

The star-formation history in the last 10 billion years from CIB cross-correlations

Baptiste Jego^{1,2*}, Jaime Ruiz-Zapatero², Carlos García-García², Nick Koukoufilippas², and Dav

¹ENS Paris-Saclay, Gif-sur-Yvette, France

²Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The Cosmic Infrared Background (CIB) traces the emission of star-forming galaxies throughout all cosmic epochs. Breaking down the contribution from galaxies at different redshifts to the observed CIB maps would allow us to probe the history of star formation. In this paper, we cross-correlate maps of the CIB with galaxy samples covering the range $z \lesssim 2$ to measure the bias-weighted star-formation rate (SFR) density $\langle b\rho_{\text{SFR}} \rangle$ as a function of time in a model independent way. This quantity is complementary to direct measurements of the SFR density ρ_{SFR} , giving a higher weight to more massive haloes, and thus provides additional information to constrain the physical properties of star formation. Using cross-correlations of the CIB with galaxies from the DESI Legacy Survey and the extended Baryon Oscillation Spectroscopic Survey, we obtain high signal-to-noise ratio measurements of $\langle b\rho_{\text{SFR}} \rangle$, which we then use to place constraints on halo-based models of the star-formation history. We fit halo-based SFR models to our data and compare the recovered ρ_{SFR} with direct measurements of this quantity. We find a qualitatively good agreement between both independent datasets, although the details depend on the specific halo model assumed. This constitutes a useful robustness test for the physical interpretation of the CIB, and reinforces the role of CIB maps as valuable astrophysical probes of the large-scale structure. We report our measurements of $\langle b\rho_{\text{SFR}} \rangle$ as well as a thorough account of their statistical uncertainties, which can be used to constrain star-formation models in combination with other data.

Key words: cosmology: large-scale structure of the Universe – galaxies: star formation

1 INTRODUCTION

Dust in star-forming galaxies absorbs the stellar ultraviolet radiation from newly-formed, short-lived, massive stars, obscuring it, and re-emits it in the infrared (IR). The combined diffuse IR emission from all such sources thus constitutes a tracer of star formation (Partridge & Peebles 1967; Knox et al. 2001), and is known as the Cosmic Infrared Background (CIB).

Since its first detection by COBE (Dwek et al. 1998) (also see Puget et al. (1996)), the CIB has been both a nuisance and a blessing for cosmology and astrophysics. On the one hand, it is the most significant extra-galactic foreground affecting observations of the Cosmic Microwave Background (CMB) temperature fluctuations at high frequencies (Planck Collaboration et al. 2011; Dunkley et al. 2013). Unlike other foregrounds, the fact that the IR spectral energy distribution (SED) is not universal across different galaxies and cosmic times, makes the CIB particularly difficult to separate in multi-frequency CMB observations. This complicates not only the cosmological analysis of small-scale CMB data, but also the recovery of unbiased maps of the CIB itself in the presence of Galactic

dust contamination (Planck Collaboration et al. 2014b, 2016; Lenz et al. 2019). Furthermore, since the CIB fluctuations trace the non-Gaussian large-scale structures that develop at low redshifts, it is an important systematic for CMB lensing analyses, which exploit the non-Gaussian modifications caused by the gravitational lensing of the otherwise Gaussian CMB fluctuations (Osborne et al. 2014; van Engelen et al. 2014; Sailer et al. 2021; Darwish et al. 2021a). The CIB is also an important contaminant in studies of the thermal Sunyaev-Zel’dovich effect, given the non-trivial contamination induced, particularly in cross-correlation with high-redshift tracers of the large-scale structure (Alonso et al. 2018; Yan et al. 2019; Tröster et al. 2022). Given the central role that small-scale primary CMB data and CMB lensing will play in cosmology in the next decade (Schmittfull & Seljak 2018; Fang et al. 2022), improving our understanding of the CIB, and developing better models that allow us to quantify its impact on these observations, is of paramount importance.

On the other hand, the CIB is in itself a treasure trove for both astrophysics and cosmology. First, it is a high-sensitivity tracer of the star-formation rate (SFR) density (i.e. the total mass of stars created per unit time and volume) (Dole et al. 2006). Understanding star formation across time and galaxy type is a central piece in the

* E-mail: baptiste.jego@ens-paris-saclay.fr

puzzle of galaxy formation and evolution (Tinsley 1980). Through observations of the infrared luminosity function at high redshifts (Grupponi et al. 2013; Magnelli et al. 2013; Marchetti et al. 2016; Davies et al. 2016a), a qualitative picture has emerged where star formation grows from the epoch of reionisation, peaks at around $z \sim 2$, and then decreases as the gas that fuels it runs out (Madau & Dickinson 2014b). However, collecting large samples of star-forming galaxies at high redshift is observationally challenging. Thus, since the CIB combines the emission of all IR sources since $z \sim 6$, it encodes information about the full history of star formation, making it a useful complementary probe for these studies.

Unfortunately, studies of the star-formation history from maps of the CIB alone (e.g. Shang et al. 2012; Viero et al. 2013; Planck Collaboration et al. 2014b; Maniyar et al. 2018; Maniyar et al. 2021) are hampered by its projected nature. Without additional information it is not possible to disentangle, in a model-independent manner, the contribution to the CIB from sources at different redshifts. Although cross-correlations with CMB lensing can be used to improve constraints on different CIB models (Planck Collaboration et al. 2014a; Maniyar et al. 2018; Cao et al. 2020; Darwish et al. 2021b; McCarthy & Madhavacheril 2021), they suffer from the same problem. This, however, can be solved through the use of *galaxy clustering tomography* (Chiang et al. 2019). Since the sources that give rise to the CIB trace the same large-scale structure as any other galaxy sample (although different galaxy types contribute differently to the CIB), the amplitude of the angular correlation between the CIB and a set of galaxies at a sufficiently narrow range of redshifts will be proportional to the contribution to the CIB map from sources at those redshifts. This technique has been widely utilised in various areas of cosmology and astrophysics to reconstruct the evolution, or the redshift dependence, of different quantities. These applications include recovering the mean gas pressure from maps of the thermal Sunyaev-Zel’dovich effect (Pandey et al. 2019; Koukoufilippas et al. 2020; Chiang et al. 2020; Yan et al. 2021), the cosmic UV background (Chiang et al. 2019), and the growth of structure (García-García et al. 2021; Krolewski et al. 2021), as well as the calibration of redshift distributions for imaging galaxy samples (the so-called “clustering redshifts” method Newman 2008; Ménard et al. 2013; Morrison et al. 2017). Recently Yan et al. (2022) used cross-correlations between CIB maps and galaxies from the Kilo-Degree Survey (KiDS) to constrain a physical model of the CIB (see also Serra et al. (2014); Wang et al. (2015); Chen et al. (2016)).

The aim of this paper is to use cross-correlations of the CIB with samples of galaxies at different redshifts to reconstruct the star-formation history since $z \sim 2$. Our approach will be somewhat different from that of Yan et al. (2022). First, we will show that, from the amplitude of the galaxy auto-correlation and the cross-correlation with the CIB on large scales, robust and model-independent constraints can be placed on the evolution of the *bias-weighted* star-formation rate density $\langle b\rho_{\text{SFR}} \rangle$. In this quantity, the contribution to the total SFR density from each source is weighted by the linear clustering bias of the dark matter halo it resides in. $\langle b\rho_{\text{SFR}} \rangle$ is therefore more sensitive to the contribution from galaxies in massive haloes than direct measurements of the SFR density ρ_{SFR} , and thus provides a complementary probe for the physics of star formation. The model-independent measurements of $\langle b\rho_{\text{SFR}} \rangle$ obtained can then be used to constrain the parameters of a halo-based SFR model. This will allow us to propagate our measurements onto constraints on ρ_{SFR} that can be compared with direct measurements of this quantity from studies of the infrared luminosity function, thus

providing a useful consistency test of the physical models used to describe star formation.

This paper is structured as follows. Section 2 lays out the theoretical framework, based on the halo model, that will allow us to measure $\langle b\rho_{\text{SFR}} \rangle$ as a function of redshift. The galaxy samples and CIB maps used in this analysis are then presented in Section 3. Section 4 presents our model-independent measurements of $\langle b\rho_{\text{SFR}} \rangle$, quantifies their robustness, and propagates them onto constraints on ρ_{SFR} and on two halo-based models for star formation. We then present our conclusions and outlook in Section 5. Throughout the paper we will use natural units, with the speed of light $c = 1$.

2 METHODS

2.1 Theory

2.1.1 Projected probes

Let $u(\hat{\mathbf{n}})$ be a cosmological field projected on the sphere, and assume that it is the radial projection of a 3D field U , i.e.

$$u(\hat{\mathbf{n}}) = \int d\chi q_u(\chi) U(\chi\hat{\mathbf{n}}, z). \quad (1)$$

Here χ is the comoving distance, $q_u(\chi)$ is the radial kernel associated with u , and z is the redshift corresponding to the comoving distance χ . The angular power spectrum of two such quantities, u and v , C_ℓ^{uv} , defined as the covariance between their harmonic-space coefficients, is then related to the power spectrum of the corresponding 3D quantities, $P_{UV}(k, z)$, via a similar radial projection:

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

Here we have used the Limber approximation (Limber 1953), which is appropriate when the radial kernels are much broader than the typical correlation length of the fields (this is true for the quantities studied here).

2.1.2 Halo model

The halo model (Seljak 2000; Peacock & Smith 2000; Cooray & Sheth 2002) is a versatile formalism to model the 3D correlations of cosmological quantities. In this framework, the power spectrum $P_{UV}(k)$ can be written as the sum of two contributions, the one-halo and two-halo terms, $P_{UV}^{1h}(k)$ and $P_{UV}^{2h}(k)$, with

$$P_{UV}^{1h}(k) \equiv \int dM n(M) \langle U(k, M) V(k, M) \rangle. \quad (3)$$

$$P_{UV}^{2h}(k) \equiv \langle bU(k) \rangle \langle bV(k) \rangle P_{\text{lin}}(k). \quad (4)$$

Here, $n(M)$ is the halo mass function (the comoving number density of haloes per mass interval), $U(k, M)$ is the profile of quantity U around a halo of mass M in Fourier space, $P_{\text{lin}}(k)$ is the linear matter power spectrum, and $\langle bU(k) \rangle$ is

$$\langle bU(k) \rangle \equiv \int dM n(M) b_h(M) \langle U(k, M) \rangle, \quad (5)$$

where b_h is the halo bias.

On large scales, both P_{UV}^{1h} and $\langle bU \rangle$ tend to a constant. This can be seen easily by writing $U(k, M)$ in the limit $k \rightarrow 0$ (i.e. when $e^{-i\mathbf{k}\cdot\mathbf{x}} \rightarrow 1$):

$$U(k \rightarrow 0, M) = \int d^3x U(\mathbf{x}, M). \quad (6)$$

Thus, on sufficiently large scales, the power spectrum can be approximated as

$$P_{UV}(k) = \langle bU \rangle \langle bV \rangle P_{\text{lin}}(k) + N_{UV}, \quad (7)$$

where $\langle bU \rangle$, $\langle bV \rangle$ and N_{UV} are now scale-independent quantities. Crucially, $\langle bU \rangle$, has a simple physical interpretation: it is the cosmic mean value of quantity U weighted by the halo bias. N_{UV} absorbs the large-scale contribution of the one-halo term, as well as any additional source of stochastic white noise in the data.

Extracting information from scales smaller than those for which the approximation above is appropriate requires tackling several modelling challenges. First, an accurate model of the scale dependence of the halo profiles (i.e. the distribution of U and V around haloes of different masses) is necessary. Secondly, the 1-halo term depends on the statistical correlation of the two quantities under study, which in principle requires a model of not only the mean profiles, but also of their statistical correlation. Although these are often implicitly assumed to be independent (i.e. $\langle UV \rangle = \langle U \rangle \langle V \rangle$), deviations from this can be highly relevant and lead to incorrect conclusions (particularly given the high precision with which small-scale correlations can be measured). In turn, the 2-halo term requires no knowledge of the joint distribution of the quantities. Finally, the halo model is known to fail at the $\gtrsim 10\%$ level on scales straddling the transition between the 1-halo- and 2-halo-dominated regimes. This is likely caused by the overly simplified assumption of linear biasing used in vanilla versions of the halo model to describe the clustering of dark matter haloes (Mead et al. 2021; Mead & Verde 2021).

Tackling these challenges is a complex problem that can be addressed through more sophisticated astrophysical modelling, often at the cost of introducing new nuisance parameters that must be marginalised over. The obvious benefit of this approach is that it enables the use of smaller scales, potentially achieving tighter constraints on the model parameters. This is indeed a common approach used in previous analyses of CIB auto- and cross-power spectra (e.g.: Maniyar et al. (2018); Yan et al. (2022)).

Here, we choose to avoid this additional complexity, and opt for a robust and model-independent measurement of the phenomenological parameters in Eq. 7, restricting our analysis to large scales. Once the bias-weighted quantities $\langle bU \rangle$ have been measured, they can be used to constrain the underlying physical model describing them (in this case the star-formation rate history). All that remains is designing this model for the two tracers under consideration: the projected galaxy overdensity and the CIB emissivity.

