

HerMES: deep number counts at 250 μm, 350 μm and 500 μm in the COSMOS and GOODS-N fields and the build-up of the cosmic infrared background

M. Béthermin, E. Le Floc'H, O. Ilbert, A. Conley, Guilaine Lagache, A. Amblard, V. Arumugam, H. Aussel, S. Berta, J. Bock, et al.

▶ **To cite this version:**

M. Béthermin, E. Le Floc'H, O. Ilbert, A. Conley, Guilaine Lagache, et al.. HerMES: deep number counts at 250 μm, 350 μm and 500 μm in the COSMOS and GOODS-N fields and the build-up of the cosmic infrared background. Astronomy and Astrophysics - A&A, 2012, 542, pp.A58. 10.1051/0004-6361/201118698 . cea-00840110

HAL Id: cea-00840110
https://cea.hal.science/cea-00840110

Submitted on 20 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Astronomy & Astrophysics manuscript no. SPIRE˙STACKING˙COUNTS˙v8
August 6, 2018
© ESO 2018

arXiv:1203.1925v2 [astro-ph.CO] 17 Apr 2013

HerMES: deep number counts at 250 μm, 350 μm and 500 μm in the COSMOS and GOODS-N fields and the build-up of the cosmic infrared background

M. Béthermin[1,2], E. Le Floc'h[1], O. Ilbert[3], A. Conley[4], G. Lagache[2], A. Amblard[5], V. Arumugam[6], H. Aussel[1], S. Berta[7], J. Bock[8,9], A. Boselli[3], V. Buat[3], C.M. Casey[10], N. Castro-Rodríguez[11,12], A. Cava[13], D.L. Clements[14], A. Cooray[15,8], C.D. Dowell[8,9], S. Eales[16], D. Farrah[17], A. Franceschini[18], J. Glenn[19,4], M. Griffin[16], E. Hatziminaoglou[20], S. Heinis[3], E. Ibar[21], R.J. Ivison[21,6], J.S. Kartaltepe[22,23], L. Levenson[8,9], G. Magdis[1], L. Marchetti[18], G. Marsden[24], H.T. Nguyen[9,8], B. O'Halloran[14], S.J. Oliver[17], A. Omont[25], M.J. Page[26], P. Panuzzo[1], A. Papageorgiou[16], C.P. Pearson[27,28], I. Pérez-Fournon[11,12], M. Pohlen[16], D. Rigopoulou[27,29], I.G. Roseboom[17,6], M. Rowan-Robinson[14], M. Salvato[30], B. Schulz[8,31], Douglas Scott[24], N. Seymour[32,26], D.L. Shupe[8,31], A.J. Smith[17], M. Symeonidis[26], M. Trichas[33], K.E. Tugwell[26], M. Vaccari[18], I. Valtchanov[34], J.D. Vieira[8], M. Viero[8], L. Wang[17], C.K. Xu[8,31], and M. Zemcov[8,9]

(Affiliations can be found after the references)

Received ??? / Accepted ???

Abstract

Aims. The Spectral and Photometric Imaging REceiver (SPIRE) onboard the *Herschel* space telescope has provided confusion limited maps of deep fields at 250 μm, 350 μm, and 500 μm, as part of the *Herschel* Multi-tiered Extragalactic Survey (HerMES). Unfortunately, due to confusion, only a small fraction of the cosmic infrared background (CIB) can be resolved into individually-detected sources. Our goal is to produce deep galaxy number counts and redshift distributions below the confusion limit at SPIRE wavelengths (∼20 mJy), which we then use to place strong constraints on the origins of the cosmic infrared background and on models of galaxy evolution.

Methods. We individually extracted the bright SPIRE sources (>20 mJy) in the COSMOS field with a method using the positions, the flux densities, and the redshifts of the 24 μm sources as a prior, and derived the number counts and redshift distributions of the bright SPIRE sources. For fainter SPIRE sources (<20 mJy), we reconstructed the number counts and the redshift distribution below the confusion limit using the deep 24 μm catalogs associated with photometric redshift and information provided by the stacking of these sources into the deep SPIRE maps of the GOODS-N and COSMOS fields. Finally, by integrating all these counts, we studied the contribution of the galaxies to the CIB as a function of their flux density and redshift.

Results. Through stacking, we managed to reconstruct the source counts per redshift slice down to ∼2 mJy in the three SPIRE bands, which lies about a factor 10 below the 5σ confusion limit. Our measurements place tight constraints on source population models. None of the pre-existing models are able to reproduce our results at better than 3-σ. Finally, we extrapolate our counts to zero flux density in order to derive an estimate of the total contribution of galaxies to the CIB, finding $10.1^{+2.6}_{-2.3}$ nW m^{-2} sr^{-1}, $6.5^{+1.7}_{-1.6}$ nW m^{-2} sr^{-1}, and $2.8^{+0.9}_{-0.8}$ nW m^{-2} sr^{-1} at 250 μm, 350 μm, and 500 μm, respectively. These values agree well with FIRAS absolute measurements, suggesting our number counts and their extrapolation are sufficient to explain the CIB. We find that half of the CIB is emitted at z =1.04, 1.20, and 1.25, respectively. Finally, combining our results with other works, we estimate the energy budget contained in the CIB between 8 μm and 1000 μm: 26^{+7}_{-3} nW m^{-2} sr^{-1}.

Key words. Cosmology: observations – Cosmology: diffuse radiation – Galaxies: statistics – Galaxies: photometry – Submillimeter: galaxies – Submillimeter: diffuse background

1. Introduction

About half of the relic energy arising from the emission of galaxies, which we refer to as the Extragalactic Background Light (EBL), is contained in the cosmic infrared background (CIB), which lies between 8 μm and 1000 μm, and peaks at around 150 μm (Hauser & Dwek 2001; Dole et al. 2006). The first absolute measurements of the CIB were performed in the nineties with the Far-Infrared Absolute Spectrophotometer (FIRAS; Puget et al. 1996; Fixsen et al. 1998; Lagache et al. 1999) and the Diffuse Infrared Background Experiment (DIRBE; Hauser et al. 1998) onboard the *COsmic Background Explorer* (*COBE*). The far-infrared emission from galaxies is mainly due to dust heated by ultraviolet photons re-radiate in the infrared. A small fraction of these far-infrared emission (∼15%) are due to accretion processes (Alexander et al. 2005; Jauzac et al. 2011). The CIB thus primarily gives a budget of infrared photons emitted by star-formation processes.

More recently, deep number counts (flux density distributions of infrared sources) have been measured in the mid- and far-infrared domain, thanks to the sensitivity of the *Spitzer* (Werner et al. 2004) and *Herschel*[1] (Pilbratt et al. 2010) space telescopes. They exhibit power-law behavior at the faint end, which can be extrapolated to zero flux density in order to estimate the contribution of all the galaxies to the

[1] Herschel is an ESA space observatory with science instruments provided by European-led Principle Investigator consortia and with important participation from NASA.

CIB (e.g. Papovich et al. 2004 at 24 μm with *Spitzer*/MIPS, Béthermin et al. 2010a at 24 μm, 70 μm, and 160 μm with *Spitzer*/MIPS, Berta et al. 2011 at 70 μm, 100 μm, and 160 μm with *Herschel*/PACS). These estimations of the total CIB agree with the absolute measurements performed by *COBE*, suggesting the CIB is now explained shortward of 160 μm. At longer wavelengths, due to source confusion (Dole et al. 2003; Nguyen et al. 2010), the *Herschel*/SPIRE instrument (Griffin et al. 2010) can directly resolve only 20%, 12%, and 6% of the CIB at 250 μm, 350 μm, and 500 μm, respectively (Oliver et al. 2010b). Due to the limited depth of these confusion-limited observations, the break and the power-law behavior of the counts at faint flux density cannot be seen. It is thus necessary to use statistical tools like $P(D)$ analysis[2] (Condon 1974; Patanchon et al. 2009) or stacking (Dole et al. 2006; Marsden et al. 2009) to study the origins of the sub-mm part of the CIB.

Using a P(D) analysis, Patanchon et al. (2009) produced deep counts from the Balloon-borne Large-Aperture Submillimeter Telescope (BLAST, Pascale et al. 2008; Devlin et al. 2009) data. They were only able to constrain one data point below 100 mJy at 250 μm, and were not sensitive to more subtle features in the shape of the counts. Using a stacking analysis, Béthermin et al. (2010b) managed to detect the peak of the Euclidian-normalized counts at 250 μm in the BLAST data, but not at longer wavelengths. Using a $P(D)$ analysis on SPIRE data, Glenn et al. (2010) managed to clearly detect this peak at 250 μm and 350 μm, but not at 500 μm. In all these cases, the uncertainties are too large to reliably detect a power-law behavior at the faint end.

A stacking analysis of SPIRE data similar to that performed on BLAST data by Béthermin et al. (2010b) could significantly reduce the uncertainties and provide more precise information on the sources which make up the CIB. Le Floc'h et al. (2009) and Berta et al. (2011) also showed that counts per redshift slice are strong constraints for galaxy evolution models (e.g. Le Borgne et al. 2009, Valiante et al. 2009, Marsden et al. 2011, Béthermin et al. 2011, Gruppioni et al. 2011, Rahmati & van der Werf 2011). Lastly, unlike a P(D) analysis, stacking allows us to measure directly the counts in redshift slices, but requires a prior catalog. Thus, here we perform a stacking analysis in the SPIRE bands, in the COSMOS and GOODS-N fields to produce deep counts per redshift slice in SPIRE bands, combining the *Herschel* Multi-tiered Extragalactic Survey (HerMES)[3] data (Oliver et al. 2011) and the ancillary data.

The paper is organized as follows. In Sect. 2, we present the different data sets used in our analysis. We then introduce the method used to measure the number counts of resolved sources (Sect. 3) and another method based on stacking to reconstruct the number counts below the confusion limit (Sect. 4). In Sect. 5, we detail the estimation of the statistical uncertainties. Sect. 6 presents a end-to-end simulation used to check the accuracy of our method. In Sect 7, we interpret our number counts and compare them with previous measurements and models of galaxy evolution. The same thing is done in Sect. 8 for the redshift distributions. In Sect. 9, we derive constraints on the CIB level and its redshift distribution from our number counts. We finally discuss our results (Sect. 10) and conclude (Sect. 11).

2. Data

2.1. SPIRE maps at 250 μm, 350 μm and 500 μm

The SPIRE instrument (Griffin et al. 2010) onboard the *Herschel Space Observatory* (Pilbratt et al. 2010) observed the COSMOS field as part of the *Herschel* Multi-tiered Extragalactic Survey (HerMES) program Oliver et al. (2011). The maps were built using an iterative map-making technique (Levenson et al. 2010). The full width at half maximum (FWHM) of the SPIRE beam (Swinyard et al. 2010) is 18.1″, 24.9″, and 36.6″ at 250 μm, 350 μm, and 500 μm, respectively. The typical instrumental noise is 1.6, 1.3, and 1.9 mJy beam^{-1} in COSMOS (1.6, 1.3, and 2.0 mJy beam^{-1} in GOODS-N) and the 1-σ confusion noise is 5.8, 6.3 and 6.8 mJy beam^{-1} in the three wavebands (Nguyen et al. 2010). The maps are thus confusion limited. The absolute calibration uncertainties in point sources are estimated to be 7% (Swinyard et al. 2010, updated in the SPIRE Observers' Manual[4]).

2.2. Ancillary data in COSMOS

Deep 24 μm imaging of the COSMOS field was performed by the *Spitzer Space Telescope* (S-COSMOS, Sanders et al. 2007). The associated catalog reaches 90% completeness at 80 μJy (Le Floc'h et al. 2009). This catalog was matched with photometric redshifts of Ilbert et al. (2009). Due to the high density of optical sources compared with the size of the MIPS beam, the cross-identification can be ambiguous in many cases. An intermediate matching was thus performed with the K and IRAC bands where the source density is smaller, which helps to discriminate between several optical counterparts in a MIPS beam (Le Floc'h et al. 2009).

In this paper, we use an updated version of the photometric redshift catalog of Ilbert et al. (2009) (v1.8). This version uses new deep H-band data. However, this catalog is not optimized for AGN. For the sources detected by XMM-*Newton* we instead use the photometric redshifts of Salvato et al. (2009), estimated with a technique specific to AGN. In addition, 10 000 sources have spectroscopic redshifts provided by the S-COSMOS team Lilly et al. (2007), which, where available, are used instead of the photometric redshifts. Details of the updated COSMOS S_{24} + z catalog will be given in Le Floc'h et al. (in prep.). In this new version, 96% of the 27 811 $S_{24} > 80\,\mu$Jy sources have redshifts (9.7% of them are spectroscopic).

2.3. Ancillary data in GOODS-N

In the GOODS-N field, we use the 24 μm catalog of Magnelli et al. (2011). This catalog was built using the IRAC catalog at 3.6 μm as a prior, and has an estimated 3-σ depth of 20 μJy, but at this depth the completeness is only ∼50%. The stacking of an incomplete catalog can bias the results (Béthermin et al. 2010b; Heinis et al. in prep., Vieira et al. in prep.). According to simulations, cutting at 80% completeness results in smaller bias. We thus cut the catalog

[2] $P(D)$ analysis is a statistical method used to estimate the number counts in a field from the pixel histogram of an extragalactic map.

[3] hermes.sussex.ac.uk

[4] http://herschel.esac.esa.int/Docs/SPIRE/pdf/spire_om.pdf

at 30 μJy (80% completeness limit) to have a more complete and reliable sample. These sources were matched with the photometric redshifts of Eales et al. (2010) (97.4% of the 2791 24 μm sources are associated with a redshift). The data fusion of these two catalogs will be explained in more detail in Vaccari et al. (in prep.).

3. Measuring the statistical properties of the resolved sources

In order to build counts per redshift slice and redshift distributions of the sources selected by their SPIRE flux densities, we require catalogs containing SPIRE flux densities and redshifts. The redshift catalogs are built from optical and near-infrared catalogs. We start from catalogs of the 24 μm sources which have optical counterparts and thus photometric redshifts. Due to the large beam of SPIRE, it is not trivial to identify the MIPS 24 μm counterpart for a given SPIRE source. To avoid this problem, we directly measure the SPIRE flux denisty of the 24 μm sources in the maps by PSF-fitting assuming a known position (Béthermin et al. 2010b; Chapin et al. 2011; Roseboom et al. 2010). Sect. 3.3 and 7 discuss the relevance of this choice of prior.

The GOODS-N field, being much smaller than COSMOS, has little impact on the statistical uncertainties (using GOODS-N+COSMOS reduces the uncertainties of 0.2% compared to COSMOS only). The inclusion of GOODS-N introduces heterogeneity to the 24 um catalogs, which are built using different methods between the two fields. For this reason, we used only the COSMOS field in the following section.

3.1. Source extraction

We use the FASTPHOT PSF-fitting routine, described in Béthermin et al. (2010b), to fit the following model to the SPIRE map:

$$m = \sum_{k=1}^{N_s} S_k \times b_{x_k,y_k} + \mu, \quad (1)$$

where m is the map, N_s the number of sources, S_k the SPIRE flux density of the k-th source, b_{x_k,y_k} a point spread function (PSF) centered on the position of the k-th source (x_k, y_k), and μ a constant background. The catalog of positions used as input to FASTPHOT is discussed below. The free parameters fit by FASTPHOT are the SPIRE fluxes of sources in the prior list S_k and the level of the constant background μ. We used the PSF based on the Neptune scan from Glenn et al. (2010) [5]. The map is not fit in one pass, but split into 100×100 pixel regions (the pixel sizes are 6.0″, 8.3″, and 12″at 250 μm, 350 μm, and 500 μm, respectively). Each region was fit independently. To limit the edge effects, we also fit simultaneously an additional region of 20 pixels around each 100×100 region. The positions of the sources in both the central and additional regions are used by FASTPHOT, but we keep in the final catalog only the photometry of the sources in the central region. The signal at 20 pixels (~6 times the FWHM) from the center of a source is negligible. A source outside of the additional region cannot thus significantly affect the photometry in the central region.

The FASTPHOT routine suffers some instabilities when two sources are too close to one another. We thus do not use the position of all the 24 μm sources in FASTPHOT. For several redshift and 24 μm flux density slices, we estimated the mean color by stacking (see Sect. 4.1). We then use these mean colors to estimate the flux density of each source in the SPIRE bands. A 24 μm source is included in the position list of FASTPHOT only if it has the highest estimated SPIRE flux density in a 0.5×FWHM radius. This process was therefore performed independently in each band. Some sources with unusually high sub-mm/mid-infrared colors could be missed by this method, but there are few objects of this kind (see Sect. 3.3).

To avoid unphysical negative flux densities for faint sources lying on negative fluctuations of the noise, we run FASTPHOT iteratively, removing from the position list the sources with negative flux densities at each iteration. Removing a source from the input catalog is equivalent to assuming its flux density is zero, which is the most probable value in this case.

3.2. Estimating photometric noise

We estimate the photometric noise from the standard deviation of the FASTPHOT residual map, finding the values 4.6, 5.5 and 5.1 mJy at 250 μm, 350 μm, and 500 μm, respectively. These values are about 20% lower than the combination the 1-σ confusion noise measured by Nguyen et al. (2010) and the instrumental noise (6.0, 6.4 and 7.0 mJy). Our method is thus more efficient than a naive blind extraction. We chose to cut our statistical analysis at 20 mJy in the three SPIRE passbands, which corresponds to about 4-σ.

In order to cross-check our estimate of the photometric noise, we inject 200 artificial point sources in the real SPIRE map and add them in the input position list of FASTPHOT. We add a random shift drawn from a 2D Gaussian with $\sigma = 2''$ to the source position in order to simulate the astrometric uncertainties of the real catalog. We then rerun FASTPHOT and compare the input and output flux densities. Fig. 1 shows the histogram of the difference between the recovered and input flux densities. We found a 1-σ photometric noise of 3.9, 5.2 and 5.1 mJy at 250 μm, 350 μm, and 500 μm, respectively. The values are similar to those estimated from the residual map. Comparing the two sets of values, we can estimate an error of about 20% on the photometric noise.

