MNRAS 000, 1–25 (2017) Preprint 21 December 2017 Compiled using MNRAS LaTeX style file v3.0

Candidate high-z proto-clusters among the *Planck* compact sources, as revealed by *Herschel*-SPIRE

J. Greenslade[1⋆], D. L. Clements[1], T. Cheng[1], G. De Zotti[2,3], D. Scott[4], E. Valiante[5], S. Eales[5], M. N. Bremer[6], H. Dannerbauer[7,8], M. Birkinshaw[6,9], D. Farrah[10], D. L. Harrison[11], M. J. Michałowski[12], I. Valtchanov[13], I. Oteo[14], M. Baes[15], A. Cooray[16] , M. Negrello[7], L. Wang[17,18], P. van der Werf[19], L. Dunne[7,14], S. Dye[20]

[1]*Astrophysics Group, Imperial College, Blackett Laboratory, Prince Consort Road, London, UK*
[2]*SISSA, Via Bonomea 265, 34136, Trieste, Italy*
[3]*INAF-Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, I-35122 Padova, Italy*
[4]*Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T1Z1, Canada*
[5]*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK*
[6]*HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK*
[7]*Instituto de Astrofísica de Canarias (IAC), E-38205 La Laguna, Tenerife, Spain*
[8]*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[9]*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[10]*Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA*
[11]*Institute of Astronomy, University of Cambridge, Madingley Road,Cambridge CB3 0HA, UK*
[12] *Astronomical Observatory Institute, Faculty of Physics, Adam Mickiewicz University, ul. Słoneczna 36, 60-286 Poznań, Poland*
[13]*Herschel Science Centre, European Space Astronomy Centre, ESA, E-28691 Villanueva de la Ca ÌČnada, Spain*
[14]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[15]*Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium*
[16]*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*
[17]*SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD, Groningen, The Netherlands*
[18]*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV, Groningen, The Netherlands*
[19]*Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands*
[20]*School of Physics and Astronomy, University of Nottingham, NG7 2RD, UK*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

By determining the nature of all the *Planck* compact sources within 808.4 deg^2 of large *Herschel* surveys, we have identified 27 candidate proto-clusters of dusty star forming galaxies (DSFGs) that are at least 3σ overdense in either 250, 350 or 500 μm sources. We find roughly half of all the *Planck* compact sources are resolved by *Herschel* into multiple discrete objects, with the other half remaining unresolved by *Herschel*. We find a significant difference between versions of the *Planck* catalogues, with earlier releases hosting a larger fraction of candidate proto-clusters and Galactic Cirrus than later releases, which we ascribe to a difference in the filters used in the creation of the three catalogues. We find a surface density of DSFG candidate proto-clusters of $(3.3 \pm 0.7) \times 10^{-2}$ sources deg^{-2}, in good agreement with previous similar studies. We find that a *Planck* colour selection of $S_{857}/S_{545} < 2$ works well to select candidate proto-clusters, but can miss proto-clusters at $z < 2$. The *Herschel* colours of individual candidate proto-cluster members indicate our candidate proto-clusters all likely all lie at z > 1. Our candidate proto-clusters are a factor of 5 times brighter at 353 GHz than expected from simulations, even in the most conservative estimates. Further observations are needed to confirm whether these candidate proto-clusters are physical clusters, multiple proto-clusters along the line of sight, or chance alignments of unassociated sources.

Key words: galaxies: clusters: general – submillimetre: galaxies – infrared: galaxies – galaxies: evolution

1 INTRODUCTION

The formation epoch of galaxy clusters remains a poorly constrained and understood component in galaxy formation and evolutionary theories. The masses and formation time of these structures in the early universe can not only place key constrains on cosmological theories and parameters (Harrison & Coles 2011), but the elliptical galaxies in the cores of these massive clusters (Kravtsov & Borgani 2012; Ma et al. 2015) are expected to go through an intense starburst phase at z $>$ 2, where a large portion of their stellar mass is rapidly built up over a timescale <1 Gyr (Eisenhardt et al. 2008; Hopkins et al. 2008; Petty et al. 2013; Granato et al. 2015). This starbursting phase should be visible in the far-infrared (FIR) and sub-mm, where cool dust in the galaxies reemit absorbed UV photons. At what point this takes place during the evolution of the cluster remains unknown, and the study and identification of clusters and proto-clusters at z $>$ 2 is important both for cosmology, and for understanding the evolutionary process within massive clusters and their members.

However, few clusters or proto-clusters containing significant numbers of dusty star-bursting galaxies have been detected and confirmed at redshift z $>$ 2 (Daddi et al. 2008; Capak et al. 2011; Walter et al. 2012; Dannerbauer et al. 2014; Yuan et al. 2014; Casey et al. 2015). The rarity of proto-clusters, their large luminosity distance, and lack of an X-ray detectable intra-cluster medium or well formed red-sequence, makes traditional cluster selection techniques ineffective at selecting clusters in the earliest stage of their evolution. The sub-mm, and to a lesser degree the FIR, also benefits from the negative k-correction, enabling reasonably easy identification of sources from redshift 2 to 8, at a fixed wavelength (Blain 2002; Casey et al. 2014).

The dusty star forming galaxies (DSFGs), are thought to play a key role in the evolution of the massive ellipticals primarily seen today in the cores of local clusters (Swinbank et al. 2006; Tacconi et al. 2008; Michałowski et al. 2010; Stevens et al. 2010; Hickox et al. 2012; Casey et al. 2014; Toft et al. 2014; Simpson et al. 2014; Dannerbauer et al. 2014; Ma et al. 2015; Wilkinson et al. 2016). The detection of a large number of physically associated DSFGs would be surprising, as the timescales on which they are expected to be sub-mm bright are only around 100 Myrs, so detecting several physically associated sources either implies some sort of large scale (>1 Mpc) starburst triggering mechanism (Hung et al. 2016; Oteo et al. 2017a), or that these sources are being externally re-fuelled, possibly by cosmic inflows (Casey 2016; Falgarone et al. 2017).

However, several overdensities of sub-mm bright proto-clusters have already been discovered (Herranz et al. 2013; Ivison et al. 2013; Clements et al. 2014; Dannerbauer et al. 2014; Casey 2016; MacKenzie et al. 2017; Oteo et al. 2017a, 2016), some of which have spectroscopic redshifts and ALMA obsevations showing further sub-mm bright members (Ivison et al. 2013, 2016; Oteo et al. 2017a), implying that either a large scale triggering event (>10Mpc) "activates" the DSFGs simultaneously, or alternatively, that the duration of the starburst event is longer (0.5-0.7Gyrs) (Granato et al. 2004; Lapi et al. 2011; Cai et al. 2013; Falgarone et al. 2017). Some evidence exists which suggests that the duty cycle of DSFGs in proto-clusters is indeed longer than those in the field (Emonts et al. 2016; Dannerbauer et al. 2017), with depletion timescales of several hundred Myrs. Overall however, it is uncertain which of these scenarios is correct, and the discovery and study of further proto-clusters and their dusty components is needed, as measurements of the gas depletion timescale imply the former solution is correct, whereas the surface density of sub-mm bright proto-clusters implies the latter is correct. Large field and all sky surveys in the sub-mm, such as *Planck* (Tauber et al. 2010; Planck Collaboration et al. 2011a) or *Herschel* (Pilbratt et al. 2010), are ideal for selecting rare overdensities of DSFGs clustered together on the sky.

Negrello et al. (2005) studied the counts of extragalactic sources expected from low angular resolution surveys such as *Planck*, and concluded that several luminous IR/sub-mm sources clustered on the scale of the instrument beam may appear as an unresolved or marginally resolved source. The individual components that make up these sources could be chance projections along the line of sight, or physically associated. Therefore many *Planck* compact objects might resolve into high-z clusters or proto-clusters of dusty sources when examined with a higher resolution instrument such as SPIRE (Griffin et al. 2010) on the *Herschel* satellite.

Planck has produced three catalogues of compact sources: The early release compact source catalogue (ERCSC, Planck Collaboration et al. 2011b); The Planck catalogue of compact sources (PCCS, Planck Collaboration et al. 2014); and the second Planck catalogue of compact sources (PCCS2, Planck Collaboration et al. 2015b), based on 1.6, 2.6 and 5[1] full surveys of the sky. In each catalogue, the compact *Planck* sources were compiled into nine separate sub-catalogues, one for each *Planck* channel, ranging from 30 to 857 GHz. The beamsizes vary both between channel and between catalogues, but are generally around 4 to 5 arcminutes for the 217, 353, 545 and 857 GHz channels we use here, corresponding to 2 to 2.5 Mpc at $z = 2$. *Herschel*-SPIRE's 350 μm band is matched to *Planck*'s 857 GHz channel, while SPIRE's 500 μm channel has a similar wavelength and passband to *Planck*'s 545 GHz band (500 μm against 550 μm).

Herschel performed several wide surveys over large areas of the extragalactic sky. SPIRE operated at 250, 350 and 500 μm, with beam FWHMs of 17.9, 24.2, and 35.4 arcseconds respectively.[2] The largest extragalactic surveys are the *Herschel*-ATLAS (H-ATLAS, 616 deg^2, Eales et al. 2010), the *Herschel* multi-tiered extragalactic survey (HerMES, including the HerMES large mode survey (HELMS), ~ 370 deg^2, Oliver et al. 2010, 2012) and the *Herschel* Stripe 82 survey (HerS, 79 deg^2, Viero et al. 2014). This provides over 1000 deg^2 of sky with approximate flux density 1σ limits of 7-10 mJy in the three SPIRE bands. By cross-matching the higher resolution *Herschel* map with the catalogues of *Planck* sources, the nature of all the *Planck* compact sources in the *Herschel* fields can be determined.

Several authors have already found plausible high redshift clusters using the *Planck* data (Herranz et al. 2013; Clements et al. 2014; Baes et al. 2014; Planck Collabo-

[1] For the highest frequency channels only.
[2] SPIRE Handbook, Version 3.1, February 8, 2017, http://herschel.esac.esa.int/Docs/SPIRE/spire_handbook.pdf .

ration et al. 2016, 2015a), with a variety of approaches. Both Herranz et al. (2013) and Clements et al. (2014) performed similar cross-matches between *Herschel* and *Planck* in order to search for clusters of DSFGs. Herranz et al. (2013) used 134 deg^2 of preliminary H-ATLAS Phase 1 data and the ERCSC and discovered a redshift 3.26 candidate cluster/proto-cluster of sub-mm sources surrounding the lensed source H12-00 (Fu et al. 2012; Clements et al. 2016). Clements et al. (2014) meanwhile, cross-matched the ERCSC with the HerMES survey, and found evidence for four further candidate proto-clusters of DSFGs, with each candidate proto-cluster having total SFRs $> 1000\mathrm{M}_{\odot}\mathrm{yr}^{-1}$.

Here we set out to investigate and characterise the nature of all the *Planck* compact sources that fall within any of the major *Herschel* fields, using H-ATLAS, HerMES and HerS with the aim of searching for further rare cluster/proto-cluster candidates and potentially other rare and unexpected sources. In general, since we are unable to confirm whether our detected clusters / proto-clusters contain a well developed intracluster medium, and since they generally span scales on the order of arcminutes, we will refer to them as proto-clusters rather than clusters unless otherwise stated. This is prudent given our uncertainties about the evolutionary state of these systems, but we do allow for the possibility that some of our proto-clusters are actually physically evolved clusters.

The rest of this paper is organised as follows: In Section 2, we describe the data sets used in this paper. In Section 3 we outline the methodology used to cross-match with *Herschel*, and present the matches we found between *Planck* and *Herschel*. In Section 4 we verify the photometry of our sources from the *Planck* and *Herschel* observations. In Section 5 we examine the colours of the *Planck*-detected sources, and discuss the likely nature of the reddest sources discovered, whilst in Section 6 we further characterise our candidate proto-clusters. In Sections 7 and 8 we discuss the implications of our findings and summarise our results. Throughout this paper, we assume a standard cosmology, with $H_0 = 67.7$km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$.

2 DATA SETS

In this section, we provide a brief overview of the construction of the three *Planck* and 17 Herschel catalogues used in this paper, as well as the limits of each catalogue and any key differences between them. A summary of the *Herschel* field properties is given in Table 1, and a map of their location on the sky is given in Fig. 1.

2.1 The *Planck* compact catalogues of sources

The ERCSC used *SExtractor* (Bertin & Arnouts 1996) on the *Planck* maps to identify sources in each band; this is based on extracting a number of connected bright pixels that are some threshold above a background measurement. The PCCS and PCCS2, on the other hand, divided the maps into multiple patches, and convolved these patches with a second-order Mexican-hat wavelet that had been locally optimised to detect point sources (López-Caniego et al. 2006). Peaks $> 5\sigma$ in the resulting convolved map were then classified as detections (Planck Collaboration et al. 2014).

We focus on the High Frequency Instrument's (HFI) 857 GHz (350μm) and 545 GHz (550μm) channels (Planck HFI Team et al. 2010), since the peak of dust emission in galaxies (around 100μm) will be redshifted into these bands between $z = 1$ and 5. The quoted FWHM beam-size varies between catalogue releases, between 4.23 to 4.63 arcminutes in the 857 GHz band, and between 4.47 and 4.83 in the 545 GHz band due to improvements in calibration and beam information (Planck Collaboration et al. 2015b). The 90% flux completeness level for the 857 GHz band is given as 680 and 790 mJy[3] at Galactic latitudes |b| > 30 for the PCCS and PCCS2 respectively. The ERCSC does not provide a 90% completeness level, but the faintest source detected at $|b| > 30^o$ is 655 mJy, with the flux density of the faintest 10σ source at $|b| > 30^o$ being 813 mJy, demonstrating that the limits of the three catalogues at 857 GHz are all typically around 700 to 800 mJy. We use the aperture photometry flux density estimate in the *Planck* catalogues, as it performs best when compared to *Herschel* (See table 12 of Planck Collaboration et al. 2015b), is likely to correctly capture emission from extended structures, and is available in all 3 catalogues.

2.2 H-ATLAS

H-ATLAS surveyed five fields: The Northern Galactic Pole (NGP, 170 deg^2), the Southern Galactic Pole (SGP 285 deg^2), and three smaller fields that lie along the equatorial plane at RAs of approximately 9, 12 and 15 hours, referred to as GAMA09, GAMA12 and GAMA15 (around 54 deg^2 each) which correspond to 3 of the fields surveyed by the Galaxy and Mass Assembly (GAMA) project (Driver et al. 2011). Maps were produced with the *Herschel* Interactive Pipeline Environment (*HIPE*, Ott et al. 2010), and the typical 1σ total noise per *Herschel* beam (confusion plus instrumental) in the final background-subtracted and filtered maps is 7.4, 9.4 and 10.2 mJy for the 250, 350 and 500 μm bands respectively (Valiante et al. 2016, Maddox et al. in preparation, Smith et al. in preparation). Sources were extracted using the Multi-band Algorithm for source detection and extraction (*MADX*, Maddox, in preparation).

2.3 HerMES

HerMES field sizes varied from 0.4 deg^2 for GOODS-North, up to 280 deg^2 for the HELMS field. The majority of the fields have 1σ total noises of 6.2 - 6.8, 7.1 - 7.5 and 8.2 - 8.9 mJy for the 250, 350 and 500 μm bands respectively, with the exception of FLS, ADFS, ELAIS-N1, ELAIS-S1, BOOTES and XMM-LSS, which have 1σ noise levels of 7.9, 8.2 and 10.1 mJy (Nguyen et al. 2010). We exclude the HELMS field from further study, since no publicly released, formally verified catalogue of detected sources is yet available for cross-matching, and the field is strongly contaminated with Galactic cirrus. We do examine the *Planck* compact sources present in HELMS using a private catalogue in section 7, but do not include them in our final results. The

[3] This is higher than for the PCCS1 (see section 3.2.3 of Planck Collaboration et al. 2015b).