2.1.3 CIB and star formation

The CIB intensity at a given observed frequency ν is related to the infrared emissivity (energy emitted per unit frequency, time and comoving volume in the rest frame of the emitter) $j_{\nu(1+z)}$ via:

$$I_{\nu}(\hat{\mathbf{n}}) = \int d\chi \frac{j_{\nu(1+z)}(\chi \hat{\mathbf{n}}, z)}{4\pi(1+z)}. \quad (8)$$

Given a model for the specific infrared luminosity of haloes of mass M , $L_{\nu}(M)$, the mean comoving emissivity becomes

$$\langle j_{\nu(1+z)} \rangle = \int dM n(M) L_{\nu(1+z)}(M), \quad (9)$$

where $n(M)$ is the halo mass function. We now proceed as in Maniyar et al. (2018); Maniyar et al. (2021). The specific infrared luminosity can be written in terms of the observed source flux S_{ν}

$$L_{\nu(1+z)} = 4\pi\chi^2(1+z)S_{\nu}. \quad (10)$$

At the same time, the total infrared luminosity L_{IR} (integrated over frequencies) is tightly correlated with the star-formation rate of the source:

$$\text{SFR} = K L_{\text{IR}}, \quad (11)$$

where the $K = 10^{-10} M_{\odot} \text{yr}^{-1} L_{\odot}^{-1}$ is the calibration constant between the far infrared luminosity and star-formation rate Kennicutt (1998); Kennicutt & Evans (2012) for a Chabrier initial mass function (Chabrier 2003), and L_{IR} is integrated over the range 8–1000 μm . Combining Eqs. 9, 10, and 11, the CIB intensity can be written as

$$I_{\nu}(\hat{\mathbf{n}}) = \int d\chi \frac{\chi^2 S_{\nu}^{\text{eff}}(z)}{K} \rho_{\text{SFR}}(\chi \hat{\mathbf{n}}, z), \quad (12)$$

where ρ_{SFR} is the star-formation rate density, given by the contribution of all haloes of different masses. $S_{\nu}^{\text{eff}}(z) \equiv S_{\nu}(z)/L_{\text{IR}}$ is the mean flux of sources at redshift z normalised to a total luminosity of L_{\odot} . We use the estimates of $S_{\nu}^{\text{eff}}(z)$ for the three *Planck* frequency channels derived by Béthermin et al. (2013, 2015, 2017), and made available by Maniyar et al. (2021)¹.

Comparing this with Eq. 1, the CIB intensity $I_{\nu}(\hat{\mathbf{n}})$ can therefore be understood as a projected tracer of the star-formation rate density ρ_{SFR} , with a radial kernel given by

$$q_{\nu}(\chi) = \frac{\chi^2 S_{\nu}^{\text{eff}}(z)}{K}. \quad (13)$$

The associated $\langle bU \rangle$ is the mean star-formation rate density weighted by halo bias:

$$\langle b\rho_{\text{SFR}} \rangle \equiv \int dM n(M) b_h(M) \text{SFR}(M), \quad (14)$$

where $\text{SFR}(M)$ is the mean star-formation rate in haloes of mass M . The mean profile, in turn, would provide a measurement of the mean star-formation rate density

$$\rho_{\text{SFR}} = \int dM n(M) \text{SFR}(M). \quad (15)$$

When interpreting the measured values of $\langle b\rho_{\text{SFR}} \rangle$, we will make use of a halo-model parametrisation similar to that of Moster et al. (2018) and Maniyar et al. (2021) to describe the SFR-mass relation. The model splits the total SFR of a given halo into the contribution from central (SFR_c) and satellite galaxies (SFR_s):

$$\text{SFR}(M, z) = \text{SFR}_c(M, z) + \text{SFR}_s(M, z). \quad (16)$$

As haloes accrete gas, part of it is transformed into stars. The model parametrises this process as:

$$\text{SFR}_c(M, z) = \eta(M, z) \text{BAR}(M, z), \quad (17)$$

where BAR is the baryon accretion rate, and η is the efficiency with which those baryons are transformed into stars. The BAR is modelled after Fakhouri et al. (2010) as

$$\text{BAR}(M, z) = \dot{M}_0 \frac{\Omega_b}{\Omega_M} \left(\frac{M}{10^{12} M_{\odot}} \right)^{1.1} (1 + 1.11z) \frac{H(z)}{H_0}, \quad (18)$$

with $\dot{M}_0 = 46.1 M_{\odot} \text{yr}^{-1}$.

We will consider two parametrisations for the efficiency $\eta(M, z)$:

¹ See https://github.com/abhimaniyar/halomodel_cib_tsz_cibxtsz.

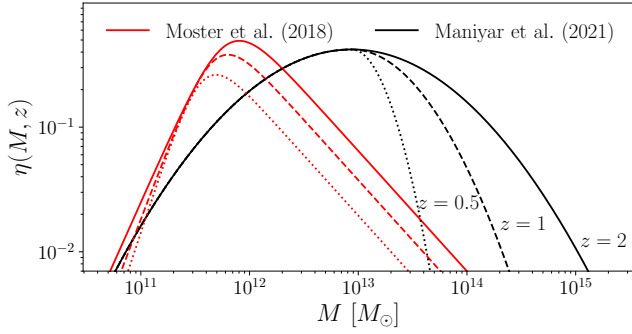


Figure 1. Mass-dependent star-formation efficiency at $z = 0.5$ (dotted lines), $z = 1$ (dashed lines) and $z = 2$ (solid lines). Results are shown for the two halo-based models described in the text: Model A, from Maniyar et al. (2021), in black, and Model B, from Moster et al. (2018), in red. The curves shown correspond to the best-fit parameters found in the two papers ($\log_{10} M_{\max} = 12.94$, $\eta_{\max} = 0.42$, $\sigma_{M,0} = 1.75$, $\tau = 1.17$ for Model A, $m_0 = 11.34$, $m_z = 0.692$, $\eta_0 = 0.005$, $\eta_z = 0.689$ for Model B).

(i) **Model A:** the model of Maniyar et al. (2021). In this case

$$\eta(M, z) = \eta_{\max} \exp \left[-\frac{(\log M_{\max} - \log M)^2}{2\sigma_M^2(M, z)} \right], \quad (19)$$

where M_{\max} is the halo mass of maximum efficiency η_{\max} . The logarithmic width σ_M , describing the range of masses with efficient star formation, depends on both mass and redshift as

$$\sigma_M(M, z) = \sigma_{M,0} - \tau \Theta(M - M_{\max}) \max(0, z_c - z), \quad (20)$$

where Θ is the Heaviside step function.

Fixing $z_c = 1.5$ as in Maniyar et al. (2021), the model is described by 4 free parameters: η_{\max} , $\log_{10} M_{\max}$, $\sigma_{M,0}$, and τ .

(ii) **Model B:** the parametrisation of Moster et al. (2018). In this case:

$$\eta(M, z) = \frac{2\eta_*}{(M_1/M)^\beta + (M/M_1)^\gamma}, \quad (21)$$

where η_* , $m \equiv \log_{10} M_1/M_\odot$, β , and γ are functions of redshift of the form:

$$x(z) = x_0 + x_z \frac{z}{1+z} \quad (22)$$

Fixing $(\beta_0, \beta_z, \gamma_0, \gamma_z)$ to the best-fit values found by Moster et al. (2018), $(3.344, -2.079, 0.966, 0)$, the model is described by 4 free parameters: η_0 , η_z , m_0 , and m_z .

In order to compare this model with Model A, note that the peak efficiency in this case is $\eta_{\max} = \eta_* f_\eta$, and it is attained at a mass $M_{\max} = M_1 f_M$, where

$$f_M = \left(\frac{\beta}{\gamma} \right)^{\frac{1}{\beta+\gamma}}, \quad f_\eta = \frac{2\gamma}{\beta+\gamma} f_M^\beta. \quad (23)$$

For the values of $(\beta_0, \beta_z, \gamma_0, \gamma_z)$ used here, both factors are roughly constant and close to 1 (f_M varies between 1.33 and 1.23, and f_η varies between 1.17 and 1.04, in the range $0 < z < 4$). Thus, to a good approximation, η_* and M_1 provide the peak efficiency and its associated halo mass.

Both models are qualitatively similar in that the star-formation efficiency peaks at a particular halo mass, and decreases for lighter and heavier haloes. This is physically well motivated, based on the effect of low gravitational potentials and supernova feedback at

low masses (Silk 2003), and an increased gas cooling time at high masses (Kereš et al. 2005). There are, however, important differences between them. On the one hand, the peak efficiency in Model A is constant as a function of time, and all time evolution is incorporated in the extent of the high-mass tail of $\eta(M, z)$. On the other, the peak efficiency is explicitly time-dependent in Model B, while the slope of the high-mass tail is time-independent. Furthermore, while the efficiency falls exponentially away from the peak mass in Model A, this drop is slower (power-law) in Model B. These differences are illustrated in Fig. 1, and they will allow us to explore the dependence of our final constraints on ρ_{SFR} from our measurements of $\langle b\rho_{\text{SFR}} \rangle$ on the assumptions of the underlying model.

Finally, the contribution from satellite galaxies will be calculated by adding the contribution of all subhaloes for each parent halo of mass M :

$$\text{SFR}_s(M, z) = \int_{M_{\min}}^M dM_{\text{sub}} \frac{dN}{dM_{\text{sub}}} \text{SFR}_{\text{sat}}(M_{\text{sub}}, M, z), \quad (24)$$

where dN/dM_{sub} is the subhalo mass function, calculated as in Tinker & Wetzel (2010), and $\text{SFR}_{\text{sat}}(M_{\text{sub}}, M, z)$ is the SFR in a satellite galaxy with subhalo mass M_{sub} in a parent halo of mass M . As in Maniyar et al. (2021), we model SFR_{sat} as

$$\text{SFR}_{\text{sat}}(M_{\text{sub}}, M, z) = \text{Min} \left[\text{SFR}_c(M_{\text{sub}}, z), \frac{M_{\text{sub}}}{M} \text{SFR}_c(M, z) \right].$$

2.1.4 Projected galaxy clustering

The projected galaxy overdensity is related to its three-dimensional counterpart via:

$$\delta_g(\hat{\mathbf{n}}) = \int d\chi H(z) p(z) \Delta_g(\chi \hat{\mathbf{n}}, z), \quad (25)$$

where $p(z)$ is the galaxy redshift distribution for the galaxy sample under consideration, and Δ_g is its three-dimensional overdensity. Therefore, the kernel for this tracer is $q_g = H(z) p(z)$. The associated $\langle bU \rangle$ is the galaxy bias b_g .

A popular halo model parametrisation for galaxy clustering is the halo occupation distribution (HOD) (Berlind & Weinberg 2002), based on a model for the number of galaxies as a function of halo mass $\bar{N}_g(M)$. In this framework, b_g would simply be given by

$$b_g = \frac{\int dM n(M) b_h(M) \bar{N}_g(M)}{\int dM n(M) \bar{N}_g(M)}. \quad (26)$$

We will fit directly for the value of b_g , and thus our treatment is insensitive to the details of the $\bar{N}_g(M)$ relation.

The observed overdensity of galaxies is also affected by lensing magnification. Gravitational lensing by foreground structures induces a small perturbation in the observed galaxy positions and in their observed flux. This causes an additive contribution to δ_g of the form

$$\delta_g^M(\hat{\mathbf{n}}) = - \int d\chi (2 - 5s) q_L(\chi) \Delta_M(\chi \hat{\mathbf{n}}, z), \quad (27)$$

where Δ_M is the 3D matter overdensity, $s = d \log_{10} N/dm$ is the logarithmic slope in the distribution of galaxy magnitudes at the limiting magnitude, and q_L is the lensing kernel

$$q_L(\chi) \equiv \frac{3H_0^2 \Omega_m}{2} (1+z) \chi \int_z^\infty dz' p(z') \frac{\chi(z') - \chi}{\chi(z')}. \quad (28)$$

Due to the cumulative nature of the weak lensing kernel, if relevant, the contribution from magnification to the galaxy-CIB cross-correlation would bias the tomographic measurement of $\langle b\rho_{\text{SFR}} \rangle$

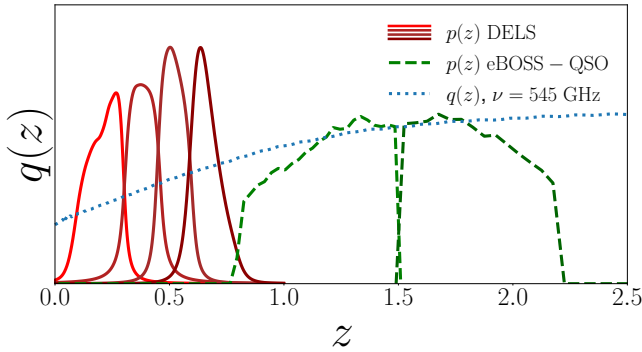


Figure 2. Redshift distributions of the 6 redshift bins used in this work. The red lines show the distribution of the 4 DELS samples, while the dashed green lines show the two eBOSS-QSO redshift bins. The dotted blue line shows the effective radial kernel for the CIB (Eq. 13) at 545 GHz. All curves have been rescaled by an arbitrary factor for visualisation purposes.

described in the next section. To verify that this contribution can indeed be ignored, we estimate the theoretical prediction with and without magnification for the cross-correlation between our highest eBOSS redshift bin (for which magnification is most relevant), and the CIB. We use $s = 0.2$ for quasars (Scranton et al. 2005). We find that magnification changes the cross-correlation only at the sub-percent level on the scales used here, and can therefore be ignored. Note that the impact of magnification on the DELS bins should be significantly smaller than for the quasar samples (see Hang et al. (2021)).

2.1.5 Tomography

A crucial aspect of projected galaxy clustering is its local nature: galaxies at redshift z trace structures at the same redshift. This makes it possible to directly measure the redshift evolution of astrophysical quantities by cross-correlating a given projected tracer (in our case CIB intensity maps) with galaxy samples at different redshifts, a procedure commonly called “tomography” (Chiang et al. 2019).

Consider the Limber equation (Eq. 2), together with the simplified power spectrum model of Eq. 7, and the radial kernels for the CIB (Eq. 13) and for galaxy clustering (Eq. 25). Assuming that the radial/redshift dependence of b_g , $\langle b\rho_{\text{SFR}} \rangle$, and $q_\nu(z)$ is slow compared with the width of the galaxy kernel $q_g(z)$, the auto-correlation of galaxies and their cross-correlation with the CIB can be approximated as

$$C_\ell^{gg} \simeq b_g^2 M_\ell^{gg} + n_{gg}, \quad C_\ell^{g\nu} \simeq b_g \langle b\rho_{\text{SFR}} \rangle M_\ell^{g\nu} + n_{g\nu}, \quad (29)$$

where b_g and $\langle b\rho_{\text{SFR}} \rangle$ are evaluated at the mean redshift of the galaxy sample, and we have defined the power spectrum templates

$$M_\ell^{u\nu} \equiv \int \frac{d\chi}{\chi^2} q_u(\chi) q_\nu(\chi) P_M \left(\frac{\ell + 1/2}{\chi}, z \right), \quad (30)$$

$$n_{uv} \equiv \int \frac{d\chi}{\chi^2} q_u(\chi) q_\nu(\chi) N_{UV}(z), \quad (31)$$

where $P_M(k, z)$ is the matter power spectrum, and N_{UV} was defined in Eq. 7. Assuming a fixed cosmology, and an accurate knowledge of the radial kernels, M_ℓ^{gg} and $M_\ell^{g\nu}$ can be precomputed and treated as fixed templates. In that case, the galaxy auto-correlation effectively constrains the value of b_g , which then can be used to constrain $\langle b\rho_{\text{SFR}} \rangle$ at the redshift of the sample through the cross-correlation.

