Title
A Multiwavelength Study of the Epoch of Reionization through Intensity Fluctuations

Author
Mitchell-Wynne, Ketron

Publication Date
2016

UNIVERSITY OF CALIFORNIA,
IRVINE

A Multiwavelength Study of the Epoch of Reionization through Intensity Fluctuations

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Ketron Mitchell-Wynne

Dissertation Committee:
Professor Asantha Cooray, Chair
Professor Michael Cooper
Professor James Bullock

2016

DEDICATION

To Koogie, and all the rest of my family.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

I would be remiss if I didn't first acknowledge the long-standing and unwavering support from my adviser, Asantha Cooray. I started working in the Cooray group seven years ago, and despite a very few and ephemeral periods of difficulty, it has overwhelmingly been exciting and rewarding. Asantha has consistently come up with stimulating and unique research projects, supplied us with the necessary resources to carry out our research, and shared the insight that leads to interesting, impactful solutions. My time spent doing research in the Cooray group has indeed been edifying. Even after leaving Irvine, I suspect I will continue to be an acolyte of the forthcoming work from Asantha's lab.

Secondly, I would like to thank the many postdocs who helped me, over many years, to become a better researcher. Joseph Smidt, Yan Gong, Alex Amblard and Paolo Serra played cricial roles in my early development as a scientist. Hooshang Nayyeri, Chang Feng, Julie Wardlow, Caitlin Casey, Francesco De Bernardis and Hai Fu have also been very helpful postdocs over time. I am also grateful for the help from my many collaborators, especially Mário Santos, Henry Ferguson, Anton Koekemoer, Tim Dolch and Steve Finkelstein. I have been fortunate to have had many worthwile conversations, often about research, with many of my fellow graduate students. I want to thank Adithya Manohara, Sean Fillingham, Chris Karwin, Cameron Thacker, Jae Calanog, Jon-Michael O'Bryan, Nicholas Timmons, Ryan Keeley, Ali Khostovan and Chris Frazer. I suppose I should also thank Mike Zemcov, as he taught me, among other things, that when I'm having trouble finding a solution to a problem it would be more apropos to "drink a glass of water and think about it" rather than immediately seeking help from others. Incidentally I did eventually find this to be true and effective.

Finally I want to thank all my friends and family, and acknowledge the significant investment my parents have made, both financially and emotionally, to get me here. Frequent meetings with my grandparents in Carlsbad have kept me level-headed, and my ongoing friendships with Peter Hurley and Ian Lockey have continued to keep me "calibrated", both personally and professionally. And Yasmin Tina Sanati has been a resplendent presence in what would otherwise be a much colder and darker Universe.

The text of this thesis is a reprint of the material as it appears in "Cross-correlation between X-ray and Optical/Near-Infrared Background Intensity Fluctuations" (ApJ, submitted), "Ultraviolet Luminosity Density of the Universe During the Epoch of Reionization" (Nature Communications, 6, 7945), and "HerMES: A Statistical Measurement of the Redshift Distribution of Herschel-SPIRE Sources Using the Cross-correlation Technique" (ApJ, 753, 23).

I am grateful for the support from NASA grants NNX16AF39G, NNX16AJ69G, NNX15AQ06A, HST-AR-13886, and GAANN P200A150121.

CURRICULUM VITAE

Ketron Mitchell-Wynne

EDUCATION

Doctor of Philosophy in Physics	**2016**
University of California, Irvine	*Irvine, CA*
Master of Science in Physics	**2015**
University of California, Irvine	*Irvine, CA*
Master of Science in Astronomy & Astrophysics (equivalence)	**2013**
University of Lisbon	*Lisbon, Portugal*
Bachelor of Science in Physics	**2010**
University of California, Irvine	*Irvine, CA*

RESEARCH EXPERIENCE

Graduate Research Assistant	**2014–2016**
University of California, Irvine	*Irvine, California*
Graduate Research Assistant	**2012–2013**
University of Lisbon	*Lisbon, Portugal*
Research Assistant	**2009–2012**
University of California, Irvine	*Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant	**2013**
University of California, Irvine	*Irvine, CA*

REFEREED JOURNAL PUBLICATIONS

1. "Angular Clustering of Strongly Lensed Submillimeter Galaxies"
   Ketron Mitchell-Wynne, Asantha Cooray, Hooshang Nayyeri. *ApJL in prep.*

2. "Cross-correlation between X-ray and Optical/Near-Infrared Background Intensity Fluctuations"
   Ketron Mitchell-Wynne et al. 2016. ApJ, 832, 104

3. "Axion Decay and Anisotropy of Near-IR Extragalactic Background Light"
   Gong, Yan et al. 2015, ApJ, 825, 104

4. "Ultraviolet Luminosity Density of the Universe During the Epoch of Reionization"
   Ketron Mitchell-Wynne et al. 2015, Nature Communications, 6, 7945

5. "Cross-Correlation of Near and Far-Infrared Background Anisotropies as Traced by Spitzer and Herschel"
   Thacker, C. et al. 2014, ApJ, 811, 125

6. "Beyond stacking: a maximum-likelihood method to constrain radio source counts below the detection threshold"
   Mitchell-Wynne, K., Santos, M. G., Afonso, J., & Jarvis, M. J. 2014, MNRAS, 437, 2270

7. "The Cosmic Infrared Background Experiment (CIBER): The Wide-field Imagers"
   Bock, J., Sullivan, I., Arai, T., et al. 2013, ApJS, 207, 32

8. "H-ATLAS: The Cosmic Abundance of Dust from the Far-infrared Background Power Spectrum"
   Thacker, C., Cooray, A., Smidt, J., et al. 2013, APJ, 768, 58

9. "HerMES: A Statistical Measurement of the Redshift Distribution of Herschel-SPIRE Sources Using the Cross-correlation Technique"
   Mitchell-Wynne, K., Cooray, A., Gong, Y., et al. 2012, ApJ, 753, 23

10. "Sub-millimetre galaxies reside in dark matter halos with masses greater than 3×10^{11} solar masses"
    A. Amblard et al. 2011, Nature 470, 510

11. "HerMES: Halo Occupation Number and Bias Properties of Dusty Galaxies from Angular Clustering Measurements"
    A. Cooray et al. 2010, A&A, Herschel Special Issue, 518 L22

12. "HerMES: SPIRE galaxy number counts at 250, 350 and 500 microns"
    S. J. Oliver et al. 2010, A&A, Herschel Special Issue, 518 L21

HONORS & AWARDS

GAANN graduate fellowship **2015**
University of California, Irvine

NASA Group Achievement Award **2014**
University of California, Irvine

International Doctorate Network in Particle Physics, Astrophysics and Cosmology (IDPASC) Graduate Fellowship **2012**
University of Lisbon

Undergraduate Research Opportunities Program (UROP) **2009**
University of California, Irvine

Summer Undergraduate Research Opportunities Program (SURP) **2009**
University of California, Irvine

ABSTRACT OF THE DISSERTATION

A Multiwavelength Study of the Epoch of Reionization through Intensity Fluctuations

By

Ketron Mitchell-Wynne

Doctor of Philosophy in Physics

University of California, Irvine, 2016

Professor Asantha Cooray, Chair

The diffuse cosmological background light that pervades our Universe continues to be one of the best probes to uncover the astrophysical environments of the early Universe. In our attempt to gain a more holistic understanding of our present cosmological situation, we seek to understand the formation and evolution of the first stars and galaxies that came into existence. It is thought that these first-light galaxies are the progenitors to our Milky Way, and all other evolved galaxies. The spatial fluctuations of the extragalactic background light trace the total emission from all stars and galaxies in the Universe, and these first-light galaxies have left a measurable imprint on the cosmic infrared backgroud (CIB) light. In the first part of this thesis I describe measurements which suggest a detection of this signal, using a large subset of data from the largest allocation of *Hubble* time to date (CANDELS). This is followed by a cross-correlation analysis of the CIB with the cosmic X-ray background (CXB), which can give some insight as to whether or not X-ray emitting sources were present during this first-light epoch. I then shift towards the far-infrared (submillimeter) regime and discuss a statistical measurement of the redshift distribution of *Herschel* galaxies. Finally I show two-point correlation measurements of strongly lensed submillimeter galaxies, and the bias that this population may be introducing into past and current clustering analyses, which can lead to erroneous dark matter halo mass measurements from such studies.

Chapter 1

Introduction to the EBL and Submillimeter Galaxies

The cosmic extragalactic background light (EBL) traces emission from a wide range of sources, extending back to cosmic times as early as 380,000 years after the Big Bang ($z = 1100$), or the period of recombination (Peebles 1968; see Fig 1.1). Studies of the redshifted emission from this early era peak in the microwave (the cosmic microwave background or CMB) and have revealed the fundamental λCDM values of the Universe, such as the age, curvature, baryonic matter density content, and dark matter and dark energy density contents (Planck Collaboration et al. 2015 and references therein). Measurements of the CMB provide compelling evidence in support of the Big Bang theory, and exemplify the tremendous amount of information that can be extracted from studies of the EBL.

After the period of recombination came the "dark ages", where the Universe continued to undergo expansion, with a baryonic matter composition of predominately neutral Hydrogen, initially distributed over a relatively smooth density field. Over time, slight perdurbations in the density field grew into larger potential wells through gravitational interactions. The

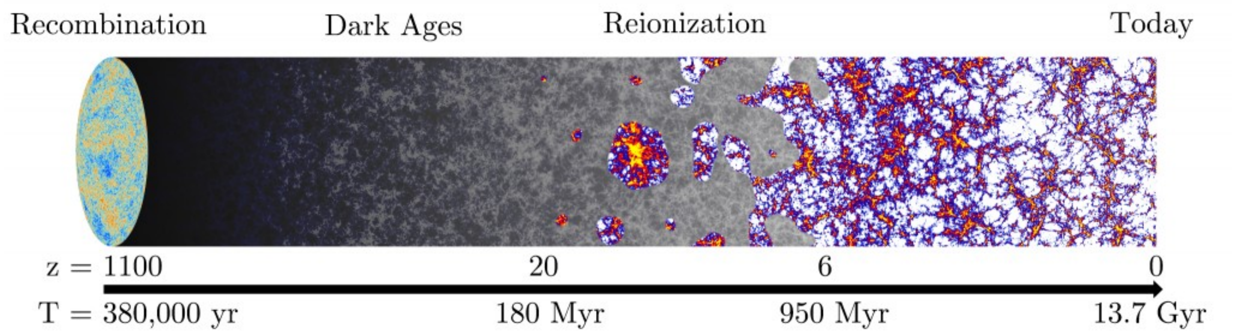

Figure 1.1 Graphical representation of the history of the Universe. We are interested in probing the first sources in the Universe, around redshift ~ 10. Image credit: Alexander Kaurov

formation and early evolution of the first galaxies in the Universe occurred some time after the dark ages, when the coalescence of gravitationally bound masses formed in complex structures, with a spatial distribution that can be traced back to primordial overdensities, or quantum fluctuations (Loeb & Barkana, 2001; Fan et al., 2006b). The ultraviolet (UV) photons from these first sources initiated the reionization of the surrounding neutral medium, thus ending the dark ages and beginning the era of a transparent cosmos, which we are increasingly familiar with today. This structure formation marks the end of the dark ages the beginning of the ”epoch of reionization” (EoR), which, much like the CMB, has left a faint but detectable imprint on the EBL, redshifted to near-infrared (NIR) wavelengths. This wavelength regime of the EBL is the cosmic infrared background (CIB), and will constitue a large part of the study discussed in this thesis.

The luminosity per unit volume of the UV photons at a rest wavelength around 1500Å (ρ_{UV}) during this reionization period is an important quantity to measure, as it traces the star formation and evolution of these ionizing sources. The traditional method to measure the UV luminosity density of the universe, ρ_{UV}, during the epoch of reionization, involves searching for candidate galaxies at $z > 6$ through their Lyman-dropout signature (Oesch et al., 2014; Zheng et al., 2012; Coe et al., 2013; Bouwens et al., 2014; Finkelstein et al., 2014) and then constructing the luminosity function of those detected galaxies based on the observed number counts. This luminosity function is then extraploated to a fainter absolute magnitude and integrated in luminosity to calculate ρ_{UV}.

