-galaxy bispectrum.

It is worth comparing these results with those found in Alonso et al. (2017) (A17), who studied the feasibility of clustering redshifts with 21 cm using the standard two-point function approach. In A17, foreground contamination was modelled in terms of a frequency decorrelation parameter ξ . It was found that, even for values of $\xi \sim 0.1$, corresponding to a physical scale similar to the foreground cutoff used here $k_{\parallel}^{\text{FG}} \sim 0.01 h \text{ Mpc}^{-1}$, it would be possible to calibrate redshift distributions to the LSST requirement. This treatment of foregrounds is similar to the SM model of equation (22), which effectively assumes that, although foregrounds dominate on scales below $k_{\parallel}^{\text{FG}}$, the remaining cosmological signal is still accessible. The results found here agree with A17 when neglecting any foreground contamination ($k_{\parallel}^{\text{FG}} = 0$), and also qualitatively when using the SM foreground suppression model, similar in spirit to the decorrelation scale model used in A17. The main difference with respect to A17 is therefore the more conservative treatment of foregrounds used here, discarding all information below $k_{\parallel}^{\text{FG}}$, and the inclusion of the horizon wedge.

Given the potential applicability of this method, the next step should be the development of robust estimators able to reconstruct the galaxy redshift distribution, as has been done for the standard clustering-redshifts approach (McQuinn & White 2013; Morrison et al. 2017), and their calibration against simulations with realistic levels of foreground contamination and photo- z uncertainties. This should also include a thorough exploration of the different systematics affecting 21-cm observations, which could affect significantly the range of modes lost to foregrounds. We leave this for future work.

Finally, given the promising results presented here, it is worth thinking of other potential applications of higher-order correlations between 21-cm IM and photometric redshift surveys, or other projected probes of the LSS (e.g. radio continuum surveys or CMB lensing maps). From a cosmological perspective, the motivation behind both of these probes is the same: overcoming the difficulty of measuring high-quality spectra for billions of galaxy in order to obtain high-sensitivity maps of the density inhomogeneities over large volumes. It would therefore be worth exploring whether these higher-order correlators are well-suited for the study of large-scale effects, such as the impact of primordial non-Gaussianity through its effects on the squeezed-limit bispectrum (de Putter 2018), or reconstructing super-sample density modes in general (Chiang & Slosar 2018b; Li et al. 2020).

ACKNOWLEDGEMENTS

We would like to thank Raul Abramo, Phil Bull, Steven Cunnington, and Anže Slosar for useful comments and discussions. We also thank the anonymous referee for comments that improved the final version of this work. CG is supported by the São Paulo Research Foundation (FAPESP), grant 2018/10396-2. DA acknowledges support from the Beecroft Trust, and the Science and Technology Facilities Council through an Ernest Rutherford Fellowship, grant reference ST/P004474. KM acknowledges support from the National Research Foundation of South Africa. This work made extensive use of the public code `CLASS` (Blas, Lesgourgues & Tram 2011), and the following python packages and libraries: `NUMPY` (Harris et al. 2020), `SCIPY` (Virtanen et al. 2020), `MATPLOTLIB` (Hunter 2007), and `PYCCCL` (Chisari et al. 2019). We thank Eline Maaikje de Weerd for making her SNR code public, on which we based ourselves to compute the summation of equation (C1).

DATA AVAILABILITY

The data and software used in this work are publicly available at <https://github.com/cmguandalin/ClusteringZ-21cm-Bispectrum>.

REFERENCES

- Abdalla F. B., Rawlings S., 2005, *MNRAS*, 360, 27
- Aghanim N. et al., 2020, *A&A*, 641, A6
- Alam S. et al., 2021, *Phys. Rev. D*, 103, 083533
- Alonso D., Ferreira P. G., 2015, *Phys. Rev. D*, 92, 063525
- Alonso D., Ferreira P. G., Santos M. G., 2014, *MNRAS*, 444, 3183

⁶The impact of such biases – and their marginalisation, as mentioned in item (iii) – is an interesting issue that should be explored. Indeed, uncertainties on the details of the non-linear and scale-dependent galaxy biases should also affect the usual clustering-redshifts approach, and a more thorough investigation of their impact, or of the scale cuts needed to avoid them, is left for future work.

