

The impact of galaxy bias on cross-correlation tomography

Sara Maleubre¹,¹★ Matteo Zennaro,¹ David Alonso,¹ Ian G McCarthy¹,² Matthieu Schaller¹,^{3,4} and Joop Schaye¹,³

¹*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

²*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*

³*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

⁴*Lorentz Institute for Theoretical Physics, Leiden University, PO Box 9506, NL-2300 RA Leiden, the Netherlands*

Accepted 2025 November 26. Received 2025 October 27; in original form 2025 August 7

ABSTRACT

The cross-correlation of galaxies at different redshifts with other tracers of the large-scale structure can be used to reconstruct the cosmic mean of key physical quantities, and their evolution over billions of years, at high precision. However, a correct interpretation of these measurements must ensure that they are independent of the clustering properties of the galaxy sample used. In this paper, we explore different prescriptions to extract tomographic reconstruction measurements and use the FLAMINGO hydrodynamic simulations to show that a robust estimator, independent of the small-scale galaxy bias, can be constructed. We focus on the tomographic reconstruction of the halo bias-weighted electron pressure $\langle bP_e \rangle$ and star formation density $\langle b\rho_{\text{SFR}} \rangle$, which can be reconstructed from tomographic analysis of Sunyaev–Zel’dovich and cosmic infrared background maps, respectively. We show that these quantities can be reconstructed with an accuracy of 1–3 per cent over a wide range of redshifts, using different galaxy samples. We also show that these measurements can be accurately interpreted using the halo model, assuming that a sufficiently reliable model can be constructed for the halo mass function, large-scale halo bias, and for the dependence of the physical quantities being reconstructed on halo mass.

Key words: methods: numerical – galaxies: formation – large-scale structure of Universe.

1 INTRODUCTION

Galaxies are biased tracers of the underlying fluctuations in the matter density. In general, the abundance of galaxies of a given type at different points in space is strongly correlated with the density of matter and, on sufficiently large scales, both overdensities can be linearly related via

$$\delta_g \simeq b_g \delta_m, \quad (1)$$

where b_g is the linear galaxy bias parameter, and δ_g and δ_m are the galaxy and matter overdensities, respectively. On small scales, however, the galaxy bias relation becomes significantly more complicated and difficult to model. Non-linear, non-local, and stochastic contributions become important, and their impact on galaxy clustering statistics is difficult to predict with high accuracy (V. Desjacques, D. Jeong & F. Schmidt 2018; N. Kokron et al. 2022; M. Zennaro et al. 2022).

Galaxies are also *local* large-scale structure (LSS) tracers: their abundance at a given redshift tracks the matter overdensity, as well as any other physical quantity correlated with it, at the same redshift. This makes it possible to use the cross-correlation of galaxies with different projected tracers of the matter fluctuations to reconstruct the redshift dependence of interesting quantities, in a technique

commonly called ‘tomography’ or ‘tomographic reconstruction’ (J. A. Newman 2008; M. McQuinn & M. White 2013; Y.-K. Chiang, B. Ménard & D. Schiminovich 2019). In broad terms, the large-scale amplitude of the cross-correlation between galaxies and a given LSS tracer U depends on the bias-weighted average of U , $\langle bU \rangle$,¹ and on the clustering properties of the galaxies used in the analysis. The autocorrelation of these galaxies can then be used to constrain the galaxy bias relation, resulting in a measurement of $\langle bU \rangle$.

This technique has been widely used for multiple science purposes. In the ‘cosmic microwave background (CMB) lensing tomography’ approach, cross-correlations with maps of the CMB lensing convergence have been used to reconstruct the growth of structure as a function of time (A. Krolewski, S. Ferraro & M. White 2021; M. White et al. 2022; G. Piccirilli et al. 2024; F. J. Qu et al. 2025). Cross-correlations with maps of the thermal Sunyaev–Zel’dovich (tSZ) effect have been used to reconstruct the bias-weighted mean electron pressure $\langle bP_e \rangle$, which can be used to study the thermodynamics of the intergalactic medium, as well as the amplitude of matter fluctuations (V. Vikram, A. Lidz & B. Jain 2017; S. Pandey et al. 2019; Y.-K. Chiang et al. 2020; N. Koukoufilippas et al. 2020; Z. Chen, P. Zhang & X. Yang 2023; J. Sánchez et al. 2023). The same approach

¹Note that this ‘bias weighting’ refers to the bias of the dark matter haloes populated by U , and not to the bias of the galaxies used in the cross-correlation.

* E-mail: sara.maleubremolinero@physics.ox.ac.uk

can also be used to measure the bias-weighted star formation rate (SFR) density ($b\rho_{\text{SFR}}$), as shown by Z. Yan et al. (2022), B. Jago et al. (2023b), Z. Yan, A. S. Maniyar & L. van Waerbeke (2024), and Y.-K. Chiang, R. Makiya & B. Ménard (2025). Cross-correlation with intensity maps at different wavelengths has also been used to reconstruct other key physical quantities, including the radio, ultraviolet, and gamma-ray background (Y.-K. Chiang et al. 2019; A. Paopiamsap et al. 2024; E. Todarello et al. 2024). Finally, the same principle is behind the so-called clustering redshifts approach: cross-correlations between galaxy samples with known redshifts, and samples with uncertain redshifts can be used to reconstruct the redshift distribution of the latter (J. A. Newman 2008; D. J. Matthews & J. A. Newman 2010; M. McQuinn & M. White 2013; B. Ménard et al. 2013; S. J. Schmidt et al. 2013).

In spite of the broad applicability of tomographic reconstruction, and its ability to recover model-independent measurements of interesting physical quantities, two outstanding questions have limited its use in cosmological analysis:

- (i) Are the estimated values of $\langle bU \rangle$ unbiased, and independent of the clustering properties of the galaxy sample used to measure it?
- (ii) Can the measured values of $\langle bU \rangle$ be interpreted from more fundamental descriptions (e.g. through the halo model)?

Answering these questions is important for several reasons. First, past efforts at tomographic reconstruction have often made use of relatively small scales in order to boost the sensitivity of the resulting measurements (Y.-K. Chiang et al. 2025) [this is particularly common in the case of clustering redshifts (Y.-K. Chiang et al. 2020; M. Gatti et al. 2022)]. However, on these scales, galaxy bias can be significantly scale-dependent, affecting the shape of galaxy auto- and cross-correlations in different ways, and potentially biasing the recovered quantities. Secondly, an unbiased measurement of e.g. $\langle bP_e \rangle$ is not immediately useful unless theoretical predictions can be built for it in terms of both cosmology and gas thermodynamics. Although this has been addressed in e.g. S. Young, E. Komatsu & K. Dolag (2021) in the particular case of tSZ tomography, a more general study has not been carried out yet.

This paper addresses these two questions, finding that they can both be answered in the affirmative assuming that suitable estimators and scale ranges are used. To do so, we will select different galaxy samples within the FLAMINGO suite of hydrodynamical simulations, taking advantage of their large simulation volume to probe the realistic correlation between matter, galaxies, and different baryonic properties over a wide range of linear and non-linear scales. We will apply a tomographic reconstruction estimator to all of these samples, using their autocorrelation and cross-correlation with various LSS fields U . In particular, we will target the matter overdensity, the thermal gas pressure, and the SFR density. We will then assess the ability of the estimator to recover consistent and unbiased measurements of $\langle bU \rangle$, and the possibility of interpreting the measurements in the context of the halo model.

This paper is structured as follows. Section 2 introduces the theoretical background of tomographic reconstruction. The simulations used, including the different galaxy samples and LSS quantities extracted from them, are described in Section 3. Our results are presented in Section 4, including an assessment of the robustness of $\langle bU \rangle$ measurements to small-scale galaxy bias, and their theoretical interpretation. Finally, Section 5 summarizes our findings.

2 TOMOGRAPHY AND GALAXY BIAS

2.1 Cross-correlation-based tomography

Consider a projected field $u(\hat{n})$ defined on the celestial sphere, where \hat{n} is a directional unit vector, tracing a three-dimensional (3D) physical quantity $U(\mathbf{x}, z)$ ² through a projection integral of the form

$$u(\hat{n}) = \int d\chi q_u(\chi) U(\chi\hat{n}, z(\chi)), \quad (2)$$

where $q_u(\chi)$ is a known radial (comoving distance) kernel. Multiple examples of such fields exist in cosmology and astrophysics. Here, we will consider the following:

- (i) The **CMB lensing** convergence field (A. Lewis & A. Challinor 2006) is a projected tracer of the matter overdensity ($U = \delta_m$):

$$\kappa(\hat{n}) = \int d\chi q_\kappa(\chi) \delta_m(\chi\hat{n}, z(\chi)), \quad (3)$$

where the lensing kernel is

$$q_\kappa(\chi) \equiv \frac{3}{2} H_0^2 \Omega_m (1+z) \chi \frac{\chi_{\text{LSS}} - \chi}{\chi_{\text{LSS}}}. \quad (4)$$

Here, Ω_m and H_0 are the cosmic non-relativistic matter fraction and the Hubble parameter, respectively, and χ_{LSS} is the comoving distance to the last-scattering surface.

- (ii) The **tSZ** effect (J. E. Carlstrom, G. P. Holder & E. D. Reese 2002) Compton- y parameter is a projected tracer of the thermal electron pressure P_e :

$$y(\hat{n}) = \int d\chi q_y(\chi) P_e(\chi\hat{n}, z(\chi)), \quad q_y(\chi) \equiv \frac{\sigma_T}{m_e c^2} \frac{1}{1+z}, \quad (5)$$

where σ_T and m_e are the Thomson scattering cross-section and the electron mass, respectively, and c is the speed of light.

- (iii) The **cosmic infrared background** (CIB) intensity at a given observed frequency I_ν is sourced by dust emission in star-forming galaxies, and is thus a tracer of the SFR density ρ_{SFR} (A. Maniyar, M. Béthermin & G. Lagache 2021):

$$I_\nu(\hat{n}) = \int d\chi q_{\text{CIB},\nu}(\chi) \rho_{\text{SFR}}(\chi\hat{n}, z(\chi)), \quad (6)$$

where the radial kernel is

$$q_{\text{CIB},\nu}(\chi) \equiv \frac{\chi^2 S_\nu^{\text{eff}}(z)}{K}. \quad (7)$$

Here, K is the calibration constant relating far-infrared luminosity and SFR (sometimes called the ‘Kennicutt constant’; R. C. Kennicutt 1998), and $S_\nu^{\text{eff}}(z)$ is the mean flux of infrared sources at redshift z normalized to unit luminosity (see e.g. A. Maniyar et al. 2021).

In addition to these, consider the projected overdensity of galaxies

$$\Delta_g(\hat{n}) = \int d\chi q_g(\chi) \delta_g(\chi\hat{n}, z(\chi)), \quad q_g(\chi) \equiv \frac{H(z)}{c} p(z), \quad (8)$$

where δ_g is the 3D galaxy overdensity, $H(z)$ is the expansion rate at redshift z , and $p(z)$ is the redshift distribution of the galaxies. It is often possible to obtain reasonably accurate measurements for galaxy redshifts, and thus well-localized samples can be selected, in which case $p(z)$ is a compact function centred at a mean redshift \bar{z}_g .

The angular power spectrum between any two projected quantities, u and v , is given by

$$C_\ell^{uv} = \int \frac{d\chi}{\chi^2} q_u(\chi) q_v(\chi) P_{UV} \left(k = \frac{\ell + 1/2}{\chi}, z(\chi) \right), \quad (9)$$

²Note that we use redshift z as an effective time variable.

where $P_{UV}(k, z)$ is the 3D power spectrum between the two 3D quantities, U and V , that the projected quantities trace.