There is a second way to quantify ρ_{UV}, without extrapolations. This involves a measurement of the EBL and, in particular, the angular power spectrum of the cosmic infrared background (CIB) intensity fluctuations. Because these intensity fluctuations are the result of emissions throughout the cosmic time, the signal we measure today is the sum of many different emission components, from nearby in our Galaxy to distant sources. If the integrated intensity from reionization can be reliably separated from that of foreground signals, we may be able to make an accounting of the total luminosity density of UV photons from reionization. Just as Lyman-dropout galaxies are detected in deep sky surveys, there is a way to achieve such a separation with fluctuations. Due to redshifting of the photons arising from sources present during reionization, their emission, as seen today, is expected to peak between 0.9 and 1.1 μm. This assumes that the reionization occurred around z ~7 to 9, consistent with optical depth to electron scattering as measured by Planck (Planck Collaboration et al., 2015). Due to absorption of ionizing UV photons, there is no contribution shortward of the redshifted Lyman break around 0.8 μm (Santos et al., 2002; Salvaterra & Ferrara, 2003). Spatial fluctuations of the EBL centered around 1 μm thus provide the best mechanism to discriminate the signal generated by galaxies present during reionization (Cooray et al., 2004; Fernandez et al., 2010) from those at lower redshifts, based on the strength of the drop-out signature in the fluctuations measured in different bands.

Intensities of the EBL as a function of wavelength are shown in Fig 1.2. Absolute intensity measurements of the EBL are extremely difficult because of significant foreground comtamination. The brightness of the CIB is on the order of 10 nW m^{-2} sr^{-1}. On a clear night, the brightness of the full moon is $\sim 10^{11}$ nW m^{-2} sr^{-1}, or 10 billion times brighter than the CIB. Omitting the moon from observations is easy enough – don't point the telescope towards it. However CIB contaminants are numerous, and largely come from atmospheric

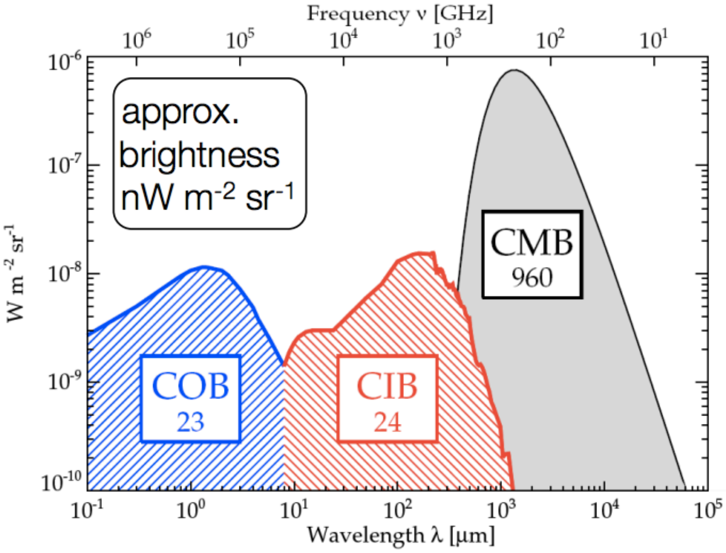

Figure 1.2 Spectral enegry distributions of the EBL. The left axis is the brightness. Taken from Dole et al. (2006).

intereference and sunlight scattered by interstellar dust (Kelsall et al., 1998), in addition to foreground stars and galaxies. The latter are effectively removed by producing a resolved foreground sorce mask. We strive to measure the spatial fluctuations of the CIB instead of absolute intensities, because in this way we can largely overcome foreground contaminants while still making meaningful insights from the EBL measurements. As an example, consider a diffuse signal, statistically described by some (clustered) spatial correlation function, with a bright, constant foreground sitting atop. With or without that constant foreground, the spatial fluctions are still there, and can be readily measured. An example of what these CIB fluctuations may look like is shown in Fig 1.3; the fluctuations in the CIB reveal themselves after all the foreground stars and galaxies are removed from the image.

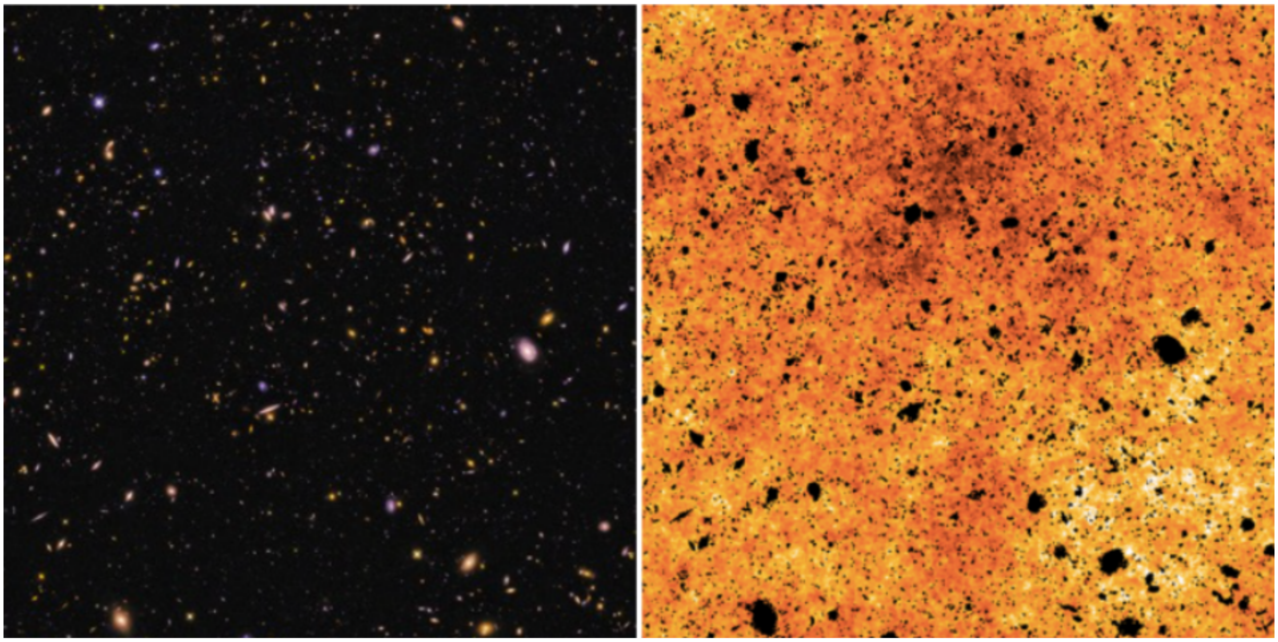

Figure 1.3 A large-area NIR mosaic (left) has a foreground source mask applied to reveal the spatially correlated CIB intensity fluctuations (right).

1.1 Review of Power Spectrum Formalism

These intensity fluctuations are described statistically by the angular power spectrum, C_ℓ, which is calculated by decomposing the sky into spherical harmonics and measuring the variance of the $a_{\ell m}$ coefficients:

$$\langle a_{\ell m}\, a^*_{\ell' m'} \rangle = \delta_{\ell\ell'}\, \delta_{mm'}\, C_\ell, \tag{1.1}$$

where the $a_{\ell m}$'s are the spherical harmonic coefficients, an observable quantity. A scalar field, $\Theta(\boldsymbol{u})$, defined over the full sky is decomposed in spherical harmonic coefficients as

$$\Theta(\boldsymbol{u}) = \sum_{\ell>0} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_\ell^m(\boldsymbol{u}), \tag{1.2}$$

or equivalently

$$a_{\ell m} = \int d\boldsymbol{u}\, \Theta(\boldsymbol{u}) Y^*_{\ell m}(\boldsymbol{u}). \tag{1.3}$$

With noiseless data and full sky coverage, the C_ℓ's are estimated as

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \tag{1.4}$$

This is the quantity we are most interested in obtaining. However, for our purposes we will not be observing the full sky, as was done with previous CMB experiments, but instead work over very small angular scales due to limitations in available datasets. In this case, $\ell \gg 1$ and $\theta \ll 1$, and it turns out that a very good approximation to the power spectrum can be found using a 2-dimensional Fourier decomposition. The trick to uncovering this correspondence is to approximate the Y_ℓ^m's as Bessel functions, $J_m(l\theta)$, and use the Jacobi-Anger expansion on the plane waves:

$$Y_l^m \approx J_m(l\theta)\sqrt{\frac{l}{2\pi}} e^{im\varphi}\,, \tag{1.5}$$

and the plane wave expansion:

$$\begin{aligned} e^{i\boldsymbol{\ell}\cdot\hat{\mathbf{n}}} &= \sum_m i^m J_m(l\theta) e^{im(\varphi-\varphi_l)} && (1.6)\\ &\approx \sqrt{\frac{2\pi}{l}} \sum_m i^m Y_l^m e^{im\varphi_l}\,. && (1.7) \end{aligned}$$

See Hu (2000) and the Appendix of Thacker et al. (2013) for the complete derivation. Fourier decomposition is a somewhat straightforward analysis to do if one has a clean dataset to work with (which is typically not the case, ever). The subsequent difficulty is formally explaining the astrophysical origin of these fluctuations.

1.2 State-of-the-art CIB and CXB Measurements

There are existing measurements of the EBL intensity fluctuations, though their origin remain uncertain. This is mostly due to the fact that the previous measurements of EBL fluctuations have until now been limited to wavelengths greater than $1.1\,\mu$m, with the best measurements performed at $3.6\,\mu$m (Kashlinsky et al., 2012; Matsumoto et al., 2011; Cooray et al., 2012b; Zemcov et al., 2014; Seo et al., 2015). These studies have been interpreted with models involving populations of sources present during reionization at redshift $z > 8$, direct collapse and other primordial blackholes at $z > 12$ (Cappelluti et al., 2013; Yue et al., 2013a), and with stellar emission from tidally stripped intergalactic stars residing in dark matter halos, or the "intra-halo light" (IHL; Cooray et al. 2012b) at $z < 3$. The IHL is diffuse stars in dark matter halos due to galaxy mergers and tidal interactions. While the relative strengths of these various contributions are still unknown, we expect the signal from high-redshift galaxies to be separable from low-redshift contributions, including those from faint nearby dwarf galaxies (Helgason et al., 2012), through a multi-wavelength fluctuation study spanning the 1 μm range, including in the optical ($\lambda \lesssim 1$ μm) and near-IR ($\lambda \gtrsim 1$ μm) wavelengths.

In addition to galaxies throughout cosmic history, the CIB also contains the signatures from diffuse sources of emission (Santos et al., 2002), such as the intra-halo light (IHL; Cooray et al. 2012a; Zemcov et al. 2014), or sources that are not traditionally counted as galaxies. For example, the near-IR background (NIRB) is generally considered to be a probe of reionization and the sources present during that epoch. Indeed it was first posited by Kashlinsky et al. (2005) that NIRB clustering measured on arcmin scales was due to zero-metallicity Population III stars during reionization. Theoretical calculations by

Cooray et al. (2012b) subsequently showed that the measured amplitude in previous works claiming Population III signatures (Kashlinsky et al., 2007; Matsumoto et al., 2011) is an order of magnitude larger than theoretical expectations (Komatsu et al., 2011). This shifted the discussion to include the possibility of a number of other sources contributing to the NIRB, including low-redshift galaxies (Sullivan et al., 2007; Thompson et al., 2007; Cooray et al., 2007; Chary et al., 2008). The contribution from low-z galaxies was then explicitly constrained by Helgason et al. (2012) by reconstructing the background signal from the known low-z galaxy populations.

There still exists an excess signal in the NIRB fluctuation power spectra, in addition to the Pop III signature and low-z galaxy population. As mentioned in the beginning, one such low redshift contribution is now identified to be IHL (Cooray et al., 2012a; Zemcov et al., 2014). An additional and important source population to consider involves direct collapse black holes (DCBHs; e.g. Yue et al. 2013a, 2014) that are hypothesized to explain the presence of massive billion to ten billion solar mass black holes in luminous quasars at $z \sim 6$ (e.g. Fan et al. 2006a). Actively accreting DCBHs are expected to leave an imprint in the X-ray background. Similarly, the rest-frame UV photons emitted by these sources will now be visible in the infrared. If these sources were present during the earliest epochs of galaxy and black hole formation, it would then make sense that the cosmic X-ray background (CXB) would be correlated with the intensity fluctuations of the near-IR background.

This cross-correlation was measured to be statistically significant at 3.6 and 4.5 μm in the *Spitzer*/Extended Groth Strip (EGS; Davis et al. 2007) with the broad-band 0.5-2 keV background fluctuations (Cappelluti et al. 2013, hereafter C13) of the same field. Cross-correlations between *Spitzer* and the harder X-ray band at 2-7 keV were found to be statistically less significant. This detection in the soft band and the non-detection in the hard band was subsequently modeled in Helgason et al. (2014) where they found that X-ray emission from AGN, normal galaxies and virialized hot gas at cannot account for the large-scale correlation. They suggest that the CIB and CXB signals are from the same, unknown source population at high-z, perhaps miniquasars as proposed in Cappelluti et al. (2012) and explained as DCBHs in Yue et al. (2013b). In order to gain more insight, and to test whether the cross-correlation is present in other legacy fields, it is essential that we consider more optical and infrared bands and extend the studies to other areas.