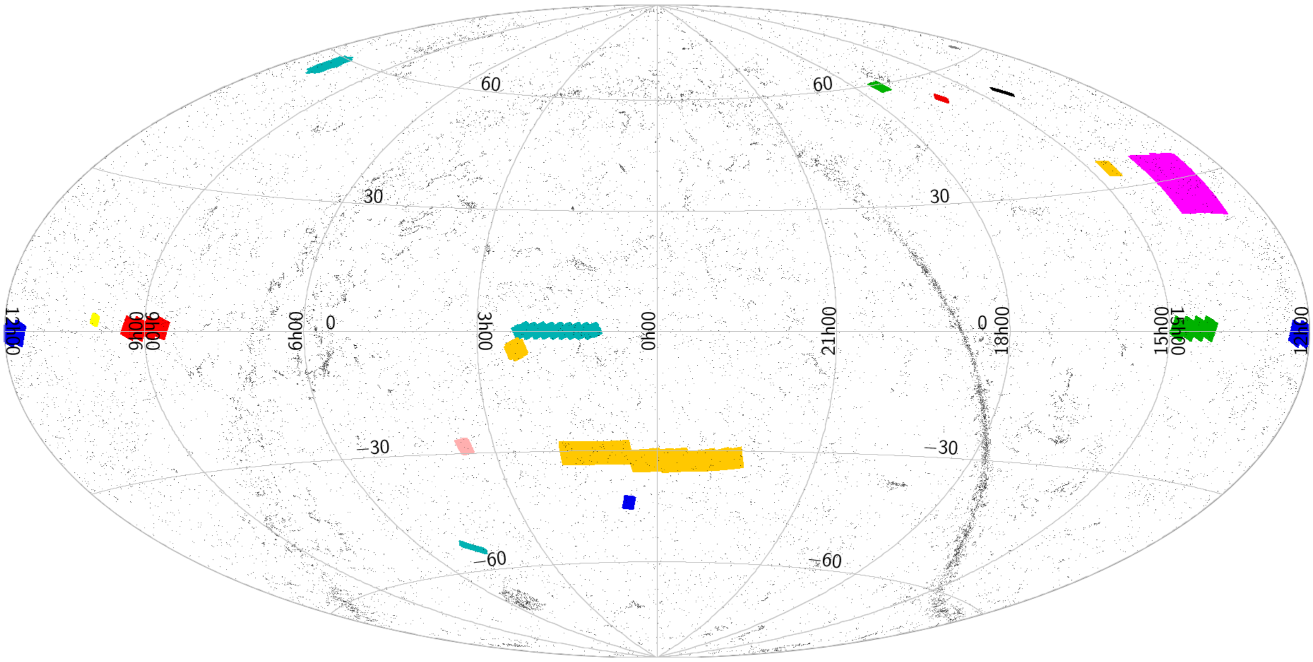

Figure 1. All-sky map showing the 17 *Herschel* fields under examination here (coloured polygons) and all 38,260 *Planck* compact sources (black dots). Some of the major fields include the NGP (Magenta, Dec of 30^o), the SGP (Yellow, Dec of -30^o), HeRS (Turquoise, centre) and the three GAMA fields centred at a Dec of 0^o and RA of 09^h (Red), 12^h (Blue) and 15^h (Green). The Milky Way is indicated by the thick band of *Planck* sources stretching across the sky.

maps used in this paper were produced using the SPIRE-HerMES Iterative Mapper (Levenson et al. 2010), and the catalogues we used were the DR4 xID250 catalogues (Wang et al. 2013).

2.4 HerS

The *Herschel* Stripe 82 Survey (HerS, Viero et al. 2014), is a 79 deg^2 survey taken along the SDSS Stripe 82 region with the SPIRE instrument on *Herschel*. Sources were extracted from the 250 μm map using *STARFINDER* requiring S/N > 3, after filtering the maps with a high pass filter to remove extended emission. Flux estimates were then extracted from the 350 and 500 μm maps, using the 250 μm source positions as a prior. The 1σ median total noise is 7.1, 7.1 and 8.4 mJy for the 250, 350 and 500 μm bands respectively.

		Planck source count	
Field	Area [deg^2]	857GHz	545GHz
NGP	170.0	82	21
SGP	285.0	91	35
GAMA09	53.4	26	13
GAMA12	53.6	15	5
GAMA15	54.6	16	13
ADFS	7.5	3	3
BOOTES	11.3	11	2
CDFS-SWIRE	10.9	5	1
COSMOS HerMES	4.4	2	1
EGS HerMES	2.7	1	0
ELAIS N1 SWIRE	12.3	6	2
ELAIS S1 SWIRE	7.9	2	0
FLS	6.7	5	3
GOODS-North	13.5	0	0
LOCKMAN-SWIRE	16.1	6	3
XMM-LSS-SWIRE	18.9	4	2
HERS	79.0	38	14
Total	808.4	313	118

Table 1. The 17 *Planck*/*Herschel* Fields under consideration in this paper, their areas, and the number of unique *Planck* sources detected within them across all three compact source catalogues.

3 SELECTION METHODS

At 857 and 545 GHz, the *Planck* beam physically corresponds to a size of a few hundred kpc at a redshift of 0.1, and around 2.5 Mpc at redshifts 1 to 3. Therefore, most sources will not be resolved in the *Planck* maps, since only local ($z \ll 1$) extragalactic sources, extended cirrus, or galaxy clusters larger than 2.5 Mpc could have emission extended on larger angular scales. By visually inspecting the *Herschel* maps at the positions of the *Planck* sources, the nature of the *Planck* sources in these regions can be studied.

3.1 Creation of the *Planck-Herschel* catalogue

As different detection pipelines used in the creation of the ERCSC, PCCS and PCCS2 could be sensitive to different source populations, we include all three as part of our analysis. We crossmatch each *Planck* catalogue with the 17 catalogues of *Herschel* sources. We use a search radius equal

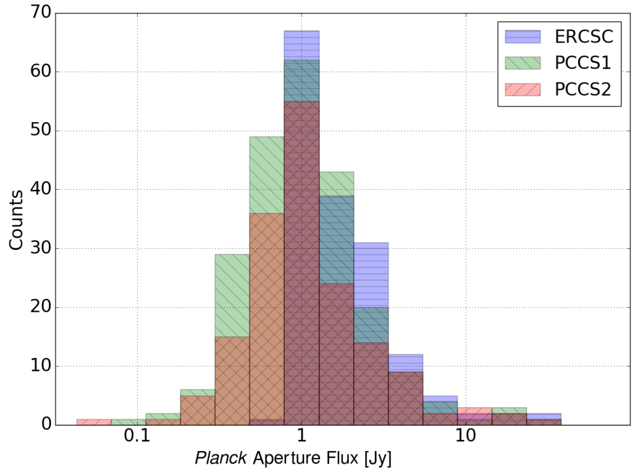

Figure 2. Aperture flux distribution of *Planck* sources that lie in one of our *Herschel* fields from the ERCSC (Blue) , PCCS1 (Green) and PCCS2 (Red).

to the *Planck* FWHM at 857 GHz in the PCCS2, which used the most up to date calibration and beam information (4.63 arcminutes). We varied this search radius between 4.00 and 5.00 arcminutes to check for consistency, as the *Planck* beam FWHM varies not only with channel, but also with *Planck* catalogue, typically between 4.2 and 4.8 arcminutes. With the exception of some minor changes in the number of *Herschel* sources detected in each *Planck* source, our conclusions remained consistent. The *Herschel* source density is high enough that there are always multiple *Herschel* sources per *Planck* beam, typically > 10. For the 857 GHz - 350 μm match, there are 160 *Planck* sources in the *Herschel* fields from the ERCSC, 229 from the PCCS1 and 168 from the PCCS2. The 545 GHz - 500 μm match finds 50 *Planck* sources from the ERCSC, 99 from the PCCS1 and 60 from the PCCS2.

In Fig. 2, we plot the aperture flux values for our *Planck* sources. While the PCCS and PCCS2 appear to be similar in terms of their flux distribution, the ERCSC distribution is skewed towards higher flux values. The ERCSC, using *SExtractor*, requires isolated, bright, connected pixels in order to flag a detection, with the minimum flux found for the whole ERCSC being 655 mJy at 857 GHz. The PCCS and PCCS2, on the other hand, require a single local peak in the *Planck* map after convolution with the filter, and so can contain $> 5\sigma$ sources with aperture fluxes as low as 69 mJy in our catalogue. However, for bright sources detected in all three catalogues, the distributions should be similar, and above an aperture flux of $\sim$ 750 mJy, we find a much better match between the source flux densities.

Cross-matching the three versions of the catalogues together to find the total number of *Planck* compact sources detected in at least one of the three catalogues, we find 313 *Planck* sources in the *Herschel* fields from the 857 GHz band and 118 unique *Planck* sources from the 545 GHz band. Combining these two catalogues together to search for unique objects, we find a total of 354 sources detected across 808.4 deg^2

We also create a catalogue of *Herschel* sources that fall within 4.63 arcminutes of each *Planck* source. We created two uniform catalogues of *Herschel* sources for the 857 GHz and 545 GHz data by selecting *Herschel* sources using a minimum flux density limit of 25.4 mJy at 250, 350 or 500 μm (i.e. a source must be at least 25.4 mJy in one of the three SPIRE bands). This is approximately 3 times the highest median total error seen in any of the *Herschel* fields. This results in 3,709 individual sources with $S_{350} > 25.4$ mJy ($\gtrsim 3\sigma$) that lie within 4.63 arcminutes of a *Planck* 736 GHz source, and 693 *Herschel* sources with $S_{500} > 25.4$ mJy that lie within 4.63 arcminutes of a *Planck* 545 GHz source.

Finally, we cross-matched our catalogue of unique compact *Planck* objects with the *Planck* Sunyaev-Zel'dovich Galaxy Cluster Catalogue (Planck Collaboration et al. 2015c), the *Planck* Galactic cold clump catalogue (Planck Collaboration et al. 2015d), and the *Planck* High Z catalogue (PHZ, Planck Collaboration et al. 2016). We found no matches with the SZ catalogue, a single match with the Galactic cold cores catalogue, PLCKERC857 G339.76-85.56, and four matches in the PHZ, PCCS1 545 G160.59-56.75, PCCS1 545 G084.81+46.34, PLCKERC545 G007.56-64.14 and PCCS1 545 G012.89-66.24.

3.2 The nature of the *Planck* sources

We visually inspected each source in the *Herschel* 350 μm maps at the position of the *Planck* objects, to identify the nature of each *Planck* source. A summary of our results is presented in Table 2, a full table of identifications is available in Appendix B, and images of the 324 that lie on the maps and away from the edge are available in Appendix A.

Most local ($z \ll 1$) galaxies can be identified by their bright, point source or extended emission in the *Herschel* maps. Cross-matching these with the NASA Extragalactic Database (NED) identifies 192 local galaxies, two QSOs and eight lens candidates that have known H-ATLAS identifications. Four times, single bright sources with $S_{350} > 50$ mJy are found to have no optical or other known counterparts in NED or elsewhere. These we assign as additional lens candidates, though these could also easily be examples of hyper luminous infrared galaxies, with $L_{fir} > 10^{13} L_{\odot}$, and are note necessarily lensed Sources were also cross-matched with SIMBAD (Wenger et al. 2000), and three stars were identified this way. Fourteen of the *Planck* sources lie just outside the map coverage, and these are included in Table 2 but not considered further.

For the remaining 131 sources, as well as examining the *Herschel* maps, we examined the Improved Reprocessing of the IRAS Survey (IRIS Miville Deschenes & Lagache 2005) maps at the positions of the *Planck* sources to search for bright emission at 100 μm, which will be present for Galactic cirrus but not for proto-clusters of DSFGs at redshifts $\gtrsim 1$. *Planck* objects with structures in the 100 μm map were conservatively catagorised as Galactic cirrus, 43 in total. This left 88 regions without an identification.

To search for proto-clusters amongst these 88, we counted the number of 250, 350 and 500 μm sources with fluxes > 25.4 mJy that lie within 4.63 arcminutes of the *Planck* position, with the flux limit chosen to compare to published number counts. Assuming our sources are Poisson distributed, number counts from Clements et al. (2010) and Valiante et al. (2016) suggest that the expected number of

Table 2. Identifications of all the *Planck* objects that fall within one of the *Herschel* survey fields under consideration here.

Type	857 GHz	545 GHz	Unique
Local galaxies	187	54	192
Galactic cirrus	37	18	43
Proto-cluster candidates	21	10	27
Lensed sources	12	2	12
Stars	3	2	3
QSOs	2	1	2
Off Map	13	3	14
No Assignment Given	38	28	61
Total	313	118	354

250, 350 and 500 μm sources are 16.5 $\pm$ 4.1, 9.1 $\pm$ 3.0 and 2.7 $\pm$ 1.7 per *Planck* beam. Any objects that show a 3σ overdensity in any of the three SPIRE bands (at least 31, 19 or 9 sources[4] in the 250, 350 or 500 μm bands, respectively) are classed as candidate proto-clusters of galaxies, 27 in total are found in this way. These over-densities are not necessarily physical associated proto-clusters, as they could also be line of sight effects of unrelated sources, multiple clusters / proto-clusters along the same line of sight (Flores-Cacho et al. 2016; Negrello et al. 2016), or they might be explained by differences in the actual distributions of the number of *Herschel* sources in the tail of the distribution compared to Poisson. The assumption of Poisson oversimplifies the complex distribution of galaxies, so in order to justify our assumption, we simulate 10,000 *Planck* beams (circles of radius 4.63 arcminutes) at random positions on the NGP 350 μm map, and count the number of *Herschel* sources with $S_{350} > 25.4$ mJy. We then and compare this to our Poisson assumption that 19 or more sources indicates an overdensity. Only 16 of 10,000 of the random positions contain at least 19 *Herschel* sources with $S_{350} > 25.4$ mJy, with an average of 8.8 ± 2.9 per *Planck* beam, in good agreement with the estimates from Valiante et al. (2016). Interpreting the 16 out of 10,000 as a probability, and converting this to an equivalent σ value in the normal distribution, this corresponds to a 2.94σ overdensity, in excellent agreement with our choice of assuming these sources are Poisson distributed. We find similar results for the 250 and 500 μm bands. Therefore, these 27 *Planck* sources are clearly overdense in *Herschel* sources, and we assign them as candidate proto-clusters, though we retain the possibility that these are line of sight effects or multiple clusters / proto-clusters along the line of sight remains. The remaining 61 sources in our maps remain unclassified, as we cannot reliably determine their nature. We thus have a total of 340 unique *Planck* compact sources across both the 857 GHz and 545 GHz bands, including local galaxies, galactic cirrus, proto-cluster candidates, lensed sources, stars, QSO's and sources we were unable to assign a classification.

[4] Using Poisson statistics

3.3 Properties of our catalogues

3.3.1 Nature of the unassaigned sources

We cannot reliably assign a catagory for several of our sources. These could be false detections by *Planck* or a series of fainter sources which we do not detect in *Herschel*. Significant differences at low S/N were seen from preliminary versions of the catalogues, which were created from preliminary versions of the *Planck* maps. (D. Harrison, private communication). Keeping the parameters for the *Planck* catalogue creation the same, sources near the detection threshold would appear / disappear, depending on the preliminary version of the map used in the creation of the *Planck* catalogue. For sources detected at a high S/N, this was very rare, whilst for sources near the detection thresholds, this was more common.

In our catalogue, for sources not assigned a counterpart, the median detection level in the PCCS and PCCS2 is 5.4 $\pm$ 0.5σ, near the detection threshold of 5σ (for our proto-clusters, this is similar at 5.4 $\pm$ 0.3σ).However, ten of the 65 are detected in multiple catalogues (six of these were detected in both the ERCSC and either the PCCS or PCCS2, thus using different detection methods). This is unlikely if these ten sources are false detections. Of these ten, five have colours that would be selected as a high redshift candidate by the Planck High Z collaboration (PHZ) in their analysis of candidate high-z sources in *Planck* (Planck Collaboration et al. 2016). This could indicate an overdensity of red compact sources, too faint to be included in our analysis. This conclusion was also hinted at when varying our search radius between 4.00 and 5.00 arcminutes; several of our unassigned sources became classified as candidate proto-clusters, and several candidate proto-clusters became unassigned. In all cases we found roughly 30 candidate proto-clusters, with the exact number depending both on our choice of search radius, and flux density limit. It is therefore likely that some of the unassigned sources are proto-clusters of DSFGs, but for the specific values we have chosen they do not pass our threshold test.