The halo model (J. A. Peacock & R. E. Smith 2000; U. Seljak 2000; A. Cooray & R. Sheth 2002) provides a useful framework to make predictions for the power spectrum of general LSS tracers. A key prediction of this model is that, on sufficiently large scales, the power spectrum $P_{UV}(k, z)$ can be approximated as

$$P_{UV}(k, z) \simeq \langle bU(z) \rangle \langle bV(z) \rangle P_{\text{mm}}(k, z), \quad (10)$$

where $P_{\text{mm}}(k, z)$ is the matter power spectrum, and we have defined the bias-weighted average:

$$\langle bU \rangle = \int dM n(M) b_h(M) \tilde{U}(M), \quad (11)$$

where $n(M)$ is the halo mass function (HMF), $b_h(M)$ is the linear halo bias, and $\tilde{U}(M)$ is the average integral of the halo profile for the physical quantity U over volume, for haloes of mass M :

$$\tilde{U}(M) \equiv 4\pi \int_0^\infty dr r^2 U(r|M), \quad (12)$$

where $U(r|M)$ is the halo profile for tracer U (e.g. gas pressure, SFR density, and matter density). $\langle bU \rangle$ is thus the cosmic average value of U weighted by the linear bias of the haloes it populates. This is an interesting quantity to measure, as it can complement other measurements of the unweighted cosmic average $\langle U \rangle$, by providing a handle on the mass dependence of U . Note that, although we have used the matter power spectrum in equation (10), the simplest halo model prediction uses the *linear* matter power spectrum instead. Both agree on sufficiently large scales, although we will explore different options for $P_{\text{mm}}(k, z)$ when building robust tomographic reconstruction estimators in Section 2.2.

Consider now the cross-correlation between galaxies at a given mean redshift \bar{z}_g , and a projected field u , C_{ℓ}^{gu} , and the galaxy autocorrelation C_{ℓ}^{gg} . As discussed above, on sufficiently large scales, and assuming that the width of the galaxy redshift distribution is much smaller than the time-scales over which cosmological quantities and biases vary significantly, we can model these two power spectra as

$$C_{\ell}^{gu} \simeq b_g \langle bU \rangle_{\bar{z}_g} T_{\ell}^{gu}, \quad C_{\ell}^{gg} \simeq b_g^2 T_{\ell}^{gg}, \quad (13)$$

where the template angular power spectra T_{ℓ}^{gu} and T_{ℓ}^{gg} are

$$T_{\ell}^{xy} \equiv \int \frac{d\chi}{\chi^2} q_x(\chi) q_y(\chi) P_{\text{mm}} \left(\frac{\ell + 1/2}{\chi}, z(\chi) \right). \quad (14)$$

Since these templates depend only on the known radial kernels, and on cosmological parameters, the amplitudes of C_{ℓ}^{gu} and C_{ℓ}^{gg} can be used, in combination, to measure $\langle bU \rangle$ at the redshift of the galaxies, in addition to the galaxy bias b_g .

Although we have defined $\langle bU \rangle$ as a bias-weighted average in equation (11), as predicted by the halo model, this quantity can be more generally interpreted as the relative clustering amplitude of the quantity U with respect to the matter overdensity. Quantitatively, this could be defined as in e.g. S. Young et al. (2021) and Z. Chen et al. (2024):

$$\langle bU \rangle = \lim_{k \rightarrow 0} \frac{P_{Um}(k, z)}{P_{\text{mm}}(k, z)}. \quad (15)$$

We will refer to this model as ‘large-scale approximation’ (LSA), hereafter.

Ultimately, we need to be able to connect the measured amplitude parameter $\langle bU \rangle$ with the clustering of both U and δ_m , with as little dependence as possible on the unknown clustering of the galaxies used to make the measurement.

2.2 Tomographic reconstruction estimators

Consider now the problem of estimating $\langle bU \rangle$ from real measurements of the galaxy autocorrelation and its cross-correlation with the relevant LSS tracer. For simplicity, we will address this problem directly at the level of the 3D power spectra $P_{xy}(k)$, since these measurements are straightforward to make in simulations, and the radial projection on to C_{ℓ} s does not introduce any additional relevant complications (assuming that the galaxy redshift distribution is compact and well known).

Let $\hat{P}_{\text{gg}}(k)$ and $\hat{P}_{\text{gu}}(k)$ be measurements of the auto- and cross-correlations, made at a discrete set of wavenumbers k . A simple ansatz to model these measurements would be

$$\hat{P}_{\text{gg}}(k) = b_g^2 \tilde{P}_{\text{mm}}(k), \quad \hat{P}_{\text{gu}}(k) = b_g \langle bU \rangle \tilde{P}_{\text{mm}}(k), \quad (16)$$

where $\tilde{P}_{\text{mm}}(k)$ is a suitable fixed template for the matter power spectrum. The values of b_g and $\langle bU \rangle$ can then be determined by fitting the measured power spectra to this model.

This model can be improved by adding a constant term in $\hat{P}_{\text{gg}}(k)$ to account for the presence of shot noise in the galaxy autocorrelation. For sufficiently dense samples, this term would only be relevant at large values of k , but in general shot noise can impact all scales. At the same time, we may include a similar term for $\hat{P}_{\text{gu}}(k)$. While in principle shot noise would not affect these cross-correlations, at intermediate scales this term would be able to absorb the small-scale contribution to $\hat{P}_{\text{gu}}(k)$ from pairs of points within the same dark matter halo (i.e. the ‘1-halo’ contribution, in the language of the halo model), as well as any other sources of uncorrelated noise in either spectrum. The extended model would then read (B. Jago et al. 2023b)

$$\hat{P}_{\text{gg}}(k) = b_g^2 \tilde{P}_{\text{mm}}(k) + N_{\text{gg}}, \quad (17)$$

$$\hat{P}_{\text{gu}}(k) = b_g \langle bU \rangle \tilde{P}_{\text{mm}}(k) + N_{\text{gu}}, \quad (18)$$

where N_{gg} and N_{gu} are now two additional free parameters to be included in the fit. We will refer to this model as ‘linear model’ or $n_{\text{max}} = 0$ (in reference of the more general model explained below).

Finally, we may further enhance the model by adding terms that could absorb the impact of non-linear biasing in both δ_g and U . Assuming these deviations to be slow-varying in k space, this could be achieved through a polynomial expansion in k . Here, we will consider adding terms of the form $\propto k^{2n} \tilde{P}_{\text{mm}}(k)$, which, for values of $n \geq 1$, lead to contributions that are relevant at increasingly small scales, where we expect deviations from linear bias to become more relevant. The most general model we will use here to estimate $\langle bU \rangle$ is thus

$$\hat{P}_{\text{gg}}(k) = b_g^2 \tilde{P}_{\text{mm}}(k) + \sum_{n=1}^{n_{\text{max}}} A_{\text{gg}}^n k^{2n} \tilde{P}_{\text{mm}}(k) + N_{\text{gg}}, \quad (19)$$

$$\hat{P}_{\text{gu}}(k) = b_g \langle bU \rangle \tilde{P}_{\text{mm}}(k) + \sum_{n=1}^{n_{\text{max}}} A_{\text{gu}}^n k^{2n} \tilde{P}_{\text{mm}}(k) + N_{\text{gu}}, \quad (20)$$

where A_{gg}^n and A_{gu}^n are additional free amplitude parameters. The simplicity of this model allows us to obtain fast analytical estimates for all free parameters, as described in Appendix A.

The final ingredient of this estimator is the template $\tilde{P}_{\text{mm}}(k)$ used in lieu of the matter power spectrum. Here, we will consider four different options:

- (i) The dark matter-only (DMO) linear matter power spectrum.
- (ii) The DMO non-linear power spectrum calculated via the HALOFIT parametrization of R. Takahashi et al. (2012).
- (iii) The DMO non-linear power spectrum calculated through the BACCOEMU emulator (R. E. Angulo et al. 2021).

(iv) The matter power spectrum estimated from the same simulation used to measure \hat{P}_{gg} and \hat{P}_{gU} . This will include, for example, the impact of baryonic effects (M. Schaller et al. 2025).

The first three options represent viable choices in a realistic data analysis scenario, and differ in the level of accuracy with which they predict the true underlying matter power spectrum. The fourth option is clearly unrealistic, as the true matter power spectrum is not directly observable, but the resulting constraints on $\langle bU \rangle$ represent the best possible scenario, unaffected by inaccuracies in the model of the matter power spectrum.

This defines the most general cross-correlation-based estimator of $\langle bU \rangle$ we will use in this work. We will apply this estimator in Section 4 to different galaxy samples, and quantify the consistency of the resulting $\langle bU \rangle$ measurements. Rather than only comparing the values found for different galaxy samples among themselves, we will compare these with a ‘true’ value, independent of any galaxy clustering properties. Instead of predicting $\langle bU \rangle$ according to the halo model, using equation (11), we will define its reference value as the value found by the estimator defined above using the matter overdensity instead of the overdensity of any galaxy sample [i.e. using $\hat{P}_{\text{mm}}(k)$ and $\hat{P}_{\text{mU}}(k)$ measured from the simulation instead of $\hat{P}_{\text{gg}}(k)$ and $\hat{P}_{\text{gU}}(k)$]. This will allow us to ascertain whether the $\langle bU \rangle$ estimated from galaxies agrees with a universal definition of $\langle bU \rangle$ that only depends on the clustering of matter and of the tracer U .

2.3 The halo model prediction

Once we have established the agreement between the value of $\langle bU \rangle$ estimated from galaxies and its reference value estimated from the matter field, we will turn to the question of whether this reference value agrees with the halo model prediction for $\langle bU \rangle$. This theoretical prediction can be calculated by solving the integral in equation (11), using existing parametrizations for the HMF and halo bias (e.g. J. Tinker et al. 2008; J. L. Tinker et al. 2010; G. Despali et al. 2016; L. Ondaro-Mallea et al. 2024), as well as a model for the halo mass dependence of U . However, the resulting estimate of equation (11) would be subject to the inaccuracies in these parametrizations (A. Knebe et al. 2011), as well as differences in the halo mass dependence of hydrodynamical quantities (e.g. thermal energy or SFR) in FLAMINGO and in existing models based on first principles or other simulation suites. Instead, we will estimate the halo model prediction by implementing equation (11) directly in the simulation:

$$\langle bU \rangle_{\text{HM}} = \frac{1}{V_{\text{box}}} \sum_h b_h \sum_{i \in h} \tilde{U}_i. \quad (21)$$

Here, the index h runs over all haloes resolved in the simulation, while i runs over all simulation particles identified as belonging to halo h . b_h is the linear halo bias for the h th halo, which we calculate in three different ways:

(i) b_h calculated directly from the simulation. We estimate the value of b_h using the estimator of $\langle bU \rangle$ described in the previous section (equations 19 and 20), with $\hat{P}_{\text{mm}}(k)$ instead of $\hat{P}_{\text{gg}}(k)$, and the matter-halo cross-spectrum, $\hat{P}_{\text{mh}}(k|M_h)$, instead of $\hat{P}_{\text{gU}}(k)$. See Section 3.2 for a full description.

(ii) b_h calculated from J. L. Tinker et al. (2010).

(iii) b_h calculated using the peak-background split (PBS) framework with the HMF in G. Despali et al. (2016).

(iv) b_h calculated using PBS and the HMF in L. Ondaro-Mallea et al. (2024).

Finally, \tilde{U}_i is the value of the U field integrated over the effective volume of the i th particle. To estimate its value for each halo, $\sum_{i \in h} \tilde{U}_i$, we proceed as follows:

(i) For the case of $U = \rho_{\text{SFR}}$, \tilde{U}_i is simply the total SFR assigned to the particle. Thus, we use the total SFR for each halo in the FLAMINGO halo catalogue. This is calculated by combining all particles assigned to the halo according to a spherical overdensity mass definition with an overdensity $\Delta = 200$ times larger than the critical density.

(ii) For the case of $U = P_e$, we calculate $\sum_{i \in h} \tilde{U}_i$ from the values of the Compton- y parameter assigned to each halo in the FLAMINGO catalogues. As described in I. G. McCarthy et al. (2017), the volume-integrated pressure in this case is given by

$$\tilde{P}_{e,h} = y_h d_A^2(z) \frac{m_e c^2}{\sigma_T}, \quad (22)$$

where d_A is the angular distance to the redshift of the corresponding snapshot, and y_h is the Compton- y parameter of the halo, calculated by summing over all particles assigned to it. Unlike in the case of SFR, the tSZ signal receives significant contribution from gas at distances beyond the virial radius of the halo (I. G. McCarthy et al. 2014). To ensure that this contribution is included in our calculation, in this case we include all particles within a sphere of radius $5 \times R_{500c}$, where R_{500c} is the spherical overdensity radius of the halo with an overdensity of $\Delta = 500$ times the critical density. Note that this may lead to double-counting of particles lying within the spheres of more than one halo, as discussed in Section 4.2.3.