- Alonso D., Bull P., Ferreira P. G., Maartens R., Santos M. G., 2015, *ApJ*, 814, 145
- Alonso D., Ferreira P. G., Jarvis M. J., Moodley K., 2017, *Phys. Rev. D*, 96, 043515
- Alonso D., Bellini E., Hale C., Jarvis M. J., Schwarz D. J., 2021, *MNRAS*, 502, 876
- Bacon D. J., Andrianomena S., Clarkson C., Bolejko K., Maartens R., 2014, *MNRAS*, 443, 1900
- Bacon D. J. et al., 2020, *PASA*, 37,e007
- Bagla J. S., Khandai N., Datta K. K., 2010, *MNRAS*, 407, 567
- Ballardini M., Matthewson W. L., Maartens R., 2019, *MNRAS*, 489, 1950
- Battye R. A., Davies R. D., Weller J., 2004, *MNRAS*, 355, 1339
- Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, *MNRAS*, 434, 1239
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Phys. Rep.*, 367, 1
- Biagetti M., Castiblanco L., Noreña J., Sefusatti E., 2021, preprint ([arXiv:2111.05887](https://arxiv.org/abs/2111.05887))
- Blake C., Bridle S., 2005, *MNRAS*, 363, 1329
- Blas D., Lesgourgues J., Tram T., 2011, *J. Cosmol. Astropart. Phys.*, 2011, 034
- Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, *ApJ*, 803, 21
- Camera S., Santos M. G., Ferreira P. G., Ferramacho L., 2013, *Phys. Rev. Lett.*, 111, 171302
- Carucci I. P., Irfan M. O., Bobin J., 2020, *MNRAS*, 499, 304
- Chan K. C., Blot L., 2017, *Phys. Rev. D*, 96, 023528
- Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, *Phys. Rev. Lett.*, 100, 091303
- Chaves-Montero J., Angulo R. E., Hernández-Monteagudo C., 2018, *MNRAS*, 477, 3892
- Chen S.-F., Lee H., Dvorkin C., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 030
- Chiang C.-T., Slosar A., 2018a, *J. Cosmol. Astropart. Phys.*, 2018, 012
- Chiang C.-T., Slosar A., 2018b, *J. Cosmol. Astropart. Phys.*, 2018, 049
- Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, *J. Cosmol. Astropart. Phys.*, 2014, 048
- Chiang C.-T., Cieplak A. M., Schmidt F., Slosar A., 2017, *J. Cosmol. Astropart. Phys.*, 2017, 022
- Chiang Y.-K., Ménard B., Schiminovich D., 2019, *ApJ*, 877, 150
- Chisari N. E. et al., 2019, *ApJS*, 242, 2
- Choudhuri S., Bull P., Garsden H., 2021, *MNRAS*, 506, 2066
- Crichton D. et al., 2022, *JATIS*, 8, 011019
- Cunnington S., Harrison I., Pourtsidou A., Bacon D., 2019a, *MNRAS*, 482, 3341
- Cunnington S., Wolz L., Pourtsidou A., Bacon D., 2019b, *MNRAS*, 488, 5452
- Cunnington S., Camera S., Pourtsidou A., 2020, *MNRAS*, 499, 4054
- Cunnington S., Watkinson C., Pourtsidou A., 2021, *MNRAS*, 507, 1623
- de Putter R., 2018, preprint ([arXiv:1802.06762](https://arxiv.org/abs/1802.06762))
- DESI Collaboration, 2016, preprint ([arXiv:1611.00036](https://arxiv.org/abs/1611.00036))
- Dalal N., Doré O., Huterer D., Shirokov A., 2008, *Phys. Rev. D*, 77, 123514
- Gil-Marín H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, *J. Cosmol. Astropart. Phys.*, 2012, 047
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Hotinli S. C., Johnson M. C., 2022, *Phys. Rev. D*, 105,063522
- Hunter J. D., 2007, *Comput Sci Eng.*, 9, 90
- Jasche J., Wandelt B. D., 2013, *MNRAS*, 432, 894
- Kaiser N., 1987, *MNRAS*, 227, 1
- Lazeyras T., Wagner C., Baldauf T., Schmidt F., 2016, *J. Cosmol. Astropart. Phys.*, 2016, 018
- Leclercq F., Jasche J., Gil-Marín H., Wandelt B., 2013, *J. Cosmol. Astropart. Phys.*, 2013, 048
- Li D., Zhu H.-M., Pen U.-L., 2019, *Phys. Rev. D*, 100, 023517
- Li P., Dodelson S., Croft R. A., 2020, *Phys. Rev. D*, 101, 083510
- Ma Z., Hu W., Huterer D., 2006, *ApJ*, 636, 21
- Maartens R., Jolicoeur S., Umeh O., Weerd E. M. D., Clarkson C., Camera S., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 065
- Masui K. W. et al., 2013, *ApJ*, 763, L20
- McQuinn M., White M., 2013, *MNRAS*, 433, 2857–2883
- McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, *ApJ*, 653, 815
- Modi C., White M., Slosar A., Castorina E., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 023
- Modi C., White M., Castorina E., Slosar A., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 056
- Morrison C. B., Hildebrandt H., Schmidt S. J., Baldry I. K., Bilicki M., Choi A., Erben T., Schneider P., 2017, *MNRAS*, 467, 3576
- Newburgh L. B. et al., 2016, in Hall H. J., Gilmozzi R., Marshall H. K., eds, Proc. SPIE Conf. Ser. Vol. 9906, Ground-based and Airborne Telescopes VI. SPIE, Bellingham, p. 99065X
- Newman J. A., 2008, *ApJ*, 684, 88
- Nicola A. et al., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 044
- Percival W. J., 2013, preprint ([arXiv:1312.5490](https://arxiv.org/abs/1312.5490))
- Peterson J. B. et al., 2009, preprint ([arXiv:0902.3091](https://arxiv.org/abs/0902.3091))
- Petrovic N., Oh S. P., 2011, *MNRAS*, 413, 2103
- Pourtsidou A., Bacon D., Crittenden R., 2017, *MNRAS*, 470, 4251
- Robertson N. C. et al., 2021, *A&A*, 649, A146
- Ross A. J., Percival W. J., Crocce M., Cabré A., Gaztañaga E., 2011, *MNRAS*, 415, 2193
- Santos M. G., Cooray A., Knox L., 2005, *ApJ*, 625, 575
- Santos M. G. et al., 2015, Proc. Sci., Advancing Astrophysics with the Square Kilometre Array (AASKA14). SISSA, Trieste, PoS#019
- Santos M. G. et al., 2018, Proc. Sci., MeerKAT Science: On the Pathway to the SKA. SISSA, Trieste, PoS#032
- Schmittfull M., Seljak U., 2018, *Phys. Rev. D*, 97, 123540
- Schmittfull M., Baldauf T., Zaldarriaga M., 2017, *Phys. Rev. D*, 96, 023505

- Scoccimarro R., Sefusatti E., Zaldarriaga M., 2004, *Phys. Rev. D*, 69, 103513
 Sefusatti E., Crocce M., Pueblas S., Scoccimarro R., 2006, *Phys. Rev. D*, 74, 023522
 Seljak U., 2009, *Phys. Rev. Lett.*, 102, 021302
 Seo H.-J., Eisenstein D. J., 2007, *ApJ*, 665, 14
 Seo H.-J., Hirata C. M., 2016, *MNRAS*, 456, 3142
 Shaw J. R., Sigurdson K., Sitwell M., Stebbins A., Pen U.-L., 2015, *Phys. Rev. D*, 91, 083514
 Spinelli M., Carucci I. P., Cunnington S., Harper S. E., Irfan M. O., Fonseca J., Pourtsidou A., Wolz L., 2022, *MNRAS*, 509, 2048
 Switzer E. R. et al., 2013, *MNRAS*, 434, L46
 The LSST Dark Energy Science Collaboration, 2021, preprint (arXiv:1809.01669)
 Virtanen P. et al., 2020, *Nature Methods*, 17, 261
 Weltman A., Walters A., 2019, preprint (arXiv:1905.07132)
 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., Tonini C., Blake C., Wyithe J., 2016, *MNRAS*, 458, 3399
 Wolz L., Murray S., Blake C., Wyithe J., 2019, *MNRAS*, 484, 1007
 Wolz L. et al., 2021, *MNRAS*, 510, 3495
 Wyithe J. S. B., Loeb A., 2008, *MNRAS*, 383, 606
 Yadav A. P. S., Komatsu E., Wandelt B. D., 2007, *ApJ*, 664, 680
 Zennaro M., Angulo R. E., Contreras S., Pellejero-Ibáñez M., Maion F., 2022, *MNRAS*, 514, 5443
 Zhu H.-M., Pen U.-L., Yu Y., Er X., Chen X., 2016, *Phys. Rev. D*, 93, 103504
 Zhu H.-M., Pen U.-L., Yu Y., Chen X., 2018, *Phys. Rev. D*, 98, 043511

APPENDIX A: CLUSTERING REDSHIFTS FROM 2- AND 3-POINT FUNCTIONS

A1 Preliminaries

We will carry out most of our calculations in Fourier space. We use the following conventions for Fourier transforms in N dimensions:

$$f(\mathbf{k}) = \int d^N \mathbf{r} f(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}, \quad f(\mathbf{r}) = \int \frac{d^N \mathbf{k}}{(2\pi)^N} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (\text{A1})$$

Let us parametrize the selection function as a linear combination of basis functions $W_a(r_{\parallel})$:

$$\phi(r_{\parallel}) = \sum_a \phi_a W_a(r_{\parallel}), \quad (\text{A2})$$

in which case the projected overdensity Δ_g can be written as a sum over ‘slices’ δ_g^a :

$$\Delta_g(\mathbf{r}_{\perp}) = \sum_a \phi_a \delta_g^a(\mathbf{r}_{\perp}), \quad \delta_g^a(\mathbf{r}_{\perp}) \equiv \int dr_{\parallel} W_a(r_{\parallel}) \delta_g(r_{\parallel}, \mathbf{r}_{\perp}). \quad (\text{A3})$$

Our aim is to quantify the accuracy with which the coefficients ϕ_a can be determined through cross-correlations.