3.3. 24 μm Dropouts

A fundamental limitation of our model is that it is not sensitive to any population of sources that are faint at 24 microns but bright in the SPIRE passbands (24 μm dropouts). No such population is known or theoretically predicted, except possibly at very high redshifts, but the possibility remains that nature has been more inventive than we have. In this section we attempt to test whether there is any evidence for such sources, and do not find any.

[5] Beam data are also available from the Herschel Science Centre at ftp://ftp.sciops.esa.int/pub/hsc-calibration/SPIRE/PHOT/Beams

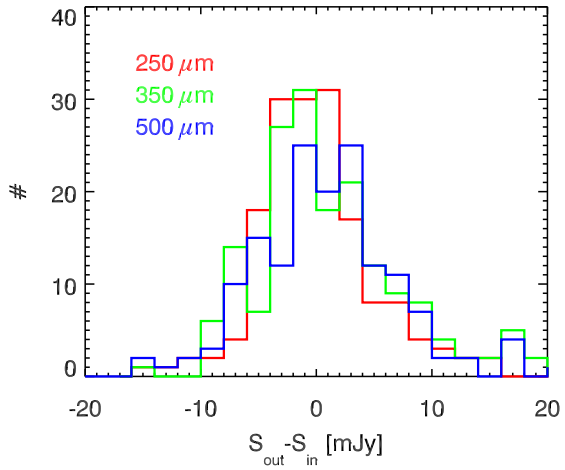


Figure 1. Simulation of the photometric uncertainties: histogram of the difference between the recovered and the input flux density of the artificial sources injected in the real SPIRE map and re-extracted with FASTPHOT.

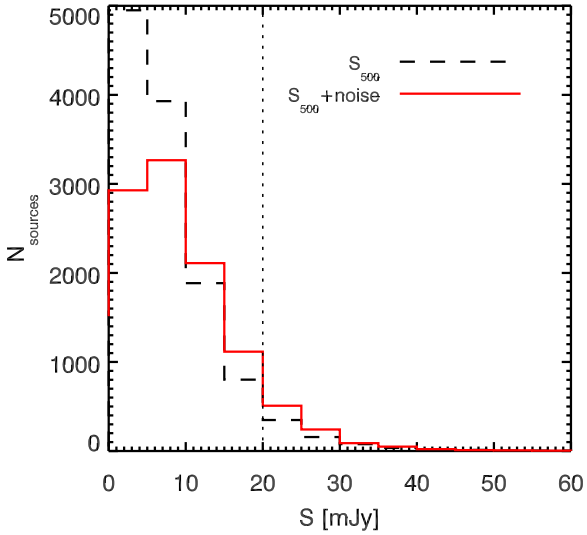


Figure 2. Effect of the photometric uncertainties in the flux density distribution at 500 μm. *Black dashed line*: distribution of the flux density measured at the position of the 24 μm sources. *Red solid line*: the same distribution after adding a 5.1 mJy random Gaussian noise to each measured flux density. Due to photometric noise, some sources have a negative flux density (put to zero in our iterative algorithm) and are not represented here. *Black dotted line*: flux density cut used in our analysis (20 mJy).

First, we study the residual SPIRE maps after removing all of the sources extracted with FASTPHOT to estimate the number of sources missed by the extraction. The density of remaining sources is quite small so that confusion is not a problem. We then search for additional sources by looking for peaks in the beam smoothed residual map to a depth of 20 mJy, and find that the fraction[6] of such possible sources is only 1.3%, 0.7%, and 0.6% at 250 μm, 350 μm, and 500 μm, respectively. Next, we have compared our prior catalog with the source list from the blindly-extracted HerMES catalog (Smith et al. 2012), which is limited to sources brighter than 20 mJy. The fraction of sources in the blind catalog without counterpart in the prior catalog strongly depends on the choice of matching radius. For a narrow radius of 0.5 FWHM, we obtained 3.2%, 2.6%, and 0.6% whereas for a large radius of 1 FWHM, we obtained 0.6%, 0.3%, and 0.0%, at 250 μm, 350 μm, and 500 μm, respectively. However, with the narrow radius, we miss some sources due to astrometric uncertainties, and with the large radius, we possible have a contamination by neighboring sources. It is expected that the fraction of dropouts decreases with the flux density. Nevertheless, this behavior is hard to constrain because of the small number of bright sources. Note however that the fraction of dropouts at the flux density limit is very close to the values obtained for the full sample because of the steep slope of the counts. Thus, we conclude that our catalog, based on the 24 μm prior, is very close to complete above 20 mJy. From a blind extracted catalog of H-GOODS data, Magdis et al. (2011) estimated the fraction of 24 μm dropouts for H-GOODS fields and shallower fields. They predict a dropout faction smaller than 2% in the COSMOS field.

Finally, we can compare our number counts measurement to other analyses that did not make use of a 24 μm prior. We find good agreement with the blind extractions of Oliver et al. (2010b) and Clements et al. (2010), but note that these analyses required significant model corrections for Eddington bias and confusion. We also find good agreement with the P(D) analysis of the SPIRE maps by Glenn et al. (2010), which, by construction, is not affected by either issue. We take these comparisons as a strong indication that we have not missed a statistically significant population, at least in terms of the redshift integrated number counts. However, we must acknowledge that the fraction of dropouts could evolve with redshift, and in particular the high redshift bins may be less complete.

3.4. Correction of the biases

Intuitively, the simplest way to compute the source counts is to measure the number of sources in a flux density bin and divide it by the width of the bin and the surface area of the field. However, due to photometric noise, this estimate is biased. In fact, for a prior-based extraction, we do not have a flux boosting phenomena (which appears at low signal to noise ratio for blind extraction, because the completeness is higher for sources lying on positive fluctuations of the noise, as discussed e.g. in Béthermin et al. 2010b), but another statistical effect, Eddington bias, biases the counts measurement, as illustrated by Fig. 2. The black dashed line shows the distribution of the 500 μm flux densities measured at the prior positions. We will assume this distribution is close to the real one, and will somewhat arbitrarily refer to it as initial distribution. The red line shows the same distribution, but adding a 1-σ 5.1 mJy Gaussian error on the flux density of each source, called measured distribution. At bright flux density ($S_{500} > 20$ mJy), we can observe an excess in the measured distribution compared to the initial one.

[6] The fraction of dropouts is defined as $\frac{N_d}{N_d+N_p}$, where N_d is the number of sources brighter than 20 mJy found in the residual map and N_p the number of sources brighter than 20 mJy extracted by FASTPHOT in the normal map.

To correct this bias, we use a Monte Carlo (MC) method as in Béthermin et al. (2010b). We compute 1000 realizations of the bias in each flux density (regular in logarithm from 20 mJy) and redshift ($0 < z < 0.5$, $0.5 < z < 1$, $1 < z < 2$, and $z > 2$) bin, and use them to compute the mean correction and its uncertainty:

- We start from the measured distribution of the 250 μm, 350 μm, or 500 μm flux density of the 24 μm sources in the prior list of FASTPHOT in a given redshift bin, and assume it is close to the initial distribution. This last hypothesis is a significant approximation, but the selection function of the procedure used to construct the prior catalog is too complex to be modeled without introducing strong assumptions about galaxy evolution.
- We draw with replacement N sources in the initial sample, where N is the number of sources in the initial sample. This bootstrap step is used to take into account the sample variance on the initial flux density distribution.
- We add a Gaussian random photometric noise to the flux density of each source. We use the values of the noise found in Sect. 3.2 plus a 20% systematic shift (different at each iteration of the MC procedure), which takes into account the systematic uncertainty on the determination of the noise.
- We compute the bias on the counts dividing the counts from the drawn sample before and after adding the photometric noise.

Table B.1 shows the corrective factor in various flux density and redshift bins and at various wavelengths. This correction can reach 40% in the fainter flux density bins and decreases at brighter flux densities.

Similar corrections are applied to the redshift distributions. However, in addition, we apply a random error to the redshift based on the uncertainties provided in the photo-z catalog (but without taking into account the catastrophic outliers) during the MC procedure. In this case, there is only one flux density bin (>20 mJy).

4. Measuring the statistical properties of sources below the confusion limit

Due to source confusion, SPIRE cannot resolve the bulk of the CIB into individual sources (Nguyen et al. 2010; Oliver et al. 2010b). Nevertheless, about 80% of the CIB is resolved at 24 μm (Papovich et al. 2004; Béthermin et al. 2010a). We thus perform a stacking analysis using the 24 μm prior to probe fainter populations and resolve a larger fraction of the sub-mm CIB.

4.1. Stacking method

Stacking is a statistical method which allows us to measure the mean flux density of a population of sources selected at another wavelength, but which are too faint to be detected individually at the working wavelength. Several methods can be used (e.g. Dole et al. 2006, Marsden et al. 2009, see the discussion in Vieira et al. in prep.). We use the following method (also used in Vieira et al. in prep.): we first subtract the mean of the SPIRE map in the region covered by the 24 μm observations. We then compute the mean signal in pixels which has a source centered on them. This provides the mean flux density of the population, because the SPIRE maps are in Jy beam^{-1}. Vieira et al. (in prep.) showed that this method is more accurate in a confusion-limited case than PSF-fitting on a stacked image. The uncertainties are estimated using a bootstrap method (Jauzac et al. 2011).

Due to the large number of sources in COSMOS, we can split our 24 μm sample into eight redshift bins ($0 < z < 0.25$, $0.25 < z < 0.5$, $0.5 < z < 0.75$, $0.75 < z < 1$, $1 < z < 1.5$, $1.5 < z < 2$, $2 < z < 3$, and $z > 3$) and logarithmic flux density slices (80 μJy$< S_{24} <$ 172 μJy, 172 μJy$< S_{24} <$ 371 μJy, 371 μJy$< S_{24} <$ 800 μJy, and 800 μJy$< S_{24} <$ 1723 μJy). In GOODS-N, we use the same redshift slices, but a single flux density slice (30 μJy$< S_{24} <$ 80 μJy). This choice of the number of bins was done to have a compromise between a fine grids in 24 μm flux density and redshift, but also a reasonable number of sources to stack in each bins to obtain a good signal-to-noise ratio. We stack the sources in each bin to compute their mean flux density in the three SPIRE bands. Fig. 3 shows the mean flux density as a function of wavelength, which, as expected, decreases rapidly in low redshift bins and peaks between 350 μm and 500 μm in the $z > 3$ bin. The mean color in each bin is computed by dividing the mean SPIRE flux density by the mean 24 μm flux density.

4.2. Scatter of the photometric properties of the stacked populations

The uncertainties given by the bootstrap method, $\sigma_{\rm boot}$, are

$$\sigma_{\rm boot} = \frac{\sqrt{\sigma^2_{\rm instr} + \sigma^2_{\rm conf} + \sigma^2_{\rm pop}}}{\sqrt{N_{\rm stack}}}, \quad (2)$$

where σ_{instr} is the instrumental noise, $\sigma_{\rm conf}$ the confusion noise, $\sigma_{\rm pop}$ is the standard deviation of the flux density of the population, and $N_{\rm stack}$ the number of stacked sources. The quantity $\sqrt{\sigma^2_{\rm instr} + \sigma^2_{\rm conf}}$ can be estimated from the standard deviation of the map, allowing us to deduce $\sigma_{\rm pop}$ from our bootstrap analysis.

While this formula is true for a Gaussian distribution, we note that the distribution of colors of the sources are probably better described by a log-normal distribution. Fig. 4 shows the distribution of the logarithm of the S_{250}/S_{24} colors of the resolved sources ($S_{250} > 20$ mJy) in the $1 < z < 1.5$ redshift bin (this redshift bin was chosen because it has the larger number of sources). The confusion noise is also non Gaussian (Glenn et al. 2010). Nevertheless, due to central limit theorem, these distributions of the mean flux density tend to be Gaussian if a sufficient number of sources are stacked. Fig. 5 illustrates this property. The red histogram is the pixel histogram of the 250 μm SPIRE map. It is not Gaussian, because the confusion noise is not. The blue one is the distribution of the mean signal in 100 pixels taken randomly in 100 000 realizations (typically the effect of the instrumental and confusion noise on a stack of 100 sources). This histogram is much closer to a Gaussian. The same thing happens for the color scatter term. The Gaussian approximation is thus very relevant here.

The scatter on the $S_{\rm SPIRE}/S_{24}$ color can be estimated by dividing $\sigma_{\rm pop}$ by the mean flux density of the population. We do not detect a significant evolution of this scatter with redshift, wavelength or 24 μm flux density. We use the median and the standard deviation of the values found in the different redshift and 24 μm flux density bins, and find a scatter

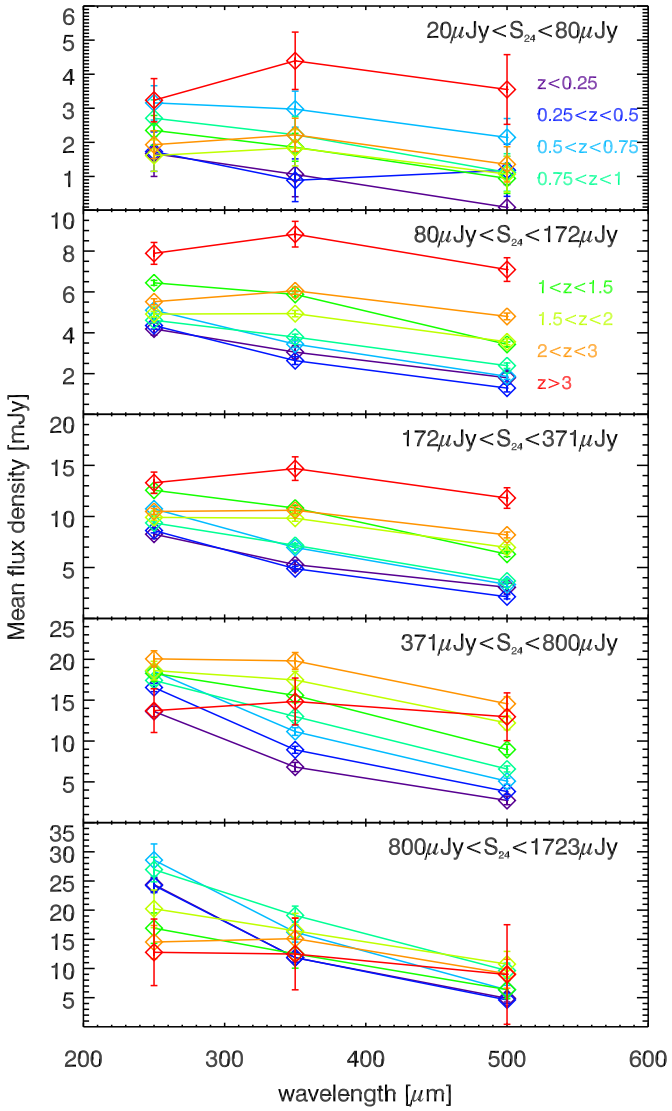


Figure 3. Mean flux density measured by stacking as a function of wavelength. The various redshift bins are represented using various colors. Each panel corresponds to each 24 μm flux density bins. The error bars are estimated with a bootstrap method.

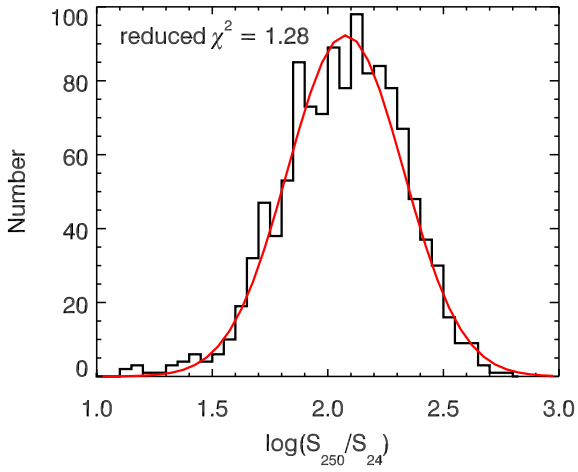


Figure 4. *Black histogram*: distribution of the logarithm of the S_{250}/S_{24} colors of the resolved sources ($S_{250} > 20$ mJy) in the 1<z<1.5 redshift bin. *Red line*: fit of the histogram by a Gaussian.

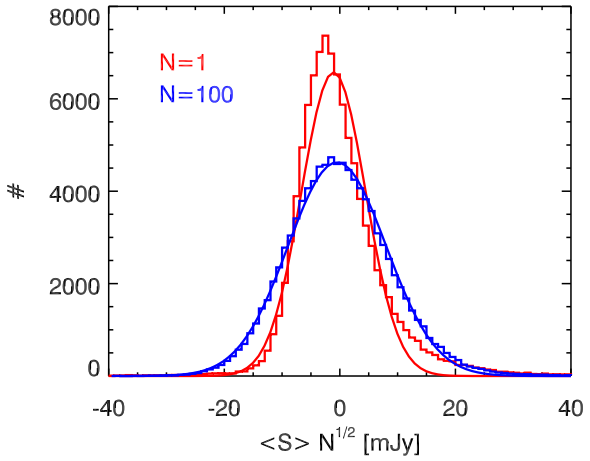


Figure 5. *Red histogram*: Pixel histogram of the 250 μm SPIRE map in COSMOS. *Blue histogram*: Histogram of the mean signal in 100 pixels taken randomly in 100 000 realizations. *Red and blue lines*: Gaussian fit of the previous histograms.

$\sigma_{\rm color} = \sigma_{\rm pop} / < S_{\rm SPIRE} >$ of 68±35% (in linear units). This agrees with the value of 62% found for the resolved sources (see Fig. 4).

4.3. Does the clustering of sources introduce a bias?

The simplest stacking method assumes implicitly that the sources in the map are not clustered, but this has been shown to be unrealistic and must be accounted for (Béthermin et al. 2010b; Viero et al. 2011; Penner et al. 2011). We have performed several tests in the COSMOS field to estimate the bias due to clustering, which we now describe.