3 SIMULATIONS

We employ the state-of-the-art FLAMINGO simulations (R. Kugel et al. 2023; J. Schaye et al. 2023), one of the largest sets of gravity-only and hydrodynamical simulations with cosmological volumes ever produced. They are calibrated such that the stellar and active galactic nucleus feedback models reproduce the observed $z = 0$ galaxy stellar mass function and cluster gas fractions, shown to reproduce cluster scaling relations and thermodynamic profiles (J. Braspennig et al. 2024). The simulations were run with the SWIFT code (M. Schaller et al. 2024), using the SPHENIX smoothed-particle hydrodynamic (SPH) scheme (J. Borrow et al. 2022).

We focus on the fiducial simulation of the set. This has a cosmology, dubbed D3A, based on the results of the DES-Y3 3×2 -point analysis, including all external constraints (T. M. C. Abbott et al. 2022). That is, it assumes a flat geometry, a total matter density parameter $\Omega_m = 0.306$, a baryon density parameter $\Omega_b = 0.0486$, and a neutrino component with total mass $M_\nu = 0.06$ eV. The power-law index of primordial scalar perturbations is $n_s = 0.967$ and the amplitude of the primordial power spectrum on a pivot scale $k_{\text{pivot}} = 0.05 \text{ Mpc}^{-1}$ is $A_s = 2.099 \times 10^{-9}$. This corresponds to a standard deviation of linear perturbations (smoothed on spheres of radius $8 h^{-1} \text{ Mpc}$) at $z = 0$ of $\sigma_{8,\text{tot}} = 0.807$ when considering perturbations in total matter or $\sigma_{8,\text{cb}} = 0.811$ when considering only cold matter perturbations (cold dark matter + baryons). The Hubble parameter at redshift zero is $H_0 = 68.1 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The fiducial simulation we employ simultaneously evolves 1800^3 dark matter particles, an equal number of initial gas particles, and 1000^3 neutrino particles in a comoving box of side $L_{\text{box}} = 1000 \text{ Mpc}$. The mass of each dark matter particle is $M_{\text{DM}} = 5.65 \times 10^9 M_\odot$, and the initial mass of the gas particle is $M_{\text{gas}} = 1.07 \times 10^9 M_\odot$. The particles are initialized at $z = 31$ using three-fluid third-order Lagrangian perturbation theory. Cosmic variance is partially suppressed by enforcing a fixed amplitude initial power spectrum (matching the

ensemble average) on scales $(kL_{\text{box}})^2 < 1025$. The simulations have been post-processed to identify dark matter haloes, their substructures, and galaxies with their respective properties. We use haloes identified with an updated version (V. J. Forouhar Moreno et al. 2025) of the Hierarchical Bound Tracing algorithm, HBT (J. Han et al. 2018). Galaxy properties are computed with a FLAMINGO-specific tool called SOAP, Spherical Overdensity and Aperture Processor (see J. Schaye et al. 2023, for details). Specifically, we use galaxy properties computed by SOAP in apertures of 50 kpc, excluding unbound particles.

The different galaxy types used in our analysis have been selected by a cumulative bin in stellar mass, M_* . We sort galaxies in the simulation in decreasing order of M_* , and select the first N galaxies, with $N = nV_{\text{box}}$. We create three bins with number densities³ $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$. Note that we used the fiducial simulation with $L_{\text{box}} = 1000 \text{Mpc}$, instead of the larger box simulation, with $L_{\text{box}} = 2800 \text{Mpc}$, in order to facilitate a direct comparison with the results presented in M. Zennaro et al. (2024), which used the same galaxy samples used here. We have tested that the results presented here are not affected by this choice, by comparing the Power-Spectra of the two simulations at the scales of interest.

3.1 $P(k)$ estimation

We calculate new power spectra, independent from FLAMINGO, for all tracers: simulation particles, haloes, and galaxies. For all of them, we use an FFT grid with $n_g = 512^3$ cubic cells covering the whole simulated volume, and a Triangular Shape Cloud (TSC) mass assignment. This corresponds to a Nyquist frequency of $k_N = 1.61 \text{Mpc}^{-1}$. Although this corresponds to a moderate spatial resolution, it is sufficient for the tomography estimators we will explore here, which employ relatively large scales. We use a logarithmic spacing from the fundamental mode to k_N , sampling $n_g/4$ bins. The final product is corrected for aliasing following R. E. Angulo et al. (2008), M. Sato & T. Matsubara (2011), and R. Takahashi et al. (2012).

To construct the U maps from simulated particles, which will then be cross-correlated with \hat{P}_g to create $\hat{P}_{gU}(k)$, we proceed as follows:

- (i) $U = \delta_m$: we just add the masses of all particle types in the simulation (dark matter, gas, stars, and neutrinos) at each grid point, and divide by the mean mass per cell.
- (ii) $U = \rho_{\text{SFR}}$: in addition to the TSC weights, each gas particle is weighted by its SFR. The total SFR in each cell is then divided by the comoving cell volume.
- (iii) $U = P_e$: the comoving thermal pressure is calculated by adding the total thermal energy of all particles in each cell and dividing by its comoving volume. As described in A. J. Mead et al. (2020), this is achieved by weighting each particle by the quantity

$$\tilde{P}_{e,i} = k_B n_{e,i} T_{e,i} \frac{V_i}{V_{\text{cell}}}, \quad (23)$$

where k_B is the Boltzmann constant, $n_{e,i}$ is the electron number density for the i th gas particle, $T_{e,i}$ is its electron temperature, and $V_i = m_i/\rho_i$ is the SPH volume element of the particle. The comoving cell volume is simply $V_{\text{cell}} = (L_{\text{box}}/n_g)^3$.

³Note that, confusingly, we use ‘ h -inverse’ units to define the galaxy samples used here, as opposed to ‘non- h ’ units used elsewhere in this work. This is to match the same galaxy samples used in M. Zennaro et al. (2024).

3.2 b_h estimation

In order to calculate the halo model prediction $\langle bU \rangle_{\text{HM}}$ from equation (21), we need to estimate the halo bias b_h as a function of mass. Although we will test the performance of parametrizations present in the literature (J. L. Tinker et al. 2010; G. Despali et al. 2016; L. Ondaro-Mallea et al. 2024), we also infer its value directly from the simulation. This will ensure that any differences between $\langle bU \rangle_{\text{HM}}$ and the value of $\langle bU \rangle$ estimated from the data are not due to inaccuracies in the halo bias parametrizations, or due to completeness issues for low-mass haloes in FLAMINGO, rather than a genuine failure of the halo model. We measure $b_h(M, z)$ by dividing the FLAMINGO halo catalogue in five different mass bins at each snapshot. We bin haloes according to their spherical overdensity mass, with an overdensity parameter $\Delta = 200$, defined with respect to the critical density. We use adjacent logarithmic mass bins with edges:

$$\log_{10}(M_{200c}/M_{\odot}) = [8, 11, 12, 13, 14, 17]. \quad (24)$$

In each mass bin, we construct maps of the halo overdensity and estimate its cross-correlation with the matter overdensity field $\hat{P}_{mh}(k, z|M)$. We then use this in combination with a measurement of the matter power spectrum $\hat{P}_{mm}(k, z)$ to estimate the halo bias b_h using the same fiducial prescription we will use in Section 4 to estimate $\langle bU \rangle$. The prescription is based on the model in equation (20), with $n_{\text{max}} = 1$, setting $b_g \rightarrow 1$ and fitting for $b_h = \langle bU \rangle$ marginalized over the model nuisance parameters A_{mh}^1 and N_{mh} , using scales up to a maximum wavenumber $k_{\text{max}} = 0.3 \text{Mpc}^{-1}$.

Once b_h has been estimated in each mass bin, we build a linear interpolating function associating the measured values of b_h with the median mass in each bin. We then assign a value of b_h to each halo in the simulation by evaluating this function at the corresponding halo mass M_{200c} . Note that the first mass bin above is highly incomplete. Nevertheless, the method used here to measure b_h accounts for the impact of this inconsistency in the effective halo bias at low masses.

4 RESULTS

In this section, we will first study the robustness of the procedure described in Section 2.2 to reconstruct $\langle bU \rangle$ from galaxy cross-correlations, and select a fiducial version of this procedure simultaneously optimizing the precision and robustness of the recovered $\langle bU \rangle$. In doing this, we will also evaluate our ability to account for the dependence of the reconstructed values on cosmological parameters. We will then apply our chosen prescription to the recovery of $\langle bP_e \rangle$ and $\langle b\rho_{\text{SFR}} \rangle$ from FLAMINGO, verifying that consistent values can be obtained with different galaxy samples, and compare the result with existing measurements from real data. Finally, we will turn to the interpretability of the recovered $\langle bU \rangle$ in terms of the halo model.

4.1 Validation of the model

4.1.1 Bias modelling and scale cuts

Our first task is to determine the range of validity of the bias model described by equations (19) and (20) and used to extract measurements of $\langle bU \rangle$, as a function of galaxy type, n_{max} , and of the smallest scale used k_{max} . To do so, we will test the model against a known solution, by setting $U = \delta_m$, the matter overdensity. Labeling the corresponding bias-weighted quantity for this case (b_m), an unbiased estimator should recover the value $\langle b_m \rangle = 1$. As a power spectrum template, $\tilde{P}_{mm}(k)$, we used the true matter power spectrum of the hydrodynamical simulation being analysed, in order to factor