There are two obvious classes of basis functions that can be considered. The first is the **real-space top-hat**. Let L be the range of radial comoving distance covered by the selection function. In this case, W_a is a set of N top-hat functions with width $\Delta x \equiv L/N$, centred at $r_{\parallel}^a \equiv (a + 1/2)\Delta x$, with $a \in [0, N)$ and unit area. One can also consider the **Fourier decomposition**. In this second case, the basis functions are characterized by a radial wavenumber k_{\parallel} and are given by:

$$W_{k_{\parallel}}(r_{\parallel}) = \exp(-ik_{\parallel}r_{\parallel}). \quad (\text{A4})$$

The Fourier basis is more manageable as the slices are simply the 3D Fourier coefficients of the galaxy overdensity $\delta_g^{k_{\parallel}}(\mathbf{k}_{\perp}) \equiv \delta_g(k_{\parallel}, \mathbf{k}_{\perp})$, and the selection function coefficients $\phi_{k_{\parallel}}$, in this basis, are related to their radial Fourier transform via

$$\phi_{k_{\parallel}} = \frac{\Delta k_{\parallel}}{2\pi} \phi(k_{\parallel}) = \frac{1}{L} \phi(k_{\parallel}), \quad (\text{A5})$$

where $\Delta k_{\parallel} \equiv 2\pi/L$ is the natural Fourier spacing. Note that we assume throughout that the redshift support of each redshift bin is narrow enough to discard any evolution effects within it.

We will consider 2- and 3-point correlators between the projected fields (Δ_g , which we will label with a slice index T , and the slices δ_h^a):

$$\langle \delta_x^a(\mathbf{k}_{\perp}) \delta_y^b(\mathbf{q}_{\perp}) \rangle \equiv (2\pi)^2 \delta^D(\mathbf{k}_{\perp} + \mathbf{q}_{\perp}) \mathcal{P}_{xy}^{ab}(k_{\perp}), \quad \langle \delta_x^a(\mathbf{k}_{\perp}) \delta_y^b(\mathbf{q}_{\perp}) \delta_z^c(\mathbf{p}_{\perp}) \rangle \equiv (2\pi)^2 \delta^D(\mathbf{k}_{\perp} + \mathbf{q}_{\perp} + \mathbf{p}_{\perp}) \mathcal{B}_{xyz}^{abc}(k_{\perp}, q_{\perp}, p_{\perp}), \quad (\text{A6})$$

where \mathcal{P} and \mathcal{B} are the 2D power spectra and bispectra, respectively, and $x, y, z = \{g, h\}$ will label the type of tracer under consideration. The

3D analogues of these quantities, P and B , are

$$\langle \delta_x(\mathbf{k})\delta_y(\mathbf{q}) \rangle \equiv (2\pi)^3 \delta^D(\mathbf{k} + \mathbf{q}) P_{xy}(k), \quad \langle \delta_x(\mathbf{k})\delta_y(\mathbf{q})\delta_z(\mathbf{p}) \rangle \equiv (2\pi)^3 \delta^D(\mathbf{k} + \mathbf{q} + \mathbf{p}) B_{xyz}(k, q, p), \quad (\text{A7})$$

where $k = |\mathbf{k}|$ is the modulus of the 3D Fourier vector. Using the Fourier radial decomposition, we can relate the different possible projected correlators, equation (A6), with their 3D counterparts:

$$\mathcal{P}_{xy}^{k_{\parallel}q_{\parallel}}(k_{\perp}) = L \delta_{k_{\parallel}+q_{\parallel}}^K P_{xy}(k), \quad \mathcal{P}_{gx}^{Tq_{\parallel}}(k_{\perp}) = \sum_{k_{\parallel}} L \delta_{k_{\parallel}+q_{\parallel}}^K \phi_{k_{\parallel}} P_{gx}(k), \quad \mathcal{P}_{gg}^{TT}(k_{\perp}) = \sum_{k_{\parallel}q_{\parallel}} L \delta_{k_{\parallel}+q_{\parallel}}^K \phi_{k_{\parallel}} \phi_{q_{\parallel}} P_{gg}(k), \quad (\text{A8})$$

$$\text{and} \\ \mathcal{B}_{xyz}^{k_{\parallel}q_{\parallel}p_{\parallel}}(k_{\perp}, q_{\perp}, p_{\perp}) = L \delta_{k_{\parallel}+q_{\parallel}+p_{\parallel}}^K B_{xyz}(k, q, p), \quad \mathcal{B}_{gxy}^{Tq_{\parallel}p_{\parallel}}(k_{\perp}, q_{\perp}, p_{\perp}) = \sum_{k_{\parallel}} L \delta_{k_{\parallel}+q_{\parallel}+p_{\parallel}}^K \phi_{k_{\parallel}} B_{gxy}(k, q, p), \quad (\text{A9})$$

$$\mathcal{B}_{ggx}^{TTp_{\parallel}}(k_{\perp}, q_{\perp}, p_{\perp}) = \sum_{k_{\parallel}q_{\parallel}} L \delta_{k_{\parallel}+q_{\parallel}+p_{\parallel}}^K \phi_{k_{\parallel}} \phi_{q_{\parallel}} B_{ggx}(k, q, p), \quad \mathcal{B}_{ggg}^{TTT}(k_{\perp}, q_{\perp}, p_{\perp}) = \sum_{k_{\parallel}q_{\parallel}p_{\parallel}} L \delta_{k_{\parallel}+q_{\parallel}+p_{\parallel}}^K \phi_{k_{\parallel}} \phi_{q_{\parallel}} \phi_{p_{\parallel}} B_{ggg}(k, q, p). \quad (\text{A9})$$

In these equations δ^D is the Dirac ‘delta’ function, and δ_k^K is the Kronecker ‘delta’ symbol, equal to 1 if $k = 0$ and 0 otherwise.

