HerMES: Halo Occupation Number and Bias Properties of Dusty Galaxies from Angular Clustering Measurements*

Asantha Cooray1,2, A. Amblard1, L. Wang3, B. Altieri4, V. Arumugam5, R. Auld6, H. Aussel7, T. Babbedge8, A. Blain2, J. Bock2,9, A. Boselli10, V. Buat10, D. Burgarella10, N. Castro-Rodriguez11,12, A. Cava11,12, P. Chianial9, D.L. Clements8, A. Conley13, L. Conversi3, C.D. Dowell2,9, E. Dwek14, S. Eales6, D. Elbaz2, D. Farrah3, M. Fox8, A. Franceschini15, W.Gear6, J. Glenn13, M. Griffin9, M. Halpern16, E. Hatziminaoglou17, E. Ibar18, K. Isaak9, R.J. Ivison18,5, A.A. Khostovan1, G. Lagache19, L. Levenson5,9, N. Lu20, S. Madden7, B. Maffei21, G. Mainetti15, L. Marchetti15, G. Marsden16, K. Mitchell-Wynne1, A.M.J. Mortier8, H.T. Nguyen9,2, B. O’Halloran8, S.J. Oliver3, A. Omont22, M.J. Page23, P. Panuzzo7, A. Papageorgiou6, C.P. Pearson24,25, I. Pérez-Fournon11,12, M. Pohlen6, J.I. Rawlings23, G. Raymond6, D. Rigopoulou24,26, D. Rizzo8, I.G. Roseboom3, M. Rowan-Robinson8, M. Sánchez Portal4, B. Schulz22,20, Douglas Scott16, P. Serra1, N. Seymour21, D.L. Shupe20, A.J. Smith3, J.A. Stevens27, M. Symeonidis23, M. Trichas8, K.E. Tugwell23, M. Vaccari15, I. Valtchanov2, D. Vieira9, L. Vigroux22, R. Ward3, G. Wright18, C.K. Xu27,20, and M. Zemcov2,9

(Affiliations can be found at the references)

Preprint online version: May 20, 2010

ABSTRACT

We measure the angular correlation function, w(θ), from 0.5 to 30 arcminutes of detected sources in two wide fields of the Herschel Multi-tiered Extragalactic Survey (HerMES). Our measurements are consistent with the expected clustering shape from a population of sources that trace the dark matter density field, including non-linear clustering at arcminute angular scales arising from multiple sources that occupy the same dark matter halos. By making use of the halo model to connect the spatial clustering of sources to the dark matter halo distribution, we estimate source bias and halo occupation number for dusty sub-mm galaxies at z ~ 2. We find that sub-mm galaxies with 250 μm flux densities above 30 mJy reside in dark matter halos with mass above (5 ± 4) × 10^{12} h M_{⊙}, while (14 ± 8)% of such sources appear as satellites in more massive halos.

Key words. Cosmology: observations — large-scale structure of Universe — galaxies: high-redshift — submillimeter: galaxies

1. Introduction

The Herschel Multi-tiered Extragalactic Survey (HerMES [Oliver et al. 2010] is the largest project being undertaken by Herschel [Pilbratt et al. 2010]. It surveys a large set of well-known extra-galactic fields (totaling 70 deg2) at various depths, primarily using the Spectral and Photometric Imaging Receiver (SPIRE) instrument (Griffin et al. 2010).

In this Letter, we focus on the clustering of these sources from 0.5 to 30 arcminute angular scales by making use of source catalogues in the two widest HerMES fields, Lockman-SWIRE and the Spitzer First Look Survey (FLS) observed during the Science Demonstration Phase (Oliver et al. 2010). Previous studies on the spatial correlations of sub-mm galaxies was limited to at most 100 sources, leading either to a limit on the clustering amplitude (Blain et al. 2004) or a marginal detection (Scott et al. 2006). While the BLAST source catalog was not used for a measurement of the angular correlation function, clustered fluctuations were detected in a power spectrum analysis of all three bands (Viero et al. 2009).