3.3.2 Diffuse and Dominated Sources

Given we are searching for proto-clusters, we take all the *Herschel* sources associated with a *Planck* 857 GHz object, and calculate the standard deviation of their *Herschel* S_{350} flux densities, σ_{350}. A large value of σ_{350} is likely due to singular bright sources, whereas a small value indicates either of multiple distinct sources as in a proto-cluster, or simply extended Galactic cirrus. We do the same for the 545 GHz *Planck* sources and the *Herschel* S_{500} flux densities, σ_{500}. We show these in Fig. 3 for all 342 *Planck* objects, as well as the result when taking 1000 random positions, and calculating the Log σ_{350} in each case as a comparison. Any source with fewer than two *Herschel* sources is not included in our analysis. There are 28 sources with 2,3 or 4 associated 350 μm detections, so the vast majority have reasonable samples from which to calculate σ_{350}. The distribution appears bi-modal, with two distinct regions below and above $\mathrm{Log}_{10}(\sigma) \approx 1.65$. This bi-modality is not seen when examining 1,000 random positions. We designate these two regions as "diffuse" ($\mathrm{Log}_{10}(\sigma) < 1.65$) and "dominated" ($\mathrm{Log}_{10}(\sigma) >$ 1.65), indicating that flux from these sources appears to be

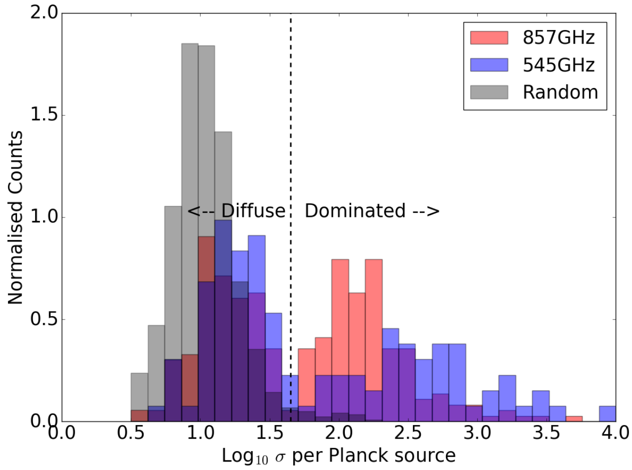

Figure 3. Log of the dispersion at 350 μm for the *Herschel* sources contained within a *Planck* object in the 857 (Red) and the 545 (Blue) GHz catalogues. The vertical black dashed line indicates the selected division between "diffuse" and "dominated" sources. In grey is the result from taking 1,000 random positions in the NGP field, showing very few sources in the "dominated" region.

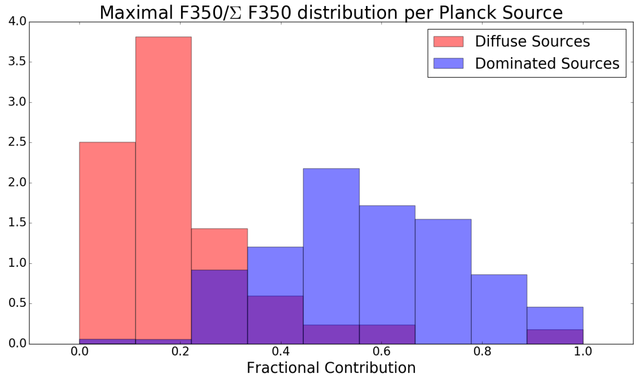

Figure 4. Fractional contribution of the brightest *Herschel* source in each *Planck* source to the total *Herschel* 350 μm flux density from all the sources associated with each *Planck* object. "diffuse" sources (red) and "dominated" sources (blue) are plotted separately.

from extended diffuse / multiple faint source emission or dominated by a single source respectively.

For the 857 GHz *Planck* sources, of the 299 sources not near the edge and with more than 1 associated *Herschel* sources, 155 sources are identified as "Dominated" and 144 identified as "Diffuse". In the 545 GHz catalogue, of the 109 sources not near the edge and with more than 1 associated *Herschel* sources, 44 are "Dominated" and 65 are "Diffuse". Overall this resulted in 159 unique "dominated" sources and 186 unique "diffuse" sources, with 9 sources having only 1 counterpart or lying near the edge of the *Herschel* map. All the cirrus sources, all the proto-cluster candidates and all but one of the not assigned sources are identified as being "diffuse". The other 156 "dominated" sources are all identified with local galaxies, lensed candidates, the QSO or stars. Of the 186 total diffuse sources, 41 are associated with local galaxies, usually because of extended emission or several bright neighbours. We also find that four of the lens sources are diffuse, though they lie on the border between diffuse and dominated.

In Figure 4 we plot the distribution of the fractional contribution from the brightest 350 μm *Herschel* source to each *Planck* 857 GHz source, divided by whether a *Planck* source is "diffuse" or "dominated". This independently shows that our intuitive explanation for the division seen in the σ_{350} seems to be the correct one; "dominated" objects tend to have one bright source dominating the flux whereas the "diffuse" objects individually have a relatively low contribution to the total flux. A similar relationship is seen in the 545 GHz data.

The clear divide in both Fig. 3 and Fig. 4 indicate that only around 60% of the *Planck* compact sources are actually compact on scales reasonably smaller than the *Planck* beam. Both figures also show that the *Planck* maps are well suited for detecting extended emission from sources such as proto-clusters of DSFGs.

Table 3. Fractional make up of the three *Planck* catalogues of compact sources at 857 GHz

Source Type	ERCSC [%]	PCCS1[%]	PCCS2[%]
Local Galaxies	56.0	61.0	80.0
Galactic Cirrus	16.7	8.8	5.0
Cluster Candidates	9.5	4.6	1.1
No Assignment Given	11.9	16.1	5.6
Lenses	1.2	3.8	3.3
QSO	0.0	0.8	0.5
Stars	0.6	1.1	1.7

3.3.3 Variations between the ERCSC, PCCS and PCCS2

The key difference between the *Planck* compact source catalogues is the use of *SExtractor* for the ERCSC and a Mexican-hat wavelet for the detection pipeline in the PCCS and PCCS2. This latter approach was designed to suppress emission on large scales, in order to reduce cirrus contamination in the catalogues, and simulations of its effectiveness were run on point sources (López-Caniego et al. 2006). However, its effect on extended, non-cirrus sources is unclear.

In Table 3, we provide the fractional composition of each 857 GHz catalogue. Though from the ERCSC to the PCCS2, the cirrus contamination of the catalogues has reduced from 16.7% to 5.0%, the fraction of proto-cluster candidates has been also reduced from 9.5% to 1.1%. Put another way, the fraction of "diffuse" sources has decreased from $\sim$ 47% in the ERCSC to 28% in the PCCS2. Though these proto-cluster candidates may not be real, may be line of sight effects, or potentially cirrus contamination, recent work has shown that several of these candidates are consistent with their being clusters in formation at z $\sim$ 2 (Herranz et al. 2013; Clements et al. 2014, 2016, Cheng et al. in preparation). The inclusion of the Mexican-hat wavelet for source detection potentially suppresses the detection of these proto-cluster candidates, as the Bootes, EGS, Lockman and CDFS proto-cluster candidates revealed by Clements et al. (2014) do not appear in the PCCS1 or PCCS2.

4 PHOTOMETRY

Having identified our 27 proto-cluster candidates, alongside numerous other source types, we now examine the photometry associated with these sources. *Planck* have previously compared their photometry against *Herschel* in order to verify that the two photometry measurements agree (Bertincourt et al. 2016). In this section, we extend this analysis to checking whether summing our selected 350 μm *Herschel* sources (i.e. $S_{350} > 25.4$ mJy) alone can adequately match the *Planck* flux densities seen in all the *Planck* compact sources.

As the band passes are well matched, a direct comparison between the 857 GHz *Planck* band and the 350 μm SPIRE band can be performed with few assumptions. Here, we follow the same procedure set out in appendix A.1. of the PCCS1 for estimating the aperture photometry, but use the *Herschel* maps instead of the *Planck* maps. We took the background subtracted maps of all of the *Herschel* fields, and integrated the SPIRE 350 μm flux density over a *Planck* 857 GHz beam by summing all the pixels that fell within 1 FWHM of the nominal *Planck* source position. The assumed FWHM was 4.63 arcminutes. Once again, this was varied between 4.0 and 5.0 arcminutes to check for consistency in the results, finding similar results. A background annulus of inner radius 1 × FWHM and outer radius of 2 × FWHM was used to estimate the median background value and this was removed from the aperture flux estimate. Any sources that fell on the edge of the map or contained null pixels within the primary or background aperture had a flux density assigned to them of zero to prevent edge effects contaminating our sample. Errors were estimated from a combination of SPIRE instrumental noise, SPIRE calibration error, and a constant confusion noise conservatively estimated at 7 mJy per SPIRE-beam, all added in quadrature. The results of this analysis, for both diffuse and dominated sources, are shown in Figure 5.

We then use the absolute relative flux density difference, defined as

$$\eta = |100 \times \frac{S_{SPIRE} - S_{Planck}}{S_{SPIRE}}|, \tag{1}$$

and use the weighted average of the *Planck* and *Herschel* aperture photometry, finding an absolute relative flux density difference between *Planck* and *Herschel* of only 4.9%, comparable to the 1 to 5% uncertainty found in Bertincourt et al. (2016).

The absolute relative flux density difference is, however, not the same for the dominated (1.8%) and diffuse (11.4%) sources. Given we are using background subtracted maps in each case, we repeat our analysis using the raw H-ATLAS maps that are publically avaliable. These *Herschel* maps have not had any background subtraction applied to them, and therefore could contain the flux that appears to be missing in several of our diffuse sources for *Herschel*. We found that absolute relative flux density difference for our dominated and diffuse sources changed to 4.8% and 3.8% respectively when we used the raw maps, both well within the *Planck* calibration uncertainty. This indicates that the missing flux from our sources, especially diffuse sources, is being removed during the background removal process on the *Herschel* maps.

The *Planck* and *Herschel* aperture photometry are generally in agreement for *Planck* objects dominated by a single *Herschel* source. Given roughly 40% of all *Planck* compact objects are expected to be diffuse in nature when examined at *Herschel* resolutions, we consider whether the detected sources alone can account for the total *Planck* flux, or whether an extended diffuse emission component is required.

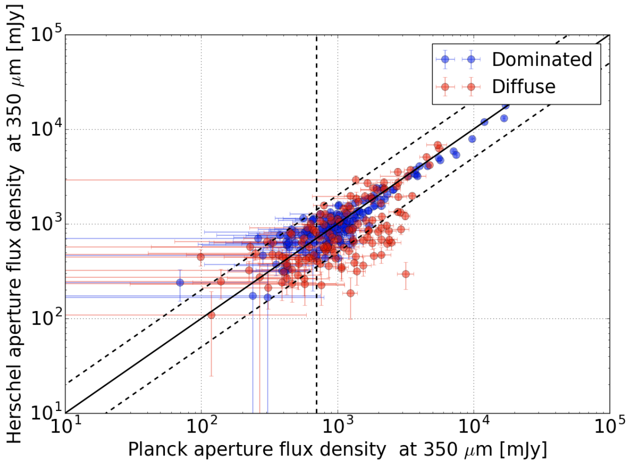

Figure 5. Comparison between the *Planck* aperture flux density and the *Herschel* aperture flux density, as calculated in the text. The red points are the those sources considered to be diffuse, and the blue those considered dominated by a single source. The solid black line shows the 1:1 ratio. The diagonal dashed lines show the limits where the *Herschel* flux is half/double that of the *Planck* flux, and the vertical dashed line shows the PCCS 90% completeness limit.

In Fig. 6, we plot the *Planck* aperture flux densities and the summed 350 μm fluxes from the detected *Herschel* sources, coloured by their source classification type. We find a 5% absolute relative flux density difference between the summed fluxes and the aperture flux for non cirrus sources, but an 77% relative flux difference for sources we have identified with galactic cirrus. Several local galaxies, with emission extended well beyond the scale of the *Herschel* beam, are poorly fit in the *Herschel* catalogues and therefore have a smaller summed-*Herschel* flux compared to the *Planck* flux.

When summing up detected *Herschel* sources, proto-cluster candidates are well matched to *Planck* but Galactic cirrus sources are not, suggesting that our selection of Cirrus sources in Section 3 was successful. This also implies that estimates of the physical properties of these proto-cluster candidates can be derived from the *Planck* flux density alone, as it represents the summed total of the individual sources that make up the proto-cluster and no diffuse emission is needed to account for the *Planck* flux.

Fig. 6 also shows that the proto-cluster candidates mostly lie near to the *Planck* detection limits, with a median *Planck* aperture flux of 886 mJy. Only eight of our 21 candidate proto-clusters detected at 857 GHz have an 857 GHz flux density > 1 Jy. For the unassigned objects, 14 of 63 have *Planck* 857 GHz flux densities > 1 Jy, and these brighter sources we often find are not well matched between *Planck* and *Herschel*; Only two of these unassigned sources have *Herschel* aperture flux densities > 1 Jy, and none of the

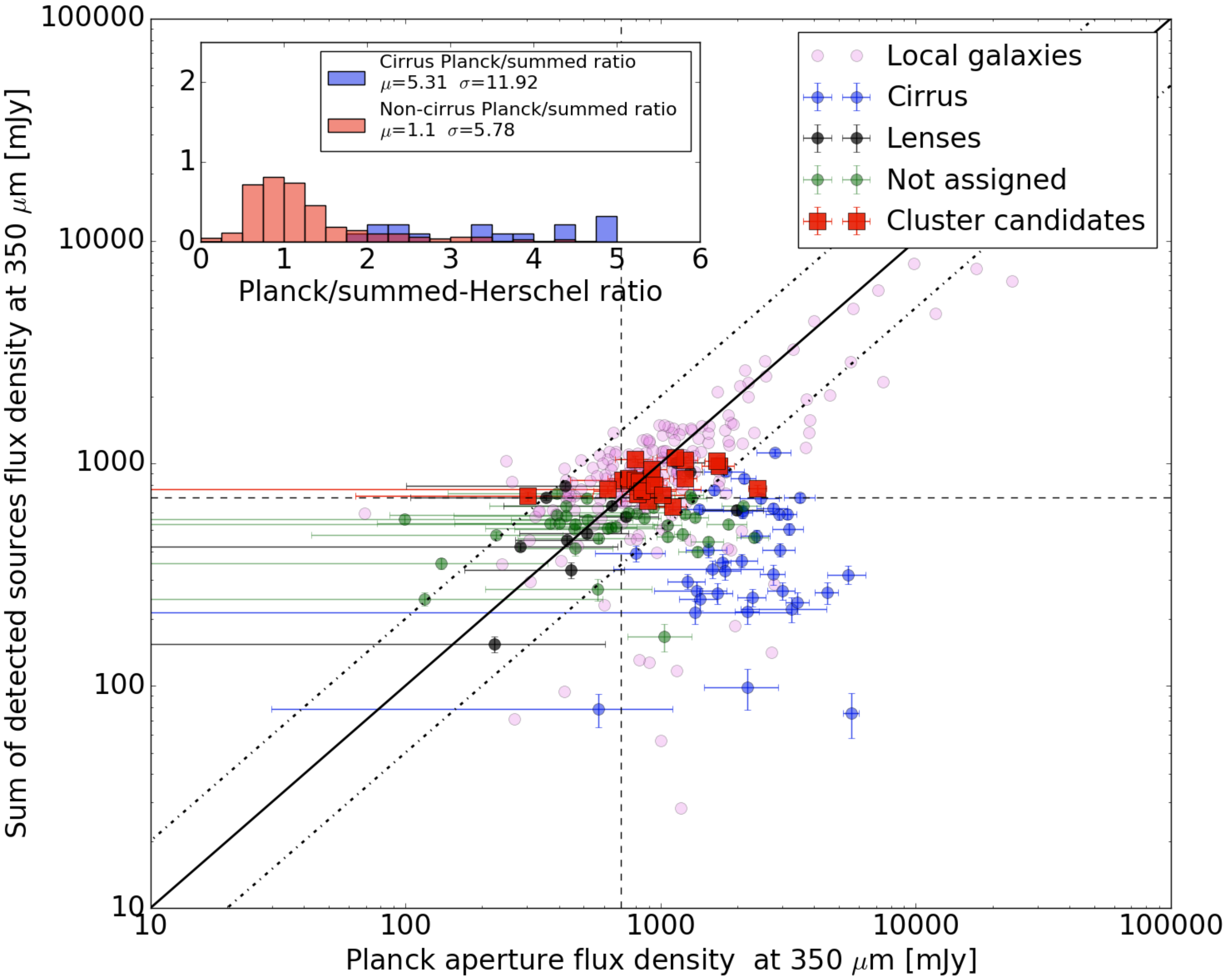

Figure 6. Comparison between the *Planck* aperture flux density and summing up the 350 μm flux density from the detected *Herschel* sources. The light pink points are local galaxies, the blue are cirrus, the red are proto-cluster candidates, the black are lens candidates and the green are those points not assigned an identification. The solid black line shows the 1:1 ratio, whereas the dot-dashed lines show where the *Planck* aperture flux is double or half the summed detected sources. The vertical and horizontal dashed lines show the nominal *Planck* 90% completeness levels from the PCCS1. Error-bars are not shown for the local galaxies to aid in clarity, but are comparable to other sources at all fluxes. The histogram in the top left corner shows the *Herschel Planck* ratio, with cirrus sources indicated in blue, and non-cirrus sources indicated in red, as well as the mean and standard deviation. The histogram has been truncated to a maximum ratio of 6 for clarity, with 19 cirrus sources with ratios beyond this.

unassigned sources have a 857 GHz flux density > 1 Jy when summing detected *Herschel* sources. Given also that Fig. 2 indicates the ERCSC, which appears to be best at detecting these proto-cluster candidates, is limited to sources with flux density > 750 mJy, it is possible that the candidates we are selecting here are the bright tail of the DSFG proto-cluster population, and there could be many more proto-clusters that lie below this limit.

