Cosmological constraints with the sub-millimetre galaxies
Magnification Bias after large scale bias corrections.

González-Nuevo J.1,2, Cueli M. M.1,2, Bonavera L.1,2, Lapi A.3,4, Migliaccio M.5,6, Argüeso F.7,2, Toffolatti L.1,2

1Departamento de Fisica, Universidad de Oviedo, C. Federico García Lorca 18, 33007 Oviedo, Spain
2Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias (ICTEA), C. Independencia 13, 33004 Oviedo, Spain
3International School for Advanced Studies (SISSA), via Bonomea 265, I-34136 Trieste, Italy
4Institute for Fundamental Physics of the Universe (IFPU), Via Beirut 2, I-34014 Trieste, Italy
5Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
6INFN, Sezione di Roma 2, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
7Departamento de Matematicas, Universidad de Oviedo, C. Federico García Lorca 18, 33007 Oviedo, Spain

Received xxx, xxxx; accepted xxx, xxxx

ABSTRACT

Context. The study of the magnification bias produced on high redshift sub-millimetre galaxies by foreground galaxies through the analysis of the cross-correlation function was recently demonstrated as an interesting independent alternative to the weak lensing shear as a cosmological probe.

Aims. In the case of the proposed observable, most of the cosmological constraints depends mainly on the largest angular separation measurements. Therefore, we aim at studying and correcting the main large scale biases that affect foreground and background galaxy samples in order to produce a robust estimation of the cross-correlation function. Then we analyse the corrected signal in order to derive updated cosmological constraints.

Methods. The large scale bias corrected cross-correlation functions are measured using a background sample of H-ATLAS galaxies with photometric redshifts > 1.2 and two different foreground samples (GAMA galaxies with spectroscopic redshifts or SDSS galaxies with photometric ones, both in the range 0.2 < z < 0.8). They are modelled using the traditional halo model description that depends on both halo occupation distribution and cosmological parameters. These parameters are then estimated by performing a Markov chain Monte Carlo under different scenarios to study the performance of this observable and the way to further improve its results.

Results. After the large scale bias corrections, we get only minor improvements with respect to the Bonavera et al. 2020 results, mainly confirming their conclusions: a lower bound on Ω_m > 0.22 at 95% C.L. and an upper bound σ_8 < 0.97 at 95% C.L. (results from the c_8_m−σ样−c_8_m sample). Neither the much higher surface density of the foreground photometric sample nor the assumption of gaussian priors for the remaining unconstrained parameters improves significantly the derived constraints. However, by combining both foreground samples into a simplified tomographic analysis, we were able to obtain interesting constraints on the Ω_m-c_8_m plane: Ω_m = 0.42^{+0.08}_{-0.14} and σ_8 = 0.81^{+0.09}_{-0.09} at 68% CL.

Key words. Galaxies: high-redshift – Submillimetre: galaxies – Gravitational lensing: weak – Cosmology: cosmological parameters

1. Introduction

The apparent excess number of high redshift sources observed near low redshift mass structures is known as Magnification Bias (see e.g. Schneider et al. [1992]): the deflections produced by the intervening gravitational field (area stretching and amplification) affecting the light rays coming from distant sources increase, in general, their chances of being included in a flux-limited sample (see for example Aretxaga et al. [2011]).

An unambiguous manifestation of this bias is the existence of a non negligible cross-correlation function between two source samples with non-overlapping redshift distributions. It has been observed in several contexts: galaxy-quasar cross-correlation function (Scranton et al. [2005], Ménard et al. [2010]), cross-correlation signal between Herschel sources and Lyman-break galaxies (Hildebrandt et al. [2013]) or the CMB (Bianchini et al. [2015], 2016) among others.

The cross-correlation signal can be enhanced by optimizing the choice of foreground and background samples. In this paper we use the sub-millimetre galaxies (SMGs) as the background sample because some of their features (steep luminosity function, very faint emission in the optical band and typical redshifts above z > 1 − 1.5) make them close to the optimal background sample for lensing studies as confirmed by a long series of publications (see for example Blain et al. [1996], Negrello et al. [2007, 2010], González-Nuevo et al. [2012], Bussmann et al. [2012], 2013], Fu et al. [2012], Wardlow et al. [2014], Calanog et al. [2014], Nayeri et al. [2016], Negrello et al. [2017], González-Nuevo et al. [2019], Bakx et al. [2020] among the most important ones).

In early works, the magnification bias produced on SMGs was already observed (Wang et al. [2011]) and measured with high significance, > 10σ (González-Nuevo et al. [2014]). In González-Nuevo et al. [2017] the measurements were further improved, allowing a more detailed study with a Halo model. It was concluded that the lenses are massive galaxies or even galaxy groups/clusters, with minimum mass of M_{lens} ∼ 10^{13} M⊙. Moreover, it was demonstrated that it is possible to split the fore-
ground sample in different redshift bins and to perform a tomographic analysis thanks to the better statistics. Finally, Bonavera et al. (2019) use the magnification bias to study the mass properties of a different type of lenses, a sample QSOs at 0.2 < z < 1.0. It was possible to estimate the halo mass where the QSOs acting as lenses are located in the sky, \( M_{\text{min}} = 10^{13.3 \pm 0.2} M_{\odot} \). These mass values indicate that we are observing the lensing effect of a cluster size halo signposted by the QSOs.

The interest in magnification bias is driven by the fact that it can be used as an additional cosmological probe to address the estimation of the parameters in the standard cosmological model. In fact, the importance of the magnification bias effect depends on the importance of changes in the gravitational deflection caused by low redshift galaxies on light travelling close to such lens, which in turn depends on cosmological distances and galaxy halo properties.

Features like the anisotropies in the CMB (e.g., Hinshaw et al. 2013), Planck Collaboration et al. (2016a, 2018a), the big bang nucleosynthesis (e.g., Fields & Olive 2000) and the SN1a observations of the Universe accelerating expansion (e.g., Betoule et al. 2014) are well handled by the current “standard cosmological model”. It is also inclusive of some Large Scale Structure (LSS) significant predictions about galaxies distributions (e.g., Peacock et al. 2001), such as baryon acoustic oscillations (BAOs) (e.g., Ross et al. 2015). Therefore, measurements based on such observables provide independent and complementary constraints on the cosmological parameters (e.g., Peacock & Dodds 1994). The success of the current model is in the fact that results from different probes are in great accordance.

However, with the increase in the quality and quantity of the measurements, some ‘tensions’ and small-scale issues have arisen that might indicate the necessity of modifications of the \( \Lambda \)CDM model. The main tensions are the value of the Hubble’s constant, \( H_0 = 74.03 \pm 1.42 \text{ km/s/Mpc} \) by Riess et al. (2019), Planck Collaboration et al. (2018a) with 67.4 ± 0.5 km/s/Mpc, and the usually degenerate relationship between the \( \Omega_m \) and \( \sigma_8 \) parameters (e.g., Heymans et al. 2013, Planck Collaboration et al. 2016, Hildebrandt et al. 2017, Planck Collaboration et al. 2018a).

In this context, Bonavera et al. (2020) (hereafter BON20) test the capability of the Magnification Bias produced on high-z SMGs as an additional independent cosmological probe in the effort to resolve the tensions. With this proof of concept analysis \( \Omega_m \) and \( H_0 \) were not well constrained. However, interesting limits were found: a lower limit of \( \Omega_m > 0.24 \) at 95% CL and an upper limit of \( \sigma_8 < 1.0 \) at 95% CL (with a tentative peak around 0.75).

Although the derived cosmological constraints from the Magnification Bias are relatively weak, it was confirmed as a new, independent observable making it a valuable new technique. Therefore it is worth making some efforts to improve further such results.