4.3.1. Method A: convolution of the 24 μm map with the SPIRE beam

A simple way to estimate the bias due to clustering is to convolve the 24 μm map with a Gaussian kernel to obtain a 24 μm map with a Gaussian PSF of the same FWHM as the SPIRE map Oliver et al. (2010a). To match resolutions, we use a Gaussian kernel with beamsize $\sigma_{\rm kernel} = \sqrt{\sigma^2_{\rm SPIRE} - \sigma^2_{\rm MIPS}}$. We measure the mean flux density of the 24 μm sources by stacking the 24 μm catalog on this convolved 24 μm map. The bias due to clustering is estimated by comparing the mean flux density measured by stacking with the mean flux density estimated from the 24 μm catalog. We find biases of 5±2%, 11±2%, and 20±5% at 250 μm, 350 μm, and 500 μm, respectively. This method is equivalent to building a simulated map assuming a single

color C for all the 24 μm objects (including the ones below the detection limit at 24 μm), measuring the mean flux density of the selected population by stacking on this convolved map, and comparing it with the mean flux density coming from the catalog ($< S_{24} > \times C$). The same color factor C is present in the mean stacked flux measured by stacking in the convolved map and in the mean flux coming from the catalog. It thus disappears when we compute the relative difference between these two quantities. We thus take $C = 1$ for simplicity. As expected, the bias due to clustering increases with the size of the beam. This estimate is exact only if the S_{SPIRE}/S_{24} color is constant, or if the properties of the angular clustering do not evolve with the color of the sources (and thus the redshift). These assumptions are not going to be exactly met, so that next we use another method to cross-check this estimate.

4.3.2. Method B: simulation based on mean colors measured by stacking

In Sect. 4.1, the $S_{\rm SPIRE}/S_{24}$ mean color as a function of the 24 µm flux density and redshift were measured by stacking. We use these mean colors and the scatter measured in Sect. 4.2 to generate mock SPIRE flux densities for the sources in the $S_{24}+z$ catalog, and then build a simulated map of the COSMOS field using the position given in the $S_{24}+z$ catalog, the estimated SPIRE flux density, and the SPIRE PSF. Random Gaussian noise was added following the noise map of the real data. We then stacked all the 24 µm sources, and compared the mean flux density measured by stacking in the simulated map and the mean flux density in the mock catalog. We find a bias of 7.0±0.9%, 10.4± 0.7%, and 20.6±1.2% at 250 µm, 350 µm, and 500 µm, respectively, in agreement with the values provided by method A. The main drawback of this method is that any bias due to sources undetected at 24 microns is not modeled.

We have also stacked sub-samples selected in redshift and/or in 24 µm flux density. The bias tends to slightly decrease with the redshift and the 24 µm flux density cut. Nevertheless, this evolution is small (below 3%), and the significance is smaller than 3-σ. We thus chose to neglect it, and assume a single value for the bias due to the clustering.

4.3.3. Method C: fitting the profile of the stacked image

For method C, we follow the Dole et al. (2006) method to produce our stacked images. In the presence of clustering, this image can be fit by the following function (Béthermin et al. 2010b; Henis et al. in prep.):

$$M = \alpha \times b + \beta \times \left(\frac{b * w}{\max(b * w)} \right), \quad (3)$$

where M is the stacked image, w the auto-correlation function (ACF), $*$ the convolution product, and b the beam function. The PSF is normalized to unity at the center to match the per-beam normalization of the maps. α and β are free parameters in the fit. The results of the fit are plotted in Fig. 6. In order to estimate the uncertainties, the fit was performed on 1000 bootstrap samples. If we measure the photometry in the central pixel of the PSF, the bias due to clustering is β/α. We found 7.7±0.5%, 10.3±0.8%, and 19.1±1.8% at 250 µm, 350 µm, and 500 µm, respectively. The uncertainty here is the standard deviation of the values found for the different bootstrap samples. As expected, we also found that α and β are significantly anti-correlated (correlation coefficients of −0.46, −0.56, and −0.57 at 250 µm, 350 µm, and 500 µm, respectively). As was the case for method B, we do not detect any significant evolution of the bias with redshift.

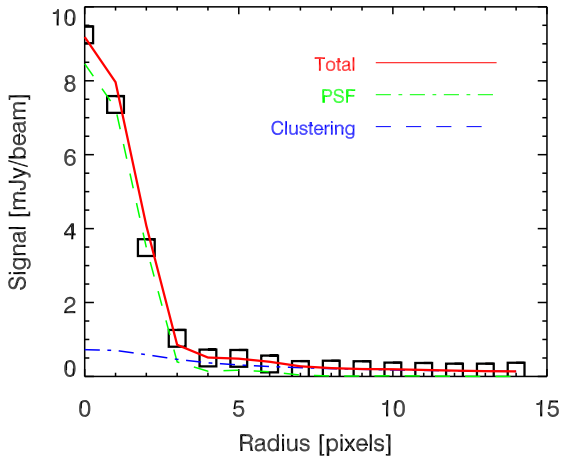


Figure 6. Radial profile of the stacked image at 250 µm of all the 27 811 $S_{24} > 80\,\mu$Jy sources in COSMOS. *Black squares*: measurements. *Red solid line*: best fit. *Green dot-dashed line*: contribution of the PSF. *Blue dashed line*: contribution of the clustering. The error bars are too small to be represented. The pixel size is 6".

wavelength	Bias due to clustering		
µm	Method A	Method B	Method C
250	5±2%	7.0±0.9%	7.7±0.5%
350	11±2%	10.4± 0.7%	10.3± 0.8%
500	20±5%	20.6±1.2%	19.1±1.8%

Table 1. Bias due to clustering as a function of the wavelength. These values are estimated with the methods presented in Sect. 4.3.

4.3.4. Correction of the bias due to clustering

Table 1 summarizes our estimates of the bias due to clustering. Our three methods give similar results. To correct for the effects of the clustering, we divide our measured mean flux densities by the mean values of 1.07, 1.10, and 1.20 at 250 µm, 350 µm, and 500 µm, respectively.

4.4. Reconstruction of the SPIRE counts

We can reconstruct the SPIRE counts using the information provided by the $S_{24}+z$ catalog, the mean color, and the standard deviation provided by the stacking analysis. In this analysis, we assume that the distribution of the SPIRE flux density for a given 24 µm flux density is log-normal (see Fig. 4 and Sect. 4.2). For a small scatter (<<1), the standard deviation of the logarithm of the flux density $\sigma_{\log-\rm norm,\,color}$ can be computed from the standard deviation of the flux density $\sigma_{\rm color}$: $\sigma_{\log-\rm norm,\,color} = \sigma_{\rm color}/\ln(10)$. For larger scatter, this

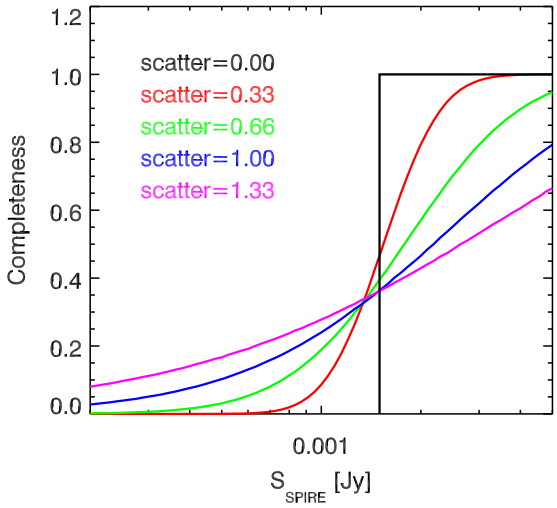


Figure 7. Completeness of the SPIRE counts reconstructed by stacking in a simple case for various values of scatter on the colors. We assume power-law counts ($dN/dS \propto S^{-1.5}$), a mean SPIRE/24 µm ratio, a log-normal scatter with different values and a flux density cut at 24 µm of 30 μJy. This figure is discussed in Sect. 4.4.

approximation is no longer valid. However, there is a bijective link between the following two pairs of parameters: the mean and the scatter of the color in linear units and the same thing in logarithmic units. We can thus deduce the two parameters of the log-normal distribution (mean and scatter) of the color from the linear mean and standard deviation measured by stacking.

We generate 1000 realizations of the SPIRE counts using the following recipe:

- We take randomly a value of the scatter (see Sect. 4.2). At each realization, we used a single value of the scatter for all the flux density and redshift bins.
- In each flux density and redshift bin, we take randomly one value of the $S_{\rm SPIRE}/S_{24}$ color following the uncertainties (see Sect. 4.1). We obtain a relationship between S_{24} and the color in each redshift slice interpolating between the centers of the 24 μm bins.
- We then compute the mean color of each source using the previous relationship.
- For each source, we draw randomly a SPIRE flux density from its 24 μm flux, its color and the scatter on it. We assume a log-normal distribution.
- We then compute the counts from the obtained SPIRE flux densities.

The final counts are computed taking the mean and the standard deviation of the different realizations.

Due to the flux density cut of the 24 µm catalogs, the SPIRE simulated catalogs are not complete at the faint end. If there was a single color for all objects, the cut of the SPIRE catalog would be the SPIRE/24 µm color multiplied by the flux density cut at 24 µm. Above this limit, the catalog would be complete (statistically speaking), and there would be no sources below this limit. However, due to the scatter of the colors, this transition is smoother. We call the ratio between the reconstructed counts (taking into account the 24 µm selection) and the input counts the completeness. Berta et al. (2011) used the Le Borgne et al. (2009) model to estimate the completeness as a function of the far-infrared flux density. We have chosen to use a similar, but more empirical, method to estimate the completeness and correct for it.

- We generate a mock 24 µm catalog following power-law counts with a typical slope in $dN/dS \propto S^{-1.5}$ (Béthermin et al. 2010a).
- We associate a SPIRE flux density with each source of the mock catalog using the real colors and scatters measured by stacking. The color of each source depends on its 24 µm flux density and redshift.
- In each SPIRE flux density bin, we compute the ratio between the total number of sources and the number of sources which are brighter than the 24 µm flux density cut.

Several realizations of the colors and scatters are used to estimate the uncertainties in this correction. Fig. 7 illustrates how the completeness values vary with the scatter of the colors in a simplified case, where we assume a single color for all the sources ($S_{\rm SPIRE}/S_{24} = 50$) and a 24 µm flux density cut of 30 μJy. As expected, without scatter, the transition happens around 1.5 mJy (50 × 0.03), and the width of the transition increases with the scatter. Table B.2 and B.3 provides the completeness corrections used in GOODS-N and COSMOS. We have cut our analysis in COSMOS at 6 mJy in the three SPIRE bands, because the completeness in the higher redshift bins is only ∼50%. Below this limit, we used the GOODS-N field where the 24 μm catalog is deeper. Following the same criterion as in COSMOS, we cut our analysis at 2 mJy in the three bands. These cuts are slightly arbitrary, because the mean $S_{\rm SPIRE}/S_{24}$ color of the sources, and consequently the completeness, vary with redshift. Nevertheless, we use the same cuts for all redshifts in order to simplify the interpretation and the discussion.

The same type of analysis was performed to compute the redshift distribution of $S_{250,\,350,\,\rm or\,500} > 6$ mJy sources in COSMOS. In this case, there is only one flux density bin (> 6 mJy).

5. Estimation of the statistical uncertainties

In Sect. 3 and 4, we explained how we derived number counts and redshift distributions above and below the confusion limit. We also discussed the uncertainties in the corrections applied to our measurements. In this section, we explain how the field-to-field variance on our measurements is estimated and how we combine these uncertainties with the errors on the corrections.

5.1. Sample variance

Our study is based on only one or two fields depending on the flux density regime. The field to field variance cannot thus be easily estimated. We have used the same method based on the clustering of the sources as in Béthermin et al. (2010a), which is briefly described here.

5.1.1. Principle

Spatially, sub-mm sources are not Poisson distributed (Blain et al. 2004; Farrah et al. 2006; Cooray et al. 2010; Magliocchetti et al. 2011), but clustered. The uncertainty, σ_N,

on the number of sources in a given bin, N, is thus not $\sqrt{N}$. In large fields, this effect is not negligible, and the clustering of the sources must be taken into account (Béthermin et al. 2010a). The uncertainties in the clustered case are (Wall & Jenkins 2003)

$$\sigma_N = \sqrt{yN^2 + N}, \tag{4}$$

with

$$y = \frac{\int_{\text{field}} \int_{\text{field}} w(\theta) d\Omega_1 d\Omega_2}{\Omega^2}, \tag{5}$$

where $w(\theta)$ is the auto-correlation function (ACF) and Ω the solid angle of the field. The effect of the clustering on the uncertainties depend only on the field (size and shape) and the ACF.

5.1.2. Estimation of the auto-correlation function

The purpose of this paper is not to study the clustering of sub-mm galaxies, but is just to compute, with a reasonable accuracy, its effect on uncertainties in the number counts. We measured the ACF of the resolved sources for the selection in redshift. This measurement is performed with the Landy & Szalay (1993) estimator:

$$w(\theta) = \frac{DD - 2 \times DR + RR}{RR}, \tag{6}$$

where DD is the number of pairs separated by an angle between $\theta - d\theta/2$ and $\theta + d\theta/2$ in the real catalog, RR the number of pairs in a Poisson distributed catalog generated with the mask used for the source extraction, and DR the number of pairs coming from a source in the real catalog and a source in the random catalog. The method used to quickly compute the number of pairs is described in Appendix A.

We fit our results with the following simple form (Magliocchetti et al. 2011):

$$w(\theta) = A\left(\left(\frac{\theta}{1\ \text{deg}}\right)^{1-\gamma} - C\right), \tag{7}$$

where γ is fixed at the standard value of 1.8. This simple form does not work at small scales (<2′), where the contribution of the clustering between the sources in the same dark matter halo is not negligible (e.g. Cooray et al. 2010). We use only the scales larger than 2′in our analysis. The integral constraint C is a factor taking into account the fact that Landy & Szalay (1993) estimator is biased for finite size survey. C depends on the size and the shape of the field and the value of γ, and can be computed from

$$C = \frac{\int_{\text{field}} \int_{\text{field}} \left(\frac{\theta}{1\ \text{deg}}\right)^{1-\gamma} d\Omega_1 d\Omega_2}{\Omega^2} \tag{8}$$

Combining Eq. 5, 7, and 8, we obtain:

$$y = A \times C. \tag{9}$$

For our masks, $C = 1.72$ in the COSMOS field and 7.16 in GOODS-N. In order to compute the effect of the clustering on our error bars on the number counts, we thus have to estimate A in the various redshift and flux bins used in our analysis.

5.1.3. Uncertainties in the resolved number counts

We measure the clustering of the resolved sources from the source list produced in Sect. 3. If we use only the source in a given flux density and redshift bin, we do not obtain sufficient signal to noise. Therefore, we compute the ACF of all $S_{\text{SPIRE}} > 20$ mJy sources in a single flux density bin, but four redshift bins (0<z<0.5, 0.5<z<1, 1<z<2, and z>2), and assume that the ACF does not evolve too much with flux density. We obtained very good fits in each redshift bins at each wavelengths (reduced $\chi^2 < 1.3$ in all bins). From these fits, we compute the value of the y parameter and the sample variance on our measurements. The uncertainties coming from the sample variance are then combined with the uncertainties coming from the correction applied to the counts (see Sect. 3.4).

Table B.4 summarizes the relative contribution of the clustering term ($\sigma_{\text{clus}} = \sqrt{yN^2}$) to the total sample variance $\left(\sigma_{\text{clus+poi}} = \sqrt{\sigma^2_{\text{clus}} + \sigma^2_{\text{poi}}} = \sqrt{yN^2 + N}\right)$. This contribution is dominant in the low flux density bins (~85%), and decreases in brighter flux density bins, where N is smaller. We also compared the sample variance with the uncertainties in the corrections. This last correction increases the uncertainties by less than 40%. We are thus dominated by sample variance.

5.1.4. Uncertaintes in the number counts measured by stacking

The clustering of the SPIRE sources below the confusion limit (< 20 mJy) measured by stacking (see Sect. 4) cannot be measured directly. In our analysis, we started from the 24 μm population as a prior. We thus use the clustering of this population to compute the effect of the clustering on the uncertainties, assuming it is close to the one of the SPIRE faint sources. The ACF was measured in the same redshift bins as for the resolved sources. We then compute the sample variance, and combine it with the uncertainties coming from the completeness corrections, the colors, and the scatter.

Table B.5 and B.6 summarize the relative contribution of the clustering to the uncertainties. As for the resolved sources, the clustering term dominates the Poisson term in the sample variance. In contrast to resolved counts, the errors coming from the completeness correction and the uncertainties in the colors and the scatter dominate the sample variance. A possible bias, due to the assumption that the 24 μm and sub-confusion limit 250 μm population have similar clustering properties, has therefore only a modest impact to our uncertainty budget.

5.1.5. Uncertainties in the redshift distributions

The ACF is difficult to measure in small redshift bins, because the number of sources is small and the signal-to-noise ratio is then poor. For this reason, we have estimated how the clustering evolves when we reduce the size of a redshift bin. To quantify this effect, we compute the ACF of the 24 μm catalog (the signal for resolved SPIRE sources only is too low) in $1 - dz/2 < z < 1 + dz/2$ bins with dz varying from 0.1 to 1. We find $A \propto dz^{-0.9}$. To compute the uncertainties in the redshift distribution, we thus use the ACF measured previously to compute the sample variance on the counts in large redshift bins ($0 < z < 0.5, 0.5 < z < 1, 1 < z < 2$, and $z > 2$), and

apply the scaling relation $y \propto A \propto dz^{-0.9}$. We then derived the sample variance, and combine it with uncertainties in the correction factor for resolved counts and the ones coming from completeness corrections, colors, and scatter for counts measured by stacking.

6. Validation on simulation

In order to check the accuracy of our methods used to measure the number counts, we have performed an end-to-end simulation. The clustering of the sources below the confusion limit is not well known, and its effect on stacking has been estimated in Sect 4.3 with three methods based on the data. We have thus chosen to use a simulation with a Poisson distribution of the sources, because it is easier to generate.