A2 Fisher matrix

The Fisher matrix for a data vector made up of a set of power spectra $\mathcal{P}^{XY}(k_{\perp})$ is given by:

$$F_{\alpha\beta}^P = \sum_{XX'} \frac{A}{4\pi} \int_0^{\infty} dk_{\perp} k_{\perp} \partial_{\alpha} \mathcal{P}^{XY}(k_{\perp}) \partial_{\beta} \mathcal{P}^{X'Y'}(k_{\perp}) \mathcal{I}^{XX'}(k_{\perp}) \mathcal{I}^{Y'Y'}(k_{\perp}), \quad (\text{A10})$$

where \mathcal{I}^{XY} is the XY element of the inverse matrix of \mathcal{P}^{XY} , and $X = \{X, Y\}$ runs over T and all possible HI slices. Likewise for bispectra (Sefusatti et al. 2006; Yadav, Komatsu & Wandelt 2007; Chen, Lee & Dvorkin 2021):

$$F_{\alpha\beta}^B = \sum_{XX'} \frac{A}{4\pi} \frac{1}{6} \int_0^{\infty} dk_{\perp} dq_{\perp} dp_{\perp} \frac{k_{\perp} q_{\perp} p_{\perp}}{\pi^2 A_T} \partial_{\alpha} \mathcal{B}^{XYZ}(k_{\perp}, p_{\perp}, q_{\perp}) \partial_{\beta} \mathcal{B}^{X'Y'Z'}(k_{\perp}, p_{\perp}, q_{\perp}) \mathcal{I}^{XX'}(k_{\perp}) \mathcal{I}^{Y'Y'}(q_{\perp}) \mathcal{I}^{ZZ'}(p_{\perp}), \quad (\text{A11})$$

where $A_T = \frac{1}{2} \sqrt{2k_{\perp}^2 q_{\perp}^2 + 2q_{\perp}^2 p_{\perp}^2 + 2p_{\perp}^2 k_{\perp}^2 - k_{\perp}^4 - q_{\perp}^4 - p_{\perp}^4}$ is the area of a triangle with sides $(k_{\perp}, p_{\perp}, q_{\perp})$, A is the survey area, and $X = \{X, Y, Z\}$.

In our case, the only parameters we will consider are the selection function coefficients $\phi_{k_{\parallel}}$. Let us denote $\partial_{k_{\parallel}} \equiv \partial/\partial\phi_{k_{\parallel}}$; then, following equations (A8) and (A9), the only non-vanishing derivatives of the different power spectra and bispectra are:

$$\partial_{p_{\parallel}} \mathcal{P}_{gh}^{Tq_{\parallel}}(k_{\perp}) = L \delta_{p_{\parallel}+q_{\parallel}}^K P_{gh}(p_{\parallel}, k_{\perp}),$$

$$\partial_{p_{\parallel}} \mathcal{P}_{gg}^{TT}(k_{\perp}) = 2L \phi_{-p_{\parallel}} P_{gg}(p_{\parallel}, k_{\perp}),$$

$$\partial_{l_{\parallel}} \mathcal{B}_{ghh}^{Tq_{\parallel}p_{\parallel}}(k_{\perp}, p_{\perp}, q_{\perp}) = L \delta_{l_{\parallel}+q_{\parallel}+p_{\parallel}}^K B_{ghh}(l_{\parallel}, q_{\parallel}, p_{\parallel}; \Delta_{\perp}),$$

$$\partial_{l_{\parallel}} \mathcal{B}_{ggh}^{TTp_{\parallel}}(k_{\perp}, p_{\perp}, q_{\perp}) = L \phi_{-l_{\parallel}-p_{\parallel}} [B_{ggh}(l_{\parallel}, -l_{\parallel}-p_{\parallel}, p_{\parallel}; \Delta_{\perp}) + B_{ggh}(-l_{\parallel}-p_{\parallel}, l_{\parallel}, p_{\parallel}; \Delta_{\perp})],$$

$$\partial_{l_{\parallel}} \mathcal{B}_{ggh}^{TTT}(k_{\perp}, p_{\perp}, q_{\perp}) = \sum_{p_{\parallel}} L \phi_{p_{\parallel}} \phi_{-l_{\parallel}-p_{\parallel}} [B_{ggg}(l_{\parallel}, -l_{\parallel}-p_{\parallel}, p_{\parallel}; \Delta_{\perp}) + B_{ggg}(p_{\parallel}, l_{\parallel}, -l_{\parallel}-p_{\parallel}; \Delta_{\perp}) + B_{ggg}(-l_{\parallel}-p_{\parallel}, p_{\parallel}, l_{\parallel}; \Delta_{\perp})],$$

since $\partial_{p_{\parallel}} \mathcal{P}_{hh}^{k_{\parallel}q_{\parallel}}(k_{\perp}) = 0$ and $\partial_{l_{\parallel}} \mathcal{B}_{hhh}^{k_{\parallel}p_{\parallel}q_{\parallel}}(k_{\perp}, p_{\perp}, q_{\perp}) = 0$ as the autocorrelations of HI maps do not depend on $\phi_{k_{\parallel}}$. Above, we defined the short-hand notation $\Delta_{\perp} \equiv (k_{\perp}, p_{\perp}, q_{\perp})$. Finally, the inverse power-spectrum elements are given by:

$$\mathcal{I}^{TT}(k_{\perp}) = \left\{ \sum_{k_{\parallel}} L |\phi_{k_{\parallel}}|^2 P_{gg}(k) \left[1 - \frac{P_{gh}^2(k)}{P_{gg}(k)P_{hh}(k)} \right] \right\}^{-1}, \quad (\text{A12})$$

$$\mathcal{I}^{Tk_{\parallel}}(k_{\perp}) = -\mathcal{I}^{TT}(k_{\perp}) \phi_{k_{\parallel}} \frac{P_{gh}(k)}{P_{hh}(k)}, \quad (\text{A13})$$

$$\mathcal{I}^{k_{\parallel}q_{\parallel}}(k_{\perp}) = \frac{\delta_{k_{\parallel}+q_{\parallel}}^K}{L P_{hh}(k)} + \mathcal{I}^{TT}(k_{\perp}) \phi_{k_{\parallel}} \phi_{q_{\parallel}} \frac{P_{gh}(k_{\parallel}, k_{\perp}) P_{gh}(q_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp}) P_{hh}(q_{\parallel}, k_{\perp})} \simeq \frac{\delta_{k_{\parallel}+q_{\parallel}}^K}{L P_{hh}(k)}. \quad (\text{A14})$$

A2.1 Power spectrum Fisher matrix

The power spectrum Fisher matrix, equation (3), can be decomposed into the following four non-zero contributions:

$$F_{k_{\parallel}q_{\parallel}}^P = F_{k_{\parallel}q_{\parallel}}^{2T2T} + 4F_{(k_{\parallel}q_{\parallel})}^{2T1T} + 2F_{k_{\parallel}q_{\parallel}}^{1T1T+} + 2F_{k_{\parallel}q_{\parallel}}^{1T1T\times}, \quad (\text{A15})$$

where $4F_{(k_{\parallel}q_{\parallel})}^{2T1T} = 2F_{k_{\parallel}q_{\parallel}}^{2T1T} + 2F_{q_{\parallel}k_{\parallel}}^{2T1T}$ and

$$F_{k_{\parallel}q_{\parallel}}^{2T2T} = \frac{A}{4\pi} \int dk_{\perp} k_{\perp} \partial_{k_{\parallel}} \mathcal{P}_{gg}^{TT}(k_{\perp}) \partial_{q_{\parallel}} \mathcal{P}_{gg}^{TT}(k_{\perp}) [\mathcal{I}^{TT}(k_{\perp})]^2 \\ = \frac{A}{\pi} L^2 \phi_{-k_{\parallel}} \phi_{-q_{\parallel}} \int dk_{\perp} k_{\perp} \mathcal{R}_{gg}(k_{\parallel}, k_{\perp}) \mathcal{R}_{gg}(q_{\parallel}, k_{\perp}) \quad (\text{A16})$$

