CROSS-CORRELATION OF NEAR- AND FAR-INFRARED BACKGROUND ANISOTROPIES AS TRACED BY
SPITZER AND HERSCHEL

CAMERON THACKER$^1$, YAN GONG$^1$, ASANTHA COORAY$^1$, FRANCESCO DE BERNARDIS$^2$, JOSEPH SMIDT$^3$, AND
KETRON MITCHELL-WYNNE$^1$

$^1$ Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
$^2$ Department of Physics, Cornell University, Ithaca, NY 14853, USA
$^3$F-2, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Received 2014 December 12; accepted 2015 April 11; published 2015 September 29

ABSTRACT

We present the cross-correlation between the far-infrared (far-IR) background fluctuations as measured with the Herschel Space Observatory at 250, 350, and 500 μm and the near-infrared (near-IR) background fluctuations with the Spitzer Space Telescope at 3.6 and 4.5 μm. The cross-correlation between the FIR and NIR background anisotropies is detected such that the correlation coefficient at a few to 10 arcminute angular scale decreases from 0.3 to 0.1 when the FIR wavelength increases from 250 to 500 μm. We model the cross-correlation using a halo model with three components: (a) FIR bright or dusty star-forming galaxies below the masking depth in Herschel maps, (b) NIR faint galaxies below the masking depth, and (c) intra-halo light (IHL), or diffuse stars in dark matter halos, that is likely dominating the large-scale NIR fluctuations. The model is able to reasonably reproduce the auto-correlations at each of the FIR wavelengths and at 3.6 μm and their corresponding cross-correlations. While the FIR and NIR auto-correlations are dominated by faint, dusty, star-forming galaxies and IHL, respectively, we find that roughly half of the cross-correlation between the NIR and FIR backgrounds is due to the same dusty galaxies that remain unmasked at 3.6 μm. The remaining signal in the cross-correlation is due to IHL present in the same dark matter halos as those hosting the same faint and unmasked galaxies.

Key words: cosmology: observations – galaxies: evolution – infrared: galaxies – large-scale structure of universe – submillimeter: galaxies

1. INTRODUCTION

The cosmic infrared background (CIB) contains the total emission history of the universe integrated along the line of sight. The CIB contains two peaks, one at optical/near-infrared (near-IR) wavelengths around 1 μm and the second at far-infrared (far-IR) wavelengths around 250 μm (Dole et al. 2006). The former is composed of photons produced during nucleosynthesis in stars while the latter is the reprocessing of some of those photons by dust in the universe. While the total intensity in the NIR background, especially at 3.6 μm, has mostly been resolved to individual galaxies, we are still far from directly resolving the total CIB intensity at 250 μm and above to individual sources. This is due to the fact that at FIR wavelengths, observations are strongly limited by the aperture sizes. With the SPIRE Instrument (Griffin et al. 2010) on board the Herschel Space Observatory(Pillbratt et al. 2010), the background has been resolved to 5%, 15%, and 22% at 250, 350, and 500 μm, respectively (Oliver et al. 2010). In order to understand some properties of the faint FIR sources, it is essential that we study the fluctuations or the anisotropies of the background.

These spatial fluctuations in the CIB are best studied using the angular power spectrum. This technique provides a method to study faint and unresolved galaxies because, while not individually detected, they trace the large-scale structure. The clustering of these galaxies is then measurable through the angular power spectrum of the background intensity variations (Amblard et al. 2011). Such fluctuation clustering measurements at FIR wavelengths have been followed up with both Herschel and Planck (Viero et al. 2013; Planck Collaboration et al. 2014), the latter of which has provided the highest signal-to-noise calculation in the FIR.

NIR background anisotropies have been studied separately in the literature and have been interpreted as due to galaxies containing Pop III stars present during reionization (Kashlinsky et al. 2005, 2007, 2012), direct collapse black holes (DCBHs) at z > 12 (Yue et al. 2013), and intra-halo light (IHL; Cooray et al. 2012; Zemcov et al. 2014). In addition to Spitzer at 3.6 μm and above, fluctuation measurements in the NIR wavelengths have come from Akari (Matsumoto et al. 2011), NICMOS on HST at 1.6 and 1.1 μm at sub-arcminute angular scales (Thompson et al. 2007), and with the Cosmic Infrared Background Experiment (CIBER) at 1.1 and 1.6 μm at tens of arcminute-to-degree angular scales (Zemcov et al. 2014). While earlier studies argued for a substantial contribution from galaxies at z > 8, including those with Pop III stars or mini-quasars (Cooray & Yoshida 2004), recent calculations find that high-redshift galaxies are not likely to be the dominant contribution (Cooray et al. 2012; Yue et al. 2014). Both Cooray et al. (2012), using Spitzer, and Zemcov et al. (2014), using CIBER, argue for the case that the signal is coming from low redshifts and proposes an origin that may be associated with IHL. The IHL signal comes from diffuse stars that are tidally stripped during galaxy mergers and tidal interactions between galaxies.

An alternative to IHL for explaining excess fluctuations at NIR wavelengths is motivated by the cross-correlation between the NIR and X-ray background (Cappelluti et al. 2012). The model for the cross-correlation involves DCBHs with a minimum redshift for their presence at 12.4$^{+0.7}_{-0.1}$ (Yue

---

$^4$ Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.
The two models involving IHL and DCBHs remain equally viable to explain the Spitzer fluctuations. Given the sharp cutoff in redshift, a model involving DCBHs will not be able to explain the presence of fluctuations at observed wavelengths bluewards of 1.6 μm. Regardless of whether the signal is IHL or DCBHs, we expect some signal from galaxies present during reionization. This is due to the fact that deep Lyman drop-out surveys with HST in surveys such as CANDELS and ERS using WFC3 have detected a substantial number of galaxies at z > 7 (Finkelstein et al. 2012; Bouwens et al. 2014). Such galaxies, including those that are below the point-source detection level in HST surveys, will contribute to the fluctuations at 3.6 μm. Their presence can also be inferred in a multi-wavelength fluctuation analysis. In addition to these two, there is also a third contribution. These are the faint dwarf galaxies that are present between us and reionization (Helgason et al. 2012). Such galaxies contribute to the NIR and, since they have flux densities below the point-source detection level in NIR maps, they remain unmasked during fluctuation power spectrum measurements. The exact relative amplitudes of each of these signals are yet to be determined.

While fluctuations have been studied separately at FIR and NIR wavelengths, no attempt has yet been made to combine those measurements. In this work, we present the first results from just such a cross-correlation. We make use of the overlap coverage in the eight square degrees Boötes field between the Spitzer Deep Wide Field Survey (SDWFS) at 3.6 μm (Ashby et al. 2009) and Herschel at 250, 350, and 500 μm. We find that the NIR and FIR fluctuations are correlated but the cross-correlation coefficient is weak at 30% or below. Such a weak cross-correlation argues for the scenario that Spitzer and Herschel are mostly tracing two different populations. To interpret the data, we model the auto and cross-correlation signals using a three-component halo model composed of FIR galaxies, IHL, and faint galaxies. While we mainly focus on interpreting the cross-correlation between FIR and NIR fluctuations at 3.6 μm, we also perform the same cross-correlation with all three of the Herschel/Spire maps against 4.5 μm data of Boötes field. We find consistent correlations and provide an overall validity check on the measurement at 3.6 μm.

The paper is structured as follows. In Section 2, we discuss the data analysis and power spectra measurements. This includes the map making process for the Herschel and Spitzer data using the Herschel Interactive Processing Environment (HIPE) and self-calibration, respectively. This section also explains the mask generation procedure, defines the cross-correlation power spectrum, and discusses sources of error. In Section 3, we describe the halo model including components for FIR galaxies, IHL, and faint galaxies, and the Markov Chain Monte Carlo (MCMC) process used to fit the data. Finally, in Sections 4 and 5, we present the results of our model fit, discuss their implications, and give our concluding thoughts.