In this respect, most of the cosmological analysis that can be performed using the measured cross-correlation function (cosmological parameters, mass function, neutrinos, ...) depends mainly on the observed data at the largest angular scales (≥ 20 arcmin). On the one hand, this data are the most uncertain ones with large error-bars. Large areas and high source densities are needed in order to derive precise measurements. On the other hand, large scale bias, that can be considered negligible at smaller scales, can affect the data, and, as a consequence, the derived cosmological results. For these reasons the main goal of this work is to deeply study and find the optimal strategy to measure and analyse a precise and unbiased cross-correlation function at cosmological scales.

The work is organised as follows. In section 2 the background and foreground samples are described and in section 3 the methodology is presented. The large scale biases and how to correct them are described in 4. The derived cosmological constraints and conclusions are discussed in sections 5 and 6 respectively. In Appendix A we show the posteriors distributions of all the cases analysed and discussed in this work.

## 2. Data

The different galaxy samples used in this work are described in this section: the background sample, consisting of SMGs sources, and the foreground samples, consisting of two independent samples with spectroscopic and photometric redshifts, respectively.

![Fig. 1. Normalised redshift distributions of the three catalogues used in this work: the background sample i.e. H-ATLAS high-z SMGs (red solid line), the GAMA spectroscopic foreground sample (blue solid line) and the SDSS photometric foreground sample (magenta dashed line).](image)

### 2.1. Background sample

The Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al. 2010) is the largest area extragalactic survey carried out by the Herschel space observatory (Pilbratt et al. 2010) covering ~ 610 deg² with PACS [43] and SPIRE [44] instruments between 100 and 500 μm. Details of the H-ATLAS map-making, source extraction and catalogue generation can be found in Ibar et al. (2010), Pascale et al. (2011), Rigby et al. (2011), Valiante et al. (2016), Bourne et al. (2016), and Maddox & Dunne (2020).

The background sample consists of H-ATLAS sources detected in the three GAMA fields (total area of ~ 147 deg²), the North Galactic Pole (NGP, ~ 170 deg²) and the part of the South Galactic Pole (SGP) that overlaps with the spectroscopic foreground sample (~ 60 deg²). A photometric redshift selection of 1.2 < z < 4.0 has been applied to ensure no overlap in the redshift distributions of lenses and background sources, and we are thus left with ~ 66000 (~ 24 per cent of the initial sample and \( z_{\text{ph,med}} = 2.20 \)). The redshifts estimation is described in detail in González-Nuevo et al. (2017), Bonavera et al. (2019) and references therein.
This is the same background sample used in González-Nuevo et al. (2017), Bonavera et al. (2019) and Bonavera et al. (2020).

2.2. Foreground samples

In this work we use two independent foreground samples. The first one is the same one used by González-Nuevo et al. (2017), BON20, and we name it as the "$z_{\text{spec}}$ Sample". It consists of a sample extracted from the GAMA II (Driver et al. 2011; Baldry et al. 2010; 2014; Liske et al. 2015) spectroscopic survey, with $\sim 150000$ galaxies for $0.2 < z_{\text{spec}} < 0.8$ ($z_{\text{spec,med}} = 0.28$).

H-ATLAS and GAMA II surveys were carried out to maximize the common area coverage. Both surveys covered the three equatorial regions at 9, 12 and 14.5 h (referred to as G09, G12 and G15, respectively) and the SGP region was partially observed also by GAMA II. Thus, the resulting common area is of about $\sim 207\text{deg}^2$, surveyed down to a limit of $r \approx 19.8$ mag.

This is the same foreground sample used in González-Nuevo et al. (2017) and BON20.

The second foreground sample is selected in the Sixteenth Data Release of the Sloan Digital Sky Survey (SDSS; Blanton et al. 2017; Ahumada et al. 2019). It consists of galaxies with photometric redshift between $0.2 < z_{\text{ph}} < 0.8$ and photometric redshift error $\Delta z_{\text{err}}(1 + z) < 1$ (photoErrorClass=1), SDSS have completely covered the H-ATLAS equatorial regions and the NGP one (a total area of $\sim 317\text{deg}^2$). The sample, denominated "$z_{\text{ph}}$ sample", comprises $\sim 962000$ galaxies in total with median value of $z_{\text{ph,med}} = 0.38$.

The reason to introduce this second foreground sample is to study the improvements in the final results by increasing the density of potential lenses. The higher uncertainty in the redshift estimation of the foreground photometric redshifts is not very important in the current analysis because we are using a single wide redshift bin.

The normalised redshift distributions of the different samples are compared in Figure [1]. As in González-Nuevo et al. (2017), the random errors in the photometric redshifts are taken into account to estimate the redshift distributions. The main effect is to broadening the distributions beyond the selection limits. Figure [1] clearly shows the gap in redshift between the background and the foreground sources. The same figure also highlights the different redshift distributions between the two foreground samples.

3. Methodology

3.1. Tiling area scheme

The H-ATLAS survey is divided in five different fields: three GAMA fields in the ecliptic (9h, 12h, 15h) and two in the North and South Galactic Poles (NGP, and SGP). The H-ATLAS scanning strategy produced a characteristic diamond repeated shape in most of their fields. Taking into account the available area in each field we have different possibilities to measure the cross-correlation function:

- The "All" field area (blue line).- It provide the best statistics, i.e. smaller statistical uncertainties, both at small and large scales. The drawback is that we are limited to 4-5 fields to minimize the cosmic variance.
- The "Tile" area (red line).- This is the straightforward shape to be selected taking into account the observational strategy. The area of each tile, 16 sq. deg, should be large enough to avoid a bias in the large scale measurements (normally limited to angular separation below 2 deg). In order to maintain a regular shape for the tiles, a small overlap among such regions is needed, typically lower than 20% of the tiles’ area. The advantage of this area scheme is in the fact that it provides around 24 different tiles, that should help to diminish the cosmic variance.
- The "mini-Tile" area (magenta line).- It is built by dividing the tiles in four equal area "mini-Tiles" (each of $2 \times 2$ sq. deg). This area scheme typically provides around 96 different tiles. However, the maximum distance allowed by such area scheme is close to the cosmological scales that we want to measure. This was the area scheme used in BON20.

Each tiling area scheme has its own strong and weak points and can be affected by different types of large scale biases. Therefore, we perform a detailed analysis in order to compare the measurements from the different tiling area schemes and derive a robust estimation of the cross-correlation function, in particular at the cosmological angular scales.

3.2. Angular cross-correlation function estimation

As described in detail in González-Nuevo et al. (2017), BON20, we used a modified version of the Landy & Szalay (1993) estimator (Hirranz 2001):

$$w_2(\theta) = \frac{D_1 D_2 - D_1 R_2 - D_2 R_1 + R_1 R_2}{R_1 R_2},$$

(1)

where $D_1, D_2, D_1 R_2, D_2 R_1$ and $R_1 R_2$ are the normalized data1-data2, data1-random2, data2-random1 and random1-random2 pair counts for a given separation $\theta$.

For each selected area, we compute the angular cross-correlation function and the statistical function (averaging between 10 different realizations using different random catalogues each time). Each final measurement corresponds to the mean value of the cross-correlation functions estimated in each individual selected area for a given angular separation bin. The uncertainties correspond to the standard error of the mean, i.e. $\sigma_{\mu} = \sigma / \sqrt{n}$ with $\sigma$ the standard deviation of the population and $n$ the number of independent areas (each selected region can be assumed as statistically independent due to the small overlap).