In the Lockman-SWIRE field we have detected 8154, 4899, and 1680 sources with flux densities above 30 mJy at 250, 350, and 500 μm, respectively, in an area of 218′ × 218′ (Oliver et al. 2010). These counts are supplemented by 3592, 2207, and 1016 sources detected in the FLS field over an area of 155′ × 135′, again down to the same flux density in each of the three bands. These numbers allow clustering estimates at the same precision level as the first-generation of clustering studies at shorter IR wavelengths with source samples from Spitzer data (e.g., Farrah et al. 2006; Magliocchetti et al. 2007; Waddington et al. 2007; Bradwin et al. 2008). Instead of simple power-law models, the correlation functions of HerMES sources have high enough signal-to-noise ratios that we are also able to constrain parameters of a halo model (e.g. Cooray & Sheth 2002).

2. w(θ) measurement

The angular correlation function, w(θ), is a measure of the probability above Poisson fluctuations of finding two galaxies with a separation θ, P(dΩ_1 dΩ_2 = N [1 + w(θ)] dΩ_1 dΩ_2, where N is the surface density of galaxies and dΩ_1 and dΩ_2 are solid angles for each galaxy, corresponding to angle θ. The angular correlation function is of great interest in cosmology as sources are expected to trace the underlying dark matter distribution and the clustering of sources can be related to that of the dark matter halos.

* Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

1 hermes.sussex.ac.uk
For clustering measurements we make use of the HerMES source catalogues (Oliver et al. 2010), based on maps made with calibrated timelines with optimized internal astrometry. The maps were produced using the standard SPIRE pipeline after undergoing calibration and other reduction procedures (Swinyard et al. 2010). In the FLS and Lockman-SWIRE fields, a small number of individual scans have been removed due to artifacts arising from the temperature drift correction. Catalogues were generated using the SUSSEXtractor source extractor in HIPE version 3.0, using a Gaussian PSF with FWHM of 18.15"′, 25.15"′, and 36.3"′ at 250, 350, and 500 µm, respectively. Two independent maps were also produced by dividing the data in time. The sources that do not appear in both sub-maps are flagged as spurious sources and removed from the catalogues. The overall astrometry has been adjusted by comparing with radio positions. We also apply a Wiener filter optimized for unresolved point sources to the maps in the wide fields to remove contamination from cirrus and then we correct source fluxes through simulations. The two fields chosen for this study are largely free from cirrus and the clustering measured with two source samples with and without the Wiener filter applied agree with each other within the errors. Moreover, the source sample used here is restricted to sources detected with an overall significance higher than 5σ, including confusion noise.

Once the catalogues are generated, we use the Landy-Szalay estimator to measure the correlation function with 
\[ \hat{w}(\theta) = \frac{\sum_n (w(\theta_i)w(\theta_j))}{\sum_n (w(\theta_i))w(\theta_j)} \]
where \( DD(\theta) = \) the number of unique pairs of real sources with separation \( \theta \), \( DR(\theta) \) is the number of unique pairs between the real catalogue and a mock sample of sources with random positions, and \( RR(\theta) \) is the number of unique pairs in the random source catalogues (Landy & Szalay 1993). We employ \( 10^3 \) random mocks with \( 10^5 \) sources in each, a larger number of sources than in real data to reduce shot-noise in the random pair counts. Since the measured \( w(\theta) \) does not automatically satisfy the integral constraint (Infante 1994), we estimate the constant necessary to correct \( w(\theta) \) following Adelberger et al. (2005). With a size of \( 10 \) deg\(^2\) for each of our two fields, we find a required value of \( \sim (2 \pm 0.2) \times 10^{-3} \) between 30′ and 40′. This correction is small compared to \( w(\theta \sim 30′) \geq 10^{-2} \).

We calculate the covariance matrix \( C_{ij} \) of the correlation function involving measurements at two different angular scales \( \theta_i \) and \( \theta_j \), using a bootstrap method similar to the one employed by Scranton et al. (2002) for measurements of angular clustering in SDSS DR1. We also calculate the same covariance analytically following the prescription of Eisenstein & Zaldarriaga (2001) and find 10% to 20% smaller off-diagonal correlations than obtained by bootstrapping the data. We compute the usual Poisson errors by taking the square root of the number of pairs, finding that the bootstrap variances are a factor of 1.5 to 2 larger and are a better representation of errors than the Poisson errors. As an example, we show the correlation matrix of \( S_{350} > 30 \) mJy
sources in the Lockman-SWIRE field in Fig. 1, where the correlation matrix is defined as $C_{ij} \equiv \sqrt{C_{ii} C_{jj}}$.