Figure 2 shows the radial kernels of the different tracers used in this analysis. The CIB kernel for the $\nu = 545$ GHz *Planck* channel, shown as a blue dotted line (with arbitrary normalisation), varies very slowly in comparison with the redshift distributions of the six galaxy samples used, justifying our use of the approximate phenomenological model of Eq. 29 to make tomographic measurements of $\langle b\rho_{\text{SFR}} \rangle^2$.

Fig. 3 shows the cross-power spectrum between the galaxy overdensity and ρ_{SFR} at $z \sim 0.8$. The solid red line shows the prediction using the halo model from Maniyar et al. (2021). This prediction can be recovered to 2% accuracy up to $k = 1 \text{ Mpc}^{-1}$ using the simplified model of Eq. 29, given by the black dash-dotted line. This model is constructed by adding the linear matter power spectrum scaled by the product of the scale-independent biases associated with both quantities (dashed blue line), and the constant large-scale limit of the 1-halo term (black dotted line). However, in order to further improve the model and alleviate any residual theoretical uncertainties associated with the 1-2-halo transition regime, we use the *non-linear* matter power spectrum, computed using the model of Takahashi et al. (2012), scaled by the scale-independent biases (blue dashed line). This effective model is able to reproduce the halo model prediction to the same accuracy on scales $k < 0.2 \text{ Mpc}^{-1}$ without the additional shot noise terms. Therefore, we will by default use the non-linear matter power spectrum as P_M when calculating the templates in Eq. 30. We verified that the results shown in Section 4.2 change by less than 0.25σ if we use the linear power spectrum instead. Thus, the conservative choice of small-scale cut $k_{\text{max}} = 0.15 \text{ Mpc}^{-1}$ used in our analysis (see Section 4), marked as a vertical dotted line in the figure, should allow us to obtain unbiased constraints on b_g and $\langle b\rho_{\text{SFR}} \rangle$ from the data. It is worth noting that, after integrating over χ in Eq. 30, a given value of ℓ will receive contributions from a range of k values. For the broadest bin used here, where the effect is largest, the variation in k for a given ℓ is $\lesssim 20\%$, and thus the conservative scale cut defined above allow us to obtain unbiased constraints on $\langle b\rho_{\text{SFR}} \rangle$ in spite of this effect.

2.2 Angular power spectra and covariances

We estimate auto- and cross-power spectra from all datasets in this study using the so-called pseudo- C_ℓ or MASTER algorithm as implemented in NaMaster (Alonso et al. 2019). The method is described briefly here, and further details can be found in Hivon et al. (2002); Alonso et al. (2019).

In all cases studied here (galaxy overdensity and CIB intensity maps), the observed field $\tilde{a}(\hat{\mathbf{n}})$ is a masked version of the true underlying field $a(\hat{\mathbf{n}})$, i.e.

$$\tilde{a}(\hat{\mathbf{n}}) = w(\hat{\mathbf{n}})a(\hat{\mathbf{n}}), \quad (32)$$

where w is the mask. w can simply be a binary map defining the footprint over which a has been observed, but in general it can be thought of as a local weight that can be designed to maximise the signal-to-noise ratio of the resulting power spectrum estimator. Consider two such fields, \tilde{a} and \tilde{b} , with masks v and w respectively. Their pseudo- C_ℓ is defined as $\tilde{C}_\ell^{ab} \equiv \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m} \tilde{b}_{\ell m}^* / (2\ell + 1)$, where $\tilde{a}_{\ell m}$ are the harmonic coefficients of \tilde{a} . Due to the convolution

² Quantitatively, the value of $\langle b\rho_{\text{SFR}} \rangle$ evaluated at the mean of each redshift bin, and the value of $\langle b\rho_{\text{SFR}} \rangle$ averaged over the bin differ by less than $\sim 1 - 2\%$. This is negligible compared with the statistical uncertainties reported here.

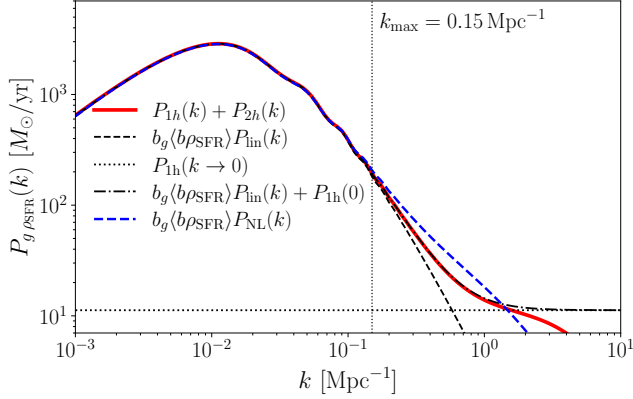


Figure 3. Halo model power spectrum at $z = 0.8$ for the galaxy- ρ_{SFR} cross-correlation. The solid red line shows the halo model prediction, while the black dash-dotted line shows the simplified phenomenological model of Eq. 7, using the $k \rightarrow 0$ limit for the 1-halo term (black dotted line), and for the bias-weighted halo profiles in the 2-halo term (black dashed line). This approximation is accurate up to 2% on scales $k \leq 1 \text{ Mpc}^{-1}$, which justifies scale cuts used in Section 4.1. The dashed blue line shows the simplified two-halo term using the non-linear matter power spectrum instead of the linear one. The halo model predictions were calculated using Model A, with parameters $\log_{10} M_{\text{max}} = 12.7$, $\eta_{\text{max}} = 0.42$, $\sigma_{M,0} = 1.75$, $\tau = 1.17$.

theorem, the $\tilde{a}_{\ell m}$ are a convolution of the harmonic coefficients of the true map and the mask. This propagates to the power spectra, so that the expectation value of \tilde{C}_{ℓ}^{ab} is

$$\langle \tilde{C}_{\ell}^{ab} \rangle = \sum_{\ell'} M_{\ell\ell'}^{vw} C_{\ell'}^{ab}. \quad (33)$$

Here, $M_{\ell\ell'}^{vw}$ is the mode-coupling matrix (MCM), which depends only on the masks v and w , and can be calculated analytically with relative ease. In order to estimate the true underlying power spectra from the observations we then need to invert the MCM. Since the MCM is not generally invertible, the pseudo- C_{ℓ} estimator often involves binning \tilde{C}_{ℓ}^{ab} into bandpowers before inverting the binned MCM. Here, we used the same bandpowers used in García-García et al. (2021). Linear ℓ bins with $\Delta\ell = 30$ were used up to $\ell = 240$, and logarithmic bins were used at higher ℓ with $\Delta \log_{10} \ell = 0.055$.

The covariance matrix of angular power spectra is sourced by three contributions: a disconnected (or ‘‘Gaussian’’) part, a super-sample covariance (SSC) due to modes larger than the survey footprint, and a connected non-Gaussian (cNG) part caused by the intrinsic non-Gaussianity of the tracers (Takada & Hu 2013). On the large scales used here the Gaussian component dominates, and we neglect the SSC and cNG contributions. We use an approximate analytical method to compute the Gaussian covariance matrix accounting for the effects of survey geometry, the so-called ‘‘narrow-kernel approximation’’ (NKA) described in García-García et al. (2019), and implemented in NaMaster. The NKA reduces the scaling of the exact calculation from $O(\ell_{\text{max}}^6)$ to a more tractable $O(\ell_{\text{max}}^3)$ by assuming that the MCM is narrowly peaked at the diagonal compared with the slow scale dependence of the true power spectrum.

Schematically, the covariance between two pseudo- C_{ℓ} s can be approximated as

$$\text{Cov}(\tilde{C}_{\ell}^{ab}, \tilde{C}_{\ell'}^{fg}) = C_{(\ell,\ell')}^{ag} C_{(\ell,\ell')}^{bf} \Xi_{\ell\ell'}(w^a w^g, w^b w^f) + (a \leftrightarrow b) \quad (34)$$

where $\Xi_{\ell\ell'}(w^a w^g, w^b w^f)$ are coupling coefficients similar to those involved in the calculation of $M_{\ell\ell'}^{vw}$, but now dependent on the products of pairs of masks (labelled $w^a/b/f/g$ here). The subscript (ℓ, ℓ') denotes the arithmetic average of the corresponding power spectra evaluated at both multipoles. The modified spectra C_{ℓ}^{ab} are related to the true underlying power spectra via (Nicola et al. 2021)

$$C_{\ell}^{ab} \equiv \frac{\langle \tilde{C}_{\ell}^{ab} \rangle}{\langle w^a w^b \rangle_{\text{pix}}}, \quad (35)$$

where $\langle \dots \rangle_{\text{pix}}$ denotes averaging over all pixels in the sky (as opposed to ensemble averaging). In general $\langle \tilde{C}_{\ell}^{ab} \rangle$ can be computed from a theoretical model and convolved with the MCM. However, since the power spectra estimated here have a high signal-to-noise ratio, we simply use the estimated pseudo- C_{ℓ} s in the formula above (i.e. replace $\langle \tilde{C}_{\ell}^{ab} \rangle \rightarrow \tilde{C}_{\ell}^{ab}$). In addition to simplicity, this has the added benefit of automatically accounting for the noise contribution to the power spectrum with no additional modelling.

This estimate of the covariance matrix was validated against jackknife resampling. As described in Section 4.1, both estimates are in reasonable agreement, although a small modification was applied to the final covariance matrix elements involving two different CIB frequency maps to accurately account for their map-level correlation.

2.3 Likelihood

We derive parameter constraints from the power spectrum data in two different ways. First, as described in Section 4.2, we derive model-independent tomographic measurements of $\langle b\rho_{\text{SFR}} \rangle$ from the galaxy auto-correlations and their cross-correlation with the CIB. We will then use these measurements, as described in Section 4.3, to place constraints on the free parameters of the halo models for the SFR described in Section 2.1.3. In both cases, we will make use of Gaussian likelihoods of the form

$$-2 \log p(\mathbf{d}|\Theta) = (\mathbf{d} - \mathbf{m}(\Theta))^T \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{d} - \mathbf{m}(\Theta)) + \text{const.}, \quad (36)$$

where \mathbf{d} is the data vector (power spectra or measurements of $\langle b\rho_{\text{SFR}} \rangle$), \mathbf{C} is the covariance of \mathbf{d} , \mathbf{m} is the theoretical model for \mathbf{d} , and Θ denotes the set of parameters used to define \mathbf{m} . The posterior parameter distribution is then given by the product of this likelihood and the parameter priors, which we discuss in Section 4.2. In all cases we will use the affine-invariant Markov-Chain Monte-Carlo (MCMC) algorithm implemented in emcee (Foreman-Mackey et al. 2013).

The assumption of a Gaussian likelihood for the power spectra stems from the central limit theorem: for sufficiently high ℓ , the estimate of C_{ℓ} involves averaging over a large number of independent m modes, and thus the distribution of C_{ℓ} values tends to a Gaussian. This is valid on the scales used in this analysis Hamimeche & Lewis (2008). The validity of the Gaussian approximation for the likelihood of the measured $\langle b\rho_{\text{SFR}} \rangle$ values will be demonstrated in Section 4.2.

All cosmological theory predictions were computed using the Core Cosmology Library (CCL, Chisari et al. (2019)³). Cosmological parameters were fixed to the *Planck* best-fit values $(\Omega_c, \Omega_b, h, n_s, \sigma_8) = (0.261, 0.049, 0.677, 0.9665, 0.8102)$. When using the halo model, we use the halo mass function parametrisation

³ The source code can be found at <https://github.com/LSSTDESC/CCL>

of Tinker et al. (2008), and the halo bias of Tinker et al. (2010). Halo masses are defined using a spherical overdensity $\Delta = 200$ times the critical density.

3 DATA

3.1 Galaxy clustering

In order to recover the star-formation history in the range $0 \lesssim z \lesssim 2$ we make use of photometric galaxy samples from the DESI Legacy Imaging Survey (Dey et al. 2019, DELS), and quasar samples (QSOs) from the extended Baryon Oscillation Spectroscopic Survey (eBOSS), derived from the Sloan Digital Sky Survey Data Release 16 (DR16).

At low redshifts, we use the photometric sub-sample of DELS selected by Hang et al. (2021)⁴. DELS combines photometry from three different telescopes: the DECam Legacy Survey (Flaugher et al. 2015; Blum et al. 2016) at declinations $\text{DEC} < 33^\circ$, the Mayall z -band Legacy Survey (Silva et al. 2016), and the Beijing-Arizona Sky Survey (Zou et al. 2019). The final survey covers 17739 deg^2 . In Hang et al. (2021), each galaxy was assigned a photometric redshift based on a multi-dimensional matching in colour space with a set of spectroscopic samples. We use these redshifts to separate the sample into the four tomographic bins used in Hang et al. (2021), following the same procedure described there to model the redshift distributions of the resulting samples. Moreover, we impose low-declination cut of $\text{DEC} > -36^\circ$, discarding a region exhibiting significant variations in completeness.

At higher redshifts we use the homogeneous eBOSS quasar sample used for the cosmological power spectrum analysis of Neveux et al. (2020); Hou et al. (2021), and presented in Ross et al. (2020). The catalogue comprises 343708 objects with measured redshifts in the range $0.8 \leq z \leq 2.2$, covering over 4800 deg^2 . We combine the measurements from the North and South Galactic Caps into one single catalogue that we split into two different bins with redshifts above and below $z = 1.5$ respectively. The redshift

⁴ The data can be found at <https://gitlab.com/qianjunhang/desi-legacy-survey-cross-correlations>. We thank the authors for the remarkable care with which these data were made publicly available.