Fig. 2. Examples of the different area selection to measure the cross-correlation function for the G09 H-ATLAS field. The "All" field area is shown in blue ( 56 sq. deg). The "tile" selection is shown in red ( 4x4 sq. deg.) and the "mini-Tile" one in magenta ( 2x2 sq. deg.)
3.3. Halo Model

As described in detail in the previous related works (González-Nuevo et al. 2017; Bonavera et al. 2019), BON20 we adopt the halo model formalism proposed by Cooray & Sheth (2002) in order to interpret a foreground-background source cross-correlation signal. An halo is defined as spherical regions whose mean over-density with respect to the background at any redshift is given by its virial value, which is estimated following Weinberg & Kamionkowski (2003) assuming a flat ΛCDM model. We used the traditional Navarro et al. (1996) density profile with the concentration parameter given in Bullock et al. (2001).

The cross-correlation between the foreground and background sources is linked to the low redshift galaxy-mass correlation through the weak gravitational lensing effect. The foreground galaxy sample traces the mass density field that causes the weak lensing, affecting the number counts of the background galaxy sample through Magnification Bias.

Following mainly Cooray & Sheth (2002) (see González-Nuevo et al. 2017 for details), we compute the correlation between the foreground and background sources adopting the standard Limber (Limber 1955) and flat-sky approximations (see e.g. Kilbinger et al. 2011) and references therein. It can be estimated as:

\[
W_{\text{fib}} = 2(\beta - 1) \int_0^\infty \frac{dN_f}{\chi^2(z)} \frac{dN_b}{dz} W_{\text{lens}}(z) \int_0^\infty \frac{dldl}{2\pi} p_{\text{gal-dm}}(l/\chi(z), z) J_0(l/\theta) \]

where

\[
W_{\text{lens}}(z) = \frac{3}{2} H_0^2 \frac{\Omega_m}{c^2} E(z) \int_z^\infty \frac{dN_f}{dz} \frac{\chi(z') - z}{\chi(z')} dN_b \]

being \(E(z) = \sqrt{\frac{\Omega_m(1+z)^3 + \Omega_\Lambda}{\Omega_m}} dN_f/dz\) and \(dN_f/dz\) as the unit-normalised background redshift distribution and \(z\) the source redshift. \(\chi(z)\) is the comoving distance to redshift \(z\). The logarithmic slope of the background sources number counts is assumed \(\beta = 3(N(S) = N_0 S^{-\beta})\) as in previous works (Lapi et al. 2011; 2012; Cai et al. 2013; Bianchini et al. 2015; 2016; González-Nuevo et al. 2017; Bonavera et al. 2019).

Small variations of its value are almost completely compensated by small changes in the \(M_{\text{min}}\) parameter.

As the Halo Occupation Distribution (HOD) we adopted the three parameters Zheng et al. (2005) model: all halos above a minimum mass \(M_{\text{min}}\) host a galaxy at their centre, while any remaining galaxy is classified as satellite. Satellites are distributed proportionally to the halo mass profile and halos host them when their mass exceeds the \(M_1\) mass. Finally, the number of satellites is a power-law function of halo mass with \(\alpha\) as the exponent, \(N_{\text{sat}}(M) = (\frac{M}{M_1})^{\alpha}\). Therefore, \(M_{\text{min}}, M_1\) and \(\alpha\) are the astrophysical free-parameters of the model.

3.4. Estimation of parameters

To estimate the different set of parameters, we performed a Markov chain Monte Carlo (MCMC) using the open source emcee software package (Foreman-Mackey et al. 2013). It is an MIT licensed pure-Python implementation of Goodman & Weare (2010) Affine Invariant MCMC Ensemble sampler. For each run, we generated at least 90000 posterior samples to ensure a good statistical sampling after convergence.

In the cross-correlation function analysis, we took into account both the astrophysical HOD parameters, and the cosmological parameters. The astrophysical parameters to be estimated are \(M_{\text{min}}, M_1\) and \(\alpha\). The cosmological parameters we want to constrain are \(\Omega_m, \sigma_8\) and \(h = H_0/100\). With the current samples, we do not have the statistical power to constrain \(\Omega_\Lambda\), \(\Omega_0\), and \(n_s\) in our analysis. As we assume a flat universe, \(\Omega_\Lambda\) is simply: \(\Omega_\Lambda = 1 - \Omega_m\). For the the other two cosmological parameters, we keep them fixed to the Planck most recent results, \(\Omega_0 = 0.0486\) and \(n_s = 0.9667\) (see Planck Collaboration et al. 2018a).

A traditional Gaussian likelihood function was used during this work.

It should be noted that only the cross-correlation data in the weak lensing regime (\(\theta \geq 0.2\) arcmin) are being taken into account for the fit since we are in the weak lensing approximation (see Bonavera et al. 2019 for a detailed discussion).

In general, we used the same flat priors for all the different analyses. They are based on the ones used in BON20. As for
the astrophysical parameters we chose: [12.0-13.5] for log $M_{\text{min}}$, [13.0-15.5] for log $M_1$, and [0.5-1.5] for $\alpha$. And for the cosmological parameters: [0.1-0.8] for $\Omega_m$, [0.6-1.2] for $\sigma_8$ and [0.5-1.0] for $h$.

4. Large Scale Biases

The cross-correlation function measurements using the different Tiling area schemes are compared in Figure 3. The left panel shows the measurements before any correction is applied. While all the different measurements agree almost perfectly within the uncertainties at small scales, there is a widespread variation of estimated values for angular separations above $\sim 10$ arcmin. But the cosmological parameters affect mainly those angular scales (see BON20 appendix figures). Therefore, we need to understand the causes that produce such high variation on our observations at those large angular scales before attempting any robust cosmological analysis.

It is well known that the distribution of galaxies in the Universe is not perfectly homogeneous. Therefore, in a field with a limited area, the number of detected galaxies will be somewhat higher or lower than the mean value obtained considering large enough areas. If this variation is not taken into account when building the random catalogues for a particular field, it will affect the DR and RR related terms in equation 1 and the estimated correlation could be stronger or weaker than the intrinsic value (see e.g. [Adelberger et al. 2005] for a detailed discussion on this topic).

To this respect, there are mainly two different biases that can affect the cross-correlation measurements at large scales: the integral constraint (IC; Roche & Eales 1999) and the surface density variation (SD; Blake & Wall 2002).

4.1. Integral constraint (IC)

When many fields are averaged, the overall effect of the large scale fluctuations tends to make the observed correlation weaker mainly at the largest observed scales. This means that the estimated cross-correlation function is biased low by a constant, the IC: $w_{\text{obs}}(\theta) = w_{\text{Ideal}}(\theta) + IC$.

Although there are possible theoretical approaches to estimate the IC for a particular scoring strategy (see e.g. [Adelberger et al. 2005]), it is commonly estimated numerically using the RR counts:

$$IC = \frac{\sum R_1 R_2(\theta) w_{x, \text{Ideal}}(\theta)}{\sum R_1 R_2(\theta)}.$$

As a first approximation of $w_{x, \text{Ideal}}(\theta)$, we assumed a power-law model, $w_{x, \text{Ideal}}(\theta) = A \theta^\gamma$. In order to be as much independent as possible of the exact value of the cosmological parameters (that mainly affect the largest angular scales), we estimated the best fit parameters for the power-law using only the observed cross-correlation function below 20 arcmin ($A = 10^{-1.54}$ arcmin and $\gamma = -0.89$). With the estimated power-law, the derived IC value for the "mini-Tiles" area was $9 \times 10^{-4}$. We verified that choosing a smaller angular separation upper limit or using different data set did not affect the derived IC value.

Moreover, assuming the best fit model of BON20 that can be considered biased low due to the fact that neglected the IC correction, the IC derived was again the same value. Therefore, we can conclude that the "mini-Tiles" estimated cross-correlation functions at the largest scales (>20 arcmin) are biased low but can be safely corrected by adding an IC $= 9 \times 10^{-4}$. Anyway, as discussed in section [5.1] this correction does not introduce any substantial difference with respect to the BON20 results on cosmological parameters.