5 COLOURS

With only a maximum of 3 photometric points from SPIRE available, any photometric redshift attempt will have large uncertainties ($\Delta z = \pm 1$) associated with it. However, the sub-mm colours of *Herschel* sources have often been used as a proxy to give a useful indication of their redshifts (Clements et al. 2014; Dowell et al. 2014; Dannerbauer et al. 2014; Asboth et al. 2016; Rowan-Robinson et al. 2016; Ivison et al. 2016). Therefore, in this section we set out to examine the *Planck* colours of our sources, and compare them to the selection used by the PHZ in their search for high-z sources, as well as using the *Herschel* colours to give an indication of the likely redshifts of our *Planck* sources. We leave a more accurate determination of the redshift to a future paper that contains additional follow up observations (Cheng et al. in preparation).

5.1 *Planck* Colours

In Figure 7, we plot the *Planck* 857/545 GHz (350/550 μm) and 545/353 GHz (550/850 μm) colours for the major populations identified in Section 3.2. We only plot sources from the 857 GHz selected catalogue, since it is the only catalogue which additionally provides aperture flux estimates at 545 and 353 GHz at the position of the *Planck* source. Planck Collaboration et al. (2016), in their selection of high-z candidates from the *Planck* maps, used a criterion with *Planck* colours of 857/545 GHz < 2 and 545/353 GHz > 1 to search for candidate high redshift galaxies/clusters of galaxies. We mark their selection area as the gray hashed region. For clarity the local galaxies that are detected at 3σ in all three of the the 857, 545 and 353 GHz bands are plotted. We also plot two of the proto-clusters detected by Clements et al. (2014) in the Bootes and EGS fields to demonstrate their colours (both of which are also detected in our analysis).

We note that many of our proto-cluster candidates fall outside the *Planck* selection region. For our identified candidate proto-clusters, 21 are included in the 857GHz *Planck* catalogue, and so are considered here. Of these 21, only twelve lie within the *Planck* selection region, with a mean S_{857}/S_{545} ratio of 2.0 ± 0.5. As the only constraint we impose upon our sources is that they are detected as a *Planck* compact source, and lie in one of the major *Herschel* fields, we could be selecting a population of lower redshift or warmer clusters / proto-clusters than found by Planck Collaboration et al. (2016).

Local galaxies and cirrus have mean 857/545 colours of 3.0 ± 1.0 and 2.8 ± 0.7 respectively, whereas the unassigned sources have a colour of 2.5 ± 1.0. For the unassigned sources, nine of the 35 have colours that would have been selected in Planck Collaboration et al. (2016) as potentially high redshift. It is therefore not unreasonable to suggest that unassigned sources with both red colours and a large, but not overdense, number of *Herschel* sources could also be high-redshift proto-clusters of *Herschel* sources. Our lens candidates have a median 857/545 GHz colour of of 1.8 ± 0.5, and our QSO has 857/545 GHz colour of 0.8 ± 0.4 at a redshift of 2.099. The three stars have a mean 857/545 GHz colour of 3.0 ± 0.4. As expected, the stars, local galaxies and cirrus all have 857/545 colours that indicate that they are at redshifts $\ll 1$, whereas the redshift 2.099 QSO, lens candidates and our proto-cluster candidates have colours that indicate they lie at redshifts > 1.

The total colour from a candidate proto-cluster will be a combination of foreground/background sources and sources associated with the proto- cluster. This is especially important, considering that overdensities of *Herschel* sources have been argued to be due to line of sight effects from multiple clusters, both theoretically (Negrello et al. 2016) and observationally (Flores-Cacho et al. 2016). In order to assess the contribution from foreground sources to the colour of a *Planck* source, we simulated the *Planck* colours of a region of sky containing a proto-cluster. Our simulated proto-clusters have, on average, 11 members which would be selected by our flux cutoff, and we include contribution from sources not associated with the proto-cluster by adding in, on average, 9 sources which would be selected by our flux cutoff randomly distributed between redshifts 1 and 3. The total number of detected sources in then around 20, which is just high enough to be selected as a candidate proto-cluster for our sample. For all sources, we drew samples from a single temperature modified blackbody function

$$S_\nu \propto \nu^\beta B_\nu(T), \tag{2}$$

where ν^β modifies the emissivity function of the dust and $B_\nu(T)$ is the Planck function at temperature T. The temperature was fixed at 29 K and β was fixed at 2, so that the background sources have an average S_{857}/S_{545} flux density ratio that matches that seen in the *Herschel* maps, in this case $S_{857}/S_{545} = 1.87$. The fluxes of each source are drawn from an exponential distribution, which roughly matches the distribution of fluxes we see in our catalogues of 350 μm detected *Herschel* sources, and our 350 μm flux is then normalised to this value. We simulate 4 proto-clusters in total, at redshifts 1, 2, 3 and 4, and for each redshift we draw 100proto- clusters using the method described above. We determined the total colour by summing the total 857 GHz flux density and dividing by the total 545 GHz flux density from all sources. The results of this are shown in Fig. 8.

We find that when there are few proto-cluster sources compared to background/foreground sources, the colours tend to the average colours of the foreground/background sources, as expected, and in this case with an average of $S_{857}/S_{545} = 1.87$. Once there are roughly equal number of proto-cluster sources and background/foreground sources however, the proto-cluster tends to dominate the colour of the source. However, that colour is dominated by the redshift of the source, with proto-clusters at redshift 3 and 4 having a lower S_{857}/S_{545} flux density ratio, proto-clusters at redshift 1 having a higher S_{857}/S_{545} flux density ratio. If a proto-cluster is roughly at the same redshift as the average redshift of the background/foreground sources, then there is no obvious difference in its colour compared to a patch of sky where there is no proto-cluster. This provides a simple explanation for the "warm" proto-cluster candidates, that they are lower redshift clusters / proto-clusters compared to the likely high-z clusters detected in the PHZ (Planck Collaboration et al. 2016). However, we note in particular that our results are very sensitive to the assumption that all our galaxies are the same temperature; even if we allow the temperature to vary by ±5 K, the standard deviation in the S_{857}/S_{545} flux density ratios of proto-clusters can double from 0.1 to 0.2 for a proto-cluster at redshift 2. This further suggests that there can be significant boosting both into and out of the selection region used by the PHZ, though the general trend remains that higher redshift proto-clusters tend to have lower S_{857}/S_{545} flux density ratios. The major benefit used in this paper compared to the PHZ is that we do not make any colour selection, and are therefore sensitive to clusters / proto-clusters at all redshifts where we would detect them by our flux cut.

5.2 *Herschel* Colours

The use of *Herschel*-SPIRE colour-colour diagrams to separate sources of different redshifts is well established (e.g. Herranz et al. 2013; Noble et al. 2013; Clements et al. 2014; Ivison et al. 2016; Negrello et al. 2016), though the precise interpretation of the results are uncertain. Typically, sources whose SED peak at longer wavelengths tend to lie at higher redshifts (Casey et al. 2014; Dowell et al. 2014; Asboth et al.

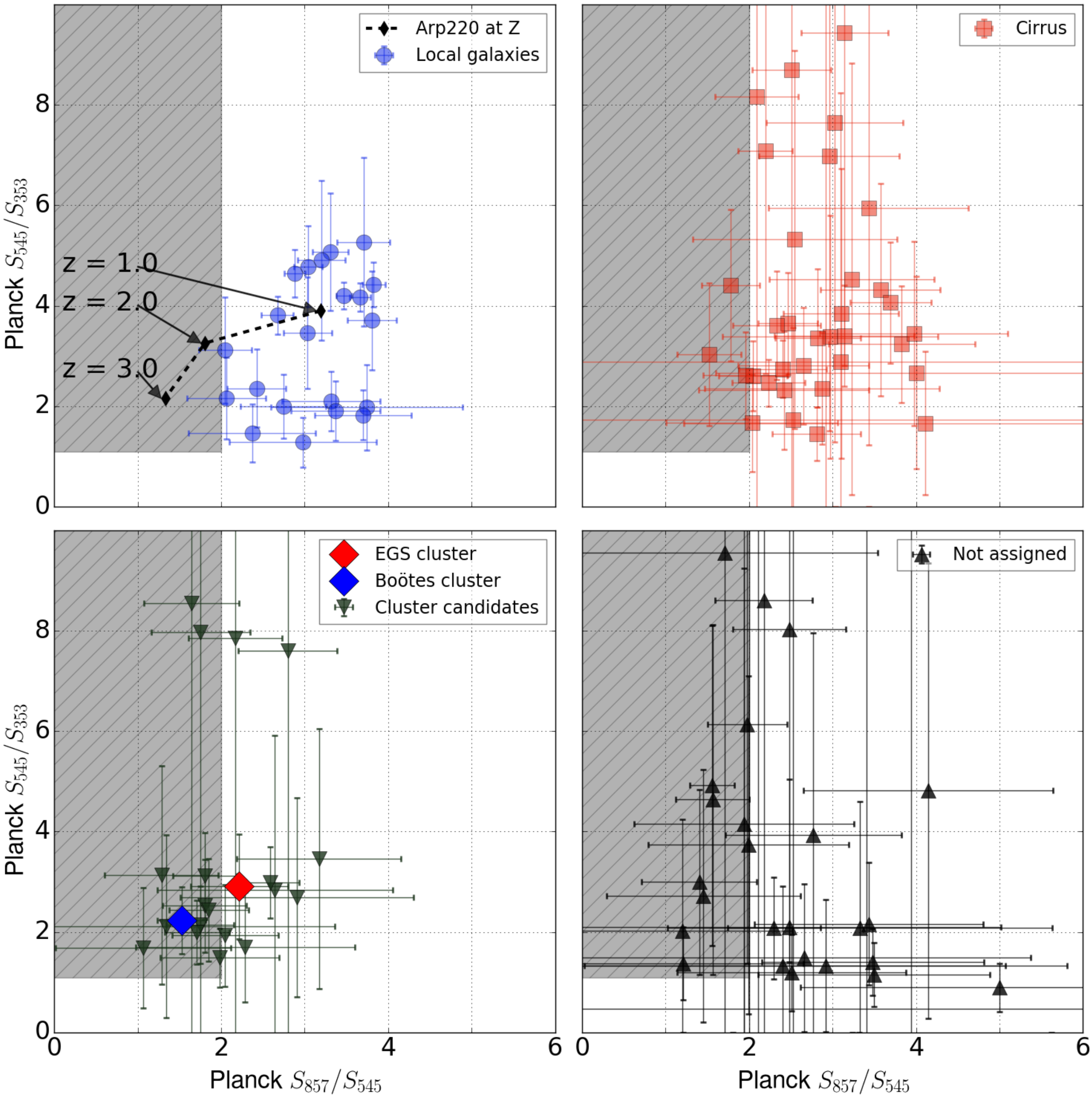

Figure 7. *Planck* 857/545 GHz and 545/353 GHz colours for the categories of source we identify as local galaxies (top left), cirrus sources (top right), cluster candidates (bottom left) and unassigned sources (bottom right) . The grey shaded region represents the selection criteria used in Planck Collaboration et al. (2016) for their selection of high redshift source candidates. The black line in the top left plot shows the *Planck* colours of Arp 220 as it would appear at z=1,2 and 3, and the blue and red diamonds in the proto-cluster candidates plot show, respectively, the Bootes and EGS proto-cluster candidates identified in Clements et al. (2014).

2016; Ivison et al. 2016), and therefore sources whose SED peak at 250, 350 and 500 μm likely indicate progressively higher redshifts.

In Fig. 9, we simulate the *Herschel* colours, again using a single temperature modified blackbody function, in an attempt to show the rough redshift a source is likely to have, given its *Herschel* colours. We fix the redshifts at 0, 2 and 4, where we expect our sources to approximately peak in the 3 SPIRE bands, and uniformly distribute the temperatures and β values between 20 and 60 K and 1 and 2.5, respectively. Figure 9 shows that the *Herschel* colours of a source can provide a good proxy for the redshift of that source.

To compare to our simulation, in Fig. 10, we plot the individual S_{250}/S_{350} and S_{350}/S_{500} *Herschel* colours for the local galaxies and proto-cluster candidate *Planck* sources. Any local galaxy extended on arcminute scales, or where ex-

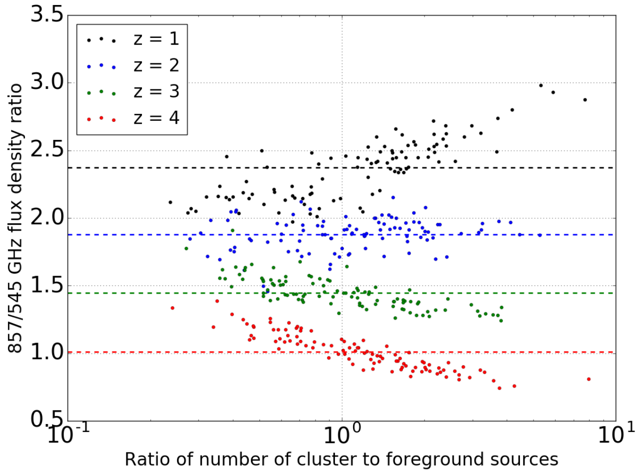

Figure 8. Estimated *Planck* 857/545 GHz flux density ratio of 400 proto-clusters, as a ratio of the number of proto-cluster to background/foreground sources. Points in black are proto-clusters at a redshift of one, in blue at a redshift of two, in green at a redshift of three and in red at a redshift of four. The dashed lines show the average colour of the 100 proto-clusters at each redshift. Large symbols show proto-clusters which would be selected by our 3σ overdensity requirement, with small labels showing proto-clusters that would not be.

traction on the *Herschel* map has clearly divided the source into multiple sources were removed. For the 250/350 μm and 350/500 μm colours of the local galaxies, we find a mean of 2.05 ± 0.43 and 2.60 ± 0.74, respectively, whereas for the proto-cluster candidates these values are 1.13 ± 0.47 and 1.57 ± 0.49. Fig. 10 clearly divides into two regions, one bluer region associated with the local galaxies, and one red region where the bulk of the *Herschel* detected proto-cluster candidates lie. Similar to the *Planck* colours, the *Herschel* colours of the proto-cluster candidates are on average redder than for the local galaxies. At the same time, we take a template starburst galaxy, Arp 220 (Donley et al. 2007), and examine the *Herschel* colours as it would appear at various redshifts. Direct comparison suggests the proto-cluster candidates lie at a redshift of $\simeq 2$. This is also in good agreement with our estimates of single temperature blackbody fits in Fig. 9, with local galaxies inhabiting the low redshift region and proto-cluster candidates inhabiting the region suggested for redshifts between 2 and 4. However, estimating the redshift of sources from its Herschel colours alone can be difficult; often the errors are large, and the variation seen in Figures 9 and 10 alone is enough to make the true redshift of a source uncertain. Given the simulations, observed error, and variation we see here, we can therefore reasonably say these sources likely lie at $z > 1$, but little more until future follow up work can constrain these sources further.

