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

J. Greenslade1*, D. L. Clements1, T. Cheng1, G. De Zotti2,3, D. Scott4, E. Valiante5, S. Eales5, M. N. Bremer6, H. Dannerbauer7,8, M. Birkinshaw6,9, D. Farrah10, D. L. Harrison11, M. J. Michalowski12, I. Valtchanov13, I. Oteo14, M. Baes15, A. Cooray16, M. Negrello7, L. Wang17,18, P. van der Werf19, L. Dunne7,14, S. Dye20

1Astrophysics Group, Imperial College, Blackett Laboratory, Prince Consort Road, London, UK
2SISSA, Via Bonomea 265, 34136, Trieste, Italy
3INAF-Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, I-35122 Padova, Italy
4Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T1Z1, Canada
5School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK
6HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK
7Instituto de Astrofísica de Canarias (IAC), E-38205 La Laguna, Tenerife, Spain
8Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
9Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
10Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA
11Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
12Astronomical Observatory Institute, Faculty of Physics, Adam Mickiewicz University, ul. Słoneczna 36, 60-286 Poznań, Poland
13Herschel Science Centre, European Space Astronomy Centre, ESA, E-28691 Villanueva de la Cañada, Spain
14Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ
15Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium
16Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
17SRON Netherlands Institute for Space Research, Landelijk 12, 9747 AD, Groningen, The Netherlands
18Rapport Astronomical Institute, University of Groningen, Postbus 800, 9700 AV, Groningen, The Netherlands
19Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands
20School 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 deg² 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 constraints on cosmological theories and parameters (Harrison & Cole 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; Michalowski 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 observations showing further sub-mm bright members (Ivison et al. 2013, 2016; Oteo et al. 2017a), implying that either a large scale triggering event ($> 10$ Mpc) “activates” the DSFGs simultaneously, or alternatively, that the duration of the starburst event is longer (0.5–0.7 Gyrs) (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 (Einonts 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. 2011), 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 unresolvable 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 (ERSCC, 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 beam sizes vary both between channel and between catalogues, but are generally around 4 to 5 arcmin 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 (HERMIES, 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 $\sigma$ 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 red-shift clusters using the Planck data (Herranz et al. 2013; Clements et al. 2014; Baes et al. 2014; Planck Collaboration et al. 2011b).

---

$^1$ For the highest frequency channels only.

we discuss Clements et al. (2017) on (Herranz et al. and Valiante et al. 2016, Maddox et al. in 3). We verify the photometry at Galactic latitudes Clements et al. 1). In this section, we provide a brief overview of the construction of the three 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 $H_0 = 67.7 \text{km s}^{-1} \text{Mpc}^{-1}$, $\Omega_m = 0.3$ and $\Omega_k = 0.7$.

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 shown in Fig. 1.

2.1 The Planck compact catalogues of sources

The ERCSC used $SE\text{Xtractor}$ (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 (350$\mu$m) and 545 GHz (550$\mu$m) channels (Planck HFI Team et al. 2010), since the peak of dust emission in galaxies (around 100$\mu$m) 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 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 $|b| > 30^\circ$ is 655 mJy, with the flux density of the faintest $10\sigma$ source at $|b| > 30^\circ$ 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 1$\sigma$ 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 $\mu$m 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 1$\sigma$ total noises of 6.2 - 6.8, 7.1 - 7.5 and 8.2 - 8.9 mJy for the 250, 350 and 500 $\mu$m bands respectively, with the exception of FLS, ADIFS, ELAIS-N1, ELAIS-S1, BOTSES and XMM-SS, which have 1$\sigma$ 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 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).
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.

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 ≪ 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.

### 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.

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

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
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 $\mu$m 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 $\mu$m 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 $\sim750$ 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 $\mu$m (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 $S_{250} > 25.4$ mJy ($\geq 3\sigma$) that lie within 4.63 arcminutes of a Planck 736 GHz source, and 693 Herschel sources with $S_{250} > 25.4$ mJy that lie within 4.63 arcminutes of a Planck 545 GHz source.

