BLAST: CORRELATIONS IN THE COSMIC FAR-INFRARED BACKGROUND AT 250, 350, AND 500 μm REVEAL CLUSTERING OF STAR-FORMING GALAXIES

Marco P. Viero[1], Peter A. R. Ade[2], James J. Bock[3], Edward L. Chapin[4], Mark J. Devlin[5], Matthew Griffin[2], Joshua O. Gundersen[6], Mark Halpern[4], Peter C. Hargrave[2], David H. Hughes[7], Jeff Klein[5], Carrie J. MacTavish[8], Gaelen Marsden[4], Peter G. Martin[1,9], Philip Mauskopf[2], Lorenzo Moncelsi[2], Mattia Negrello[10], Calvin B. Netterfield[1,11], Luca Olmi[12,13], Enzo Pascale[2], Guillaume Patanchon[14], Marie Rex[5], Douglas Scott[4], Christopher Semisch[5], Nicholas Thomas[6], Matthew D. P. Truch[5], Carole Tucker[2], Gregory S. Tucker[15], and Donald V. Wiebe[4,11]

[1] Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada; viero@astro.utoronto.ca
[2] Department of Physics & Astronomy, Cardiff University, 5 The Parade, Cardiff, CF24 3AA, UK
[3] Jet Propulsion Laboratory, Pasadena, CA 91109-8099, USA
[4] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada
[5] Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
[6] Department of Physics, University of Miami, 1320 Campo Sano Drive, Coral Gables, FL 33146, USA
[7] Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE), Aptdo. Postal 51 y 72000 Puebla, Mexico
[8] Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[9] Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada
[10] Department of Physics and Astronomy, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
[11] Department of Physics, University of Toronto, 60 St. George Street, Toronto, ON M5S 1A7, Canada
[12] University of Puerto Rico, Rio Piedras Campus, Physics Department, Box 23343, UPR station, San Juan, PR, USA
[13] INAF—Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125, Firenze, Italy
[14] Université Paris Diderot, Laboratoire APC, 10, rue Alice Domon et Léonie Duquet 75205 Paris, France
[15] Department of Physics, Brown University, 182 Hope Street, Providence, RI 02912, USA

ABSTRACT

We detect correlations in the cosmic far-infrared background due to the clustering of star-forming galaxies in observations made with the Balloon-borne Large Aperture Submillimeter Telescope, at 250, 350, and 500 μm. We perform jackknife and other tests to confirm the reality of the signal. The measured correlations are well fitted by a power law over scales of $5'$–$25'$, with $\Delta I / I = 15.1\% \pm 1.7\%$. We adopt a specific model for submillimeter sources in which the contribution to clustering comes from sources in the redshift ranges $1.3 \leqslant z \leqslant 2.2$, $1.5 \leqslant z \leqslant 2.7$, and $1.7 \leqslant z \leqslant 3.2$, at 250, 350, and 500 μm, respectively. With these distributions, our measurement of the power spectrum, $P(k_\theta)$, corresponds to linear bias parameters, $b = 3.8 \pm 0.6$, 3.9 ± 0.6, and 4.4 ± 0.7, respectively. We further interpret the results in terms of the halo model, and find that at the smaller scales, the simplest halo model fails to fit our results. One way to improve the fit is to increase the radius at which dark matter halos are artificially truncated in the model, which is equivalent to having some star-forming galaxies at $z \geqslant 1$ located in the outskirts of groups and clusters. In the context of this model, we find a minimum halo mass required to host a galaxy is $\log(M_{\min}/M_\odot) = 11.5^{+0.4}_{-0.1}$, and we derive effective biases $b_{\rm eff} = 2.2 \pm 0.2$, 2.4 ± 0.2, and 2.6 ± 0.2, and effective masses $\log(M_{\rm eff}/M_\odot) = 12.9 \pm 0.3$, 12.8 ± 0.2, and 12.7 ± 0.2, at 250, 350 and 500 μm, corresponding to spatial correlation lengths of $r_0 = 4.9$, 5.0, and $5.2 \pm 0.7h^{-1}$Mpc, respectively. Finally, we discuss implications for clustering measurement strategies with *Herschel* and *Planck*.

Key words: galaxies: evolution – galaxies: high-redshift – infrared: galaxies – large-scale structure of universe – submillimeter

Online-only material: color figures

1. INTRODUCTION

With the discovery of the cosmic far-infrared background (CIB; Puget et al. 1996; Fixsen et al. 1998) and subsequent studies from the mid-infrared to the millimeter, it has been established that the peak epoch of star formation lies between $1 \lesssim z \lesssim 3$ (Dickinson et al. 2003; Hopkins 2004).

Where star formation occurs with respect to the underlying dark matter distribution is less well understood. In the local universe, the sites of most active star formation occur far from the densest environments, a property which reverses at $z \sim 1$ (Elbaz et al. 2007). This points to the environment contributing to the mechanisms that trigger or quench star formation.

Measurements of the clustering of star-forming galaxies on large and small scales can be used to directly relate star for-

mation to the environment. In the regime of linear growth of structure, the clustering amplitude relative to that of the underlying dark matter is described by a bias parameter, *b*, which describes how strongly star formation traces the underlying dark matter. On smaller scales, the distribution of galaxies hosted by a dark matter halo is described by the halo occupation distribution (HOD; Peacock & Smith 2000). It contains information on the abundance of star-forming sources within individual halos. Targeting the submillimeter band is particularly efficient for observing star formation directly. The submillimeter background results from the thermal emission of interstellar dust in high-redshift star-forming galaxies, which is heated by optical and ultraviolet radiation from stars and to a lesser extent active galactic nuclei (see Blain et al. 2002). A substantial effort has been devoted to surveys of these galaxies (e.g., Smail et al. 1997;

Hughes et al. 1998; Barger et al. 1998; Borys et al. 2003); thus, in principle, a clustering signal could be measured from these sources directly.

However, direct measurement of the clustering properties of resolved submillimeter galaxies has been elusive. Due to limited mapping speeds, the areal coverage of even the most ambitious submillimeter surveys has been relatively small (e.g., SHADES mapped approximately a quarter square degree at 850 μm; Coppin et al. 2006). In addition, due to steeply falling counts and modest resolutions of single-dish submillimeter telescopes, source confusion has made it difficult to resolve any sources other than those with very high signal-to-noise ratio. These sources span a relatively wide redshift range, roughly $1 \leqslant z \leqslant 4$, peaking at $z \sim 2.4$ (Chapman et al. 2005), which has the effect of washing out the angular clustering signal, further complicating measurements. This difficulty was confirmed by Scott et al. (2006), who re-analyzed all the SCUBA fields and found tentative evidence of strong angular clustering, but with errors too large to adequately constrain the spatial correlation length. Furthermore, relatively large beams have made it difficult to efficiently identify direct counterparts (Barger et al. 1999; Ivison et al. 2000) in order to obtain spectroscopic or photometric redshifts. Blain et al. (2004) attempted to measure the spatial clustering properties combining 73 sources with spectroscopic information, but again were only able to tentatively measure the clustering length. Other SCUBA galaxy clustering measurements, some tentatively detecting clustering, have been made by Webb et al. (2003) and Blake et al. (2006); meanwhile Almaini et al. (2003) claim to have found evidence for strong angular clustering between X-ray and submillimeter populations, although Borys et al. (2004) were unable to confirm the result using another, seemingly less biased, estimator. Using a nearest neighbor analysis, Greve et al. (2004) found that most significant MAMBO (1.2 mm) sources come in pairs, separated by $\sim 23''$. Put together, these limitations have made it extremely difficult to measure the clustering signal robustly.

To study the clustering properties of submillimeter galaxies, it is more powerful to consider the statistics of the *unresolved* CIB, which contains the full intensity, rather than working from a limited catalog. In other words, instead of measuring a correlation among numbers of galaxies, use the background fluctuations of the total intensity to measure correlations among brightnesses of galaxies. Devlin et al. (2009) and Marsden et al. (2009) have demonstrated that the CIB is composed of emission by discrete sources. Since submillimeter galaxies are optically thin, this signal will be proportional to the total star formation rates of those sources. Correlations in the CIB will have a contribution in excess of white noise—which arises from Poisson sampling of a background made up of a finite number of sources—in the presence of clustering, with an amplitude that should be detectable with current surveys (Scott & White 1999; Haiman & Knox 2000; Knox et al. 2001; Magliocchetti et al. 2001; Perrotta et al. 2003; Amblard & Cooray 2007; Negrello et al. 2007). Initial attempts to detect correlations by Peacock et al. (2000), for the Hubble Deep Field observed by SCUBA at 850 μm, and by Lagache & Puget (2000) for a 0.25 deg^2 *Infrared Space Observatory* (*ISO*) field at 170 μm, were only able to measure a signal consistent with the Poisson contribution. More recently, Grossan & Smoot (2007) and Lagache et al. (2007) reported the weak detection of a clustering component in 160 μm data from ~ 9 deg^2 *Spitzer* fields.

In this paper, we report the detection of correlations in the submillimeter part of the CIB due to the clustering of star-forming galaxies, in a 6 deg^2 field centered on the Great Observatories Origins Deep Survey South field (GOODS-South; Giavalisco et al. 2004). These data were collected by the Balloon-borne Large Aperture Submillimeter Telescope (BLAST; Devlin et al. 2009), which is designed to bracket the peak of redshifted thermal emission from dust by observing at 250, 350, and 500 μm. Operating above most of the atmosphere, BLAST is able to make observations in bands which are difficult or impossible to observe from the ground. A detailed description of the instrument and calibration can be found in Pascale et al. (2008) and Truch et al. (2009).

This paper is organized as follows. In Part I, we describe how we make the measurement—from map preparation to power spectrum calculation. We address each contribution to the total power spectrum, and how they are removed to uncover the clustering signal. We show that the observed spectra are consistent across BLAST bands and correspond to a modest bias. In Part II, we assess the plausibility of the detection by fitting models to the data: beginning with a simple linear bias model, followed by a more detailed halo model (Mo & White 1996; Cooray & Sheth 2002) and *HOD* (Peacock & Smith 2000). When required we adopt the concordance model, a flat ΛCDM cosmology with $\Omega_M = 0.274$, $\Omega_\Lambda = 0.726$, $H_0 = 70.5$ km s^{-1} Mpc^{-1}, and $\sigma_8 = 0.81$ (Hinshaw et al. 2009).