On the other hand, the estimated IC for the "Tiles" area is $IC = 5 \times 10^{-4}$, considering both the power-law fit and the BON20 best fit model. As expected, the correction is smaller than in the "mini-Tiles" case taking into account the larger area of the "Tiles". The IC in the "Tiles" case affects marginally only the measurements above ~ 40 arcmin. Considering the large uncertainties at those angular scales, it can be almost considered a negligible correction for the $z_{\text{spec}}$ sample measured using the "Tiles" area. However, in the seek of precision, we decided to apply it in any case.

On the other hand, the IC results are completely negligible in the case of using the "all" field area scheme, as expected.

4.2. Surface density variations

The results using the "Tiles" area differ for the $z_{\text{spec}}$ and the $z_{\text{ph}}$ samples. This difference remains after the IC correction because it is the same for both cases. Moreover, the discrepancy is even stronger in the "All" scenario case (since the "All" measurements are almost the same between both samples, we are focusing only in the $z_{\text{ph}}$ for simplicity). This is a clear indication that an additional large scale bias is affecting the measurements when larger areas are considered. The fact that the $z_{\text{ph}}$ sample is more affected is probably related to the much higher density of sources in this sample.

If there is an additional variation of the source density of the foreground or the background sample that is not taken into account when building the random catalogues, it can produce a spurious enhancement of the measured correlation. As explained by [Blake & Wall 2002], the number of close pairs depended on the local surface density while the random pairs are related to the global average surface density. Then, systematic fluctuations produce $DD > RR$ that means a higher correlation (e.g. consider just the simplest estimator of the auto-correlation: $w(\theta) = DD/RR - 1$). Therefore, if present, the surface density variation produces the opposite effect with respect to the IC, that is what we are observing with the $z_{\text{ph}}$ sample.

4.2.1. Instrumental Noise variation

For the background sample, there is a well known surface density variation related to the instrumental noise due to the scanning strategy (see Figure 4 top panel). The overlap between the "Tiles" reduces the instrumental noise that allows fainter SMGs to be detected with respect to the rest of the field. For the auto-correlation analysis it was demonstrated that the potential effect can be considered negligible [Amyrosiadias et al. 2019]. Moreover, our results indicate that the relatively low surface density of the $z_{\text{spec}}$ sample makes this effect also negligible. In other words, the number of additional pairs due to the fainter background sources in those areas is not relevant enough to affect the measurements for the $z_{\text{spec}}$ sample. However, the much higher surface density of the $z_{\text{ph}}$ sample could produce a relevant enough enhancement of background-foreground pairs in those regions and, therefore, inducing a large scale surface density variation for the "Tiles" and "All" area schemes (we can consider the "mini-Tile" measurements simply dominated by the IC correction and neglect this other type of large scale bias even for the $z_{\text{ph}}$ sample).
In order to correct the instrumental noise surface density bias, we adopted the same procedure to generate random catalogues used in Amvrosiadis et al. (2019) for the auto-correlation analysis of the SMGs. First, a flux was chosen randomly from the flux densities of our background sample. Then the simulated galaxy is situated in a random position on the field. At this position the local noise was estimated as the instrumental noise and the confusion noise (see Table 3 of Valiante et al. 2016, for the GAMA fields). The estimated local noise is used to introduce a random Gaussian perturbation in the flux density. Finally, the simulated galaxy was kept in the sample if its flux density was greater than four times the local noise, the same detection limit used to produce the official H-ATLAS catalogue. This process was repeated for each random galaxy until the completion of the random catalogue.

These newly generated random catalogues correspond only to the background sample, i.e. it was only applied to build the $R_1$ random catalogues (used to estimate the $D_2R_1$ and $R_1R_2$ terms).

When the instrumental noise variation is considered, the cross-correlation functions showed a small correlation toward lower values at the largest angular scales (not shown individually in Figure 3). Although this result confirms that this bias is not negligible, it also highlights that it is not enough to explain the stronger correlation observed in the "Tiles" scheme for the $z_{ph}$ sample and the "All" one for both samples.

Therefore, we studied additional sources of surface density variations in the foreground samples.

4.2.2. Surface density variation of the foreground samples.

There are different causes of surface density variations in large area galaxy surveys, such as scanning strategy, sensitivity variation with time and foreground contamination. Moreover, the sample selection can amplify or reduce these variations, for example a region where the conditions for spectroscopic observations are different from the mean field ones. The detailed correction of these possible variations is complicated and requires a deep knowledge of the particular details of the instrument and the pipeline used for the production of the catalogue.

For the purpose of this work we adopted a simple approach to investigate the existence and correction of surface density variations in the foreground samples. As we can only observe a discrepancy at the largest angular scales, we decided to focus just on this range.

First, we created a surface density map by adding $+1$ to the pixel value at the position of each galaxy on the sample. Then we smoothed the map using a Gaussian kernel with a certain standard deviation (see discussion later in this section). Next, we apply the H-ATLAS survey masks (so that we can neglect border effects due to the smoothing step). These surface density maps are then used to generate the Random catalogues, $R_2$, for the foreground samples (used to estimate $D_1R_2$ and $R_1R_2$ terms in equation 1). The bottom panel in Figure 4 shows an example of smoothed surface density map built using the $z_{ph}$ sample, with a standard deviation of 180 arcmin, for the G15 field. As expected, the overall density map at those angular scales is almost homogeneous. However, there are some variations that might be biasing our measurements: the source density in the second "Tile" from the left is higher than the fourth one.

However, the exact value to be used as the Gaussian kernel dispersion is an unknown quantity. Using values smaller than 180 arcmin, the resulting density map starts to mimic the two-halo correlation of the foreground data. This means that the obtained $R_2$ contain part of the real auto-correlation and will remove part of this power from the estimated cross-correlation. For this reason and considering that the cross-correlation function decreases steeply for $\theta \sim 100$ arcmin, we can set a Gaussian dispersion of $> 150$ arcmin as a lower limit. On the other hand, for dispersion values above 180 arcmin, the surface density variation along the area becomes almost negligible in the derived $R_2$. Therefore, we can consider a dispersion of $< 200 – 220$ arcmin as an upper limit. Overall, we decided to proceed using a dispersion of 180 arcmin as a representative value, but taking into account that it is arbitrarily chosen. At the same time, given the uncertainties of the measurements at the relevant angular scales, small variations around the chosen deviation value became only a second order effect in our large scale measurements.

When both surface density variations are taken into account to generate the random catalogues the large scale bias observed in the "Tiles" scheme for the $z_{ph}$ sample or the "All" field area one for both samples disappear.

The right panel of Figure 3 shows the estimated cross-correlation functions using different tiling area schemes for the two samples after all the large scale bias corrections. The difference between the mean values at each angular scale is much smaller than the uncertainties. Considering this good agreement, we are confident that the measurements can be considered robust in all the angular scales commonly used for the cosmological analysis.

4.2.2. Surface density variation of the foreground samples.

There are different causes of surface density variations in large area galaxy surveys, such as scanning strategy, sensitivity variation with time and foreground contamination. Moreover, the sample selection can amplify or reduce these variations, for example a region where the conditions for spectroscopic observations are different from the mean field ones. The detailed correction of these possible variations is complicated and requires a deep knowledge of the particular details of the instrument and the pipeline used for the production of the catalogue.

For the purpose of this work we adopted a simple approach to investigate the existence and correction of surface density variations in the foreground samples. As we can only observe a discrepancy at the largest angular scales, we decided to focus just on this range.