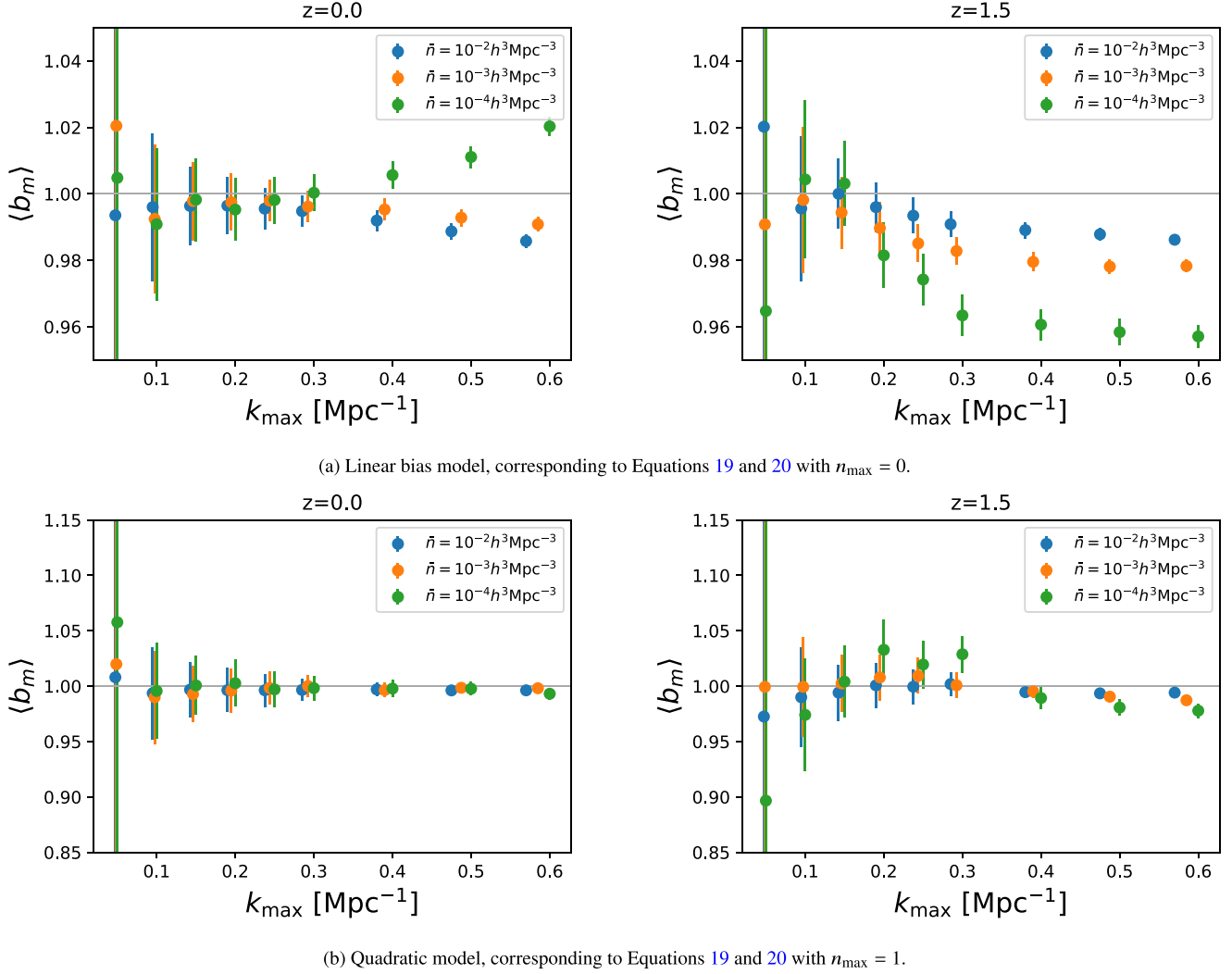


Figure 1. Deviation from theoretical prediction in the value of $\langle b_m \rangle$ as a function of k_{\max} , and for two different redshifts ($z = 0, 1.5$ in the left and right panels, respectively). Results are shown for three stellar mass-selected samples, with number densities $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$ (blue, orange, and green points, respectively). The top and bottom panels show the results for two different choices of n_{\max} , which defines the bias model used to fit the data according to equations (19) and (20).

out uncertainties related with this choice, which we will study in detail in the next section.

Fig. 1 shows the value of $\langle bU \rangle$ estimated from our three different galaxy samples (blue, orange, and green points for the samples with number densities $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$, respectively). Results are shown as a function of k_{\max} , for the linear bias model with $n_{\max} = 0$ (top panels), and for the quadratic model ($n_{\max} = 1$, bottom panels). Results are shown for the snapshots at $z = 0$ and $z = 1.5$ (left and right, respectively). We find that, as expected, the quadratic model is significantly more stable than the linear model, particularly at low redshifts. We find that the former is able to obtain unbiased results across different samples up to scales $k_{\max} > 0.3-0.5 \text{ Mpc}^{-1}$, whereas the linear model breaks down dramatically at $k_{\max} = 0.15-0.25 \text{ Mpc}^{-1}$. This, as we will see in Section 4.2, is also comparable to the performance achieved when recovering $\langle b\rho_{\text{SFR}} \rangle$ and $\langle bP_e \rangle$.

Increasing n_{\max} leads to increased stability in the recovered value of $\langle bU \rangle$, and robustness against deviations from a purely linear bias relation in both δ_g and U . However, this comes at the cost of larger statistical uncertainties. It is therefore interesting to study whether

an optimal value of n_{\max} can be found that minimizes the error in the resulting measurement of $\langle bU \rangle$ within the range of scales for which the estimator remains sufficiently unbiased. Fig. 2 shows the results of inferring $\langle b_m \rangle$ as a function of k_{\max} for different values of $n_{\max} \in \{0, 1, 2\}$. The results shown were obtained for the galaxy sample with $\bar{n} = 10^{-2} h^3 \text{Mpc}^{-3}$ at an intermediate redshift $z = 0.8$, but the same qualitative results are obtained for other galaxy-redshift combinations. The top panel shows the difference in the recovered value of $\langle b_m \rangle$ with respect to the theoretical expectation of $\langle b_m \rangle = 1$ as a fraction of the statistical uncertainties. The horizontal dashed line marks the points at which the estimator recovers a bias of $\sim 0.5\sigma$,⁴ with the vertical dotted lines signalling the scale cut k_{\max} at which

⁴Note that this benchmark (a bias in $\langle b_m \rangle$ equal to half the statistical uncertainties) is relatively arbitrary, as $\sigma(\langle b_m \rangle)$ is largely determined by the size of the simulation box. We use it here only to highlight the trade-off between precision and accuracy when selecting a value of n_{\max} . Elsewhere in this work, we will use the fractional bias on $\langle bU \rangle$ as a metric to quantify the performance of a given estimator.

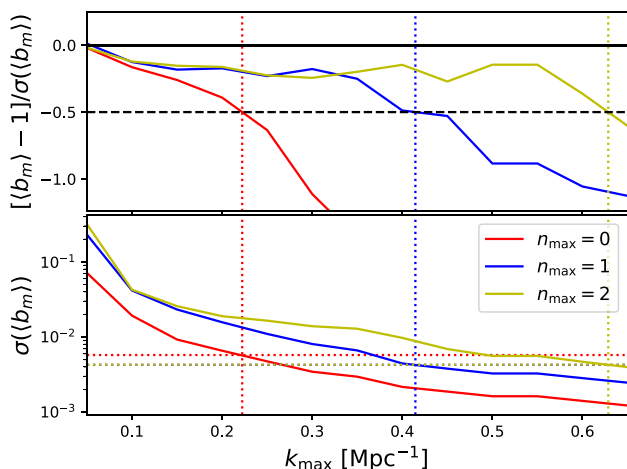


Figure 2. Top panel: deviation from the theoretical prediction in the value of $\langle b_m \rangle$, relative to the statistical uncertainties in this quantity, as a function of k_{\max} . Results are shown for bias models with $n_{\max} = 0, 1$, and 2 (red, blue, and yellow contours, respectively). The dotted lines mark the scales at which the bias in $\langle b_m \rangle$ exceeds 0.5σ . Bottom panel: uncertainty on the recovered $\langle b_m \rangle$ as a function of k_{\max} for the same bias models, with the dotted lines marking the smallest uncertainty achieved by each model within the range of k_{\max} for which it remains unbiased. All results are shown for the galaxy sample with density $\bar{n} = 10^{-2} h^3 \text{Mpc}^{-3}$ at an intermediate redshift $z = 0.8$.

this happens. The bottom panel then shows the error with which $\langle b_m \rangle$ is recovered, with the dotted lines showing the error achieved at the corresponding k_{\max} . We can see that, while the uncertainties for $n_{\max} = 0$ are consistently lower than for $n_{\max} = 1$ or 2 at any given k_{\max} , the fact that $n_{\max} = 1$ can recover unbiased results while pushing to smaller scales allows it to outperform $n_{\max} = 0$. In turn, while $n_{\max} = 2$ allows us to use even smaller scales while remaining unbiased, it does not outperform $n_{\max} = 1$. Given this, we choose $n_{\max} = 1$ as the fiducial bias model defining our estimator for $\langle bU \rangle$, balancing precision and robustness against non-linear biasing. With this choice, as we will show also in the cases of $\langle bP_e \rangle$ and $\langle b\rho_{\text{SFR}} \rangle$, we will in general be able to obtain largely unbiased constraints using scales up to $k_{\max} = 0.3 \text{Mpc}^{-1}$.

4.1.2 Power spectrum template

Another variable defining our estimator of $\langle bU \rangle$ is the choice of power spectrum template $\tilde{P}_{\text{mm}}(k)$ used to fit the data. The previous analysis has been done using the spectrum measured directly from the simulation $\tilde{P}_{\text{mm}}^{\text{sim}}(k)$. This is obviously not available when applying this estimator to real data, and a model predicting $\tilde{P}_{\text{mm}}(k)$ must be used. In Fig. 3, we test the robustness of our results with respect to the choice of template, exploring the three other templates described in Section 2.2: the linear power spectrum, the HALOFIT prediction for the non-linear matter power spectrum, and the result of the BACCOEMU emulator for the same quantity. The results shown were obtained for our fiducial bias model with $n_{\max} = 1$, and we present them for the $z = 0$ and $z = 1.5$ snapshots (top and bottom panels).

We find that both HALOFIT and BACCOEMU are able to obtain unbiased constraints on $\langle b_m \rangle$ for all galaxy samples up to relatively small scales (certainly within the range $k_{\max} \leq 0.3$). Even the linear matter power spectrum recovers relatively accurate constraints, although the results are substantially less stable, particularly for the most massive galaxy sample explored ($\bar{n} = 10^{-4} h^3 \text{Mpc}^{-3}$) at high redshifts. Thus, we find that a relatively naive prediction for

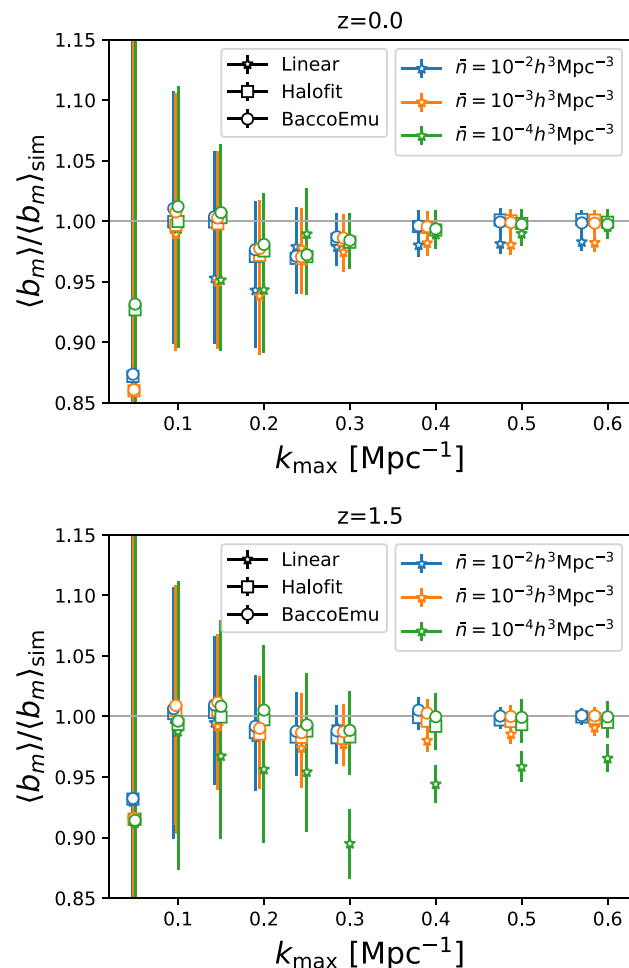


Figure 3. Estimation of $\langle b_m \rangle$ in the quadratic model (i.e. $n_{\max} = 1$) with $P(k)$ from linear theory, and non-linear HALOFIT and BACCOEMU results (star, squared, and circle, respectively), as a function of k_{\max} for two different redshifts ($z = 0, 1.5$). We study galaxies selected in order of decreasing stellar mass, with the different colours (blue, orange, and green) corresponding to different number densities, namely $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$.

the matter power spectrum, such as HALOFIT, is sufficient to recover robust constraints on $\langle bU \rangle$, with the free parameters of our bias model able to absorb the small-scale inaccuracies of this prediction. It is also interesting to note that, although the simulated data contain baryonic effects, here we use the power spectrum of only the cold matter components (dark matter and baryons) in the absence of baryonic feedback. The bias model is again flexible enough to absorb these effects and yield an unbiased estimate of $\langle bU \rangle$.

4.1.3 Dependence on the fiducial cosmology

Although we have shown that the estimator is relatively robust to the choice of power spectrum template, this has been under the assumption that this template was created assuming a cosmological model that matches that of the data. This is not possible in practice, and it is therefore important to characterize the dependence of the estimated $\langle bU \rangle$ on the true cosmological parameters, which in general differ from those used to generate the $P(k)$ template. In this section, we provide a prescription to account for the potential dependence on cosmological parameters of the $\langle bU \rangle$ estimator if desired. This is similar in spirit to the procedure used to interpret

Table 1. Cosmological parameters used to quantify the cosmology dependence of the tomographic estimator developed here. The fiducial cosmology corresponds to that of the FLAMINGO simulation analysed here.