PART I. MEASURING CORRELATIONS IN THE CIB

2. BACKGROUND CORRELATIONS AND THE CORRELATION FUNCTION: OVERVIEW

Galaxy clustering can be expressed in a number of ways, the most common being the two-point correlation function, $w(\theta)$, which measures the number of pairs at a given distance in excess of what would be expected of a Poisson distribution. Alternatively, the clustering of galaxies can be expressed as a power spectrum in excess of Poisson noise, $\tilde{P}(k_\theta)$, which can be expressed in dimensionless units as the fractional variance per logarithmic increment of wavenumber (see Peacock 1999), i.e.,

$$\tilde{\Delta}^2_{k_\theta} \equiv 2\pi k_\theta^2 \tilde{P}(k_\theta), \tag{1}$$

where k_θ is the angular wavenumber, which is also known as σ in the literature, and is expressed in inverse angular scale as $k_\theta = 1/\lambda$. It is related to the multipole index, ℓ, by $\ell = 2\pi k_\theta$. The tildes over $\Delta^2_{k_\theta}$ and $P(k_\theta)$ denote that these quantities refer to galaxy *locations* rather than intensities.

Naturally, the correlation function and the power spectrum of galaxy clustering are related; they form a Hankel transform pair. Explicitly, for small surveys,

$$\tilde{\Delta}^2_{k_\theta} = (2\pi k_\theta)^2 \int_0^\infty w(\theta) J_0(2\pi k_\theta \theta) \theta \, d\theta,$$
$$w(\theta) = \int_0^\infty \tilde{\Delta}^2_{k_\theta} J_0(2\pi k_\theta \theta) \, dk_\theta / k_\theta. \tag{2}$$

For small separation correlations of local galaxies, angular clustering is often described as a power law, $w(\theta) = (\theta/\theta_0)^{-\epsilon}$, where $\epsilon \simeq 0.8$ is the canonical slope (e.g., Giavalisco et al. 1998). The power spectrum analog for small areas (see Peacock et al. 2000) is

$$\tilde{\Delta}^2_{k_\theta}(k_\theta) = (2\pi k_\theta \theta_0)^\epsilon 2^{1-\epsilon} \frac{\Gamma(1 - \epsilon/2)}{\Gamma(\epsilon/2)}, \tag{3}$$

which is equal to $3.35(k_\theta \theta_0)^{0.8}$ for $\epsilon \simeq 0.8$.

As previously stated, this holds for correlations of galaxy *locations*. On the other hand, the power spectrum of background fluctuations, which is what we calculate, measures correlations of galaxy *intensities*. The redshift distribution of the cumulative flux contributed by the background sources is represented as

$$\frac{d\mathcal{S}}{dz} = \int_0^\infty S \frac{dN}{dS\,dz}(S, z)\, dS, \qquad (4)$$

where $dN/(dS\,dz)$ is the number density of sources per unit flux density and redshift interval. The measured power spectrum of background fluctuations, $P(k_\theta)$, is the two-dimensional, *flux weighted*, projection of the three-dimensional galaxy clustering spectrum, $P_{3D}(k_\theta)$. For $k_\theta \gg 1$ and a flat cosmology, their relationship can be approximated as

$$P(k_\theta) = \int_{z_{\min}}^{z_{\max}} P_{3D}(2\pi k_\theta/x(z), z) \left(\frac{d\mathcal{S}}{dz}(z)\right)^2 \frac{1}{dV_c(z)}\, dz, \quad (5)$$

(e.g., Tegmark et al. 2002), where $x(z)$ is the comoving radial distance and $dV_c(z)$ is the comoving volume element, i.e., $dV_c(z) = x(z)^2 dx/dz$ (assuming a flat cosmology), where $dN/(dS\,dz)$ is the number density of sources per unit flux density and redshift interval.

$P(k_\theta)$ can expressed in dimensionless units as

$$\Delta_{k_\theta}^2 \equiv 2\pi k_\theta^2 P(k_\theta)\big/ I_\nu^2, \qquad (6)$$

where I_ν is the intensity of the background in Jy. Although $\Delta_{k_\theta}^2$ and $\tilde{\Delta}_{k_\theta}^2$ have the same units and similar meaning, they differ because Equation 6 deals with an intensity weighted power spectrum.

In addition to the clustering signal, the total power spectrum has contributions from instrumental noise, Poisson noise from individual background galaxies, and cirrus emission. We will address each contribution individually.

3. METHODS

3.1. Map Preparation

BLAST observed a wide 8.7 deg^2 patch, centered on the GOODS-South field ($3^h32^m35^s$, $-28°15'$; hereafter BGS-Wide), with mean 1σ sensitivities of 36, 31, and 20 mJy beam^{-1} at 250, 350, and 500 μm, as well as a deep, nested field of 0.8 deg^2, centered on ($3^h32^m30^s$, $-27°48'$; hereafter BGS-Deep) with mean 1σ sensitivities of 11, 9, and 6 mJy beam^{-1}, respectively.[16] A 6 deg^2 region, centered on BGS-Wide, is selected from the map, because of its uniformity in observed depth. The BLAST bolometers are prone to drifts on timescales greater than ~ 10 s. To retain as much large angular scale signal as possible, a fast scan rate is preferred because larger scales will survive the high-pass filtering, designed to remove noise below the 0.1 Hz $1/f$ knee. For this analysis, we use only the data from the wide region of the map, where the scan rate is 0.1 deg s^{-1}, and the rms of the maps is 2.4, 1.9, 1.0 MJy sr^{-1}, at 250, 350, and 500 μm, which is applicable to calculating uncertainties of *point-source* flux densities. The data for the nested deep region, whose scan rate is only 0.05 deg s^{-1}, are not included. Large-scale noise is removed by high-pass filtering the time-streams at 0.2 Hz. Correlated noise—a drift of multiple detectors in unison—is not removed because it would inevitably
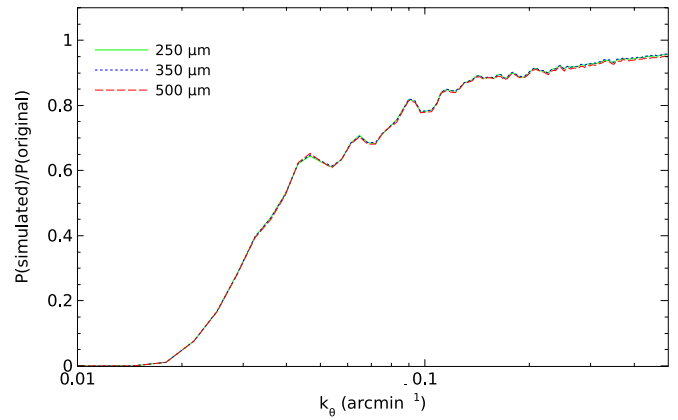
---

[16] BLAST maps and catalogs are publicly available at http://www.blastexperiment.info

Figure 1. Transfer functions calculated with a Monte Carlo simulation involving 500 mock-maps observed with the BLAST simulator. The lines are identical beyond the scale of the beams, which implies that the map-maker is linear, as we would expect.

(A color version of this figure is available in the online journal.)

suppress large-scale signal as well. Instead, a cross-correlation of a subset of the maps is used to remove large-scale correlated noise (described below). Although fully optimized map-makers are available (e.g., SANEPIC; Patanchon et al. 2008), the long time to convergence (typically 24 hr on 10 processors for this particular map) makes it impractical to use them for our Monte Carlo simulations. Full analysis with SANEPIC maps, including the nested deep field, will be the subject of a future paper. Here, we instead use OPTBIN (Pascale et al. 2009) a fast, naive, map-maker whose transfer function is calculated using a Monte Carlo simulation (see Section 3.2 and Figure 1 for details), and we use SANEPIC maps for consistency checks. Parallel power spectrum analysis with both map-makers show agreement well within the simulated uncertainties.

3.2. Power Spectrum Calculation

The map intensity, S_{map}, can be written as

$$S_{\mathrm{map}} = (T \otimes [S_{\mathrm{sky}} \otimes B + N])W, \qquad (7)$$

where we use $\otimes$ to represent a convolution, S_{sky} is the true sky surface brightness, T is the transfer function of the map-maker, B is the measured instrumental beam, N is the instrumental noise, and W is the "aperture function," which is zero beyond the region of interest. The autocorrelation of a map will contain a contribution from detector noise. To suppress this instrumental noise, cross-power spectra are taken among a set of four maps which are made by dividing the time-stream into four roughly equal parts and then making four separate maps (hereafter referred to as sub-maps). The time-streams, which are made up of numerous chunks, are divided into every fourth chunk (e.g., 1, 5, 9,..., and 2, 6, 10,..., etc.), so that the sub-maps have as similar coverage as possible. The number of sub-maps chosen maximizes the number that can be made while maintaining uniformity in hits, retaining some cross-linking, and avoiding holes in the maps. The rms of the resulting sub-maps is 4.6, 3.6, and 2.0 MJy sr^{-1}, at 250, 350, and 500 μm. In the cross-spectrum, noise which is uncorrelated between sub-maps averages to zero. Consequently, the spectrum does not depend on modeling the potentially complicated or non-stationary noise.

We prepare the maps before calculating the power spectrum by removing their means, apodizing them with a Welch window (Press 2002, Chapter 13.4), and zero-padding them with a

width on each side equal to half the map. The cross-correlation two-dimensional power spectrum of each pair of maps is calculated. The azimuthal average of the amplitudes (which in two-dimensional *k*-space appears to be isotropic) is taken to find the one-dimensional power spectrum, $P(k)$. The resulting spectra are averaged and divided by the power spectrum of the beam and the transfer function. The transfer function is calculated with a Monte Carlo simulation from simulated maps made with the BLAST simulator,[17] and is shown in Figure 1. At angular scales, which are large compared to the high-pass filter (<0.03 arcmin^{-1}) and approaching the cutoff of the beam (>0.9 arcmin^{-1}), the transfer function is unreliable. We are interested in the scales bracketed by these limits.

3.2.1. Jackknife Tests

We have performed jackknife tests in which a correlation is calculated between two distinct difference maps. Specifically, we take the cross-correlation of: (sub $-$ map 1 $-$ sub $-$ map 2) and (sub $-$ map 3 $-$ sub $-$ map 4). If the cross-correlation measures only signal, then taking a difference between sub-maps should cancel sky signal and results in a cross-correlation power spectrum consistent with zero. And indeed, our results are consistent with zero, in the range of interest, in all three bands.

As a further check, the cross-power spectrum is compared to the difference of the auto-power spectrum (of a map made from the entire time-stream) and the measured noise. The noise is estimated from odd–even pixel jackknife maps, and is approximately white over the angular scales of interest. The resulting two spectra are in excellent agreement.

3.2.2. Poisson Noise

Poisson (or shot) noise arises from the finite number of galaxies per unit area. It can be calculated analytically from the source counts as

$$P_{\text{shot}} = \int_0^\infty S^2 \frac{dN}{dS} dS. \tag{8}$$

Alternatively, Poisson noise levels can be estimated from Monte Carlo simulations using mock-maps that are populated with uncorrelated sources whose fluxes are drawn from the measured BLAST counts (Patanchon et al. 2009). This has the added advantage that realistic uncertainties can be calculated as well. Since the counts are so steep, care must be taken to reproduce the bright end of the counts faithfully, so that extremely rare bright sources do not unrealistically appear in the realizations. We find that the two methods of estimating the level of shot noise agree within the 1σ uncertainties.