Finally, we cross-matched our catalogue of uniform 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 $\mu$m 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 $S_{250} > 50$ 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 $L_{IR} > 10^{13}L_\odot$, 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 $\mu$m, which will be present for Galactic cirrus but not for proto-clusters of DSFGs at redshifts $\geq 1$. Planck objects with structures in the 100 $\mu$m 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 $\mu$m 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
250, 350 and 500 µm sources are 16.5 ± 4.1, 9.1 ± 3.0 and 2.7 ± 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 \(S_{350} > 25.4 \text{ mJy}\). 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 \(S_{350} > 25.4 \text{ mJy}\), 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 \(\sigma\) value in the normal distribution, this corresponds to a 2.94\(\sigma\) 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 over-dense 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.

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

\(^4\) Using Poisson statistics

3.3 Properties of our catalogues

3.3.1 Nature of the unassigned sources

We cannot reliably assign a category 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 ± 0.5\(\sigma\), near the detection threshold of 5\(\sigma\) (for our proto-clusters, this is similar at 5.4 ± 0.3\(\sigma\)). However, ten of the 65 are detected in multiple catalogues (six of these were detected in both the ERCSC and either the PCSS or PCSS2, 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 \(S_{50}\) flux densities, \(\sigma_{50}\). A large value of \(\sigma_{50}\) 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 \(S_{50}\) flux densities, \(\sigma_{50}\). 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 \(\sigma_{50}\) 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 \(\sigma_{50}\). The distribution appears bi-modal, with two distinct regions below and above Log10(\(\sigma\)) ≈ 1.65. This bi-modality is not seen when examining 1,000 random positions. We designate these two regions as “diffuse” (Log10(\(\sigma\)) < 1.65) and “dominated” (Log10(\(\sigma\)) > 1.65), indicating that flux from these sources appears to be
Planck selected Herschel proto-clusters

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.

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

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

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 ~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 ∼ 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 $\times$ FWHM and outer radius of 2 $\times$ 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

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

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 available. 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.

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 a 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...
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 1.8 ± 0.5, and our QSO has 857/545 GHz colour of 0.9 ± 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 $<1$, 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

$$S_\nu \propto \nu^\beta B_\nu(T),$$

where $\nu^\beta$ 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 $\beta$ 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 100 proto-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; Ashboth et al. 2014).
Planck selected Herschel proto-clusters

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 candidate 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 $\mu$m 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 $\beta$ 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-
alone is enough to make the true redshift of a
and and
5
42x382
1
42x-548
1
42x371
Image 42x562 to 281x741
42x-26
comes from a number of discrete, individual sources, and
starburst galaxy, Arp 220 (Donley et al. 2007)
photometry of these objects indicates that the flux density
for the local galaxies. At the same time, we take a template
of the proto-cluster candidates are on average redder than
9
6 THE CANDIDATE CLUSTERS
Out of the 279 unique Planck sources we have identified, 27
appear to be >3σ 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 ~ 2. These ob-
servations could correspond to a physical cluster of DSFGs,
a series of line of sight sources stretching from z ~ 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 sur-
veyed 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 statis-
tics. In Fig. 11, we sample the NGP field with 1,000 ran-
don 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 his-
tograms 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 re-
spect 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.

Given we here examine roughly 800 deg² of sky, and ac-

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 can-
didates 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 ~ 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 Fig-
ures 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.

5 These probabilities have been converted to their corresponding
σ value in the Normal distribution.
Planck selected Herschel proto-clusters

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 $\beta$ 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 $\beta$ are indicated for clarity.

Figure 9.

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.

Figure 10.

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 associated proto-clusters or multiple clusters / proto-clusters along the line as sight, as opposed to chance overdensities along the line of sight. We find a surface density of candidate proto-clusters of $(3.3 \pm 0.7) \times 10^{-2}$ sources deg$^{-2}$. In their assessment of the number of Planck detectable clusters, Clements et al. (2014) find a surface density of $(4.4 \pm 2.2) \times 10^{-2}$ 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 initial 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 $(0.18 \pm 0.01)$ 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, with a surface density rising to $(0.11 \pm 0.02)$ 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 HERMES (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 $S_{250}$, $S_{350}$, 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.

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:

$$\delta_{\text{gal}}(x) = \frac{n_{\text{gal}}(x) - \langle n_{\text{gal}} \rangle}{\langle n_{\text{gal}} \rangle}$$

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 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.