Name	σ_8	Ω_m	n_s
Fiducial	0.8069	0.3060	0.967
Low σ_8	0.7499	0.3060	0.967
Low Ω_m	0.8071	0.27	0.967
High Ω_m	0.8071	0.344	0.967
High n_s	0.8069	0.3060	1.0

baryon acoustic oscillation and redshift space distortion observables in cosmological galaxy clustering analyses (e.g. J. E. Bautista et al. 2021).

Let us consider a simple scenario in which the only parameter mismatch is in the value of the amplitude of matter fluctuations, σ_8 , such that, on the large scales we use to measure $\langle bU \rangle$, $\tilde{P}_{\text{mm}}(k) = (\tilde{\sigma}_8/\sigma_8)^2 P_{\text{mm}}(k)$, where $P_{\text{mm}}(k)$ is the true underlying power spectrum, and $\tilde{\sigma}_8$ is the value of σ_8 assumed to generate $\tilde{P}_{\text{mm}}(k)$. On large scales, we can equate the gg and gU spectra to

$$\hat{P}_{\text{gg}}(k) = \hat{b}_g^2 \tilde{P}_{\text{mm}}(k) = b_g^2 P_{\text{mm}}(k), \quad (25)$$

$$\hat{P}_{\text{gU}}(k) = \hat{b}_g \langle b\hat{U} \rangle \tilde{P}_{\text{mm}}(k) = b_g \langle bU \rangle P_{\text{mm}}(k), \quad (26)$$

where hatted quantities (e.g. \hat{P}) are those estimated from the data. Solving for $\langle b\hat{U} \rangle$, we find that its dependence on the true value of σ_8 is, unsurprisingly,

$$\langle b\hat{U} \rangle = \frac{\sigma_8}{\tilde{\sigma}_8} \langle bU \rangle. \quad (27)$$

In practice, a general shift in cosmological parameters will result in changes in both the amplitude and the shape of the power spectrum template. We can, however, assume that the additional terms in our generic bias model (N_{xy} and A_{xy}^n in equation 20) will be able to absorb most of the change in the power spectrum shape while recovering the small-scale behaviour of the measured spectra, while b_g and $\langle bU \rangle$ capture the overall amplitude of the power spectrum on large scales. With this intuition in mind, we will use the following ansatz to capture the cosmology dependence of the estimated value of $\langle bU \rangle$, by generalizing equation (27) to

$$\langle b\hat{U} \rangle = \frac{\sigma(k_L)}{\tilde{\sigma}(k_L)} \langle bU \rangle, \quad (28)$$

where $\sigma(k_L)$ is the variance of the matter overdensity on wavenumbers smaller than k_L :

$$\sigma(k_L) = \left[\frac{1}{2\pi^2} \int_0^{k_L} dk k^2 P_{\text{mm}}(k) \right]^{1/2}, \quad (29)$$

and $\tilde{\sigma}(k_L)$ is the same quantity calculated for the power spectrum template $\tilde{P}_{\text{mm}}(k)$ used in the estimator. We will use a default value of $k_L = 0.07 h \text{ Mpc}^{-1}$.

To validate this ansatz, we reanalyse the simulation data used to infer $\langle b_m \rangle$ using matter power spectrum templates generated using cosmologies that differ from the fiducial model (i.e. the fiducial FLAMINGO cosmology). In particular, we tested cosmologies with a lower value of σ_8 , lower and higher values of Ω_m , and a higher value of the scalar spectral index n_s . The specific values used are listed in Table 1. Note that these parameter variations are substantially larger than the current uncertainties on these parameters from CMB data. The results of this analysis are shown in Fig. 4. The figure shows the values of $\langle b_m \rangle$, for the baseline marker (the matter density), recovered using power spectrum templates for each cosmology (star markers),

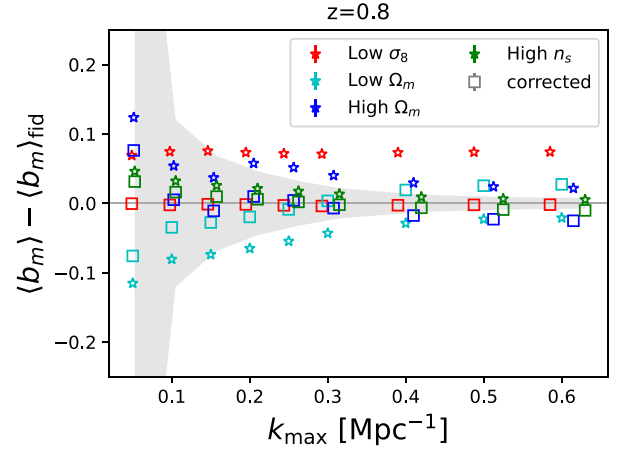


Figure 4. Comparison between the inferred values of $\langle b_m \rangle$ obtained when using a $\tilde{P}_{\text{mm}}(k)$ with a different cosmology from the fiducial (that of the galaxy measurements). The analysis uses $n_{\text{max}} = 1$, $\tilde{P}_{\text{mm}}(k)$ from BACCOEMU, and only the matter distribution (baseline fit) is shown. The corrected values are obtained by applying equation (28) to the starred values, showing the validity of the model up to the limit scales ($k_{\text{max}} < 0.3 \text{ Mpc}^{-1}$). The shaded area corresponds to the uncertainty in $\langle b_m \rangle_{\text{fid}}$.

as well as the same measurements corrected for the cosmology-dependent factor in equation (28) (empty square markers). We can see that the large scatter between different cosmologies is significantly reduced after applying the cosmology-dependent scaling factor, recovering the value of $\langle b_m \rangle$ inferred with the unbiased power spectrum template within 1σ in all cases, within the regime of validity of our estimator ($k_{\text{max}} \leq 0.3 h \text{ Mpc}^{-1}$). Interestingly, the low- Ω_m cosmology displays the largest disagreement after this correction. It is likely possible to develop a more precise scheme to account for the cosmology dependence of our measurements. However, since the main aim of this paper is to present a scheme for tomographic reconstruction that is robust to the clustering properties of the galaxy samples used, we leave such a study for future work.

4.2 Tomographic reconstruction of baryonic properties

Having tested the robustness of the tomographic estimator of $\langle bU \rangle$ in the previous section, we now apply it to the reconstruction of baryonic properties, namely the bias-weighted electron pressure $\langle bP_e \rangle$ and SFR density $\langle b\rho_{\text{SFR}} \rangle$. The two aims of this section will be as follows:

- To test the consistency of the estimated $\langle bU \rangle$ across different galaxy samples, and its agreement with the underlying true value (see below). We will do this in Sections 4.2.1 and 4.2.2.
- To study the possibility of interpreting the recovered values of $\langle bU \rangle$ from first principles, using the halo model, and to quantify the limitations of this approach. This will be discussed in Section 4.2.3.

To assess the unbiasedness of $\langle bU \rangle$ estimated from galaxy cross-correlations in Sections 4.2.1 and 4.2.2, we will consider its true underlying value to be that obtained by applying the same estimator to correlations involving the matter overdensity. While this definition is not provided in terms of fundamental ingredients (describing e.g. the detailed relation between matter, gas pressure, and SFR), the value of $\langle bU \rangle$ thus constructed is a well-defined quantity that can be constructed from either a simulation or a sufficiently accurate model of the matter–matter and matter– U power spectra. Section 4.2.3 will then determine the halo model’s accuracy in predicting this quantity.

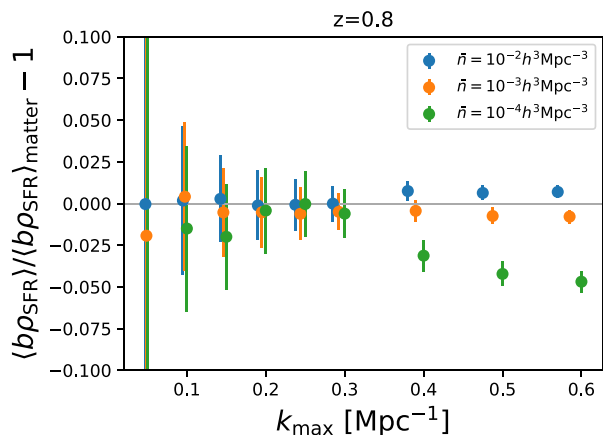


Figure 5. Estimation of $\langle b_{\rho_{\text{SFR}}} \rangle$, using the quadratic model ($n_{\text{max}} = 1$), as a function of k_{max} for two different redshifts ($z = 0, 1.5$). We study galaxies selected in order of decreasing stellar mass, with the different colours (red, blue, and green) corresponding to different number densities, namely $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$.

Based on the results shown in the previous section, in what follows we will use our tomographic estimator with $n_{\text{max}} = 1$ [i.e. including scale-dependent biasing terms proportional to $k^2 \tilde{P}_{\text{mm}}(k)$], employing as template the matter power spectrum measured in FLAMINGO. For this choice, we determined that the estimator remains robust up to scales $k_{\text{max}} \simeq 0.3 \text{Mpc}^{-1}$ in the case of $\langle b_m \rangle$, and we will now study whether this is also the case for $\langle b_P \rangle$ and $\langle b_{\rho_{\text{SFR}}} \rangle$.

4.2.1 Estimating $\langle b_{\rho_{\text{SFR}}} \rangle$

We start by applying the tomographic reconstruction estimator to cross-correlations between the galaxy overdensity and 3D maps of the SFR density (as a proxy for CIB cross-correlations). Fig. 5 shows the relative difference between the value of $\langle b_{\rho_{\text{SFR}}} \rangle$ reconstructed from galaxy cross-correlations and the true value estimated from the cross-correlation with the matter overdensity. Results are presented as a function of the maximum wavenumber included in the analysis k_{max} , for our three galaxy samples, with densities $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$, and at an intermediate redshift $z = 0.8$. We find that our estimator is able to recover a value of $\langle b_{\rho_{\text{SFR}}} \rangle$ that is consistent across galaxy samples and unbiased within 1σ when using scales $k < k_{\text{max}} = 0.3 \text{Mpc}^{-1}$. This regime of validity is similar to that found in the case of $\langle b_m \rangle$, in Section 4.1. Furthermore, in this range of scales, we are able to obtain an unbiased estimate of $\langle b_{\rho_{\text{SFR}}} \rangle$, accurate at the ~ 1 per cent level, within 1σ for the statistical uncertainties corresponding to the FLAMINGO simulation and, more importantly, much smaller than the statistical uncertainties in real data (see Fig. 9 and B. Jogo et al. 2023b).

This result is also consistent across different redshifts. Fig. 6 shows the value of $\langle b_{\rho_{\text{SFR}}} \rangle$ found using our fiducial estimator with $k_{\text{max}} = 0.3 \text{Mpc}^{-1}$ for different FLAMINGO snapshots, spanning the redshift range $z < 1.5$. The results are shown for the three different galaxy samples (blue, green, and orange colours), as well as the true value estimated from the cross-correlation with the matter overdensity (in black). The estimates of $\langle b_{\rho_{\text{SFR}}} \rangle$ from different galaxy samples agree with each other and with the truth across all redshifts, with per cent-level accuracy. The mean relative deviation is 1.2 per cent, with the largest deviation, corresponding to $\bar{n} = 10^{-4} h^3 \text{Mpc}^{-3}$ sample at the highest redshift being 3 per cent.