$$\begin{aligned}
 F_{k_{\parallel}q_{\parallel}}^{2T1T} &= \frac{A}{4\pi} \sum_{p_{\parallel}} \int dk_{\perp} k_{\perp} \partial_{k_{\parallel}} \mathcal{P}_{gg}^{TT}(k_{\perp}) \partial_{q_{\parallel}} \mathcal{P}_{gh}^{Tp_{\parallel}}(k_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{Tp_{\parallel}}(k_{\perp}) \\
 &= -\frac{A}{4\pi} 2L^2 \phi_{-k_{\parallel}} \phi_{-q_{\parallel}} \int dk_{\perp} k_{\perp} \mathcal{R}_{gg}(k_{\parallel}, k_{\perp}) \mathcal{R}_{gh}^2(q_{\parallel}, k_{\perp})
 \end{aligned} \tag{A17}$$

$$\begin{aligned}
 F_{k_{\parallel}q_{\parallel}}^{1T1T+} &= \frac{A}{4\pi} \sum_{p_{\parallel}l_{\parallel}} \int dk_{\perp} k_{\perp} \partial_{k_{\parallel}} \mathcal{P}_{gh}^{Tp_{\parallel}}(k_{\perp}) \partial_{q_{\parallel}} \mathcal{P}_{gh}^{Tl_{\parallel}}(k_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{p_{\parallel}l_{\parallel}}(k_{\perp}) \\
 &= \frac{AL}{4\pi} \delta_{k_{\parallel}+q_{\parallel}}^K \int dk_{\perp} k_{\perp} \mathcal{R}_{gh}^2(k_{\parallel}, k_{\perp}) + \frac{A}{4\pi} L^2 \phi_{-k_{\parallel}} \phi_{-q_{\parallel}} \int dk_{\perp} k_{\perp} \mathcal{R}_{gh}^2(k_{\parallel}, k_{\perp}) \mathcal{R}_{gh}^2(q_{\parallel}, k_{\perp}),
 \end{aligned} \tag{A18}$$

$$\begin{aligned}
 F_{k_{\parallel}q_{\parallel}}^{1T1T\times} &= \frac{A}{4\pi} \sum_{p_{\parallel}l_{\parallel}} \int dk_{\perp} k_{\perp} \partial_{k_{\parallel}} \mathcal{P}_{gh}^{Tp_{\parallel}}(k_{\perp}) \partial_{q_{\parallel}} \mathcal{P}_{gh}^{Tl_{\parallel}}(k_{\perp}) \mathcal{I}^{Tp_{\parallel}}(k_{\perp}) \mathcal{I}^{Tl_{\parallel}}(k_{\perp}) \\
 &= \frac{A}{4\pi} L^2 \phi_{-k_{\parallel}} \phi_{-q_{\parallel}} \int dk_{\perp} k_{\perp} \mathcal{R}_{gh}^2(k_{\parallel}, k_{\perp}) \mathcal{R}_{gh}^2(q_{\parallel}, k_{\perp})
 \end{aligned} \tag{A19}$$

where

$$\mathcal{R}_{gg}(k_{\parallel}, k_{\perp}) \equiv P_{gg}(k_{\parallel}, k_{\perp}) \mathcal{I}^{TT}(k_{\perp}), \quad \mathcal{R}_{gh}^2(k_{\parallel}, k_{\perp}) \equiv \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})} \mathcal{I}^{TT}(k_{\perp}). \tag{A20}$$

Putting everything together:

$$\begin{aligned}
 F_{k_{\parallel}q_{\parallel}}^P &= \frac{AL^2}{2\pi} \left[\frac{\delta_{k_{\parallel}+q_{\parallel}}^K}{L} \int dk_{\perp} k_{\perp} \mathcal{R}_{gh}^2(k_{\parallel}, k_{\perp}) + 2\phi_{-k_{\parallel}} \phi_{-q_{\parallel}} \int dk_{\perp} k_{\perp} (\mathcal{R}_{gg}(k_{\parallel}, k_{\perp}) + \mathcal{R}_{gh}^2(k_{\parallel}, k_{\perp})) (\mathcal{R}_{gg}(q_{\parallel}, k_{\perp}) + \mathcal{R}_{gh}^2(q_{\parallel}, k_{\perp})) \right] \\
 &\simeq \delta_{k_{\parallel}+q_{\parallel}}^K \frac{AL}{2\pi} \int dk_{\perp} k_{\perp} \mathcal{I}^{TT}(k_{\perp}) \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})},
 \end{aligned} \tag{A21}$$

where the second equality holds to lowest order in P_{gh}/P_{hh} and neglecting any information coming from $\mathcal{P}_{gg}^{TT}(k_{\perp})$. The latter approximation also allows us to isolate the constraints coming only from the HI cross-correlation, avoiding self-calibration constraints that would be degenerate with e.g. uncertainties on the galaxy bias or the impact of systematic uncertainties in the galaxy autocorrelation.

A2.2 The bispectrum Fisher matrix

The bispectrum Fisher matrix can be decomposed into the following non-zero contributions:

$$F_{k_{\parallel}q_{\parallel}}^B = F_{k_{\parallel}q_{\parallel}}^{3T3T} + 6F_{(k_{\parallel}, q_{\parallel})}^{3T2T} + 6F_{(k_{\parallel}, q_{\parallel})}^{3T1T} + 3F_{k_{\parallel}q_{\parallel}}^{2T2T+} + 6F_{k_{\parallel}q_{\parallel}}^{2T2T\times} + 12F_{(k_{\parallel}, q_{\parallel})}^{2T1T+} + 6F_{(k_{\parallel}, q_{\parallel})}^{2T1T\times} + 3F_{k_{\parallel}q_{\parallel}}^{1T1T+} + 6F_{k_{\parallel}q_{\parallel}}^{1T1T\times}, \tag{A22}$$

where, e.g. $6F_{(k_{\parallel}, q_{\parallel})}^{3T2T} \equiv 3F_{k_{\parallel}, q_{\parallel}}^{3T2T} + 3F_{q_{\parallel}, k_{\parallel}}^{3T2T}$, and:

$$F_{k_{\parallel}q_{\parallel}}^{3T3T} = \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggg}^{TTT}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ggg}^{TTT}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{TT}(q_{\perp}) \mathcal{I}^{TT}(p_{\perp}), \tag{A23}$$

$$F_{k_{\parallel}q_{\parallel}}^{3T2T} = \sum_{p_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggg}^{TTT}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ggh}^{TTp_{\parallel}}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{TT}(q_{\perp}) \mathcal{I}^{Tp_{\parallel}}(p_{\perp}), \tag{A24}$$

$$F_{k_{\parallel}q_{\parallel}}^{3T1T} = \sum_{p_{\parallel}r_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggg}^{TTT}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ghh}^{Tp_{\parallel}r_{\parallel}}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{Tp_{\parallel}}(q_{\perp}) \mathcal{I}^{Tr_{\parallel}}(p_{\perp}), \tag{A25}$$