Chiang et al. (2013)
galaxy clusters; PCCSS1 545 G058.72+82.59 (PCCSS1 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. PCCSS1 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 PCCSS1 545 G084.40+81.05 is associated with the estimated redshift 0.43 galaxy cluster NSCJ131812+335831 (Lopes et al. 2004). Given our earlier estimates on the redshift of our sources 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’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 compare 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 $3\sigma$ 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 $\mu$m flux densities from Granato et al. (2015), and comparing the distribution of the 350 $\mu$m 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 the $z = 1$ flux densities generally being the largest. 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 $z \approx 1$, 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 $\mu$m 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 explanation 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 $\mu$m 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) $Q = 1/b$ and $Q = 1/b^2$, 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.

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 lie 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 $\text{deg}^2$ field imaged with Herschel-SPIRE at 250, 350 and 500 $\mu$m. 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 $857\,\text{GHz}$ and 40 $545\,\text{GHz}$ 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 catalogue 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+58.51 (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 > 2σ in the 250, 350 or 500 μm bands, and 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 = 0.20. 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 (S150 > 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⊙ yr⁻¹, and a likely total cluster SFR of up to several thousands of M⊙ 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 ~ 1, SFR is efficiently quenched in denser environments, but the mechanisms for this quenching remain uncertain, and Figure 2 of Casey (2016) shows that there is clearly a downturn below z ~ 1 between the theoretical 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 virial-
sation state or the presence of a evolved intrachannel 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 viri-

alisation or presence of an evolved intrachannel 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 intrachannel medium. Finding clusters / proto-clusters of DSFGs, and determining particularly the virialisation 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 over-
density 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 Sur-

vey, 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 ~ 10× 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 con-

clude 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 cat-

alogues, 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 (> 2σ 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 PCCS2. 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 $S_{250}$ 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 $S_{353}/S_{253} < 2$ 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 contam-

ination 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 $Q \propto 1$.

Without redshift confirmation, there remains the pos-

sibility 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 DS-

FGs exist, given their expected lifetimes of $\sim 100$ 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 virialised, fur-
ther 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 CECAM and hosted by the Laboratoire d’Astrophysique de Marseille. HerMES DR3 was made possible through support of the Herschel Extragalactic Legacy Project, HELP (http://herches.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 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 POŁONEZ 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 acknowledges support from the European Research Council in the form of Advanced Investigator grant COSMICISM and also the Consolidator Grant CosmicDust.

REFERENCES

BertinCourt B., et al., 2016, Astronomy & Astrophysics, 588, A107
Blain A., 2002, Physics Reports, 369, 111
Casey C. M., Narayanan D., Cooray A., 2014, Physics Reports, 541, 45161
Clements D. L., et al., 2010, Astronomy and Astrophysics, 518, L8
Eales S., et al., 2010, Publications of the Astronomical Society of the Pacific, 122, 499

MNras 000, 1–25 (2017)
APPENDIX A: IMAGES OF THE PLANCK COMPACT SOURCES

To comply with arXiv 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 $\mu$m 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 $\mu$m bands.
Table B1. Candidate protoclusters from the *Planck* 857GHz catalogues of compact sources. The $\sigma$ values provide the strength of the overdensity at 250, 350 and 500 $\mu$m. A table containing the properties and identifications of all the *Planck* compact sources is available online.