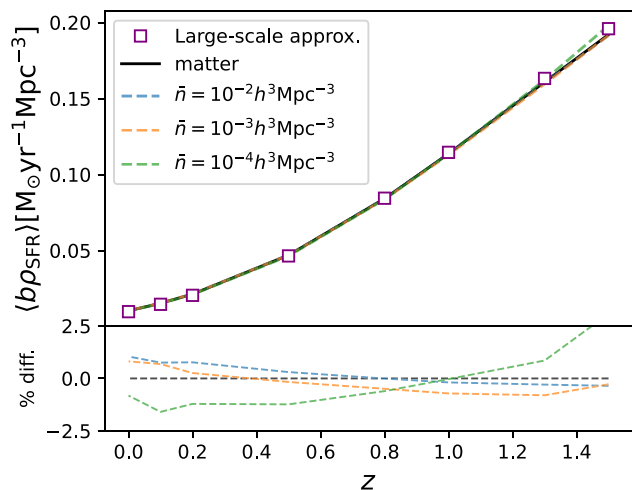


Figure 6. Best-fitting results of $\langle b_{\rho_{\text{SFR}}} \rangle$ from the quadratic model ($n_{\text{max}} = 1$) as a function of redshift, with $P(k, z) = P_{\text{mm, sim}}$, at scales $k_{\text{max}} = 0.3 \text{Mpc}^{-1}$. We study galaxies selected in order of decreasing stellar mass, with the different colours (blue, orange, and green) corresponding to different number densities, namely $\bar{n} = \{10^{-2}, 10^{-3}, 10^{-4}\} h^3 \text{Mpc}^{-3}$. We compare the results obtained from these samples with those found using the matter overdensity. Purple squares correspond to values computed from the LSA model, using scales $k < 0.15 \text{Mpc}^{-1}$. The *bottom panel* shows the percent relative difference between the values recovered from the matter field and the different galaxy samples.

For context, Fig. 6 shows the result of estimating $\langle b_{\rho_{\text{SFR}}} \rangle$ using the LSA of equation (15). In this case, we measure $P_{\text{mm}}(k)$ and $P_{m_{\rho_{\text{SFR}}}}(k)$ from the simulation, take the ratio of both measurements, and estimate $\langle b_{\rho_{\text{SFR}}} \rangle_{\text{LSA}}$ as the average of this ratio on scales $k < 0.15 \text{Mpc}^{-1}$. We see that, in this case, the LSA is in fact an excellent approximation to the value of $\langle b_{\rho_{\text{SFR}}} \rangle$ calculated through our fiducial estimator. This is not a trivial statement, as scale-dependent and stochastic biasing in the SFR density field could induce a bias in the value of $\langle b_{\rho_{\text{SFR}}} \rangle$ inferred through the LSA estimator. This result is not necessarily surprising: the SFR density is dominated by low-mass haloes, where these effects (scale-dependent bias, non-Poissonian stochasticity) are likely smaller than in large-mass haloes (N. Kokron et al. 2022).

4.2.2 Estimating $\langle b_P \rangle$

We repeat the same procedure on cross-correlations involving maps of the thermal electron pressure, constructed as described in Section 3.1, in order to estimate $\langle b_P \rangle$. This is interesting, not only because this is a quantity that can be measured from cross-correlations with maps of the tSZ effect, but also because inhomogeneities in the thermal gas pressure and the SFR density have markedly different clustering properties that could affect the performance of the tomographic reconstruction estimator. While star formation efficiency peaks at intermediate halo masses [$\log(M/M_{\odot}) \sim 12$; B. P. Moster, T. Naab & S. D. M. White 2018; A. Maniyan et al. 2021; B. Jogo et al. 2023a], thermal energy scales strongly with halo mass, and thus thermal pressure maps are dominated by the contribution of a small number of massive haloes (B. Bolliet 2018). Both quantities thus probe different cosmic density environments, affecting their effective scale-dependent and stochastic bias relation.

In Fig. 7, we show the relative difference between the values of $\langle b_P \rangle$ recovered from our fiducial estimator for the three different galaxy samples, and the true value recovered using the matter instead

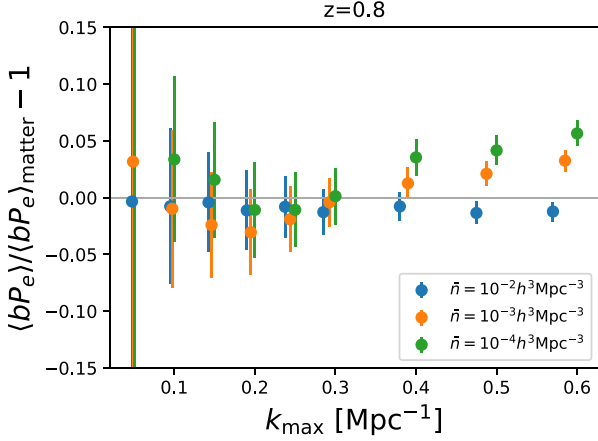


Figure 7. Same as Fig. 5, but for $\langle bP_e \rangle$.

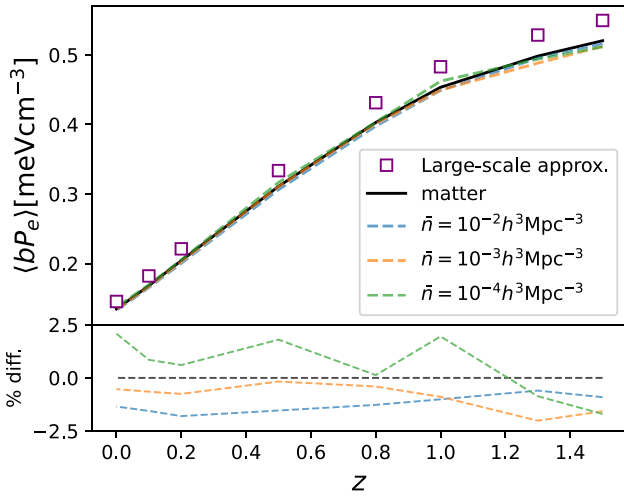


Figure 8. Same as Fig. 6, but for $\langle bP_e \rangle$.

of galaxy overdensity. The results are shown as a function of k_{\max} for the snapshot at redshift $z = 0.8$. As in the case of $\langle b_m \rangle$ and $\langle b\rho_{\text{SFR}} \rangle$, we find that the estimator is robust against non-linear galaxy biasing up to scales $k > 0.3 \text{ Mpc}^{-1}$, beyond which the estimates from different galaxy samples start to differ significantly. It is interesting to note that the relative statistical uncertainties of the $\langle bP_e \rangle$ measurements are noticeably larger than those of $\langle b\rho_{\text{SFR}} \rangle$ (see Fig. 5). This may be a consequence of the aforementioned dominant contribution to P_e from sparse massive haloes, raising the effective shot noise contribution to the cross-spectrum measurement.

This result holds across different redshifts, as demonstrated by Fig. 8, which shows the value of $\langle bP_e \rangle$ recovered for the different galaxy samples, as well as the true value, for the different redshift snapshots studied here, using scales up to $k_{\max} = 0.3$. The average relative difference with respect to the truth across all the recovered values is approximately 1.2 per cent, with the largest difference being 2 per cent, again corresponding to the $\bar{n} = 10^{-4} h^3 \text{ Mpc}^{-3}$ sample at the highest redshift probed. As in the previous section, we also show the prediction based on the LSA, calculated as described above. In this case, however, we find that the LSA significantly overpredicts the value of $\langle bP_e \rangle$ at all redshifts. This is not entirely surprising: as mentioned above, the strong influence of massive haloes in the pressure map means that the matter-pressure cross-correlation receives a large 1-halo contribution (see e.g. fig. 2 of A. J. Mead

et al. 2020), which acts as an effective shot noise-like term on large scales. Unlike our estimator, the LSA does not account for this component, leading to an overestimation of the true large-scale amplitude governed by $\langle bP_e \rangle$.

4.2.3 Theoretical interpretation of tomographic reconstruction

So far, we have compared our estimator of $\langle bU \rangle$ using galaxy cross-correlation against its true value, having defined the latter as ‘the result of applying our estimator to correlations involving the matter overdensity’. Although this is a well-defined quantity, requiring only a physical model for the $P_{\text{mm}}(k)$ and $P_{\text{m}U}(k)$ power spectra, in terms of both cosmology and the astrophysics of the U field, it is interesting to explore whether a more direct theoretical prediction of $\langle bU \rangle$ can be constructed, independent of the specific definition of our estimator.

This can potentially be achieved via the halo model, which predicts $\langle bU \rangle$ to be given by equation (11). The halo model prediction depends on three quantities, all of which are, to some extent, uncertain:

- (i) The HMF $n(M)$, which depends not only on the chosen halo mass definition, but also on the specific halo finder used to estimate it from simulations (A. Knebe et al. 2011; S. Maleubre et al. 2024; V. J. Forouhar Moreno et al. 2025).
- (ii) The linear halo bias $b_h(M)$, which compounds the same sources of uncertainty as $n(M)$ in addition to potential departures from the predictions of the PBS, or the specific method used to estimate it in simulations.
- (iii) A mass-observable relation describing $\bar{U}(M)$ (defined in equation 12). In most cases, this will involve a model for the astrophysics of U , which should be sufficiently accurate or flexible, but will also be affected by uncertainties in the choice of halo definition and its boundaries.

We would like to separate the potential inaccuracies of the halo model prediction caused by uncertainties in the $n(M)$ and $b_h(M)$ parametrizations, as well as those due to incomplete astrophysical models describing $\bar{U}(M)$, from those arising from non-linear halo biasing and assembly bias. While the former sources of error may be addressed by a more careful design of the models describing these quantities, the latter point to problems in the core assumptions of the vanilla halo model that would require the introduction of new ingredients. To that end, we will aim to use the abundance, linear clustering bias, and mass-observable relations measured directly in FLAMINGO, instead of choosing any particular parametrization. We will achieve this by writing equation (11) as a sum over the contribution of all haloes in the simulation, as given in equation (21), with the halo bias b_h measured directly from the simulation as described in Section 3.2.

Fig. 9 shows the halo model prediction for $\langle bP_e \rangle$ and $\langle b\rho_{\text{SFR}} \rangle$ (top and bottom panels, respectively) as a function of redshift, together with the measurements of $\langle bU \rangle$ extracted using our fiducial estimator with $k_{\max} = 0.3 \text{ Mpc}^{-1}$ (results are shown for the matter density). In the case of $\langle b\rho_{\text{SFR}} \rangle$, we see that the halo model using the halo bias measured directly in FLAMINGO is able to accurately predict our measurements at all redshifts. The largest relative difference with respect to the measurements is < 2 per cent at the highest redshift. In the case of $\langle bP_e \rangle$, we see that the halo model is again precise at low redshifts, but becomes increasingly inaccurate beyond $z \simeq 0.6$, consistently overpredicting the measurements.

As we discussed in Section 2.3 (see also I. G. McCarthy et al. 2017), a significant contribution to the pressure map comes from gas in the outskirts of haloes, which motivated us to calculate the

total pressure integrated over the halo volume defined as a sphere with radius $R = 5 \times R_{500}$. However, if the density of haloes that significantly contribute to the pressure map is sufficiently high, this procedure has the risk of double-counting the contribution from simulation particles simultaneously assigned to different haloes within their $5 \times R_{500}$ spheres. The total comoving volume occupied by these spheres at $z = 1.5$ is approximately three times larger than at $z = 0$, significantly increasing the fraction of doubly counted particles. We verified that using smaller radii to calculate the volume-integrated pressure had a strong effect on the result, particularly at high redshifts. We thus conclude that the definition of halo boundaries may be a major source of uncertainty for the halo model predictions of $\langle bP_e \rangle$, particularly at high redshifts.

It is also interesting to place these differences in the context of the theoretical uncertainties associated with the halo model when using existing parametrizations to model its different ingredients. As a way to quantify this, Fig. 9 shows the halo model predictions obtained by replacing the halo bias measured directly in FLAMINGO with various fitting functions. Namely, we use the parametrization of b_h presented in J. L. Tinker et al. (2010), as well as the halo bias derived from the mass function parametrizations of G. Despali et al. (2016) and L. Ondaro-Mallea et al. (2024), assuming a relation between both quantities given by the PBS (H. J. Mo & S. D. M. White 1996). We see that using different bias parametrizations leads to differences in the prediction of $\langle bU \rangle$ that are larger than the differences between our fiducial halo model predictions and our measurements. A more detailed discussion of the differences between these halo bias prescriptions is presented in Appendix B. Therefore, the inaccuracy of the halo model in predicting $\langle bU \rangle$ is likely dominated by the theoretical uncertainty in existing parametrizations of the halo model ingredients, rather than in the core assumptions of the halo model itself.