2. DATA ANALYSIS AND POWER SPECTRA MEASUREMENTS

We discuss the analysis pipeline we implemented for the cross-correlation of Herschel and Spitzer fluctuations. The study was conducted in the Boötes field making use of the Herschel/Spire instrument data taken as part of the Herschel Multi-tiered Extragalactic Survey (HerMES) and Spitzer/IRAC imaging data taken as part of the SDWFS (Ashby et al. 2009). We describe the data sets, map making, source masking, and power spectrum measurement details in the sub-sections below.

2.1. Map Making

2.1.1. FIR Maps with Herschel

We make use of the publicly available SPIRE instrument data of the Boötes field taken as part of HerMES from the ESA Herschel Science Data Archive. Our map making pipeline makes use of the Level 1 time-ordered scans that have been corrected for cosmic rays, temperature drifts, and bolometer time response as part of the standard data reduction by ESA. In those timelines, we remove a baseline polynomial on a scan-by-scan basis to normalize gain variations. The SPIRE maps are generated using the HIPE (Ott 2010). For this work, we make use of the HIPE (Cantalupo et al. 2010) algorithm that is native to HIPE. It is a maximum likelihood estimate and follows the approach used in Thacker et al. (2013). We do not use the map-maker that was developed for anisotropy power spectrum measurement in (Amblard et al. 2010) due to improvements that have been made to the HIPE and its built-in map-makers since the Amblard et al. study in 2010.

The end products of the map making process are single-tile maps covering roughly 8 square degrees and containing two sets of 80 scans in orthogonal directions with a pixel scale of 6 arcsec pixel$^{-1}$ in all SPIRE bands (250, 350, and 500 μm). These maps, using publicly available data from HerMES, are generated at a pixel scale one-third of the beam size consistent with our previous work as well. Such a sampling allows the point-source fluxes to be measured more accurately with adequate sampling of the point-spread function (PSF). For this study, we are forced to compromise between an accurate sampling of the PSF, point-source detection, and flux density measurement to obtain an accurate combination of Herschel and Spitzer data. Since the native scale of Spitzer maps is at 1.2 arcsec pixel$^{-1}$, we do not want to increase the SPIRE pixel scale substantially. We will return to this issue below when we discuss our point-source masks that were applied to the maps to remove detected sources.

2.1.2. Spitzer

The SDWFS (Ashby et al. 2009) maps of the Boötes field consist of four observations taken over 7–10 days from 2004 to 2008. For this analysis, we primarily concentrate on the cross-correlation with the 3.6 μm channel and interpret that cross-correlation with a model. Using four sets of data taken at different roll angles, we can ensure our fluctuation measurements are robust to detector systematics. In addition, since the measurements were taken in different years and months, multiple jack-knife tests enable us to reduce and quantify the error from astrophysical systematics such as the zodiacal light (ZL). As an overall check on the validity of our results and measurements, we also perform cross-correlation with 4.5 μm data. The cross-correlation is consistent, finding that there are no significant differences in the model between 3.6 and 4.5 μm.

5 Observation IDs listed in acknowledgments.  
6 http://sha.ipac.caltech.edu/applications/Spitzer/SHA/ Program number GO40839 (PI: D. Stern).
To further control the errors and to obtain a uniform background measurement across a very wide area, the maps are mosaiced using a Self-Calibration algorithm (Arendt et al. 2000). The algorithm is able to match the sky background levels with free gain parameters between adjacent overlapping frames using a least-squares fitting technique. We make use of the cleaned basic calibrated data publicly available from the Spitzer data archive for the map making process. These frames have asteroid trails, hot pixels, and other artifacts removed. Our original maps are made at a pixel scale of 1.2 arcsec pixel\(^{-1}\). Unlike our previous work (Cooray et al. 2012), we have repixelized the maps to a 6 arcsec pixel\(^{-1}\) scale to match the pixel scale of our Herschel maps.

The Spitzer maps span an area of about 10 deg\(^2\), which is larger than the Herschel SPIRE coverage at 8 deg\(^2\). For this study, we extract the overlapping area in both Spitzer and Herschel as outlined in Figure 1.

2.1.3. Astrometry

To ensure that both the Herschel and Spitzer images are registered to the same astrometric frame, we first checked for any offset in the Spitzer astrometry against the public catalog of SDWFS sources and corrected the astrometry using GAIA. We then made Herschel maps with the same astrometry as Spitzer. As a test on our overall astrometric calibration, we used Source Extractor (SExtractor; Bertin & Arnouts 1996) to detect all of the sources in both the Spitzer and Herschel images. We then matched the sources in the two catalogs within a radius of 18 arcsec (corresponding to the FWHM of the beam in 250 \(\mu\)m). Next, we took these matches and subtracted their R.A. and decl. such that a perfect match would have a \(\Delta\)R.A. and \(\Delta\)decl. of zero. Figure 2 shows a scatter plot of these

values in pixels rather than R.A. and decl. Perfectly aligned images would show a scatter centered at zero with equal spread in both directions. Our analysis shows a slight offset that is less than a pixel and within our tolerance as the effect on the cross-correlation is negligible since an offset less than a pixel will be binned to the nearest pixel and thus produces no offset.

2.2. Detected Source Masking

We remove the detected sources from our maps so that the cross-correlation is aimed at the unresolved fluctuations in both sets of data. We first generate two masks, one for the Spitzer data and one for the Herschel data (in all three bands). We combine these two masks to a common mask. This guarantees that both sets of data are free of as much foreground contamination as possible, minimizing spurious correlations by having a masked area in one map that is unmasked in the other. With a common mask we are also able to better handle the mode-coupling effects which masking introduces. The final combined mask (a zoomed in portion can be seen in Figure 3) removes 61.4\% of the pixels.

2.2.1. Spitzer

Since our model is based on sources below the detection threshold, we must pay particular attention to generating a relatively deep mask for Spitzer. Otherwise, unmasked Spitzer sources will correlate with the faint Herschel sources leading to a cross-correlation; in fact, as we find later, a significant fraction of the cross-correlation is due to faint Spitzer sources correlating with Herschel sources that are responsible for the SPIRE confusion noise.
We produce the Spitzer mask using a combination of catalogs from the NOAO Deep Wide Field Survey (NDWFS) in the B$_w$, R, and I bands, and the SEXTRACTOR catalog that was generated for Spitzer maps at a 3$\sigma$ detection threshold. As detailed in our previous work (Cooray et al. 2012), we start by iteratively running SEXTRACTOR with the same parameters used to generate the SDWFS catalogs (Ashby et al. 2009). This catalog is then combined with the NDWFS catalogs in the bands listed above. As with the Herschel mask, we apply a flux cut, remove all pixels that are 5$\sigma$ from the mean, and convolve everything with the PSF. Finally, we combine this mask with the Herschel mask.

Since this image has been rebinneed to a larger pixel size by a factor of five at 6 arcsec from our previous work at 1.2 arcsec in Cooray et al. (2012), significant blending occurs from nearby galaxies that are resolved in the 1.2 arcsec pixelized image. One effect of this is to increase the power at high-$\ell$ by increasing the shot-noise associated with unmasked galaxies. An aggressive mask for Spitzer, consistent with Cooray et al. (2012), cannot be used for this study since that will lead to a small fraction of unmasked pixels for the cross-correlation study with Herschel. Our main limitation here is not the depth of Spitzer imaging data, but the large PSF of Herschel maps. In Figure 4, we compare the power spectrum from Cooray et al. (2012) and the new power spectrum of Spitzer imaging data with a mask that retains more of the faint sources.