The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) has opened up a new window of opportunity in measuring the intensity fluctuations at near-infrared (NIR) and optical wavelengths, ranging between 0.6 and 1.6 μm. Statistical measurements (namely the angular power spectrum) of the background fluctuations at this wavelength range were only recently measured (Mitchell-Wynne et al. 2015, hereafter MW15) and are interpreted to have contributions from IHL, faint low-z galaxies (Helgason et al., 2012), possibly diffuse Galactic light (DGL) from dust in our Milky Way, and first-light galaxies during the reionization epoch in the two 1.25 and 1.6 μm *Hubble*/WFC3 bands. Angular power spectra measurements have also been made over a 10 deg^2 field in the IRAC 3.6 and 4.5 μm bands, with similar interpretations of IHL and faint low-z residual galaxies (Cooray et al., 2012a), with no contribution from high-redshift sources. The measurements with IRAC by independent groups (e.g., Kashlinsky et al. 2007; 2012) agree with each other, though interpretations of those clustered fluctuations remain different.

1.3 Motivations in the Submillimter

Submillimeter galaxies (SMGs) are the most ardently star forming systems in the Universe, with the brightest of them having infrared luminosities in excess of 10^{13} $L_\odot$ (Casey et al., 2014). Understanding the role SMGs play in the evolutionary process of galaxies is then extremely important in our quest to gain a holistic picture of our place in the cosmos. The properties of these dusty, star-forming galaxies detected by the Spectral and Photometric Imaging Receiver (SPIRE; Griffin et al. 2010) aboard the *Herschel Space Observatory*[1] (Pilbratt et al. 2010) at submilllimeter (submm) wavelengths provide important clues to the nature of dusty star-formation and the role of galaxy mergers in triggering such star-formation in distant galaxies. However, the redshift distribution of these galaxies has yet to be determined observationally.

The low spatial resolution of *Herschel*-SPIRE observations complicate the identification of counterparts at optical and NIR wavelengths. Moreover, the optical emission from these star-bursting galaxies is highly extinct and could potentially bias optical spectroscopy observations to low-redshift bright galaxies. Instead of optical or IR spectroscopy, millimeter (mm) and submm wave spectroscopy can be pursued targeting fine-structure and molecular lines such as CO and [CII]. Such measurements, unfortunately, are currently limited to a handful of the brightest sources – mostly the rarely lensed sub-mm galaxies (e.g., Lupu et al. 2010; Scott et al. 2011; Riechers et al. 2011; Harris et al. 2012), as existing instrumental capabilities do not allow large CO or [CII] surveys of typical sub-mm galaxies.

There have been a few other approaches to obtain the $N(z)$ of sources at these wavelengths. A statistical approach based on photometry alone, using SPIRE colors, was considered in Amblard et al. (2010; see also Lapi et al. 2011), but such techniques are subject to uncertainties on the assumed spectral energy distribution (SED) of the galaxies at submm wavelengths. These generally involve isothermal SED models, where the redshift distribution is degenerate with the assumed dust temperature distribution. Marsden et al. (Marsden et al., 2009) employed stacking methods to effectively measure the CIB as a function of redshift, and Béthermin et al. (Béthermin et al., 2012b) have recently measured deep source counts as a function of redshift, also via stacking, which is compared to with our results, described in Chapter 4.

---

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

There we pursue a second statistical approach to measure the SPIRE galaxy redshift distribution using the spatial clustering of the submm population relative to clustering of galaxies with an a priori known redshift distribution (Schneider et al. 2006; Newman 2008; Zhang et al. 2010). The unknown submm redshift distribution can be estimated via the strength of its cross-correlation relative to galaxy samples of known redshifts. Modeling also requires that the clustering bias factors of all galaxies be determined jointly through a combination of auto and cross-correlation functions. The key advantage of this technique is that it does not require cross-identification of SPIRE sources in optical and IR catalogs.

Cross- and auto-correlations in Chapters 4 and 5 are done with the two-point angular correlation function, $w(\theta)$. One can think of this as the discrete equivalent of the power spectrum C_ℓ. In the previous section we discussed the statistical evaluation of diffuse signals; $w(\theta)$ will be implemented with discrete galaxy positions. The angular correlation function $w(\theta)$ is the excess probability, P, of finding a pair of galaxies at some angular separation θ

$$dP = N[1 + w(\theta)]d\Omega, \tag{1.8}$$

over some solid angle $d\Omega$ with a mean number density of galaxies per unit area N. On relatively small angular scales, we again use the flat sky approximation. With this approximation, $w(\theta)$ and the power spectrum $P(k)$ are related by a Hankel transform:

$$w(\theta) = \int_0^\infty P(k) J_0(k\theta) \frac{dk}{k}, \tag{1.9}$$

with the Bessel function J_0 and wave number k. $P(k)$ is simply related to C_ℓ by converting the wavenumbers to multipoles, where the conversion is ℓ $[2\pi \text{ rad}^{-1}] = 2{\cdot}60{\cdot}180{\cdot}k$ $[\text{arcmin}^{-1}]$.

1.4 Gravitational Lensing in the Submm

Strong gravitational lensing of faint, high-redshift submillimeter (submm) galaxies (SMGs) provides us with an opportunity to study the nature and evolution of these sources at $z \gtrsim 1.5$ (Bussmann et al., 2013), where their fluxes would be otherwise too faint to detect without the flux boosting of gravitational lensing. SMGs have a sufficiently large and negative K-correction that high-z galaxies observed at submm wavelengths have roughly constant flux densities between $z \sim 1 - 8$ (Blain & Longair, 1993; Casey et al., 2014). The typical intrinsic flux density of a high-z 500 μm source is 10 mJy (Ivison et al., 2010; Béthermin et al., 2012a,c) and the number counts drop off rapidly at higher flux densities (Clements et al., 2010; Oliver et al., 2010). These characteristics together make the 500 μm channel ideal for easy identification of lensed sources. However because these events are so rare, a necessary requirement for detecting a statistically sizeable sample of lensed galaxies is wide-area surveys.

Negrello et al. (2007, 2010) and Wardlow et al. (2013) have demonstrated that a simple flux cut in the longest band of the Spectral and Photometric Imaging REceiver (SPIRE) instrument (Griffin et al., 2010) onboard the *Herschel Space Observatory* (*Herschel*, Pilbratt et al. 2010) can be used to classify lensed galaxies with nearly a 100% efficiency. In addition to lensed galaxies, this simple criterion, $S_{500} > 100$ mJy, also picks up low-z spirals (Serjeant & Harrison, 2005) and radio-loud blazars (de Zotti et al., 2005). These interlopers are readily removed by cross-matching potential lens sources with shallow surveys at optical and radio wavelengths, e.g. Wardlow et al. (2013); Nayyeri et al. (2016).

The widest-area *Herschel* surveys, the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012) and the *Herschel* Astrophysical Terahertz Large-Area Survey (H-ATLAS; Eales et al. 2010), cover hundreds of square degrees of sky with SPIRE. Here we focus on lensed sources detected in the HerMES Large Mode Survey (HeLMS), which covers $\sim$

270 deg^2, and the *Herschel* Stripe 82 Survey (HerS; Viero et al. 2014), which overlaps with HeLMS and covers $\sim$ 80 deg^2. Altogether we perform statistics on 72 lensed candidate sources over a total area of $\sim$ 350 deg^2.

In this work we consider the angular clustering of *Herschel* lensed sources and the effect their clustering has on previous measurements of the overall *Herschel* SMG population. Cooray et al. (2010) used the clustering of submm galaxies over 10 deg^2 and found what appears to be clear evidence of a 1-halo term at arcmin angular scales. With model fits, they infer the minimum dark matter halo mass a submm galaxie can reside in. However they did not include the (unknown) effects of the underlying lensed distribution in their $w(\theta)$ measurements. Although lenses are identified with $S_{500\,\mu m} > 100$ mJy, there still exists a distribution extending down to lower flux densities, which effectively boosts galaxies with lower intrinsic luminosities above the lower threshold flux cut (typically 3-5σ) performed in clustering studies.

Chapter 2

UV luminosity density of the universe during the epoch of reionization

Here we present results from a multi-wavelength fluctuation study using data from the Hubble Space Telescope (HST) that span across the interesting wavelength range centered at 1 μm, in an attempt to measure ρ_{UV} during the EoR. We begin by describing our data selection, data reduction, map making procedure, and analysis process with which we obtained multi-band angular power spectra of optical and near-IR background fluctuations. Our analysis begins with the calibrated, flat-fielded exposures from HST using two instruments Advanced Camera for Surveys (ACS) and Wide-Field Camera 3 (WFC3). The data are publicly available on the Barbara A. Mikulski Archive for Space Telescopes (MAST).[1] Observation IDs can be found in table 2.1.

---

[1] https://archive.stsci.edu/hst/search.php

2.1 Survey and Field Selection

The Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) is a multi-wavelength survey of multiple extragalactic fields, conducted through the largest allocation of HST orbits. The fields surveyed by CANDELS are some of the well-known and standard extragalactic fields in cosmology. They have ample ancillary data including previous observations with HST. The CANDELS survey, along with archival data, are together a dataset which is sufficiently wide and deep to measure optical and infrared background light fluctuations between 0.606 and 1.60 μm. Although the CANDELS observations cover multiple fields, we restrict our analysis to those of the Great Observatories Origins Deep Survey-South (GOODS-S; Giavalisco et al. 2004), a deep field with an area of approximately 120 square arcminutes. As part of the analysis presented here, we also considered and studied fluctuations in other CANDELS fields but found many of them, e.g. the UKIDSS Ultra-Deep Survey field (UDS; Lawrence et al. 2007; Koekemoer et al. 2011) area, to be lacking adequate exposures and to have significant gaps between the WFC3 frames that were taken as part of CANDELS. While such poor dithering patterns, as can be seen in Fig. 18 of Grogin et al. (2011), are a significant draw-back for fluctuation studies since one cannot calibrate the full mosaic to a consistent background level without introducing artifical gradients to the background intensity. Such dithering patterns were pursued by CANDELS to maximize the total area covered with WFC3 in areas such as UDS. The increase in area is of benefit to studies that aim to detect rare galaxies, such as Lyman-break drop-outs at $z > 5$.

We also encountered one other additional issue during our survey and data selection process. After the replacement of the ACS CCD Electronics Box during the fourth Hubble servicing mission (SM4), ACS imaging data are plagued with horizontal striping dominated by 1/f noise. A rudimentary fluctuation analysis in the UDS field clearly shows the post-SM4 bias striping is contaminating the angular power spectrum to such an extent that the systematics dominate the measurement. This effect is amplified in CANDELS areas such as UDS and EGS as the position angle of the ACS observations of those fields is constant. Attempts to mitigate the striping bias with a de-striping algorithm (Grogin et al., 2010a), while incorporating the redidual striping in a $T(\ell)$ (see section 2.5.3) specific to theCANDELS wide area ACS observations proved unsuccessful. A further complication in UDS is the tendency of ACS frames to introduce a Moiré pattern (correlated noise) when the pixel scale is modified in a low signal-to-noise area. This directly propigates into the angular power spectrum as a high-amplitude noise. Through simulations we found out that with an increased number of ACS frames taken with multiple position angles effectively remove the Moiré pattern upon repixelization and ameliorates the bias-striping issue. In essence, a more robust dithering pattern fixes many of the potential systematic issues that can contaminate fluctuation studies.

Given the poor dithering patterns in Hubble wide-fields, our final selection was reduced to a choice between GOODS-North (Giavalisco et al., 2004) and GOODS-South. The choice between the two is somewhat arbitrary as both fields have multi-band coverage. To improve signal-to-noise ratio in the fluctuation measurements in the optical, especially to study the drop-out signal coming from high-redshift faint galaxies, we also needed a wide-band filter. Finally, our systematic study showed that the F606W dither pattern in GOODS-S is far better than the dither pattern in GOODS-N due to some position angle issues as all of the CANDELS observations were done in the parallel mode. The GOODS-N dataset has a larger number of frames with clear overall offsets resulting from scattered light than does GOODS-S. GOODS-S also has the largest area that overlaps in five HST bands with ACS and WFC3, allowing not only auto power spectra in a single wavelength but also cross power spectra between multiple wavelengths. With the best dither pattern, least systermatics, and the widest overlap in multiple bands, GOODS-S easily became the field of choice for the present study.