First, we created a surface density map by adding $+1$ to the pixel value at the position of each galaxy on the sample. Then we smoothed the map using a Gaussian kernel with a certain standard deviation (see discussion later in this section). Next, we apply the H-ATLAS survey masks (so that we can neglect border effects due to the smoothing step). These surface density maps are then used to generate the Random catalogues, $R_2$, for the foreground samples (used to estimate $D_1R_2$ and $R_1R_2$ terms in equation 1). The bottom panel in Figure 4 shows an example of smoothed surface density map built using the $z_{ph}$ sample, with a standard deviation of 180 arcmin, for the G15 field. As expected, the overall density map at those angular scales is almost homogeneous. However, there are some variations that might be biasing our measurements: the source density in the second "Tile" from the left is higher than the fourth one.

However, the exact value to be used as the Gaussian kernel dispersion is an unknown quantity. Using values smaller than 180 arcmin, the resulting density map starts to mimic the two-halo correlation of the foreground data. This means that the obtained $R_2$ contain part of the real auto-correlation and will remove part of this power from the estimated cross-correlation. For this reason and considering that the cross-correlation function decreases steeply for $\theta \sim 100$ arcmin, we can set a Gaussian dispersion of $> 150$ arcmin as a lower limit. On the other hand, for dispersion values above 180 arcmin, the surface density variation along the area becomes almost negligible in the derived $R_2$. Therefore, we can consider a dispersion of $< 200 – 220$ arcmin as an upper limit. Overall, we decided to proceed using a dispersion of 180 arcmin as a representative value, but taking into account that it is arbitrarily chosen. At the same time, given the uncertainties of the measurements at the relevant angular scales, small variations around the chosen deviation value became only a second order effect in our large scale measurements.

When both surface density variations are taken into account to generate the random catalogues the large scale bias observed in the "Tiles" scheme for the $z_{ph}$ sample or the "All" field area one for both samples disappear.

The right panel of Figure 3 shows the estimated cross-correlation functions using different tiling area schemes for the two samples after all the large scale bias corrections. The difference between the mean values at each angular scale is much smaller than the uncertainties. Considering this good agreement, we are confident that the measurements can be considered robust in all the angular scales commonly used for the cosmological analysis.

As a final summary, to minimise the number of corrections applied to the data, we recommend to apply just the IC correction to the "mini-Tile" measurements for both samples and to the "Tile" one in the $z_{ph}$ case. In the other cases, the surface density correction is the most relevant one to be considered.

5. Cosmological constraints

Once the cross-correlation measurements are corrected for the different large scale biases discussed in the previous section, we
focus our analysis in their application to the estimate of some relevant parameters as done in BON20: the astrophysical parameters \((M_{\text{min}}, M_1\) and \(\alpha\)) and the cosmological ones \((\Omega_m, \sigma_8\) and \(h\)).

The higher number of independent smaller sky areas allows to minimise the error contribution given by the cosmic variance resulting in smaller uncertainties. For this reason and considering the almost perfect agreement between the "All" tiling scheme and the "Tiles" ones for both samples, we decided to focus just on the second case in order to simplify the discussion. Therefore, we focus on just four cases (all of them corrected for the relevant large scale biases): "mini-Tiles" and "Tiles" tiling schemes for both samples (\(z_{\text{spec}}\) and \(z_{\text{ph}}\)).

In the case of \(\Omega_m\), the new results moved the mean, \(\sim 0.45\), toward lower, more traditional values. This indicates that the large scale corrections helped to increase slightly the recovered values at the largest angular scales and to reduce their uncertainties. As a consequence, the highest \(\Omega_m\) values become less probable based on our current measurements. However, similar lower limits as in BON20 are confirmed, e.g., \(>0.22\) for the \(z_{\text{spec}}\) cases.

On the other hand, the results for \(M_1\) and \(\sigma_8\) are different depending on the sample used. However, the results based on the same sample but using different Tile schemes are consistent between them.

For \(M_1\), using the \(z_{\text{spec}}\) sample, we find a preference for \(\log(M_1/M_{\odot}) \geq 13.8\) but only at 68\% CL, whereas it shows a clear peak around \(\log(M_1/M_{\odot}) \sim 13.6 - 13.7\), using the \(z_{\text{ph}}\) one. In both cases these results are consistent with the BON20 ones.

In a similar way, \(\sigma_8\) mean estimated value moves from \(\sim 0.8\), obtained with the \(z_{\text{spec}}\) sample, to \(\sim 1.0\) using the \(z_{\text{ph}}\) one. Therefore, with the \(z_{\text{spec}}\) sample, the same as in BON20, we obtain similar \(\sigma_8\) constraints, but not confirmed by the \(z_{\text{ph}}\) ones.

Taking into account that the measurements of the cross-correlation function are almost the same between both samples (see again right panel of Figure 3), this discrepancy in some of the recovered parameters can only be related to the fact that both samples have different redshift distributions. In fact, Gonzalez-Nuevo et al. (2017) perform a tomographic analysis of the cross-correlation function using four different redshift bins, between \(0.1 < z < 0.8\), and study the evolution of the same HOD parameters. While the \(M_1\) parameter remains almost constant with redshift, there is a clear evolution of an increasing \(M_{\text{min}}\) values with redshift. The results of \(\alpha\) are inconclusive as it is unconstrained in most of the redshift bins. By using a single wide redshift bin, we are deriving an average of the astrophysical parameters weighted by the sample redshift distribution. Therefore, by analysing samples with different redshift distributions, it is expected to estimate different astrophysical parameter values, at least for those showing an evolution with redshift as \(M_{\text{min}}\).

5.2. Gaussian priors for the unconstrained parameters

As discussed in the previous section, there are two parameters that remain unconstrained with the current data sets: \(\alpha\) and \(h\). In this section, we study the potential improvements on the results by assuming external constraints on these two parameters. This additional information is introduced in the MCMC as Gaussian priors. For all the analysis in this section we used only the \(z_{\text{spec}}\) sample with the "mini-Tile" scheme.

In the case of \(\alpha\) we adopted a normal distribution with mean 1.0 and a dispersion of 0.1 (very similar to the Gaussian priors also used in BON20). The results are summarize in Table 3 and the derived posterior distribution are shown in Figure 3. In general, adopting a Gaussian prior for \(\alpha\) parameters produces almost no variation with respect to the default case. Only the most related parameters, \(\log M_1\) and \(\sigma_8\) move slightly toward lower values with a reduction on their dispersion of \(\sim 9\) and \(\sim 21\%\), respectively.

For the Hubble constant, we adopted the two popular values given by the local estimation, \(74.03 \pm 1.42\) km/s/Mpc (Riess et al. 2019), and the CMB one, \(67.4 \pm 0.5\) km/s/Mpc (Planck Collaboration et al. 2018b). The results obtained in these two cases are summarized in Table 4, while the derived posterior distributions are compared in Figure A.4. The only relevant variation with re-

![Figure 5](image-url)
Table 1. Results obtained from the $z_{\text{spec}}$ cross-correlation data-sets (the "mini Tiles" and the "Tiles"). From left to right, the columns are the parameters, the priors and the results (the mean, $\mu$ with the upper and lower limit at the 68 $\%$ CL, the $\sigma$ and the peak of the of the posterior distribution) for each data set. Those parameters without a value indicates they are unconstrained, i.e. that there is no constraint at 68$\%$ CL.