$$F_{k_{\parallel}q_{\parallel}}^{2T2T+} = \sum_{p_{\parallel}r_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggh}^{TTp_{\parallel}}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ggh}^{TTp_{\parallel}}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{TT}(q_{\perp}) \mathcal{I}^{p_{\parallel}r_{\parallel}}(p_{\perp}), \tag{A26}$$

$$F_{k_{\parallel}q_{\parallel}}^{2T2T\times} = \sum_{p_{\parallel}r_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggh}^{TTp_{\parallel}}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ghg}^{Tr_{\parallel}T}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{Tr_{\parallel}}(q_{\perp}) \mathcal{I}^{Tp_{\parallel}}(p_{\perp}), \tag{A27}$$

$$F_{k_{\parallel}q_{\parallel}}^{2T1T+} = \sum_{p_{\parallel}r_{\parallel}s_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggh}^{TTp_{\parallel}}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ghh}^{Tr_{\parallel}s_{\parallel}}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{Tr_{\parallel}}(q_{\perp}) \mathcal{I}^{p_{\parallel}s_{\parallel}}(p_{\perp}), \tag{A28}$$

$$F_{k_{\parallel}q_{\parallel}}^{2T1T\times} = \sum_{p_{\parallel}r_{\parallel}s_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggh}^{TTp_{\parallel}}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{hhg}^{r_{\parallel}s_{\parallel}T}(\Delta_{\perp}) \mathcal{I}^{Tr_{\parallel}}(k_{\perp}) \mathcal{I}^{Ts_{\parallel}}(q_{\perp}) \mathcal{I}^{Tp_{\parallel}}(p_{\perp}), \tag{A29}$$

$$F_{k_{\parallel}q_{\parallel}}^{1T1T+} = \sum_{p_{\parallel}r_{\parallel}s_{\parallel}t_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggg}^{Tp_{\parallel}r_{\parallel}}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ggg}^{Ts_{\parallel}t_{\parallel}}(\Delta_{\perp}) \mathcal{I}^{TT}(k_{\perp}) \mathcal{I}^{p_{\parallel}s_{\parallel}}(q_{\perp}) \mathcal{I}^{r_{\parallel}t_{\parallel}}(p_{\perp}), \tag{A30}$$

$$F_{k_{\parallel}q_{\parallel}}^{1T1T\times} = \sum_{p_{\parallel}r_{\parallel}s_{\parallel}t_{\parallel}} \int \mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \partial_{k_{\parallel}} B_{ggg}^{Tp_{\parallel}r_{\parallel}}(\Delta_{\perp}) \partial_{q_{\parallel}} B_{ggg}^{s_{\parallel}t_{\parallel}}(\Delta_{\perp}) \mathcal{I}^{Ts_{\parallel}}(k_{\perp}) \mathcal{I}^{Tp_{\parallel}}(q_{\perp}) \mathcal{I}^{r_{\parallel}t_{\parallel}}(p_{\perp}), \tag{A31}$$

with

$$\mathcal{D}(k_{\perp}, q_{\perp}, p_{\perp}) \equiv \frac{A}{4\pi} \frac{1}{6} dk_{\perp} dq_{\perp} dp_{\perp} \frac{k_{\perp} q_{\perp} p_{\perp}}{\pi^2 A_T}. \tag{A32}$$

As before, to lowest order in P_{gh}/P_{hh} , and ignoring any information from B^{TTT} or $B^{TTk_{\parallel}}$ (i.e. keeping only data combinations that are linear in the unknown parameters $\phi_{k_{\parallel}}$), the only non-zero component is $F_{k_{\parallel}q_{\parallel}}^{1T1T+}$ which, in this approximation, becomes:

$$F_{k_{\parallel}q_{\parallel}}^B \simeq \delta_{k_{\parallel}+q_{\parallel}}^K \frac{AL}{4\pi} \frac{1}{6} \int dk_{\perp} dq_{\perp} dp_{\perp} \frac{k_{\perp} q_{\perp} p_{\perp}}{\pi^2 A_T} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dp_{\parallel}}{2\pi} \frac{B_{ghh}^2(k_{\parallel}, p_{\parallel}, -k_{\parallel} - p_{\parallel}, \Delta_{\perp})}{P_{hh}(p_{\parallel}, q_{\perp}) P_{hh}(-k_{\parallel} - p_{\parallel}, p_{\perp})}. \quad (\text{A33})$$

A.2.3 Parametrizing redshift uncertainty

To condense the clustering redshift constraints into a small number of parameters for which the forecast uncertainties can be compared with existing requirements for future photometric redshifts, let us model the redshift distribution $\phi(r_{\parallel})$ as a simple Gaussian centred at r_{\parallel}^* with standard deviation σ_{\parallel} , treating both as free parameters. The Fourier coefficients of this distribution are

$$\phi_{k_{\parallel}} = \frac{1}{L} e^{-\frac{k_{\parallel}^2 \sigma_{\parallel}^2}{2}} e^{-ik_{\parallel} r_{\parallel}^*}. \quad (\text{A34})$$

It is then possible to propagate the Fisher matrix of the redshift distribution amplitudes $\phi_{k_{\parallel}}$ (equations (A21) and (A33) for the power spectrum and bispectrum respectively) into that of the two free parameters r_{\parallel}^* and σ_{\parallel} :

$$F_{\theta_1 \theta_2} = \sum_{k_{\parallel} q_{\parallel}} \partial_{\theta_1} \phi_{k_{\parallel}} \partial_{\theta_2} \phi_{q_{\parallel}} F_{k_{\parallel} q_{\parallel}}. \quad (\text{A35})$$

Doing so, and after taking the continuum limit in all radial wavenumbers (k_{\parallel} etc.) we obtain the following result:

For the power spectrum:

$$\text{Var}_P^{-1}(r_{\parallel}^*) = \frac{A}{2\pi} \int dk_{\perp} k_{\perp} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dk_{\parallel}}{2\pi} k_{\parallel}^2 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})} - C_{r_{\parallel}^*}(k_{\parallel}, q_{\parallel}, k_{\perp}), \quad (\text{A36})$$

$$\text{Var}_P^{-1}(\sigma_{\parallel}) = \frac{A}{2\pi} \int dk_{\perp} k_{\perp} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dk_{\parallel}}{2\pi} \sigma_{\parallel}^2 k_{\parallel}^4 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})} + C_{\sigma_{\parallel}}(k_{\parallel}, q_{\parallel}, k_{\perp}), \quad (\text{A37})$$

where

$$C_{\theta}(k_{\parallel}, q_{\parallel}, k_{\perp}) = \frac{A}{\pi} \int dk_{\perp} k_{\perp} [\mathcal{I}^{TT}(k_{\perp})]^2 \left\{ \int \frac{dk_{\parallel}}{2\pi} \frac{dq_{\parallel}}{2\pi} \mathcal{K}_{\theta}(k_{\parallel}, \sigma_{\parallel}) \mathcal{K}_{\theta}(q_{\parallel}, \sigma_{\parallel}) e^{-\sigma_{\parallel}^2(k_{\parallel}^2 + q_{\parallel}^2)} \mathcal{R}(k_{\parallel}, k_{\perp}) \mathcal{R}(q_{\parallel}, k_{\perp}) \right\}, \quad (\text{A38})$$

with

$$\mathcal{R}(k_{\parallel}, k_{\perp}) \equiv P_{gg}(k_{\parallel}, k_{\perp}) - \frac{P_{gh}^2(k_{\parallel}, k_{\perp})}{P_{hh}(k_{\parallel}, k_{\perp})}, \quad \mathcal{K}_{\theta=r_{\parallel}^*} = k_{\parallel}, \quad \mathcal{K}_{\theta=\sigma_{\parallel}} = \sigma_{\parallel} k_{\parallel}^2. \quad (\text{A39})$$