Due to the steep nature of the source counts, the Poisson noise is dominated by the contributions of the fainter population. Furthermore, Marsden et al. (2009) find only 15% of the total sky intensity is associated with a 3σ catalog. We find that removal of only the brightest sources results in an approximately 5% reduction in Poisson noise at 250 μm, and that more aggressive cuts lead to removal of correlations beyond just Poisson noise. We find behavior consistent with this in our simulations. Therefore, we subtract only five sources above 500 mJy at 250 μm, two sources above 400 mJy at 350 μm, and no sources at 500 μm. To subtract sources, first we make a

---

[17] We simulate the BLAST pipeline by "observing" an input map, creating a set of time-streams, filtering them, and making a map from the filtered time-streams.
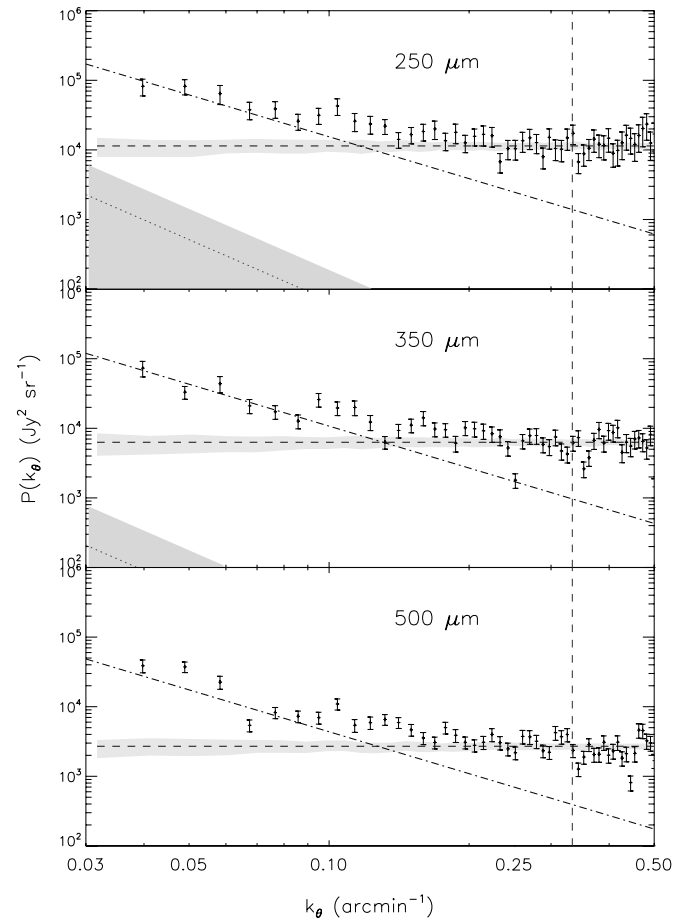

Figure 2. Power spectra of 6 deg^2 regions selected from the BGS-Wide maps at 250, 350, and 500 μm are shown with 1σ uncertainties. Color-corrected fits to galactic cirrus measured from IRIS 100 μm maps are shown as dotted lines with shaded region representing 1σ uncertainties, sloping down from the left side (500 μm cirrus is too low to make it onto the region plotted). Poisson noise contributions are found with Monte Carlo simulations, and are shown as horizontal dashed lines with similar error regions. Scale invariant, k^{-2}, power spectra are shown as dot-dashed lines. Vertical dashed line at $k_\theta = 0.33$ arcmin^{-1} represents the scale at which the variance in the FIDEL catalog—which Marsden et al. (2009) show resolves most of the CIB—is no longer Poisson noise dominated. For angular scales greater than 0.33 arcmin^{-1}, signal in excess of Poisson noise is attributed to the clustering of star-forming galaxies.

source list by performing a noise-weighted convolution of the maps with the effective BLAST point-spread function (PSF) and identify local maxima in the smoothed map. We then subtract a scaled effective PSF with amplitude taken from the source list. We perform the same operation on our mock-maps, from which we calculate Poisson levels of $11.4 \pm 1.0 \times 10^3$, $6.3 \pm 0.5 \times 10^3$, and $2.7 \pm 0.2 \times 10^3$ Jy^{2}sr^{-1} at 250, 350, and 500 μm. These are shown as dashed lines in Figure 2.

3.2.3. Estimates of Galactic Foregrounds

Gautier et al. (1992) show that the power spectrum of Galactic cirrus can be approximated by a power law,

$$P_{\text{cirrus}}(k_\theta) = P_0 \left(\frac{k_\theta}{k_0} \right)^{-\alpha}, \tag{9}$$

where k_θ is the angular wavenumber in inverse arcminutes, and P_0 is the power spectrum value at $k_0 = 0.01$ arcmin^{-1}. We measure the cirrus component at 100 μm from the cross-correlation

Table 1
CIB Normalized Clustering Power Spectra, $P(k_\theta)/I_\nu^2$

k_θ (arcmin^{-1})	BLAST 250 (steradian^{-1})	BLAST 350 (steradian^{-1})	BLAST 500 (steradian^{-1})
0.044	$(1.39 \pm 0.52) \times 10^{-7}$	$(1.33 \pm 0.47) \times 10^{-7}$	$(2.47 \pm 1.17) \times 10^{-7}$
0.070	$(5.94 \pm 2.02) \times 10^{-8}$	$(4.98 \pm 1.83) \times 10^{-8}$	$(5.71 \pm 2.06) \times 10^{-8}$
0.113	$(2.99 \pm 0.95) \times 10^{-8}$	$(2.62 \pm 0.90) \times 10^{-8}$	$(2.98 \pm 1.00) \times 10^{-8}$
0.183	$(9.81 \pm 7.51) \times 10^{-9}$	$(9.94 \pm 6.92) \times 10^{-9}$	$(6.91 \pm 4.62) \times 10^{-9}$

Notes. The errors do not include uncertainties in the CIB. CIB values are listed in Section 4.

of the three co-added IRIS (HCON[18]) maps (reprocessed *IRAS*: Miville-Deschênes & Lagache 2005) for a ~ 15 deg^2 region surrounding the BGS-Wide field. The amplitude of the observed power spectrum has a contribution from cirrus emission, which is highly variable on the sky. The GOODS region was specifically chosen because it is a low cirrus region, with a mean intensity of 1.39 ± 0.18 MJy sr^{-1}. At scales $0.008 < k_\theta < 0.03$ arcmin^{-1}, which is larger than the scales probed by BLAST, we find that the cirrus is well approximated by a power law, $P_0 = (0.47 \pm 0.18) \times 10^6$ Jy^{2}sr^{-1}and $\alpha = 2.91 \pm 0.11$, values which are similar to those found in low cirrus emission regions measured by Lagache et al. (2007) and Miville-Deschênes et al. (2007). We assume that the power spectrum continues to smaller angular scales, and scale it to the BLAST bands using the average dust emission color $(I_{BLAST}/I_{100})^2$, which is found by assuming cirrus emission behaves as a modified blackbody, $\nu^\beta B(\nu)$, where $B(\nu)$ is the Planck function, and β is the emissivity index (Draine & Lee 1984). The scaled power spectrum and errors are calculated with a Monte Carlo simulation, varying temperature (17.5 ± 1.5 K) and β (1.9 ± 0.2) (Boulanger et al. 1996). The resulting power-law approximations and uncertainties are illustrated as dotted lines in Figure 2, which show that we are negligibly affected by Galactic cirrus on all recovered angular scales.

3.2.4. Uncertainties

To estimate uncertainties in our angular power spectra, clustered signal plus noise simulated maps are analyzed with the same pipeline as the astronomical data. We include clustering in the simulations in case the amplitude or shape of the transfer function depends on the input, but we found that not to be the case. We follow Almaini et al. (2005, see appendix for algorithm) to introduce correlations to the simulated maps, with an angular correlation length, $\theta_0 = 3''0$. This angle was chosen from the measured upper limit of Peacock et al. (2000). Realizations with stronger clustering clearly do not *look* like BLAST maps.

4. BASIC CLUSTERING RESULTS

Figure 2 shows an unambiguous signal in excess of Poisson noise on scales of 0.04–0.2 arcmin^{-1}, which cannot be explained by Galactic cirrus, and which we interpret as correlations from clustered star-forming galaxies. The vertical dashed line indicates the angle at which the distribution of 24 μm selected FIDEL galaxies (Magnelli et al. 2009) begin to show variance in excess of Poisson. Indeed, the similarity to Figure 3 of Marsden et al. (2009) is striking.
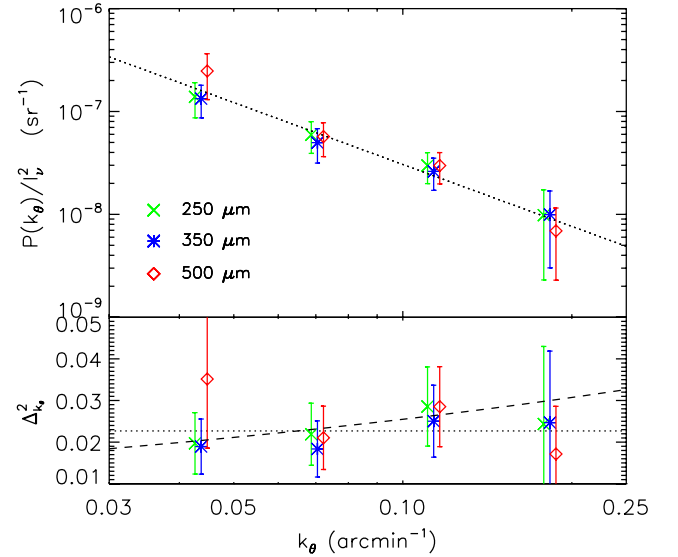
Figure 3. Top panel: power Spectra of the clustering component, after removal of cirrus and Poisson noise, and normalized by $(I_\nu^{CIB})^2$. The best-fit power spectrum, proportional to k^{-2}, is shown as a dotted line. Bottom panel: $\Delta_{k_\theta}^2$ shown along with best-fit power spectra (horizontal dotted line); as well as best-fit parameters to Equation 3 (dashed line) for 250 μm data. The 250 and 500 μm points are offset horizontally for clarity, by factors of -0.025 and $+0.025$, respectively. Clearly, the power spectrum signal of clustered star-forming galaxies is well fit by a power-law power spectrum proportional to k^{-2}.

(A color version of this figure is available in the online journal.)