2.2.2. Herschel

In the Herschel data, the mask is generated by a simple three step process in each of the SPIRE bands. They are then combined into a single mask. The zeroth step is to crop the region in which Herschel and Spitzer overlap into a rectangular area as shown in Figure 1. First, we apply a flux cut at 50 mJy beam$^{-1}$, removing the brightest galaxies and staying consistent with our previous work on the clustering of fluctuations in the SPIRE bands (Amblard et al. 2011; Thacker et al. 2013). Next, to include the extended nature of the sources, we expand the mask by convolving with the PSF. Finally, we

Figure 3. Zoomed in image, roughly 0.7 square degrees, of Herschel 250 $\mu$m (left image) and Spitzer 3.6 $\mu$m (right image) data of the Boötes/SDWFS field. The upper images are the unmasked maps showing all sources, while the lower images have the mask applied to remove bright detected sources from the cross-correlation study. The mask over the whole area removes 61.4% of the pixels leaving the remainder for the study presented here.
remove pixels with no data, which arise mostly in 350 and 500 μm due to the enforced 6 arcsec pixel size (oversampling the PSF), and pixels with corrupt data by applying a $5\sigma$ clip to the images. The final mask for this study is obtained by multiplying the individual Herschel and Spitzer masks for the union between the two maps. The auto power spectra we show in Figure 5 use this combined mask between the two sets of data. A comparison of the shot-noise levels at high-$\ell$ for the Herschel power spectrum with models for source counts reveals that our final mask has an effective depth in the flux density cut at 250 μm of 29.5 mJy beam$^{-1}$. Figure 5 shows the differences in the auto power spectra with our combined mask versus power spectra measured with a mask based on a flux density cut for sources at 50 mJy.

2.3. Angular Power Spectra and Sources of Error

Our power spectra and cross-power spectra measurements follow the procedures we used for past work (Cooray et al. 2012; Thacker et al. 2013). Here, we summarize the key ingredients related to the uncertainties of the power spectrum measurements.

2.3.1. Instrumental Noise

One benefit of a cross-correlation is that instrumental noise is minimized, especially between two different detectors or imaging experiments. The maps can be thought of as the signal, $S$, plus a noise component, $N$. So in a cross-correlation, we have $M_1 \times M_2 = (S_1 + N_1) \times (S_2 + N_2)$. Since the noise is expected to be random and because these are two different detectors, the noise should be uncorrelated, thus dropping out of the cross-correlation.

To get a handle on the instrumental noise component and an estimate of systematic errors like ZL or cirrus contamination, we perform jack-knife tests. These tests involve taking half maps for Herschel and single epoch maps for Spitzer and cross-correlating their differences. For example, in Spitzer, we would take epochs (1–2) cross-correlated with epochs (3–4), where 1–4 are the four epochs of the SDWFS survey. We average all of the combinations of this and find that it is stationary about zero. Finally, the variance arising from this is an estimate of the residual noise and is added to the total error budget of the Spitzer auto power spectrum.

While it is assumed that the detector noise components of Herschel and Spitzer do not correlate, we still need to place a limit on any residual noise correlations in the cross-correlation. This is accomplished by taking null tests where we cross-correlate a Herschel map with all combinations of epoch-differenced maps of Spitzer. The variance of this multi epoch cross-correlations are added to our error budget of the Herschel and Spitzer cross-correlation power spectrum.

2.3.2. Beam Correction

At high $\ell$, there is a substantial drop in power due to the finite resolution of the detector and this drop in power must be accounted for. This is corrected for by taking the power
spectrum of the PSF and normalizing to one at low $\ell$. For cross spectra, we need a correction for the two separate beams. For that, we simply use the geometric mean of the two beams. For the Herschel data, we use the beam calculated by Amblard et al. (2011) based on observations of Neptune. The Spitzer beam used was calculated in Cooray et al. (2012) directly from the PSF as described there.

2.3.3. Mode-coupling Correction

Unlike the case of the cosmic microwave background (CMB), infrared background power spectra involve aggressive masks that remove a substantial fraction of the pixels. A consequence of such masking is the breaking up of larger Fourier modes into smaller modes. This results in a shift of power from low-$\ell$ to high-$\ell$. The effects of masking, in the context of CMB studies, are discussed in Hivon et al. (2002). The MASTER method is computationally efficient and is used for all CMB power spectra estimates by the cosmological community. We have extended the MASTER technique from an analytical calculation to a numerical technique in Cooray et al. (2012) where we generate the mode-coupling matrix by simulation effects of the mask at each $\ell$-mode. That is to say, we simulate a sky at one fixed $\ell$ mode and we mask that simulated map with the same mask as the data to produce the Fourier transform. That Fourier transform shows how the power leaks from that particular mode to all of the other modes due to the effects of masking. We repeat this process for all $\ell$ modes probed by the data and at each $\ell$ mode take the power spectrum of the masked map. This results in the mode-coupling matrix shown in Figure 6, where at each $\ell$ one can see how the power leaks to different $\ell'$ modes. Thus, by construction the matrix transforms unmasked power to masked power. We simply invert the matrix to obtain a transformation to an unbiased power spectrum. Figure 7 illustrates the effect of masking has on power spectra and the ability of the mode-coupling matrix to correct for that mask.7

2.3.4. Map Making Transfer Function

Unfortunately, the map making process does not result in a perfect representation of the sky. The map making process can induce fictitious correlations or suppress them based on deficiencies in the scan pattern, the technique that was used to process the timeline data including any filtering that was applied. We capture all of the modifications associated with the map making process by the transfer function, $T(\ell)$.

For Herschel data, we calculate the transfer function by making 100 Gaussian random maps and reading them into timeline data using the same pointing information as the actual data. Then, we produce artificial maps from these timelines and take the average of the power spectra of each of these maps to the power spectrum in each of the input maps. For the Spitzer data, we follow a similar approach but instead of reading data into timelines we create small tiles that are re-mosaiced using the same self-calibration algorithm as the actual data. These simulated tiles include Gaussian fluctuations and instrumental noise consistent with the corresponding tile in the actual data. Due to the dithering and tiling pattern that was used for actual observations in SDWFS with Spitzer, the map making transfer function is essentially one at all angular scales of interest. This is substantially different for Herschel/SPIRE with a transfer function that departs from one at low and high $\ell$ as shown in Figure 8. The difference at high $\ell$ is due to the scan pattern that leaves stripes in the data; the difference at low $\ell$ or large angular scales is due to the filtering that was applied to timeline data to remove gain drifts during individual scans. That is the process used to make maps and to ensure the map making process has not generated a bias in the power spectrum.

7 Footnote 1 of Kashlinsky et al. (2014) raises this issue of whether a heavily masked sky is affected by the application of the MASTER algorithm (Hivon et al. 2002). The technique we have used here and developed in Cooray et al. (2012) extends the MASTER algorithm to numerically simulate, mode-by-mode, the effect of the mask. It is not subject to the masking size of the map. Figure 7 here and Figure S9 of Cooray et al. (2012) demonstrate that the method is adequate and works to the level needed for power spectrum measurements. The disagreements highlighted by Kashlinsky et al. (2014) do not result from the correction related to the mask.
2.3.5. Power Spectrum Estimate

To put these corrections to use, we must first define the cross-correlation. Given two maps $H$ and $S$, the cross-correlation is calculated as

$$C_{\ell l H l l S l l} = \|H_{xy}\| \|S_{xy}\| \times \sum_i w(i, l) \times \bar{H}(i, l) \bar{S}^*(i, l),$$

where $H$ and $S$ are the two-dimensional (2D) Fourier transforms of their respective maps and $w(i, l)$ is the mask in Fourier space. The auto-spectra follows from the case when $H = S$.