| Param       | Priors $U[a,b]$ | mini-Tiles $\mu$ $\pm$68CL | Tiles $\mu$ $\pm$68CL |
|-------------|----------------|-----------------------------|------------------------|
| $\log(M_{\text{min}}/M_\odot)$ | [12.0, 14.0] | 12.57$^{+0.23}_{-0.17}$ 0.20 12.61 | 12.61$^{+0.19}_{-0.15}$ 0.18 12.56 |
| $\log(M_1/M_0)$ | [12.5, 15.5] | 14.26$^{+1.24}_{-0.38}$ 0.78 15.03 | 14.37$^{+1.13}_{-0.37}$ 0.74 14.71 |
| $\alpha$    | [0.5, 1.5]    | $-$ $-$ $-$ $-$ $-$ $-$ $-$ |
| $\Omega_m$  | [0.1, 0.8]    | 0.45$^{+0.13}_{-0.21}$ 0.16 0.38 | 0.42$^{+0.14}_{-0.24}$ 0.18 0.31 |
| $\sigma_8$  | [0.6, 1.2]    | 0.84$^{+0.11}_{-0.18}$ 0.14 0.83 | 0.82$^{+0.08}_{-0.20}$ 0.14 0.75 |
| $h$         | [0.5, 1.0]    | $-$ $-$ $-$ $-$ $-$ $-$ $-$ |

Table 2. Results obtained from the $z_{\text{ph}}$ cross-correlation data-sets (the "mini Tiles" and the "Tiles"). From left to right, the columns are the parameters, the priors and the results (the mean, $\mu$ with the upper and lower limit at the 68 $\%$ CL, the $\sigma$ and the peak of the of the posterior distribution) for each data set. Those parameters without a value indicates they are unconstrained, i.e. that there is no constraint at 68$\%$ CL.

| Param       | Priors $U[a,b]$ | mini-Tiles $\mu$ $\pm$68CL | Tiles $\mu$ $\pm$68CL |
|-------------|----------------|-----------------------------|------------------------|
| $\log(M_{\text{min}}/M_\odot)$ | [12.0, 14.0] | 12.60$^{+0.20}_{-0.13}$ 0.18 12.67 | 12.61$^{+0.20}_{-0.13}$ 0.17 12.66 |
| $\log(M_1/M_0)$ | [12.5, 15.5] | 13.81$^{+0.53}_{-1.09}$ 0.76 13.60 | 13.95$^{+0.74}_{-0.95}$ 0.76 13.74 |
| $\alpha$    | [0.5, 1.5]    | 0.96$^{+0.15}_{-0.46}$ 0.27 0.77 | 0.96$^{+0.15}_{-0.46}$ 0.28 0.73 |
| $\Omega_m$  | [0.1, 0.8]    | 0.46$^{+0.11}_{-0.18}$ 0.14 0.38 | 0.46$^{+0.12}_{-0.19}$ 0.15 0.39 |
| $\sigma_8$  | [0.6, 1.2]    | 0.99$^{+0.12}_{-0.11}$ 0.11 0.98 | 0.98$^{+0.16}_{-0.10}$ 0.12 1.00 |
| $h$         | [0.5, 1.0]    | 0.71$^{+0.06}_{-0.21}$ 0.14 0.50 $-$ $-$ $-$ |

spect to the default case is that the $\sigma_8$ distribution moves again slightly toward lower values with a reduction on their dispersion of $\sim 29\%$.

When comparing between both $h$ priors cases, the results are almost identical. However, as also indicated in BON20, higher values of $h$ seem to perform slightly better: the $\Omega_m$ posterior distribution becomes thinner and moves towards lower, more traditional, values. However, the current uncertainties do not allow us to derive stronger conclusions on this particular topic.

Overall, adopting more restrictive priors on the unconstrained parameters does not improve remarkably the results in general. The parameter that seems to benefit more from the reduction of uncertainty in both cases is $\sigma_8$. This is probably due to the fact that it is the parameter that mostly depends on the intermediate angular scales and, therefore, it is the one mostly affected by changes induced both by the smallest scales ($\alpha$’s main influence) and by the largest scales ($h$’s main influence), see appendix in BON20.

5.3. Combining both data sets

The $z_{\text{ph}}$ sample has much better statistics with respect to the $z_{\text{spec}}$ one, but we do not see a relevant improvement in the obtained constraints. In addition, even if the measured cross-correlation function is almost the same, each sample provides different results in some of the studied parameters. This is probably linked to the different redshift functions. On this respect, González-Nuevo et al. [2017] tomographic analysis of the cross-correlation function show a strong evolution with redshift at least for the $\log M_\text{min}$ parameter. As explained before, by using a single wide redshift bin, the derived astrophysical parameters are the average of the evolving values measured by González-Nuevo et al. [2017] weighted by the particular sample redshift distribution. As we saw, the different averaged astrophysical parameter values between the two samples is affecting also the recovered values of some of the cosmological parameters. In particular $\sigma_8$ changes from 0.84 for the $z_{\text{spec}}$ sample to 0.99 for the $z_{\text{ph}}$ one.

A proper tomographic analysis is beyond the scope of this paper, but we can try a simple but interesting analysis: taking into
Table 3. Results obtained from the $z_{\text{spec}}$ cross-correlation data-set using the "mini-Tiles" scheme but assuming a Gaussian priors for the $\alpha$ parameter. From left to right, the columns are the parameters, the priors ($\mathcal{U}$,$\mathcal{N}$) for Uniform priors, and $\mathcal{N}[\mu,\sigma]$ for the Normal ones) and the results (the mean, $\mu$ with the upper and lower limit at the 68% CL, the $\sigma$ and the peak of the of the posterior distribution).

| Params          | Priors          | $\mu$       | $\sigma$ | peak |
|-----------------|-----------------|-------------|----------|------|
| log$(M_{\min}/M_\odot)$ | $\mathcal{U}[12.0, 14.0]$ | 12.53$^{+0.16}_{-0.04}$ | 0.21 | 0.10 | 1.00 |
| log$(M_1/M_\odot)$ | $\mathcal{U}[12.5, 15.5]$ | 14.31$^{+0.47}_{-0.38}$ | 0.71 | 1.00 | 1.00 |
| $\alpha$        | $\mathcal{N}[1.0, 0.1]$ | 0.99$^{+0.08}_{-0.05}$ | 0.10 | 0.10 | 1.00 |
| $\Omega_m$      | $\mathcal{U}[0.1, 0.8]$ | 0.46$^{+0.10}_{-0.16}$ | 0.15 | 0.15 | 0.37 |
| $\sigma_8$      | $\mathcal{U}[0.6, 1.2]$ | 0.76$^{+0.01}_{-0.16}$ | 0.10 | 0.10 | 0.64 |
| $h$             | $\mathcal{U}[0.5, 1.0]$ | 0.75$^{+0.05}_{-0.09}$ | 0.14 | 0.14 | 0.66 |

Fig. 6. Comparison of the posterior distributions for the astrophysical parameters, log$(M_{\min}/M_\odot)$, log$(M_1/M_\odot)$ and $\alpha$, derived from both samples, $z_{\text{spec}}$ and $z_{\text{ph}}$, using the "mini-Tiles" scheme (solid lines) and when combined in a tomographic analysis (dotted lines).

5.4. Comparison with other results

The weak gravitational lensing results available in the literature are usually related with a different and complementary observable, the shear. In this section we compare with measurements by cosmic shear of galaxies, focusing on the most constraining, and the CMB lensing by Planck (Planck Collaboration et al. 2018b). In particular the results from the following surveys (with different redshift ranges and affected by different systematic effects) are being taken into account for the comparison: the Canada-France-Hawaii Telescope Lensing Survey presented in CFHTLenS (Joudaki et al. 2017), the Kilo Degree Survey and VIKING based on 450 deg$^2$ data (KV450, Hildebrandt et al. 2020), the first-year lensing data from the Dark Energy Survey (DES, Troxel et al. 2018) and the Subaru Hyper Suprime-Cam first-year data (HSC, Hamana et al. 2020).

For the comparison, we use publicly released MCMC results. Moreover, the different results we are comparing with have different priors. Since we are not interested in an in-depth comparison, we do not adjust them to our fiducial set-up.