6.1. Description of the simulation

Our simulation is based on Béthermin et al. (2011) model, which is a parametric model based on the Lagache et al. (2004) spectral energy distribution (SED) library (two populations: normal and starburst galaxies). This model uses a simple broken power-law evolutionary behavior of the characteristic luminosity and density of the luminosity function (LF). The free parameters of the model were determined by fitting observed counts (including the Herschel resolved counts published by Oliver et al. (2010b)), LFs, and the CIB. The model has not been modified since the publication of the associated paper. Note that this model includes the contribution of strongly-lensed sources to the counts.

A mock catalog, containing the 24 μm, 250 μm, 350 μm, and 500 μm flux densities and the redshift of the sources, was generated following the model. We then build a map of the COSMOS field from this mock catalog, the SPIRE noise map, and the SPIRE PSF. We then redo all the analysis described Sect. 3 and 4 using the $S_{24} + z$ mock catalog and the simulated SPIRE maps. The SEDs of this model were not calibrated following the correlation between stellar mass and the star formation rate (roughly proportional to the infrared luminosity) and are thus not valid below 8 μm rest-frame. We thus cannot use this simulation at redshifts larger than 2.

6.2. Results

The Fig. 8 shows the results of this simulation. The recovered counts (triangles and diamonds) nicely reproduce the shape of the counts. The flux density regime probed by stacking is well reproduced (reduced $\chi^2 = 1.4$ for 81 degrees of freedom). Paradoxically, the resolved counts are not as well reproduced, with some points deviating at more than 3-σ. At bright flux densities, our recovered counts are systematically lower than our results. It could be due to the fact that the extraction technique shares the flux of a bright source between several prior positions.

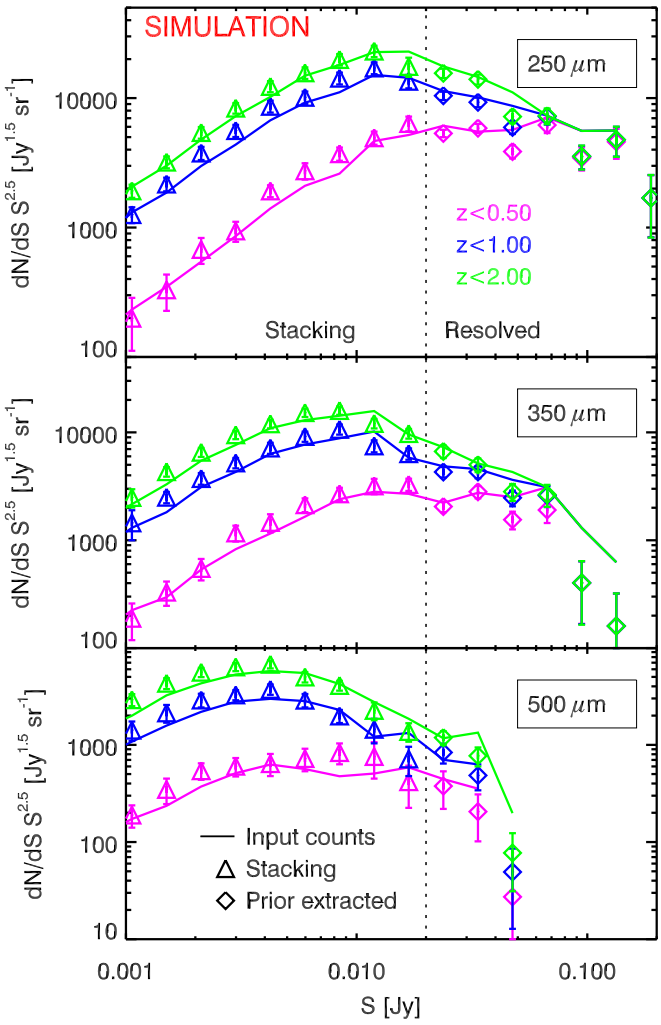


Figure 8. Validation of our method of measurement of the counts at 250 μm (top), 350 μm (center), and 500 μm (bottom) from a simulation based on Béthermin et al. (2011) model. *Solid lines:* input counts from the simulated catalog for various cuts in redshift. *Diamonds*: resolved number counts measured using the same method as for the real data. *Triangles*: number counts measured by stacking using the same method as for the real data.

7. Number counts

7.1. Results

From the extraction with priors presented in Sect. 3, we build number counts per redshift slice down to 20 mJy at all three SPIRE wavelengths. Thanks to the stacking of the 24 µm sources in the COSMOS and GOODS-N fields, we reach 6 mJy and 2 mJy, respectively. We checked that the counts deduced from stacking analysis are in agreement with resolved counts above 20 mJy, but they have larger uncertainties than the resolved ones. The COSMOS and GOODS-N counts deduced by stacking analysis are also in agreement where they overlap, but with much smaller uncertainties in COSMOS due to the size of this field. We thus use GOODS-N points only at faint flux densities, which the COSMOS data does not constrain. Fig. 9 and 10 show our results. The points obtained by stacking in the GOODS-N and COSMOS fields disagree at 2σ in the $1 < z < 2$ bin at all SPIRE wavelengths (see Fig. 10), which could be due

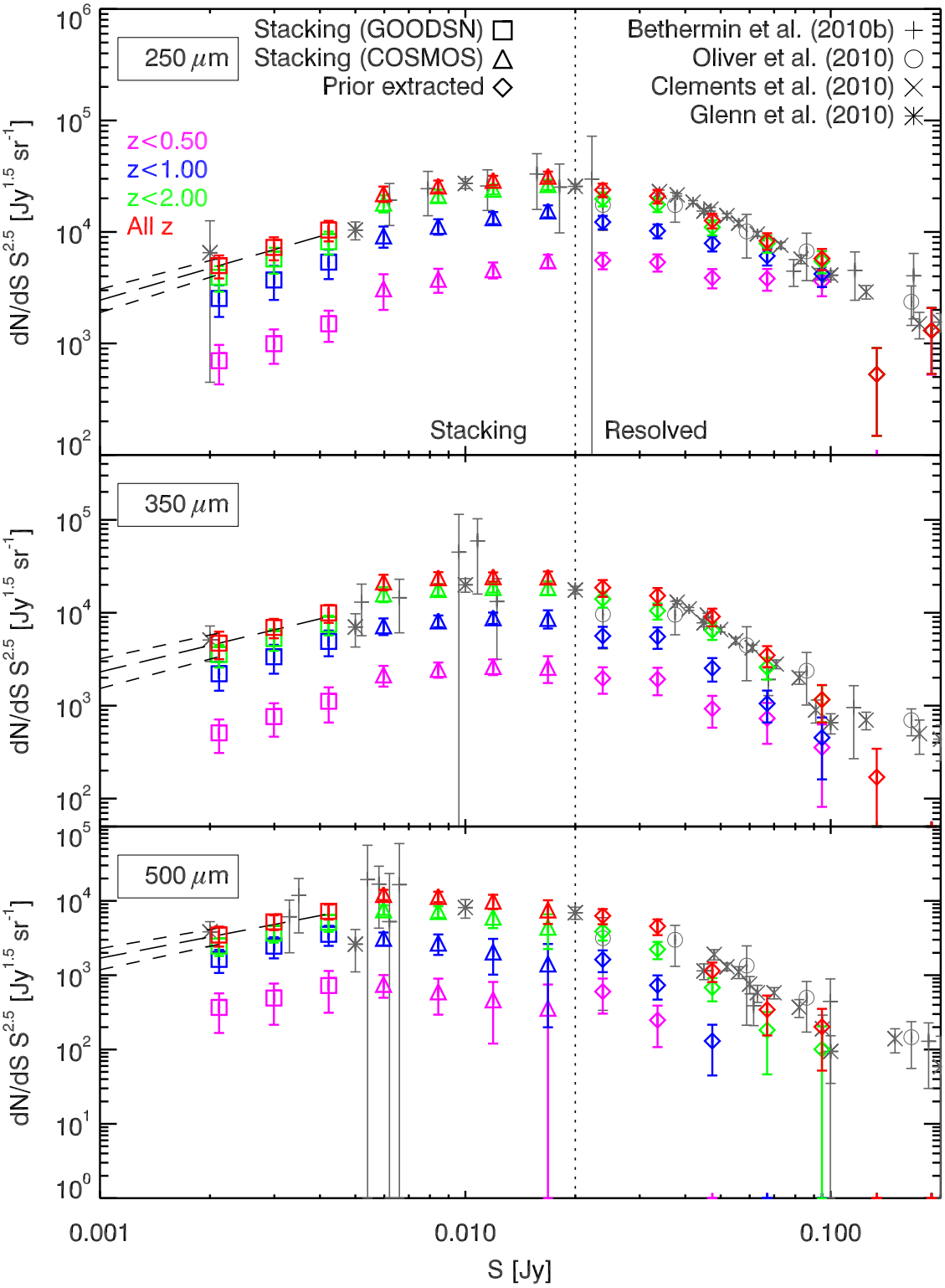


Figure 9. Number counts at 250 µm (top panel), 350 µm (middle panel), and 500 µm (bottom panel). The contribution of $z < 0.5$, $z < 1$, $z < 2$, and all sources are plotted in violet, blue, green and red, respectively. *Squares*: points from stacking in GOODS-N. *Triangles*: points from stacking in COSMOS. *Diamonds*: points measured from source extraction using priors. *Vertical dotted line*: 4-σ confusion limit. *Long and short dashed lines*: extrapolation of the counts and 1-σ confidence region (see Sect. 9.2.1). *Plus symbols*: Béthermin et al. (2010b) measurements using BLAST data. *Circles*: Oliver et al. (2010b) measurements from resolved sources in the HerMES science demonstration phase data (*Herschel*/SPIRE). *Crosses*: Clements et al. (2010) measurements from resolved sources in the SPIRE H-ATLAS science demonstration phase data (*Herschel*/SPIRE). *Asterisks*: Glenn et al. (2010) measurements from $P(D)$ analysis of the HerMES science demonstration phase data (*Herschel*/SPIRE).

Flux mJy	Normalized counts ($dN/dS \times S^{2.5}$) $Jy^{1.5}sr^{-1}$				
	All	$0 < z < 0.5$	$0.5 < z < 1$	$1 < z < 2$	$z > 2$
			Stacking (GOODS-N)		
2.1	4989 ± 824	699 ± 220	1828 ± 553	1392 ± 445	1068 ± 356
3.0	7261 ± 1199	994 ± 250	2714 ± 941	2013 ± 542	1538 ± 442
4.2	10405 ± 1337	1499 ± 307	3856 ± 1035	2812 ± 598	2237 ± 512
			Stacking (COSMOS)		
6.0	22037 ± 3228	3086 ± 1082	6116 ± 1575	8918 ± 2273	3915 ± 1265
8.4	25787 ± 2704	3759 ± 906	7463 ± 1389	9969 ± 1851	4594 ± 1064
11.9	29044 ± 2168	4574 ± 728	8834 ± 1284	10727 ± 1345	4907 ± 843
16.8	31574 ± 2527	5503 ± 723	9899 ± 1516	11185 ± 1579	4985 ± 1035
			Resolved (COSMOS)		
23.8	23851 ± 1595	5558 ± 733	6694 ± 760	7516 ± 1076	4082 ± 519
33.6	20926 ± 1654	5349 ± 804	4875 ± 690	7379 ± 1151	3323 ± 536
47.4	12653 ± 1345	3878 ± 762	4036 ± 752	3114 ± 688	1623 ± 433
67.0	8319 ± 1337	3813 ± 928	2281 ± 681	1777 ± 613	447 ± 289
94.6	5780 ± 1431	3606 ± 1142	596 ± 432	1279 ± 681	298 ± 304
133.7	528 ± 532	-	528 ± 532	-	-
188.8	1306 ± 1084	1306 ± 1084	-	-	-

Table 2. Number counts at 250 μm. The errors take into account the statistical uncertainties, including the clustering effect, and the uncertainties in the completeness corrections. For the points measured by stacking, we also take into account the uncertainties in the colors and the scatter. The uncertainties in the SPIRE absolute calibration are neglected here.

Flux mJy	Normalized counts ($dN/dS \times S^{2.5}$) $Jy^{1.5}sr^{-1}$				
	All	$0 < z < 0.5$	$0.5 < z < 1$	$1 < z < 2$	$z > 2$
			Stacking (GOODS-N)		
2.1	4709 ± 1342	510 ± 159	1695 ± 572	1341 ± 450	1162 ± 1117
3.0	6949 ± 1167	761 ± 234	2598 ± 870	1912 ± 481	1676 ± 564
4.2	9964 ± 1396	1115 ± 363	3802 ± 1016	2587 ± 520	2459 ± 715
			Stacking (COSMOS)		
6.0	21510 ± 3858	2143 ± 525	5083 ± 1309	8691 ± 2354	5592 ± 2711
8.4	23820 ± 3174	2458 ± 431	5715 ± 996	9627 ± 1997	6020 ± 2216
11.9	24402 ± 2274	2638 ± 462	6175 ± 1001	9875 ± 1445	5713 ± 1366
16.8	24229 ± 3158	2579 ± 790	6115 ± 1648	9953 ± 2208	5581 ± 1325
			Resolved (COSMOS)		
23.8	18652 ± 1605	1967 ± 434	3660 ± 643	8442 ± 1215	4581 ± 703
33.6	15285 ± 1448	1927 ± 484	3600 ± 708	4995 ± 838	4760 ± 811
47.4	9092 ± 1187	927 ± 346	1606 ± 467	3929 ± 828	2628 ± 620
67.0	3487 ± 828	728 ± 390	327 ± 234	1527 ± 553	904 ± 416
94.6	1163 ± 630	354 ± 351	98 ± 160	710 ± 497	-
133.7	170 ± 273	-	-	-	170 ± 273

Table 3. Number counts at 350 μm.

to field-to-field variance.

The depth and small error bars of our counts enable us to detect with high significance the peak of the Euclidian normalized counts near 15, 10 and 5 mJy at 250 μm, 350 μm, and 500 μm, respectively. This maximum was seen at 250 μm and 350 μm by Glenn et al. (2010). With our new results, we are able to detect this maximum at 500 μm. We also start to see a power-law behavior below this peak, which was seen previously only up to 160 μm (Papovich et al. 2004; Béthermin et al. 2010a; Berta et al. 2011). Nevertheless, the significance of this detection is hard to estimate because of the correlation between the points obtained by stacking.

7.2. Comparison with the previous measurements

We have compared our total counts with previous measurements (cf. Fig. 9). At high flux densities (S>20 mJy), our counts agree with the counts measured from resolved sources of Béthermin et al. (2010b) in BLAST, Oliver et al. (2010b) in SPIRE/HerMES SDP fields, and Clements et al. (2010) in the SPIRE/H-ATLAS SDP field. Our measurements are also in agreement with the stacking analysis of Béthermin et al. (2010b) of the BLAST data. Our new stacking analysis of the SPIRE

Flux mJy	Normalized counts ($dN/dS \times S^{2.5}$) $Jy^{1.5}sr^{-1}$				
	All	$0 < z < 0.5$	$0.5 < z < 1$	$1 < z < 2$	$z > 2$
Stacking (GOODS-N)					
2.1	3465 ± 864	368 ± 178	1262 ± 394	836 ± 217	998 ± 716
3.0	5216 ± 783	493 ± 244	1966 ± 500	1220 ± 351	1535 ± 424
4.2	7244 ± 1089	726 ± 360	2823 ± 631	1592 ± 529	2102 ± 615
Stacking (COSMOS)					
6.0	12170 ± 1764	750 ± 263	2381 ± 562	4444 ± 817	4594 ± 1435
8.4	11446 ± 1716	596 ± 322	2107 ± 783	4481 ± 872	4260 ± 1210
11.9	9917 ± 2089	465 ± 380	1586 ± 1000	3976 ± 1363	3888 ± 1167
16.8	7540 ± 2665	358 ± 443	1055 ± 1192	3003 ± 1817	3122 ± 1478
Resolved (COSMOS)					
23.8	6298 ± 675	602 ± 220	1023 ± 278	2258 ± 343	2413 ± 460
33.6	4548 ± 656	248 ± 138	483 ± 191	1493 ± 329	2322 ± 516
47.4	1143 ± 343	-	130 ± 117	549 ± 238	463 ± 218
67.0	343 ± 251	-	-	182 ± 185	160 ± 169
94.6	202 ± 230	-	-	100 ± 162	101 ± 163

Table 4. Number counts at 500 μm.

data reduces the uncertainties by about a factor 5 compared for the BLAST data. Finally, we agree with the $P(D)$ analysis of Glenn et al. (2010), except for the 6 mJy points at 250 μm and 500 μm which disagrees by about 2 σ with our measurements. Due to the number of points compared (21), we expect to have about 2 points with 2 σ difference, so this is not significant.

The good agreement between the counts produced by the stacking and the $P(D)$ analysis confirms the accuracy of these two statistical methods. It also suggests that the galaxies seen in the mid-IR are a good tracer of the sources responsible for the sub-mm counts, and justifies a posteriori our choice to use the 24 μm sources as a prior. The mid-IR faint and far-IR bright population thus constitute a small contribution to the number counts.

7.3. Comparison with the models

In Fig. 10, we compare our results with a set of recent (≥2009) evolutionary models:

- The Béthermin et al. (2011) model was presented in Sect. 6.1.
- The Marsden et al. (2011) model is also a parametric model similar to the Béthermin et al. (2011) one, but using a different SED library, and taking into account the scatter in the temperature of the cold dust in the different galaxies.
- Le Borgne et al. (2009) carried out a non-parametric inversion of the counts assuming a single population (Chary & Elbaz 2001) to determine the evolution of the luminosity function with redshift.
- The Valiante et al. (2009) model used a large library of starburst and AGNs templates. This model takes into account the scatter in the temperature of the sources. The parameters of the model were tuned manually.
- The Gruppioni et al. (2011) model uses 5 separately evolving populations, including 3 populations of AGN.
- The Rahmati & van der Werf (2011) model is based on a modified Dale & Helou (2002) library. This model takes into account the scatter in the temperature of the sources. It was fit to the 850 μm counts and redshift distribution.