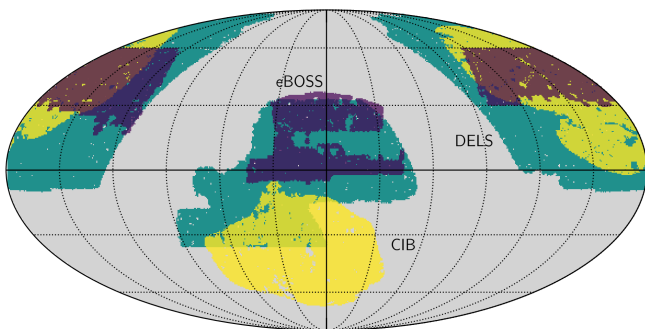


Figure 4. Sky footprint of the 545 GHz CIB and galaxy surveys used in this analysis in Equatorial coordinates.

Bin	Redshift range	Mean redshift	Density (deg^{-2})
DELS			
1	[0.10, 0.30)	0.21	808
2	[0.30, 0.15)	0.37	651
3	[0.45, 0.60)	0.50	760
4	[0.60, 0.80]	0.63	409
eBOSS			
5	[0.80, 1.50)	1.12	34
6	[1.50, 2.20]	1.87	35

Table 1. Mean redshift values and approximate redshift ranges for the 6 redshift bins shown in Figure 2. For DELS, the bin edges are in photo- z space, while the mean redshift corresponds to the mean of the true redshift distribution. The 4th column shows the angular number density of each sample.

distribution of each bin is estimated directly from the data as a histogram of the measured spectroscopic redshifts.

We construct maps of the galaxy overdensity from both samples using the methods described in García-García et al. (2021). These include the construction of galaxy masks, the use of random catalogues for eBOSS, the correction of large-scale systematic fluctuations for DELS, and the estimation of the noise bias for both samples. Details can be found in Section 3 of García-García et al. (2021). The impact of large-scale systematics in the galaxy auto-correlation is minimal given the scale cuts used for this analysis (see Section 4.2). Likewise, our analysis is not sensitive to inaccuracies in the estimation of the shot-noise contribution, since we fully marginalise over a noise-like white component (n_{xy} in Eq. 29) for all power spectra analysed. All maps are constructed and processed using HEALPix (Górski et al. 2005), using a resolution parameter $N_{\text{side}} = 2048$, corresponding to a pixel size of $\sim 1.7'$. Our final galaxy sample thus comprises 6 different redshift bins, shown in Fig. 2, and described in Table 1.

3.2 CIB maps

We use the full mission CIB maps and associated sky masks constructed by Lenz et al. (2019). The maps were created from the *Planck* 353, 545 and 857 GHz intensity maps. Galactic neutral hydrogen data from the HI4PI Survey (Bekhti et al. 2016) were used to create template maps of galactic dust. These templates were then used to subtract potential galactic dust contamination in the maps.

We apply the sky mask constructed by Lenz et al. (2019), based on the 20% *Planck* Galactic plane mask, removing small-scale Galactic filamentary structure via HI column densities above $N_{\text{HI}} > 2.5 \times 10^{20} \text{ cm}^{-2}$. The combined mask was then apodised using a $15'$ FWHM kernel.

When estimating the galaxy-CIB cross-correlations, we deconvolve the effective window function provided with the data products of Lenz et al. (2019). This accounts for the *Planck* Reduced Instrument Model (RIMO) beam, the window function of the SMICA CMB map, and the pixel window function for $N_{\text{side}} = 2048$ (see section 2.2 of Lenz et al. (2019) for further details).

As discussed in Lenz et al. (2019), the use of small sky patches to remove the Galactic dust component from the *Planck* maps leads to an over-subtraction of the underlying CIB fluctuations that results in a lack of power on large scales. The effect is limited to scales $\ell \lesssim 70$, and therefore we will omit all such scales from our analysis.

We applied color corrections as described in [Planck Collaboration et al. \(2014b\)](#).

The footprints of the three datasets used here are shown in Fig. 4. The CIB maps overlap with the DELS and eBOSS footprints over $\sim 13\%$ and $\sim 6\%$ of the sky respectively.

4 RESULTS

4.1 Power spectra and covariance

We measure the galaxy auto-correlation power spectrum for the six redshift bins considered, as well as their cross-correlations with the three CIB frequency maps, for a total of 24 auto- and cross-spectra. As an example, the galaxy clustering auto-correlations and the cross-correlations with the 545 GHz CIB map are shown in Fig. 5. The measured power spectra are shown in black together with their statistical uncertainties. The light grey bands show the fiducial scale cuts used in our analysis. We impose a default low- ℓ cut of $\ell_{\min} = 100$. This is mostly motivated by the loss of large-scale power in the CIB maps caused by the removal of Galactic contamination on different large-area patches, as reported by [Lenz et al. \(2019\)](#). Although this should not affect the galaxy auto-correlation, observational systematics leading to artificial contamination of the galaxy number counts across the sky may lead to extra power on the largest scales. $\ell_{\min} = 100$ is likely a conservative choice in this case ([Hang et al. 2021](#); [García-García et al. 2021](#)), but we choose to keep the same cut for both power spectra for simplicity. On small scales, we remove all multipoles above $\ell_{\max} = k_{\max}\bar{\chi} - 1/2$, where $\bar{\chi}$ is the radial comoving distance to the mean redshift of each galaxy sample, and $k_{\max} = 0.15 \text{ Mpc}^{-1}$. As we argued in Section 2.1.5, we expect the phenomenological model of Eq. 29 to provide a good fit to the data on these scales. To test the robustness of our results to this choice, we will also present results for a less conservative cut of $k_{\max} = 0.2 \text{ Mpc}^{-1}$ (shown by the dark grey bands in the Figure). After these scale cuts, the complete data vector, comprising six galaxy auto-correlations and 6×3 galaxy-CIB cross-correlations contains 200 elements.

Figure 5 also shows, in red, the best-fit theoretical prediction following the model of Eq. 29. The residuals between data and prediction normalised by the diagonal errors are shown at the bottom of each panel. As discussed and quantified in the next section, overall we find that the model is able to provide a good fit to the data.

Figure 6 shows the correlation matrix, defined in terms of the covariance C_{ij} as $r_{ij} \equiv C_{ij}/\sqrt{C_{ii}C_{jj}}$. The blocks along the diagonal correspond to the six different redshift bins, with the galaxy auto-correlation followed by the 3 CIB cross-correlations. In each block, the cross-covariance between $C_{\ell}^{g_s}$ and the $C_{\ell}^{g_v}$ is significantly smaller than the cross-covariance between CIB channels. The latter is at the level of $\sim 95\%$. This is because the three *Planck* maps trace essentially the same underlying CIB fluctuations with little frequency decorrelation. This makes an accurate calculation of the full covariance crucial since this high level of correlation requires the model residuals between different frequencies to follow the same trend almost exactly. The overlaps between the DELS redshift distributions shown in Figure 2 also lead to an appreciable cross-covariance between the power spectra in adjacent redshift bins.

We validate the analytical estimate of the covariance matrix described in Section 2.2 through jackknife resampling, dividing the footprint into 413 jackknife regions. In general, we find good agreement between both estimates of the covariance matrix, and we

$k_{\max} = 0.15 \text{ Mpc}^{-1}$			
Bin	\bar{z}	b_g	$\langle b\rho_{\text{SFR}} \rangle$ ($M_{\odot}\text{yr}^{-1}\text{Mpc}^{-3}$)
DELS			
1	0.21	1.068 ± 0.044	-0.005 ± 0.008
2	0.37	1.382 ± 0.026	0.026 ± 0.006
3	0.50	1.306 ± 0.017	0.041 ± 0.005
4	0.63	1.733 ± 0.019	0.048 ± 0.005
eBOSS			
5	1.12	2.057 ± 0.149	0.110 ± 0.014
6	1.87	2.165 ± 0.191	0.260 ± 0.032
$\chi^2/N_{\text{d.o.f}} = 201.2/188 = 1.07, \quad \text{PTE} = 0.245$			

Table 2. Summary of the obtained values for the galaxy bias and the bias weighted star-formation rate within each redshift bin as presented in Figure 2, when all parameters are sampled jointly. We display the mean values and the 1σ uncertainties. The last row show the χ^2 per degree of freedom value and PTE for our constraints.

are able to recover the diagonal errors within $\sim 4 - 5\%$ (see top panel of Fig. 7). However, we find that the jackknife estimate recovers cross-correlation coefficients between power spectra involving different CIB frequencies that are consistently lower than the analytical prediction by $\sim 2\%$ (see bottom panel of Fig. 7). Although this is a small difference, the tight correlation between different CIB frequencies makes it necessary to accurately characterise these off-diagonal covariance elements. We find that, otherwise, the goodness of fit to the full dataset combining all frequencies was unacceptable, even if the same model parameters were able to provide a good fit to the three different frequencies separately, which clearly points to an inconsistent covariance matrix. To correct for this, we therefore multiply all off-diagonal covariance matrix elements involving two different CIB frequencies by an overall factor of 0.98. The difference is likely due to inhomogeneity in the map noise across the footprint, which is not accounted for in the analytical covariance.

4.2 Model-independent measurements of $\langle b\rho_{\text{SFR}} \rangle$

We obtain tomographic measurements of the bias-weighted mean SFR density, $\langle b\rho_{\text{SFR}} \rangle$, by fitting the phenomenological model in Eq. 29 to the power spectrum data described in the previous section. The model consists on one galaxy bias parameter b_g and one value of $\langle b\rho_{\text{SFR}} \rangle$ for each redshift bin, as well as a free white-noise amplitude (n_{xy}) for each individual power spectrum. Although the white noise levels for different CIB frequencies should be tightly correlated, we choose to allow each of them to vary freely, in order to account for different levels of point-source subtraction, and non-Poissonian stochastic contributions in galaxy clustering. The full model therefore has a total of $6 \times N_{\text{bin}} = 36$ free parameters ($\langle b\rho_{\text{SFR}} \rangle$, b_g and one white noise amplitude for the four different power spectra in each galaxy redshift bin). Note, however, that since the white noise amplitudes are linear parameters, it is possible to marginalise over them analytically, significantly reducing the number of degrees of freedom we have to explore to $2 \times N_{\text{bin}} = 12$. This is achieved by modifying the inverse covariance matrix used in the Gaussian likelihood of Eq. 36 as ([Rybicki & Press 1992](#))

$$(\mathbf{C}_{\mathbf{d}}^{-1})' = \mathbf{C}_{\mathbf{d}}^{-1} - \mathbf{C}_{\mathbf{d}}^{-1}\mathbf{T}(\mathbf{T}^T\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{C}_{\mathbf{d}}^{-1}, \quad (37)$$

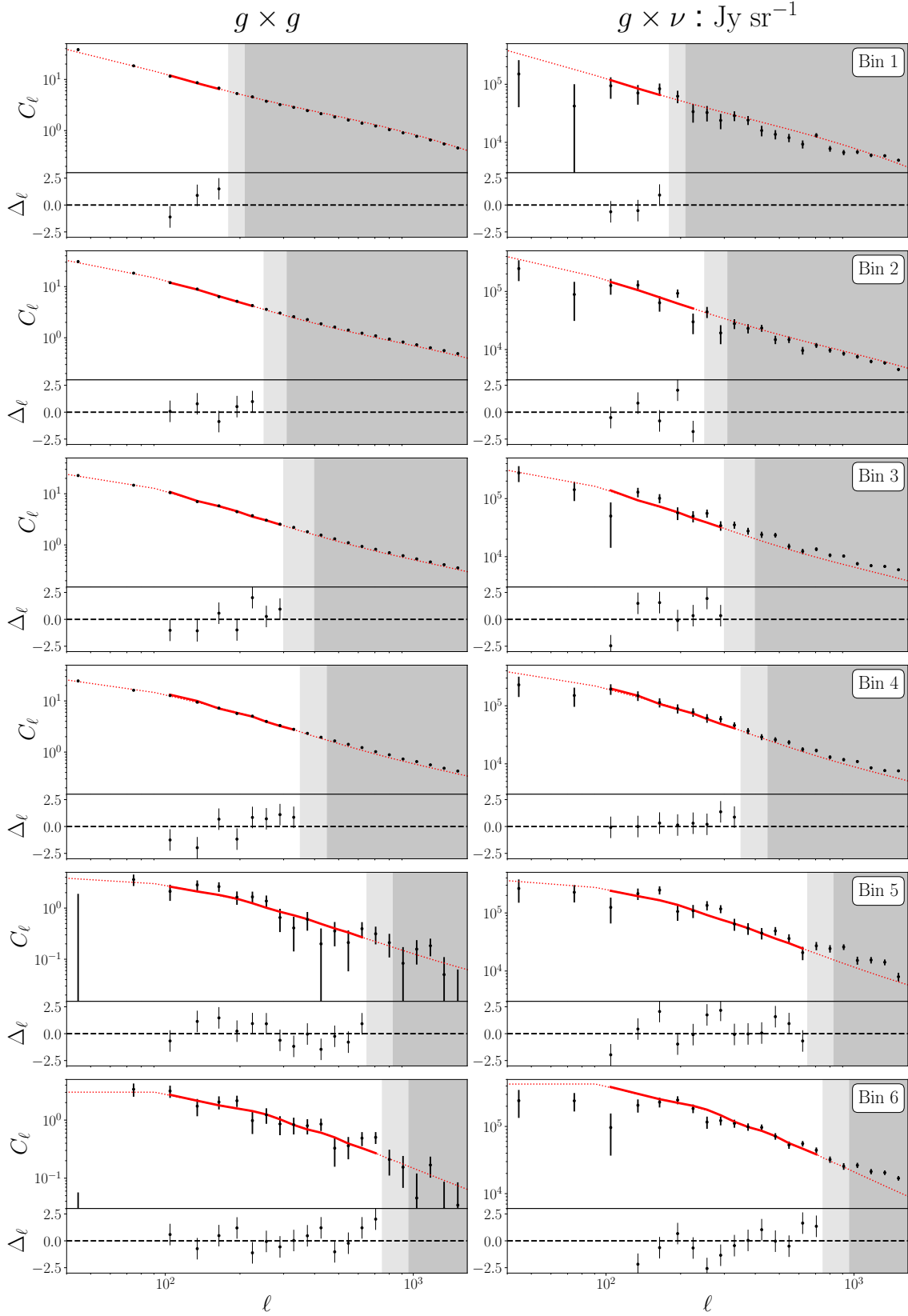


Figure 5. Galaxy auto-correlation spectra (left column) and galaxy-CIB cross-correlation spectra (right column) for the six tomographic bins presented in Fig. 2. Each panel shows the measurements and their uncertainties (black points with error bars), and their best-fit predictions (red lines), which are extrapolated outside of the range of scales used in our analysis. Our fiducial scale cuts are marked by the light grey vertical bands, while the dark grey bands mark the alternative, less conservative scale cuts corresponding to $k_{\max} = 0.2 \text{ Mpc}^{-1}$. The bottom part of each panel shows the difference between the data and the best-fit curve normalised by the 1σ uncertainties.