In particular, we compare the constraints in the $\Omega_m$ - $\sigma_8$ plane: cosmic shear measures the combination $\sigma_8\Omega_m^{0.21}$ and CMB lensing the $\sigma_8\Omega_m^{0.25}$ one. Such combinations highlight degeneracy directions, shown in the marginalised posterior contours (68% and 95% C.L.) in Figure 7 for the data-sets described above. To have a direct comparison with literature, the contours account that the results from both samples are independent and have different redshift distributions, we can try to constrain the cosmological parameters using both samples at the same time. We perform a joint analysis allowing different astrophysical parameters constraints for each sample but keeping the same cosmological parameters.

Therefore, we run an additional MCMC analysis but this time with nine parameters to be constrained (three astrophysical ones for each sample and three common cosmological ones). We used for both samples the "mini-Tiles" scheme as it needs the simplest large scale bias correction. The results are summarized in Table 3 and the derived posterior distributions for the nine parameters are shown in Figure A.5.

Regarding the astrophysical parameters (see Figure 6) the main changes of the combined analysis with respect to the individual ones are the following. Imposing a common cosmological parameters values seems not to affect the log $M_{\min}$ constraints for the $z_{\text{ph}}$ sample but it produces a shift towards slightly lower mean values for the $z_{\text{spec}}$ one (from 12.57 to 12.48). This is probably due to the more peaked redshift distribution of the $z_{\text{ph}}$ sample. In the case of the $M_1$ parameter, there is only a small reduction in the parameter uncertainty. Finally, there is no improvement in the $\alpha$ parameter, that remains unconstrained.

As expected, the most relevant changes are within the cosmological parameters. While for the $\Omega_m$ the estimated posterior distribution is simply less skewed toward high values (from an associated gaussian standard deviation of 0.16 to 0.12), the $\sigma_8$ parameter is the most affected. Its posterior distribution become almost gaussian with a mean value of $\sigma_8 = 0.81^{+0.09}_{-0.09}$ and a standard deviation of 0.09. However, for the Hubble constant the results give only an upper limit very similar to the one derived from the $z_{\text{ph}}$ sample alone ($h < 0.8$ at 68% CL).

A more detailed discussion on the $\Omega_m$ and $\sigma_8$ results will be presented in the next subsection.
Table 4. Results obtained from the $z_{spec}$ cross-correlation data-set using the “mini Tiles IC” scheme but assuming two different values for the Hubble constant. The first column is the name of the parameters and then, for each case, the columns are, from left to right, the priors ($U[a,b]$ for Uniform priors, and $N[\mu,\sigma]$ for the Normal ones) and the results (the mean, $\mu$ with the upper and lower limit at the 68% CL, the $\sigma$ and the peak of the of the posterior distribution). Those parameters without a value indicates they are unconstrained, i.e. that there is no constraint at 68% CL.

| Param | Prior | $H_0 = 74\, km/s/Mpc$ | $H_0 = 67\, km/s/Mpc$ |
|-------|-------|-------------------|-------------------|
|       |       | $\mu \pm 68%CL$ | $\mu \pm 68%CL$ |
|       |       | $\sigma$ | peak | $\sigma$ | peak |
| $\log(M_{\text{min}}/M_\odot)$ | $U[12.0, 14.0]$ | $12.54^{+0.13}_{-0.05}$ | 0.19 | $12.58$ | $12.57^{+0.12}_{-0.06}$ | 0.19 | 12.61 |
| $\log(M_1/M_\odot)$ | $U[12.5, 15.5]$ | $14.29^{+1.02}_{-0.01}$ | 0.76 | $14.83$ | $14.29^{+1.12}_{-0.01}$ | 0.77 | 14.93 |
| $\alpha$ | $U[0.5, 1.5]$ | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ |
| $\Omega_m$ | $U[0.1, 0.8]$ | $0.44^{+0.01}_{-0.15}$ | 0.15 | $0.35$ | $0.49^{+0.02}_{-0.15}$ | 0.15 | 0.41 |
| $\sigma_8$ | $U[0.6, 1.2]$ | $0.75^{+0.01}_{-0.11}$ | 0.10 | $0.69$ | $0.76^{+0.01}_{-0.15}$ | 0.10 | 0.68 |
| $h$ | $N[0.74, 0.014]$ | $0.74^{+0.01}_{-0.01}$ | 0.014 | 0.74 | $N[0.67, 0.005]$ | $0.67^{+0.02}_{-0.003}$ | 0.005 | 0.67 |

Fig. 7. Comparison with external datasets (see text for more details).

of these plots (0.68 and 0.95) are different from those used in the corner plots of this work (0.393 and 0.865, corresponding to the relevant 1-sigma and 2-sigma levels in the 1D histograms in the upper part of the same corner plots).

We also show Planck CMB temperature and polarisation angular power spectra (dark blue) that, although in certain agreement with the HSC (cyan) and DES (green) constraints, presents the tension issues with the CFHTLenS (red) and KV450 (orange) data.

The relevant cosmological constraints derived in this paper are shown in Figure 7 for both samples, $z_{spec}$ and $z_{ph}$, using the “mini-Tiles” scheme. The left panel shows the results from the analysis of each sample individually (grey filled contours for the $z_{ph}$ sample and black dashed curves for the $z_{spec}$ one) while the right panel shows the results from the combination of both samples as described in section 5.3.

With respect to the previous BON20 constraints, by analysing each sample individually, the correction of the large scale bias has shifted the constraints on the $\Omega_m$ parameter toward lower values, more in agreement with the rest of the results from other studies. However, even when combining the two data-sets, the Hubble constant remains unconstrained. There is only a mild preference for the lowest values allowed by the flat prior, which is analogous to the one we found from the $z_{ps}$ sample alone.

As displayed in Figure 7, it is very relevant to underline that when both samples are analyzed together, the constraints in the $\Omega_m, \sigma_8$ plane becomes much more restrictive: $\Omega_m = 0.42^{+0.08}_{-0.14}$ and $\sigma_8 = 0.81^{+0.09}_{-0.06}$. It can also be noted, that their almost per-
magnification bias through the cross-correlation function. Therefore, the large scale biases can affect the cosmological constraint derived from the analysis of the cross-correlation measurements. Therefore, the large scale biases can affect the cosmological constraints derived from the analysis of the cross-correlation function by González-Nuevo et al. (2017), are averaged quantities weighted by the specific redshift distribution of the selected sample.

Therefore, taking into account that the measurements of the cross-correlation function from both foreground samples are independent, we make use of the different redshift distributions to perform a simplified tomographic analysis combining both samples into a single MCMC run. We jointly performed the estimation of the cosmological parameters for both samples, but allowed different values of the astrophysical parameters for each sample. In this way, the effect of having different redshift distributions is included in the astrophysical parameters allowing us to determine with higher precision the cosmological parameters. In fact, the improvements on the \( \Omega_m - \sigma_8 \) plane are evident in the right panel of Figure 7. The cosmological constraints obtained with this independent technique are starting to become competitive with respect to the other lensing results and its particular characteristics make it an interesting possibility in breaking the strong degeneracy between \( \Omega_m \) and \( \sigma_8 \).

As a general conclusion, we showed that we are probably reaching the limits of the constraints that can be derived using just a single redshift bin, although there are still some ways to improve the results. However, the most promising advances with the study of the SMGs magnification bias are probably going to be obtained by performing a more complex tomographic analysis.

Acknowledgements. JGN, MMC, LB, FA and LT acknowledge the PGC 2018 project PGC2018-101948-B-B00 (MICINN/FEDER). LB and JGN also acknowledge PAPII-19-EMERG-11 (Universidad de Oviedo). MM is supported by the program for young researchers ’Rita Levi Montalcini’ year 2015. A.L. acknowledges support from PRIN MIUR 2017 prot. 20173ML3WW002, ‘Opening the ALMA window on the cosmic evolution of gas, stars and supermassive black holes’, the MIUR grant ‘Finanziamento annuale individuale attivitá base di ricerca’, and the EU H2020-MSCA-ITN-2019 Project 866744 ‘BD4BEST: Big Data Applications for black hole Evolution STudies’. We deeply acknowledge the CINECA award under the ISCRA initiative, for the availability of computing resources.