Note that the Béthermin et al. (2011), Gruppioni et al. (2011) and Rahmati & van der Werf (2011) models were already tuned using recent *Herschel* data, including the GOODS-N observations used here.

None of these models manages to fully reproduce our measurements. The Béthermin et al. (2011), Gruppioni et al. (2011) and Rahmati & van der Werf (2011) models are close to the data, and broadly reproduce the shape of the counts, but still deviate from the measurements by 3-σ. The Le Borgne et al. (2009) and Valiante et al. (2009) models underestimate the contribution of $z < 1$ sources to the counts. The Marsden et al. (2011) model overestimates the counts at high z ($z > 1$). Not surprisingly, models which use the most recent *Herschel* data and the redshift-dependent observables (redshift distributions, luminosity functions, etc.) provide the best match to our findings.

8. Redshift distributions

8.1. Results

From the brighter sources extracted using the 24 μm prior (see Sect. 3), we have built the redshift distribution of the sources brighter than 20 mJy at 250 μm, 350 μm, and 500 μm (see Fig. 11 and Table 5). We find that the distribution of the resolved 250 μm sources is almost flat up to $z \sim 1$ and decreases significantly at higher redshift. At 350 μm, the distribution peaks near $z \sim 1$, and the distribution is flatter at high redshift. At 500 μm, the contribution of $z < 1.5$ sources is smaller than at shorter wavelengths. Between $z = 1.5$ and $z = 3$, our measurements are compatible with a flat distribution, however the uncertainties are very large.

At 250 μm and 350 μm, we clearly see an excess in the $0.2 < z < 0.4$ and $1.8 < z < 2.0$ bins. The structure at $z = 0.3$ in COSMOS is well known (Scoville et al. 2007). The excess near $z = 1.9$ could also be explained by a large-scale structure. Fig. 12 shows the position of the sources in a thin redshift slice between $z = 1.85$ and $z = 1.9$. The sources are strongly

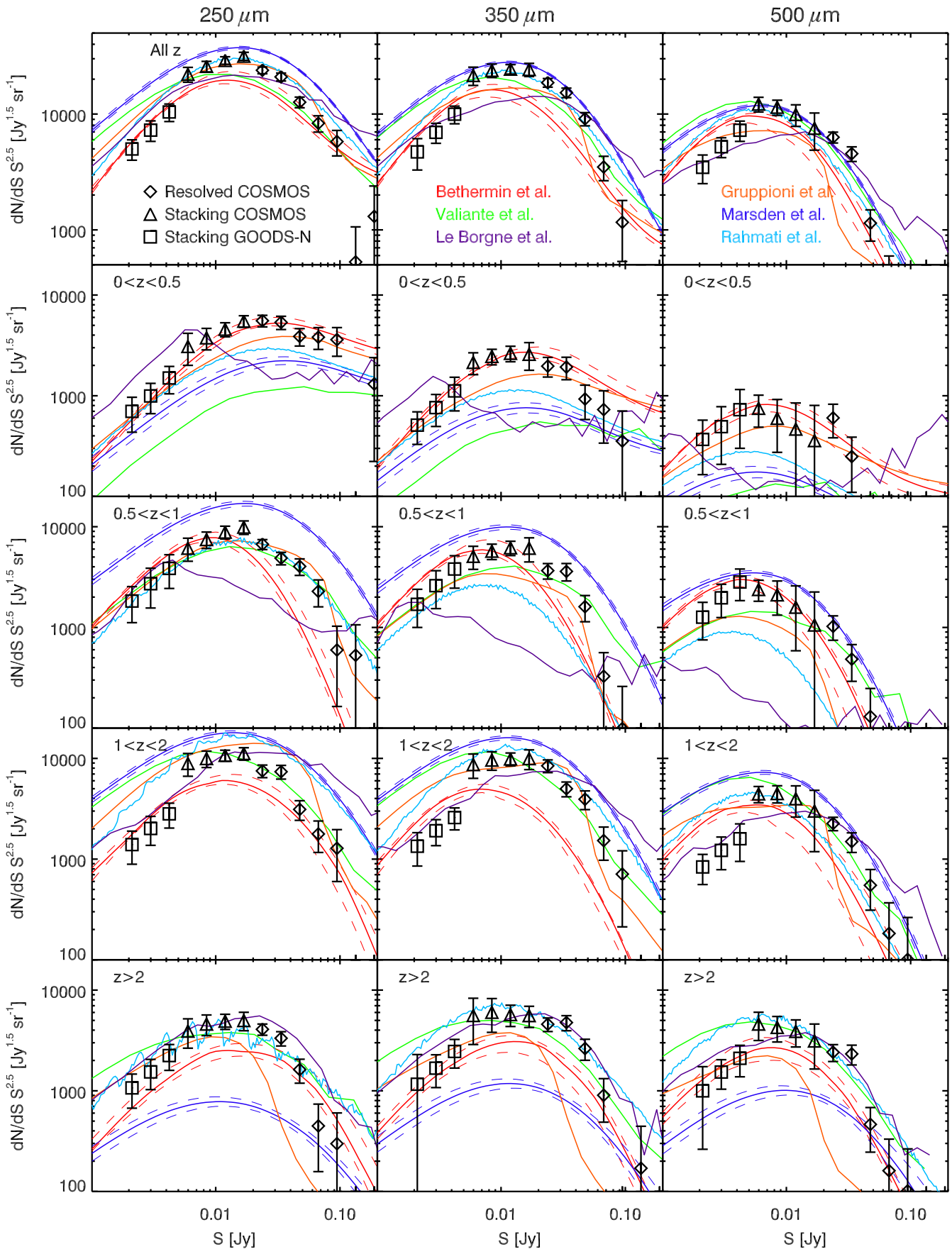


Figure 10. Comparison between the observed number counts and the models at 250 μm (left), 350 μm (middle), and 500 μm (right), for various redshift selections: all redshifts (top), $0 < z < 0.5$, $0.5 < z < 1$, $1 < z < 2$, and $z > 2$ (bottom). *Squares*: points from stacking in GOODS-N. *Triangles*: points from stacking in COSMOS. *Diamonds*: points measured from source extraction with priors. We have overplotted the models from Béthermin et al. (2011) in red, Valiante et al. (2009) in green, Le Borgne et al. (2009) in violet, Gruppioni et al. (2011) in orange, Rahmati & van der Werf (2011) in light blue and Marsden et al. (2011) in dark blue.

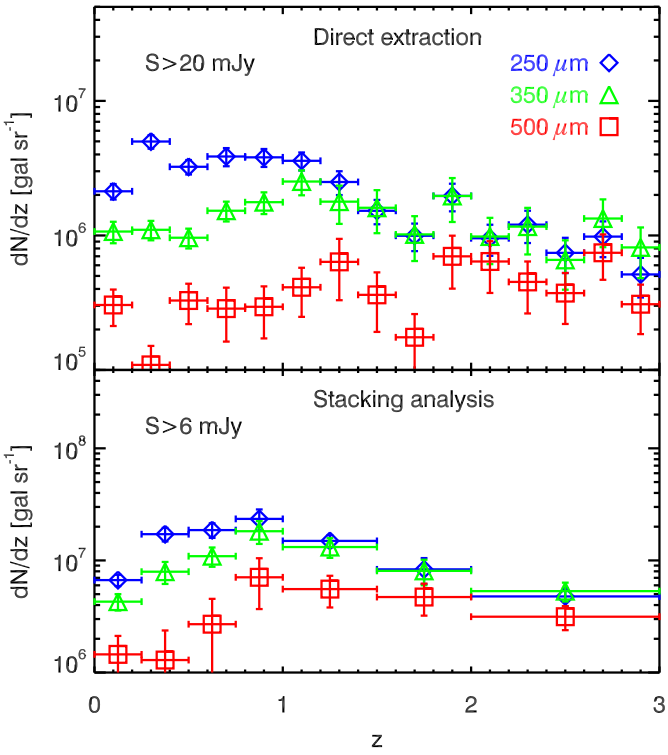


Figure 11. Redshift distribution of S_{SPIRE} > 20 mJy (upper panel) and S_{SPIRE} > 6 mJy (lower panel) sources at 250 µm (blue), 350 µm (green), and 500 µm (red) in the COSMOS field.

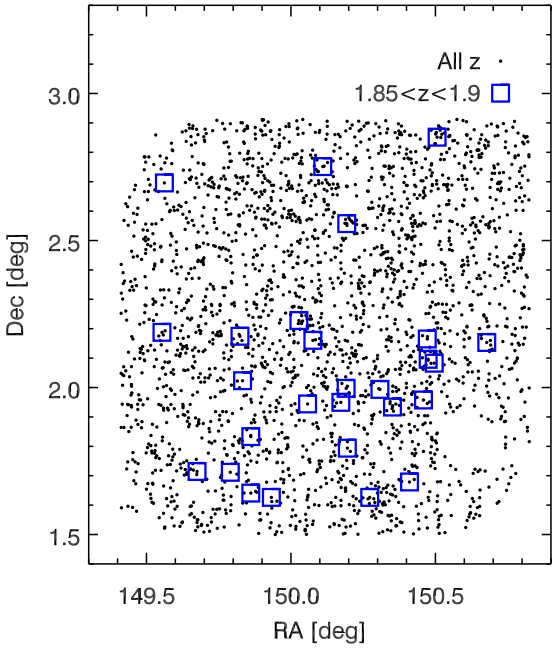


Figure 12. Spatial distribution of S_{250} >20 mJy sources in COSMOS. *Black dots*: all redshifts. *Blue boxes*: only sources in the 1.85< z <1.9 range.

concentrated in a 0.7°×0.7° region, corresponding to a physical size of about 20 Mpc. It could be linked with the three candidate clusters of galaxies at $z \sim 1.8$ found by Chiaberge et al. (2010) in the same field. Nevertheless, this overdensity could also be an artifact of the photometric redshifts. An effect of the polycyclic aromatic hydrocarbon (PAH) features redshifting into the 24 µm band is possible althought less likely, because this should affect neighboring redshift bins, due to the band width of *Spitzer* at 24 µm ($\lambda/\Delta\lambda \sim 3$).

We also used the stacking analysis presented in Sect. 4.4 to estimate the redshift distribution of the S_{SPIRE} >6 mJy sources in COSMOS. We find a smaller relative contribution of $z < 1$ sources than for the 20 mJy flux density cut at 250 μm and 350 μm. The behavior at $z > 1$ is similar to that found for the 20 mJy flux density cut.

8.2. Comparison with other measurements

Chapin et al. (2011) studied the redshift distribution of isolated BLAST sources. Their redshift distributions cannot be normalized by the surface area (because of the isolation ctriterion), and the flux density cuts are different; nevertheless, the trends of their distributions and their evolution from 250 µm to 500 µm agrees with our findings.

Amblard et al. (2010) also produced a redshift distribution of $S_{350} > 35$ mJy sources in H-ATLAS from a *Herschel* color-color diagram and using assumptions about the FIR/sub-mm SED of the sources. They found a strong peak at $z = 2$, in complete disagreement with our distribution. This could be due to the fact that they required a 3 σ detection at 250 µm and 500 µm, which correspond to a $S_{250} > 21$ mJy and $S_{500} > 27$ mJy. The 3-σ criterion at 500 µm tends to select high-redshift sources, because of the shape of the SEDs, the flux of low-z sources decreases rapidly between 250 μm and 500 μm. The method is also strongly dependent on the dust temperatures of the sources assumed in their analysis, due to the degeneracy between dust temperature and redshift for thermal sources.

8.3. Comparison with the models

We compared the measured redshift distributions with the predictions of the same models as in Sect. 7.3 (Fig. 13). Again no model manages to reproduce accurately the redshift distributions of the bright resolved sources ($S > 20$ mJy). Note however that Gruppioni et al. (2011) model reasonably fits the data at 250 μm and 350 μm at $z < 2.5$. All the models without strong lensing predict a strong break in the redshift distributions at z~2.5, which is not present in the data. The Béthermin et al. (2011) model, which includes strong lensing, predicts a more consistent slope, although the normalization is not correct. It could be interpreted as a clue that the high redshift tail is due to lensed galaxies (see e.g. Vieira et al. 2010; Negrello et al. 2010), but the contribution of lensed galaxies in the Béthermin et al. (2011) model is negligible for a flux density cut of 20 mJy[7].The redshift distribution of the faint sources ($S > 6$ mJy) are globally better modeled, a broad agreement being found with the Béthermin et al. (2011) and Gruppioni et al. (2011) models, which are fitted using the most recent data. The strong disagreement with the models at bright

[7] Note however that the lensed objects dominates the redshift distribution at $z > 2$ for a flux density cut of 100 mJy.

redshift range	dN/dz (in $10^4\times$ gal.sr^{-1})		
	$S_{250} > 20$ mJy	$S_{350} > 20$ mJy	$S_{500} > 20$ mJy
$0.0 < z < 0.2$	212±28	106±19	30±9
$0.2 < z < 0.4$	498±57	110±18	10±4
$0.4 < z < 0.6$	323±39	96±16	32±10
$0.6 < z < 0.8$	386±59	152±25	28±12
$0.8 < z < 1.0$	380±58	176±32	29±12
$1.0 < z < 1.2$	358±54	251±52	41±16
$1.2 < z < 1.4$	249±50	178±56	63±30
$1.4 < z < 1.6$	152±31	160±56	36±16
$1.6 < z < 1.8$	99±23	101±37	17±8
$1.8 < z < 2.0$	196±45	196±70	69±29
$2.0 < z < 2.2$	95±24	98±37	63±26
$2.2 < z < 2.4$	120±32	116±44	45±18
$2.4 < z < 2.6$	74±21	65±26	37±15
$2.6 < z < 2.8$	98±29	133±51	74±27
$2.8 < z < 3.0$	51±16	81±33	30±12
	$S_{250} > 6$ mJy	$S_{350} > 6$ mJy	$S_{500} > 6$ mJy
$0.0 < z < 0.2$	667±76	428±71	145±66
$0.2 < z < 0.5$	1713±236	793±176	129±107
$0.5 < z < 0.8$	1865±277	1090±213	269±184
$0.8 < z < 1.0$	2346±500	1821±417	707±339
$1.0 < z < 1.5$	1493±183	1316±262	555±174
$1.5 < z < 2.0$	832±216	807±206	471±149
$2.0 < z < 3.0$	477±88	531±101	314±76
$z > 3.0$	11±1	11±1	9±1

Table 5. Redshift distribution of the SPIRE sources in COSMOS for various flux density cuts at the three SPIRE wavelengths.

flux densities suggests that the bright end of the luminosity function and/or the SEDs of the brightest objects are not well modeled by the current studies. Our measurements therefore provide significant new constraints for such models.

9. Cosmic infrared background

9.1. Contribution of the 24 µm-selected sources to the CIB

The differential contribution of the 24 µm-selected sources to the CIB at longer wavelengths as a function of redshift is a relatively unbiased measurement and places tight constraints on evolution models. Measurements were performed at 70 and 160 µm in *Spitzer* data by Jauzac et al. (2011), and at 250 µm, 350 µm, and 500 µm by Viera et al. (in prep.). The latter were performed in GOODS-N, on a small area. We performed the same analysis in COSMOS, obtaining smaller uncertainties and a better resolution in redshift. To compute this overall observable, we have estimated the total surface brightness due to the $S_{24} > 80\,\mu$Jy sources in redshift slices. The results of this stacking analysis is shown in Fig. 14 and given in Table 6.

As in Vieira et al. (in prep.), we find a peak near $z = 1$. The relative contribution of the $z < 1$ sources decreases with wavelength, and the contribution of $z > 1$ sources increases. We observed 2 peaks at $z \sim 0.3$ and $z \sim 1.9$, probably associated with the overdensities discussed in Sect. 8.1. We compared our results with the predictions of the three models which can take into account the 24 µm selection among the six previously-compared ones. The Béthermin et al. (2011) model broadly reproduces our measurements. Nevertheless,

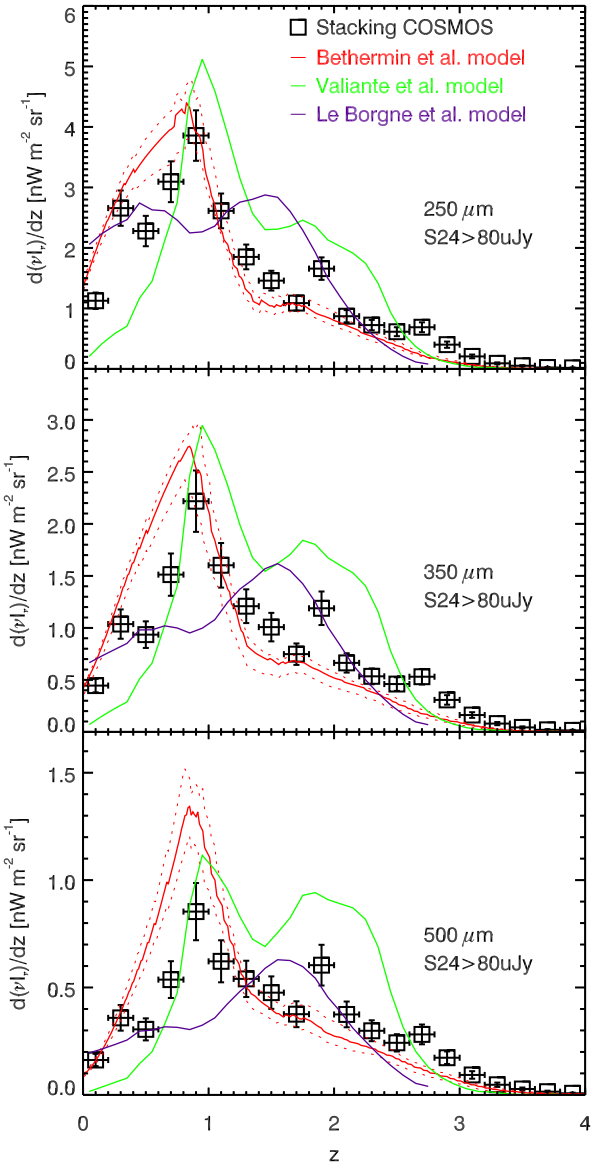


Figure 14. Contribution of $S_{24} > 80\,\mu$Jy sources to the CIB as a function of redshift at 250 µm (top), 350 µm (center), and 500 µm (bottom). We overplotted the predictions of the Béthermin et al. (2011) (red), Valiante et al. (2009) (green), and Le Borgne et al. (2009) (violet) models.

it overpredicts by 3-σ the observed values below $z = 1$ at 500 µm . The Valiante et al. (2009) model predicts a large bump near $z = 2$, which is not seen. A smaller bump is predicted by the Le Borgne et al. (2009) model. However, this model tends to underestimate the contribution of $z < 1$ sources and overestimate the contribution of $z > 1$ ones. The large bumps in the CIB contribution predicted by the Valiante et al. (2009) and Le Borgne et al. (2009) models around $z \sim 2$, caused by PAH features, are not seen in our data, although there is a single elevated point at $z \sim 1.8$ in all bands. We are unsure as to the cause of this observed feature, but note that, given the width of the MIPS 24 μm filter, we would expect any significant PAH contribution to affect multiple redshift bins instead of a single point.