Filter	Proposal	IDs							
F775W	9575								
F850LP	9500	9978	10086						
ACS	9425	9803	10189	10258					
WFC3	11359	12060	12061	12062					
F606W	9500	9978	11563	12007	12060	12062	12099	12461	12534

Table 2.1 Proposal ID's for each filter. The ACS and WFC3 rows show the proposals which are common between all the bands in each instrument. For each proposal we did not necessarily use all the frames, specifically those from deep surveys.

2.2 Data reduction

We assembled our own collection of falt-fieled (FLT) frames, comprised of some or all of the data from ten different HST proposals (Beckwith et al., 2006; Giavalisco et al., 2004; Grogin et al., 2011; Koekemoer et al., 2011; Riess et al., 2004; Rix et al., 2004; Thompson et al., 2005) for ACS and WFC3. These data are also supplemented by the Early Release Science observations (Windhorst et al., 2011) (ERS). ACS observation dates span up to a decade for some bands (2002-2012), and WFC3 frames are from the period 2009 to 2012. Exposure times for ACS range between 180s-1469s per frame, while 403s-903s per frame for WFC3. The five bands selected for this study are F606W, F775W, F850LP, F125W, and F160W. They provide 363, 234, 428, 420 and 420 frames, respectively. The data we used are initially calibrated and flat-fielded. Altogether the data we have used represents a formidable reduction effort. The map-making process and associated simulations required a considerable amount of CPU time.

All data were downloaded from the MAST archive and initially reduced with PyRAF version 2.1.1. MAST queries are reprocessed "on-the-fly", which entails using the most recent calibration files. Thus FLT frames we retrieved from the archive had standard calibrations of bias and dark frame subtraction, along with flat-field correction, already performed. We identified cosmic rays in the the FLT frames with the CRCLEAN PyRAF module; sub-arcsecond

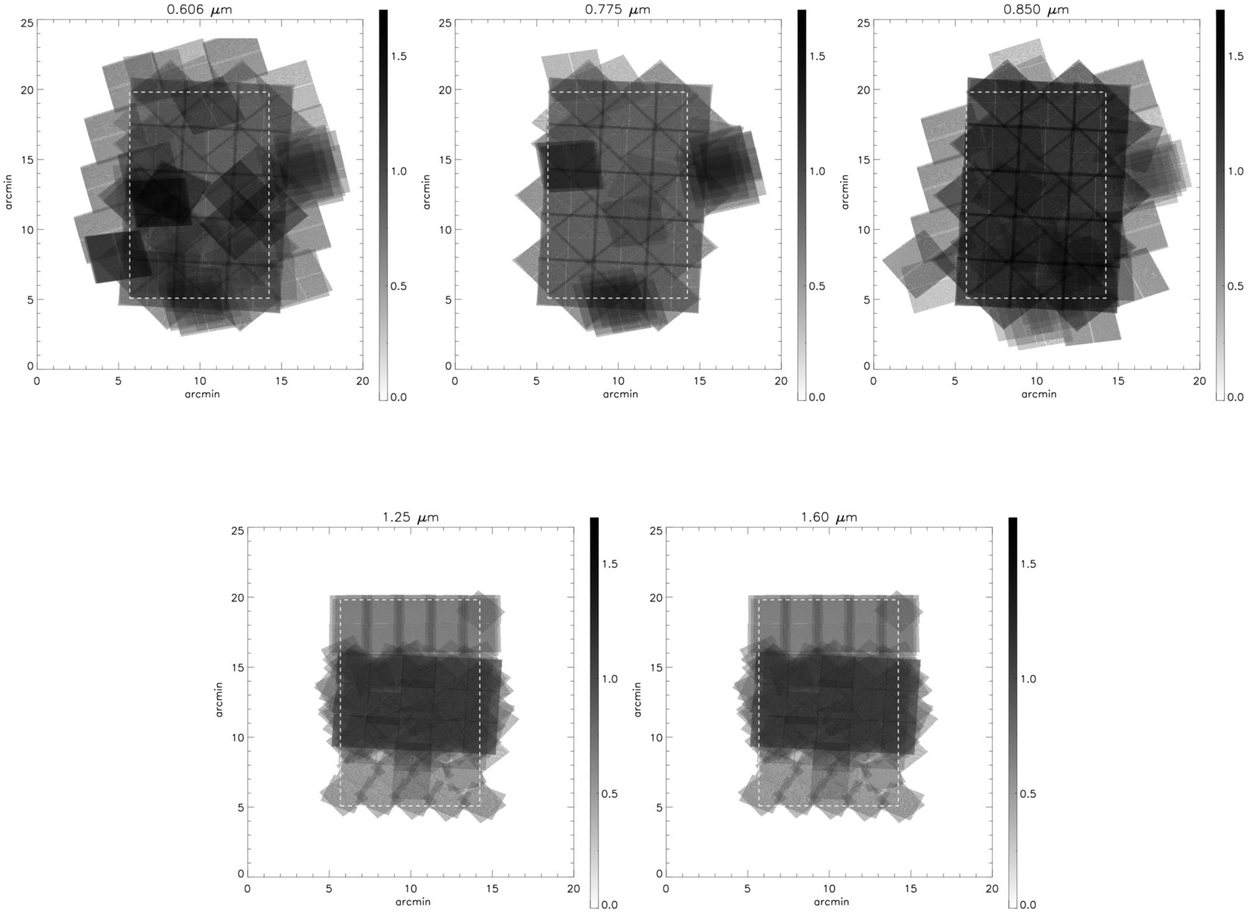

Figure 2.1 Tile layouts for each band plotted as $\log_{10}(N_{\mathrm{flt}})$, where N_{flt} is the number of FLT frames overlayed. This represents half of the total data. The dashed white line corresponds to the cropped area where the fluctuation analysis is performed.

astrometric alignment against the publicly available CANDELS mosaics[2] was achieved with TWEAKREG. All ACS data were charge transfer efficiency (CTE) corrected, and post-SM4 ACS frames were de-striped prior to CTE correction. Following the discussion above, only a small fraction of our total ACS dataset ($< 27\%$) contained post-SM4 frames. The dither patterns can be seen in figure 2.1.

---

[2] http://candels.ucolick.org/data_access/GOODS-S.html

2.3 Map making

In order to obtain angular power spectra over large angular scales, individual exposures must be combined into one or more mosaiced images. We generate our own mosaics using the self-calibration technique (SelfCal; Fixsen et al. 2000), instead of using the publicly available mosaiced images produced by astrodrizzle (Hack et al., 2012) (available at `http://candels.ucolick.org/data_access/Latest_Release.html`). Foreground emissions, predominately that of Zodiacal light, are particularly pernicious at infrared wavelengths (Abraham et al., 1997; Kelsall et al., 1998), so care must be taken when producing mosaics which combine observations taken at different times, especially at the WFC3/IR channels. Individual WFC3 exposures have clear overall offsets between frames taken at different times, which will lead to a fictitious anisotropy signal in the optical and IR background if care is not taken to properly model and remove those offsets.

The SelfCal algorithm is designed to produce mosaics which remove or minimize offsets between exposures. By dithering, or, observing the same area of sky with different parts of the detector, SelfCal builds a system of linear equations involving detector properties and sky brightness. If one can solve this system of linear equations, the sky brightness, detector gain, and offset parameters can be deduced using a likelihood. In essence, the SelfCal algorithm allows us to produce consistent mosaics from multiple observations by deducing all of the calibration parameters from the observations themselves.

Analytically, following a previous notation in Fixsen et al. (2000), we model our data, D^i, as a function of fixed detector gain G^p, actual sky intensity S^α, detector offsets per pixel F^p, and variable detector or sky offset per frame F^q as

$$D^i(G^p, S^\alpha, F^p, F^q) = G^p S^\alpha + F^q + F^p, \tag{2.1}$$

where the last term will account for a large-scale offset between frames caused by Zodiacal Light, atmospheric airglow or any other temporal stray light. We model our data in the same way as Arendt et al. (2002a), an example where HST data has already been self-calibrated; details, including the solution to Eq. (3.1), can be found there.

We de-weight bad pixels and cosmic rays, and iterate three times in order to find the SelfCal solution. Our input FLT frames are geometrically distorted with a pixel size of $0.''0498 \times 0.''0502$ for ACS, and $0.''1354 \times 0.''1210$ for WFC3. We remove the distortion in the map making procedure and produce mosaics with a slightly larger pixel size of $0.''140$ (geometrically square). For the ACS mosaics, we re-grid the input FLT frames to the WFC3 pixel scale as they are read into Selfcal. Generating SelfCal mosaics from the native ACS pixel scale for the amount of frames listed above requires upwards of 0.5 TB of memory and increases the computation time for each mosaic from ~ 8 hours to $\gtrsim 3$ days. We ultimately generate hundreds of mosaics (see section 2.5.3), so such a large computational time per moaic is unacceptable. We lose no information on this step with re-griding to a larger pixel size as we are not particularly interested in the fluctuations at the smallest scales and we have verified that re-pixelization does not introduces spurious effects by comparing power spectra from mosaics generated with re-gridded input pixel scales to the native input pixel scales in a few of the mosaics.

Note that the algorithm we use is same as the one used to generate self-calibrated maps of IRAC in the Spitzer fluctuation studies (Kashlinsky et al., 2005; Cooray et al., 2012b). The same method was implemented by the Herschel SPIRE Instrument Science team to generate wide area mosaics of the Herschel-SPIRE data, resulting in far-infrared fluctuations (Amblard et al., 2011; Viero et al., 2013). The algorithm originates from the time of FIRAS (Fixsen et al., 2000) and has wide applications. In the future we expect it will be used to combine frames and produce stable wide-area mosaics from JWST, Euclid, and WFIRST, among others.

We generate two maps per band so we can use the differences and sums to study systematics and noise biases, as was done in previous studies (Cooray et al., 2012b). The data are sorted by observation date and every other frame was used for each half map. HST data generally have two or more exposures per pointing, so this results in two maps per band of the same or similar dither pattern and exposure time per pointing. One map from each band can be found in figure 2.2. Multiple maps of the same band enable us to do cross-correlations, which ensures a removal of uncorrelated noise in the auto-spectra. This jack-knife process is similar to all other analysis related to large-scale structure and CMB angular power spectra from maps.

2.4 Power Spectrum Evaluation

Per the description in § 1.1, Fourier decomposition will be used to estimate the C_ℓ's. We discretely estimate the raw C_ℓ's (void of any fourier masking) as the average of the square FFT at all points within the annulus defined by a given ℓ bin i. If M_A is the FFT of map A, then our raw C_ℓ of bin i is given as

$$\widetilde{C_{\ell,i}} = \sum_{\ell_i}^{\ell_{i+1}} M_A(\ell_x, \ell_y)\, M_A^*(\ell_x, \ell_y), \tag{2.2}$$

where the annulus for bin i is simply defined as $\ell_i \leq \sqrt{\ell_x^2 + \ell_y^2} \leq \ell_{i+1}$. For a cross-correlation measurement one just replaces $M_A^*(\ell_x, \ell_y)$ with $M_B^*(\ell_x, \ell_y)$ in equation 2.2.