While these corrections overconstrain $\theta = \sigma_{\parallel}$ for the reasons already mentioned, they are negligible for $\theta = r_{\parallel}^*$. Therefore, we neglect them for the remaining of this work, and employ Equations 4 and 5 for the power-spectrum constraints shown in Section 4 and Appendix B. **For the bispectrum:**

$$\text{Var}_B^{-1}(r_{\parallel}^*) \simeq \frac{1}{6} \frac{A}{4\pi} \int dk_{\perp} dp_{\perp} dq_{\perp} \frac{k_{\perp} p_{\perp} q_{\perp}}{\pi^2 A_T} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dp_{\parallel}}{2\pi} \frac{dk_{\parallel}}{2\pi} k_{\parallel}^2 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{B_{ghh}^2(k_{\parallel}, p_{\parallel}, -k_{\parallel} - p_{\parallel}; k_{\perp}, p_{\perp}, q_{\perp})}{P_{hh}(p_{\parallel}, q_{\perp}) P_{hh}(-k_{\parallel} - p_{\parallel}, p_{\perp})}, \quad (\text{A40})$$

$$\text{Var}_B^{-1}(\sigma_{\parallel}^*) \simeq \frac{1}{6} \frac{A}{4\pi} \int dk_{\perp} dp_{\perp} dq_{\perp} \frac{k_{\perp} p_{\perp} q_{\perp}}{\pi^2 A_T} \mathcal{I}^{TT}(k_{\perp}) \int \frac{dp_{\parallel}}{2\pi} \frac{dk_{\parallel}}{2\pi} \sigma_{\parallel}^2 k_{\parallel}^4 e^{-k_{\parallel}^2 \sigma_{\parallel}^2} \frac{B_{ghh}^2(k_{\parallel}, p_{\parallel}, -k_{\parallel} - p_{\parallel}; k_{\perp}, p_{\perp}, q_{\perp})}{P_{hh}(p_{\parallel}, q_{\perp}) P_{hh}(-k_{\parallel} - p_{\parallel}, p_{\perp})}, \quad (\text{A41})$$

and $\text{Cov}(r_{\parallel}^*, \sigma_{\parallel}) = 0$ in both cases, in agreement with equations (4), (5), (8), and (9). Notice that, in equations (A40) and (A41), we are already neglecting the galaxy autocorrelations and high-order corrections as in equation (A33).

APPENDIX B: ADDITIONAL SKA ANALYSIS

In this appendix we extend the results presented in Section 4.1 for an SKA-like survey (Fig. B1). We consider the same redshift bin, centred at $z = 0.8$, with fixed linear bias $b \approx 1.4$ and number density $\bar{n} \approx 0.004$ for the photometric galaxies, and explore how the cross-correlations depend on the width of the Gaussian kernel $\sigma_{z,0}$, foreground cutoff scale k_{FG} and maximum wavenumber of integration k_{max} .

For the SKA survey, the beam attenuation (as modelled by S_b) is a relevant feature, damping the high- k_{\perp} modes as in Fig. 1. As a consequence, we are unable to access a larger number of perpendicular modes deep inside the non-linear regime, where the bispectrum amplitude is larger, if compared to the interferometer mode. In this case, the bispectrum constraints are degraded and they cannot outperform the power-spectrum method for the smooth foreground removal, as opposed to Fig. 3. For the conservative foreground cutoff model, the qualitative results remain: whilst there is a divergence around $k_{\parallel}^{\text{FG}} \sim 0.02 h \text{ Mpc}^{-1}$ for the power spectrum, with the bispectrum we can push the clustering-redshifts method up to $k_{\parallel}^{\text{FG}} \sim 0.2 h \text{ Mpc}^{-1}$.

When it comes to the effects of the other parameters, the bispectrum is well behaved as a function of the Gaussian width $\sigma_{z,0}$, and we also observe a convergence with respect to the maximum scale of integration, at around $k_{\text{max}} \sim 1 h \text{ Mpc}^{-1}$.