The CIB has an amplitude I_ν^{CIB} measured to be 0.71 ± 0.17, 0.59 ± 0.14, and 0.38 ± 0.10 MJy sr^{-1} at 250, 350, and 500 μm, respectively (Marsden et al. 2009). Figure 3 shows the clustering component of the power spectrum normalized by I_ν^{CIB} in the three BLAST bands, where the 250 μm and 500 μm data have been displaced slightly for visual clarity. The contributions of cirrus and Poisson noise have been subtracted, and the data rebinned in logarithmic intervals. These results are also listed in Table 1. We use the BLAST estimates for the CIB because they are the most precise estimate available of the CIB in these wavelength bands. Doing so has the additional benefit in this case that calibration uncertainties completely vanish in the ratio. The fit to a single power law, and also the agreement in amplitude across the BLAST bands, once the power spectra are normalized to the sky intensity, are excellent.

The relative variance of the CIB, $\Delta_{k_\theta}^2$, formed by dividing $P(k_\theta)$ by $2\pi k_\theta^2$, is shown in the bottom panel of Figure 3, where the 250 μm and 500 μm data have been displaced slightly for visual clarity. (Amplitudes have been additionally converted from arcmin^{-2} to steradian^{-1} in the figure and in the discussion below.) The best fits of Equation (3) are shown in Table 2. The data are consistent (to within 1σ) with $\Delta_k^2 = $ Constant, with the same amplitude in all three bands, even taking into
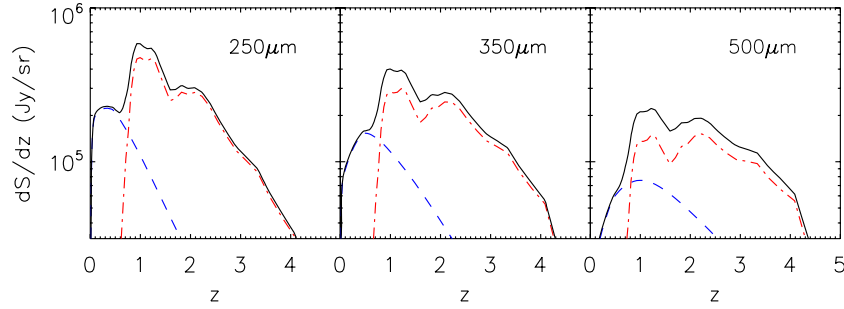
Figure 4. Redshift distribution of the cumulative flux contributed by the background sources at the BLAST bands, according to the Lagache et al. (2004) model. The dashed (blue) and dot-dashed (red) curves are for "regular" and "star-forming" *IRAS* galaxies, respectively, while the solid line is the total.

(A color version of this figure is available in the online journal.)

Table 2
Best-fit Values Obtained for θ_0 and ϵ

BAND	θ_0 (arcsec)	ϵ
$250\,\mu$m	0.017 ± 0.020	0.27 ± 0.19
$350\,\mu$m	0.008 ± 0.006	0.26 ± 0.19
$500\,\mu$m	0.0	0.0

account that the error bars shown contain a large component of uncertainty which is common mode between channels, and thus overestimate the anticipated scatter. The best-fit amplitude for a power law with slope of -2 is shown as a dotted line in the bottom panel, corresponding to $\Delta_{k_\theta} = \delta I/I = 15.1\% \pm 1.7\%$. This is directly analogous to the square root of the "band power" measured in cosmic microwave background (CMB) anisotropy experiments, where typically $\delta T/T \sim 10^{-5}$. From this point of view the CIB is much clumpier than the CMB, as one would expect since the galaxies are observed after a much longer period of linear growth. The dotted line in the upper panel of Figure 3 is $2\pi k_\theta^2 \Delta_{k_\theta}^2$, corresponding exactly to the fit in the lower panel.

The values of the parameters ϵ and θ_0 (see Equation (3)) that best fit the data are shown in Table 2. The large uncertainty quoted for θ_0 is partly due to the awkwardness of this parameterization near zero slope. The bottom panel of Figure 3 shows the best fit for $250\,\mu$m data as a dashed line. Such a small θ_0, despite significant power in excess of Poisson noise, requires that the sources which make up the CIB are distributed over a wide range of redshifts, and has implications for future clustering measurement strategies (e.g., with *Herschel* or *Planck*), which we discuss in Section 11.

We have calculated the correlations between different bands (e.g., 250 × 350) using the same pipeline and set of sub-maps. The cross-spectra are normalized by the square root of the auto-spectra of the two bands, so that the final curve would be unity at all scales for identical maps, and zero at all scales for two completely different maps. Results show that the cross-correlations are 0.95 ± 0.06, 1.06 ± 0.09, and 0.92 ± 0.04, for 250 × 350, 350 × 500, and 250 × 500, respectively, over the range of angular scales $0.04 < k_\theta < 0.5$ arcmin^{-1}. Neighboring bands are more correlated with each other than are the 250 and $500\,\mu$m bands. While we find the same spectrum in all three bands, the phases are different, as one would expect if the three BLAST bands have different selection functions and sample the galaxies in the CIB at different redshifts. While the cross-band correlation provides a powerful tool for testing the redshift distributions and spectral energy densities of source population models, a more detailed analysis is required before making any strong conclusions. This will be the subject of a future study using the SANEPIC maps and including the BGS-Deep data.

We have measured the variance projected onto two dimensions, but galaxies are of course distributed in three dimensions. Knowledge of the redshift distribution of the galaxies in the CIB allows interpretation of these power spectra in terms of a bias factor quantifying the comparison to cold dark matter (CDM) spectra, which we find in the next section. As the angular scales we probe get smaller, non-linear growth eventually sets in, which we examine by fitting halo models to these spectra later in the paper.

PART II. MODEL FITTING

5. INTRODUCTION

To interpret our detection of correlations requires comparison to an underlying model whose parameters the data constrain. Such a model could contain details of the complete source population, including number counts, i.e., intensity distributions and redshift distributions, as well as a framework for describing the linear and non-linear clustering regimes. This latter part might involve, for example, a HOD that accounts for the galaxy distribution in a given dark matter halo as a function of luminosity (i.e., the so-called conditional luminosity function; Cooray 2006). For a background of primarily unresolved sources, the conditional luminosity function model is an improvement on the simple HOD; however, its increase in complexity comes with an increase of free parameters.

Given the very good fit to a constant $\Delta_{k_\theta}^2$, we first explore the physical meaning of the BLAST power spectra in the context of a purely linear model. We assume that the galaxies which comprise the CIB have a power spectrum which is scaled from the power spectrum of dark matter at every redshift, $P(k) = b^2 P_{\rm DM}(k)$. Because $P_{\rm DM}$ is redshift dependent, estimating b^2 requires knowledge of the redshift distribution of the galaxies, which comprise the BLAST signals. To find this distribution, we adopt the model of Lagache et al. (2004), described in Section 6. It is worth noting that since we are really measuring an emission-weighted bias at rest-frame far-IR wavelengths, there will be some degeneracy between the value of b and changes in the redshift distribution or far-IR spectral shapes assumed for the sources.

We extend the fit into the non-linear regime to study whether the BLAST data can constrain parameters of the HOD. We can check whether our correlations are consistent with the model (e.g., by judging whether the same bias fits all three BLAST wave bands, and that the cross-band correlations of simulations made with the model agree with the data), but we will not explore how the model might be improved. It is important to understand that if the distribution of source redshifts or counts were different, then we would infer a different bias level.
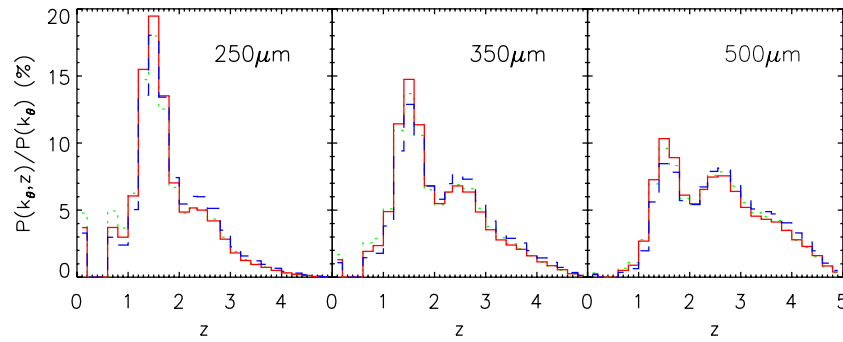
Figure 5. Contribution to the angular power spectrum from different redshift slices, inferred from Figure 4, at specific angular scales probed by the data. The different curves correspond to the following scales: dotted (green)—$k_\theta = 0.04$ arcmin^{-1}; solid (red)—$k_\theta = 0.1$ arcmin^{-1}; dashed (blue)—$k_\theta = 0.18$ arcmin^{-1}. Whereas the redshift distribution of sources (Figure 4) has a significant contribution from $z < 1$, the contribution from those sources to the angular power spectrum is negligible.

(A color version of this figure is available in the online journal.)

However, since the model we have adopted agrees with a large body of multi-wavelength observations, we are confident that these prescriptions are a reasonable first attempt.

6. SOURCE POPULATION MODEL

Knowledge of the full redshift distribution of BLAST sources would allow us to estimate the redshift distribution of the measured clustering signal, i.e., the redshift range probed by the power spectrum of correlations due to clustering described by the $P \times dS/dz$ distribution. Devlin et al. (2009), and in more detail Patanchon et al. (2009), find the number counts of the BLAST sources, and Pascale et al. (2009) divides the redshift distribution roughly into four redshift bins. Precisely understanding the redshift distribution of those sources is a work in progress (e.g., E. L. Chapin et al. 2010, in preparation). In the meantime, we adopt the model of Lagache et al. (2004)—a model which approximates the counts and redshift distributions of the populations expected to make up the CIB—to describe the underlying source population.[19] Specifically, it is a phenomenological model which extrapolates the local $60\,\mu$m luminosity function of *IRAS* sources (divided into "regular" and "star-forming" components) to longer wavelengths, assuming a set of spectral energy distribution templates. The extrapolation to higher redshift is parameterized as a mixture of luminosity and density evolution, and is constrained to be consistent with most of the available data on source counts, redshift distributions, and the far-infrared background intensity. It becomes less reliable with greater extrapolation to longer wavelengths, and therefore should be considered an approximation; however, it does reproduce the BLAST (Patanchon et al. 2009) counts at a level which is sufficient for our purposes. Furthermore, since the clustering signal is dominated by the faint source population, it is critical that the model is consistent with the level of the background at the BLAST wavelengths, and that the cross-band correlations of simulated maps agree with those measured, which we find to be the case.