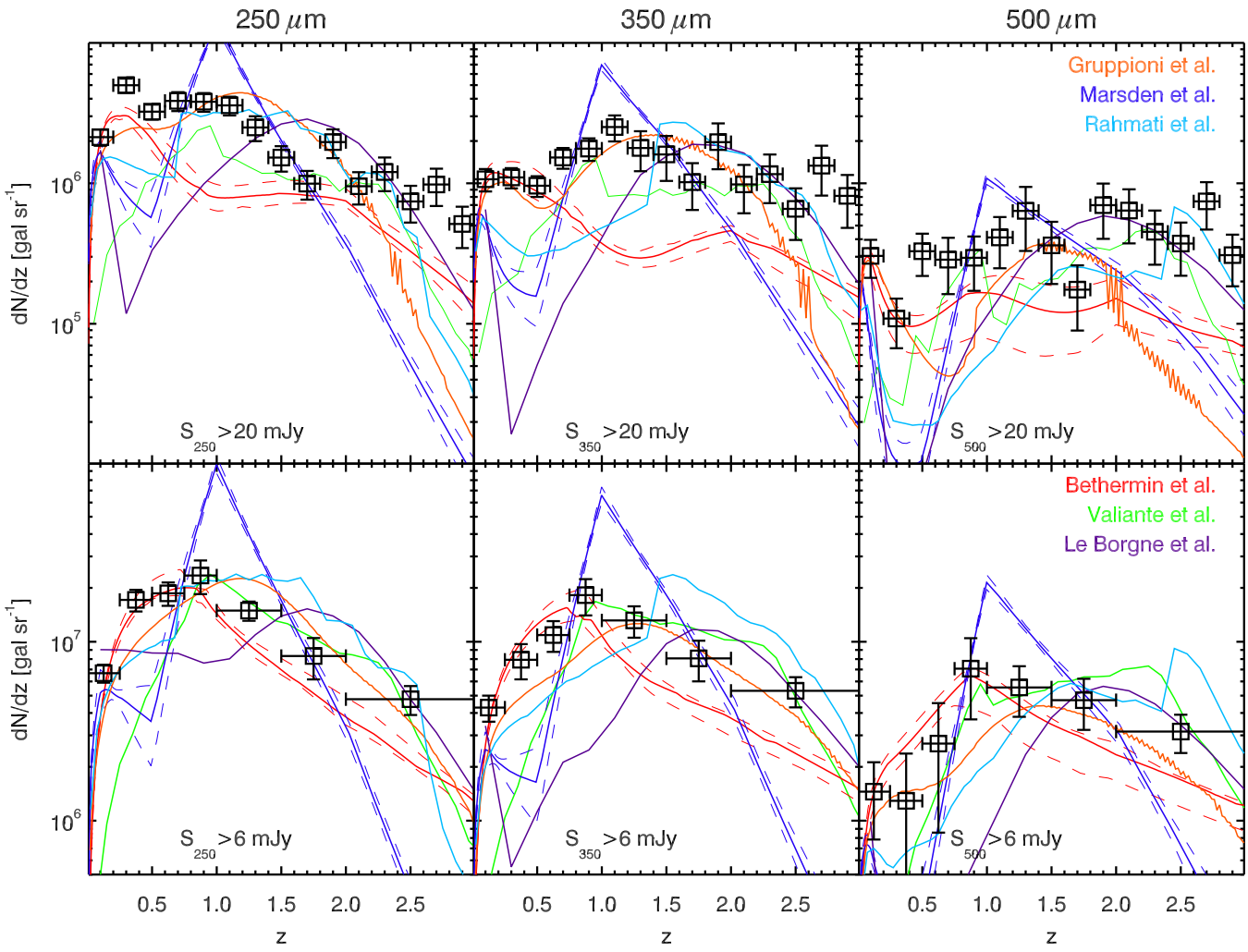


Figure 13. Redshift distribution of the $S_{\rm SPIRE} > 20$ mJy (upper panels) and $S_{SPIRE} > 6$ mJy (lower panels) sources at 250 µm (left), 350 µm (center), and 500 µm (right). We overplot the models of Béthermin et al. (2011) in red, Valiante et al. (2009) in green, Le Borgne et al. (2009) in violet, Gruppioni et al. (2011) in orange, Rahmati & van der Werf (2011) in light blue and Marsden et al. (2011) in dark blue.

9.2. Properties of the CIB

The contribution of the 24 µm sources to the CIB is an interesting quantity for models. Nevertheless, we want to have constraints on the total contribution of the galaxies to the CIB, even if the uncertainties are larger. These constraints can be derived by integrating and extrapolating our new SPIRE number counts.

9.2.1. Estimate of the contribution to the CIB from galaxies

We integrated our counts for different cuts in flux density density, assuming the data points are connected by power-laws. The contribution to the CIB of the sources brighter than the brightest constrained flux bin is less than 2% (Béthermin et al. 2010b), and is neglected. We estimated our error bars using a Monte Carlo method. We used the distribution of recovered values of the CIB to compute the confidence interval. We adopted this method down to faintest flux density probed by stacking. In order to take into account cosmic variance, we combined the statistical uncertainties with the 15% level of the large scale fluctuations measured by Planck collaboration XVIII (2011).

We also extrapolated the contribution of the sources fainter than the limit of our counts. The typical faint-end slope of the infrared counts[8] lies in a range between -1.45 and -1.65 (Papovich et al. 2004; Béthermin et al. 2010a; Berta et al. 2011). This is also the case for our input redshift catalog, even if we select only a redshift slice. We thus assumed a slope of -1.55±0.10 to estimate the contribution of the flux density fainter than the limit of the stacking analysis. The errors are estimated using a MC process, which takes into account the uncertainties in the faint-end slope. By integrating our number counts extrapolated down to zero flux density, we find a total contribution of the galaxies to the CIB of $10.13^{+2.60}_{-2.33}$ nW m^{-2} sr^{-1}, $6.46^{+1.74}_{-1.57}$ nW m^{-2} sr^{-1}, and $2.80^{+0.93}_{-0.81}$ nW m^{-2} sr^{-1} at 250 µm, 350 µm, and 500 µm, respectively. These values agree at better than 1 σ with the FIRAS absolute measurements performed by

[8] The slope α of the counts is defined by $dN/dS \propto S^{\alpha}$.

redshift range	$d(\nu I_\nu)/dz$ (in nW m^{-2} sr^{-1})		
	250 μm	350 μm	500 μm
$0.0 < z < 0.2$	1.127±0.137	0.446±0.057	0.163±0.024
$0.2 < z < 0.4$	2.657±0.295	1.037±0.123	0.359±0.048
$0.4 < z < 0.6$	2.279±0.254	0.935±0.111	0.305±0.042
$0.6 < z < 0.8$	3.094±0.339	1.512±0.175	0.537±0.068
$0.8 < z < 1.0$	3.857±0.421	2.218±0.253	0.853±0.104
$1.0 < z < 1.2$	2.612±0.288	1.602±0.185	0.622±0.077
$1.2 < z < 1.4$	1.851±0.206	1.208±0.140	0.540±0.066
$1.4 < z < 1.6$	1.459±0.165	1.008±0.118	0.476±0.060
$1.6 < z < 1.8$	1.088±0.125	0.748±0.090	0.376±0.048
$1.8 < z < 2.0$	1.658±0.187	1.189±0.139	0.604±0.074
$2.0 < z < 2.2$	0.871±0.102	0.664±0.080	0.375±0.048
$2.2 < z < 2.4$	0.725±0.086	0.534±0.066	0.299±0.039
$2.4 < z < 2.6$	0.615±0.075	0.460±0.058	0.242±0.033
$2.6 < z < 2.8$	0.689±0.082	0.531±0.066	0.282±0.037
$2.8 < z < 3.0$	0.401±0.051	0.308±0.041	0.173±0.024
$3.0 < z < 3.2$	0.207±0.031	0.162±0.024	0.093±0.015
$3.2 < z < 3.4$	0.092±0.016	0.080±0.014	0.047±0.009
$3.4 < z < 3.6$	0.052±0.011	0.043±0.009	0.028±0.006
$3.6 < z < 3.8$	0.023±0.007	0.020±0.006	0.015±0.005
$3.8 < z < 4.0$	0.018±0.006	0.015±0.005	0.007±0.003

Table 6. Differential contribution of $S_{24} > 80\mu$Jy sources to the CIB as a function of redshift.

Fixsen et al. (1998) and Lagache et al. (2000).

We estimated the fraction of the CIB resolved into individual sources ($S > 20$ mJy) using our estimation of the total CIB coming from our extrapolation of the number counts down to zero flux density. We found 15%, 11% and 5% at 250 μm, 350 μm, and 500 μm, respectively. When we go down to 2 mJy (the limit of the stacking analysis), we resolve 73%, 69%, and 55% of the CIB, respectively.

Fig. 15 shows the cumulative contribution to the CIB as a function of the flux density cut. We have compared our results with the fraction resolved by previous shallower analyses (Béthermin et al. 2010b; Oliver et al. 2010b), and find a 1σ agreement.

9.2.2. CIB build-up as a function of redshift

From our cumulative number counts as a function of redshift (see Sect. 7.1), we can extrapolate the CIB emitted below a given redshift, following the methods presented in Sect. 9.2.1. The results are presented in Fig. 16 and Table 8. The redshift at which half of the CIB is emitted is 1.04, 1.20, and 1.25, at 250 μm, 350 μm, and 500 μm, respectively. For comparison, Le Floc'h et al. (2009) measured value $z = 1.08$ at 24 μm. Berta et al. (2011) performed the same type of measurement, but considering only the resolved CIB, and found $z = 0.58$, $z = 0.67$, and $z = 0.73$, at 70 μm, 100 μm, and 160 μm, respectively. As expected, the CIB at longer wavelengths is emitted at higher redshift. The predictions of different models are also shown. The Marsden et al. (2011) and Valiante et al. (2009) models strongly overpredict the contribution of $z > 1$ sources. The Le Borgne et al. (2009) and Rahmati & van der Werf (2011) models slightly underpredict the contribution of $z < 2$ sources at 350 μm and 500 μm. The Gruppioni et al. (2011) model agrees at 1σ with the measurements, except a 1.5σ underprediction at z∼1 at 500 μm. The Béthermin et al. (2011) models agrees at 1σ with this measurement. Note, however, that it underestimates by 1σ the contribution of $z < 2$ sources to the CIB at 250 μm and 350 μm.

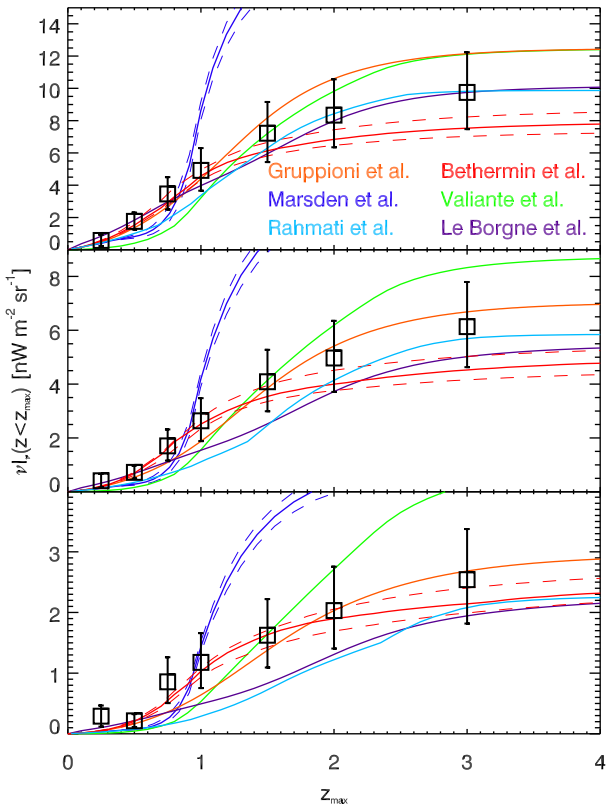


Figure 16. Cumulative contribution to the CIB as a function of redshift at 250 μm (top), 350 μm (center), and 500 μm (bottom), and comparison with the models of Béthermin et al. (2011) in red, Valiante et al. (2009) in green, Le Borgne et al. (2009) in violet, Gruppioni et al. (2011) in orange, Rahmati & van der Werf (2011) in light blue and Marsden et al. (2011) in dark blue.

z_{max}	$\nu I_\nu(z < z_{max})$ (in nW m^{-2} sr^{-1})		
	250 μm	350 μm	500 μm
0.2	$0.6^{+0.4}_{-0.3}$	$0.4^{+0.3}_{-0.2}$	$0.3^{+0.2}_{-0.2}$
0.5	$1.8^{+0.5}_{-0.5}$	$0.7^{+0.3}_{-0.2}$	$0.2^{+0.1}_{-0.1}$
0.8	$3.5^{+1.0}_{-1.0}$	$1.7^{+0.6}_{-0.5}$	$0.9^{+0.4}_{-0.3}$
1.0	$4.9^{+1.4}_{-1.3}$	$2.6^{+0.8}_{-0.8}$	$1.2^{+0.5}_{-0.4}$
1.5	$7.2^{+1.9}_{-1.8}$	$4.1^{+1.2}_{-1.1}$	$1.6^{+0.6}_{-0.5}$
2.0	$8.3^{+2.2}_{-2.0}$	$5.0^{+1.4}_{-1.3}$	$2.0^{+0.7}_{-0.6}$
3.0	$9.7^{+2.5}_{-2.3}$	$6.1^{+1.7}_{-1.5}$	$2.5^{+0.8}_{-0.7}$

Table 8. CIB build-up as a function of redshift at 250 μm, 350 μm, and 500 μm.

9.3. Spectral energy distribution of the CIB and total integrated CIB

Combining the total extrapolated CIB measured from deep surveys at various wavelengths, we can produce a fully-empirical SED of the CIB (see Fig. 17). We used the values coming from resolved counts at 16 μm (Teplitz et al. 2011), 24 μm (Béthermin et al. 2010a), 100 μm and 160 μm (Berta et al. 2011), as well as counts measured by stacking analyses at 70 μm (Béthermin et al. 2010a), our new results at 250 μm,

	Resolved ($S > 20$ mJy)		Stacking ($S > 2$ mJy)		Total extrapolated	Absolute measurements	
Wavelength μm	level nW m^{-2} sr^{-1}	fraction	level nW m^{-2} sr^{-1}	fraction	level nW m^{-2} sr^{-1}	Fixsen et al. (1998) nW m^{-2} sr^{-1}	Lagache et al. (2000) nW m^{-2} sr^{-1}
250	$1.55^{+0.30}_{-0.30}$	15%	$7.40^{+1.42}_{-1.43}$	73%	$10.13^{+2.60}_{-2.33}$	10.40±2.30	11.75±2.90
350	$0.77^{+0.16}_{-0.16}$	11%	$4.50^{+0.90}_{-0.90}$	69%	$6.46^{+1.74}_{-1.57}$	5.40±1.60	6.43±1.59
500	$0.14^{+0.03}_{-0.03}$	5%	$1.54^{+0.34}_{-0.34}$	55%	$2.80^{+0.93}_{-0.81}$	2.40±0.60	2.70±0.67

Table 7. Summary of the contribution to the CIB for various flux density cuts, and comparison with the absolute measurements (which themselves have large uncertainties).