Finally, the final cross-power spectrum estimate is obtained from the measured power spectrum, $\bar{C}_{\ell l}$, with the above described corrections as

$$C_{\ell} = \frac{M_{\ell l}^{-1} \bar{C}_{\ell l}}{T_{\ell l} b_i^H b_{ij}^S}.$$  

To generate the Spitzer auto-spectrum (Figure 4) we take the average of the permutations of different cross-correlations involving the four epochs of SDWFS. For example, we calculate $(1 + 2) \times (3 + 4)$ averaged with all other permutations where $1-4$ are the four epochs. When compared to our previous work, Spitzer auto-spectrum deviates at high $\ell$. The higher shot-noise is because we have repixelized the image from 1.2 arcsec pixel$^{-1}$ in our prior work to 6 arcsec pixel$^{-1}$ here. This also leads to a more shallow mask than our previous work.

Figure 9 shows the cross-correlations with the best-fit model (described below) shown for comparison. Again, we make use of the four different Spitzer epochs in the cross-correlation. The cross-correlation is calculated as the average of $\langle \text{epoch}_i + \text{epoch}_j \rangle / 2 \times \text{Herschel}$, where $i \neq j$.

In addition to the instrumental noise term and errors coming from the uncertainties in the beam function and map making transfer function, we also account for the cosmic variance in our total error budget. More details are available in the Appendix C.

3. HALO MODEL

In this section, we outline our model for the cross-correlation. The underlying model involves three key components: FIR galaxies, IHL, and NIR galaxies.

3.1. Model for FIR Background Fluctuations from Star-forming Dusty Galaxies

We use the conditional LF (CLF) models to calculate the power spectrum of FIR galaxies (Lee et al. 2009; De Bernardis & Cooray 2012). The probability density of a halo/subhalo
with mass $M$ to host a galaxy observed in a FIR band with luminosity $L$ (i.e., $L_{250}$, $L_{350}$ and $L_{500}$ in this work) is given by

$$P(L|M) = \frac{1}{\sqrt{2\pi} \ln(10) L_{\Sigma}} \exp \left\{ -\frac{\log_{10}[L/L_c(M,z)]^2}{2\Sigma^2} \right\},$$

(3)

where $\Sigma = 0.3$ (Lee et al. 2009) is the variance of $\log_{10}L(M,z)$, and $L(M,z)$ is the mean luminosity given a halo/subhalo mass $M$ at redshift $z$ which takes the form

$$L(M,z) = L_0 \left( \frac{M}{M_0} \right)^\alpha \exp \left[ -\left( \frac{M}{M_0} \right)^\beta \right].$$

(5)

where $\beta = 0$ when we perform the fitting process since we find $\beta$ is close to zero as a free parameter (De Bernardis & Cooray 2012).

The one-halo and two-halo terms of the three-dimensional (3D) power spectrum for FIR galaxies are given by

$$P_{1h}(k,z) = \int dMn(M,z) \frac{\langle N_g(N_g - 1) \rangle_{\bar{n}_g}}{\bar{n}_g^2} u^p(k|M,z),$$

(10)

$$P_{2h}(k,z) = \left[ \int dMb(M,z)n(M,z) \frac{\langle N_g \rangle_{\bar{n}_g}}{\bar{n}_g} u(k|M,z) \right]^2 \times P_{\text{lin}}(k,z),$$

(11)

where $u(k|M,z)$ is the Fourier transform of the Navarro–Frenk–White halo density profile for halos of mass $M$ at redshift $z$ (Navarro et al. 1997), and the power index $p$ takes $p = 1$ when $\langle N_g(N_g - 1) \rangle_{\bar{n}_g} < 1$ and $p = 2$ otherwise (Cooray & Sheth 2002). $b(M,z)$ is the halo bias (Sheth & Tormen 1999) and $P_{\text{lin}}$ is the linear matter power spectrum. $\bar{n}_g$ is the galaxy mean number density, which is expressed as

$$\bar{n}_g = \int dMn(M,z)\langle N_g(M) \rangle.$$
assumed to be the SED of old elliptical galaxies which are composed of old and red stars (Krick & Bernstein 2007), and we normalize $F_{\ell}^{\text{ISO}} = 1$ at 2.2 μm.

The one-halo and two-halo terms of the 2D IHL angular cross-power spectrum at observed wavelengths $\lambda$ and $\lambda'$ can be estimated by

$$C_{\ell,\text{IHL}}^{\lambda,\lambda'} = \frac{1}{(4\pi)^2} \int dz \left( \frac{d\chi}{dz} \right) \left( \frac{a}{\chi} \right)^2 \times \int dMn(M,z)u^2(k|M,z)\tilde{E}_{\text{IHL}}^{\lambda} \tilde{E}_{\text{IHL}}^{\lambda'}.$$  \hspace{1cm} (17)

$$C_{\ell,\text{IHL}}^{\lambda,\lambda'} = \frac{1}{(4\pi)^2} \int dz \left( \frac{d\chi}{dz} \right) \left( \frac{a}{\chi} \right)^2 P_{\text{lin}}(k,z) \times \int dMb(m,z)n(M,z)u(k|M,z)\tilde{E}_{\text{IHL}}^{\lambda} \times \int dMb(m,z)n(M,z)u(k|M,z)\tilde{E}_{\text{IHL}}^{\lambda'}. \hspace{1cm} (18)$$

The total 2D IHL angular cross-power spectrum is then given by

$$C_{\ell,\text{IHL}}^{\lambda} = C_{\ell,\text{IHL}}^{\lambda,\lambda'1h} + C_{\ell,\text{IHL}}^{\lambda,\lambda'2h}. \hspace{1cm} (19)$$

3.3. Model for NIR Background Fluctuations from Known Galaxy Populations

We follow Helgason et al. (2012) in estimating the NIR background fluctuations from known galaxy populations. We use their empirical fitting formulae of LFs for measured galaxies in the UV, optical, and NIR bands out to $z \sim 5$. Then, we estimate the mean emissivity $\tilde{J}_e(z)$ at the rest-frame frequencies and redshift to the observed frequency. We adopt the HOD model to calculate the 3D galaxy power spectrum using Equations (10) and (11) with $\langle \tilde{N}_e \rangle = \langle N_e \rangle + \langle N_i \rangle$ and $\langle N_g(N_g - 1) \rangle \simeq 2 \langle N_e \rangle \langle N_i \rangle + \langle N_i \rangle^2$, where

$$\langle N_e \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log_{10} M - \log_{10} M_{\text{min}}}{\sigma_M} \right) \right], \hspace{1cm} (20)$$

and

$$\langle N_i \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log_{10} M - \log_{10} 2M_{\text{min}}}{\sigma_M} \right) \right] \left( \frac{M}{M_i} \right)^{\alpha_i}. \hspace{1cm} (21)$$

Here, $M_{\text{min}}$ denotes the mass at which a halo has a 50% probability of hosting a central galaxy and $\sigma_M$ is the transition width. We assume the satellite term has a cutoff mass with twice the mass of central galaxy and grows as a power law with slope $\alpha_i$ and normalized by $M_i$. We set $M_{\text{min}} = 10^9 M_\odot$, $\sigma_M = 0.2$, $M_i = 5 \times 10^{10} M_\odot$, and $\alpha_i = 1$ (Helgason et al. 2012). The 2D angular cross-power spectrum $C_{\ell,\text{NIR}}^{\lambda}$ can be calculated by Equation (13) with $\tilde{J}_e(z)$ and $P_{\text{gal}}(k,z)$ of unresolved galaxies.