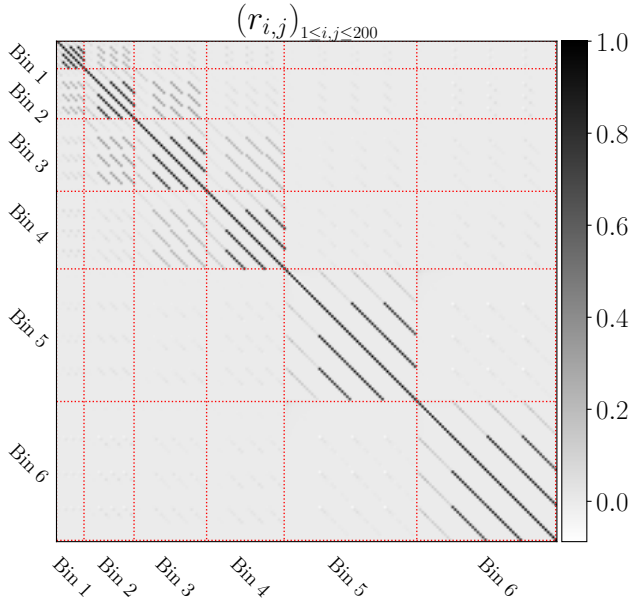


Figure 6. Correlation matrix $(r_{i,j})_{1 \leq i, j \leq 200}$ of the total data vector. The six blocks along the diagonal correspond to the different redshift bins used in the analysis. Within these blocks, the first sub-block contains the covariance of the galaxy auto-correlation, while the other blocks correspond to the different cross-correlations with the CIB. The different CIB frequency channels are strongly ($\sim 95\%$) correlated.

where \mathbf{T} is a matrix containing, as columns, the templates corresponding to the different linear parameters of the model (in this case, a constant vector for each of the different power spectra included in the data vector). We used flat, uninformative priors on all parameters.

Figure 8 shows the resulting measurements of $\langle b\rho_{\text{SFR}} \rangle$. Our fiducial measurements, simultaneously using all redshift bins and CIB frequencies, with a small-scale cut of $k_{\text{max}} = 0.15 \text{ Mpc}^{-1}$ are shown in black. The associated numerical values can be found in Table 2. We obtain $\sim 10\%$ constraints on $\langle b\rho_{\text{SFR}} \rangle$ at $z \geq 0.5$, and a $\sim 20\%$ measurement in the second redshift bin ($z \approx 0.37$). The conservative scale cuts used here leave only 3 data points in the cross-correlation with the first redshift bin, and we are only able to provide an upper bound on $\langle b\rho_{\text{SFR}} \rangle$ (the posterior distribution peaks on negative values), due to the degeneracy between this parameter and the white noise amplitude. The model is able to reproduce the data with a good χ^2 , with a probability-to-exceed PTE = 0.245. The full posterior distribution of the $\langle b\rho_{\text{SFR}} \rangle$ parameters is shown in Fig. 9. Although the measurements are largely statistically independent from one another, we can appreciate a non-negligible correlation between the measurements in adjacent DELS bins (e.g. bins 2-3 and 3-4). More importantly, we find that the posterior distribution is highly Gaussian. The dot-dashed black lines in the Figure show a multivariate Gaussian distribution with the same mean and covariance as that obtained from the MCMC chains, which reproduces the true posterior almost exactly. This will enable us to use the measured values of $\langle b\rho_{\text{SFR}} \rangle$ to constrain halo-based models of the star-formation history in Section 4.3. For convenience, we provide

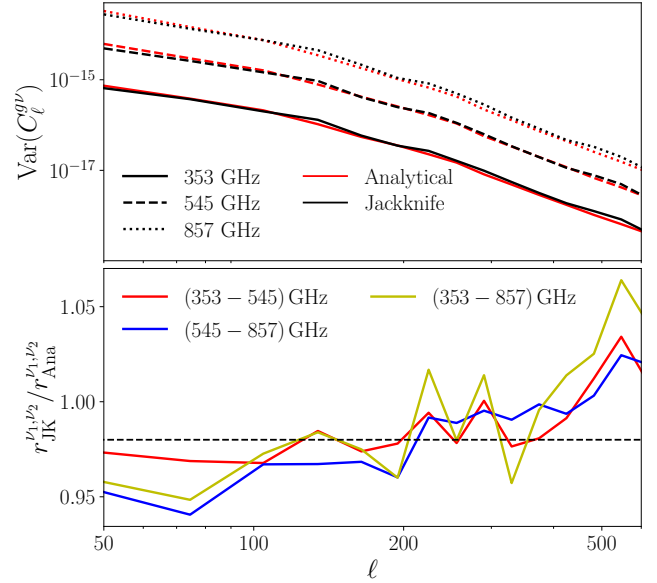


Figure 7. *Top:* variance of the galaxy-CIB cross-correlation for the fourth DELS redshift bin with the three CIB frequencies (solid, dashed, and dotted lines respectively). Results are shown in red for the analytical covariance, and in black for the jackknife estimate. *Bottom:* ratio between jackknife and analytical estimates for the correlation coefficient of pairs of cross-power spectra involving different CIB frequencies. Results are shown for the (353 – 545) GHz, (545 – 857) GHz and (353 – 857) GHz pairs in red, blue and yellow respectively. We find that, on average, the jackknife estimate predicts a $\sim 2\%$ lower correlation coefficient between frequencies. We correct our analytical estimate for this factor (shown as a horizontal dashed line).

the correlation matrix of these measurements:

$$\mathbf{r} = \begin{pmatrix} 1.00 & 0.14 & -0.01 & 0.01 & 0.02 & -0.03 \\ 0.14 & 1.00 & 0.26 & 0.02 & -0.02 & -0.01 \\ -0.01 & 0.26 & 1.00 & 0.27 & 0.01 & 0.00 \\ 0.01 & 0.02 & 0.27 & 1.00 & 0.02 & -0.01 \\ 0.02 & -0.02 & 0.01 & 0.02 & 1.00 & 0.06 \\ -0.03 & -0.01 & 0.00 & -0.01 & 0.06 & 1.00 \end{pmatrix}. \quad (38)$$

To validate these measurements, we repeat our analysis changing the number of redshift bins and CIB frequencies used, as well as the small-scale cut. The results for each of these alternative analyses are listed in Tables A1, A3, and A2 of Appendix A. First, in order to quantify the effect of the overlap between different redshift bins, we obtain new measurements of $\langle b\rho_{\text{SFR}} \rangle$ for each bin separately. We recover constraints that are almost equal to our fiducial measurements, and the model is able to reproduce the data for each redshift bin individually, with $\chi^2 p$ -values above 0.1. These results are shown as grey points in Fig. 8.

The model used here relies on having an accurate description of the mean infrared spectral energy distribution (S_{ν}^{eff} in Eq. 13), and inaccuracies in it would lead to inconsistent measurements of $\langle b\rho_{\text{SFR}} \rangle$ for different frequencies. To test for this, we repeat our measurement of $\langle b\rho_{\text{SFR}} \rangle$ three more times, using only one of the three CIB frequencies at a time. The results are presented in Table A3, and as orange/red/brown points in Fig. 8 for the 353/545/857 GHz channels respectively. Overall we find good agreement between the three frequencies. The 353 GHz channel achieves consistently larger uncertainties than the other two, and the recovered values of $\langle b\rho_{\text{SFR}} \rangle$ are also consistently higher, although compatible with

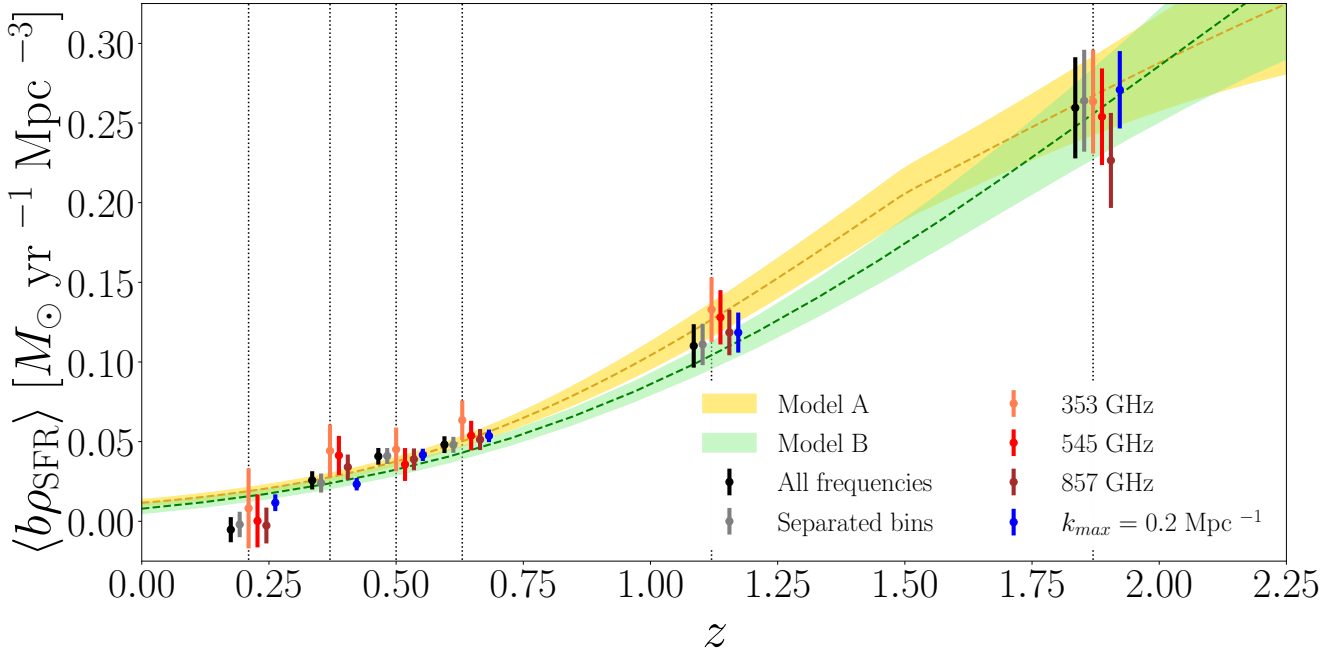


Figure 8. Inferred values of $\langle b\rho_{\text{SFR}} \rangle$ as a function of redshift for different analysis choices, with their 1σ error bars. Our fiducial measurements are shown in black. The results assuming less restrictive scale cuts are shown in blue, and the measurements carried out in each redshift bin independently are shown in grey. The constraints found from different frequency maps independently are shown in orange, red, and burgundy (for 353, 545, and 857 GHz respectively). The yellow and green bands show the prediction and $1-\sigma$ confidence interval for $\langle b\rho_{\text{SFR}} \rangle$ obtained from the two halo models described in Section 2.1.3 (Model A and Model B respectively).

the fiducial measurements within 1σ . Although this could be a hint for a small inconsistency between the CIB data and the effective spectra of Béthermin et al. (2013, 2015, 2017), it may also be an indication of contamination by other extragalactic foregrounds at lower frequencies, particularly the thermal Sunyaev-Zel’dovich effect. Since we do not observe a definite trend in the difference between the 353 GHz constraints and the rest as a function of redshift, and since the differences lie within the 68% uncertainties, we did not investigate this further, and proceeded to combine all frequencies in our fiducial analysis.

Finally we repeat the analysis relaxing the small-scale cut to $k_{\text{max}} = 0.2 \text{ Mpc}^{-1}$, which allows us to include one or two additional bandpowers in each redshift bin (see Fig. 5). Overall, the recovered values of $\langle b\rho_{\text{SFR}} \rangle$ are again in good agreement with our fiducial results, within 1σ uncertainties. With these additional data points, the uncertainties on $\langle b\rho_{\text{SFR}} \rangle$ reduce by $\sim 20 - 30\%$. More importantly, the extra data point for the first DELS redshift bins allows us to break the degeneracy with the white noise amplitude and obtain a measurement of $\langle b\rho_{\text{SFR}} \rangle$ at $z \approx 0.2$. The value recovered is in reasonable agreement with the predictions of the best-fit halo models derived in the next Section when extrapolated to lower redshifts. The model is still able to provide a reasonable fit to the data with this less conservative scale cut (PTE = 4%). Our measurements of $\langle b\rho_{\text{SFR}} \rangle$ are therefore robust to the choice of scale cut made for our analysis.