The redshift distribution of the source population model at each wavelength is shown in Figure 4. Most of the clustering signal is coming from relatively high redshifts. The medians of the redshift distributions in Figure 5 are $z = 1.61$, 1.88, and 2.42, at 250, 350, and 500 μm, respectively, and the upper and lower quartiles of the distributions are $z = (1.3, 2.2)$, $(1.5, 2.7)$, and $(1.7, 3.2)$, respectively. At the representative scale of 0.1

---

arcmin^{-1} (red solid line in Figure 5), 95% of the background originates from sources $1.1 < z$, $1.2 < z$, and $1.4 < z$ in the three bands.

7. LINEAR BIAS MODEL

We first carry out a simple fit to a scaled version of the linear theory power spectrum of the dark matter P_{DM}. As can be seen in Figure 6, a simple biasing prescription provides a good fit to the BLAST data. The required bias levels are $b = 3.8 \pm 0.6$, 3.9 ± 0.6, and 4.4 ± 0.7 at 250, 350, and 500 μm, respectively, with a reduced $\chi^2_{min} \sim 0.4$ (with 10 degrees of freedom) in all three bands.

More detailed modeling could be attempted. In principle, the cross-band measurements and wavelength dependence of the measured correlation amplitudes could be used to estimate the variation of bias with redshift, as discussed by Knox et al. (2001). However, we leave this to a future study.

8. HALO MODEL

As illustrated in Figure 6, a simple biasing prescription provides a good fit to our data. The main drawback of this simple fit is that it may not realistically account for the 1-halo, non-linear clustering component. Our data are expected to bracket the physical scales corresponding to the transition from linear to non-linear clustering regimes, and though these two contributions appear to have combined to look very much like a scaled linear bias, to conclude that there is no contribution from a 1-halo term may be unphysical. Therefore, in this section we use a particular implementation of the "halo model," which assigns galaxies to halos as a function of mass, and consequently probes into the territory of non-linear fluctuations, to explore how the 1-halo term could contribute power on small scales. This also allows us to discuss our results in the context of other measurements of galaxy clustering.

The halo model of large-scale structure has proven to be a powerful tool for describing the clustering properties of cosmic objects (for a review, see Cooray & Sheth 2002). Its main ingredient is the parameterization of the *HOD* (Peacock & Smith 2000), which describes how galaxies populate dark matter halos as a function of halo mass. The power spectrum of galaxies is written as the sum of two components: the 1-halo term, P_{1h}, which describes pairs of objects within the same dark matter halo, and the 2-halo term, P_{2h}, which accounts for pairs of objects in different halos, resulting in $P(k, z) = P_{1h}(k, z) + P_{2h}(k, z)$. The number of pairs of galaxies

---

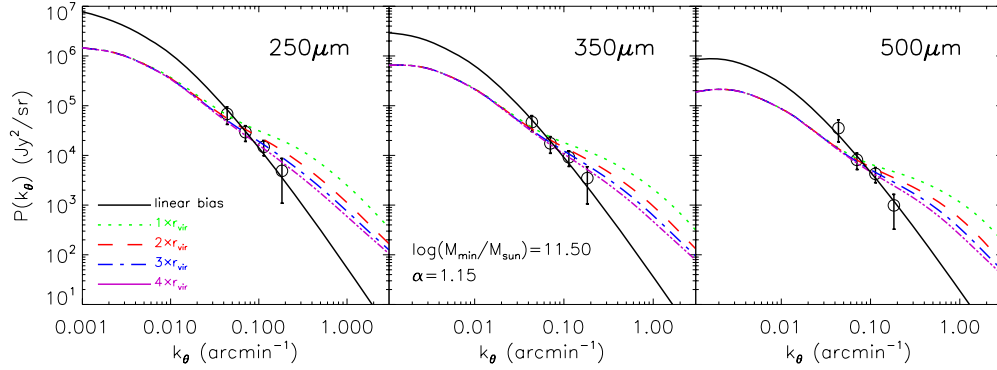[19] Documentation and *IDL* files can be found at http://www.ias.u-psud.fr/irgalaxies/model.php

Figure 6. Power spectrum of background correlations (circles with error bars) overlaid with the best-fit linear bias (solid line), as well as predictions obtained for halo models different values of r_{cut}. See Table 3 for the fit parameters. Although the power spectra shown here have different amplitudes in the three channels, $P(k_\theta)/I_\nu$ is the same in all three bands, as shown in Figure 3. The data are fitted best by a model which includes a linear term only.

(A color version of this figure is available in the online journal.)

within an individual halo is related to the variance of the HOD, $\sigma^2(M, z) = \langle N_{gal}(N_{gal} - 1)\rangle$, while the number of pairs of galaxies in separate halos is simply the square of the mean halo occupation number, $N(M, z) = \langle N_{gal}\rangle$, (HON, hereafter). We model the HON using a central-satellite formalism (see, e.g., Zheng et al. 2005): this assumes that the first galaxy to be hosted by a halo lies at its center, while any remaining galaxies are classified as *satellites* and are distributed in proportion to the halo mass profile. Different HODs for central and satellite galaxies are then applied. For central galaxies, the mean HON, N_{cen}, is described by a step function such that halos above a minimum mass threshold M_{min} contain a single central galaxy and halos below this threshold contain no galaxies. For satellite galaxies, a power law in mass describes their mean HON (e.g., Zehavi et al. 2005)

$$N_{sat}(M) = \left(\frac{M}{M_1}\right)^\alpha, \quad (10)$$

where M_1 is the mass-scale at which a halo hosts exactly one satellite galaxy (in addition to the central galaxy). Both semi-analytic models (e.g., Berlind et al. 2000) and hydrodynamic simulations (e.g., Berlind et al. 2003) show that the distribution of galaxies within a halo is close to Poisson in the high-occupancy regime, i.e., when $N_{sat} \gg 1$, and (strongly) sub-Poissonian in the low-occupancy regime. In order to agree with these results, satellite galaxies are assumed to be Poisson distributed at fixed halo mass. The distinction between central and satellite galaxies then automatically accounts for the sub-Poissonian behavior of the HOD in the low-occupancy regime (Zheng et al. 2005).

The 1- and 2-halo power spectra are

$$P_{1h}(k, z) = \int_{\mathcal{M}} n_{halo}(M, z)\big[2N_{cen}(M)N_{sat}(M)u_{DM}(k, z|M)$$
$$+ N_{sat}^2(M)u_{DM}^2(k, z|M)\big]dM\big/n_{gal}^2(z),$$

$$P_{2h}(k, z) = P_{DM}(k, z)\bigg[\int_{\mathcal{M}} n_{halo}(M, z)N_{gal}(M, z)$$
$$\times b(M, z)u_{DM}(k, z|M)dM\bigg]^2\bigg/n_{gal}^2(z). \quad (11)$$

The meaning of the symbols here is as follows: P_{DM} is the linear power spectrum of dark matter, derived using the recipes of Eisenstein & Hu (1998) for the matter transfer function; n_{halo}

is the halo-mass function (see Sheth et al. 2001); b is the linear bias parameter; u_{DM} is the normalized dark matter halo density profile in Fourier space; and n_{gal} is the mean number of galaxies per unit comoving volume at redshift z,

$$n_{gal}(z) = \int_{\mathcal{M}} n_{halo}(M, z)\left[1 + \left(\frac{M}{M_1}\right)^\alpha\right]dM. \quad (12)$$

The expression for the 1-halo term implicitly assumes that the distribution of galaxies traces that of the dark matter, for which we have adopted the profile of Navarro et al. (1997, ; hereafter NFW), with the same concentration parameter as Bullock et al. (2001). Since the NFW profile formally extends to infinity, it is necessary to artificially truncate the distribution at some radius, r_{cut}. Typically, this is chosen to be the virial radius of the halo; however, this may not necessarily be realistic. We address this by first adopting the assumption that $r_{cut} = r_{vir}$, and then exploring the consequences of relaxing that requirement, so that galaxies are allowed to lie further out. On large scales, where clustering is predominantly linear, $u_{DM} \sim 1$, so that the 2-halo power spectrum simplifies to $P_{2h} = b_{eff}^2(z)P_{DM}(k, z)$, where $b_{eff}(z)$ is the effective large-scale bias,

$$b_{eff}(z) = \int_{\mathcal{M}} n_{halo}(M, z)N_{gal}(M)b(M, z)\, dM/n_{gal}(z). \quad (13)$$

Our model has two free parameters, M_{min} and α, which we vary through $0 \leqslant \alpha \leqslant 2$ and $10 \leqslant \log(M_{min}/M_\odot) \leqslant 16$, with steps of 0.05 in both $\log M$ and α. The best-fit values of the parameters are determined through a χ^2 minimization technique by fitting the observed power spectrum at each of the three BLAST bands simultaneously.

Throughout we assume that both M_{min} and α remain constant in time, although, in principle, they are functions of redshift. Whether these parameters evolve with redshift would be difficult to constrain from our data alone; nevertheless, our assumption is consistent with what is observed for other classes of high-redshift sources (e.g., quasars, see Porciani et al. 2004). For each M_{min}–α pair, the mass-scale M_1 is fixed by requiring that at every redshift, z, the number density of the background sources derived from the halo model formalism matches that predicted by the adopted source population model, i.e.,

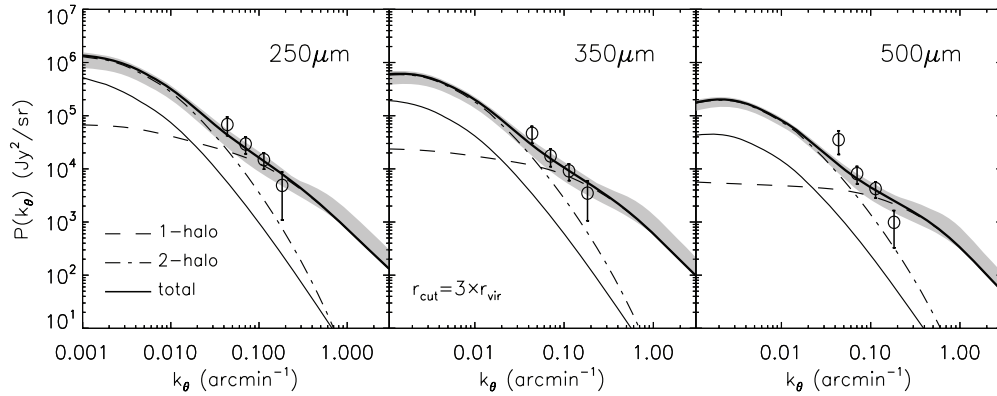$$\int_0^\infty \frac{dN}{dS\, dz}(S, z)\, dS = n_{gal}(z)\, dV_c(z). \quad (14)$$

Figure 7. Power spectrum of background correlations from clustering of extragalactic sources measured in the BLAST maps (circles with error bars) overlaid with the best-fit halo model (thick solid line), under the assumption that dark matter halos are NFW spheres truncated to $3\times$ the virial radius, and the galaxies within the halo follow the underlying dark matter distribution. The shaded region shows the 99% confidence region in the $M_{\rm min}$–α parameter space. The dashed and the dot-dashed curves show the 1- and 2-halo contributions to the power spectrum, respectively. For comparison, the power spectrum obtained under the assumption that galaxies are unbiased tracers of the underlying dark matter distribution, i.e., $P_{\rm 3D}(k, z) = P_{\rm DM}(k, z)$, is plotted as a lighter solid line. According to the model, the BLAST data occupy a range of angular scales that should be sensitive to both the linear and non-linear clustering terms.