Since we measure $w(\theta)$ directly in the real catalogues, it is likely to be affected by a variety of effects including source blending, flux boosting, and map-making artifacts, among others. To account for all these effects, we compute the transfer function necessary to correct $w(\theta)$ through a set of simulations with number counts and source clustering consistent with data and in a field similar to Lockman-SWIRE. These input maps are then processed by the SPIRE Instrument Simulator (Sibthorpe et al. 2009) for an observational program exactly the same as HerMES observations of Lockman-SWIRE. The output time-ordered data from the simulator are processed as identically as the real data. The transfer function is defined as the ratio of the correlation function in output catalogues to that of the known input used to generate the simulations. The transfer function is generally consistent with unity at angular scales of 1' or more, but varies 25% (at 250 μm) to 50% (500 μm) at angular scales of 3'. We correct the measured $w(\theta)$ in data based on the average of the ratio between input and output $w(\theta)$ computed from about 10 simulations. Due to the finite set of simulations we used the correction is only known to the level of 20% between 1' and 3' angular scales. In comparison, at the same 1' scale, $w(\theta)$ is uncertain at the 30% level at 250 μm and the 50% level at 500 μm. Thus the uncertainty introduced by the error in the transfer function is insignificant compared to the overall uncertainties of the measured correlation function in data.

Fig. 3. (a) Angular clustering of 250 μm sources in Lockman-SWIRE field divided into flux densities between 20 and 35 mJy and above 35 mJy. (b) Angular clustering for the combined 250 μm and 350 μm sample again in the Lockman-SWIRE field with $S_{350}/S_{250} < 0.85$ or $> 0.85$.

Fig. 4. Approximate redshift distribution of sources in the Lockman-SWIRE field with $S_{250} > 30$ mJy (thick solid line in black) and for the two cases based involving colour cuts with $S_{350}/S_{250} > 0.85$ (magenta dashed line) and $S_{350}/S_{250} < 0.85$ (orange dot-dashed line) (see text for details). The thin solid green line and the two shaded regions in the background show example predictions for the $S_{250} > 30$ mJy sample and the two colour cuts, respectively, using models from Le Borgne et al. (2009; thin green line) and Valiante et al. (2009; shaded regions).

3. Halo Modeling of Angular Clustering

In terms of the underlying three-dimensional power spectrum of sources as a function of redshift, $P_s(k, z)$, the projected angular correlation function is

$$w(\theta) = \int dr n^2(r) \int \frac{k dk}{2\pi} P_s(k, r) J_0(k r \theta),$$

where $r(z)$ is the radial comoving distance, $J_0(x)$ is the zeroth order Bessel function and $n(r)$ is the radial distribution of sources normalized to unity: $\int dr n(r) = 1$.

To model $P_s(k)$, we make use of the halo model with both 1- and 2-halo terms (Cooray & Sheth 2002). While the 2-halo term captures the large-scale clustering with the linear power spectrum scaled by the source bias, the 1-halo term captures the non-linear clustering arising from having multiple sources within a single dark matter halo. In this case, multiple sources within a halo are subdivided to a single source at the halo center and one or more satellites. For the central and satellite sources, we describe halo occupation numbers as

$$\langle N_{\text{cen}}(M) \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right],$$

$$\langle N_{\text{sat}}(M) \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log 2 M_{\text{min}}}{\sigma_{\log M}} \right) \right] \left( \frac{M}{M_{\text{sat}}} \right)^{\alpha_s},$$

respectively, where erf$(x)$ is the error function, $M_{\text{min}}$ is the minimum halo mass above which all halos host a central galaxy, and the scatter in the relation between galaxy halo mass and luminosity is captured by $\sigma_{\log M}$. We take a fixed value of 0.3 for $\sigma_{\log M}$, motivated by modeling of clustering at near-IR wavelengths. $M_{\text{sat}}$ is the mass scale at which one satellite galaxy per halo is found, in addition to the central galaxy, and $\alpha_s$ is the power-law slope of the satellite occupation number with halo mass. In addition to these parameters, we also compute the linear bias factor of the sources, which should be interpreted as the average bias factor of the source sample given the redshift distribution $(b_z)$, and the satellite fraction $f_s$, the fraction of sources...
Table 1. Halo model results using the Lockman-SWIRE $w(\theta)$