3.4. Model Comparison

In this work, we have three angular auto power spectra $C_{\ell}^{\text{HER}}$ in the FIR from Herschel/SPIRE at 250, 350, and 500 μm, one angular auto power spectrum $C_{\ell}^{\text{SPI}}$ from Spitzer at 3.6 μm, and three cross-power spectra $C_{\ell}^{\text{cross}}$ between Herschel and Spitzer data. Using the models discussed in the previous sections, we fit all of these auto and cross-power spectra with a single model.

As mentioned, we consider three components in our model to fit each data set. For $C_{\ell}^{\text{HER}}$, we have $C_{\ell}^{\text{HER}} = C_{\ell}^{\text{FIR}} + C_{\ell}^{\text{HER,shot}}$, where $C_{\ell}^{\text{FIR}} = C_{\ell}^{\text{FIR,obs}} + C_{\ell}^{\text{FIR,shot}}$ are the observed frequencies at 250, 350, or 500 μm, and $C_{\ell}^{\text{HER,shot}}$ is the shot-noise term. In a similar way, for $C_{\ell}^{\text{SPI}}$, we take $C_{\ell}^{\text{SPI}} = C_{\ell}^{\text{FIR,obs}} + C_{\ell}^{\text{SPI}} + C_{\ell}^{\text{SPI,shot}}$ at 3.6 μm.

For $C_{\ell}^{\text{cross}}$, it is expressed as $C_{\ell}^{\text{cross}} = C_{\ell}^{\text{FIR,obs}} + C_{\ell}^{\text{SPI}} + C_{\ell}^{\text{FIR,obs} \times \text{SPI}}$. The details of these cross-power spectra are outlined in the appendix.

Since the shot-noise is a constant and dominant at high $\ell$, we derive the shot-noise terms from the data at the high $\ell$ and fix them in the fitting process. The values of the shot-noise terms we find are $C_{\ell}^{\text{HER,shot}} = 8.4 \times 10^{-7}, 3.0 \times 10^{-7}$ and $8.0 \times 10^{-8}$ nW$^2$ m$^{-4}$ sr$^{-1}$, and $C_{\ell}^{\text{SPI,shot}} = 1.7 \times 10^{-8}, 3.0 \times 10^{-9}$ and $1.4 \times 10^{-9}$ nW$^2$ m$^{-4}$ sr$^{-1}$ at 250, 350, and 500 μm, respectively.

For $C_{\ell}^{\text{cross}}$, we scale the shot-noise of the NIR background $C_{\ell}^{\text{NIR,shot}}$ given by Equation (13) in Helgason et al. (2012) to match the high-$\ell$ data of Spitzer at 3.6 μm. This is done by adjusting the minimum apparent magnitude $m_{\text{min}}$, and we find $m_{\text{min}} = 23$, which gives $C_{\ell}^{\text{NIR,shot}} = 3.1 \times 10^{-10}$ nW$^2$ m$^{-4}$ sr$^{-1}$.

We employ the MCMC method to perform the fitting process. The Metropolis–Hastings algorithm is adopted to determine the probability of accepting a new MCMC chain point (Metropolis et al. 1953). We use the $\chi^2$ distribution to calculate the likelihood function $\mathcal{L} \propto \exp\left(-\frac{\chi^2}{2}\right)$. For the three data sets, we have $\chi^2 = \chi^2_{\text{HER}} + \chi^2_{\text{SPI}} + \chi^2_{\text{cross}}$, and $\chi^2$ is given by

$$\chi^2 = \sum_{i=1}^{N_{\text{data}}} \frac{(C_{\ell}^{\text{obs}} - C_{\ell}^{\text{th}})^2}{\sigma_{\ell}^2}, \hspace{1cm} (22)$$

where $N_{\text{data}}$ is the number of the data points, $C_{\ell}^{\text{obs}}$ and $C_{\ell}^{\text{th}}$ are the angular power spectra from observation and theory, and $\sigma_{\ell}$ is the error for each data point.

For simplicity, we take $\alpha_{\text{HER}} = 0$, $\alpha_{\text{SPI}} = 6$, $M_{\text{min}} = 10^9 h^{-1} M_\odot$, and $M_{\text{max}} = 10^{14} h^{-1} M_\odot$ when we calculate the integral over the redshift and halo mass in the power spectra. We use a uniform prior probability distribution for the free parameters in our model. The parameters and their ranges are as follow: $\log_{10} L_0 \in (-9, 1)$, $\alpha \in (-5, 5)$, $p \in (0, 5)$ for the FIR galaxy model, and $\alpha_{\text{HER}} \in (-5, 0)$, $p_{\text{SPI}} \in (-5, 5)$ for the IHL model. We generate 12 parallel MCMC chains for each data set and collect about 120,000 chain points after the chains reach convergence. After the burn-in process and thinning the chains, we merge all of the chains together and get about 10,000 chain points to illustrate the probability distribution of the free parameters. The details of our MCMC method can be found in Gong & Chen (2007).

4. RESULTS AND DISCUSSION

In this section, we discuss the results from the data analysis and the MCMC model fits to the auto and cross-power spectrum data. We will then outline estimates of derived quantities from the model fitting results, such as the redshift distribution of the FIR and NIR intensity and the cosmic dust density.
4.1. Power Spectra

In Figure 9, we show the angular auto and cross-power spectra of Herschel and Spitzer in the Boötes field. The top, middle, and bottom panels are for Herschel 250, 350, and 500 μm, respectively. The blue solid curve is the total best-fit power spectrum for the Herschel power spectrum, which is comprised of two components, i.e., FIR galaxies (blue dashed) and total shot-noise (not shown). Similarly, the green solid line is the best fit for the Spitzer auto power spectrum, which has contributions from IHL (green dotted) and NIR faint galaxies (green dashed). As described in Cooray et al. (2012), the shot-noise of the Spitzer power spectrum can be described by the faint galaxies, though the clustering of such galaxies fall short of the fluctuation power spectrum at the tens of arcminutes angular scales. The red solid line is the total cross-correlation, including a shot-noise term not shown here, while the red dotted and dashed lines separate the cross-correlation to the main terms given by IHL correlating with faint FIR dusty galaxies and faint NIR galaxies correlating with faint FIR dusty galaxies, respectively. In terms of the cross-correlations, we find that these two terms are roughly comparable.

Our best-fit model has \( \chi^2/(\text{dof}) = 2.7, 2.5, 3.0 \) for 250, 350, and 500 μm, respectively. This is the model for all of the measurements, both auto and cross-correlations at each of the FIR wavelengths. We do not separately model fit the auto-correlations independent of the cross-correlations.

Additionally, as a check that the cross-correlation does not drop out at 4.5 μm, we have included the cross-correlation between Herschel and Spitzer 4.5 μm as open circles shown in Figure 9. While no modelling was performed on this data to minimize complexity, it stands as validation that the correlation between Herschel and Spitzer 3.6 μm is not spurious.

Another comparison of the data and model is shown in Figure 10. Here, we calculate the correlation coefficient separately from the data and compare to the correlation coefficient of the best-fit model. Displaying the information in this way is a valuable check of the relative strength of correlation. As the FIR wavelength is increased from 250 to 500 μm, we find less of a correlation between Herschel and Spitzer. Not only does the total correlation decrease, from our model fit we also find that the correlation between IHL and dusty FIR galaxies, as a fraction of the total correlation, is also decreased. Note that Figure 10 does not contain any additional information beyond what is shown in Figure 9. The cross-correlation coefficient and the model curves are computed directly from the same set of curves (and error bars for the measurements) as shown in Figure 9.