Finally, Fig. 9 also shows some of the existing measurements of $\langle b\rho_{\text{SFR}} \rangle$ from B. Jago et al. (2023a), and $\langle bP_e \rangle$ from N. Koukoufilippas et al. (2020) and Y.-K. Chiang et al. (2020). The FLAMINGO predictions for both of these quantities appear to be in qualitatively good agreement with these measurements. Although a comparison of $\langle bP_e \rangle$ measurements against predictions from hydrodynamical simulations was presented in Z. Chen et al. (2024), this is the first time, to our knowledge, that a similar comparison has been made for $\langle b\rho_{\text{SFR}} \rangle$. A more detailed study of the predictions from other simulations in the FLAMINGO suite, as well as other simulations and theoretical frameworks, against existing measurements could therefore be used to place constraints on both cosmology and the astrophysics of galaxy formation and the intergalactic medium. Once applied to real data, the methodology proposed and validated in this work will ensure that any measurements used in this comparison will be robust against uncertainties in the clustering properties of the galaxies used to extract them.

5 CONCLUSIONS

Tomographic reconstruction, the analysis of the cross-correlation between galaxies and projected tracers of other LSS probes, in conjunction with the autocorrelation of these galaxies, can be used to obtain precise measurement of the cosmic average of key physical properties weighted by the halo bias. The resulting measurements can be used to constrain both cosmological parameters and models describing the baryonic components of the cosmic density field. In this paper, we have investigated the possibility of designing a tomographic reconstruction estimator that is robust against the complex clustering properties of the galaxies used in the analysis, as

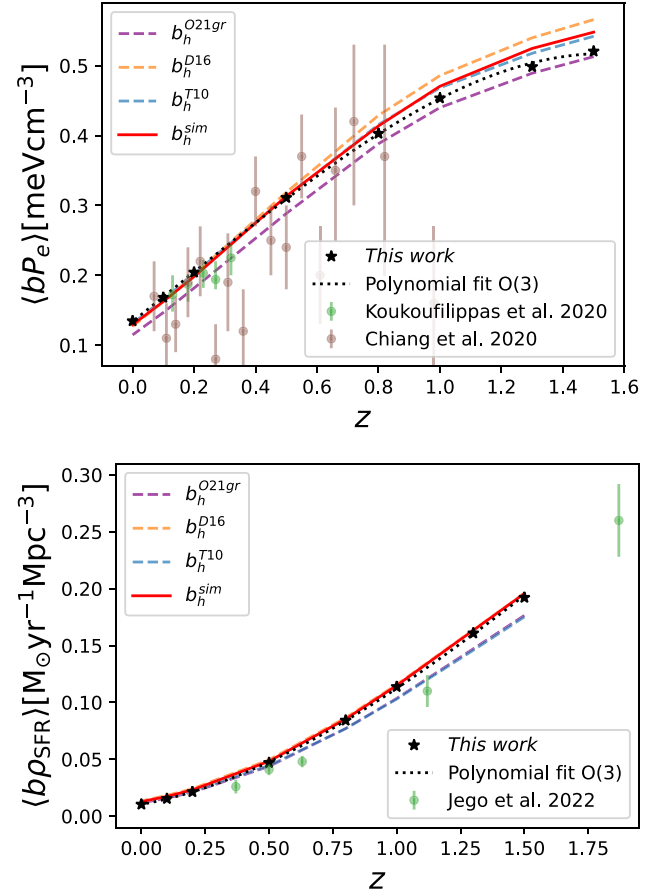


Figure 9. Comparison of $\langle bU \rangle = \langle bP_e \rangle$ and $\langle b\rho_{\text{SFR}} \rangle$ as a function of redshift for galaxies (black) and their polynomial linear fit (black dotted line), halo bias-weighted models: J. L. Tinker et al. (2010) – blue, G. Despali et al. (2016) – orange, and L. Ondaro-Mallea et al. (2024) – purple, and different real data (see refs in legend).

well as some of the challenges that must be tackled when interpreting the resulting measurements.

We have demonstrated that such an estimator can in fact be constructed, able to recover unbiased estimates of $\langle bU \rangle$ with per cent-level accuracy. The method consists of modelling both the galaxy autocorrelation and its cross-correlation with the tracer U as a linear combination of three terms: a linearly biased component, a white noise-like component, and a scale-dependent bias component proportional to k^2 . The amplitudes of the linearly biased contributions for both $P_{\text{gg}}(k)$ and $P_{\text{g}U}(k)$, extracted using only scales $k < k_{\text{max}} = 0.3 \text{ Mpc}^{-1}$, can then be combined to obtain a reliable measurement of $\langle bU \rangle$ that is independent of the small-scale clustering of the galaxies used. We have also shown that the recovered constraints on $\langle bU \rangle$ are robust against uncertainties in the template used to describe the matter power spectrum, and simple parametrizations such as HALOFIT are sufficient to obtain accurate measurements. Furthermore, we have demonstrated that the dependence of the estimated $\langle bU \rangle$ on the cosmological model chosen to generate the matter power spectrum template can be taken into account through a simple rescaling of the measurements (see Section 4.1.3).

Applying this estimator to cross-correlations between galaxies and maps of the electron thermal pressure and the SFR density, we have shown that unbiased measurements of $\langle bP_e \rangle$ and $\langle b\rho_{\text{SFR}} \rangle$ can be obtained. Furthermore, we have shown that, while these measurements

can be accurately interpreted within the halo model framework, the reliability of the halo model predictions depends critically on the accuracy with which its ingredients (namely halo abundances, linear halo bias, and a physical model describing the halo-observable relation) may be parametrized. Alternatively, accurate predictions may be obtained given a model for the matter–matter and matter– U power spectrum, derived from either simulations or first principles.

The results presented here should be broadly applicable to the tomographic reconstruction of other astrophysical quantities, such as the cosmic radio (E. Todarello et al. 2024), infrared (Y.-K. Chiang et al. 2025), ultraviolet (Y.-K. Chiang et al. 2019), gamma-ray (A. Paopiamsap et al. 2024), neutrino (A. Ouellette & G. Holder 2024), and gravitational wave (D. Alonso et al. 2024) backgrounds. This estimator could also be applied to the study of CMB lensing tomography, as an alternative approach to reconstruct the growth history independently of the small-scale galaxy bias. Finally, the same principles could be applied within the clustering redshifts approach, using the cross-correlation of spectroscopic and photometric galaxies, to reconstruct the bias-weighted redshift distribution of the latter.

ACKNOWLEDGEMENTS

We would like to thank Raul Angulo, and Jaime Salcido for useful comments. SM, MZ, and DA acknowledge support from the Beecroft Trust.

This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

This work was also supported by the Science and Technology Facilities Council (grant number ST/Y002733/1).

DATA AVAILABILITY

Any power spectrum measurement used for this paper and not already available in the FLAMINGO data release⁵ can be provided upon request.

REFERENCES

- Abbott T. M. C. et al., 2022, *Phys. Rev. D*, 105, 023520
- Alonso D., Nikjoo M., Renzini A. I., Bellini E., Ferreira P. G., 2024, *Phys. Rev. D*, 110, 103544
- Angulo R. E., Baugh C. M., Frenk C. S., Lacey C. G., 2008, *MNRAS*, 383, 755
- Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2021, *MNRAS*, 507, 5869
- Bautista J. E. et al., 2021, *MNRAS*, 500, 736
- Bolliet B., 2018, preprint ([arXiv:1806.04786](https://arxiv.org/abs/1806.04786))
- Borrow J., Schaller M., Bower R. G., Schaye J., 2022, *MNRAS*, 511, 2367
- Braspenning J. et al., 2024, *MNRAS*, 533, 2656
- Carlstrom J. E., Holder G. P., Reese E. D., 2002, *ARA&A*, 40, 643
- Chen Z., Zhang P., Yang X., 2023, *ApJ*, 953, 188
- Chen Z. et al., 2024, *Phys. Rev. D*, 109, 063513
- Chiang Y.-K., Ménard B., Schiminovich D., 2019, *ApJ*, 877, 150
- Chiang Y.-K., Makiya R., Ménard B., Komatsu E., 2020, *ApJ*, 902, 56
- Chiang Y.-K., Makiya R., Ménard B., 2025, *ApJ*, 992, 65
- Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1
- Desjacques V., Jeong D., Schmidt F., 2018, *Phys. Rep.*, 733, 1
- Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, *MNRAS*, 456, 2486
- Forouhar Moreno V. J., Helly J., McGibbon R., Schaye J., Schaller M., Han J., Kugel R., Bahé Y. M., 2025, *MNRAS*, 543, 1339
- Gatti M. et al., 2022, *MNRAS*, 510, 1223
- Han J., Cole S., Frenk C. S., Benítez-Llambay A., Helly J., 2018, *MNRAS*, 474, 604
- Jego B., Alonso D., García-García C., Ruiz-Zapatero J., 2023a, *MNRAS*, 520, 583
- Jego B., Ruiz-Zapatero J., García-García C., Koukoufilippas N., Alonso D., 2023b, *MNRAS*, 520, 1895
- Kennicutt R. C., Jr, 1998, *ARA&A*, 36, 189
- Knebe A. et al., 2011, *MNRAS*, 415, 2293
- Kokron N., DeRose J., Chen S.-F., White M., Wechsler R. H., 2022, *MNRAS*, 514, 2198
- Koukoufilippas N., Alonso D., Bilicki M., Peacock J. A., 2020, *MNRAS*, 491, 5464
- Krolewski A., Ferraro S., White M., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 028
- Kugel R. et al., 2023, *MNRAS*, 526, 6103
- Lewis A., Challinor A., 2006, *Phys. Rep.*, 429, 1
- McCarthy I. G., Brun A. M. C. L., Schaye J., Holder G. P., 2014, *MNRAS*, 440, 3645
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, *MNRAS*, 465, 2936
- McQuinn M., White M., 2013, *MNRAS*, 433, 2857
- Maleubre S., Eisenstein D. J., Garrison L. H., Joyce M., 2024, *MNRAS*, 527, 5603
- Maniyar A., Béthermin M., Lagache G., 2021, *A&A*, 645, A40
- Matthews D. J., Newman J. A., 2010, *ApJ*, 721, 456
- Mead A. J., Tröster T., Heymans C., Van Waerbeke L., McCarthy I. G., 2020, *A&A*, 641, A130
- Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint ([arXiv:1303.4722](https://arxiv.org/abs/1303.4722))
- Mo H. J., White S. D. M., 1996, *MNRAS*, 282, 347
- Moster B. P., Naab T., White S. D. M., 2018, *MNRAS*, 477, 1822
- Newman J. A., 2008, *ApJ*, 684, 88
- Ondaro-Mallea L., Angulo R. E., Zennaro M., Contreras S., Aricò G., 2024, *MNRAS*, 532, 438
- Ouellette A., Holder G., 2024, *Phys. Rev. D*, 110, 103025
- Pandey S. et al., 2019, *Phys. Rev. D*, 100, 063519
- Paopiamsap A., Alonso D., Bartlett D. J., Bilicki M., 2024, *Phys. Rev. D*, 109, 103517
- Peacock J. A., Smith R. E., 2000, *MNRAS*, 318, 1144
- Piccirilli G., Fabbian G., Alonso D., Storey-Fisher K., Carron J., Lewis A., García-García C., 2024, *J. Cosmol. Astropart. Phys.*, 2024, 012
- Qu F. J. et al., 2025, *Phys. Rev. D*, 111, 103503
- Sánchez J. et al., 2023, *MNRAS*, 522, 3163
- Sato M., Matsubara T., 2011, *Phys. Rev. D*, 84, 043501
- Schaller M. et al., 2024, *MNRAS*, 530, 2378
- Schaller M., Schaye J., Kugel R., Broxterman J. C., van Daalen M. P., 2025, *MNRAS*, 539, 1337
- Schaye J. et al., 2023, *MNRAS*, 526, 4978
- Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, *MNRAS*, 431, 3307
- Seljak U., 2000, *MNRAS*, 318, 203
- Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, *ApJ*, 761, 152
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
- Todarello E. et al., 2024, *MNRAS*, 530, 2994
- Vikram V., Lidz A., Jain B., 2017, *MNRAS*, 467, 2315

⁵https://flamingo.strw.leidenuniv.nl/power_spectra.html

- White M. et al., 2022, *J. Cosmol. Astropart. Phys.*, 2022, 007
 Yan Z., van Waerbeke L., Wright A. H., Bilicki M., Gu S., Hildebrandt H., Maniyar A. S., Tröster T., 2022, *A&A*, 665, A52
 Yan Z., Maniyar A. S., van Waerbeke L., 2024, *J. Cosmol. Astropart. Phys.*, 2024, 058
 Young S., Komatsu E., Dolag K., 2021, *Phys. Rev. D*, 104, 083538
 Zennaro M., Angulo R. E., Contreras S., Pellejero-Ibáñez M., Maion F., 2022, *MNRAS*, 514, 5443
 Zennaro M. et al., 2024, preprint (arXiv:2412.08623)

APPENDIX A: PARAMETER INFERENCE

This appendix describes, in detail, the procedure used to obtain estimates of $\langle bU \rangle$ from the auto- and cross-correlation measurements extracted from the simulations.