6 THE CANDIDATE CLUSTERS

Out of the 279 unique *Planck* sources we have identified, 27 appear to be $> 3\sigma$ overdensities of *Herschel* sources. The photometry of these objects indicates that the flux density comes from a number of discrete, individual sources, and their colours indicate that they likely lie at $z \sim 2$. These observations could correspond to a physical cluster of DSFGs, a series of line of sight sources stretching from $z \sim 2$ to ~ 4, or multiple clusters / proto-clusters along the line of sight. In this section, we attempt to quantify these proto-cluster candidates further, and examine whether the large area surveyed can explain these sources through fluctuations in the number counts alone.

6.1 Probability of observing $> N$ sources by chance

If our candidate proto-clusters are actually only line of sight or number count fluctuations, then it should be possible to model the probability of finding one using Poisson statistics. In Fig. 11, we sample the NGP field with 1,000 random *Planck* beams of radius 4.63 arcminutes, and count the number of 250, 350, and 500μm sources with fluxes greater than 25.4mJy in each of the three respective bands. We then plot the normalised version of this sample, as well as histograms of the numbers of *Herschel* sources associated with each of our candidate proto-clusters from the 857 GHz band. Our candidate proto-clusters are clearly overdense with respect to our random samples of 1,000 positions. The mean number of associated *Herschel* sources for our proto-cluster candidates is 29.1, 20.6 and 10.7 for the 250, 350 and 500 μm bands respectively, corresponding to a 2.9, 3.5 and 4.0σ overdensity respectively [5].

Given we here examine roughly 800 deg^2 of sky, and according to Poisson statistics, we may expect to find around 89.0 patches where there are 26 or more 250 μm sources, 33 regions where there are 10.2 or more 350 μm sources, and 1.3 regions where there are 11 or more 500 μm sources. If all our proto-clusters were only this overdense, this might explain our results, however, many of our proto-clusters host far stronger overdensities, with 14 of our proto-cluster candidates containing ≥ 36, 23 or 12 250, 350 and 500 μm sources respectively (with maximal numbers of associated *Herschel* sources of 43, 32 and 17 for the three bands). Over 800 deg^2 of sky, we would therefore expect to see 0.5, 1.5 and 0.3 patches containing ≥ 36, 23 or 12 250, 350 and 500 μm sources, if they were Poisson distributed. We in fact see 4 patches at least this overdense in the 250 μm band, 8 at least this overdense in the 350 μm band, and 10 at least this overdense in the 500 μm band, which cannot be explained solely by the large area surveyed in this paper. Our candidate proto-clusters are therefore likely to be physically associated or be the product of several clusters / proto-clusters or overdensities along the line of sight.

We would still expect some level of contaimination from unassociated sources. Under the assumption that the *Herschel* sources are a mix of proto-cluster members and Poisson distributed unassociated sources, for an expected μ sources, the probability that there are N proto-cluster sources out of M detected sources is given by:

$$p(N|M,\mu) = \frac{\left[\frac{\mu^{(M-N)}}{(M-N)!}\right]}{\Sigma_{i=0}^{i=M}\left[\frac{\mu^{(i)}}{(i)!}\right]}, \tag{3}$$

[5] These probabilities have been converted to their corresponding σ value in the Normal distribution.

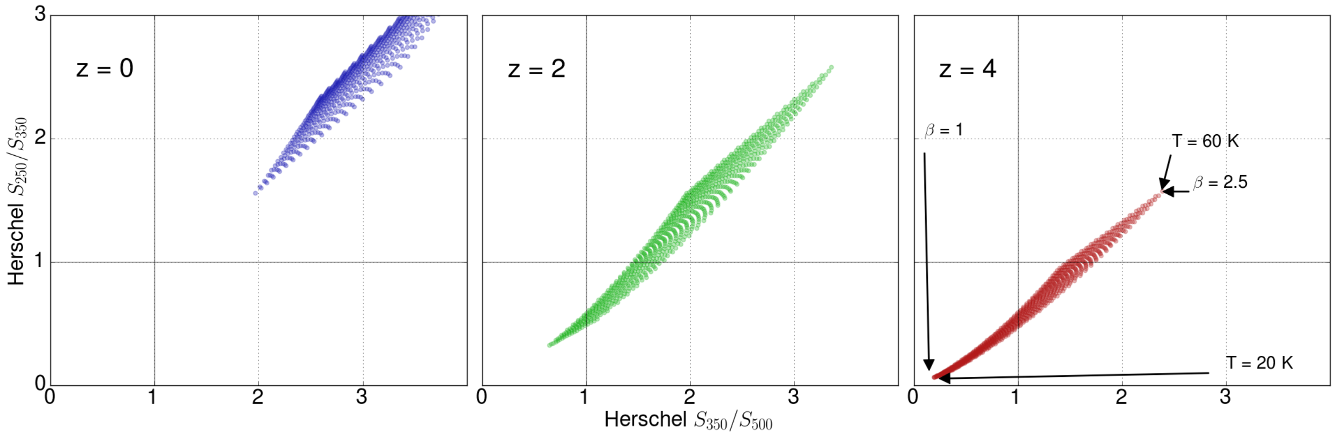

Figure 9. The *Herschel* S_{250}/S_{350} vs S_{350}/S_{500} colours for modified blackbodies (see text for more detail) at redshifts of 0, 2 and 4, allowing T and β to vary between 20 and 60 K and 1 and 2.5, respectively. In the far right plot, the regions of maximum and minimum T and β are indicated for clarity.

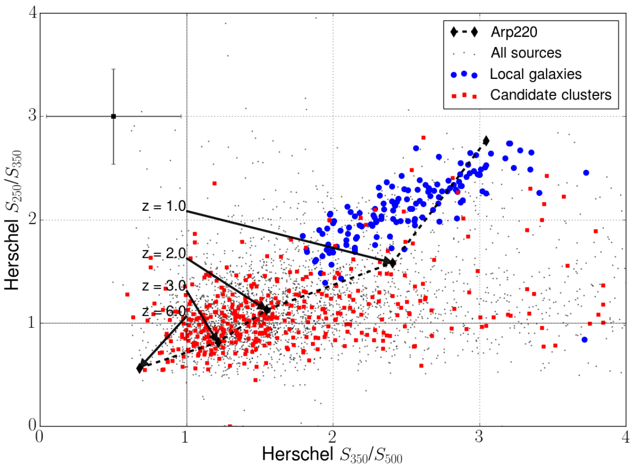

Figure 10. *Herschel* S_{250}/S_{350} and S_{350}/S_{500} colours of local galaxies (blue circles) or all the Herschel sources associated with the proto-cluster candidates (red squares). The small black circles include all *Herschel* sources detected for all of our *Planck* sources. Typical errors are given on the left (black square). The dashed black line with the black diamonds shows the *Herschel* colours of the local ULIRG Arp 220, as it would appear at z = 2,4 and 6.

the derivation of which is given in Appendix C. For our mean of 20.6 350 μm *Herschel* sources associated with each cluster, this suggests that on average, around 11 of the sources would be associated with the proto-cluster, with only a 0.7% chance of having 3 or fewer proto-cluster members.

Though we do not have accurate redshifts for our sources, we can get some idea if they lie at similar redshifts by examining where the individual *Herschel* sources for a single proto-cluster candidate lie in colour-colour space. In Fig. 12, we plot the *Herschel* colours for the *Herschel* components of three of our *Planck* sources; the Bootes proto-cluster identified by Clements et al. (2014), a candidate proto-cluster PCCS1 857 G085.48+43.36 identified in this work, and a cirrus source. The Bootes proto-cluster and the ELAIS-N1 proto-cluster show clear clustering in the colour colour plot, whereas the cirrus source shows a much larger spread. From Fig. 9, this clustering in colour-colour space indicates it is likely these sources are physically associated, but the uncertainties are large enough that we cannot rule out the possibility that these multiple clusters / proto-clusters along the line of sight.

6.2 Properties of the proto-cluster candidates

Given our previous analysis, in the following sections we assume that all 27 of our candidate proto-clusters are physically assocated proto-clusters or multiple clusters / proto-clusters along the line as sight, as opposed to chance over-densities along the line of sight. We find a surface density of candidate proto-clusters of $(3.3\pm0.7)\times10^{-2}$ sources deg^{-2}. In their assessment of the number of *Planck* detectable clusters, Clements et al. (2014) find a surface density of $(4.4\pm2.2)\times10^{-2}$ sources deg^{-2}, in good agreement with our results here. Planck Collaboration et al. (2016) in the PHZ, searched directly on the *Planck* maps, discovering a total of 2,151 candidate high-z sources across around 10,000 deg^2 of the cleanest part of the sky, with intial follow up suggesting 94% of these are overdensities of sources (Planck Collaboration et al. 2015a). Given the different selection functions used in the PHZ and this paper, it is difficult to make a direct comparison, but this would correspond to a approximate surface density of (0.18 ± 0.01) sources deg^{-2}, roughly 5 times larger than found here. This can be somewhat offset if we include our sources where do not not assign a classification, as our surface density rises to (0.11 ± 0.02) sources deg^{-2}, in closer agreement with the PHZ. Further follow up of the PHZ sources, especially at the fainter end, is needed to investigate the discrepancies.

The number counts within individual fields mostly agree with the estimated number counts given here, with 10 out of an expected 11 from the SGP, seven out of an expected six for the NGP, zero out of four for HERS (which has large amounts of Galactic cirrus), and roughly one in each of the smaller HerMES fields. The GAMA fields are lack-

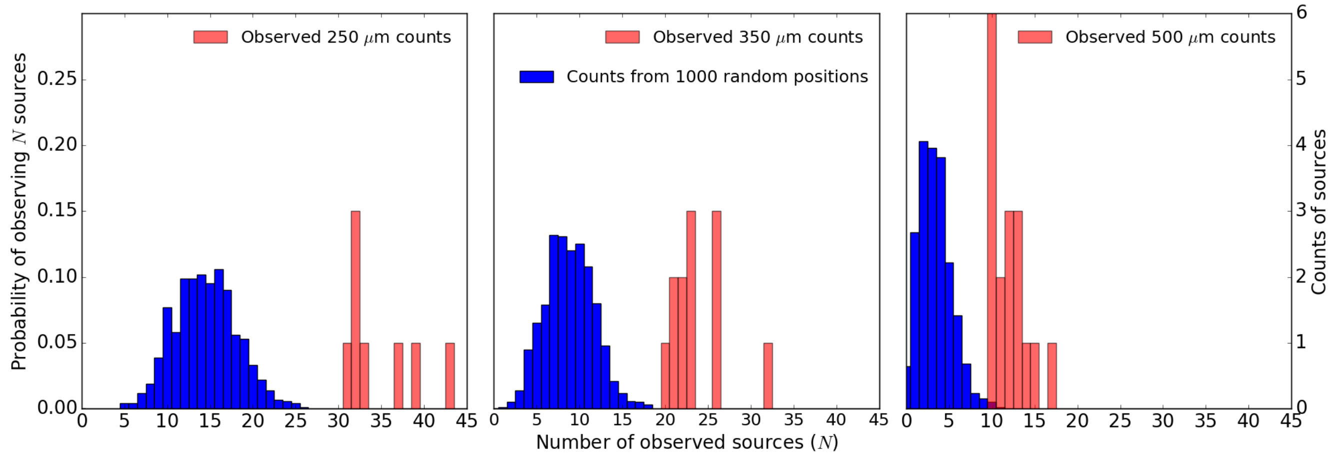

Figure 11. (Blue) histograms of the result when 1,000 random *Planck* beams are placed on the NGP map and the number of sources with S_{350}, S_{250}, or $S_{500} > 24.5$ mJy are counted for: (left) The 250 μm band, (middle) the 350 μm band and (right) the 500 μm band. (Red) histograms of the observed numbers of 250, 350 and 500 μm sources for our candidate proto-clusters, which are considered overdense in their respective bands.

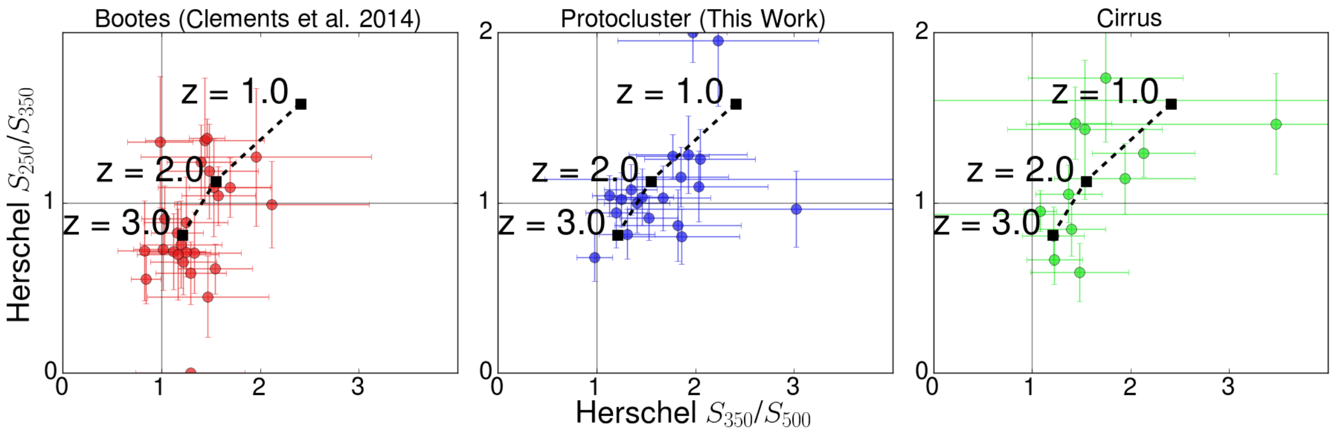

Figure 12. *Herschel* S_{350}/S_{500} S_{250}/S_{350} plot showing the colours for three *Planck* sources; The Bootes clump identified by Clements et al. (2014) on the left; a candidate cluster seen in the ELAIS-N1 field in the centre, and a source identified with Galactic cirrus on the right. The black dashed line and squares indicate the *Herschel* S_{350}/S_{500} and S_{250}/S_{350} of the local ULIRG Arp 220, as it would appear at z = 1.0, 2.0 and 3.0.

ing in sources, with no proto-cluster candidates detected in any of them. The lack of objects in GAMA could be due to the large amount of foreground cirrus present in GAMA09 and GAMA15, which could obscure a number of candidate proto-clusters.

Many confirmed proto-clusters are found to be extended on scales of tens of arcminutes (Casey 2016). The smaller *Planck* beam implies that we are detecting highly compact systems of DSFGs, compared to generic proto-clusters which tend to show less of a density contrast with respect to the background (Casey 2016). For instance, the Bootes proto-cluster candidate appears to be at a redshift of z $\sim$ 2.3. Pearson et al. (2013) estimate the redshift distribution of sources in the phase 1 release of H-ATLAS, where they find there should be roughly 10-100 *Herschel* sources per square degree with $F_{350} > 35$ mJy at a redshift $\sim$ 2, or roughly 0.2-1.5 sources per *Planck* beam. Using the definition of Chiang et al. (2013) of density contrast:

$$\delta_{gal}(\mathbf{x}) = \frac{n_{gal}(\mathbf{x}) - < n_{gal} >}{< n_{gal} >} \quad (4)$$

and a simple photo-z fitter, which fits our *Herschel* sources to a SED template of Arp 220, we find 12 sources with $F_{350} > 35$ whose photo-z is consistent within 1σ of z= 2.3, giving a density contrast between $\delta(12) = 7 - 60$, depending on whether one uses a low or high estimate of the density of *Herschel* sources at z $= 2.3$. The low density contrast estimate is still consistent with these sources being proto-clusters, but for density contrasts of > 10 this becomes more difficult to understand; the large density contrasts imply that these are systems which are well on their way to collapse and virialization. However two of our candidate proto-clusters appear to be associated with known galaxy clusters; PCCS1 545 G058.72+82.59 (PCCS1 857 G058.53+82.57) lies 4.3 arcminutes away from the core of galaxy cluster GHO 1319+3023 (Gunn et al. 1986) at a redshift of 0.4, PCCS1 545 G027.38+84.85 (PLCKERC857 G027.36+84.83) is associated with the redshift 0.43 galaxy cluster GMBCG J198.59994+26.5688 (Hao et al. 2010) and PCCS1 545 G084.40+81.05 is associated with the estimated redshift 0.43 galaxy cluster NSCS J131812+335831 (Lopes et al. 2004). Given our earlier estimates on the redshift of our souces being at $z > 1$, it is possible that our cluster of DSFGs is being lensed by a foreground cluster, rather than that they are physically associated with the foreground cluster. Three of our proto-clusters, PLCKERC857 G017.86-68.67, PLCKERC857 G149.81+50.11 and PLCKERC857 G095.44+58.94, also appear to host QSOs that are mostly, not emitting in the FIR. Again, whether or not these QSOs are associated with the cluster of DSFGs is uncertain, but they they're redshifts are typically between $z = 1$ to 2, so could be signposting the true redshifts of our proto-clusters.