| Band   | Flux density | $N_{gal}$ | $\langle z \rangle$ | $\log[M_{min}/M_\odot]$ | $\log[M_{sat}/M_\odot]$ | $\alpha_s$ | $\langle b_s \rangle$ | $f_s$ |
|--------|--------------|-----------|---------------------|--------------------------|--------------------------|------------|--------------------|------|
| 250$\mu$m | $S \geq 30$mJy | 8154 | 2.1$^{+0.4}_{-0.1}$ | 12.6$^{+0.3}_{-0.4}$ | 13.1$^{+0.3}_{-0.5}$ | 1.3 $\pm$ 0.4 | 2.9 $\pm$ 0.4 | 0.14 $\pm$ 0.08 |
| 350$\mu$m | $S \geq 30$mJy | 4899 | 2.3$^{+0.4}_{-0.1}$ | 12.9$^{+0.4}_{-0.5}$ | > 13.1 | < 1.8 | 3.2 $\pm$ 0.5 | < 0.20 |
| 500$\mu$m | $S \geq 30$mJy | 1680 | 2.6$^{+0.5}_{-0.1}$ | 13.5$^{+0.5}_{-0.6}$ | > 13.5 | < 1.6 | 3.6 $\pm$ 0.8 | < 0.24 |
| Combined $S_{250}/S_{250} \geq 0.85$ | 3333 | 2.5 $\pm$ 0.4 | 13.4$^{+0.3}_{-0.4}$ | > 13.4 | < 1.9 | 3.4 $\pm$ 0.6 | < 0.19 |
| Combined $S_{350}/S_{250} \leq 0.85$ | 3194 | 1.7$^{+0.5}_{-0.6}$ | 12.8$^{+0.3}_{-0.4}$ | > 12.9 | < 1.9 | 2.6 $\pm$ 0.6 | < 0.26 |

See text below eq. (2) for definitions of $M_{min}$, $M_{sat}$, $\alpha_s$, $\langle b_s \rangle$, and $f_s$. The redshift range is an approximate estimate based on the colour-colour diagram of the source sample through a comparison to isothermal, modified black-body SEDs with a wide range of dust temperatures and emissivity parameters (see, Fig. 4 for an example involving $S_{250} > 30$ mJy and for the two colour cuts).

4. Results & Discussion

In Fig. 2, we summarize our first set of results related to $w(\theta)$ measurements for each of the three SPIRE bands and for sources with $S > 30$mJy. We show correlation functions measured for sources in both Lockman-SWIRE and FLS fields here. We find no statistical difference in the correlation functions of sources detected in the two fields down to the flux density cut-off of 30 mJy. To test for evolutionary hints in clustering, in Fig. 3(a) we split the 250$\mu$m source sample into two bins in flux density, while in Fig. 3(b) we split the combined sample to two colour bins.

To model $w(\theta)$ we need to establish the redshift distribution of the source samples. Given the lack of adequate spectroscopic redshifts, we make use of sub-mm colours to generate a qualitative redshift distribution (e.g., Hughes et al. 2002). First we generate 10$^5$ isothermal SED models using modified black-body spectra with a broad range in dust temperature (10K to 60K) and dust emissivity $\beta$. We also include a 15% Gaussian scatter to model predictions to account for uncertainties in the observed fluxes (Swinyard et al. 2010). We grid the models along the redshift direction in the colour-colour plane to several bins and simply convert the number of observed data points in each bin in the colour-colour plane to a distribution function in redshift (see Fig. 4 for an example). While these distributions are generally consistent with certain model predictions (e.g., Le Borgne et al. 2009; Valiante et al. 2009) and with sub-mm galaxy data (e.g., Chapman et al. 2005), the redshift distribution we recover is strongly sensitive to the SEDs used and should only be considered as an approximate.