<table>
<thead>
<tr>
<th>Name</th>
<th>RA</th>
<th>DEC</th>
<th>Associations</th>
<th>Planck 857 Flux [mJy]</th>
<th>$\sigma_{250}$</th>
<th>$\sigma_{350}$</th>
<th>$\sigma_{500}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCCS1 857 G014.92-58.26</td>
<td>330.635</td>
<td>-32.177</td>
<td>Cluster Candidate</td>
<td>302 ± 239</td>
<td>1.88</td>
<td>1.96</td>
<td>4.36</td>
</tr>
<tr>
<td>PCCS1 857 G354.81-79.56</td>
<td>3.049</td>
<td>-33.228</td>
<td>Cluster Candidate</td>
<td>1014 ± 233</td>
<td>1.67</td>
<td>2.23</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC857 G257.09-87.10</td>
<td>15.233</td>
<td>-29.122</td>
<td>Cluster Candidate</td>
<td>2403 ± 198</td>
<td>3.52</td>
<td>3.02</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC857 G014.99-59.64</td>
<td>338.260</td>
<td>-32.139</td>
<td>Cluster Candidate</td>
<td>945 ± 115</td>
<td>3.12</td>
<td>1.96</td>
<td>4.36</td>
</tr>
<tr>
<td>PLCKERC857 G239.13-78.19</td>
<td>25.333</td>
<td>-31.786</td>
<td>Cluster Candidate</td>
<td>1110 ± 145</td>
<td>2.72</td>
<td>1.68</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC857 G017.86-68.67</td>
<td>348.790</td>
<td>-30.591</td>
<td>Cluster Candidate</td>
<td>1657 ± 167</td>
<td>2.92</td>
<td>3.76</td>
<td>5.05</td>
</tr>
<tr>
<td>PLCKERC857 G007.34-65.24</td>
<td>345.366</td>
<td>-35.103</td>
<td>Cluster Candidate</td>
<td>920 ± 132</td>
<td>3.32</td>
<td>3.52</td>
<td>3.25</td>
</tr>
<tr>
<td>PCCS1 857 G252.98-85.59</td>
<td>16.749</td>
<td>-29.910</td>
<td>Cluster Candidate</td>
<td>619 ± 817</td>
<td>1.88</td>
<td>2.76</td>
<td>3.25</td>
</tr>
<tr>
<td>PCCS1 857 G058.69+81.03</td>
<td>202.258</td>
<td>30.712</td>
<td>Cluster Candidate</td>
<td>845 ± 298</td>
<td>1.45</td>
<td>1.68</td>
<td>4.01</td>
</tr>
<tr>
<td>PCCS1 857 G058.53+82.57</td>
<td>200.607</td>
<td>30.124</td>
<td>Cluster Candidate</td>
<td>827 ± 197</td>
<td>1.45</td>
<td>2.23</td>
<td>4.36</td>
</tr>
<tr>
<td>PLCKERC857 G062.48+78.89</td>
<td>204.276</td>
<td>32.142</td>
<td>Cluster Candidate</td>
<td>814 ± 109</td>
<td>0.77</td>
<td>2.50</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC857 G063.13+78.00</td>
<td>205.172</td>
<td>32.621</td>
<td>Cluster Candidate</td>
<td>912 ± 169</td>
<td>2.10</td>
<td>3.52</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC857 G027.36+84.83</td>
<td>198.608</td>
<td>26.510</td>
<td>Cluster Candidate</td>
<td>792 ± 93</td>
<td>2.52</td>
<td>3.27</td>
<td>4.01</td>
</tr>
<tr>
<td>PLCKERC857 G042.54+81.51</td>
<td>202.358</td>
<td>28.224</td>
<td>Cluster Candidate</td>
<td>1697 ± 251</td>
<td>1.67</td>
<td>3.76</td>
<td>4.01</td>
</tr>
<tr>
<td>PLCKERC857 G149.81+50.11</td>
<td>158.364</td>
<td>59.196</td>
<td>Cluster Candidate</td>
<td>1249 ± 131</td>
<td>5.38</td>
<td>4.47</td>
<td>4.71</td>
</tr>
<tr>
<td>PCCS1 857 G089.66+36.10</td>
<td>256.828</td>
<td>60.449</td>
<td>Cluster Candidate</td>
<td>886 ± 194</td>
<td>3.32</td>
<td>2.76</td>
<td>2.01</td>
</tr>
<tr>
<td>PCCS1 857 G085.48+43.36</td>
<td>244.657</td>
<td>55.771</td>
<td>Cluster Candidate</td>
<td>748 ± 128</td>
<td>2.72</td>
<td>3.27</td>
<td>3.64</td>
</tr>
<tr>
<td>PLCKERC857 G095.44+58.94</td>
<td>216.128</td>
<td>52.936</td>
<td>Cluster Candidate</td>
<td>1141 ± 176</td>
<td>4.65</td>
<td>5.80</td>
<td>2.44</td>
</tr>
<tr>
<td>PCCS1 857 G238.20+42.27</td>
<td>150.845</td>
<td>1.469</td>
<td>Cluster Candidate</td>
<td>712 ± 269</td>
<td>4.28</td>
<td>3.76</td>
<td>2.01</td>
</tr>
<tr>
<td>PLCKERC857 G060.37+66.55</td>
<td>218.579</td>
<td>35.559</td>
<td>Cluster Candidate</td>
<td>1241 ± 151</td>
<td>1.00</td>
<td>4.47</td>
<td>5.70</td>
</tr>
<tr>
<td>PLCKERC857 G261.89-40.69</td>
<td>70.168</td>
<td>-53.734</td>
<td>Cluster Candidate</td>
<td>795 ± 134</td>
<td>3.32</td>
<td>4.47</td>
<td>3.64</td>
</tr>
</tbody>
</table>
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 available online.