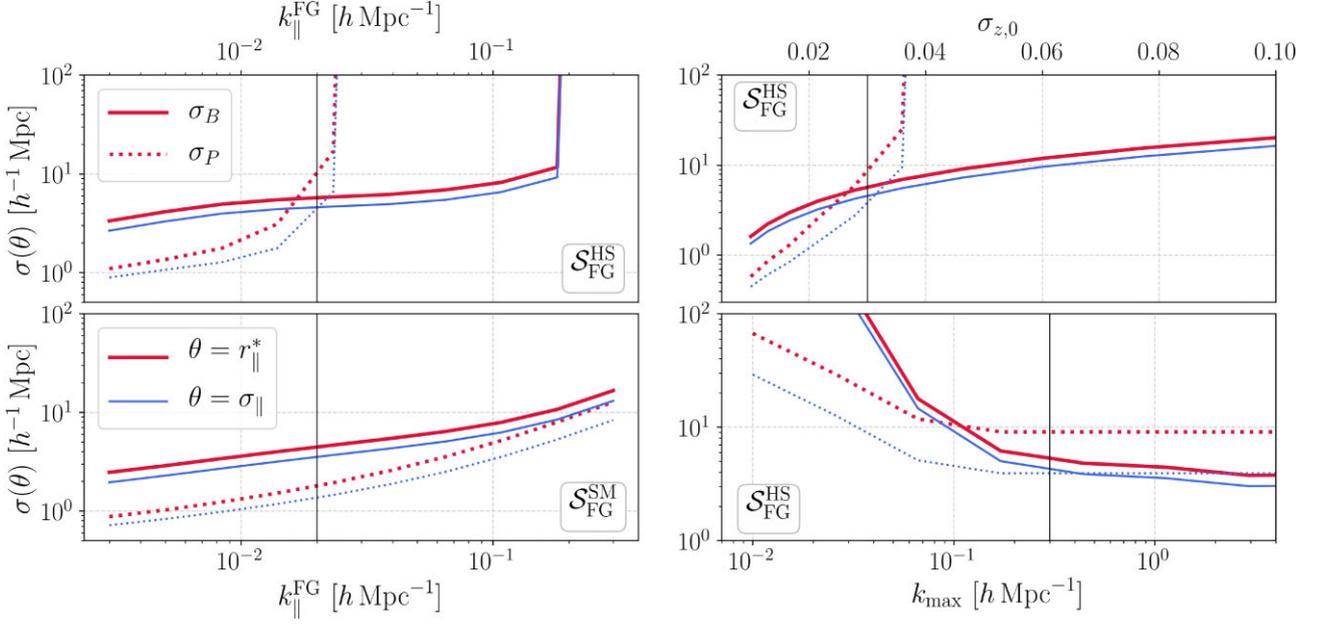


Figure B1. As in Fig. 3, we show the power-spectrum (*dotted*) and bispectrum (*solid*) constraints obtained by cross-correlating an SKA-like (single-dish) survey and a LSST-like, on the two Gaussian photo- z parameters, $\theta = r_{\parallel}^*$ (*thick red*) and $\theta = \sigma_{\parallel}$ (*thin blue*). The dependences on the foreground cutoff/damping scale $k_{\parallel}^{\text{FG}}$ (*left, top, and bottom*), photo- z scatter $\sigma_{z,0}$ (*right, top*), and maximum wavenumber k_{max} (*right, bottom*) are shown assuming no second-order bias correction for the bispectrum. In the bottom-left-hand panel, foregrounds are treated with the smooth exponential damping, $\mathcal{S}_{\text{FG}}^{\text{SM}}(k_{\parallel})$ in equation (22). We take as the fiducial modelling for the other plots the Heaviside model, $\mathcal{S}_{\text{FG}}^{\text{HS}}(k_{\parallel})$, by fixing $k_{\parallel}^{\text{min}} = k_{\parallel}^{\text{FG}}$. Vertical lines mark the fiducial values used in our analyses.

APPENDIX C: SIGNAL-TO-NOISE RATIO

As mentioned in Section 4.2, at higher redshifts the signal from the bispectrum becomes weaker as the density field becomes more Gaussian. It is possible to quantify this effect by looking at the bispectrum signal-to-noise ratio (SNR) as a function of redshift (Scoccimarro, Sefusatti & Zaldarriaga 2004; Maartens et al. 2020):

$$\left[\frac{S}{N}(z)\right]^2 = \sum_{k_1, k_2, k_3} \frac{B^2(z; k_1, k_2, k_3)}{\text{Var}[\hat{B}(z; k_1, k_2, k_3)]}, \quad \text{where} \quad \sum_{k_1, k_2, k_3} = \sum_{k_1=k_{\text{min}}}^{k_{\text{max}}} \sum_{k_2=k_{\text{min}}}^{k_1} \sum_{k_3=k_{\text{min}}}^{k_2}. \quad (\text{C1})$$

Under the Gaussian assumption for the bispectrum covariance (Chan & Blot 2017),

$$\text{Var}[\hat{B}(z; k_1, k_2, k_3)] = \frac{(2\pi)^3 s_{123} P(k_1)P(k_2)P(k_3)}{V 8\pi^2 k_1 k_2 k_3 (\Delta k)^3 \beta(\Delta)}, \quad (\text{C2})$$

where V is the survey volume, $s_{123} = 6, 2, 1$ for equilateral, isosceles and scalene triangles, respectively, and

$$\Delta = \frac{k_3^2 - k_1^2 - k_2^2}{2k_1 k_2}, \quad \beta(\Delta) = \begin{cases} 0.5, & \Delta = \pm 1, \\ 1, & 0 < \Delta < 1. \\ 0, & \text{otherwise} \end{cases} \quad (\text{C3})$$

Therefore:

$$\left[\frac{S}{N}(z)\right]^2 = \frac{V}{\pi} \sum_{k_1, k_2, k_3} k_1 k_2 k_3 (\Delta k)^3 \frac{\beta(\Delta)}{s_{123}} \frac{B^2(z; k_1, k_2, k_3)}{P(z; k_1)P(z; k_2)P(z; k_3)}, \quad (\text{C4})$$

where we took Δk to be the fundamental mode of the survey, i.e. $\Delta k \equiv k_f = 2\pi/V^{1/3}$.

The SNR of the matter bispectrum, equation (C4), is shown in Fig. C1 for the cases in which the survey volume is fixed to $V = 1 h^{-3} \text{ Gpc}^3$, and when the volume is redshift dependent: $V(z) = \frac{4\pi f_{\text{sky}}}{3} (r_+^3 - r_-^3)$, with $r_{\pm} = \chi(\bar{z} \pm \frac{\Delta z}{2})$ (assuming bins of $\Delta z = 0.1$ centred at \bar{z}), and $4\pi f_{\text{sky}} = \Omega_{\text{sky}} = 13800 \times (\pi/180)^2 \text{ rad}^2$, mimicking an LSST-like survey. For the redshift-dependent case (thin-blue curve in Fig. C1), we compute the bispectrum using equation (12), which better represents the effects on non-linearities. For the fixed survey volume, we compute the signal with equation (12) (thick red), and at the tree-level limit (dashed black), i.e. when the functions $\{\tilde{a}, \tilde{b}, \tilde{c}\} \rightarrow 1$.

From the qualitative point of view, there is no known reason to expect an increase in the SNR for realistic situations (presence of foregrounds and instrumental noise for the case of the HI field, or the presence of photo- z errors and shot noise, for the case of galaxies). In fact, the analysis of Cunnington et al. (2021) based on simulations shows that foreground removal reduces the bispectrum's SNR on equilateral configurations by 8 per cent, while the beam reduces it by 62 per cent. Since foreground cleaning removes the largest scales, we expect the

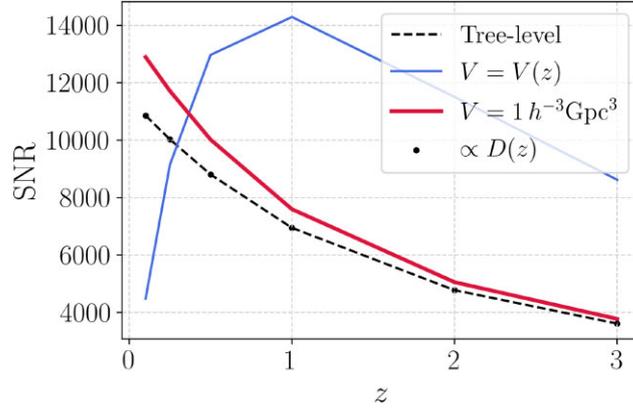


Figure C1. SNR for the matter bispectrum, as given by equation (C4). For the redshift-dependent case (*thin blue*), we model the signal via equation (12). For this case, as the volume for each redshift bin increases ($z < 1$), the SNR becomes larger; however, for $z \geq 1$, the decrease in the bispectrum amplitude is more relevant and causes the SNR to drop. For the fixed survey volume, however, there is a monotonic decrease in the SNR proportional to the growth factor $D(z)$ (*dots*), as $B^2/P^3 \sim D^2(z)$ for the bispectrum induced by gravity at tree-level (*black-dashed curve*), which scales as $P^2(k)$. For the signal given by equation (12) (*thick red*), there are deviations from a pure $D(z)$ scaling, since terms beyond the linear growth must be considered. We can also see that the tree-level limit underestimates the amplitude of the signal, especially at lower redshifts where non-linearities are stronger.

impact on squeezed configurations to be even larger. Finally, the fingers-of-god effect coming from non-linear redshift-space distortions is another source of damping to the bispectrum for both galaxies and HI, reducing the signal even further.

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