The best-fit values and 1σ errors of the model parameters are shown in Table 1. We also show one of the model parameter confidence contours in Figure 11. We find the MCMC forces only a shot-noise term, i.e., a straight line, to fit the data if we use data points out to \( \ell \sim 10^5 \) because the errors at high \( \ell \) (small scales) are very small compared to the errors at large scales. For these reasons, our results are obtained by fitting to the power spectrum data at \( \ell < 10^4 \), and ignoring all data points at \( \ell > 10^4 \) in the fitting process.

4.2. Intensity Redshift Distribution

Using the fitting results of the best-fit model parameters and their errors, we estimate the redshift distribution of \( d(\nu_\nu^{\text{FIR}})/dz \) for FIR dusty galaxies in the three SPIRE bands as

\[
\frac{d(\nu_\nu^{\text{FIR}})}{dz} = \frac{c}{H(z)(1 + z)^2J_\nu(z)},
\]
where $c$ is the speed of light, $H(z)$ is the Hubble parameter, and $\bar{J}_\nu(z)$ is the mean emissivity given by Equation (8). Also, the redshift distribution of $d(\nu L_\nu)/dz$ for the IHL component can be obtained by

$$
\frac{d(\nu L_\nu^{\text{IHL}})}{dz} = \frac{c}{4\pi H(z)(1+z)^2} \int dMn(M, z)\bar{L}_{\text{IHL}},
$$

(24)

where $\bar{L}_{\text{IHL}}(M, z)$ is the mean IHL luminosity given by Equation (14), and $n(M, z)$ is the halo mass function.

In Figure 12, we show $d(\nu L_\nu)/dz$ as a function of the redshift for Herschel galaxies at 250, 350, and 500 $\mu$m, and Spitzer IHL and faint unresolved galaxies at 3.6 $\mu$m. We find that the redshift distribution has a turnover between $z = 1$ and 2 for 250 $\mu$m, which indicates that the intensity is dominated by the sources at $z = 1 \sim 2$ for 250 $\mu$m. For 350 $\mu$m, the turnover of $d(\nu L_\nu)/dz$ is not obvious and is shifted to higher redshift around $z = 3$. For 500 $\mu$m, we find that the sources at $z > 3$ dominate the intensity. The increase in redshift with wavelength for dusty sources that we find here is consistent with the well-known results in the literature (Amblard et al. 2011; Bethermin et al. 2011; Viero et al. 2013). On the other hand, the $d(\nu L_\nu)/dz$ of Spitzer IHL and residual galaxies at 3.6 $\mu$m has maximum values at $z = 0$ and decreases quickly with increasing redshift. This shape, which is substantially different from Herschel dusty galaxies, is the main reason that the cross-correlation signal between Herschel and Spitzer is below 0.5 at 250 $\mu$m with a decrease to a value below 0.2 at 500 $\mu$m. The decrease with increasing wavelength is consistent with the overall model description. If naively interpreted, the small cross-correlation with Herschel could have been argued as evidence for a very high-redshift origin for the Spitzer fluctuations, similar to the arguments for the origin of Spitzer–Chandra cross-correlation (Cappelluti et al. 2012). Our modeling suggest the opposite: Spitzer fluctuations are very likely to be dominated by a source at $z < 1$ while intensity fluctuations in Herschel originate from $z = 1$ and above at 250, 350, and 500 $\mu$m.

### 4.3. Cosmic Dust Density

As an application of our models, we also derive the fractional cosmic dust density $\Omega_{\text{dust}}$ of galaxies from the fitting results. Following Thacker et al. (2013), the $\Omega_{\text{dust}}$ for galaxies is given by

$$
\Omega_{\text{dust}}(z) = \frac{1}{\rho_0} \int dL \Phi(L, z) M_{\text{dust}}(L),
$$

(25)

where $\rho_0$ is the current cosmic critical density, $\Phi(L, z)$ is the LF, and $M_{\text{dust}}$ is the dust mass with IR luminosity $L$ with temperature assumed to be 35 K that we fix in the power spectrum model. We make use of Equation (4) of Fu et al. (2012) to estimate $M_{\text{dust}}$ from IR luminosity (Thacker et al. 2013). Also, we assume that the dust opacity $\kappa_d$ takes the form $\kappa_d = A_d \nu^{\beta_d}$, where $\beta_d = 2$ is the dust emissivity spectral index, and $A_d$ is the normalization factor which is estimated by $\kappa_d = 0.07 \pm 0.02$ m$^2$ kg$^{-1}$ at 850 $\mu$m (Dunne et al. 2000; James et al. 2002).

We show $\Omega_{\text{dust}}$ as a function of redshift in Figure 13. The blue shaded region is the 1$\sigma$ range of the $\Omega_{\text{dust}}$ from the FIR. We also show the other measurements for comparison. The orange squares are from the H-ATLAS dust mass function as measured in Dunne et al. (2010), and the other data are based on the extinction measurements of the SDSS and 2dF surveys (Fukugita & Peebles 2004; Driver et al. 2007; Menard et al. 2010; Fukugita 2011; Menard & Fukugita 2012). We find that our $\Omega_{\text{dust}}$ result from galaxies is consistent with other measurements, which have $\Omega_{\text{dust}}$ increasing at higher redshift.

### 4.4. Alternative Models

As discussed in the Introduction, an alternative and viable model for NIR fluctuations involves the DCBHs with a minimum redshift for their presence at a redshift of 12 or so. If
DCBHs replace IHL, note that only 50% of the cross-correlation between NIR and Herschel fluctuations can be explained with faint galaxies that are also dusty and contribute to the confusion noise in Herschel/Spire maps. Such galaxies cannot be made brighter at FIR wavelengths to explain the cross-correlation as that will overproduce the Herschel auto-correlation functions. The alternative approach will be to introduce some FIR emission to DCBHs. Since their minimum redshift is around 12, the 250–500 μm emission we detect originates from rest-frame 20 to 40 μm. The SED model for DCBHs presented in Yue et al. (2013) does show emission at these wavelengths.

The overall DCBH population that explains the background fluctuations at 3.6 μm and the X-ray/IR cross-correlation (Cappelluti et al. 2012) produces slightly below $10^{-3}$ nW m$^{-2}$ sr$^{-1}$ of the total background intensity at 250–500 μm (Figure 4 of Yue et al. 2013). At 3.6 μm they account for around 0.5 nW m$^{-2}$ sr$^{-1}$ and such an intensity level is necessary to explain IRAC fluctuations. For comparison, the total FIR background at 250–500 μm is at the level of 10 nW m$^{-2}$ sr$^{-1}$. It is unclear if DCBHs, with an intensity contribution at 0.01% of the total background intensity at FIR wavelengths, can explain the additional cross-correlation, beyond faint low-redshift galaxies, between NIR and FIR. As the current model stands, we do find the intensity level produced by DCBHs to be below the level necessary. We encourage further modeling and ways to increase the rest-frame mid-IR intensity of DCBHs by at least a factor of 100 so that if the total intensity is at a level of 1% of the background, it can start to become a viable model to explain the remainder of the cross-correlation that is currently explained by the IHL model.

5. SUMMARY

We have calculated the cross-correlation power spectrum of the CIB at FIR and NIR wavelengths using Herschel and Spitzer in the Boötes field. We measured the correlation coefficient to be between 10%–40% with the highest correlation seen in the 250 μm band and the lowest in 500 μm.