Our model fitting results are summarized in Table 1. We assume WMAP 5-year best-fit $\Lambda$CDM cosmology (Komatsu et al. 2009). While in Figs. 2 and 3 we only show the errors from the variances (\sqrt{\sigma^2}), parameter results shown in Table 1 account for the covariance matrix when model fitting to measurements (e.g., Fig. 1 where we show the correlation matrix). Down to the 30 mJy flux density cut, we find average bias factors of 2.9 $\pm$ 0.4 and 3.6 $\pm$ 0.8 for 250 and 500 $\mu$m sources, respectively. Fitting a power-law to all data, the correlation lengths, $r_0$, are 4.5 $\pm$ 0.5 Mpc (250 $\mu$m) and 6.3 $\pm$ 0.7 Mpc (500 $\mu$m). While 250 $\mu$m sources are more likely to be found in halos with mass ($\approx 4 \times 10^{12} M_\odot$), we find that the bright 500 $\mu$m sources in our sample occupy halos of $(3.1 \pm 2.8) \times 10^{13} M_\odot$. The difference is because at a given redshift the 500 $\mu$m sources are at a higher luminosity. Our modeling allows us to establish that (14 $\pm 8$)% of the sources appear as satellites in massive halos than the minimum mass scale. In the case of 350 $\mu$m and 500 $\mu$m source samples, we have failed to accurately determine the parameters related to satellite occupation number. As a test on the validity of our results to uncertainties in $n(z)$, we also considered two extreme possibilities by placing all sources either at $z \sim 1$ or $z \sim 3$ and found parameters within 1$\sigma$ uncertainties of the estimates quoted in Table 1. This is mostly due to the fact that $n(z)$ we use for model fitting, with an example shown in Fig. 4, is broad with a tail to low redshifts.

Our measurements show some evidence for non-linear clustering at arcminute angular scales. Compared to 250$\mu$m, an increase in clustering at arcminute angular scales is less clear at 350 and 500$\mu$m due to the increase in the beam size. In comparison, angular power spectra of BLAST fluctuations did not convincingly reveal a 1-halo term (Viero et al. 2009) and clustering was found to be even below the linear term at smallest angular scales probed. The increase in arcminute-scale angular clustering we see here demonstrates the crucial role played by superior angular resolution of SPIRE. Beyond this initial study, future work involving understanding the large-scale structure distribution of sub-mm galaxies will pursue additional cross-clustering studies of sub-mm sources with optical and shorter IR wavelengths, and studies of unresolved fluctuations. On the modeling side an approach based on the conditional luminosity function could be used to extract additional details on the spatial distribution of Herschel sources.

Acknowledgements. Cooray, Amblard, Serra, Khostovan, and Mitchell-Wynne are supported by NASA funds for US participants in Herschel through an award from JPL. SPIRE has been developed by a consortium of institutes led by Cardiff University (UK) and including Univ. Leibniz (Germany); NAOC (China); CEA, LAM (France); IPCI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); and Caltech/JPL, IPAC, Univ. Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC (UK); and NASA (USA). The data presented in this paper will be released through the HeDaM Database in Marseille.

References

Adelberger, K.L., Steidel, C.C., Pettini, M. et al. 2005, ApJ, 619, 697
Blaizot, A., Chapman, S., Smail, I., Ivison, R. 2004, ApJ, 611, 725
Brodwin, M., Dey, A., Brown, M.J.I. et al. ApJ, 687, L65
Chapman, S.C., Blain, A.W., Smail, I., Ivison, R.J. 2005, ApJ, 622, 772
Cooray, A., Sheth, R. K., 2002, Physics Reports, 372, 1
Farrah, D., Londsdale, C., Borys, C. et al. 2006, ApJ, 643, L139
Franceschini, A., Danese, L., de Zotti, G., Xu, C. 1988, MNRAS, 233, 175
Griessmeier, A., Danese, L., de Zotti, G., Xu, C. 1988, MNRAS, 233, 175
Griffin M.J., Abergel, A., Ade, A.P.R. et al. 2010, this volume
Huges, D., Aretxaga, I., Chapin, E.L. et al. 2002, MNRAS, 335, 871

2 hedam.oamp.fr/HerMES