6.3 Simulations of DSFGs in clusters

Granato et al. (2015) simulate the FIR/sub-mm properties of high-redshift clusters and proto-clusters by combining hydrodynamical simulations with GRASIL-3D, a radiative transfer code that accounts for dust reprocessing in arbitrary geometries. In Fig. 13 we comapre the number counts of clusters of DSFGs from the Herschel data with the predicted number counts obtained by Granato et al. (2015), assuming their 24 simulated clusters, which all had a final virial mass at z = 0 above $1 \times 10^{15} h^{-1} M_{\odot}$, are representative of the cluster population we detect here. We impose a 3σ S/N cut for each band considered, and use the aperture photometry estimate from *Planck*. Again, we assume that all our 27 candidate proto-clusters are actual physical clusters of sources. Our observations indicate that our detected clumps are more numerous, or are brighter, than predicted from these simulations. The flux density from our proto-clusters appear to be on average ~ 5 times greater than predicted.

In Figure 14, we show that this is likely due to our observed sources being brighter than expected in simulations, by reproducing the histogram of expected 350, 550 and 850 μm flux densities from Granato et al. (2015), and comparing the distribution of the 350 μm flux densities of the proto-clusters identified in this work. Since some of the flux from our proto-cluster candidates will come from sources not associated with the proto-cluster, we attempt to remove this foreground contribution. We place 1000 *Planck* beams at random positions on each of the *Herschel* maps, calculate the total flux density in those beams following the same prescription in section 4, and take the median value of the aperture fluxes over those 1000 beams as the typical foreground contamination. The median value varies between maps, but is usually of the order of 100 to 300 mJy. These values are then removed from the *Herschel* aperture flux densities for each of the proto-clusters, and the results plotted in Fig. 14. The difference between our observed flux densities and the simulated clusters is exacerbated at higher redshifts, as the simulated flux densities tend to decrease (Granato et al. 2015). The original plots in Granato et al. (2015) split the data into three separate redshift bins at z = 1,2 and 3, with

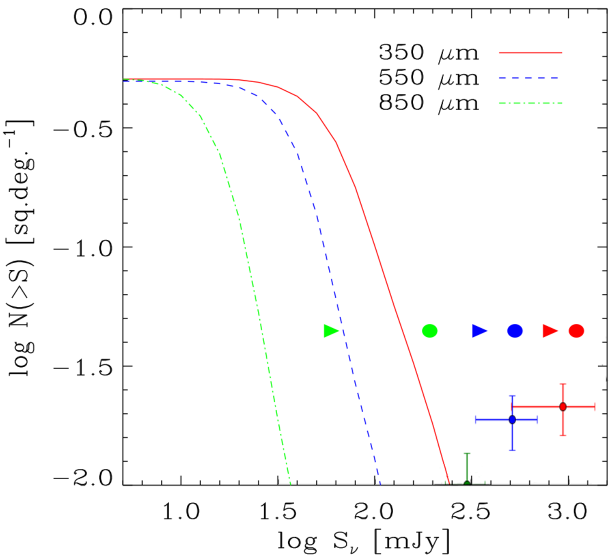

Figure 13. Expected cumulative number counts of clusters reproduced from Granato et al. (2015). The solid, dashed and dot-dashed lines show the predicted number counts of sources at 850, 550 and 350 μm. The circles and right pointing triangles are the results from Clements et al. (2014), and points with errorbars are the results from this work.

the z = 1 flux densities generally being the greatest. Therefore, to be conservative, we compare our results to those at redshift z = 1, under the extreme assumption that all our candidate proto-clusters exist at this redshift. Even in the extreme case that all our candidate proto-clusters lie at z $\simeq$ 1, the observed flux densities appear systematically higher than the simulated flux densities, with a median flux density of proto-cluster candidates of 500 mJy at 350 μm observed compared to 100 mJy simulated.

It is difficult to match the observed proto-cluster flux densities to the simulated. We earlier demonstrated that the flux from these sources comes almost entirely from multiple, detected, discrete sources rather than cirrus or fainter sources. Additionally, we remove any foreground or background contaminent and compare our sources to those simulated clusters with the highest flux densities. Even with these constrains, we still find our proto-clusters are around a factor of 5 $\times$ brighter in comparison to the simulations. If these proto-clusters are confirmed to be real, physical associations, then these results demonstrate that current models of cluster formation struggle to reproduce the FIR/sub-mm flux densities seen in observations by a factor of 5, and likely underestimate the SFR in clusters / proto-clusters during their formation. One possible explination is that these DSFGs are not tracing only the most massive clusters, and that clusters with lower final virial masses could match our observations, but redshift and mass confirmations would be required before this can be tested.

6.4 Evolution of large scale structure

According to the formalism of Negrello et al. (2005), the number counts of clusters should be sensitive to the evolution of the amplitude Q of the three-point correlation func-

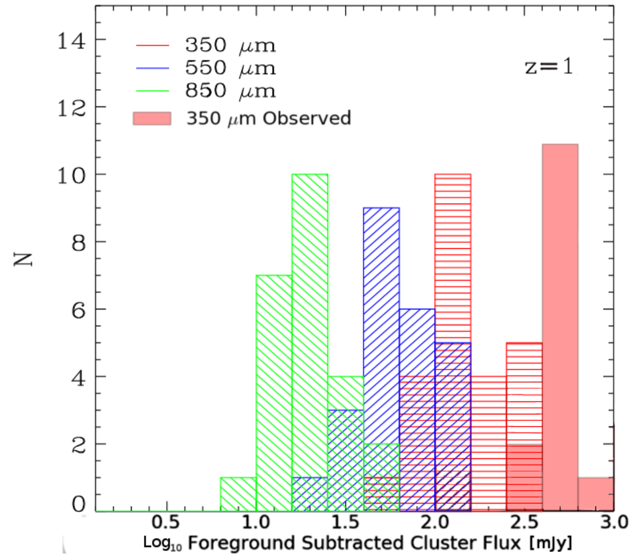

Figure 14. A histogram of the estimated flux densities of clusters at z = 1 reproduced from Granato et al. (2015). The red, green and blue hashed bins represent the histograms from the simulation of clusters as they would appear in the *Planck* HFI bands. The solid red histogram gives the foreground subtracted candidate proto-cluster flux densities from this work if placed at z = 1.

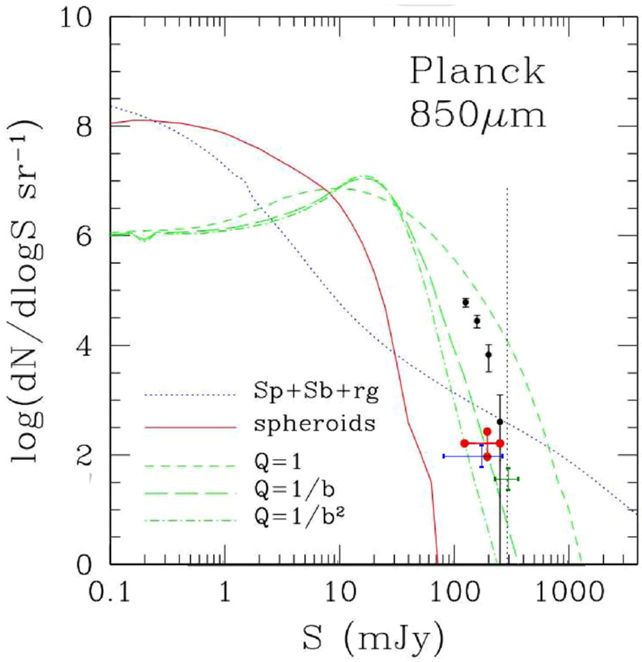

Figure 15. Predicted number counts for over-densities of 850 μm sources taken from Negrello et al. (2005). The dashed green lines gives the predictions for the number counts as the three point correlation function evolves according to Q = 1 (no evolution) Q = $\frac{1}{b}$ and Q = $\frac{1}{b^2}$, where b is the linear bias factor between galaxies and dark matter. The black points are from the simulations from Negrello et al. (2005), the red point gives the results from Clements et al. (2014) and the blue point gives the results from this paper using all our proto-clusters. The green point is our result if we restrict ourselves only to proto-clusters detected at the $> 3\sigma$ level.

tion. Under the tentative assumption that the galaxies in our proto-cluster candidates are in fact all physically associated, we compare the counts of our candidate proto-clusters to the predictions made by Negrello et al. (2005) in Figure 15. We plot both our result when no restrictions are imposed (in blue) and when we impose the constraint that only the 7 proto-cluster candidates detected to at least 3σ in the 353 GHz channel (850 μm) are included (in green). Without constraints, we find a median flux for our proto-clusters of (169 $\pm$ 95) mJy, and a number density of 135 $\pm$ 24 sources per steradian. With constraints, we find a median flux for our proto-clusters of (294 $\pm$ 65) mJy, and a number density of 39.4 $\pm$ 13 sources per steradian.

Both Clements et al. (2014) and this work strongly exclude the Q = 1 analytical model, which corresponds to the amplitude Q of the three-point correlation not evolving with redshift. Using 3σ detected proto-clusters at 353 GHz suggests that the Q $\propto \frac{1}{b^2}$ model is also incorrect. However, caution should be taken as these are only candidate proto-clusters rather than confirmed, and with only 7 proto-clusters detected to 3σ the number counts remain low. Addionally, flux contamination from unassociated sources (which we account for in Section 6.3 but cannot account for here) is likely to play a role. If we assume that roughly 25% of the flux can be from unassociated sources, similar to what was found in Section 6.3, our conclusions remain similar, with a better agreement between the $Q = \frac{1}{b}$ model and our $> 3\sigma$ detected proto-clusters.

7 DISCUSSION

We have identified 27 candidate proto-clusters from a cross-match of *Planck* compact source catalogues and *Herschel* maps. The numbers of sources are difficult to explain if none of them are associated with each other, and their colours indicate they all likely like at z > 1. We have also found several proto-cluster candidates with lower S_{857}/S_{545} flux ratios than expected. We have shown this could be from a large number of foreground contaminants, but it is also possible that there exists a warmer population of clusters / proto-clusters of DSFGs. In this section, we discuss these results in the context of the literature, as well as briefly discussing the natures of the other *Planck* compact sources we have identified.

7.1 The HeLMS field

The HerMES Large Mode Survey (HeLMS, P.I. Marco Viero) is a shallow 280 deg^2 field imaged with *Herschel*-SPIRE at 250, 350 and 500 μm. In comparison to the other extra-galactic fields under consideration here, it is highly contaminated by cirrus. No publicly released, verified catalogue exists for this complex field, but using a private catalogue (Marco Viero, Private Communication) we find 130 857 GHz and 40 545 GHz *Planck* compact sources in this field, and 137 unique sources. The maps are highly cirrus contaminated, with 64 (46%) of sources being identified with Galactic cirrus, 61 (45%) local galaxies, 2 QSOs (LBQS 0106+0119 and CRATES J2323-0316), 1 candidate proto-cluster, and 9 sources with no clear identification. Even in the PCCS2, where we found that only 2.5% of sources were associated with Galactic cirrus, we find 27% of PCCS2 sources in the HeLMS field are associated with Galactic cirrus.

Overall, given the reasonably small differences found between the other H-ATLAS, HerMES and Hers fields, and the large differences found between them and HeLMS, we ascribe the differences in our results to the complex nature of the HeLMS field, and the preliminary nature of the catalogues currently available.

7.2 The nature of the *Planck* compact sources

Fig. 3, 4 and Table 2 all indicate that the *Planck* compact sources resolve into a range of different phenomena. Furthermore, almost half of the *Planck* compact sources are actually extended on the scale of *Herschel*, so filters designed purely for point like sources can miss a range of sources, as shown in Table 3. Given this, the ERCSC, PCCS and PCCS2 should not simply be considered deeper versions of the same catalogue, but catalogues that specifically probe different source types owing to the different filters and extraction methods used in their creation.

The *Planck* compact source catalogues appear to host several stars. Two of the three stars present in our catalogue, Mira (also known as o Ceti), and R Sculptoris, are both Asymptotic Giant Branch (AGB) stars, known to produce large amounts of dust (Mayer et al. 2011, 2014), whilst the third, α PsA o, is known to host a dusty debris disk (Acke et al. 2012).

One of our lens candidates, like our proto-cluster candidates, appear to host an overdensity of *Herschel* sources. These could indicate the presence of a physical cluster or proto-cluster. PLCKERC857 G047.32+82.53 (H-ATLAS J132426.9+284452, Negrello et al. 2016) at a redshift of 1.676 (George et al. 2013; Bussmann et al. 2013; Timmons et al. 2015) has a 3.1σ over-density of 500 μm sources. We also note that PCCS2 857 G270.56+58.54 (H12-00, Herranz et al. 2013; Fu et al. 2012; Clements et al. 2016) hosts 2.8σ over-density of 350 μm sources and Clements et al. (2016) find an overdensity of SCUBA-2 850 μm sources associated with H12-00, and is unclassified in this work, though it is selected as a candidate proto-cluster if we use a slightly smaller beamsize of 4.33 arcminutes. Furthermore, H12-00 is also independently selected in Canameras et al. (2015), where they specifically search for and follow up the brightest gravitational lensed sources discovered with *Planck*. Whether DSFGs are good tracers of the most massive dark matter overdensities at $z > 2$ continues to be widely debated (Blain et al. 2004; Chapman et al. 2009; Dannerbauer et al. 2014; Miller et al. 2015; Casey et al. 2015; Hung et al. 2016), but if they do, then these lensed sources could make excellent signposts for the locations of further clusters / proto-clusters.

H12-00 (Clements et al. 2016) does not qualify as a proto-cluster candidate using our criterion in Section 3. Given that follow up work on H12-00 demonstrates its likely cluster nature (Clements et al. 2016), the large number of "red" unassigned sources, and a number of sources that are on the edge of being selected as candidate proto-clusters, it is entirely possible that *Planck* is detecting a far larger number of protocluster sources, but that the specific quantifiable criteria used here mean that they are not assigned as such during the selection process. An examination of the unassigned sources reveals almost half (28 of 61) have a $> 2\sigma$ in the 250, 350 or 500 μm bands, but are not 3σ overdense. Additionally, three of the four matches we found with the PHZ were unclassified. The final source, PHz G160.57-56.79 / PCCS1 857 G160.59-56.74, we identify as the local galaxy 2MASX J02094125+0015587 at a redshift of $z = 0.2020$. This is somewhat surprising, and could hint that our selection probes a different population of proto-clusters / overdensities of sources. It is possible that several of the "red" unassigned sources could be due to CIB fluctuations which, due to clustering, have a strong super-Gaussian tail so can appear as high S/N sources (See Figure 10 of De Zotti et al. 2015). However, Fig. 6 shows that the flux density of many of the unassigned sources is entirely accounted for by discrete, detected sources. These could still be line of sight chance alignments, but it does show that these are unlikely to be fluctuations in the background sources too faint to be detected by *Herschel* and are, at best, fluctuations in the number counts of bright ($S_{350} > 25.4$ mJy) *Herschel* sources.