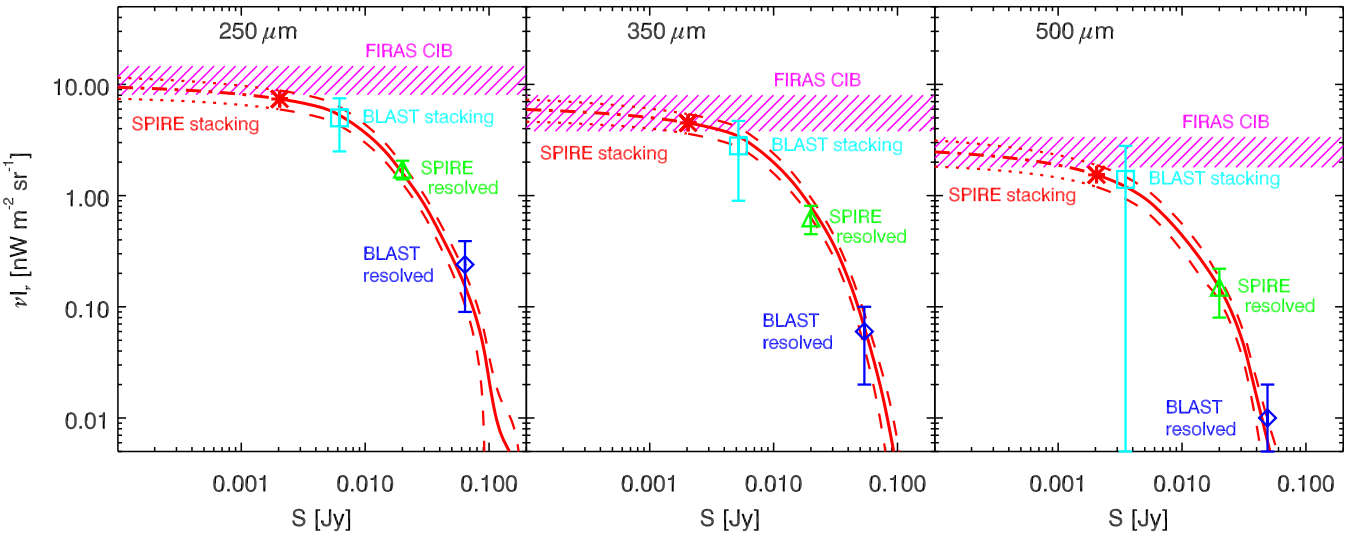


Figure 15. Cumulative contribution to the CIB as a function of the flux density cut at 250 µm (left), 350 µm (center), and 500 µm (right). *Red*: cumulative contribution from our counts. The asterisks represents the fraction resolved at the limit used for our analysis. *Cyan*: contribution of the BLAST sources probed by stacking (Béthermin et al. 2010b). *Green*: contribution of the sources resolved by SPIRE (Oliver et al. 2010b). *Blue*: contribution of the sources resolved by BLAST (Béthermin et al. 2010b). *Violet hatched region*: FIRAS absolute measurement of the CIB; a region is hatched here if it is in the 1-σ confidence region of Fixsen et al. (1998) or Lagache et al. (2000).

350 µm, and 500 µm, and also resolved sources in lensed areas at 850 µm (Zemcov et al. 2010). From these values, we then estimate the total CIB integrated between 8 µm and 1000 µm: 27^{+7}_{-3} nW m^{-2}.sr^{-1}. We use power-laws to interpolate between the data points. To account for the fact that the different data points were estimated in similar fields and are thus likely to be significantly correlated, we assume a perfect correlation between each wavelengths to obtain conservative uncertainties.

We also derive the contribution to the total CIB from different redshift slices. We use the extrapolated values deduced from the counts per redshift slice of Le Floc'h et al. (2009), Berta et al. (2011) and our SPIRE measurements. The Berta et al. (2011) counts were integrated following the same method as for our SPIRE counts. Fig. 17 (colored lines) shows how the CIB SED is built up as a function of redshift. The contribution of the various redshift slices to the CIB integrated between 8 μm and 1000 μm is given Table 9.

Redshift slice	integrated CIB intensity nW m^{-2} sr^{-1}
$0 < z < 0.5$	6.3 ± 1.5
$0.5 < z < 1$	7.9 ± 2.2
$1 < z < 2$	7.7 ± 2.8
$z > 2$	4.7 ± 2.0

Table 9. Contribution of the various redshift slices to the CIB integrated between 8 μm and 1000 μm.

10. Discussion

10.1. Deep source counts in the 250-500 µm range

With our new stacking analysis, we have confirmed, with a completely independent method, the deep counts produced by Glenn et al. (2010) using a $P(D)$ analysis. Unlike for $P(D)$ analysis, our stacking approach allows binning in redshift, providing new information on the SPIRE sources.

Our knowledge on the number counts in this wavelength interval has dramatically improved in few last years. Before BLAST and *Herschel*, the source counts were very poorly constrained by ground-based observations, e.g. 350 µm had only three reported ~20 mJy sources (Khan et al. 2007). Now, thanks to *Herschel*, they are well constrained between 2 mJy and 1 Jy.

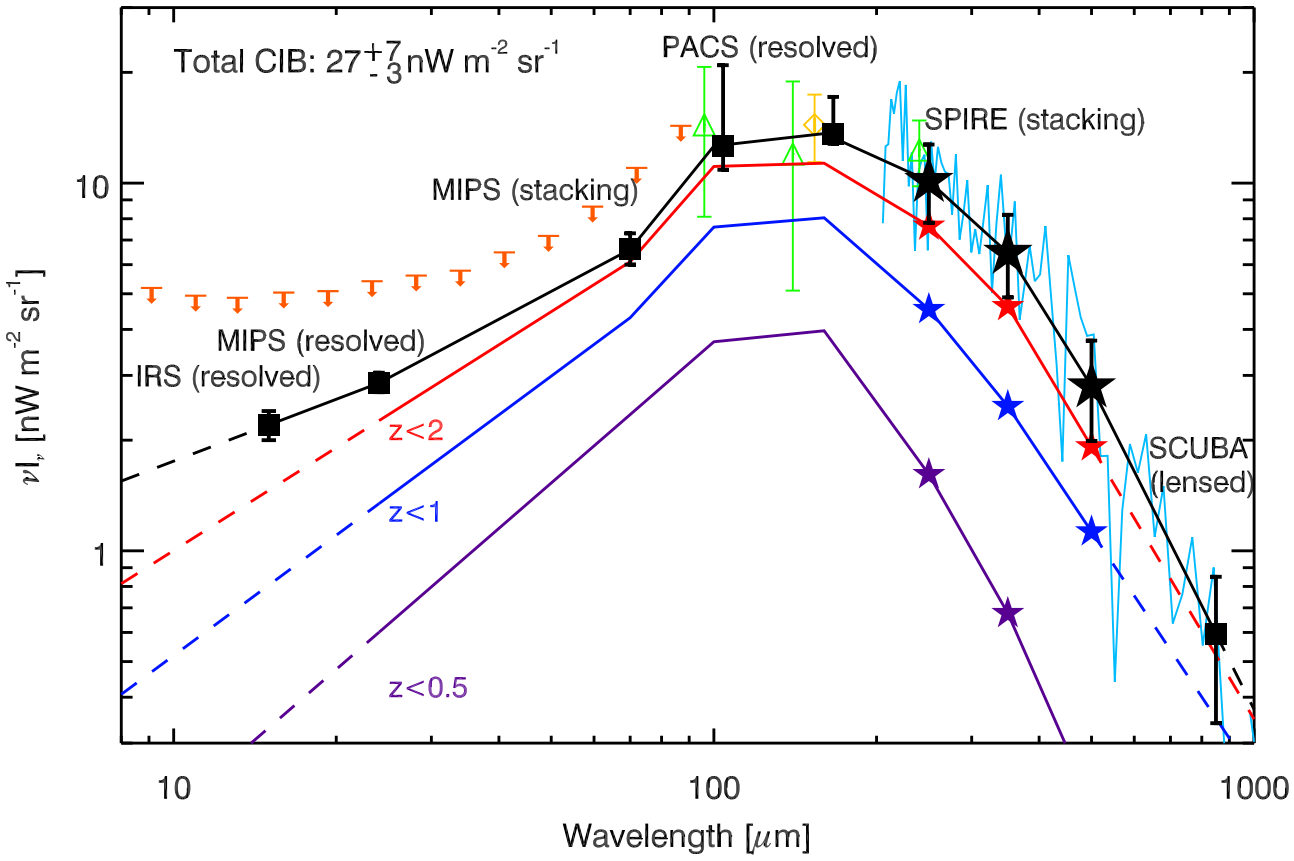


Figure 17. Spectral energy distribution of the CIB. *Black filled stars*: our total extrapolated CIB at 250 µm, 350 µm, and 500 µm. *Black filled squares*: total extrapolated CIB from deep number counts at 16 µm (Teplitz et al. 2011), 24 µm and 70µm (Béthermin et al. 2010a), 100 µm and 160 µm (Berta et al. 2011), and 850 µm (Zemcov et al. 2010). *Colored solid lines*: contribution of the z<0.5 (purple), $z < 1$ (dark blue), and $z < 2$ (red) sources to the CIB from the counts measured by Le Floc'h et al. (2009) at 24 µm, Berta et al. (2011) at 70 µm, 100 µm, and 160 µm), and in this paper at 250 µm, 350 µm, and 500 µm. *Colored filled stars*: our total extrapolated CIB at 250 µm, 350 µm, 500 µm for various cuts in redshift. The *colored stars* indicate our new points. The *dashed lines* correspond to the extrapolation of these contributions below 24 µm and above 500 µm. *Cyan solid line:* Absolute CIB spectrum measured by *COBE*/FIRAS (Lagache et al. 2000). *Green triangles*: absolute CIB measurements performed by *COBE*/DIRBE at 100 µm, 140 µm, and 240 µm (updated in Dole et al. 2006). *Yellow diamond:* absolute measurements of Pénin et al. (2011b) at 160 µm with *Spitzer*/MIPS. *Orange arrows*: upper limits derived from opacity of the Universe to TeV photons (Mazin & Raue 2007). The Berta et al. (2011), Pénin et al. (2011b), and *COBE*/FIRAS points have been slightly shifted in wavelength for clarity.

10.2. New statistical constraints for the models

The number counts alone are not sufficient to constrain evolution models. In this paper, we have compared different models fit to number counts. Some of these models reproduce the number counts using incorrect redshift distributions; while here we show that in fact, all the models are ruled out by our measurements. This highlights how redshift information is crucial in this context. The importance of the redshift distributions of the sources and of the CIB was also pointed out by Le Floc'h et al. (2009), Jauzac et al. (2011), Béthermin et al. (2011) and Berta et al. (2011) among others. Le Floc'h et al. (2009) measured the counts and the redshift distribution at 24 µm with *Spitzer* in the COSMOS field. Berta et al. (2011) produced a large collection of observables in the PACS bands (70 µm, 100 µm, and 160 µm). Here, we provide the same type of observables in the SPIRE bands, using a stacking analysis to reach a depth similar to Berta et al. (2011), despite having a stronger confusion. The combination of these three datasets will provide very stringent constraints for the next generation of evolution models.

10.3. Origin of the sub-mm part of the cosmic infrared background

Thanks to the depth and the precision of our new measurements, we can now study the sub-mm part of the CIB from an empirical point of view. As predicted by most of the models, the mean redshift of the CIB increases with wavelength (e.g. Lagache et al. 2005). This confirms that the CIB in the sub-mm domain is dominated by the high-redshift populations. The extrapolation of our counts down to zero flux density provides an estimation of the sub-mm CIB in agreement with the absolute measurements. Our reconstruction of the properties of the

SPIRE sources from the mid-infrared and optical data can thus explain how the sub-mm CIB was emitted.

In addition, these redshift distributions will help to interpret the CIB fluctuations measured by *Herschel* (Amblard et al. 2011, also see calibration, flux cut and galactic cirrus discussion in Planck collaboration XVIII (2011)) and *Planck* (Planck collaboration XVIII 2011). In fact, PACS and SPIRE redshift distributions constrain the emissivities of the infrared galaxies as a function of redshift, and will help to break degeneracies between these emissivities and the mass of the dark matter halos hosting the star-forming galaxies in the fluctuation models (e.g. Planck collaboration XVIII 2011; Pénin et al. 2011a).

11. Conclusion

Thanks to the sensitivity of SPIRE and the high-quality of the ancillary data in the GOODS and COSMOS fields, we have determined new statistical constraints on the sub-mm galaxies. The main results of this work are:

- We produced deep counts (down to 2 mJy), which confirm the previous measurements performed by stacking (Béthermin et al. 2010b) and $P(D)$ analysis (Glenn et al. 2010), and significantly reduce the uncertainties in the measurements. In addition, we provide number counts per redshift slice at these wavelengths.
- We measured the redshift distribution of the sources below the confusion limit using a stacking analysis.
- We compared our results with the predictions of the most recent evolutionary models, which do not manage to accurately reproduce our new points. These new constraints will thus be very useful for building a new generation of models.
- From our source counts, we also derived new estimates of the CIB level at 250 μm, 350 μm, and 500 μm, in agreement and with an accuracy competitive with the FIRAS absolute measurements. We also derived constraints on the redshift distribution of the CIB.
- Finally, combining our results with other work, we have estimated the CIB integrated between 8 μm and 1000 μm, produced by galaxies, to be 27^{+7}_{-3} nW m^{-2} sr^{-1}.

Acknowledgements. We thank the COSMOS and GOODS teams for releasing publicly their data. Thanks to Georges Helou for suggesting that the distribution of the colors is more likely log-normal than normal. MB thank Hervé Dole for his advices about stacking, and Elizabeth Fernandez for providing a mock catalog from the Béthermin et al. model. SPIRE has been developed by a consortium of institutes led by Cardiff Univ. (UK) and including Univ. Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); Caltech, JPL, NHSC, Univ. Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC, UKSA (UK); and NASA (USA). MB acknowledge financial support from ERC-StG grant UPGAL 240039. SJO acknowledge support from the Science and Technology Facilities Council [grant number ST/F002858/1] and [grant number ST/I000976/1]. MV was supported by the Italian Space Agency (ASI Herschel Science Contract I/005/07/0). The data presented in this paper will be released through the *Herschel* Database in Marseille HeDaM (hedam.oamp.fr/HerMES).

References

Alexander, D. M., Bauer, F. E., Chapman, S. C., et al. 2005, ApJ, 632, 736
Amblard, A., Cooray, A., Serra, P., et al. 2011, Nature, 470, 510
Amblard, A., Cooray, A., Serra, P., et al. 2010, A&A, 518, L9+
Bavouzet, N. 2008, PhD thesis, Université Paris-Sud 11
Berta, S., Magnelli, B., Nordon, R., et al. 2011, A&A, 532, A49
Béthermin, M., Dole, H., Beelen, A., & Aussel, H. 2010a, A&A, 512, A78+
Béthermin, M., Dole, H., Cousin, M., & Bavouzet, N. 2010b, A&A, 516, A43+
Béthermin, M., Dole, H., Lagache, G., Le Borgne, D., & Penin, A. 2011, A&A, 529, A4+
Blain, A. W., Chapman, S. C., Smail, I., & Ivison, R. 2004, ApJ, 611, 725
Chapin, E. L., Chapman, S. C., Coppin, K. E., et al. 2011, MNRAS, 411, 505
Chary, R. & Elbaz, D. 2001, ApJ, 556, 562
Chiaberge, M., Capetti, A., Macchetto, F. D., et al. 2010, ApJ, 710, L107
Clements, D. L., Rigby, E., Maddox, S., et al. 2010, A&A, 518, L8
Condon, J. J. 1974, ApJ, 188, 279
Cooray, A., Amblard, A., Wang, L., et al. 2010, A&A, 518, L22+
Dale, D. A. & Helou, G. 2002, ApJ, 576, 159
Devlin, M. J., Ade, P. A. R., Aretxaga, I., et al. 2009, Nature, 458, 737
Dole, H., Lagache, G., & Puget, J. 2003, ApJ, 585, 617
Dole, H., Lagache, G., Puget, J., et al. 2006, A&A, 451, 417
Eales, S. A., Raymond, G., Roseboom, I. G., et al. 2010, A&A, 518, L23+
Farrah, D., Lonsdale, C. J., Borys, C., et al. 2006, ApJ, 641, L17
Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, ApJ, 508, 123
Glenn, J., Conley, A., Béthermin, M., et al. 2010, MNRAS, 409, 109
Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3+
Gruppioni, C., Pozzi, F., Zamorani, G., & Vignali, C. 2011, MNRAS, 416, 70
Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, ApJ, 508, 25
Hauser, M. G. & Dwek, E. 2001, ARA&A, 39, 249
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Jauzac, M., Dole, H., Le Floc'h, E., et al. 2011, A&A, 525, A52+
Khan, S. A., Shafer, R. A., Serjeant, S., et al. 2007, ApJ, 665, 973
Lagache, G., Abergel, A., Boulanger, F., Désert, F. X., & Puget, J.-L. 1999, A&A, 344, 322
Lagache, G., Dole, H., Puget, J.-L., et al. 2004, ApJS, 154, 112
Lagache, G., Haffner, L. M., Reynolds, R. J., & Tufte, S. L. 2000, A&A, 354, 247
Lagache, G., Puget, J., & Dole, H. 2005, ARA&A, 43, 727
Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64
Le Borgne, D., Elbaz, D., Ocvirk, P., & Pichon, C. 2009, A&A, 504, 727
Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222
Levenson, L., Marsden, G., Zemcov, M., et al. 2010, MNRAS, 409, 83
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Magdis, G. E., Elbaz, D., Dickinson, M., et al. 2011, A&A, 534, A15
Magliocchetti, M., Santini, P., Rodighiero, G., et al. 2011, MNRAS, 416, 1105
Magnelli, B., Elbaz, D., Chary, R. R., et al. 2011, A&A, 528, A35+
Marsden, G., Ade, P. A. R., Bock, J. J., et al. 2009, ApJ, 707, 1729
Marsden, G., Chapin, E. L., Halpern, M., et al. 2011, MNRAS, 417, 1192
Mazin, D. & Raue, M. 2007, A&A, 471, 439
Negrello, M., Hopwood, R., De Zotti, G., et al. 2010, Science, 330, 800
Nguyen, H. T., Schulz, B., Levenson, L., et al. 2010, A&A, 518, L5
Oliver, S., Bock, J., Altieri, B., et al. 2011, submitted to MNRAS
Oliver, S., Frost, M., Farrah, D., et al. 2010a, MNRAS, 405, 2279
Oliver, S. J., Wang, L., Smith, A. J., et al. 2010b, A&A, 518, L21+
Papovich, C., Dole, H., Egami, E., et al. 2004, ApJS, 154, 70
Pascale, E., Ade, P. A. R., Bock, J. J., et al. 2008, ApJ, 681, 400
Patanchon, G., Ade, P. A. R., Bock, J. J., et al. 2009, ApJ, 707, 1750
Pénin, A., Doré, O., Lagache, G., & Béthermin, M. 2011a, A&Ain press
Pénin, A., Lagache, G., Noriega-Crepo, A., et al. 2011b, sub. A&A
Penner, K., Pope, A., Chapin, E. L., et al. 2011, MNRAS, 410, 2749
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1+
Planck collaboration XVIII. 2011, A&A, 536, A18
Puget, J., Abergel, A., Bernard, J., et al. 1996, A&A, 308, L5+
Rahmati, A. & van der Werf, P. P. 2011, MNRAS, 418, 176
Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, MNRAS, 409, 48
Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, ApJ, 690, 1250
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Scoville, N., Aussel, H., Benson, A., et al. 2007, ApJS, 172, 150
Smith, A. J., Wang, L., Oliver, S. J., et al. 2012, MNRAS, 419, 377
Swinyard, B. M., Ade, P., Baluteau, J.-P., et al. 2010, A&A, 518, L4+
Teplitz, H. I., Chary, R., Elbaz, D., et al. 2011, AJ, 141, 1
Valiante, E., Lutz, D., Sturm, E., Genzel, R., & Chapin, E. L. 2009, ApJ, 701, 1814
Vieira, J. D., Crawford, T. M., Switzer, E. R., et al. 2010, ApJ, 719, 763
Viero, M. P., Moncelsi, L., Buitrago, F., et al. 2011, sub. ApJ
Wall, J. V. & Jenkins, C. R. 2003, Practical Statistics for Astronomers, ed. Wall, J. V. & Jenkins, C. R.
Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1
Zemcov, M., Blain, A., Halpern, M., & Levenson, L. 2010, ApJ, 721, 424