Recent results from Cooray et al. (2012) and Zemcov et al. (2014) suggest that the NIR background anisotropies have a mostly low-redshift origin at $z < 1$, arising from IHL and faint draft galaxies. Meanwhile, the FIR signal is dominated by dusty galaxies peaking at a redshift of ~1 and above (Amblard et al. 2010; Thacker et al. 2013; Viero et al. 2013). By cross-correlating Herschel with Spitzer, we are able to provide a check on the robustness of such a model. We find that not only can such a model fit the auto-correlations they were designed to fit, but they can also explain the cross-correlation signal, including the wavelength dependence of the cross-correlation coefficient. We encourage alternative models, especially those involve sources at high redshifts, such as DCBHs, to explain the measured cross-correlation.

This work was supported by NSF AST-1313319 and GAANN at UCI. The public Herschel data in the Herschel Science Archive was taken by HerMES with observation IDs 1342187711–1342187713, 1342188090, 1342188650–1342188651, 1342188681–1342188682, and 1342189108. HIPE was a joint development by the Herschel Science Ground Segment Consortium, consisting of ESA, the NASA Herschel Science Center, and the HIFI, PACS, and SPIRE consortia. We thank Rick Arendt for sharing his self-calibration pipeline for Spitzer data. We thank ESA for maintaining the Herschel Science Archive and for making available reduced and calibrated time-ordered data of every SPIRE scan that was done during the Herschel mission.

APPENDIX A
CROSS POWER SPECTRA

Here, we discuss the 2D angular cross-power spectra of Herschel and Spitzer in our model. The total cross-power spectrum $C_{\ell}^{\text{cross}}$ is composed of four components, which are

$$C_{\ell}^{\text{cross}} = C_{\ell}^{\text{FIR} \times \text{NIR}} + C_{\ell}^{\text{FIR} \times \text{IHL}}.$$ (26)

Here, $C_{\ell} = \nu_{\text{obs}}^\ell \nu_{\text{obs}}^\ell C_{\ell}^{\nu_{\text{obs}}^\nu_{\text{obs}}}$, where $\nu_{\text{obs}}$ is the observed frequency at 250, 350, or 500 μm of the Herschel survey, and $\nu_{\text{obs}}$ is the observed frequency at 3.6 μm of the Spitzer survey.

The one-halo and two-halo terms of $C_{\ell}^{\nu_{\text{obs}}} \nu_{\text{obs}}$ are given by

$$C_{\ell, \text{FIR} \times \text{NIR}}^{\nu_{\text{obs}}} = \int dz \left( \frac{dz}{dz} \right) \left( \frac{\hat{s}}{d\hat{s}} \right) \int dM \left( M, z \right) \int d\xi \left( \frac{\hat{s}}{d\hat{s}} \right) \sqrt{\frac{N_{\ell}(N_{\ell} - 1)}{N_{\ell}(N_{\ell} - 1)}} \sqrt{\frac{N_{\ell}(N_{\ell} - 1)}{N_{\ell}(N_{\ell} - 1)}}$$

$$\times \sqrt{u_{\nu}} \sqrt{u_{\nu}}.$$ (27)
\[
C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.1h}} = \int d\chi \left( \frac{d\chi}{dz} \right) \left( \frac{a^2}{\chi} \right) \hat{J}_\ell(z) \hat{J}_\ell(z) \\
\times \int dMb(M,z) n(M,z) \frac{\langle N_{g} \rangle_v}{\hat{n}_g(z)} u(k|M,z) \\
\times \int dMb(M,z) n(M,z) \frac{\langle N_{g} \rangle_v}{\hat{n}_g(z)} u(k|M,z) \\
\times P_{\text{lin}}(k,z).
\]

(28)

Here, \( \nu \) and \( \nu' \) denote the FIR and NIR bands for Herschel and Spitzer, respectively. Then, we have \( C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.1h}} = C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.1h}} + C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.2h}} \).

The one-halo and two-halo terms of \( C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.1h}} \) are

\[
C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.1h}} = \frac{1}{4\pi} \int d\chi \left( \frac{d\chi}{dz} \right) \left( \frac{a^2}{\chi} \right) \hat{J}_\ell(z) \\
\times \int dMb(M,z) \sqrt{\langle N_{g} \rangle_v (\langle N_{g} \rangle_v - 1)} \nu \\
\times \sqrt{u_{\nu}(k|M,z) \hat{E}_{\text{FIR}}^\nu(M,z) u(k|M,z),}
\]

(29)

\[
C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.2h}} = \frac{1}{4\pi} \int d\chi \left( \frac{d\chi}{dz} \right) \left( \frac{a^2}{\chi} \right) \hat{J}_\ell(z) \\
\times \int dMb(M,z) n(M,z) \frac{\langle N_{g} \rangle_v}{\hat{n}_g(z)} u(k|M,z) \\
\times \int dMb(M,z) n(M,z) u(k|M,z) \hat{E}_{\text{NIR}}^\nu(M,z) \\
\times P_{\text{lin}}(k,z).
\]

(30)

So we get \( C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02}} = C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.1h}} + C_{\ell,\text{FIR} \times \text{NIR}}^{c_{02.2h}} \).

APPENDIX B

FLAT SKY APPROXIMATION

Using the formalism of Hivon et al. (2002), we show how to get the flat sky approximation which allows the use of Fourier transforms instead of spherical harmonics to calculate the power spectrum. The flat sky approximation requires \( \theta \ll 1 \) and \( \ell \gg 1 \). We start by defining the weighted sum over the multipole moments as

\[
M(\ell) = \sqrt{\frac{4\pi}{2\ell + 1}} \sum_{m=-\ell}^{\ell} i^{-m} M_{\ell m} \text{e}^{i m \phi}.
\]

(31)

For this derivation, we need the following approximations for when \( \theta \ll 1 \) and \( \ell \gg 1 \). The first is known as the Jacobi–Anger expansion of the plane wave,

\[
e^{i \ell r} = \sum_{m} i^{m} J_{m}(\ell \theta) \text{e}^{i m (\phi - \phi_0)}.
\]

(32)

where \( J_{m}(\ell \theta) \) are Bessel functions and we use the fact that since \( \theta \) is small \( r \approx \theta \).

Next, to show that \( Y_{\ell m} \) is approximately \( \sqrt{\ell / 2\pi} J_{m}(\ell \theta) \text{e}^{i m \phi} \), we start with the general Legendre equation and perform a change of variables \( \cos(\theta) \Rightarrow \cos \left( \frac{\theta'}{\ell} \right) \).

We use the fact that \( \theta' / \ell \ll 1 \) to simplify, and multiply the equation by \( \theta'^{2} / \ell^{2} \) to get it into the correct form. Now, one can recognize the equation as Bessel's equation with solution \( J_{m}(\theta') = J_{m}(\ell \theta) \).