7.3 The nature of our proto-cluster candidates

DSFGs have now been found in a range of cluster environments, from extremely large proto-clusters on angular scales > 10 arcminutes (Dannerbauer et al. 2014; Casey et al. 2015; Casey 2016), to scales similar to those of the *Planck* HFI beam (Herranz et al. 2013; Clements et al. 2014; Planck Collaboration et al. 2016, 2015a, This work), to > 10 sources on ~ 20 arcsecond scales (Oteo et al. 2017a). The existence of many physically associated DSFGs is surprising; simulations expect these sources to be physically unassociated (Hayward et al. 2013; Cowley et al. 2014), and without a mechanism for triggering several DSFGs simultaneously or a longer duty cycle (Emonts et al. 2016; Dannerbauer et al. 2017; Oteo et al. 2017b), we would not expect to observe several physically associated DSFGs at once (Casey 2016). Similar to the PHZ, we can be confident but not certain that the compact candidate proto-clusters we have detected are physically associated or multiple clusters. However follow up of other apparent overdensities of DSFGs (Flores-Cacho et al. 2016; Wang et al. 2016; Oteo et al. 2017a), suggests many of these objects are indeed physically associated. We leave redshift estimates and therefore SFR estimates for our proto-clusters to a future paper (Cheng et al. in preparation), but if the DSFG members of these proto-cluster candidates are similar to other DSFGs, their likely SFR will be of the order of 100 $M_\odot$ yr^{-1}, and a likely total cluster SFR of up to several thousands of $M_\odot$ yr^{-1} Dannerbauer et al. (2014); Casey (2016); Oteo et al. (2017a). Both Scoville et al. (2013) and Darvish et al. (2016) find that below z $\sim$ 1, SFR is efficiently quenched in denser environments, but the mechanicsms for this quenching remain uncertain, and Figure 2 of Casey (2016) shows that there is clearly a downturn below z $\sim$ 1 between the theoeretical and observed SFR density of clusters. In this paper, we have found DSFGs, with elevated associated SFRs, in clusters / proto-clusters over a range of scales, from the arcsecond to the arcminute. This indicates that it is unlikely that it is simply the scale or size of the structure that determines its SFR density. Given we do not see DSFGs in local clusters, it could be that is it the virialisation state or the presence of a evolved intracluster medium which determines whether the presence of multiple DSFGs is likely to occur. No clusters containing significant numbers of DSFGs have so far been confirmed to be viralised, though the Spiderweb Galaxy structure may contain DSFGs in a viralised sub-halo (Dannerbauer et al. 2014). If it is the virialisation or presence of an evolved intracluster medium that prevents or quenches DSFGs, it would suggest that none of the proto-cluster candidates detected here, and indeed none of the confirmed clusters containing significant numbers of DSFGs, are yet viralised or posses an evolved intracluster medium. Finding clusters / proto-clusters of DSFGs, and determining particularly the viralisation and environmental state around and within them, may therefore be key to understanding the mechanisms behind the quenching of galaxies in different environments.

However, it should be stressed that the candidate proto-clusters detected in this work remain candidates, and not only is there a need to confirm that the DSFGs detected lie at the same redshift, but further work should be undertaken to confirm that there is also an optical/NIR overdensity at these positions, confirming that this is indeed a cluster rather than associated sources that come from looking down a filament. Additionally, work should be done to characterise these clusters, particularly at what evolutionary stage they are at (i.e. have they virialised?). Both the PHZ and this work provide complementary targets for these systems on scales around a few arcminutes, with the PHZ generally selecting the brightest and reddest, and this work selecting both fainter and warmer candidates. Additional sources of cluster / proto-cluster candidates can come from selection directly on the *Herschel* maps (Valtchanov et al. 2013).

The diversity of DSFGs in clusters is further suggested by the difference we find with the PHZ; we only find four sources in common, none of which we identify as a candidate cluster. When applying their flux and colour cuts directly on our catalogue, we only find 15 objects, 3 (20%) are cluster candidates (including the Bootes cluster candidates identified by Clements et al. (2014)), 3 (20%) are cirrus, 4 (26%) are local galaxies (UGC 09215, UGC 08017, NGC 5056 and CGCG 160-170), 1 (6.6%) is a lens candidate (H-ATLAS J132426.9+284452), and 4 (26%) we were not able to assign a identity. Of these 15, 11 are detected only in the ERCSC.

Similar work has also been undertaken by Baes et al. (2014) in the 84 deg^2 of the *Herschel* Virgo Cluster Survey, where they find that most *Planck* compact sources are dominated by local late type galaxies, with few sources being classified as galactic cirrus, spurious detections, and no sources classified as candidate clusters. This can largely be explained by the small areas examined; As Baes et al. note, they are directly examining a local cluster, whereas our total area surveyed is $\sim 10\times$ larger across most of the accessible extragalactic sky. Given also that they use the PCCS1, which both they and we note is devoid of sources without a bright local counterpart (see Table 3), we therefore conclude that sky variance and use of PCCS1 can explain the apparently low numbers of cluster candidates in the *Herschel* Virgo Cluster Survey.

8 CONCLUSIONS

Through a cross-match of the *Planck* compact source catalogues, and 808.4 deg^2 of *Herschel* fields from H-ATLAS, HerMES and Hers, we have identified 27 proto-clusters of DSFGs that are at least 3σ overdense in either 250, 350 or 500 μm sources. Additionally, we have identified, 192 local galaxies, 43 regions of galactic cirrus, 12 candidate lensed sources, 3 stars and 2 QSOs which also make up the *Planck* compact source catalogues. A further 61 sources we are unable to assign a classification, but many host a large number of *Herschel* sources ($> 2\sigma$ in the 250, 350 or 500 μm bands), and other have colours indicative of a high redshift origin. It is possible that many of these unassigned sources are also proto-clusters of DSFGs, though it is more difficult to rule out fluctuations in the number counts as an explanation.

We find there is significant differences between the three released versions of the catalogues, with the ERCSC hosting a larger fraction of candidate proto-clusters than the PCCS or PCCS. We ascribe this to the filters used in the creation of the three catalogues, with the PCCS and PCCS2 Mexican-hat wavelet filter likely suppressing extended emission from both Galactic cirrus and candidate proto-clusters.

We verified that there is good agreement between *Planck* and *Herschel* aperture photometry for all sources, and further find that with the exception of Galactic cirrus, simply summing up the detected *Herschel* source with S_{350} or $S_{500} > 25.4$ mJy results in a good match between *Planck* and *Herschel* photometry.

The *Planck* colours of our proto-cluster candidates indicate that a selection criteria of $S_{857}/S_{545} < 2$ performs well for selecting out candidate proto-clusters. However, we have also found a number of warmer proto-cluster candidates, which would be missed by such a selection, though we have shown this can be also explained by a significant contamination of low redshift $z < 1$ DSFGs. The *Herschel* colours of our sources indicate they all likely lie at z > 2, and the small scatter of points in the *Herschel* colour-colour plots can indicate a physical cluster / proto-cluster, though the uncertainties are large.

We find a surface density of candidate proto-clusters of $(3.3 \pm 0.7) \times 10^{-2}$ sources deg^{-2}, in good agreement with previous similar studies. Crossmatching our catalogue with the PHZ, we find only four matches, none of which we identify as a candidate proto-cluster.

Finally, we compare our results to simulations, finding both that our proto-clusters are a factor of 5 times brighter at 353 GHz than expected from simulations, even in the most conservative estimates, and that the amplitude of the three-point correlation function Q likely evolves with $Q \propto \frac{1}{b}$.

Without redshift confirmation, there remains the possibility that none of these objects are physical clusters / proto-clusters. However, given the number we have found alongside other groups, if they are clusters / proto-clusters it is a challenge to explain how groups of > 20 associated DSFGs exist, given their expected lifetimes of ~ 100 Myrs. Such proto-clusters of DSFGs are being found from arcminute to arcsecond scales, yet we do not see this in the local Universe, indicating that star formation is quenched in clusters at low redshifts, but does take place in clusters / proto-clusters at higher redshifts, possibly due to a clusters virilisation state. Since we do not know if these sources are virialisaed, further characterisation, particularly of the environment and state of virialisation, should be a key focus for follow up observations. Given also that we expect DSFGs such as these to evolve into the brightest cluster members at the cores of galaxy clusters, they likely play a vital role in the earliest stages of cluster formation and evolution.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their insightful and useful comments. Based on observations obtained with Planck (http://www.esa.int/Planck), an ESA science mission with instruments and contributions directly funded by ESA Member States, NASA, and Canada. 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 website is http://www.h-atlas.org/ U.S. participants in Herschel ATLAS acknowledge support provided by NASA through a contract issued from JPL. This research has made use of data from HerMES project (http://hermes.sussex.ac.uk/). HerMES is a Herschel Key Programme utilising Guaranteed Time from the SPIRE instrument team, ESAC scientists and a mission scientist. The HerMES data was accessed through the Herschel Database in Marseille (HeDaM - http://hedam.lam.fr) operated by CeSAM and hosted by the Laboratoire d'Astrophysique de Marseille. HerMES DR3 was made possible through support of the Herschel Extragalactic Legacy Project, HELP (http://herschel.sussex.ac.uk). This work made extensive use of the Starlink Table/VOTable Processing Software, TOPCAT (Taylor 2005). GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken 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 programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT 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 website is http://www.gama-survey.org/ . This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration, 2013) (The Astropy Collaboration et al. 2013). DLC and JG acknowledge support from STFC, in part through grant numbers ST/N000838/1 and ST/K001051/1. H.D. acknowledges financial support from the Spanish Ministry of Economy and Competitiveness (MINECO) under the 2014 Ramóy Cajal program MINECO RYC-2014-15686. This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. M.J.M. acknowledges the support of the National Science Centre, Poland through the POLONEZ grant 2015/19/P/ST9/04010. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 665778. LD acknowledge support from the European Research Council in the form of Advanced Investigator grant COSMICISM and also the Consolidator Grant CosmicDust.

REFERENCES

Acke B., et al., 2012, Astronomy & Astrophysics, Volume 540, id.A125, 9 pp., 540
Asboth V., et al., 2016, Monthly Notices of the Royal Astronomical Society, Volume 462, Issue 2, p.1989-2000, 462, 1989
Baes M., et al., 2014, Astronomy & Astrophysics, Volume 562, id.A106, 16 pp., 562
Bertin E., Arnouts S., 1996, Astronomy and Astrophysics Supplement Series, 117, 393
Bertincourt B., et al., 2016, Astronomy & Astrophysics, 588, A107
Blain A., 2002, Physics Reports, 369, 111
Blain A. W., Chapman S. C., Smail I., Ivison R., 2004, The Astrophysical Journal, 611, 52
Bussmann R. S., et al., 2013, The Astrophysical Journal, Volume 779, Issue 1, article id. 25, 26 pp. (2013)., 779
Cai Z.-Y., et al., 2013, The Astrophysical Journal, Volume 768, Issue 1, article id. 21, 24 pp. (2013)., 768
Canameras R., et al., 2015, Astronomy & Astrophysics, Volume 581, id.A105, 18 pp., 581
Capak P. L., et al., 2011, Nature, 470, 233
Casey C. M., 2016, The Astrophysical Journal, Volume 824, Issue 1, article id. 36, pp. (2016)., 824
Casey C. M., Narayanan D., Cooray A., 2014, Physics Reports, 541, 45161
Casey C. M., et al., 2015, The Astrophysical Journal, 808, L33
Chapman S. C., Blain A., Ibata R., Ivison R. J., Smail I., Morrison G., 2009, The Astrophysical Journal, 691, 560
Chiang Y.-K., Overzier R., Gebhardt K., 2013, The Astrophysical Journal, Volume 779, Issue 2, article id. 127, 16 pp. (2013)., 779
Clements D. L., et al., 2010, Astronomy and Astrophysics, 518, L8
Clements D. L., et al., 2014, Monthly Notices of the Royal Astronomical Society, 439, 1193
Clements D. L., et al., 2016, Monthly Notices of the Royal Astronomical Society, 461
Cowley W. I., Lacey C. G., Baugh C. M., Cole S., 2014, Monthly Notices of the Royal Astronomical Society, 446, 1784
Daddi E., et al., 2008, The Astrophysical Journal, Volume 694, Issue 2, pp. 1517-1538 (2009)., 694, 1517
Dannerbauer H., et al., 2014, Astronomy & Astrophysics, 570, A55
Dannerbauer H., et al., 2017, eprint arXiv:1701.05250
Darvish B., Mobasher B., Sobral D., Rettura A., Scoville N., Faisst A., Capak P., 2016, The Astrophysical Journal, Volume 825, Issue 2, article id. 113, 16 pp. (2016)., 825
Donley J. L., Rieke G. H., Perez-Gonzalez P. G., Rigby J. R., Alonso-Herrero A., 2007, The Astrophysical Journal, Volume 660, Issue 1, pp. 167-190., 660, 167
Dowell C. D., et al., 2014, The Astrophysical Journal, 780, 75
Driver S. P., et al., 2011, Monthly Notices of the Royal Astronomical Society, 413, 971
Eales S., et al., 2010, Publications of the Astronomical Society of the Pacific, 122, 499
Eisenhardt P. R. M., et al., 2008, The Astrophysical Journal, 684, 905
Emonts B. H. C., et al., 2016, Science, Volume 354, Issue 6316, pp. 1128-1130 (2016)., 354, 1128