Table 5. Results obtained from combining both sample, the \( z_{\text{spec}} \) and the \( z_{\text{ph}} \), cross-correlation data-sets using the “mini Tiles” scheme. The adopted priors are the same as in Table 4. From left to right, the columns are the parameters and the results (the mean, \( \mu \) with the upper and lower limit at the 68% CL, the \( \sigma \) and the peak of the of the posterior distribution). Those parameters without value indicates they are unconstrained, i.e. that there is no constraint at 68% CL.

| Params | \( \mu \) | \( \sigma \) | peak |
|--------|----------|----------|------|
| \( \log(M_{\text{min}}/M_\odot) \) \( z_{\text{spec}} \) | 12.48$^{+0.21}_{-0.16}$ | 0.18 | 12.51 |
| \( \log(M_1/M_\odot) \) \( z_{\text{spec}} \) | 14.37$^{+1.13}_{-0.36}$ | 0.74 | 15.5 |
| \( \alpha \) \( z_{\text{spec}} \) | – | – | – |
| \( \log(M_{\text{min}}/M_\odot) \) \( z_{\text{ph}} \) | 12.60$^{+0.21}_{-0.12}$ | 0.19 | 12.67 |
| \( \log(M_1/M_\odot) \) \( z_{\text{ph}} \) | 13.69$^{+0.46}_{-0.13}$ | 0.71 | 13.49 |
| \( \alpha \) \( z_{\text{ph}} \) | 0.97$^{+0.05}_{-0.04}$ | 0.27 | 0.88 |
| \( \Omega_m \) | 0.42$^{+0.08}_{-0.04}$ | 0.12 | 0.37 |
| \( \sigma_8 \) | 0.81$^{+0.09}_{-0.09}$ | 0.09 | 0.81 |
| \( h \) | 0.72$^{+0.09}_{-0.22}$ | 0.14 | 0.5 |

In any case, the constraints derived in this work confirm the main conclusions from BON20. Finally, we note that the data here discussed cannot be used to place useful constraints on the Hubble constant yet.

6. Conclusions

As discussed in detail in BON20 (see their Figure A.1) the cosmological parameters depend mainly on the largest angular separation measurements. Therefore, the large scale biases can affect the cosmological constraint derived from the analysis of the magnification bias through the cross-correlation function.

In this work, we study and correct the large scale biases that affect our samples in order to produce a robust estimation of the cross-correlation function. The result is a remarkable agreement between the different cross-correlation measurements, calculated independently of the used Tiling scheme or foreground samples.

Then we analyse these results to estimate cosmological constraints after correcting the different large scale biases. We get minor improvements with respect to the BON20 results, mainly confirming their conclusions: a lower bound on \( \Omega_m > 0.22 \) at 95% C.L. and an upper bound \( \sigma_8 < 0.97 \) at 95% C.L. (results from the \( z_{\text{spec}} \) sample using the “mini-Tile” scheme). Therefore, the large scale biases are a systematic that need to be corrected in order to derive robust and consistent results between different foreground samples or Tiling schemes, but does not help much to improve the precision of the derived constraints.

In addition, we compare the estimates derived using two different and independent foreground samples: one consisting of foreground galaxies with spectroscopic redshifts, the \( z_{\text{spec}} \) sample, and another one with only photometric redshifts, the \( z_{\text{ph}} \) one. Analysing only one single broad redshift bin, we conclude that the higher errors of the photometric redshifts do not have a relevant role in our outcomes. The \( z_{\text{ph}} \) sample here considered has \( \sim 6 \) times more sources than the \( z_{\text{spec}} \) one. Its better surface density makes it more sensitive to some large scale biases but helps to reduce the uncertainty in the measured cross-correlation function at intermediate and small angular scales. On the other hand, our current results show that the uncertainty is still dominated by the cosmic variance rather than by the surface density of the specific foreground sample at the largest angular scales.

However, the constraints obtained making use of the \( z_{\text{ph}} \) sample, which provides a more accurate cross-correlation measurements, are generally consistent with those derived using the \( z_{\text{spec}} \) ones, with similar uncertainties.

Moreover, adopting gaussian priors for the unconstrained parameters (i.e. \( \alpha \) and the Hubble constant, similarly to BON20) does not improve much the results. Therefore, we are probably reaching the accuracy limit of the cosmological constraints that can be achieved with the analysis of a single redshift bin. Increasing the total area in order to decrease further the cosmic variance is probably an interesting improvement to be considered in the future.

Although the measured cross-correlation function is almost the same between both foregrounds samples, we find different constraints for \( \log(M_1) \) and \( \sigma_8 \) parameters. This is caused by the different redshift distributions between both samples. With a single wide redshift bin, the derived astrophysical parameters, that evolve with time as shown in the tomographic analysis of the cross-correlation function by González-Nuevo et al. (2017), are averaged quantities weighted by the specific redshift distribution of the selected sample.

Parameters of interest make it an interesting possibility in breaking the strong degeneracy between \( \Omega_m - \sigma_8 \).}

pendicular direction with respect to the other lensing results can help to break the typical degeneracy.

In any case, the constraints derived in this work confirm the main conclusions from BON20. Finally, we note that the data here discussed cannot be used to place useful constraints on the Hubble constant yet.
high performance computing resources and support. In particular the projects ‘SIS19_lapi’, ‘SIS20_lapi’ in the framework ‘Convenzione triennale SISSA-CINECA’.

The Herschel-ATLAS is a project with Herschel, which is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. The H-ATLAS web-site is http://www.h-atlas.org. GAMA is a joint European–Australasian project based around a spectroscopic campaign using the Anglo–Australian Telescope. The GAMA data input catalogue is based on data input from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KIDS, VISTA VIKING, WISE, Herschel-ATLAS, GAMA II and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA web-site is: http://www.gama-survey.org/.

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 web site 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 of Washington, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, 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.

In this work, we made extensive use of getDist (Lewis 2019), a Python package for analysing and plotting MC samples. In addition, this research has made use of the python packages tools for Python (Jones et al. 2001).
Appendix A: Posterior distributions of the MCMC results

Posterior distributions for the different analyses discussed during the article. The contours for all these plots are set to 0.393 and 0.865. Notice that the relevant 1-sigma and 2-sigma levels for a 2D histogram of samples is 39.3% and 86.5% not 68% and 95%. Otherwise, there is not a direct comparison with the 1D histograms above the contours.
Fig. A.1. MCMC results for the \(z_{\text{spec}}\) data sets (the "mini-Tiles" scheme results in red and the "Tiles" ones in blue).
Fig. A.2. MCMC results for the $z_{ph}$ data sets (the “mini-Tiles” scheme results in red and the “Tiles” ones in blue).
Fig. A.3. MCMC results for the $z_{\text{spec}}$ sample and "mini-Tile" scheme assuming a Gaussian prior for $\alpha$. 
Fig. A.4. MCMC results for the $z_{spe}$ sample and “mini-Tile” scheme assuming two different Gaussian priors for $h$. We adopted the two popular values given by the local estimation, $74.03 \pm 1.42$ km/s/Mpc (blue; Riess et al. 2019), and the CMB one, $67.4 \pm 0.5$ km/s/Mpc (red; Planck Collaboration et al. 2018b).
Fig. A.5. MCMC results combining the $z_{\text{spec}}$ and the $z_{\text{ph}}$ samples in a tomographic run using the "mini-Tile" scheme.