4.3 Constraints on halo-based SFR models

We use the tomographic measurements of $\langle b\rho_{\text{SFR}} \rangle$ presented in the previous section to derive constraints on the two halo models for the SFR history described in Section 2.1.3. For this, we use again the Gaussian likelihood of Eq. 36, where \mathbf{d} is the set of six measured

values of $\langle b\rho_{\text{SFR}} \rangle$, $\mathbf{C}_{\mathbf{d}}$ is the covariance of these measurements estimated from the MCMC chains, and \mathbf{m} is the theoretical model, given by Eq. 14. The model has 4 free parameters for both models:

- **Model A** has $\Theta = \{\eta_{\text{max}}, \log_{10} M_{\text{max}}, \sigma_{M,0}, \tau\}$ with the following top-hat priors:

$$0 \leq \eta_{\text{max}} \leq 1.2, \quad (39)$$

$$10.5 \leq \log_{10} M_{\text{max}} \leq 14, \quad (40)$$

$$0 \leq \sigma_{M,0} \leq 4, \quad (41)$$

$$0 \leq \tau \leq 3. \quad (42)$$

- **Model B** has $\Theta = \{m_0, m_z, \eta_0, \eta_z\}$ with the following top-hat priors:

$$9 \leq m_0 \leq 14, \quad (43)$$

$$-6 \leq m_z \leq 6, \quad (44)$$

$$0 \leq \eta_0 \leq 1, \quad (45)$$

$$0 \leq \eta_z \leq 1. \quad (46)$$

Since aiming to constrain 4 parameters with only 6 data points will likely lead to overfitting, we do not expect to be able to obtain competitive constraints on all of these parameters. Our main aim, however, is to use these models to map our measurements of $\langle b\rho_{\text{SFR}} \rangle$ onto constraints on the evolution of the star-formation rate density $\rho_{\text{SFR}}(z)$.

The posterior distribution of the parameters of Model A is shown in Fig. 10 (left panel). Our constraints are in good agreement with those found by Yan et al. (2022) using the same CIB maps and galaxy clustering data from KiDS. These are shown as red vertical bands in the diagonal panels of the figure. Our measurements seem to favour slightly lower values of τ than those of Yan et al. (2022), albeit with a long tail towards large τ that makes them broadly

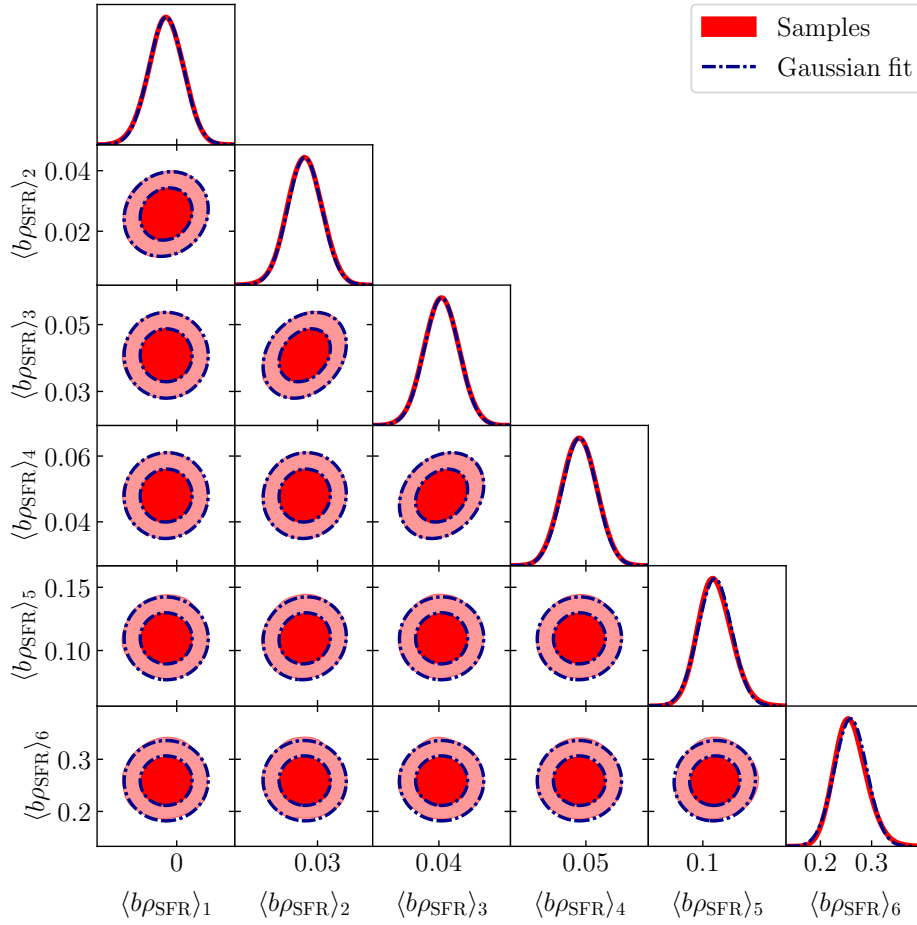


Figure 9. Posterior probability distributions for our measurements of $\langle b\rho_{\text{SFR}} \rangle$ in each redshift bin. The 1- and 2- σ contours calculated from the MCMC chains are shown in red. The blue dash-dotted line shows the same contours for a multivariate normal distribution with the same mean and covariance as the chains. This justifies the use of a Gaussian likelihood when using these measurements to place constraints on halo-based models of the SFR in Section 4.3.

compatible. The peak efficiency $\eta_{\text{max}} \sim 0.4$ is found to be in good agreement with both Yan et al. (2022) and the CIB constraints from Maniyar et al. (2021). On the other hand, our data favour a peak star-formation mass of $\log_{10} M_{\text{max}} \sim 11.5$, significantly lower than that found by Maniyar et al. (2021) ($\log_{10} M_{\text{max}} = 12.94 \pm 0.02$), but compatible with the constraints of Yan et al. (2022).

The posterior distribution for Model B is shown in Fig. 10 (right panel). Our constraints are in good agreement with the best-fit parameters of Moster et al. (2018)⁵, shown as dashed red lines. The data are able to constrain the peak mass parameters to

$$m_0 = 11.50 \pm 0.85, \quad m_z = 0.5 \pm 1.3, \quad (47)$$

although the peak efficiency parameters are largely unconstrained within their prior due to the degeneracy with m_0 and m_z . The peak mass is roughly constant, with $\log_{10} M_{\text{max}} \approx 11.6 \pm 0.8$ within the redshift range of our data. This is consistent with the results found with Model A.

In order to place constraints on the star-formation rate history from our measurements of $\langle b\rho_{\text{SFR}} \rangle$, we randomly select 1000 samples from the MCMC chains for both models and calculate $\langle b\rho_{\text{SFR}} \rangle$

and ρ_{SFR} as a function of redshift via Eqs. 14 and 15 respectively. The yellow and green bands in Fig. 8 show the 68% confidence interval for the predicted evolution of $\langle b\rho_{\text{SFR}} \rangle$ extracted from Model A and Model B respectively. The corresponding constraints on the star-formation rate density are shown in Fig. 11. The figure also shows various direct measurements of ρ_{SFR} from the infrared luminosity function (points with error bars), as well as the prediction from the best-fit model found by Maniyar et al. (2021) from the CIB auto-correlation. The direct ρ_{SFR} measurements are from Sanders et al. (2003); Takeuchi et al. (2003); Magnelli et al. (2011, 2013); Gruppioni et al. (2013); Marchetti et al. (2016) and Davies et al. (2016a), and were extracted from Madau & Dickinson (2014a) and Davies et al. (2016b). Overall, our constraints are in rough agreement with these independent direct measurements with two main caveats. At low redshifts, our constraints on Model A predict a marginally higher SFR density compared with the direct measurements, whereas the predictions of Model B agree well with the data. This may be due to the fact that, unlike Model A, Model B allows for both the peak efficiency and the maximum efficiency mass to depend on redshift. The conversion from $\langle b\rho_{\text{SFR}} \rangle$ to ρ_{SFR} at low-redshift prediction is therefore somewhat sensitive to the details of the underlying model. On the other hand, at high redshift both models predict a marginally higher ρ_{SFR} compared with the direct measurements. This is likely driven by the large amplitude of the

⁵ The constraints found by Moster et al. (2018) using a large dataset are much tighter than those found here, and thus we only show the final best-fit values.

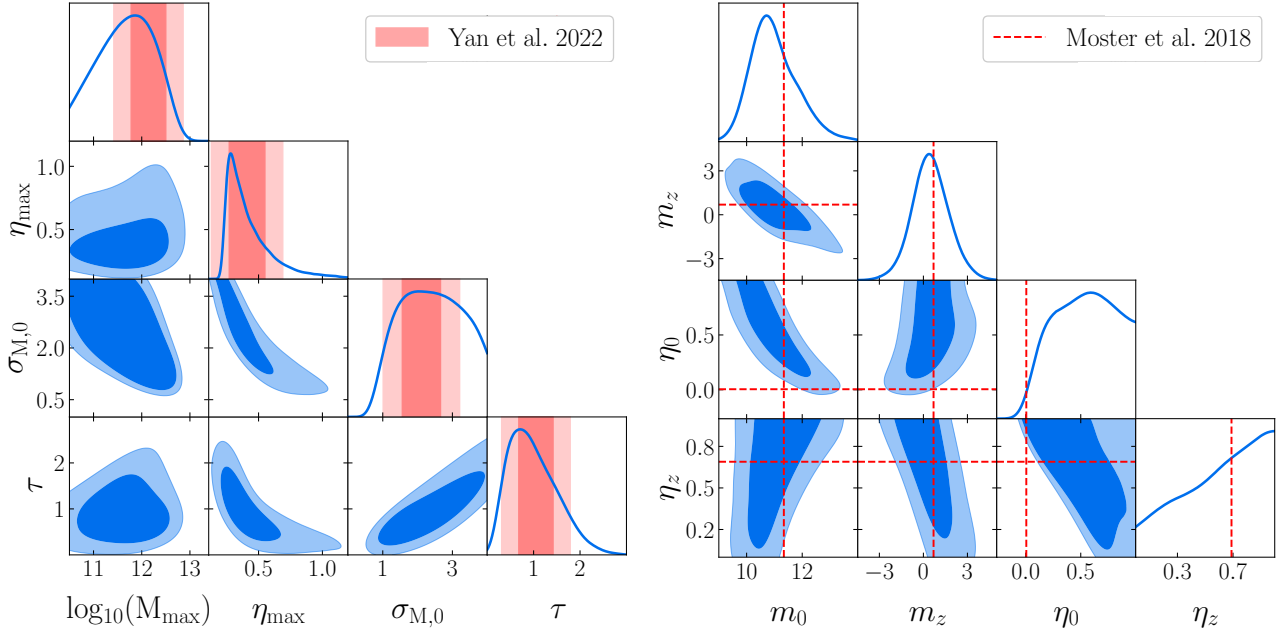


Figure 10. Constraints on the two halo-based SFR models (Model A and Model B in the left and right panels, respectively). The red bands in the left panel show the 1σ and 2σ constraints on Model A obtained by Yan et al. (2022). The red lines in the right panel show the best-fit parameters obtained by Moster et al. (2018) for Model B.

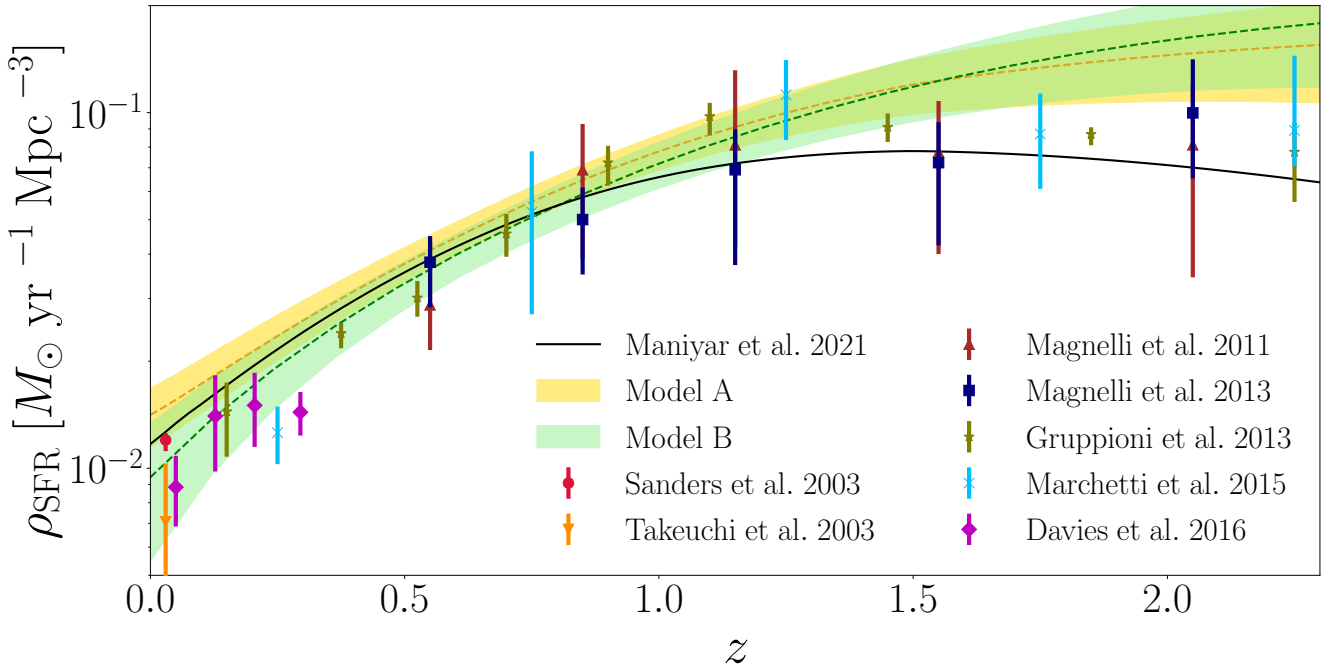


Figure 11. Star-formation rate density as a function of redshift. The points with error bars correspond to the direct measurements from the infrared luminosity function, extracted from Madau & Dickinson (2014a); Davies et al. (2016b). The solid black shows the best-fit prediction from the CIB analysis of Maniyar et al. (2021). The two bands show the $1\text{-}\sigma$ constraints obtained from our measurements of $\langle b\rho_{\text{SFR}} \rangle$ using the two halo-based SFR models described in Section 2.1.3 (Model A in yellow, Model B in green).

cross-correlation with the high-redshift eBOSS quasar sample, and therefore cross-correlations with other high-redshift tracers would be valuable to confirm this result.