<table>
<thead>
<tr>
<th>Name</th>
<th>RA</th>
<th>DEC</th>
<th>Associations</th>
<th>$S_{545}$ [mJy]</th>
<th>$\sigma_{250}$</th>
<th>$\sigma_{350}$</th>
<th>$\sigma_{500}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCCS1 545 G019.76-58.74</td>
<td>337.311</td>
<td>-29.670</td>
<td>Cluster Candidate</td>
<td>225±164</td>
<td>2.92</td>
<td>1.68</td>
<td>3.25</td>
</tr>
<tr>
<td>PCCS1 545 G354.79-79.57</td>
<td>3.060</td>
<td>-33.226</td>
<td>Cluster Candidate</td>
<td>335±158</td>
<td>1.22</td>
<td>1.39</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC545 G015.45-60.34</td>
<td>339.084</td>
<td>-31.902</td>
<td>Cluster Candidate</td>
<td>607±87</td>
<td>2.72</td>
<td>3.27</td>
<td>4.01</td>
</tr>
<tr>
<td>PCCS1 545 G058.72+82.59</td>
<td>200.566</td>
<td>30.136</td>
<td>Cluster Candidate</td>
<td>349±161</td>
<td>1.22</td>
<td>1.39</td>
<td>3.64</td>
</tr>
<tr>
<td>PCCS1 545 G027.38+84.85</td>
<td>198.561</td>
<td>26.515</td>
<td>Cluster Candidate</td>
<td>433±136</td>
<td>1.45</td>
<td>2.23</td>
<td>4.01</td>
</tr>
<tr>
<td>PCCS1 545 G084.40+81.05</td>
<td>199.569</td>
<td>33.968</td>
<td>Cluster Candidate</td>
<td>387±230</td>
<td>1.88</td>
<td>1.39</td>
<td>3.25</td>
</tr>
<tr>
<td>PLCKERC545 G146.77+52.48</td>
<td>164.280</td>
<td>59.031</td>
<td>Cluster Candidate</td>
<td>606±73</td>
<td>4.09</td>
<td>4.93</td>
<td>1.06</td>
</tr>
<tr>
<td>PCCS1 545 G222.38-54.39</td>
<td>53.046</td>
<td>-27.117</td>
<td>Cluster Candidate</td>
<td>182±185</td>
<td>2.10</td>
<td>3.02</td>
<td>2.01</td>
</tr>
<tr>
<td>PLCKERC545 G060.36+66.56</td>
<td>218.577</td>
<td>35.554</td>
<td>Cluster Candidate</td>
<td>872±114</td>
<td>0.77</td>
<td>4.47</td>
<td>5.70</td>
</tr>
<tr>
<td>PLCKERC545 G260.28-39.75</td>
<td>71.996</td>
<td>-52.643</td>
<td>Cluster Candidate</td>
<td>471±46</td>
<td>3.71</td>
<td>3.27</td>
<td>3.25</td>
</tr>
</tbody>
</table>
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,

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

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 $\mu$, the probability of observing $M$ sources is given by:

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

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:

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

However, this needs to be renormalised as the maximum possible observed field galaxies is now $M$ rather than $\infty$ (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:

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

where we simply sum over all of the possible arrangements of $N_{\text{field}} + N_{\text{cluster}}$ 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:

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

Which can be further simplified to:

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

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