For each band we have two half maps, A and B. To measure the inherent noise in the data, we compute the noise power spectrum as the auto power spectrum of (A−B). To measure the raw auto-spectrum, we compute the cross spectrum of the two half maps, A×B, which eliminates any uncorrelated noise in our power spectrum estimate. In genearal, the standard

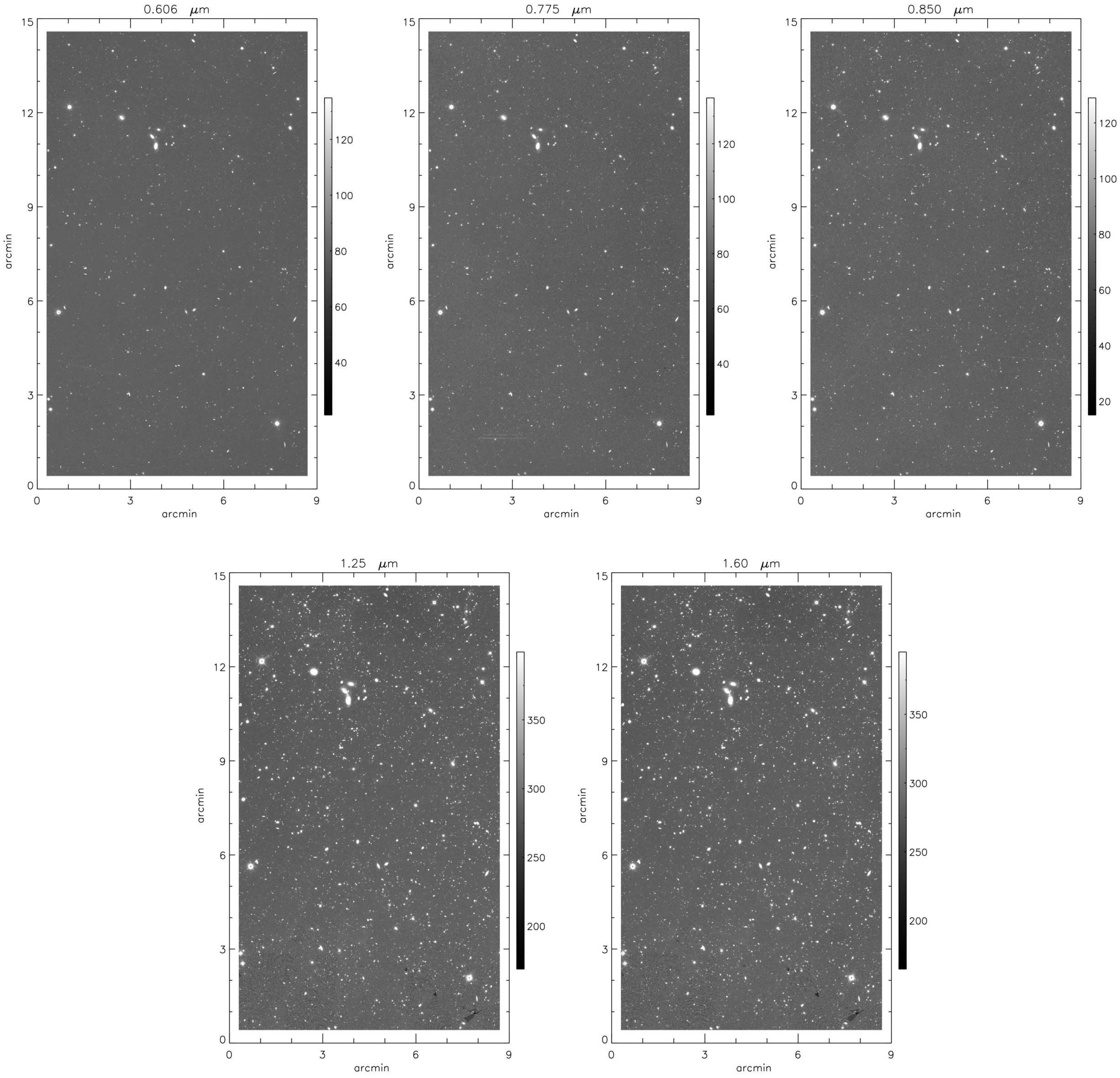

Figure 2.2 Mosaiced maps, representing one half of the total data in each band, in units of nW m^{-2} sr^{-1}.

deviation at each multipole is

$$\delta C_\ell = \sqrt{\frac{2}{f_{\rm sky}(2\ell+1)\Delta\ell}}\,(C_\ell^{\rm auto} + C_\ell^{\rm noise}), \tag{2.3}$$

where $\Delta\ell$ is the bin width for the given C_ℓ, and f_{sky} is the fractional sky coverage from all the pixels used in the FFT (excluding zeros). In our case, we have some error associated with the absolute calibration of the maps (see section 2.5.5), so we take the total error budget of our raw power spectra as the quadratic sum of the calibration errors with the variance in equation 2.3.

Figure 3.6 shows raw auto-spectra for all five bands, accompanied with each noise power spectrum and cosmic variance, $\delta C_\ell^{\rm cosmic\, variance}$, which is just equation 2.3 without the noise power spectrum term. These raw spectra show that our estimates are predominately far from being noise-dominated. A view of these cross-correlations in Fourier space can be seen in 2.5.

2.5 From raw to corrected C_ℓ's

While equation 2.2 allows an estimate on the raw C_ℓ'sa number of corrections must be taken into account before we can claim those fluctuations as solely astrophysical. We want to estimate the angular power spectrum of the unresolved intensity in our maps, so foreground sources must be completely removed. We remove sources with a source mask $S(x, y) = \delta_{ij}$ where $i \neq j$ if the pixel at location (x, y) contains a foreground source. This masking will introduce a mode-coupling between our C_ℓ bins, shifting the power between them, which is corrected for using a mode-coupling matrix $M_{\ell\ell'}$. We also need to quantify, in C_ℓ space, whether or not SelfCal, or the varying observational patterns, are biasing our power spectra in any way, which we correct for using a transfer function $T(\ell)$. Finally, we need to correct

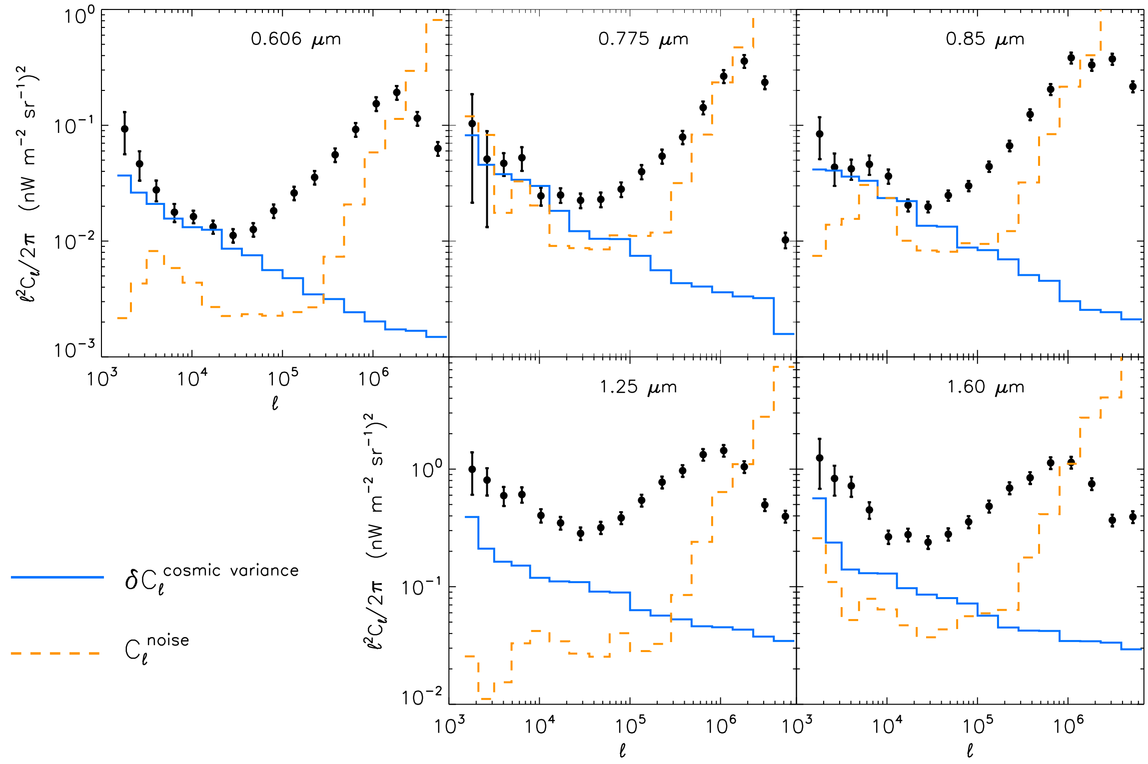

Figure 2.3 Raw auto-spectra for each filter. The dashed yellow lines represent the noise power spectrum for each band and the blue solid line represents the cosmic variance, both of which contribute to the error bars. Additional errors, which are constant at all multipoles, arise from the absolute flux calibration (see section 2.5.5 for details).

for the impact that HST's finite resolution has on our power spectra, which we do with a beam transfer function $B(\ell)$. Thus our corrected C_ℓ's are defined asHivon et al. (2002)

$$C_\ell = M^{-1}_{\ell\ell'} \widetilde{C_\ell}\, T(\ell)\, B(\ell)^{-2}. \tag{2.4}$$

2.5.1 Masking of resolved foreground sources

We utilized existing multi-wavelength catalogs of detected sources from the ultraviolet to mid-infrared (CITO/MOSAIC, VLT/VIMOS, VLT/ISAAC, VLT/HAWK-I, and *Spitzer*/IRAC) Guo et al. (2013). In addition, all the sources from the CANDELS, HUDF and ERS surveys (F435W, F606W, F775W, F814W, F850LP, F098M, F105W, F125W, and F160W) are also present in the catalog. The 50% completeness limit for F160W in the catalog is $m_{\mathrm{AB}} = 25.9, 26.6$ and 28.1 for the CANDELS wide, deep and HUDF regions respectively; the 5 σ limiting magnitudes are 27.4, 28.2 and 29.7. For each source detected by any of the aforementioned instruments, we apply an elliptical mask with parameters corresponding to the SExtractor Kron elliptical aperture. This catalog simplifies our masking procedure and ensures we are masking sources detected at other wavelengths, which may otherwise be undetected in our five bands.

In addition to the source mask generated from sources detected by other instruments, we also generate our own internal masks. We run SExtractor on a coadded map in each of our five filters and apply the same elliptical masking procedure to incorporate the shapes of sources. Next, we take the union of all five internal source masks, plus the source mask we made from the pre-existing catalog. After applying this union mask to each band, we clip 5 σ outliers and visually inspect each masked map. Any residual sources are masked by hand. We verified that all sources detected above 5σ in any of the bands, including deep IRAC data at 3.6 μm, are masked. This process yields 53% of the pixels unmasked for the

FFT computation. Note that tests can be performed which expand and shrink the source mask to further test the IHL model (Arendt et al., 2010).

2.5.2 Mode-mode coupling correction

The mode-mode coupling matrix $M_{\ell\ell'}$ is computed in the same way as has alredy been done a number of time previously (Cooray et al., 2012b; Zemcov et al., 2014). Our source mask introduces fictitious correlations in our power spectrum by shifting power between modes, as the mask, by definition, contains patches of localized zeros. Instead of generating $M_{\ell\ell'}$ analyticallyHivon et al. (2002), we generate our $M_{\ell\ell'}$ matrix computationally, one fourier mode at a time. The algorithm to get the i^{th} ($i \leq N_{\mathrm{bins}}$) row in the $M_{\ell\ell'}$ is essentially the following: we produce a large number (N) of Gaussian maps, retain all the modes in the i^{th} ℓ bin and set $C_{\ell'} = \delta_{\ell\ell'}$, and compute the average of the N power spectra of the masked maps to see how it affects this mode. In this way we generate $M_{\ell\ell'}$ mode-wise. We set $N = 100$, as was done in the previous papers. The matrix itself can be seen in fig.(2.6).

Since questions have been raised on the ability of the mode-coupling matrix to correct for the masking effects in the power spectrum, without any proofKashlinsky et al. (2014), we perform simulations to validiate our masking correction. We generate 90 Gaussian maps with a known input power spectrum, then mask the map and re-compute the power spectrum, and subsequently correct the masked power spectrum to recover the input power spectrum. This is a straight forward procedure which readily verifies our $M_{\ell\ell'}$ correction, as is shown in Fig. 2.7.