9. HALO MODEL FITS

Under the assumption that the dark matter halos are described by an NFW profile truncated at the virial radius, i.e., where $r_{\rm cut} = r_{\rm vir}$, and that galaxies within the halo trace the underlying dark matter distribution, the best fit to the observed angular power spectrum gives $\log(M_{\rm min}/M_\odot) = 11.5^{+0.4}_{-0.1}$ and $\alpha \leqslant 1.0$, with $\chi^2_{\rm min} = 16.3$ with 10 degrees of freedom; i.e., the model is marginally consistent at the 2σ level.

While this model is not formally ruled out by the data, as shown by a dotted line in Figure 6, it poorly reproduces the shape of the observed power spectrum, which exhibits a steeper slope. If we were to weight this model to fit the large-scale power spectrum (e.g., $k \lesssim 0.08$ arcmin^{-1}), then it would over-predict the power on small scales. Arguing that perhaps the model describes the small scale, 1-halo term correctly, but is under-predicting the large scales, is not a good description because: (1) the 2-halo term is less sensitive to the underlying assumptions than the 1-halo term (for a discussion see Tinker et al. 2009); and (2) the observed small-scale power spectrum is still too steep to be accounted for by the shallower 1-halo term. Another possibility is that the redshift distribution of the background cumulative flux predicted by the adopted source count model is incorrect. In order to reproduce the observed shape of the angular power spectrum, the bulk of the background would have to originate from sources at $z < 1$, which is ruled out by Devlin et al. (2009); Marsden et al. (2009); and Pascale et al. (2009). However, a full investigation of the leeway in changing the redshift distribution (and degeneracies with other changes to the model) are beyond the scope of the present study.

In light of this, we explore the possibility that the discrepancy between the predicted power spectrum and the observed one is related to the modeling of the 1-halo term. There are two obvious ways to modify the shape and normalization of the 1-halo power spectrum. One is to allow the dark matter halos, although still following an NFW profile, to be truncated at a scale $r_{\rm cut} > r_{\rm vir}$. Thus, satellite galaxies are distributed over a larger volume. This idea is not new; Magliocchetti & Porciani (2003) find that in order to adequately fit the 1- and 2-halo term to the 2dF Galaxy Redshift Survey data set, it is necessary that the galaxies are allowed to reside out to twice the virial radius. Furthermore, from semi-analytic models Diaferio et al. (1999) show that blue (and hence star-forming) galaxies tend to reside

Table 3
Best-fit Values Obtained for $M_{\rm min}$ and α for Different Choices of the Radius $r_{\rm cut}$

$r_{\rm cut}/r_{\rm vir}$	$\chi^2_{\rm min}$	$\log(M_{\rm min}/M_\odot)$	α
1	16.3	$11.50^{+0.40}_{-0.05}$	$0.95^{+0.05}_{-0.95}$
2	13.6	$11.50^{+0.40}_{-0.05}$	$1.00^{+0.10}_{-1.00}$
3	11.5	$11.50^{+0.40}_{-0.05}$	$1.10^{+0.05}_{-1.10}$
4	9.7	$11.50^{+0.40}_{-0.05}$	$1.15^{+0.05}_{-0.75}$

Note. The minimum-χ^2 (with 10 degrees of freedom) are also shown.

in the outskirts of their host halos, while red galaxies are found closer to the halo center.

The second possibility is that the distribution of galaxies within the halos does not follow that of the underlying dark matter. For example, a power-law distribution $\rho(r) \propto r^{-\gamma}$ with $\gamma < 2$ would make the 1-halo angular power spectrum steeper than that predicted by an NFW profile. Similar arguments have been made by Watson et al. (2009), who found that by allowing the inner slope of the density profile to vary, they fit the small-scale clustering of luminous red galaxies quite well.

Here we only explore the first possibility, examining $r_{\rm cut} = 1, 2, 3$, and $4 \times r_{\rm vir}$. The results are shown in Figure 6, and summarized in Table 3. The best-fit model angular power spectrum for the case $r_{\rm cut} = 3 \times r_{\rm vir}$ is shown in Figure 7, while the corresponding HON and large-scale effective bias are shown in Figure 8. Note that while the value of the reduced $\chi^2_{\rm min}$ approaches unity for increasing values of $r_{\rm cut}$, the best-fit values of $M_{\rm min}$ are negligibly affected by the changes, while α only marginally increases with increasing $r_{\rm cut}$. The effective mass of the halo, $M_{\rm eff}$, is the weighted mean over the halo-mass distribution:

$$M_{\rm eff}(z) = \int_{\mathcal{M}} n_{\rm halo}(M, z) N(M, z) M \, dM / n_{\rm gal}(z). \quad (15)$$

The results for the large-scale effective bias b, and mass $\log(M/M_\odot)$, at the respective medians of the redshift distributions of the sources contributing to the background in each band (see Figure 5) are 2.2 ± 0.2, 2.4 ± 0.2, and 2.6 ± 0.2; and 12.9 ± 0.3, 12.8 ± 0.2, and 12.7 ± 0.2, at 250, 350 and 500 μm, respectively. They are minimally affected by the change in $r_{\rm cut}$, changing by less than 8% for the bias, and 3% for the mass, over the full range of $r_{\rm cut}$ explored.
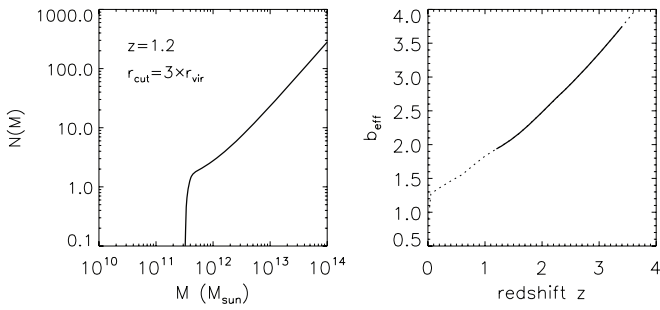
Figure 8. *Left panel*: Halo Occupation Number (HON) as a function of the mass of the halo at a representative redshift $z \sim 1.2$, assuming the best-fit values $M_{\rm min} = 10^{11.5}\,M_{\odot}$ and $\alpha = 0.95$. *Right panel*: redshift dependence of the large-scale effective bias, resulting from the best fit to the observed power spectrum of correlations due to clustering. The redshift interval over which 68% of the signal originates, $0.7 \lesssim z \lesssim 2.5$, is shown as a thick solid curve.

Figure 9 shows the contribution to the total clustering power spectrum from sources in different redshift slices, within the assumed source population model. As expected, the power is dominated by the contribution from sources in the range $0.7 < z < 1.5$, with an increasing contribution from sources at $z = 1.5$–3.0 for increasing wavelengths, as is expected from Figure 5. This is consistent with the findings of Devlin et al. (2009); Marsden et al. (2009); and Pascale et al. (2009), who through stacking show that of the sources making up the CIB, the fraction at $z > 1.2$ increases from 40% at 250 μm, to 50% and 60% at 350, and 500 μm.

Finally, we can use the model to interpret the clustering in terms of a three-dimensional spatial correlation length, r_0. To do that, we Fourier transform the best-fit power spectrum to find the spatial correlation function, $\xi(r)$, from which r_0 is then simply the linear comoving scale at which the correlation function equals 1 at each redshift. This is a model-dependent approach to estimating r_0, and as such should be considered an approximation. The more typical approach, which involves finding the angular correlation length, the redshift distribution, and deprojecting the signal by inverting the Limber equation, would result in very large uncertainties. The model-dependent r_0 is only mildly sensitive to the choice of $r_{\rm cut}$, varying by 10% over the full range. The model-dependent values for r_0 are illustrated in Figure 10 as a solid line with a shaded area representing 3σ uncertainties. The three BLAST points are plotted at the flux-weighted median redshifts of the unique redshift distributions probed by the three bands (i.e., the distributions shown in Figure 4), and should not be interpreted as the locations of all the sources contributing power in those bands. At these effective redshifts, (i.e., $z = 1.60$, 1.86, and 2.15), we find $r_0 = 4.9$, 5.0, and $5.2 \pm 0.6\,h^{-1}$ Mpc at 250, 350, and 500 μm, respectively.

10. DISCUSSION

10.1. Comparison with other Observations

Comparisons with other measurements of clustering must be made and interpreted with care because not everyone uses the same definition of bias, the same parameterization for the halo model, etc. Nevertheless it is interesting to put our measurements into the context of the large body of literature on the clustering of galaxies selected in different ways, in order to understand how they might be related.

In Figure 10, we compare the correlation lengths versus redshift of star-forming populations selected using a wide variety of techniques. It should be noted that in some cases

these techniques select overlapping populations (see Reddy et al. 2005, for a nice discussion on the overlap between color selected samples). This list is by no means exhaustive, and is meant only to be an illustration. The reported values of r_0 were converted assuming a fixed slope $\gamma = 1.8$, such that $r_{0,1.8} = (r_{0,\gamma})^{-\gamma/1.8}$. The BLAST best-fit model estimates are shown as a line and shaded 3σ confidence region. The three BLAST points are located at the median locations of their respective redshift distributions (see Figure 5). In addition, the simulated clustering lengths of dark matter halos of given mass and redshift are measured from the *Millennium Simulation*[20] (Springel et al. 2005), by fitting a single power law with slope of -1.8 to the correlation function, and are shown as dotted lines.

It is immediately clear that the galaxies which make up the background are not as strongly clustered as the more luminous sets of resolved sources (with the exception of those identified by their Lyman break, i.e., BM, BX, and LBG). Furthermore, the strength of the clustering increases with increasing luminosity (e.g., Gilli et al. 2007; Brodwin et al. 2008). Since each of the techniques used to select the populations of galaxies that lie above the BLAST lines has an IR component, it is tempting to conclude that all of these populations contribute to the total submillimeter CIB. The relative contribution of each of these populations could be explored through stacking, as Marsden et al. (2009) have done for BzKs. They found that although the BzKs make up about a quarter of the sources which completely resolve the CIB, they contribute $\sim$32%, $\sim$34%, and $\sim$42%, at 250, 350, and 500 μm, to the total BLAST intensity. This indication that the resolved sources pick out parts of the background highlights the complementary nature of the clustering measurements of the CIB and resolved sources in forming a complete picture of the environments of star-forming galaxies.