Let $\mathbf{d} \equiv (\hat{P}_{\text{gg}}(k), \hat{P}_{\text{gU}}(k))$ be a vector containing the measurements of the galaxy autospectrum and the cross-spectrum with another LSS tracer U . We will assume these measurements to have been made in a discrete set of k bins, \vec{k} covering a range of scales $k_{\text{min}} < k_i < k_{\text{max}}$. Let \mathbf{C} be the covariance matrix of these measurements. For simplicity, we will estimate \mathbf{C} to be diagonal across different k bins, and to take the form

$$\text{Cov}(\hat{P}_{wx}(k_i), \hat{P}_{yz}(k_j)) = \frac{\delta_{ij}}{N_k} [P_{wy}(k_i)P_{xz}(k_i) + P_{wz}(k_i)P_{xy}(k_i)], \quad (\text{A1})$$

where N_k is the number of independent Fourier modes over which the fields involved have been measured that fall within the corresponding k bin. This estimate of the statistical uncertainties does not represent the actual scatter of the power spectrum measurements made from the FLAMINGO simulations. On the one hand, the expression above is only accurate for Gaussian fields. On the other hand, the initial conditions of the fiducial FLAMINGO simulation used here were generated with fixed amplitudes for modes with $k < 0.05 \text{ Mpc}^{-1}$ to reduce the impact of large-scale cosmic variance. Nevertheless, although the estimate above is likely an underestimation of the true statistical uncertainties, it should provide a sufficiently accurate scheme to weight the different k modes entering the estimator, as well as the level of correlation between the different power spectra.

To extract constraints on the free parameters of our model $\vec{\theta}$, we may define the Gaussian log-likelihood

$$-2 \log p(\mathbf{d}|\vec{\theta}) = \chi^2(\vec{\theta}) \equiv (\mathbf{d} - \mathbf{t}(\vec{\theta}))^T \mathbf{C}^{-1} (\mathbf{d} - \mathbf{t}(\vec{\theta})). \quad (\text{A2})$$

Here, $\mathbf{t}(\vec{\theta})$ is the theoretical prediction. Inspecting equations (19) and (20), we may write \mathbf{t} as

$$\mathbf{t}(\vec{\theta}) = \mathbf{T} \vec{\theta}, \quad (\text{A3})$$

where the parameter vector is

$$\vec{\theta} = (N_{\text{gg}}, N_{\text{gU}}, A_{\text{gg}}^0, A_{\text{gU}}^0, A_{\text{gg}}^1, A_{\text{gU}}^1, \dots), \quad (\text{A4})$$

and the template matrix \mathbf{T} takes the form

$$\mathbf{T} = \begin{pmatrix} \vec{1} & \vec{0} & \vec{P}(\vec{k}) & \vec{0} & \vec{k}^2 \vec{P}(\vec{k}) & \vec{0} & \dots \\ \vec{0} & \vec{1} & \vec{0} & \vec{P}(\vec{k}) & \vec{0} & \vec{k}^2 \vec{P}(\vec{k}) & \dots \end{pmatrix}, \quad (\text{A5})$$

where $\vec{1}$ and $\vec{0}$ are arrays of ones and zeros with the same size as \vec{k} , and $\vec{P}(\vec{k})$ and $\vec{k}^2 \vec{P}(\vec{k})$ are arrays containing the values of the power spectrum template $\hat{P}_{\text{mm}}(k)$, and of $k^2 \hat{P}_{\text{mm}}(k)$ in each k bin. Thus, the first two columns of \mathbf{T} correspond to the white noise

components of \hat{P}_{gg} and \hat{P}_{gU} . The next two columns correspond to the linear bias contributions, with $A_{\text{gg}}^0 = b_{\text{g}}^2$ and $A_{\text{gU}}^0 = b_{\text{g}} \langle bU \rangle$. All subsequent columns correspond to the polynomial expansion terms in equations (19) and (20), aimed at absorbing the impact of non-linear biasing.

Since the log-likelihood is a quadratic function of the parameters, we can obtain analytical estimates of the best-fitting parameter values $\vec{\theta}_{\text{BF}}$ and their statistical uncertainties [i.e. their covariance $\mathbf{C}_{\theta} \equiv \langle (\vec{\theta} - \vec{\theta}_{\text{BF}})(\vec{\theta} - \vec{\theta}_{\text{BF}})^T \rangle$]. These are given by

$$\mathbf{C}_{\theta}^{-1} = \mathbf{T}^T \mathbf{C}^{-1} \mathbf{T}, \quad \vec{\theta}_{\text{BF}} = \mathbf{C}_{\theta} \mathbf{T}^T \mathbf{C}^{-1} \mathbf{d}. \quad (\text{A6})$$

Once $\vec{\theta}_{\text{BF}}$ is calculated, the values of b_{g} and $\langle bU \rangle$ can be estimated from the best-fitting amplitudes A_{gg}^0 and A_{gU}^0 as

$$b_{\text{g}} = \sqrt{A_{\text{gg}}^0}, \quad \langle bU \rangle = \frac{A_{\text{gU}}^0}{\sqrt{A_{\text{gg}}^0}}. \quad (\text{A7})$$

The uncertainties on the measured b_{g} and $\langle bU \rangle$ can also be calculated analytically in terms of the covariance of $\vec{\theta}$, using linearized error propagation, assuming that the relative error of the estimated b_{g} is small:

$$\text{Var}(b_{\text{g}}) = \frac{\text{Var}(A_{\text{gg}}^0)}{4A_{\text{gg}}^0}, \quad (\text{A8})$$

$$\begin{aligned} \text{Var}(\langle bU \rangle) &= \frac{\text{Var}(A_{\text{gU}}^0)}{(A_{\text{gg}}^0)^2} + \frac{(A_{\text{gU}}^0)^2}{4(A_{\text{gg}}^0)^4} \text{Var}(A_{\text{gg}}^0) \\ &\quad - \frac{A_{\text{gU}}^0}{(A_{\text{gg}}^0)^3} \text{Cov}(A_{\text{gg}}^0, A_{\text{gU}}^0), \end{aligned} \quad (\text{A9})$$

$$\text{Cov}(b_{\text{g}}, \langle bU \rangle) = \frac{\text{Cov}(A_{\text{gg}}^0, A_{\text{gU}}^0)}{2A_{\text{gg}}^0} - \frac{A_{\text{gU}}^0}{4(A_{\text{gg}}^0)^2} \text{Var}(A_{\text{gg}}^0). \quad (\text{A10})$$

This approximation can be easily avoided if necessary. To do so, we may simply treat the measured values of A_{gg}^0 and A_{gU}^0 as data, fitting them to a model with free parameters b_{g} and $\langle bU \rangle$ using a Gaussian likelihood with a covariance matrix constructed from \mathbf{C}_{θ} . Since this is a simple two-dimensional model, the resulting likelihood is easy to sample efficiently through a variety of methods. We find that the approximate expressions above are sufficiently accurate for all the cases explored in this paper.

APPENDIX B: IMPORTANCE OF THE ACCURACY OF b_{h}

This appendix contains extra figures showing the differences between halo-bias models as a function of redshift and halo-mass bins explored in this paper, in connection with the results in Section 3.2.

Fig. B1 shows the differences in the value of the halo bias $\langle b_{\text{h}} \rangle$, estimated from three different models in the literature (J. L. Tinker et al. 2010; G. Despali et al. 2016; L. Ondaro-Mallea et al. 2024) as well as measured directly from the simulation, for the five different halo-mass bins explored in this paper. We can see that there is no consistent result among them, and that their variations are in line with those observed in Fig. 9. Given the fact that we cannot really justify the use of one model over the other, we decided to apply our method for estimating $\langle bU \rangle$ to $\langle b_{\text{h}} \rangle$, and use the measured bias from the simulation directly in our analysis. As seen in the main text, this is also the prescription leading to the best results.

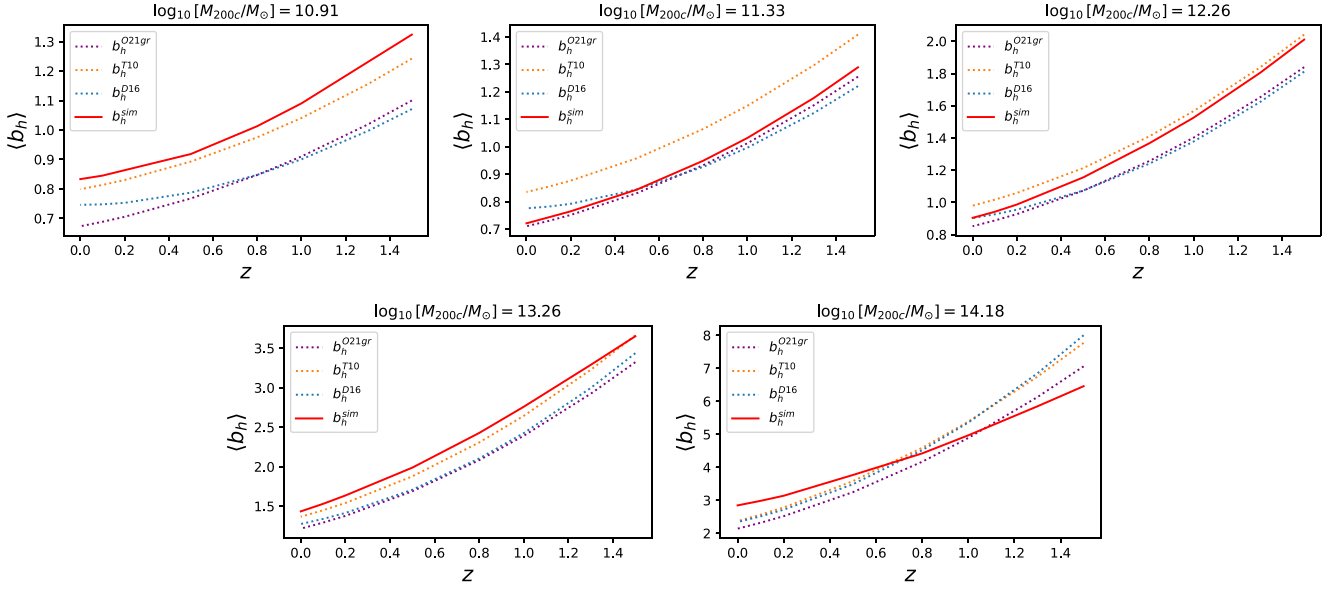


Figure B1. Halo-bias results from J. L. Tinker et al. (2010) – blue, G. Despali et al. (2016) – orange, and L. Ondaro-Mallea et al. (2024) – purple, as well as the result estimated directly from the simulation following an adapted version of equations (19) and (20) – red. We show the values of $\langle b_h \rangle$ as a function of redshift for the five mass bins studied in the main text (and label with the value of the median mass of each bin).

This paper has been typeset from a \LaTeX file prepared by the author.