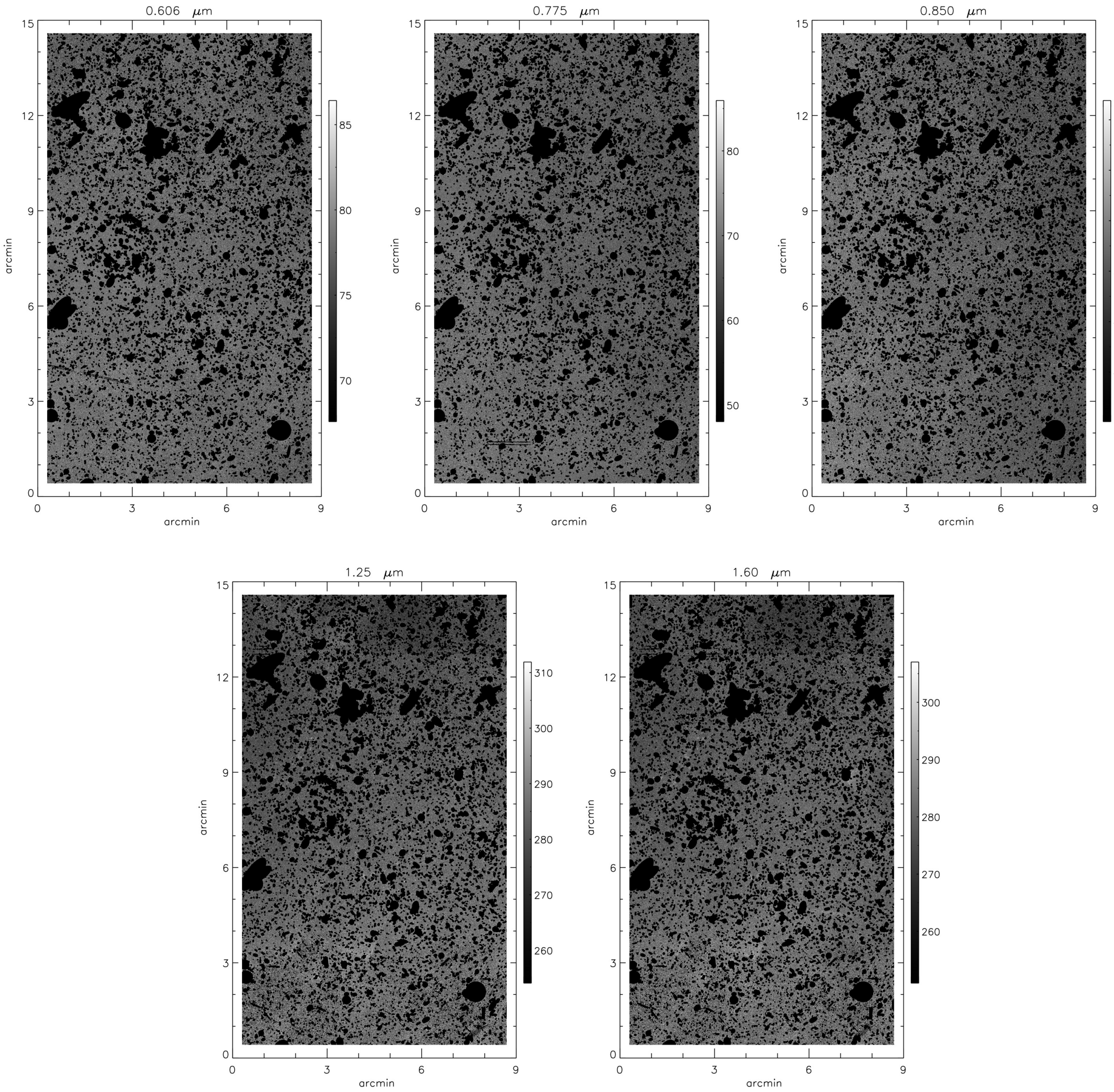

Figure 2.4 Final masked mosaics, again representing one half of the total data in each band, in units of nW m^{-2} sr^{-1}. Sections that appear to harbor excess flux are unresolved faint sources below the 5σ clipping of the background.

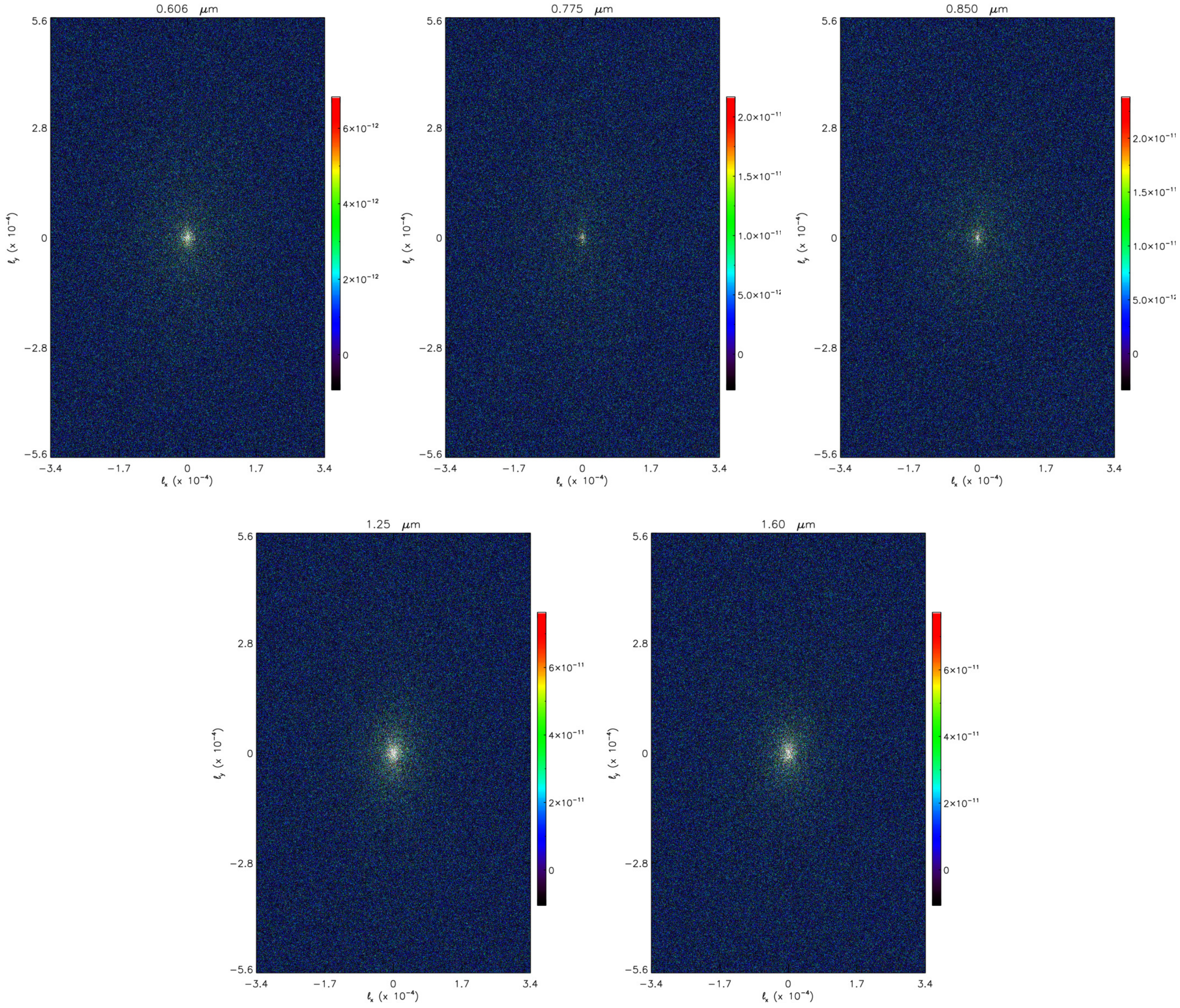

Figure 2.5 Masked mosaics in fourier space, in units of nW m^{-2} sr^{-2}.

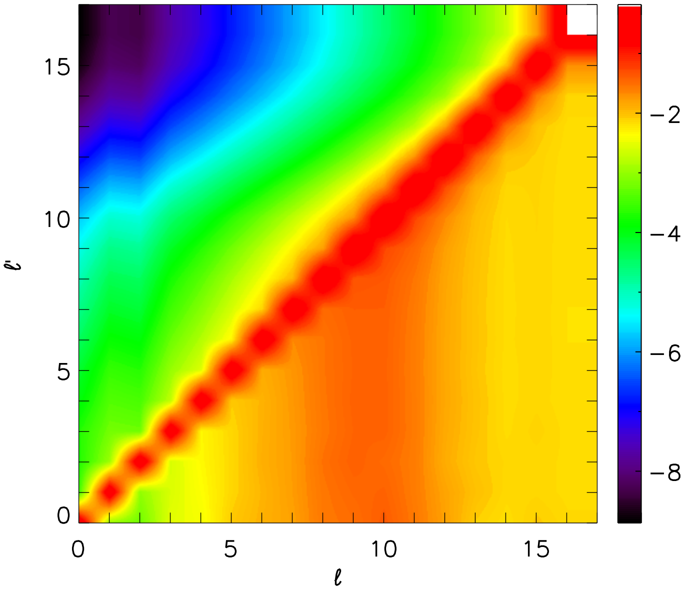

Figure 2.6 Logarithm of the mode-coupling matrix.

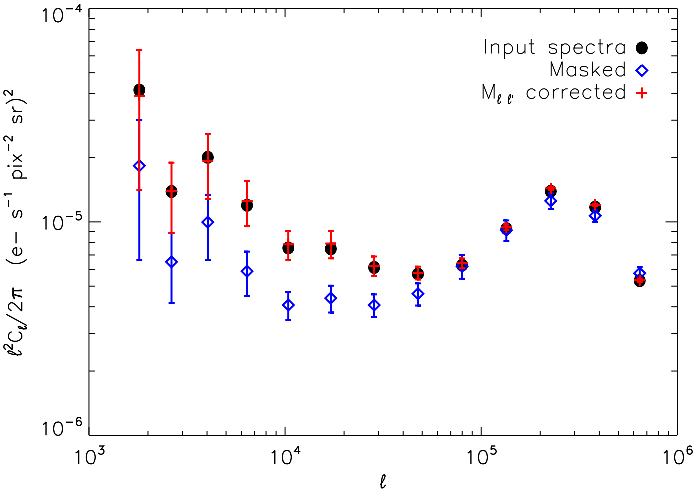

Figure 2.7 $M_{\ell\ell}$ validity simulations. The red crosses show that the input power spectrum is recovered after applying a mask to the image and correcting for the mode-mode coupling with the $M_{\ell\ell}$ matrix.

2.5.3 Transfer Function

Generating the transfer function is a computationally expensive and rather mundane process. What we want to do is quantify, in C_ℓ space, what our map making procedure is doing to our C_ℓ's. Again, we follow a similar process of what has been done before (Cooray et al., 2012b):

1. Produce a large map, at least as large as the mosaic, with the same pixel scale as the input FLT frames, and inject only a known power spectrum.
2. Replace the signal in each pixel of the input FLT frames with that from the larger map with a known power spectrum.
3. Add Gaussian noise with an rms proportional to the inverse square root exposure time, unique to each FLT frame.
4. Add an overall offset equal to the median of the given FLT frame.
5. Feed these simulated tiles to SelfCal, while maintaining the same parameters from the previous map making of real data; i.e. write distortion-free mosaics with the same output pixel scale ($0''.14$) while iterating SelfCal the same number of times. For ACS, we re-grid the simulated tiles to the WFC3 scale within this step, as was done with the real data.
6. Crop the output mosaics in the same locations as the real data and compute the power spectrum.
7. Compare the output power sectrum with the input power spectrum to obtain the transfer function, $T(\ell) = C_\ell^{\text{input}}/C_\ell^{\text{output}}$.