Of potential concern is the fact that the eBOSS and DELS samples correspond to different galaxy types, which likely reside

in significantly different environments and have different infrared properties. As emphasized in Section 1, these differences should be irrelevant in large-scale limit of the halo model, as long as both samples trace the same underlying large-scale structure. Nevertheless, potential inaccuracies in the halo model could invalidate this as-

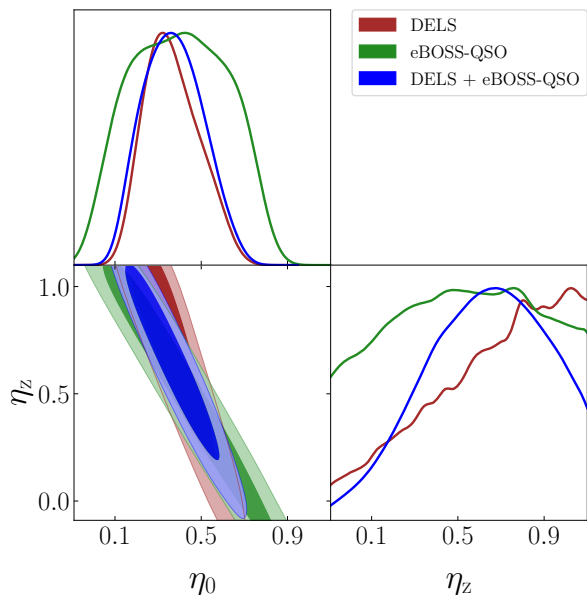


Figure 12. Posterior probability distributions for the parameters η_0 and η_z in model B when considering $\langle b\rho_{\text{SFR}} \rangle$ obtained for the 4 first bins (red), the 2 last bins (green) or the 6 bins (blue) from Figure 9 as data points. We set $m_0 = 11.339$ and $m_z = 0.692$ according to the best fit values in [Moster et al. \(2018\)](#).

sumption. To explore this possibility, we have re-analysed the four DELS and two eBOSS samples separately within the context of Model B, and compared the constraints resulting from each dataset. Since in this case we work with significantly fewer data points, we simplified Model B by fixing the mass parameters to the best-fit values from [Moster et al. \(2018\)](#) ($m_0 = 11.339$, $m_z = 0.692$), and varied only (η_0, η_z) . The results are shown in Fig. 12, together with the constraints from the full dataset. We find that the two different galaxy samples recover results that are in good agreement with each other and with the combined constraints. It is therefore unlikely that the differences in the astrophysical origin of both samples lead to a significant bias in the results presented here.

These caveats aside, the fact that our measurements of $\langle b\rho_{\text{SFR}} \rangle$ are compatible with direct measurements of the SFR density is an interesting and non-trivial result, beyond the fact that both constraints come from independent datasets. The contribution of haloes of different masses to $\langle b\rho_{\text{SFR}} \rangle$ is weighted by their linear bias, and thus $\langle b\rho_{\text{SFR}} \rangle$ is more sensitive than ρ_{SFR} to star formation in the most massive haloes. An agreement between measurements of both quantities within a given model thus provides non-trivial support for the physical interpretation of the dependence of star formation on halo mass. The measurements of $\langle b\rho_{\text{SFR}} \rangle$ presented here, and those that will be achievable with future datasets, can therefore be used to strengthen constraints on star-formation and galaxy evolution models, two areas of astrophysics of critical importance for cosmology in the next decade.

5 CONCLUSIONS

We have presented a tomographic analysis of Cosmic Infrared Background maps through cross-correlation with a suite of six galaxy samples spanning the redshift range $z \lesssim 2$ comprising galaxies from the DESI Legacy Survey and quasars from eBOSS. The CIB is a good tracer of the obscured star-formation history, and thus the main

motivation behind this exercise was to place direct constraints on the evolution of star formation in the last ~ 10 billion years. Our analysis has been carried out in two steps.

First, we used a phenomenological parametrisation of the angular power spectra, inspired on the halo model, which allows us to directly measure the mean star-formation rate density weighted by halo mass, $\langle b\rho_{\text{SFR}} \rangle$, in a virtually model-independent manner (see Eq. 29). Although this involves restricting our analysis to the largest scales, thus discarding significant amounts of information, we have argued that the resulting constraints are robust to a wide range of theoretical uncertainties that a full modelling of the small-scale C_ℓ s would be sensitive to. These include accurately describing the scale dependence of the mean and variance of the galaxy overdensity in haloes, the covariance between galaxy overdensity and SFR density in haloes (or, equivalently, characterising star formation in the target galaxy sample), the transition between the 1-halo and 2-halo-dominated regimes, and the different contributions (Poisson or otherwise) to the shot-noise component of all auto- and cross-correlations involved.

We find that this phenomenological model provides a good fit to the data on the scales probed. We thus obtained measurements for $\langle b\rho_{\text{SFR}} \rangle$ at six different redshifts, reported in Table 2. The measurements are robust with respect to the choice of scale cuts, as well as the combination of datasets used (i.e. the choice of redshift bins and frequency maps). Furthermore, we have shown that the posterior distribution for these $\langle b\rho_{\text{SFR}} \rangle$ measurements is well approximated by a multivariate normal. It is therefore easy to use these measurements to derive constraints on any halo-based parametrisation of star formation using a simple Gaussian likelihood. Since $\langle b\rho_{\text{SFR}} \rangle$ is more sensitive to the contribution to the SFR density from massive haloes than direct measurements of ρ_{SFR} , a combination of both measurements can be used to put constraints on the halo mass dependence of star formation.

As a second step, we use the measured values of $\langle b\rho_{\text{SFR}} \rangle$ as data to constrain the parameters of two halo-based models for the SFR history, characterising the mass and redshift dependence of the efficiency with which gas is converted into stars (see Fig. 1). Besides allowing us to place constraints on the models themselves, this procedure makes it possible to map our constraints on $\langle b\rho_{\text{SFR}} \rangle$ onto constraints on ρ_{SFR} itself. The result is shown in Fig. 11. Our constraints are in good qualitative agreement with independent measurements of the SFR density. This supports the physical interpretation of the halo-based SFR models of [Moster et al. \(2018\)](#); [Maniyar et al. \(2021\)](#), and the role of the CIB as a cosmic tracer of star formation. We have shown, however, that the predicted SFR density at low redshifts depends on the details of the underlying model, whereas our data seem to consistently predict a slightly higher ρ_{SFR} at $z \sim 2$. Thus our measurements of $\langle b\rho_{\text{SFR}} \rangle$ contain information that is complementary to that of direct measurements of the SFR density, and which can be used to improve our understanding of star formation in haloes.

A number of caveats associated with our analysis must be noted. First and foremost, as Fig. 5 shows, the cross-correlation between the CIB and our galaxy samples is detected at very high significance over all scales and redshifts explored. Nevertheless, the simplified model used here requires us to discard most of these data. A more sophisticated model would allow us to derive significantly tighter constraints on both models of the star-formation history, and on the astrophysical properties of the galaxy samples used. As we have noted above, several major modelling challenges must be overcome in order to construct such a model given the high statistical significance of the measurement. Ensuring that the model

is able to recover unbiased constraints on its parameters would likely require the use of simulations incorporating a somewhat realistic description of star formation and galaxy evolution. The rewards of developing more sophisticated models describing this and other cross-correlations (Hurier et al. 2015, 2017; Krolewski et al. 2021; McCarthy & Madhavacheril 2021; Krolewski & Ferraro 2022; Tröster et al. 2022) are nevertheless worth the effort. For now, we have limited ourselves to providing these conservative but robust constraints, and leave a more complex analysis, modelling the small-scale power spectra directly for future work.

Our model also relies on a number of simplified assumptions. First, we have assumed perfect knowledge of the mean infrared spectral energy distributions as tabulated by Béthermin et al. (2013, 2015, 2017). Since our measurement is based on the large-scale bias extracted from the 2-halo term, it is insensitive to the scatter in this spectrum. However, any mis-modelling of the redshift dependence of the mean SED would directly lead to a biased recovery of the evolution of $\langle b\rho_{\text{SFR}} \rangle$ (although the fact that we recover compatible results from the three *Planck* frequencies is reassuring). Furthermore, we have assumed perfect knowledge of the relationship between infrared luminosity and SFR, through a simple proportionality constant (Eq. 11) that assumes particular universal initial mass function (Chabrier 2003). This assumption would affect our results and the direct measurements of ρ_{SFR} in the same manner. The model also assumes that the SFR at a given time is proportional to the baryon accretion rate at the same time, whereas the dependence is likely more complicated. Our results are likely insensitive to most of these assumptions within the uncertainties reported, but they may become relevant in the presence of more data.

In terms of systematics, the presence of other sources of extragalactic contamination in the CIB maps, which would correlate with the galaxy distribution, could bias the measured values of $\langle b\rho_{\text{SFR}} \rangle$. The most harmful contaminants in this case are thermal Sunyaev-Zel'dovich and emission from radio point sources. The fact that we recover compatible results from the three *Planck* frequencies reassures that any residual contamination is likely negligible within the range of scales and frequencies explored here. This may not be the case if a more ambitious analysis was carried out, using smaller-scale data or other frequencies, and a more thorough quantification of the impact from possible sources of contamination would have been necessary. On the galaxy side, Galactic extinction may cause a non-cosmological modulation in the observed number counts of galaxies. Since Galactic dust is also a contaminant of the CIB maps, the cross-correlation between both tracers may be affected by this systematic. In our case, the effect should be negligible given the conservative large-scale cut we have imposed, but a more careful study of this cross-contamination may be required for more sensitive data, or if the large-scale correlation of both tracers were to be exploited (as would be the case if targeting ultra-large scale cosmological science cases such as primordial non-Gaussianity Tucci et al. (2016)). Finally, photometric redshift uncertainties leading to an incorrect determination of the galaxy redshift distribution may also impact the recovered constraints on $\langle b\rho_{\text{SFR}} \rangle$ (Hang et al. 2021). In our case this is only a potential concern for the DELS sample. Uncertainties in the sample's $p(z)$ can arise qualitatively in two ways: shifts in the mean of the distribution and changes in its width. Given the broad kernel of the CIB, our constraints should be impervious to moderate mean shifts. An incorrect determination of the width, however, would affect the amplitude of the galaxy auto-correlation, leaving the cross-correlation with the CIB almost untouched (Alonso et al. 2021). This would lead to an incorrect determination of the galaxy bias, which would then translate into a biased $\langle b\rho_{\text{SFR}} \rangle$. However, by

examining the cross-bin correlations of the DELS sample used here, Hang et al. (2021) were able to tightly constrain the overlap between bins (which would change with the $p(z)$ width), and therefore this effect should be small in our analysis. A careful calibration of the redshift distribution would be needed, however, when attempting this measurement with other photometric samples.

Further progress in the study of the CIB, and in its use as a cosmological tool, will become possible with the advent of next-generation experiments. On the CIB side, high-resolution data from experiments such as CCAT-prime (Stacey et al. 2018) or the Simons Observatory (Ade et al. 2019) will improve the quality, sensitivity, and frequency coverage of current CIB observations. On the galaxy side, future deep photometric surveys such as the Rubin Observatory Legacy Survey of Space and Time (LSST, LSST Science Collaboration et al. 2009) will enable us to trace the large-scale structure tomographically to redshifts $z \sim 3$, significantly improving our ability to trace the evolution of star formation at earlier times. Likewise, spectroscopic surveys such as DESI (Levi et al. 2019) will provide denser quasar samples, enabling the same type of constraints with better control over redshift uncertainties. With a better understanding of the star-formation history, it will also be possible to use the CIB itself as a cosmological tool, allowing us to map the Universe's structures on very large scales with a sensitivity higher than will be achievable by future CMB lensing experiments. However, a large number of modelling and data analysis challenges need to be tackled before we can fully exploit the rich information encoded in the CIB.

ACKNOWLEDGEMENTS

We would like to thank Jens Chluba, Martin Rey, Aprajita Verma, and Ziang Yan for useful discussions. BJ is supported by the ENS Paris-Saclay. CGG is supported by European Research Council Grant No: 693024 and the Beecroft Trust. DA is supported by the Science and Technology Facilities Council through an Ernest Rutherford Fellowship, grant reference ST/P004474. JRZ and NK are supported by an STFC doctoral studentship. NK also acknowledges funding from the University of Oxford, Department of Physics. We also made extensive use of computational resources at the University of Oxford Department of Physics, funded by the John Fell Oxford University Press Research Fund.

We made extensive use of the `numpy` (Oliphant 2006; Van Der Walt et al. 2011), `scipy` (Virtanen et al. 2020), `astropy` (Astropy Collaboration et al. 2013, 2018), `healpy` (Zonca et al. 2019), `GetDist` Lewis (2019), and `matplotlib` (Hunter 2007) python packages.

The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ág (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

NOIRLab is operated by the Association of Universities for

Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenossische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l’Espai (IEEC/CSIC), the Institut de Fisica d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universitat Munchen and the associated Excellence Cluster Universe, the University of Michigan, NSF’s NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program “The Emergence of Cosmological Structures” Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 11433005).

The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration.

The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center

for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

DATA AVAILABILITY

The code developed for this work as well as the derived data sets produced (power spectra and covariances) are available upon request. The catalogues and maps used were made publicly available by the authors, as described in the text.