Putting this all together, we decompose the maps into spherical harmonics and introduce some simplifications to obtain

\[
M(\ell) = \sum_{m} M_{\ell m} Y_{\ell m},
\]

\[
= \sum_{m} M_{\ell m} \sqrt{\frac{\ell}{2\pi}} J_{m}(\ell \theta) \text{e}^{i m \phi},
\]

\[
= \sum_{m} M_{\ell m} \sqrt{\frac{\ell}{2\pi}} J_{m}(\ell \theta) \text{e}^{i m (\phi - \phi_0)},
\]

\[
\approx \int d\ell^{2} M(\ell) e^{i \ell r}.
\]

(33)

APPENDIX C

CROSS-CORRELATION POWER SPECTRUM AND COSMIC VARIANCE

Analogous to the convolution theorem, correlations obey a similar relation except with a complex conjugate. Below, \( \mathcal{F} \) denotes 2D Fourier transforms and \( * \) denotes cross-correlation, where \( M_{1} \) are 2D maps:

\[
\mathcal{F}(M_{1} * M_{2}) = \mathcal{F}(M_{1}) \cdot \mathcal{F}(M_{2})^*.
\]

(34)

Since we can use Fourier transforms as a suitable basis to obtain a power spectrum (see Appendix B), we guarantee that the cross-correlation above is also a power spectrum. Putting it together, for a specific \( \ell \) bin and including Fourier masking, we obtain

\[
\langle C_{\ell} \rangle = \sum_{i \in \ell} \frac{M_{\ell m}^{w}(\ell_{i}, \ell_{i}) M_{\ell m}^{w}(\ell_{i}, \ell_{i})}{\sum_{i \in \ell} M_{\ell m}^{w}(\ell_{i}, \ell_{i})}.
\]

(35)

Here, \( M_{\ell m}^{w}(\ell_{i}, \ell_{i}) \) is a Fourier mask that has a value of one for modes we keep and zero for modes we mask.

The binning is done in annular rings so \( h \geq l_{i}^{2} + l_{j}^{2} \) and \( l_{2} \leq l_{i}^{2} + l_{j}^{2} \).

Cosmic variance, \( \delta C_{\ell} \), is the expected variance on the power spectrum estimate at each \( \ell \) mode and is given by

\[
\delta C_{\ell} = \sqrt{\frac{2}{f_{\text{sky}}(2\ell + 1) \Delta \ell \left( C_{\ell}^{\text{auto}} + N_{\ell} \right)}}.
\]

(36)

where \( N_{\ell} \) is the noise power spectrum generated through jackknife and similar techniques, \( f_{\text{sky}} \) is the fraction of the sky covered by the map, and \( \Delta \ell \) is the width of the \( \ell \)-bin.

For the cross-correlation spectrum the cosmic variance is

\[
\delta C_{\ell} = \sqrt{\frac{1}{f_{\text{sky}}(2\ell + 1) \Delta \ell} \left[ \left( C_{\ell}^{\text{auto}} + N_{\ell} \right) \left( C_{\ell}^{\text{auto}} + N_{\ell} \right) + \left( C_{\ell}^{\text{A} \times \text{B}} \right)^{2} \right]},
\]

(37)

where \( C_{\ell}^{\text{A} \times \text{B}} \) is the cross-correlation power spectrum.
# References

Amblard, A., Cooray, A., Serra, P., et al. 2010, *A&A*, 518, L9

Amblard, A., Cooray, A., Serra, P., et al. 2011, *Natur*, 470, 510

Arendt, R. G., Fixsen, D. J., & Moseley, S. H. 2000, *ApJ*, 536, 506

Ashby, M. L. N., Stern, D., Brodwin, M., et al. 2009, *ApJ*, 701, 428

Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393

Bethermin, M., Dole, H., Lagache, G., et al. 2011, *A&A*, 529, A4

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, arXiv:1403.4259

Cantalupo, C. M., Bprrill, J. D., Jaffee, A. H., et al. 2010, *ApJS*, 187, 212

Cappelluti, N., Kashlinsky, A., Arendt, R. G., et al. 2012, *ApJ*, 769, 68

Carollo, D., Beers, T. C., Chiba, M., et al. 2010, *ApJ*, 712, 692

Chapman, J. F., & Wardle, M. 2006, *MNRAS*, 371, 513

Cooray, A., & Sheth, R. K. 2002, *PR*, 372, 1

Cooray, A., & Yoshida, N. 2004, *MNRAS*, 351, L71

Cooray, A., Smidt, J., De Bernardis, F., et al. 2012, *Natur*, 490, 514

Courteau, S., Widrow, L. M., McDonald, M., et al. 2011, *ApJ*, 739, 20

Dole, H., Lagache, G., Puget, J.-L., et al. 2006, *A&A*, 451, 417

Driver, S. P., Popescu, C. C., Tiffs, R. J., et al. 2007, *MNRAS*, 379, 1022

Dunne, L., Eales, S. A., Edmunds, M. G., et al. 2000, *MNRAS*, 315, 115

Dunne, L., Gomez, H., da Cunha, E. S., et al. 2010, *MNRAS*, 417, 1510

Finkelstein, S., Papovich, C., Ryan, R. E., et al. 2012, arXiv:1206.0735

Fu, H., Julio, E., Cooray, A., et al. 2012, *ApJ*, 753, 12

Fukugita, M. 2011, arXiv:1103.4191

Fukugita, M., & Peebles, P. J. E. 2004, *ApJ*, 616, 643

Gong, Y., & Chen, X. 2007, *PhRvD*, 76, 123007

Gonzalez, A. H., Zabludoff, A. I., & Zaritsky, D. 2005, *ApJ*, 618, 195

Griffin, M. J., Abergel, M. J., Abreu, A., et al. 2010, *A&A*, 518, L3

Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, *ApJ*, 752, 113

Hivon, E., Gorski, K. M., Netterfield, C. B., et al. 2002, *ApJ*, 567, 2

Hwang, H. S., Elbaz, D., Magnis, G., et al. 2010, *MNRAS*, 409, 75

James, A., Dunne, L., Eales, S., et al. 2002, *MNRAS*, 335, 753

Kashlinsky, A., Arendt, R. G., Ashby, M. L. N., et al. 2012, *ApJ*, 753, 63

Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2005, *Natur*, 438, 45

Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2007, *ApJL*, 654, L5

Kashlinsky, A., Mather, J. C., Helgason, K., et al. 2014, arXiv:1412.5561

Krick, J. E., & Bernstein, R. A. 2007, *ApJ*, 134, 466

Lee, K.-S., Giavalisco, M., Conroy, C., et al. 2009, *ApJ*, 695, 368

Lin, Y. T., Mohr, J. J., & Stanford, S. A. 2004, *ApJ*, 610, 745

Matsumoto, T., Seo, H. J., Jeong, W.-S., et al. 2011, *ApJ*, 742, 124

Menard, B., & Fukugita, M. 2012, *ApJ*, 754, 116

Menard, B., Scranton, R., Fukugita, M., & Richards, G. 2010, *MNRAS*, 405, 1025

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *JChPh*, 21, 1087

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493

Ott, S. 2010, in ASP Conf. Ser. 434, Astronomical Data Analysis Software and Systems XIX, ed. Y. Mizumoto, K.-I. Morita & M. Ohishi (San Francisco, CA: ASP), 139

Pillbrratt, G. L., Riedlinger, J. R., Passvogel, T., et al. 2010, *A&A*, 518, L1

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, *A&A*, 571, A16

Purcell, C. W., Bullock, J. S., & Zentner, A. R. 2008, *MNRAS*, 391, 550

Schmidt, S., Menard, B., Scranton, R., et al. 2014, *MNRAS*, 446, 2696

Shang, C., Haiman, Z., Knox, L., et al. 2011, arXiv:1109.1522

Sheth, R. K., & Tormen, G. 1999, *MNRAS*, 308, 119

Thacker, C., Cooray, A., Smidt, J. S., et al. 2013, *ApJ*, 768, 58

Thompson, R., Eisenstein, D., Fan, X., et al. 2007, *ApJ*, 666, 658

Viero, M. P., Wang, L., Zemcov, M., et al. 2013, *ApJ*, 772, 77

Yue, B., Ferrara, A., Salvaterra, R., Xu, Y., & Chen, X. 2013, *MNRAS*, 433, 1556

Yue, B., Ferrara, A., Salvaterra, R., Xu, Y., & Chen, X. 2014, *MNRAS*, 440, 1263

Zemcov, M., Smidt, J., Arau, T., et al. 2014, *Sci*, 346, 732