The 4^{th} step of adding an overall offset to each of the tiles is unique to our analysis, and should in principle be a good indicator of how well SelfCal is performing in offset removal,

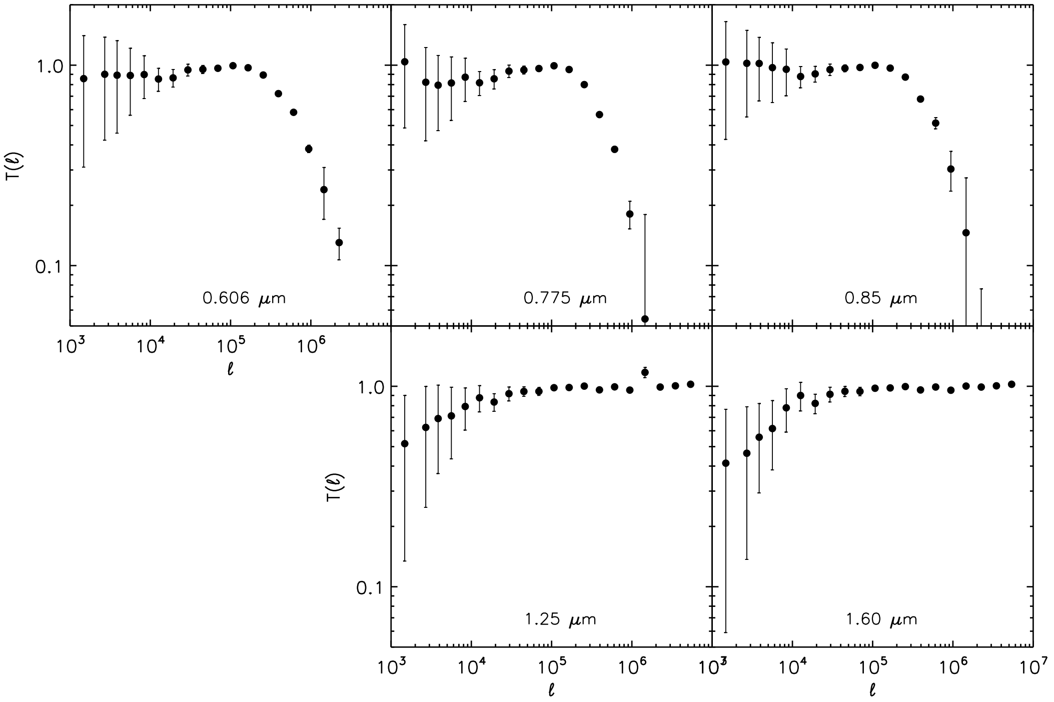

Figure 2.8 T(ℓ) from a minimum of 50 simulations in each band. These simulations incorporate the effects of the map-making algorithm, tiling pattern, varying exposure depths, residual (temporal) offsets, and cropping effects, specific to each filter.

unique to each observation. This procedure was performed 50 times or more in each band to produce $T(\ell)$, with the error bars being the standard deviation at each bin. We find an increased number of simulations, on the order of 200, is still insufficient to reduce the error bars by any significant amount. It appears that the errors at low multipoles are the result of a sampling variance and will not improve with more simulations. If there are residual offsets, we think they are accounted for in our transfer functions, which can be seen in figure 2.8

2.5.4 Beam Correction

Small scale fluctuations will be suppressed because of the finite resolution inherent in any optical device. This drop in power at high-ℓ is clearly evident in the raw power spectra in

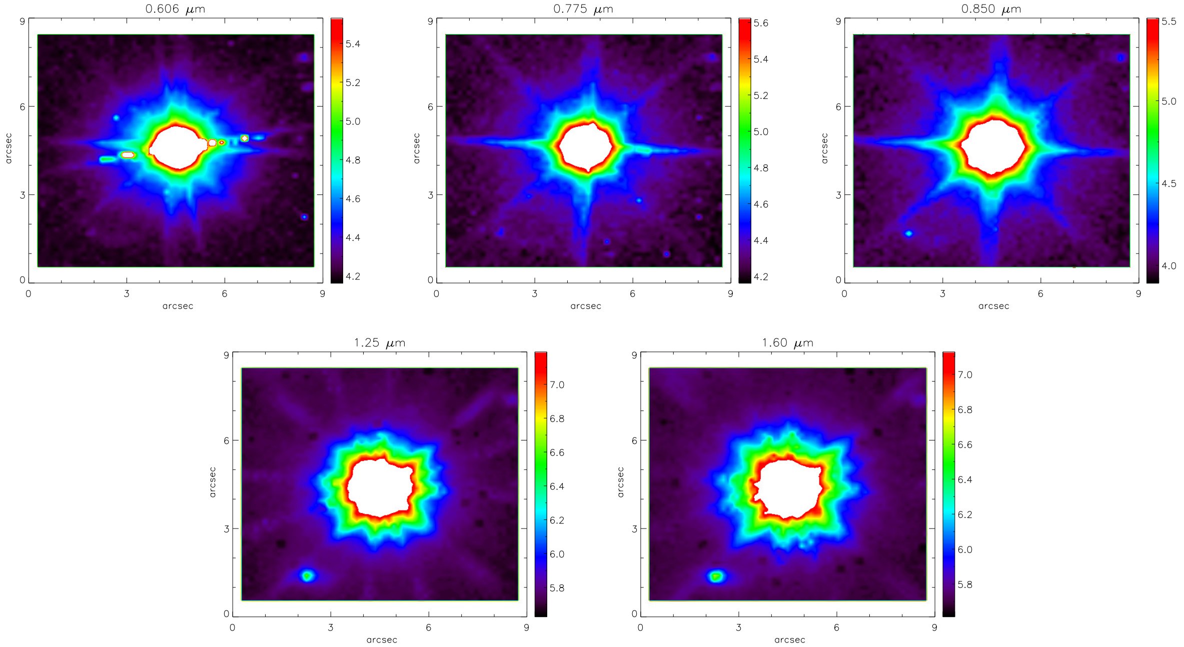

Figure 2.9 The PSF in each band, as measured from stacking sources in our SelfCal mosaics. The scaling is logarithmic in nW m^{-2} sr^{-1}.

figure 3.6. We correct this with a "beam transfer function", $B(\ell)$, which is just the PSF in harmonic space. We measure the PSFs of our mosaics by stacking sources across the image; the varying diffraction patterns in figure 2.9 are indicative of varying position angles throughout the ovservations. We make a $B(\ell)$ for each filter by taking the FFT of each of the PSFs in figure 2.9, extrapolating at low-ℓ, and normalizing to one. The $B(\ell)$ for each band can be seen in figure 2.10.

2.5.5 Absolute Flux Calibration

SelfCal achieves relative calibration between frames, so in general, the absolute flux calibration, or gain, of SelfCal output maps needs to be determined from a standard flux reference. The multi-wavelength catalog used in the masking procedure is in principle a good enough reference, however the photometry in the public CANDELS source catalogs was obtained

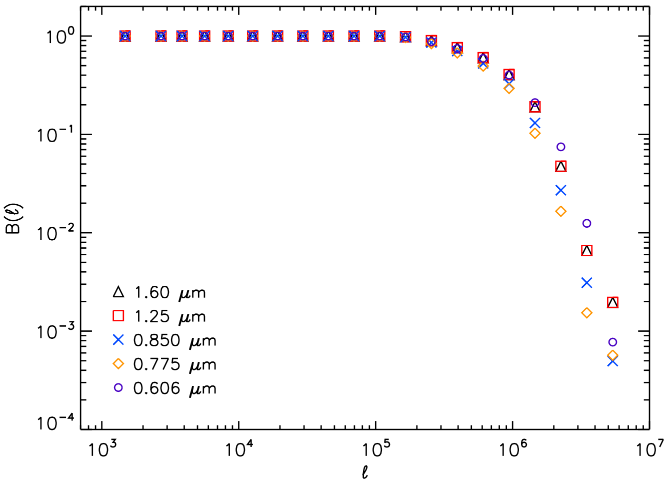

Figure 2.10 Filter specific $B(\ell)$'s.

with a private version of SExtractor, and, and has aperture corrections applied to the flux densities which were extracted from PSF-matched maps. We also generate internak catalogs from public CANDELS MultiDrizzledKoekemoer & et al. (2002) mosaics using the same procedure as we used on our mosaics.

In each band, we re-pixelize the MultiDrizzle maps to our SelfCal pixel scale, and perform source extraction with SExtractor on both our mosaics and the MultiDrizzle mosaics. We use the same parameter files for all the source extractions. We then astrometrically match the resultant catalogs in each band and keep those sources that are common within a radius of $0''.1$ (our FLT frames were aligned in tweakreg with MultiDrizzle mosaics as the astrometric reference, so our astrometry is similar to the MultiDrizzle maps at sub-arcsecond scales). Counts in $e-/s$ are converted to μJy using the current HST magnitude zero pointsKoekemoer et al. (2011).

The absolute flux calibration of each map is found by plotting a histogram of the flux ratios of the CANDELS source flux to our SelfCal source flux. The correction factor, unique to each half map, is definded as the abscissa corresponding to the peak of a fitted Gaussian,

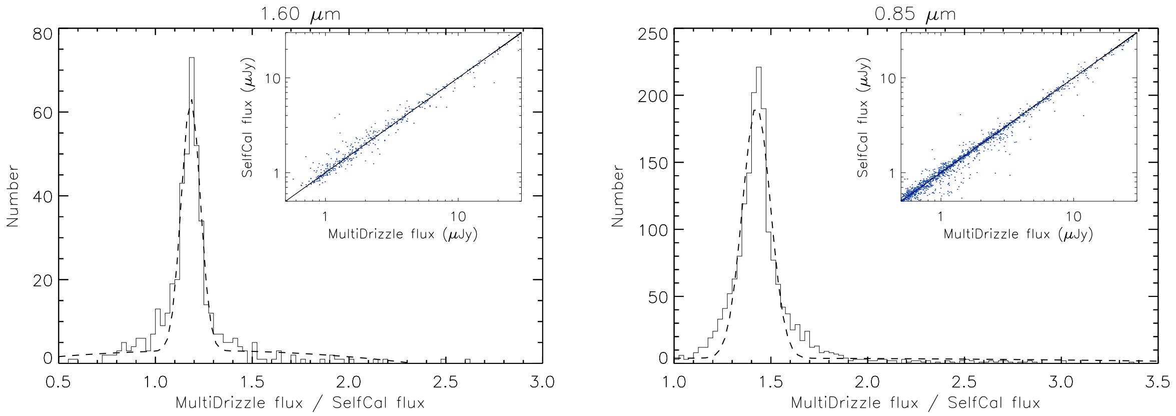

Figure 2.11 Calibration factors for one half map of WFC3 and ACS. The histograms show the uncorrected flux ratios of the CANDLES sources and our SelfCal sources. The calibration factors are calculated by fitting a Gaussian (dashed line) to the histograms. The insets show the corrected SelfCal source flux against the CANDELS source flux, with a sold line indicating the one-to-one correspondence.

Filter	Half-1	Half-2
F606W	1.41 ± 0.09	1.41 ± 0.09
F775W	1.42 ± 0.09	1.42 ± 0.09
F850LP	1.43 ± 0.07	1.42 ± 0.08
F125W	1.19 ± 0.05	1.19 ± 0.08
F160W	1.18 ± 0.05	1.19 ± 0.08

Table 2.2 Calibration factors for each half-map. The errors are incorporated into the power spectra.

with an associated error as the standard deviation of the Gaussian. Examples for WFC3 and ACS can be seen in Fig. 2.11. The re-calibration factors that match our self-calibrated mosaics to the same scale as MultiDrizzle mosaics are given in Table 2.2.

2.6 Corrected Auto- and Cross-Spectra

After applying all the corrections described above, we can now safely impute the power spectra signals to be of astrophysical origin. Our measurements continue to show the significant excess in the fluctuation amplitude at 30 arcsecond and larger angular scales, when compared to the clustering of faint, low-redshift galaxies (Helgason et al., 2012). The large-scale fluctuations correlate between filters (Figure 2.12). The excess in the amplitude of fluctuations relative to faint low-redshift galaxies is consistent with previous measurements at 3.6 μm (Kashlinsky et al., 2012; Cooray et al., 2012b). Our HST-based power spectra probe deeper into the fluctuations and have shapes departing from the fluctuations measured with the CIBER sounding rocket experiment (Zemcov et al., 2014). Due to the shallow depth of the CIBER imaging data, the measured fluctuations there are dominated by the shot noise of the residual galaxies at the arcminute angular scales that we probe here with Hubble data.