REFERENCES

- Ade P., et al., 2019, *J. Cosmology Astropart. Phys.*, 2019, 056
 Alonso D., Hill J. C., Hložek R., Spergel D. N., 2018, *Phys. Rev. D*, 97, 063514
 Alonso D., Sanchez J., Slosar A., LSST Dark Energy Science Collaboration 2019, *MNRAS*, 484, 4127
 Alonso D., Bellini E., Hale C., Jarvis M. J., Schwarz D. J., 2021, *MNRAS*, 502, 876
 Astropy Collaboration et al., 2013, *A&A*, 558, A33
 Astropy Collaboration et al., 2018, *AJ*, 156, 123
 Bekhti N. B., et al., 2016, *Astronomy & Astrophysics*, 594, A116
 Berlind A. A., Weinberg D. H., 2002, *The Astrophysical Journal*, 575, 587
 Béthermin M., Wang L., Doré O., Lagache G., Sargent M., Daddi E., Cousin M., Aussel H., 2013, *A&A*, 557, A66
 Béthermin M., et al., 2015, *A&A*, 573, A113
 Béthermin M., et al., 2017, *A&A*, 607, A89
 Blum R. D., et al., 2016, in *American Astronomical Society Meeting Abstracts #228*, p. 317.01
 Cao Y., Gong Y., Feng C., Cooray A., Cheng G., Chen X., 2020, *ApJ*, 901, 34
 Chabrier G., 2003, *PASP*, 115, 763
 Chen C.-C., et al., 2016, *ApJ*, 831, 91
 Chiang Y.-K., Ménard B., Schiminovich D., 2019, *ApJ*, 877, 150
 Chiang Y.-K., Makiya R., Ménard B., Komatsu E., 2020, *ApJ*, 902, 56
 Chisari N. E., et al., 2019, *ApJS*, 242, 2
 Cooray A., Sheth R., 2002, *Physics reports*, 372, 1
 Darwish O., Sherwin B. D., Sailer N., Schaan E., Ferraro S., 2021a, arXiv e-prints, [p. arXiv:2111.00462](https://arxiv.org/abs/2111.00462)
 Darwish O., et al., 2021b, *MNRAS*, 500, 2250
 Davies L. J. M., et al., 2016a, *MNRAS*, 461, 458

- Davies L. J., et al., 2016b, *Monthly Notices of the Royal Astronomical Society*, 461, 458
- Dey A., et al., 2019, *AJ*, 157, 168
- Dole H., et al., 2006, *A&A*, 451, 417
- Dunkley J., et al., 2013, *J. Cosmology Astropart. Phys.*, 2013, 025
- Dwek E., et al., 1998, *ApJ*, 508, 106
- Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, *MNRAS*, 406, 2267
- Fang X., Eifler T., Schaan E., Huang H.-J., Krause E., Ferraro S., 2022, *MNRAS*, 509, 5721
- Flaugher B., et al., 2015, *AJ*, 150, 150
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- García-García C., Alonso D., Bellini E., 2019, *J. Cosmology Astropart. Phys.*, 2019, 043
- García-García C., Ruiz-Zapatero J., Alonso D., Bellini E., Ferreira P. G., Mueller E.-M., Nicola A., Ruiz-Lapuente P., 2021, *J. Cosmology Astropart. Phys.*, 2021, 030
- Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759
- Grupponi C., et al., 2013, *MNRAS*, 432, 23
- Hamimeche S., Lewis A., 2008, *Phys. Rev. D*, 77, 103013
- Hang Q., Alam S., Peacock J. A., Cai Y.-C., 2021, *MNRAS*, 501, 1481
- Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, *ApJ*, 567, 2
- Hou J., et al., 2021, *MNRAS*, 500, 1201
- Hunter J. D., 2007, *Computing in Science & Engineering*, 9, 90
- Hurier G., Douspis M., Aghanim N., Pointecouteau E., Diego J. M., Macias-Perez J. F., 2015, *A&A*, 576, A90
- Hurier G., Singh P., Hernández-Monteaugudo C., 2017, arXiv e-prints, p. arXiv:1711.10774
- Kennicutt Robert C. J., 1998, *ARA&A*, 36, 189
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
- Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2
- Knox L., Cooray A., Eisenstein D., Haiman Z., 2001, *ApJ*, 550, 7
- Koukoufilippas N., Alonso D., Bilicki M., Peacock J. A., 2020, *MNRAS*, 491, 5464
- Krolewski A., Ferraro S., 2022, *J. Cosmology Astropart. Phys.*, 2022, 033
- Krolewski A., Ferraro S., White M., 2021, *J. Cosmology Astropart. Phys.*, 2021, 028
- LSST Science Collaboration et al., 2009, arXiv e-prints, p. arXiv:0912.0201
- Lenz D., Doré O., Lagache G., 2019, *ApJ*, 883, 75
- Levi M., et al., 2019, in *Bulletin of the American Astronomical Society*, p. 57 (arXiv:1907.10688)
- Lewis A., 2019, arXiv e-prints, p. arXiv:1910.13970
- Limber D. N., 1953, *ApJ*, 117, 134
- Madau P., Dickinson M., 2014a, arXiv preprint arXiv:1403.0007
- Madau P., Dickinson M., 2014b, *ARA&A*, 52, 415
- Magnelli Elbaz, D. Chary, R. R. Dickinson, M. Le Borgne, D. Frayer, D. T. Willmer, C. N. A. 2011, *A&A*, 528, A35
- Magnelli B., et al., 2013, *A&A*, 553, A132
- Maniyar A. S., Béthermin M., Lagache G., 2018, *A&A*, 614, A39
- Maniyar A., Béthermin M., Lagache G., 2021, *A&A*, 645, A40
- Marchetti L., et al., 2016, *MNRAS*, 456, 1999
- McCarthy F., Madhavacheril M. S., 2021, *Phys. Rev. D*, 103, 103515
- Mead A. J., Verde L., 2021, *MNRAS*, 503, 3095
- Mead A. J., Brieden S., Tröster T., Heymans C., 2021, *MNRAS*, 502, 1401
- Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, arXiv e-prints, p. arXiv:1303.4722
- Morrison C. B., Hildebrandt H., Schmidt S. J., Baldry I. K., Bilicki M., Choi A., Erben T., Schneider P., 2017, *MNRAS*, 467, 3576
- Moster B. P., Naab T., White S. D. M., 2018, *MNRAS*, 477, 1822
- Neveux R., et al., 2020, *MNRAS*, 499, 210
- Newman J. A., 2008, *ApJ*, 684, 88
- Nicola A., García-García C., Alonso D., Dunkley J., Ferreira P. G., Slosar A., Spergel D. N., 2021, *J. Cosmology Astropart. Phys.*, 2021, 067
- Oliphant T. E., 2006, *A guide to NumPy*. Vol. 1, Trelgol Publishing USA
- Osborne S. J., Hanson D., Doré O., 2014, *J. Cosmology Astropart. Phys.*, 2014, 024
- Pandey S., et al., 2019, *Phys. Rev. D*, 100, 063519
- Partridge R. B., Peebles P. J. E., 1967, *ApJ*, 147, 868
- Peacock J., Smith R., 2000, *Monthly Notices of the Royal Astronomical Society*, 318, 1144
- Planck Collaboration et al., 2011, *A&A*, 536, A18
- Planck Collaboration et al., 2014a, *A&A*, 571, A18
- Planck Collaboration et al., 2014b, *A&A*, 571, A30
- Planck Collaboration et al., 2016, *A&A*, 596, A109
- Puget J. L., Abergel A., Bernard J. P., Boulanger F., Burton W. B., Desert F. X., Hartmann D., 1996, *A&A*, 308, L5
- Ross A. J., et al., 2020, *MNRAS*, 498, 2354
- Rybicki G. B., Press W. H., 1992, *ApJ*, 398, 169
- Sailer N., Schaan E., Ferraro S., Darwish O., Sherwin B., 2021, *Phys. Rev. D*, 104, 123514
- Sanders D. B., Mazzarella J. M., Kim D.-C., Surace J. A., Soifer B. T., 2003, *The Astronomical Journal*, 126, 1607
- Schmittfull M., Seljak U., 2018, *Phys. Rev. D*, 97, 123540
- Scranton R., et al., 2005, *ApJ*, 633, 589
- Seljak U., 2000, *Monthly Notices of the Royal Astronomical Society*, 318, 203
- Serra P., Lagache G., Doré O., Pullen A., White M., 2014, *A&A*, 570, A98
- Shang C., Haiman Z., Knox L., Oh S. P., 2012, *MNRAS*, 421, 2832
- Silk J., 2003, *MNRAS*, 343, 249
- Silva D. R., et al., 2016, in *American Astronomical Society Meeting Abstracts #228*, p. 317.02
- Stacey G. J., et al., 2018, in *Marshall H. K., Spyromilio J., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 10700, Ground-based and Airborne Telescopes VII*, p. 107001M (arXiv:1807.04354), doi:10.1117/12.2314031
- Takada M., Hu W., 2013, *Phys. Rev. D*, 87, 123504
- Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, *ApJ*, 761, 152
- Takeuchi T. T., Yoshikawa K., Ishii T. T., 2003, *The Astrophysical Journal*, 587, L89
- Tinker J. L., Wetzel A. R., 2010, *ApJ*, 719, 88
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *The Astrophysical Journal*, 688, 709
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
- Tinsley B. M., 1980, *Fundamentals Cosmic Phys.*, 5, 287
- Tröster T., et al., 2022, *A&A*, 660, A27
- Tucci M., Desjacques V., Kunz M., 2016, *MNRAS*, 463, 2046
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science & Engineering*, 13, 22
- Viero M. P., et al., 2013, *ApJ*, 772, 77
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Wang L., et al., 2015, *MNRAS*, 449, 4476
- Yan Z., Hojjati A., Tröster T., Hinshaw G., van Waerbeke L., 2019, *ApJ*, 884, 139
- Yan Z., et al., 2021, arXiv e-prints, p. arXiv:2102.07701
- Yan Z., van Waerbeke L., Wright A. H., Bilicki M., Gu S., Hildebrandt H., Maniyar A. S., Tröster T., 2022, arXiv e-prints, p. arXiv:2204.01649
- Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, *Journal of Open Source Software*, 4, 1298
- Zou H., et al., 2019, *ApJS*, 245, 4
- van Engelen A., Bhattacharya S., Sehgal N., Holder G. P., Zahn O., Nagai D., 2014, *ApJ*, 786, 13

APPENDIX A: $\langle b\rho_{\text{SFR}} \rangle$ CONSTRAINTS FROM ALTERNATIVE ANALYSIS CHOICES

Table A1 shows the constraints on $\langle b\rho_{\text{SFR}} \rangle$ obtained by treating each redshift bin independently. The constraints found with a less conservative scale cut of $k_{\text{max}} = 0.2 \text{ Mpc}^{-1}$ are shown in Table A2. Finally, the results found including only one CIB frequency map at a time are shown in Table A3. The correlation matrix for the

Bin	b_g	$\langle b\rho_{\text{SFR}} \rangle$ ($M_{\odot}\text{yr}^{-1}\text{Mpc}^{-3}$)	$\chi^2/N_{\text{d.o.f}}$	PTE
1	1.073 ± 0.044	-0.002 ± 0.008	$1.75/10 = 0.175$	0.998
2	1.377 ± 0.026	0.024 ± 0.006	$19.98/18 = 1.109$	0.334
3	1.307 ± 0.017	0.041 ± 0.005	$34.18/26 = 1.315$	0.131
4	1.733 ± 0.019	0.048 ± 0.005	$23.38/30 = 0.779$	0.799
5	2.069 ± 0.146	0.111 ± 0.013	$56.23/50 = 1.125$	0.253
6	2.158 ± 0.192	0.264 ± 0.032	$61.43/54 = 1.138$	0.227

Table A1. Summary of the result values for the galaxy bias and the bias weighted star-formation rate within each redshift bin as presented in 2, when all parameters are sampled bin by bin.

Bin	b_g	$\langle b\rho_{\text{SFR}} \rangle$ ($M_{\odot}\text{yr}^{-1}\text{Mpc}^{-3}$)
1	1.125 ± 0.030	0.012 ± 0.005
2	1.366 ± 0.018	0.023 ± 0.004
3	1.307 ± 0.012	0.042 ± 0.004
4	1.739 ± 0.014	0.053 ± 0.004
5	1.970 ± 0.135	0.118 ± 0.013
6	2.321 ± 0.152	0.271 ± 0.024

$\chi^2/N_{\text{d.o.f}} = 270.1/232 = 1.16$, PTE = 0.044

Table A2. Summary of the result values for the galaxy bias and the bias weighted star-formation rate for the least restrictive value of $k_{\text{max}} = 0.20 \text{ Mpc}^{-1}$.

measurements presented in the Table A2 is

$$r = \begin{pmatrix} 1.00 & 0.11 & 0.02 & 0.03 & 0.00 & 0.01 \\ 0.11 & 1.00 & 0.28 & 0.04 & 0.01 & -0.01 \\ 0.02 & 0.28 & 1.00 & 0.25 & 0.00 & -0.01 \\ 0.03 & 0.04 & 0.25 & 1.00 & 0.01 & 0.00 \\ 0.00 & 0.01 & 0.00 & 0.01 & 1.00 & 0.05 \\ 0.01 & -0.01 & -0.01 & 0.00 & 0.05 & 1.00 \end{pmatrix}. \quad (\text{A1})$$

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

Bin	b_g	$\langle b\rho_{\text{SFR}} \rangle$ ($M_{\odot}\text{yr}^{-1}\text{Mpc}^{-3}$)
353 GHz		
1	1.071 ± 0.045	0.008 ± 0.025
2	1.382 ± 0.026	0.044 ± 0.017
3	1.312 ± 0.016	0.045 ± 0.014
4	1.736 ± 0.019	0.063 ± 0.012
5	2.064 ± 0.149	0.133 ± 0.020
6	2.244 ± 0.186	0.263 ± 0.033

$\chi^2/N_{\text{d.o.f}} = 84.6/88 = 0.96$, PTE = 0.583

545 GHz		
1	1.073 ± 0.045	0.000 ± 0.017
2	1.382 ± 0.026	0.041 ± 0.012
3	1.311 ± 0.017	0.036 ± 0.010
4	1.737 ± 0.019	0.054 ± 0.009
5	2.062 ± 0.151	0.128 ± 0.017
6	2.244 ± 0.187	0.254 ± 0.030

$\chi^2/N_{\text{d.o.f}} = 89.5/88 = 1.02$, PTE = 0.440

857 GHz		
1	1.073 ± 0.045	-0.003 ± 0.011
2	1.382 ± 0.026	0.034 ± 0.008
3	1.310 ± 0.017	0.039 ± 0.007
4	1.736 ± 0.019	0.051 ± 0.007
5	2.075 ± 0.148	0.119 ± 0.014
6	2.246 ± 0.185	0.227 ± 0.030

$\chi^2/N_{\text{d.o.f}} = 94.7/88 = 1.08$, PTE = 0.296

Table A3. Summary of the result values for the galaxy bias and the bias weighted star-formation rate within each redshift bin as presented in 2, when all parameters are sampled frequency per frequency.