Falgarone E., et al., 2017, Nature, Volume 548, Issue 7668, pp. 430-433 (2017)., 548, 430
Flores-Cacho I., et al., 2016, Astronomy & Astrophysics, 585, A54
Fu H., et al., 2012, The Astrophysical Journal, 753, 134
George R. D., et al., 2013, Monthly Notices of the Royal Astronomical Society: Letters, Volume 436, Issue 1, p.L99-L103, 436, L99
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, The Astrophysical Journal, 600, 580
Granato G. L., Ragone-Figueroa C., Dominguez-Tenreiro R., Obreja A., Borgani S., De Lucia G., Murante G., 2015, Monthly Notices of the Royal Astronomical Society, 450, 1320
Griffin M. J., et al., 2010, Astronomy and Astrophysics, 518, L3
Gunn J. E., Hoessel J. G., Oke J. B., 1986, The Astrophysical Journal, 306, 30
Hao J., et al., 2010, The Astrophysical Journal Supplement, Volume 191, Issue 2, pp. 254-274 (2010)., 191, 254
Harrison I., Coles P., 2011, Monthly Notices of the Royal Astronomical Society, 421, L19
Hayward C. C., Behroozi P. S., Somerville R. S., Primack J. R., Moreno J., Wechsler R. H., 2013, Monthly Notices of the Royal Astronomical Society, Volume 434, Issue 3, p.2572-2581, 434, 2572
Herranz D., et al., 2013, Astronomy & Astrophysics, 549
Hickox R. C., et al., 2012, Monthly Notices of the Royal Astronomical Society, 421
Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, The Astrophysical Journal Supplement Series, 175, 356
Hung C.-L., et al., 2016, The Astrophysical Journal, Volume 826, Issue 2, article id. 130, 10 pp. (2016)., 826
Ivison R. J., et al., 2013, The Astrophysical Journal, Volume 772, Issue 2, article id. 137, 15 pp. (2013)., 772
Ivison R. J., et al., 2016, APJ, 832
Kravtsov A., Borgani S., 2012, Annual Review of Astronomy and Astrophysics, 50, 353
Lapi A., et al., 2011, The Astrophysical Journal, Volume 742, Issue 1, article id. 24, 21 pp. (2011)., 742
Levenson L., et al., 2010, Monthly Notices of the Royal Astronomical Society, 409, 83
Lopes P. A. A., de Carvalho R. R., Gal R. R., Djorgovski S. G., Odewahn S. C., Mahabal A. A., Brunner R. J., 2004, The Astronomical Journal, 128, 1017
López-Caniego M., Herranz D., González-Nuevo J., Sanz J. L., Barreiro R. B., Vielva P., Argüeso F., Toffolatti L., 2006, Monthly Notices of the Royal Astronomical Society, 370, 2047
Ma C.-J., et al., 2015, The Astrophysical Journal, 806, 257
MacKenzie T. P., et al., 2017, Monthly Notices of the Royal Astronomical Society, Volume 468, Issue 4, p.4006-4017, 468, 4006
Mayer A., et al., 2011, Astronomy & Astrophysics, Volume 531, id.L4, 4 pp., 531
Mayer A., et al., 2014, Asymmetrical Planetary Nebulae VI conference, Proceedings of the conference held 4-8 November, 2013. Edited by C. Morisset, G. Delgado-Inglada and S. Torres-Peimbert. Online at http://www.astroscu.unam.mx/apn6/PROCEEDINGS/, id.59
Michałowski M., Hjorth J., Watson D., 2010, Astronomy and Astrophysics, 514, A67
Miller T. B., Hayward C. C., Chapman S. C., Behroozi P. S., 2015, Monthly Notices of the Royal Astronomical Society, 452, 878
Miville Deschenes M. A., Lagache G., 2005, The Astrophysical Journal Supplement Series, 157, 302:323
Negrello M., Gonzalez-Nuevo J., Magliocchetti M., Moscardini L., De Zotti G., Toffolatti L., Danese L., 2005, Monthly Notices of the Royal Astronomical Society, 358, 869
Negrello M., et al., 2016, Monthly Notices of the Royal Astronomical Society, Volume 465, Issue 3, p.3558-3580, 465, 3558
Nguyen H. T., et al., 2010, Astronomy and Astrophysics, Volume 518, id.L5, 5 pp., 518
Noble A. G., et al., 2013, Monthly Notices of the Royal Astronomical Society: Letters, 436, L40
Oliver S. J., et al., 2010, Astronomy and Astrophysics, 518, L21
Oliver S. J., et al., 2012, Monthly Notices of the Royal Astronomical Society, 424, 1614
Oteo I., Zwaan M. A., Ivison R. J., Smail I., Biggs A. D., 2016, The Astrophysical Journal, Volume 837, Issue 2, article id. 182, 9 pp. (2017)., 837
Oteo I., et al., 2017a, eprint arXiv:1709.02809
Oteo I., et al., 2017b, eprint arXiv:1701.05901
Ott S., Centre H. S., Agency E. S., 2010, Astronomical Data Analysis Software and Systems XIX, 434, 139
Pearson E. A., et al., 2013, Monthly Notices of the Royal Astronomical Society, Volume 435, Issue 4, p.2753-2763, 435, 2753
Petty S. M., et al., 2013, The Astronomical Journal, Volume 146, Issue 4, article id. 77, 17 pp. (2013)., 146
Pilbratt G. L., et al., 2010, Astronomy and Astrophysics, 518, L1
Planck Collaboration P., et al., 2011a, Astronomy & Astrophysics, Volume 536, id.A1, 16 pp., 536
Planck Collaboration P., et al., 2011b, Astronomy & Astrophysics, 536, A7
Planck Collaboration P., et al., 2014, Astronomy & Astrophysics, 571, A28
Planck Collaboration P., et al., 2015a, Astronomy & Astrophysics, Volume 582, id.A30, 29 pp., 582
Planck Collaboration P., et al., 2015b, Astronomy & Astrophysics, Volume 594, id.A26, 39 pp., 594
Planck Collaboration P., et al., 2015c, Astronomy & Astrophysics, Volume 594, id.A27, 38 pp., 594
Planck Collaboration P., et al., 2015d, Astronomy & Astrophysics, Volume 594, id.A28, 28 pp., 594
Planck Collaboration P., et al., 2016, Astronomy & Astrophysics, Volume 596, id.A100, 28 pp., 596, A100
Planck HFI Team P., et al., 2010, Astronomy and Astrophysics, 520, A9
Rowan-Robinson M., et al., 2016, Monthly Notices of the Royal Astronomical Society, Volume 461, Issue 1, p.1100-1111, 461, 1100
Scoville N., et al., 2013, The Astrophysical Journal Supplement, Volume 206, Issue 1, article id. 3, 26 pp. (2013)., 206
Simpson J. M., et al., 2014, The Astrophysical Journal, 788, 125
Stevens J. A., Jarvis M. J., Coppin K. E. K., Page M. J., Greve T. R., Carrera F. J., Ivison R. J., 2010, Monthly Notices of the Royal Astronomical Society, 405, no
Swinbank A. M., Chapman S. C., Smail I., Lindner C., Borys C., Blain A. W., Ivison R. J., Lewis G. F., 2006, Monthly Notices of the Royal Astronomical Society, 371, 465
Tacconi L. J., et al., 2008, The Astrophysical Journal, 680, 246
Tauber J. A., et al., 2010, Astronomy and Astrophysics, 520, A1
Taylor M. B., 2005, Astronomical Data Analysis Software and Systems XIV - ASP Conference Series, 347, 29
The Astropy Collaboration A., et al., 2013, Astronomy & Astrophysics, Volume 558, id.A33, 9 pp., 558
Timmons N., et al., 2015, The Astrophysical Journal, Volume 805, Issue 2, article id. 140, 6 pp. (2015)., 805
Toft S., et al., 2014, The Astrophysical Journal, Volume 782, Issue 2, article id. 68, 12 pp. (2014)., 782
Valiante E., et al., 2016, Monthly Notices of the Royal Astronomical Society, Volume 462, Issue 3, p.3146-3179, 462, 3146
Valtchanov I., et al., 2013, Monthly Notices of the Royal Astronomical Society, Volume 436, Issue 3, p.2505-2514, 436, 2505
Viero M. P., et al., 2014, The Astrophysical Journal Supplement Series, 210, 22
Walter F., et al., 2012, Nature, 486, 233
Wang L., et al., 2013, Monthly Notices of the Royal Astronomical Society, Volume 444, Issue 3, p.2870-2883, 444, 2870

Wang T., et al., 2016, The Astrophysical Journal, Volume 828, Issue 1, article id. 56, 15 pp. (2016)., 828
Wenger M., et al., 2000, Astronomy and Astrophysics Supplement, v.143, p.9-22, 143, 9
Wilkinson A., et al., 2016, Monthly Notices of the Royal Astronomical Society, Volume 464, Issue 2, p.1380-1392, 464, 1380
Yuan T. T., et al., 2014, The Astrophysical Journal Letters, Volume 795, Issue 1, article id. L20, 6 pp. (2014)., 795

APPENDIX A: IMAGES OF THE *PLANCK* COMPACT SOURCES

To comply with arXivs size limits, only 1 of the 4 grids of pictures is included here. The full published article will include all of the images.

APPENDIX B: TABLE OF THE*PLANCK* COMPACT SOURCES

Here we include a list of all our *Planck* compact sources that lie on the maps

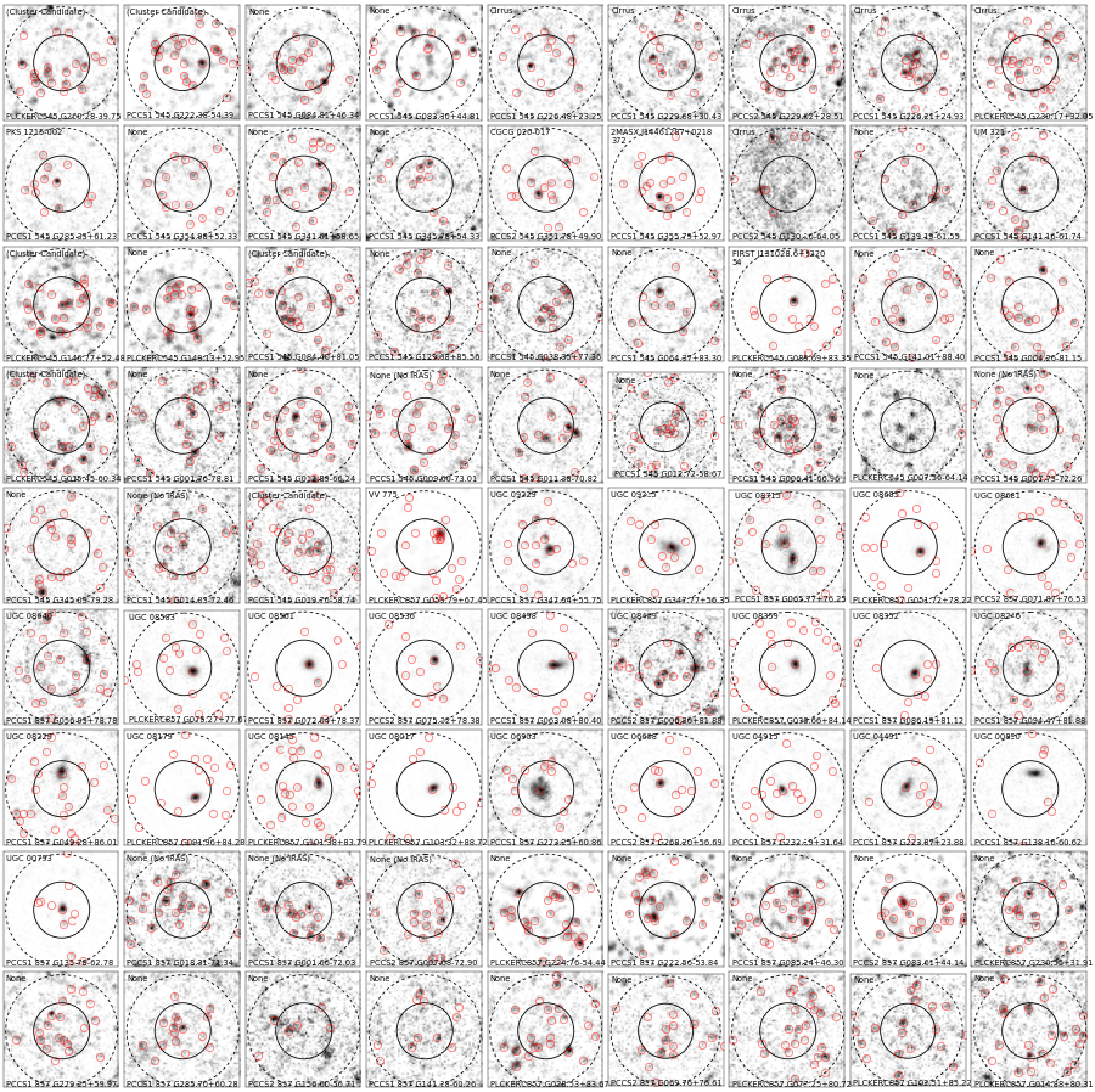

Figure A1. The 350 μm Herschel map for all of our sources, with the Planck beam in solid black circle, the aperture photometry in dashed black circle, and the red circles indicate the positions of sources which have a flux density > 25.4 mJy in either the 250, 350 or 500 μm bands.

Table B1. Candidate protoclusters from the *Planck* 857GHz catalogues of compact sources. The σ values provide the strength of the overdensity at 250, 350 and 500 μm. A table containing the properties and identifications of all the *Planck* compact sources is avaliable online.

Name	RA	DEC	Associations	Planck 857 Flux [mJy]	σ_{250}	σ_{350}	σ_{500}
PCCS1 857 G014.92-58.26	336.635	-32.177	Cluster Candidate	302 ± 239	1.88	1.96	4.36
PCCS1 857 G354.81-79.56	3.049	-33.228	Cluster Candidate	1014 ± 233	1.67	2.23	3.25
PLCKERC857 G257.09-87.10	15.233	-29.122	Cluster Candidate	2403 ± 198	3.52	3.02	3.25
PLCKERC857 G014.99-59.64	338.260	-32.139	Cluster Candidate	943 ± 115	3.12	1.96	4.36
PLCKERC857 G239.13-78.19	25.333	-31.786	Cluster Candidate	1110 ± 145	2.72	1.68	3.25
PLCKERC857 G017.86-68.67	348.790	-30.591	Cluster Candidate	1657 ± 167	2.92	3.76	5.05
PLCKERC857 G007.34-65.24	345.366	-35.103	Cluster Candidate	920 ± 132	3.32	3.52	3.25
PCCS1 857 G252.98-85.59	16.749	-29.910	Cluster Candidate	619 ± 817	1.88	2.76	3.25
PCCS1 857 G058.69+81.03	202.258	30.712	Cluster Candidate	845 ± 298	1.45	1.68	4.01
PCCS1 857 G058.53+82.57	200.607	30.124	Cluster Candidate	827 ± 197	1.45	2.23	4.36
PLCKERC857 G062.48+78.89	204.276	32.142	Cluster Candidate	814 ± 109	0.77	2.50	3.25
PLCKERC857 G063.13+78.00	205.172	32.621	Cluster Candidate	912 ± 169	2.10	3.52	3.25
PLCKERC857 G027.36+84.83	198.608	26.510	Cluster Candidate	792 ± 93	2.52	3.27	4.01
PLCKERC857 G042.54+81.51	202.358	28.224	Cluster Candidate	1697 ± 251	1.67	3.76	4.01
PLCKERC857 G149.81+50.11	158.364	59.196	Cluster Candidate	1249 ± 131	5.38	4.47	4.71
PCCS1 857 G089.66+36.10	256.828	60.449	Cluster Candidate	886 ± 194	3.32	2.76	2.01
PCCS1 857 G085.48+43.36	244.657	55.771	Cluster Candidate	748 ± 128	2.72	3.27	3.64
PLCKERC857 G095.44+58.94	216.128	52.936	Cluster Candidate	1141 ± 176	4.65	5.80	2.44
PCCS1 857 G238.20+42.27	150.845	1.469	Cluster Candidate	712 ± 269	4.28	3.76	2.01
PLCKERC857 G060.37+66.55	218.579	35.559	Cluster Candidate	1241 ± 151	1.00	4.47	5.70
PLCKERC857 G261.89-40.69	70.168	-53.734	Cluster Candidate	795 ± 134	3.32	4.47	3.64

Table B2. Candidate protoclusters from the *Planck* 545GHz catalogues of compact sources. A table containing the properties and identifications of all the *Planck* compact sources is avaliable online.

Name	RA	DEC	Associations	S_{545} [mJy]	σ_{250}	σ_{350}	σ_{500}
PCCS1 545 G019.76-58.74	337.311	-29.670	Cluster Candidate	225±164	2.92	1.68	3.25
PCCS1 545 G354.79-79.57	3.060	-33.226	Cluster Candidate	335±158	1.22	1.39	3.25
PLCKERC545 G015.45-60.34	339.084	-31.902	Cluster Candidate	607±87	2.72	3.27	4.01
PCCS1 545 G058.72+82.59	200.566	30.136	Cluster Candidate	349±161	1.22	1.39	3.64
PCCS1 545 G027.38+84.85	198.581	26.515	Cluster Candidate	433±136	1.45	2.23	4.01
PCCS1 545 G084.40+81.05	199.569	33.968	Cluster Candidate	387±230	1.88	1.39	3.25
PLCKERC545 G146.77+52.48	164.280	59.031	Cluster Candidate	606±73	4.09	4.93	1.06
PCCS1 545 G222.38-54.39	53.046	-27.117	Cluster Candidate	182±185	2.10	3.02	2.01
PLCKERC545 G060.36+66.56	218.577	35.554	Cluster Candidate	872±114	0.77	4.47	5.70
PLCKERC545 G260.28-39.75	71.996	-52.643	Cluster Candidate	471±46	3.71	3.27	3.25

APPENDIX C: PROBABILITY OF N CLUSTER GALAXIES IN M DETECTED SOURCES

For a given proto-cluster candidate, we observe M sources. Under the assumption that some of these are physically associated with a proto-cluster, whilst some are not, M is a combination of both the number of field and proto-cluster galaxies,

$$M = N_{field} + N_{cluster}.$$

Furthermore, we assume that the field galaxies are distributed in a Poisson manner, and can be described by Poisson statistics. For a Poisson process with a mean and variance of μ, the probability of observing M sources is given by:

$$[\frac{\mu^M}{M!}]exp(-\mu)$$

If we assume that N of our M sources are associated with the proto-cluster, then $M-N$ sources will be associated with the field, and the probability of observing these $M-N$ field galaxies is:

$$[\frac{\mu^{(M-N)}}{(M-N)!}]exp(-\mu)$$

However, this needs to be renormalised as the maximum possible observed field galaxies is now M rather than ∞ (It is impossible to observe $M+1$ field galaxies out of M total galaxies). This can be done using the Poisson cumulative distribution function, given by:

$$\Sigma_{i=0}^{i=M}[\frac{\mu^i}{i!}]exp(-\mu)$$

where we simply sum over all of the possible arrangements of $N_{field} + N_{cluster}$ to give a total of M sources. This is now the correct normalisation factor, as it allows for the full range of possibilities stretching from no sources are associated with the proto-cluster, to all the sources are associated with the proto-cluster. The full equation becomes:

$$p(N|M,\mu) = \frac{[\frac{\mu^{(M-N)}}{(M-N)!}]exp(-\mu)}{\Sigma_{i=0}^{i=M}[\frac{\mu^{(i)}}{(i)!}]exp(-\mu)}.$$

Which can be further simplified to:

$$p(N|M,\mu) = \frac{[\frac{\mu^{(M-N)}}{(M-N)!}]}{\Sigma_{i=0}^{i=M}[\frac{\mu^{(i)}}{(i)!}]}.$$

This paper has been typeset from a TEX/LATEX file prepared by the author.