Although the correlation length of the background galaxies appears to change very little with redshift, the bias is a strong function of redshift (see Figure 8). While both the bias and the correlation length are indicators of galaxy clustering, their behavior is not in contradiction because the clustering strength of the host dark matter halos is rapidly increasing with decreasing redshift as well, thus in this sense, it is the bias that is a more telling description of how star formation relates to structure formation. This strong evolution of the bias is confirmed by Lagache et al. (2007), who found a redshift-independent bias parameter, $b \sim 2.4$, for the sources which make up the CIB at 160 μm,[21] at $z \sim 1$. Our earlier result (see Section 7), for galaxies which lie at higher redshifts, was $b \simeq 4$. Thus, in the scenario of a strongly evolving bias parameter, from anti-biased in the local universe, to highly biased at $z \sim 1$ and beyond (e.g., Sheth et al. 2006; Elbaz et al. 2007), our result is consistent.

10.2. Clustering in Context

Le Floc'h et al. (2005) show that IR-luminous galaxies (LIRGs and ULIRGS) represent ~ 70% of the IR energy density, and are responsible for most of the star formation, at $z \sim 0.5$–1.0 and beyond. Stacking work (e.g., Marsden et al.

---

[20] Note that Millenium Simulation cosmology is $\sigma_8 = 0.9$ and $\Omega_{\rm M} = 0.25$, which has a minimal impact on the prediction. Catalogs can be found at http://www.mpa-garching.mpg.de/Millennium.

[21] Note that the value $b = 1.7$ reported in Lagache et al. (2007) is for $\sigma_8 = 1.1$, and not $\sigma_8 = 0.8$, as quoted in the text. The correct value of the measured bias parameter is 2.4 ± 0.2 (G. Lagache 2009, private communication).
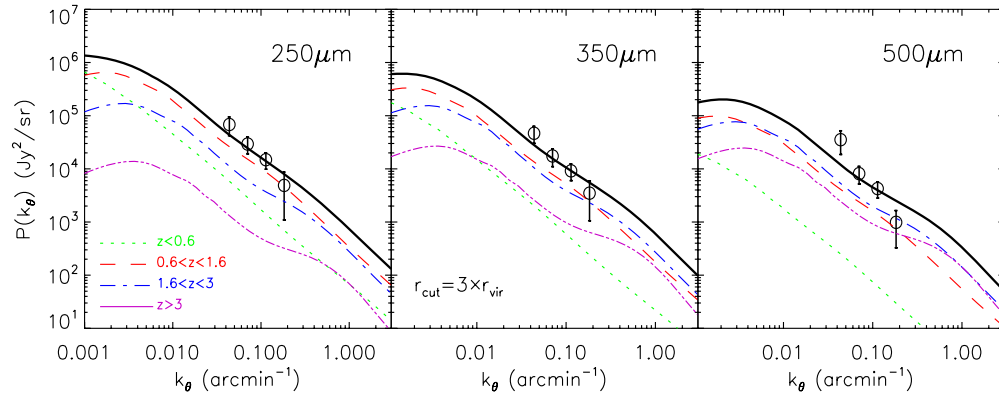
Figure 9. Contributions to the total clustering power spectrum from sources in increasing redshifts slices. BLAST measurements and the best fit to the halo model (with $r_{cut} = 3 \times r_{vir}$) are shown as circles with error bars and a solid line, respectively. Overlaid are the contributions from: dotted line (green), $0 < z < 0.7$; dashed line (red), $0.7 < z < 1.5$; dot-dashed line (blue), $1.5 < z < 3.0$; and triple-dot-dashed line (magenta), $z > 3$. It is clear that at $250\,\mu$m, the bulk of the signal comes from galaxies in the redshift range 0.7–1.5, and that the contribution from galaxies in the redshift range 1.5–3.0 increases with increasing wavelength.

(A color version of this figure is available in the online journal.)
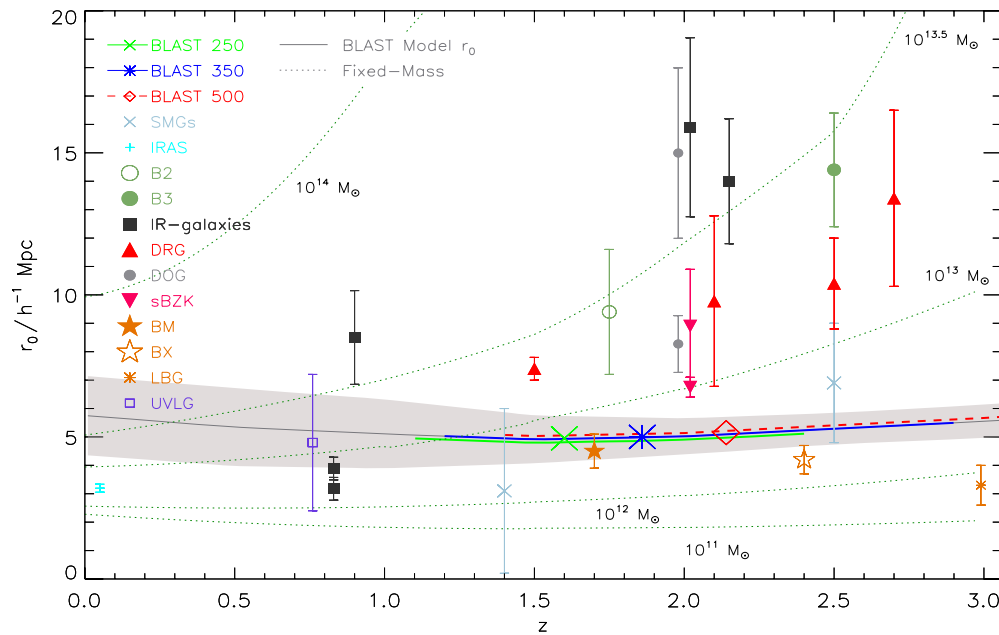
Figure 10. Comoving correlation length vs. redshift for star-forming galaxies selected with a variety of techniques. Other data taken from: *IRAS*—Saunders et al. (1992); SMG—Webb et al. (2003); Blain et al. (2004); B2 and B3—Farrah et al. (2006); IR—Magliocchetti et al. (2007, 2008); Gilli et al. (2007); DRG—Grazian et al. (2006); Quadri et al. (2008); DOG—Brodwin et al. (2008); BzK—Blanc et al. (2008); Hartley et al. (2008); BM and BX—Adelberger et al. (2005); LBG—Giavalisco et al. (1998); Adelberger et al. (2005); and UVLG—Basu-Zych et al. (2009). The dotted fixed mass lines show the predicted clustering lengths of halos of a given mass at a given redshift, found with the *Millennium Simulation* (Springel et al. 2005). The model-dependent BLAST values for r_0 are shown as a solid line with a shaded area representing 3σ uncertainties. The three BLAST points are plotted at the median redshifts of the distributions from which the signal originates (i.e., the distributions shown in Figure 5). The ranges from which 90% of the power originates are illustrated as corresponding colored lines. In the context of the model, the clustering strength of BLAST galaxies is compatible with the Webb et al. (2003) and Blain et al. (2004) estimates for clustering of submillimeter galaxies, but less stronger than that of the other resolved populations of galaxies.

(A color version of this figure is available in the online journal.)

2009) shows that most of the objects responsible for the CIB are fainter than the flux density limit of the BLAST catalogs (Devlin et al. 2009), corresponding to LIRG-like luminosities for those sources. Therefore, by identifying the locations of active star formation, we are also identifying the locations of the formation of the majority of stars in the present day universe. With this in mind we ask: where are active star-forming galaxies preferentially found through cosmic time? From Figure 10, it appears that the most active star-forming galaxies are found occupying halos whose mass increases roughly from 10^{12} to $10^{13.5}$ $M_\odot$ over the redshift range $z = 0$–2, after which it appears to remain roughly constant. This appears to be consistent with

downsizing (Cowie et al. 1996), the scenario in which the sites of most active star formation shifts to ever larger galaxies at higher redshifts. However, according to the model, which is only constrained for redshifts greater than ~ 1.1, the galaxies which make up the background do not exhibit the same trend. While the bias is still a strong increasing function of redshift, the clustering strength of the host halos remains roughly constant, corresponding to typical host halos that become slightly smaller. Since this is very much a model-dependent claim, it may be indicative of a flaw in the model (perhaps assuming M_{min}–α remain constant with z is incorrect); future studies should clarify this picture.

A striking feature is the sharp cutoff at $M \geqslant 10^{13.5}\ M_{\odot}$, which appears to hold out to $z \sim 2.5$. As was pointed out by Brodwin et al. (2008), this appears to be inconsistent with models which claim that star formation should be quenched in halos with masses greater than $10^{12}\ M_{\odot}$ due to shock heating (Birnboim & Dekel 2003; Dekel & Birnboim 2006; Cattaneo et al. 2008). Dekel et al. (2009) attempt to resolve this dilemma with a model where cold streams penetrate the shock-heated media. On the other hand, if the shock radius roughly follows the virial radius (Birnboim & Dekel 2003), then finding satellites actively forming stars outside of the shock-heated volume would satisfy both the model and the observations.

11. CONCLUSIONS

We report the detection of correlations in excess of Poisson noise in the CIB, over scales of approximately $5'$–$25'$, with BLAST at 250, 350, and 500 μm, at a level with respect to the CIB of $\Delta I / I = 15.1\% \pm 1.7\%$.

The CIB is made almost entirely out of individual sources distributed over a wide range of redshifts. We find that within the context of a reasonable model for the source population, the signal originates from galaxies in the redshift ranges $1.3 \leqslant z \leqslant 2.2$, $1.5 \leqslant z \leqslant 2.7$, and $1.7 \leqslant z \leqslant 3.2$, with median redshifts $z = 1.61, 1.88$, and 2.42, at 250, 350, and 500 μm, respectively. Fitting to the linear theory power spectrum, we find that the BLAST galaxies responsible for the CIB fluctuations have a bias parameter, $b = 3.8 \pm 0.6, 3.9 \pm 0.6$, and 4.4 ± 0.7.

We further interpret our results in terms of the halo model. We find that the simplest prescription does not fit very well. One way to improve the fit is to increase the radius at which we artificially truncate dark matter halos to well outside the virial radius. This may imply that the star-forming galaxies that we are seeing at $z \sim 1$ are preferentially found in the outskirts of groups and clusters. This is consistent with related phenomena that have been observed at other wavelengths (Magliocchetti et al. 2004; Marcillac et al. 2007), as well as in simulations (Diaferio et al. 1999).