[1] Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu - CNRS - Université Paris Diderot, CE-Saclay, pt courrier 131, F-91191 Gif-sur-Yvette, France
e-mail: matthieu.bethermin@cea.fr
[2] Institut d'Astrophysique Spatiale (IAS), bâtiment 121, Université Paris-Sud 11 and CNRS (UMR 8617), 91405 Orsay, France
[3] Laboratoire d'Astrophysique de Marseille, OAMP, Université Aix-marseille, CNRS, 38 rue Frédéric Joliot-Curie, 13388 Marseille cedex 13, France
[4] Center for Astrophysics and Space Astronomy 389-UCB, University of Colorado, Boulder, CO 80309, USA
[5] NASA, Ames Research Center, Moffett Field, CA 94035, USA
[6] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[7] Max-Planck-Institut für Extraterrestrische Physik (MPE), Postfach 1312, 85741, Garching, Germany
[8] California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[9] Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[10] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[11] Instituto de Astrofísica de Canarias (IAC), E-38200 La Laguna, Tenerife, Spain
[12] Departamento de Astrofísica, Universidad de La Laguna (ULL), E-38205 La Laguna, Tenerife, Spain
[13] Departamento de Astrofísica, Facultad de CC. Físicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain
[14] Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[15] Dept. of Physics & Astronomy, University of California, Irvine, CA 92697, USA
[16] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK
[17] Astronomy Centre, Dept. of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK
[18] Dipartimento di Astronomia, Università di Padova, vicolo Osservatorio, 3, 35122 Padova, Italy
[19] Dept. of Astrophysical and Planetary Sciences, CASA 389-UCB, University of Colorado, Boulder, CO 80309, USA
[20] ESO, Karl-Schwarzschild-Str. 2, 85748 Garching bei München, Germany
[21] UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[22] Hubble Fellow
[23] National Optical Astronomy Observatory, 950 North Cherry Avenue, Tucson, AZ 85719, USA
[24] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada
[25] Institut d'Astrophysique de Paris, UMR 7095, CNRS, UPMC Univ. Paris 06, 98bis boulevard Arago, F-75014 Paris, France
[26] Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK
[27] RAL Space, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK
[28] Institute for Space Imaging Science, University of Lethbridge, Lethbridge, Alberta, T1K 3M4, Canada
[29] Department of Astrophysics, Denys Wilkinson Building, University of Oxford, Keble Road, Oxford OX1 3RH, UK
[30] Max Planck Institut für Plasma Physik and Excellence Cluster, 85748 Garching, Germany
[31] Infrared Processing and Analysis Center, MS 100-22, California Institute of Technology, JPL, Pasadena, CA 91125, USA
[32] CSIRO Astronomy & Space Science, PO Box 76, Epping, NSW 1710, Australia
[33] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[34] Herschel Science Centre, European Space Astronomy Centre, Villanueva de la Cañada, 28691 Madrid, Spain

Appendix A: Measurement of the auto-correlation function

To compute the uncertainties in our counts, we measure the auto-correlation function (ACF) using a new method, based on the stacking of a density map, containing in each pixel the number of sources centered on it. If we stack this map at the position of the sources, the expected stacked image $M(\theta)$ will be (Bavouzet 2008; Béthermin et al. 2010b)

$$M(\theta) = (1 + w(\theta)) \times \rho_S, \quad \text{(A.1)}$$

where ρ_S is the source density (in sources pixel^{-1}) and $w(\theta)$ the ACF. Note that this method is not fully accurate because a relative error of 10^{-3} on ρ_S affects $w(\theta)$ by an absolute error of the same amount.

We can generalize this method to compute the Landy & Szalay (1993) estimator ($w(\theta) = (DD - 2 \times DR + RR)/RR$, see Sect. 5.1.2). In this case, we use two density maps, one for the real sources, and one for a simulated catalog, called hereafter "real" and "random" maps, respectively. DD is estimated by stacking of the real map at the positions of the real catalog, DR by stacking of the random map at the position of the real sources, and RR by stacking of the random map at the position of the random sources. We then compute the Landy & Szalay (1993) estimator from the three stacked maps: DD, DR, and RR. This provides an estimate for $w(\theta, \phi)$, where (θ, ϕ) are polar coordinates. To reduce the noise, we compute the mean in several annuli.

This method has a computation time proportional to $N_{\rm sources}$, instead of $N^2_{\rm sources}$ for the naive one. Nevertheless, the computation time is also proportional $N^2_{\rm pixels}$. A small number of pixels reduces the range of scales which can be probed. We thus use successive rebinning of our density maps to accelerate the computation of the ACF over a wide range of scales.

Appendix B: Additional tables

Flux	Correction factor			
mJy	$0 < z < 0.5$	$0.5 < z < 1$	$1 < z < 2$	$z > 2$
		250 μm		
23.8	0.89±0.08	0.79±0.08	0.81±0.08	0.86±0.09
33.6	0.91±0.10	0.81±0.10	0.89±0.09	0.87±0.13
47.4	0.92±0.15	0.94±0.15	0.82±0.14	0.85±0.20
67.0	0.96±0.20	0.96±0.27	0.92±0.29	0.94±0.55
94.6	1.02±0.30	0.96±0.65	1.16±0.60	0.89±1.04
133.7	-	0.97±1.19	-	-
188.8	0.96±0.83	-	-	-
		350 μm		
23.8	0.70±0.10	0.67±0.09	0.80±0.07	0.81±0.07
33.6	0.82±0.15	0.82±0.12	0.73±0.10	0.88±0.11
47.4	0.79±0.24	0.79±0.19	0.93±0.17	0.85±0.17
67.0	0.95±0.44	0.67±0.37	0.85±0.26	0.86±0.35
94.6	1.07±1.85	0.61±2.00	1.10±0.74	-
133.7	-	-	-	0.62±1.72
		500 μm		
23.8	0.72±0.18	0.63±0.14	0.65±0.11	0.75±0.10
33.6	0.74±0.32	0.66±0.21	0.77±0.18	0.90±0.16
47.4	-	0.76±0.50	0.73±0.28	0.63±0.23
67.0	-	-	0.94±1.08	0.85±0.86
94.6	-	-	0.62±2.09	0.64±1.49

Table B.1. Correction factor applied to the resolved counts in the various flux density and redshift bins.

Flux	Completeness correction factor							
mJy	$0 < z < 0.25$	$0.25 < z < 0.5$	$0.5 < z < 0.75$	$0.75 < z < 1$	$1 < z < 1.5$	$1.5 < z < 2$	$2 < z < 3$	$z > 3$
				250 μm				
2.1	1.13±0.22	1.27±0.31	2.63±0.17	3.20±0.04	1.75±0.44	1.55±0.15	1.71±0.24	2.55±0.33
3.0	1.05±0.13	1.11±0.18	1.81±0.28	2.22±0.24	1.35±0.32	1.24±0.10	1.36±0.19	1.75±0.43
4.2	1.02±0.07	1.04±0.10	1.38±0.22	1.65±0.26	1.15±0.20	1.09±0.06	1.17±0.12	1.34±0.30
				350 μm				
2.1	1.26±0.22	1.17±0.27	2.11±0.35	1.83±0.33	1.61±0.37	1.62±0.35	1.72±0.25	9.15±6.24
3.0	1.10±0.12	1.07±0.17	1.54±0.32	1.35±0.30	1.24±0.30	1.26±0.28	1.33±0.20	3.74±0.30
4.2	1.03±0.06	1.02±0.09	1.25±0.24	1.14±0.21	1.08±0.19	1.10±0.17	1.14±0.14	2.06±0.32
				500 μm				
2.1	1.34±0.54	1.59±0.31	1.40±0.31	1.04±0.09	1.40±0.37	1.06±0.12	1.18±0.27	1.55±0.19
3.0	1.12±0.38	1.25±0.22	1.12±0.16	1.08±0.04	1.17±0.27	1.02±0.06	1.06±0.17	1.24±0.13
4.2	1.04±0.24	1.09±0.12	1.03±0.06	1.02±0.02	1.06±0.17	1.00±0.03	1.02±0.09	1.09±0.07

Table B.2. Completeness correction factor applied to the counts by stacking in GOODS-N in the various flux and redshift bins.

Flux	Completeness correction factor							
mJy	$0 < z < 0.25$	$0.25 < z < 0.5$	$0.5 < z < 0.75$	$0.75 < z < 1$	$1 < z < 1.5$	$1.5 < z < 2$	$2 < z < 3$	$z > 3$
				250 μm				
6.0	1.41±0.29	1.20±0.25	1.44±0.30	1.49±0.28	1.67±0.21	1.24±0.36	1.36±0.20	2.41±0.07
8.4	1.19±0.17	1.06±0.14	1.18±0.22	1.24±0.18	1.28±0.22	1.07±0.26	1.14±0.14	1.70±0.27
11.9	1.09±0.09	1.02±0.07	1.06±0.13	1.10±0.11	1.12±0.15	1.02±0.17	1.05±0.08	1.34±0.24
16.8	1.04±0.04	1.00±0.03	1.02±0.07	1.04±0.05	1.05±0.09	1.00±0.10	1.02±0.04	1.15±0.17
				350 μm				
6.0	1.40±0.37	1.12±0.11	1.17±0.15	1.33±0.20	1.52±0.27	1.35±0.19	1.67±0.28	3.03±0.59
8.4	1.18±0.28	1.04±0.05	1.06±0.09	1.14±0.12	1.21±0.26	1.13±0.14	1.30±0.25	2.06±0.13
11.9	1.07±0.18	1.02±0.02	1.02±0.04	1.05±0.06	1.08±0.20	1.05±0.08	1.13±0.16	1.59±0.35
16.8	1.02±0.10	1.00±0.01	1.01±0.02	1.02±0.03	1.03±0.14	1.02±0.04	1.05±0.09	1.33±0.36
				500 μm				
6.0	1.11±0.30	1.05±0.18	1.23±0.29	1.20±0.35	1.25±0.25	1.14±0.25	1.30±0.26	2.22±0.52
8.4	1.04±0.22	1.01±0.10	1.12±0.20	1.08±0.25	1.12±0.19	1.05±0.15	1.11±0.22	1.51±0.53
11.9	1.02±0.15	1.00±0.05	1.06±0.14	1.03±0.16	1.06±0.14	1.01±0.09	1.04±0.16	1.20±0.36
16.8	1.00±0.11	1.00±0.02	1.03±0.09	1.01±0.10	1.03±0.09	1.00±0.05	1.01±0.10	1.08±0.22

Table B.3. Completeness correction factor applied to the counts by stacking in COSMOS in the various flux and redshift bins.

Flux	$\sigma_{\rm clus}/\sigma_{\rm poi+clus}$				$\sigma_{\rm clus+poi}/\sigma_{\rm tot}$			
mJy	$0 < z < 0.5$	$0.5 < z < 1$	$1 < z < 2$	$z > 2$	$0 < z < 0.5$	$0.5 < z < 1$	$1 < z < 2$	$z > 2$
				250 μm				
23.8	0.86	0.85	0.92	0.77	0.90	0.90	0.95	0.86
33.6	0.79	0.72	0.88	0.64	0.86	0.85	0.91	0.81
47.4	0.65	0.59	0.68	0.41	0.81	0.79	0.83	0.77
67.0	0.54	0.39	0.48	0.18	0.77	0.74	0.76	0.72
94.6	0.44	0.17	0.33	0.11	0.74	0.72	0.70	0.72
133.7	-	0.12	-	-	-	0.71	-	-
188.8	0.17	-	-	-	0.72	-	-	-
				350 μm				
23.8	0.87	0.89	0.93	0.88	0.92	0.94	0.95	0.92
33.6	0.80	0.84	0.84	0.83	0.88	0.90	0.91	0.89
47.4	0.58	0.62	0.73	0.65	0.81	0.82	0.83	0.82
67.0	0.44	0.26	0.46	0.36	0.75	0.78	0.77	0.76
94.6	0.25	0.12	0.26	-	0.70	0.79	0.70	-
133.7	-	-	-	0.10	-	-	-	0.78
				500 μm				
23.8	0.83	0.83	0.70	0.85	0.90	0.91	0.86	0.91
33.6	0.60	0.61	0.52	0.77	0.82	0.84	0.80	0.86
47.4	-	0.30	0.27	0.39	-	0.77	0.77	0.81
67.0	-	-	0.13	0.19	-	-	0.72	0.74
94.6	-	-	0.07	0.12	-	-	0.78	0.78

Table B.4. Sources of uncertainty on the counts measured from the resolved sources. *Columns 2 to 5*: relative contribution of the clustering term to the total sample variance (Poisson+clustering) of resolved counts ($\sigma_{\rm clus}/\sigma_{\rm poi+clus}$). *Columns 6 to 9*: relative contribution the sample variance term to the total uncertainties in resolved counts ($\sigma_{\rm clus+poi}/\sigma_{\rm tot}$). $\sigma_{\rm tot}$ contains both the uncertainties in the corrections and the sample variance.

Flux	$\sigma_{\rm clus}/\sigma_{\rm poi+clus}$				$\sigma_{\rm clus+poi}/\sigma_{\rm tot}$			
mJy	$0<z<0.5$	$0.5<z<1$	$1<z<2$	$z>2$	$0<z<0.5$	$0.5<z<1$	$1<z<2$	$z>2$
				250 μm				
2.1	0.79	0.95	0.84	0.77	0.56	0.63	0.47	0.44
3.0	0.77	0.94	0.82	0.74	0.66	0.58	0.54	0.50
4.2	0.75	0.93	0.80	0.72	0.74	0.68	0.64	0.60
				350 μm				
2.1	0.75	0.94	0.84	0.78	0.63	0.60	0.46	0.16
3.0	0.73	0.94	0.81	0.76	0.66	0.61	0.60	0.46
4.2	0.70	0.93	0.78	0.73	0.65	0.71	0.73	0.53
				500 μm				
2.1	0.69	0.93	0.77	0.76	0.44	0.62	0.55	0.22
3.0	0.65	0.92	0.75	0.74	0.45	0.70	0.53	0.51
4.2	0.62	0.91	0.70	0.71	0.46	0.74	0.50	0.51

Table B.5. *Columns 2 to 5*: relative contribution of the clustering term to the total sample variance (Poisson+clustering) on counts measured by stacking in GOODS-N ($\sigma_{\rm clus}/\sigma_{\rm poi+clus}$). *Columns 6 to 9*: relative contribution the sample variance term to the total uncertainties in resolved counts ($\sigma_{\rm clus+poi}/\sigma_{\rm tot}$). $\sigma_{\rm tot}$ contains both the uncertainties in the completeness corrections, the uncertainties in the mean colors and the scatter, and the sample variance.

Flux	$\sigma_{\rm clus}/\sigma_{\rm poi+clus}$				$\sigma_{\rm clus+poi}/\sigma_{\rm tot}$			
mJy	$0<z<0.5$	$0.5<z<1$	$1<z<2$	$z>2$	$0<z<0.5$	$0.5<z<1$	$1<z<2$	$z>2$
				250 μm				
6.0	0.91	0.97	0.95	0.88	0.26	0.46	0.28	0.21
8.4	0.89	0.97	0.93	0.84	0.39	0.64	0.40	0.31
11.9	0.85	0.95	0.90	0.78	0.61	0.83	0.61	0.45
16.8	0.81	0.93	0.85	0.69	0.78	0.80	0.57	0.42
				350 μm				
6.0	0.88	0.97	0.95	0.91	0.38	0.46	0.27	0.14
8.4	0.84	0.96	0.93	0.87	0.56	0.69	0.36	0.19
11.9	0.78	0.93	0.89	0.80	0.61	0.75	0.53	0.32
16.8	0.68	0.89	0.84	0.71	0.39	0.47	0.37	0.36
				500 μm				
6.0	0.74	0.94	0.91	0.90	0.32	0.52	0.41	0.22
8.4	0.60	0.89	0.87	0.83	0.25	0.34	0.41	0.26
11.9	0.46	0.80	0.79	0.74	0.22	0.23	0.26	0.27
16.8	0.33	0.64	0.65	0.61	0.20	0.16	0.18	0.21

Table B.6. Sources of uncertainty on the counts measured by stacking. *Columns 2 to 5*: relative contribution of the clustering term to the total sample variance (Poisson+clustering) on counts measured by stacking in COSMOS ($\sigma_{\rm clus}/\sigma_{\rm poi+clus}$). *Columns 6 to 9*: relative contribution the sample variance term to the total uncertainties in resolved counts ($\sigma_{\rm clus+poi}/\sigma_{\rm tot}$). $\sigma_{\rm tot}$ contains both the uncertainties in the completeness corrections, the uncertainties in the mean colors and the scatter, and the sample variance.