For a HOD with "satellite" galaxies occupying halos out as far as $r_{\rm cut} = 3r_{\rm vir}$, we find parameters $\log(M_{\rm min}/M_{\odot} = 11.5^{0.4}_{-0.1}$, and $\alpha = 1.1^{+0.8}_{-0.1}$, resulting in effective biases $b_{\rm eff} = 2.2 \pm 0.2, 2.4 \pm 0.2$, and 2.6 ± 0.2, and effective masses $\log(M_{\rm eff}/M_{\odot}) = 12.9 \pm 0.3, 12.8 \pm 0.2$, and 12.7 ± 0.2 at 250, 350, and 500 μm, corresponding to spatial correlation lengths of $r_0 = 4.9, 5.0$, and $5.2 \pm 0.7 h^{-1}$Mpc, respectively.

In the context of the model, we see that star formation is highly biased at $z \gtrsim 1$, unlike in the local universe, where analogous galaxy populations, such as *IRAS* galaxies, are found to be mildly anti-biased (e.g., $M_{\rm eff} \lesssim 10^{11}\ M_{\odot}$ and $b = 0.86$; Saunders et al. 1992).

We find relatively small values for θ_0, further confirming that the sources which make up the CIB are distributed over a wide range of redshifts, which has implications for planning future submillimeter clustering measurements. For example, as Knox et al. (2001) have argued, to match the precision of the bias measurement made from correlations in the background by using discrete sources will be very challenging. To achieve $\sim 5\%$ accuracy would effectively require thousands of sources with exact redshifts, over tens of square degrees. When redshifts are only approximately known, this increases to hundreds of thousands of sources over hundreds of square degrees, which is only possible with instruments whose beams are smaller than

$5''$. Thus, while measuring the clustering from resolved sources has numerous advantages—for example, studying the clustering properties of a subset of galaxies—accurately measuring the large-scale bias is best achieved through correlation analysis of the background fluctuations. This will be a focus of future studies with BLAST, as well as with the *Herschel* and *Planck* satellites, and SCUBA-2.

REFERENCES

Adelberger, K. L., Steidel, C. C., Pettini, M., Shapley, A. E., Reddy, N. A., & Erb, D. K. 2005, ApJ, 619, 697
Almaini, O., Dunlop, J. S., Willott, C. J., Alexander, D. M., Bauer, F. E., & Liu, C. T. 2005, MNRAS, 358, 875
Almaini, O., et al. 2003, MNRAS, 338, 303
Amblard, A., & Cooray, A. 2007, ApJ, 670, 903
Barger, A. J., Cowie, L. L., Sanders, D. B., Fulton, E., Taniguchi, Y., Sato, Y., Kawara, K., & Okuda, H. 1998, Nature, 394, 248
Barger, A. J., Cowie, L. L., Smail, I., Ivison, R. J., Blain, A. W., & Kneib, J.-P. 1999, AJ, 117, 2656
Basu-Zych, A. R., et al. 2009, ApJ, 699, 1307
Berlind, A. A., Narayanan, V. K., & Weinberg, D. H. 2000, ApJ, 537, 537
Berlind, A. A., et al. 2003, ApJ, 593, 1
Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349
Blain, A. W., Chapman, S. C., Smail, I., & Ivison, R. 2004, ApJ, 611, 725
Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P., & Frayer, D. T. 2002, Phys. Rep., 369, 111
Blake, C., Pope, A., Scott, D., & Mobasher, B. 2006, MNRAS, 368, 732
Blanc, G. A., et al. 2008, ApJ, 681, 1099
Borys, C., Chapman, S., Halpern, M., & Scott, D. 2003, MNRAS, 344, 385
Borys, C., Scott, D., Chapman, S., Halpern, M., Nandra, K., & Pope, A. 2004, MNRAS, 355, 485
Boulanger, F., Abergel, A., Bernard, J.-P., Burton, W. B., Desert, F.-X., Hartmann, D., Lagache, G., & Puget, J.-L. 1996, A&A, 312, 256
Brodwin, M., et al. 2008, ApJ, 687, L65
Bullock, J. S., Dekel, A., Kolatt, T. S., Kravtsov, A. V., Klypin, A. A., Porciani, C., & Primack, J. R. 2001, ApJ, 555, 240
Cattaneo, A., Dekel, A., Faber, S. M., & Guiderdoni, B. 2008, MNRAS, 389, 567
Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, ApJ, 622, 772
Cooray, A. 2006, MNRAS, 365, 842
Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1
Coppin, K., et al. 2006, MNRAS, 372, 1621
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., et al. 2009, Nature, 457, 451
Devlin, M. J., et al. 2009, Nature, 458, 737
Diaferio, A., Kauffmann, G., Colberg, J. M., & White, S. D. M. 1999, MNRAS, 307, 537
Dickinson, M., Papovich, C., Ferguson, H. C., & Budavári, T. 2003, ApJ, 587, 25
Draine, B. T., & Lee, H. M. 1984, ApJ, 285, 89
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Elbaz, D., et al. 2007, A&A, 468, 33
Farrah, D., et al. 2006, ApJ, 641, L17
Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, ApJ, 508, 123
Gautier, T. N., III, Boulanger, F., Perault, M., & Puget, J. L. 1992, AJ, 103, 1313

Giavalisco, M., Steidel, C. C., Adelberger, K. L., Dickinson, M. E., Pettini, M., & Kellogg, M. 1998, ApJ, 503, 543
Giavalisco, M., et al. 2004, ApJ, 600, L93
Gilli, R., et al. 2007, A&A, 475, 83
Grazian, A., et al. 2006, A&A, 453, 507
Greve, T. R., Ivison, R. J., Bertoldi, F., Stevens, J. A., Dunlop, J. S., Lutz, D., & Carilli, C. L. 2004, MNRAS, 354, 779
Grossan, B., & Smoot, G. F. 2007, A&A, 474, 731
Haiman, Z., & Knox, L. 2000, ApJ, 530, 124
Hartley, W. G., et al. 2008, MNRAS, 391, 1301
Hinshaw, G., et al. 2009, ApJS, 180, 225
Hopkins, A. M. 2004, ApJ, 615, 209
Hughes, D. H., et al. 1998, Nature, 394, 241
Ivison, R. J., Smail, I., Barger, A. J., Kneib, J.-P., Blain, A. W., Owen, F. N., Kerr, T. H., & Cowie, L. L. 2000, MNRAS, 315, 209
Knox, L., Cooray, A., Eisenstein, D., & Haiman, Z. 2001, ApJ, 550, 7
Lagache, G., Bavouzet, N., Fernandez-Conde, N., Ponthieu, N., Rodet, T., Dole, H., Miville-Deschênes, M.-A., & Puget, J.-L. 2007, ApJ, 665, L89
Lagache, G., & Puget, J. L. 2000, A&A, 355, 17
Lagache, G., et al. 2004, ApJS, 154, 112
Le Floc'h, E., et al. 2005, ApJ, 632, 169
Magliocchetti, M., Moscardini, L., Panuzzo, P., Granato, G. L., De Zotti, G., & Danese, L. 2001, MNRAS, 325, 1553
Magliocchetti, M., & Porciani, C. 2003, MNRAS, 346, 186
Magliocchetti, M., Silva, L., Lapi, A., de Zotti, G., Granato, G. L., Fadda, D., & Danese, L. 2007, MNRAS, 375, 1121
Magliocchetti, M., et al. 2004, MNRAS, 350, 1485
Magliocchetti, M., et al. 2008, MNRAS, 383, 1131
Magnelli, B., Elbaz, D., Chary, R. R., Dickinson, M., Le Borgne, D., Frayer, D. T., & Willmer, C. N. A. 2009, A&A, 496, 57
Marcillac, D., Rigby, J. R., Rieke, G. H., & Kelly, D. M. 2007, ApJ, 654, 825
Marsden, G., et al. 2009, ApJ, 707, 1729
Miville-Deschênes, M.-A., & Lagache, G. 2005, ApJS, 157, 302
Miville-Deschênes, M.-A., Lagache, G., Boulanger, F., & Puget, J.-L. 2007, A&A, 469, 595
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Negrello, M., Perrotta, F., González, J. G.-N., Silva, L., de Zotti, G., Granato, G. L., Baccigalupi, C., & Danese, L. 2007, MNRAS, 377, 1557
Pascale, E., et al. 2008, ApJ, 681, 400
Pascale, E., et al. 2009, ApJ, 707, 1740
Patanchon, G., et al. 2008, ApJ, 681, 708
Patanchon, G., et al. 2009, ApJ, 707, 1750
Peacock, J. A. 1999, Cosmological Physics (Cambridge, UK: Cambridge Univ. Press)
Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144
Peacock, J. A., et al. 2000, MNRAS, 318, 535
Perrotta, F., Magliocchetti, M., Baccigalupi, C., Bartelmann, M., De Zotti, G., Granato, G. L., Silva, L., & Danese, L. 2003, MNRAS, 338, 623
Porciani, C., Magliocchetti, M., & Norberg, P. 2004, MNRAS, 355, 1010
Press, W. H. 2002, Numerical Recipes in C++ : The Art of Scientific Computing (Cambridge: Cambridge Univ. Press)
Puget, J.-L., Abergel, A., Bernard, J.-P., Boulanger, F., Burton, W. B., Desert, F.-X., & Hartmann, D. 1996, A&A, 308, L5
Quadri, R. F., Williams, R. J., Lee, K.-S., Franx, M., van Dokkum, P., & Brammer, G. B. 2008, ApJ, 685, L1
Reddy, N. A., Erb, D. K., Steidel, C. C., Shapley, A. E., Adelberger, K. L., & Pettini, M. 2005, ApJ, 633, 748
Saunders, W., Rowan-Robinson, M., & Lawrence, A. 1992, MNRAS, 258, 134
Scott, S. E., Dunlop, J. S., & Serjeant, S. 2006, MNRAS, 370, 1057
Scott, D., & White, M. 1999, A&A, 346, 1
Sheth, R. K., Jimenez, R., Panter, B., & Heavens, A. F. 2006, ApJ, 650, L25
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Smail, I., Ivison, R. J., & Blain, A. W. 1997, ApJ, 490, L5
Springel, V., et al. 2005, Nature, 435, 629
Tegmark, M., et al. 2002, ApJ, 571, 191
Tinker, J. L., Wechsler, R. H., & Zheng, Z. 2009, arXiv:0902.1748
Truch, M. D. P., et al. 2009, ApJ, 707, 1723
Watson, D. F., Berlind, A. A., McBride, C. K., & Masjedi, M. 2009, arXiv:0908.3678
Webb, T. M., et al. 2003, ApJ, 587, 41
Zehavi, I., et al. 2005, ApJ, 630, 1
Zheng, Z., et al. 2005, ApJ, 633, 791