At angular scales of tens of arcminutes and above, CIBER detected an up-turn in the fluctuations with an amplitude well above the level expected from instrumental systematics and residual flat-field errorsZemcov et al. (2014). However, as shown in Figure 5, the combination of CIBER and Hubble fluctuations is consistent with a power-law clustering signal out to the largest angular scales probed by CIBER. If the power-law signal is of the form $C_l \propto l^{\alpha}$, the best-fit slope to combined CIBER and Hubble measurements at 1.6 μm is $\alpha = -3.05 \pm 0.07$. This slope is consistent with Galactic dust, which in emission at 100 μm has a power-law of -2.89 ± 0.22 (Ref. Amblard et al. (2011)). At the largest angular scales we could be detecting interstellar light scattered off of Galactic dust or diffuse Galactic light (DGL). The overall amplitude of fluctuations we measure at 1.6 μm is consistent with 10% DGL fluctuations at tens of arcminute angular scales, given the 100 μm surface brightness of GOODS-S of ~ 0.5 MJy/sr, and existing DGL intensity measurements (Figure 6). Future measurements in the optical wavelengths over a wider area are necessary to confirm the Galactic nature of fluctuations at angular scales greater than one degree. The corrected HST auto-spectra, along with those from other surveys, can be found in Fig. 2.13. In this section we also show the cross-correlations between all five bands, and the coherence, or correlation coefficient, as defined by

$$C_\ell^{\mathrm{AxB}} \sqrt{C_\ell^{\mathrm{A}}\, C_\ell^{\mathrm{B}}}. \tag{2.5}$$

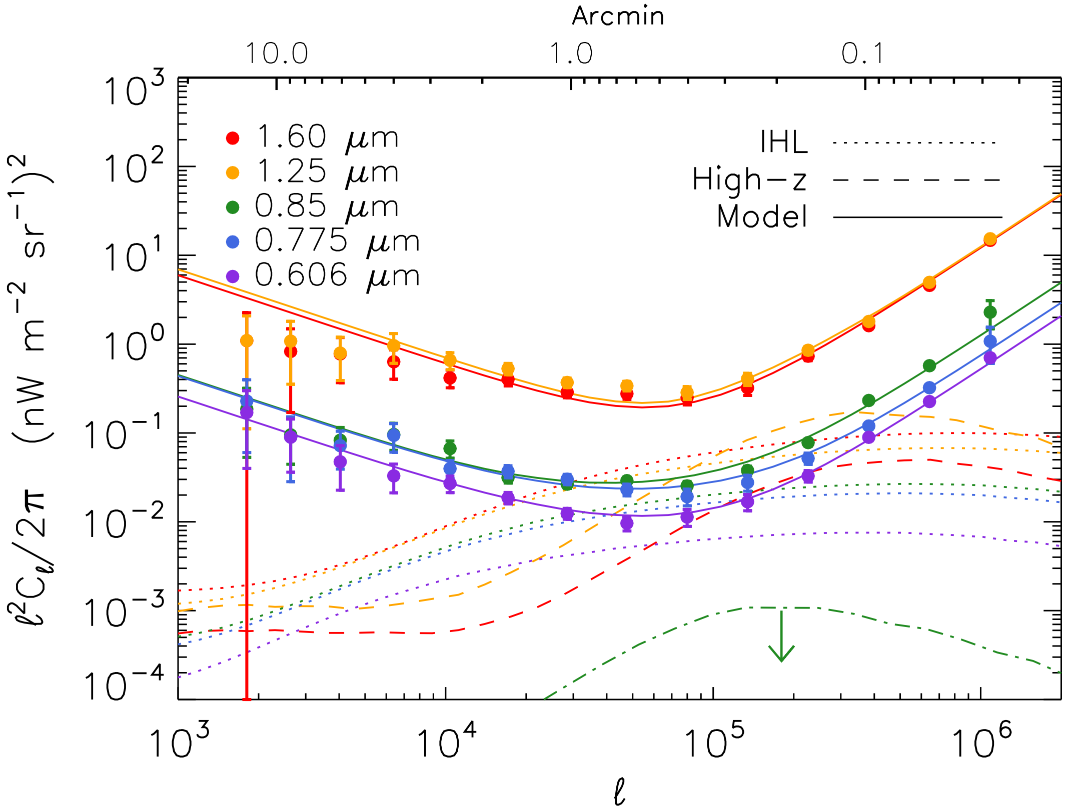

Figure 2.12 Angular power spectra of optical and near-infrared background intensity fluctuations. Multi-wavelength auto power spectra of optical to near-IR intensity fluctuations in the GOODS-South field using Hubble Space Telescope data (see Online Methods for data selection details). The error bars are calculated by adding in quadrature the errors from the beam transfer function, map-making transfer function and calibration errors, to the standard deviation at each multipole, δC_ℓ, described in Equation 11. Thus the 1 σ uncertainties account for all sources of noise and error, including map-making, calibration, detector noise, and cosmic variance associated with finite size of the survey. We show the best-fit model which makes use of four components: (a) $z > 8$ high-redshift galaxies, (b) intra-halo light (IHL; Cooray et al. 2012b, (c) faint low-redshift galaxies Helgason et al. (2012), and (d) diffuse Galactic light. At 1.25 and 1.6 μm, the best-fit high-redshift galaxy signal is shown as dashed lines. The signal is zero in the optical bands. We show the upper limit (denoted by a downard facing arrow) of fluctuations generated by $6 < z < 8$ galaxies as a dot-dashed line.

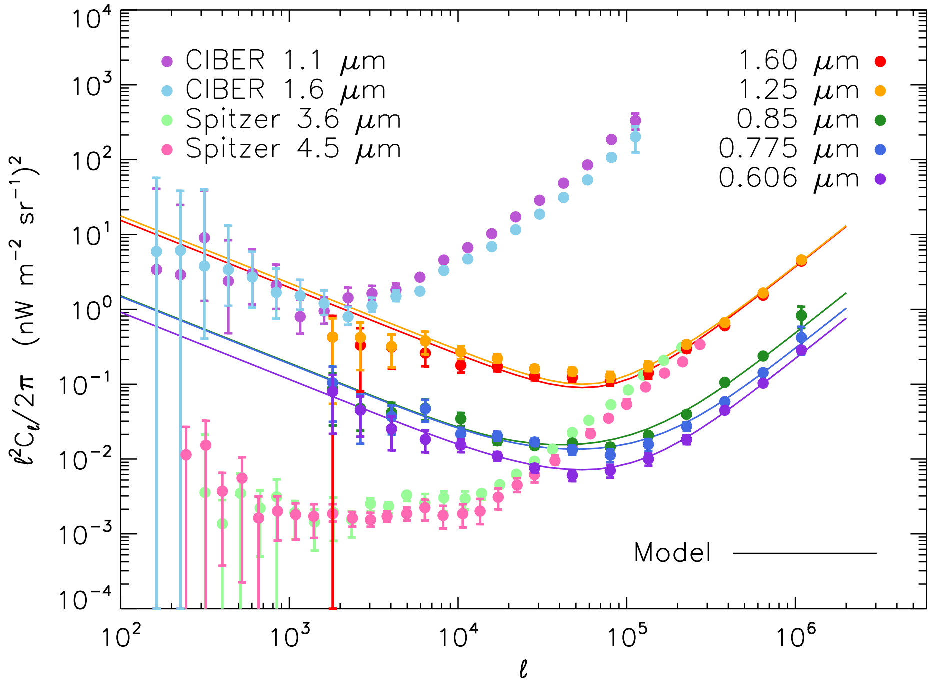

Figure 2.13 CANDELS corrected power spectra plotted against the CIBER (Zemcov et al., 2014), Spitzer (Cooray et al., 2012b) and NICMOS (Thompson et al., 2007) measurements (see also Donnerstein 2015 for more recent NICMOS measurements). The power spectrum resulting from the NICMOS analysis was measured from a MultiDrizzle map and has not been corrected for the transfer function and mode-coupling matrix resulting from source masking as discussed in our Methods section. Therefore we show it as a comparison but do not use it in our modeling. The error bars are 1 σ uncertainties that account for all sources of noise and error, including map-making, calibration, detector noise, and cosmic variance associated with finite size of the survey.CANDELS corrected power spectra plotted against the CIBER and Spitzer (Cooray et al., 2012b) measurements.

Questions have been raised on the validity of CIBER data reductionKashlinsky et al. (2014) due to coherence values that are larger than 1. We do not find this to be the case with the Hubble analysis. We show three examples in Fig. 2.14 for clarity.

Excess signal is detected in the cross-correlations between all five bands. Ten cross-correlations are shown in figure 2.15. The origin of uncharacteristic signatures in some of these measurements are unclear, but we find that the majority of the features are from inter-instrument cross-correlations coming from residual systematics that are unaccounted for. The error bars

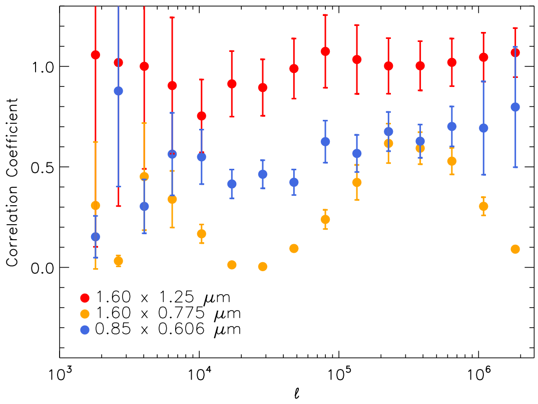

Figure 2.14 The coherence, or the correlation coefficient, in various bands is statistically between 0 and 1, despite potential systematics showing up in the cross-correlations.

ℓ	$\ell^2\ C_\ell^{1.6}\ /\ 2\pi$	$\ell^2\ C_\ell^{1.25}\ /\ 2\pi$	$\ell^2\ C_\ell^{0.850}\ /\ 2\pi$	$\ell^2\ C_\ell^{0.775}\ /\ 2\pi$	$\ell^2\ C_\ell^{0.606}\ /\ 2\pi$
1.81×10^3	1.10 ± 1.17	1.10 ± 0.98	$(1.86 \pm 1.33) \times 10^{-1}$	$(2.30 \pm 1.69) \times 10^{-1}$	$(1.70 \pm 1.30) \times 10^{-1}$
2.64×10^3	$(8.30 \pm 6.60) \times 10^{-1}$	1.08 ± 0.72	$(9.55 \pm 5.10) \times 10^{-2}$	$(9.02 \pm 6.19) \times 10^{-2}$	$(9.01 \pm 5.38) \times 10^{-2}$
4.04×10^3	$(7.78 \pm 4.09) \times 10^{-1}$	$(7.94 \pm 4.04) \times 10^{-1}$	$(8.30 \pm 3.25) \times 10^{-2}$	$(7.21 \pm 3.29) \times 10^{-2}$	$(4.74 \pm 2.47) \times 10^{-2}$
6.40×10^3	$(6.33 \pm 2.30) \times 10^{-1}$	$(9.61 \pm 3.53) \times 10^{-1}$	$(9.61 \pm 3.23) \times 10^{-2}$	$(9.40 \pm 3.34) \times 10^{-2}$	$(3.29 \pm 1.17) \times 10^{-2}$
1.04×10^4	$(4.18 \pm 0.94) \times 10^{-1}$	$(6.60 \pm 1.43) \times 10^{-1}$	$(6.70 \pm 1.47) \times 10^{-2}$	$(3.98 \pm 0.91) \times 10^{-2}$	$(2.71 \pm 0.57) \times 10^{-2}$
1.71×10^4	$(4.01 \pm 0.63) \times 10^{-1}$	$(5.29 \pm 0.77) \times 10^{-1}$	$(3.13 \pm 0.41) \times 10^{-2}$	$(3.70 \pm 0.57) \times 10^{-2}$	$(1.86 \pm 0.27) \times 10^{-2}$
2.85×10^4	$(2.89 \pm 0.39) \times 10^{-1}$	$(3.71 \pm 0.46) \times 10^{-1}$	$(2.66 \pm 0.28) \times 10^{-2}$	$(3.02 \pm 0.39) \times 10^{-2}$	$(1.24 \pm 0.17) \times 10^{-2}$
4.76×10^4	$(2.76 \pm 0.37) \times 10^{-1}$	$(3.39 \pm 0.41) \times 10^{-1}$	$(2.90 \pm 0.27) \times 10^{-2}$	$(2.30 \pm 0.34) \times 10^{-2}$	$(9.62 \pm 1.74) \times 10^{-3}$
7.99×10^4	$(2.50 \pm 0.43) \times 10^{-1}$	$(2.85 \pm 0.46) \times 10^{-1}$	$(2.54 \pm 0.30) \times 10^{-2}$	$(1.92 \pm 0.40) \times 10^{-2}$	$(1.13 \pm 0.24) \times 10^{-2}$
1.34×10^5	$(3.19 \pm 0.56) \times 10^{-1}$	$(4.03 \pm 0.63) \times 10^{-1}$	$(3.76 \pm 0.45) \times 10^{-2}$	$(2.75 \pm 0.55) \times 10^{-2}$	$(1.67 \pm 0.34) \times 10^{-2}$
2.26×10^5	$(7.29 \pm 0.80) \times 10^{-1}$	$(8.50 \pm 0.89) \times 10^{-1}$	$(7.76 \pm 0.69) \times 10^{-2}$	$(5.16 \pm 0.74) \times 10^{-2}$	$(3.27 \pm 0.47) \times 10^{-2}$
3.82×10^5	1.62 ± 0.09	1.80 ± 0.11	$(2.33 \pm 0.13) \times 10^{-1}$	$(1.20 \pm 0.10) \times 10^{-1}$	$(8.96 \pm 0.75) \times 10^{-2}$
6.44×10^5	4.59 ± 0.13	4.96 ± 0.16	$(5.73 \pm 0.48) \times 10^{-1}$	$(3.24 \pm 0.22) \times 10^{-1}$	$(2.26 \pm 0.13) \times 10^{-1}$
1.09×10^6	$(1.49 \pm 0.02) \times 10^{+1}$	$(1.55 \pm 0.05) \times 10^{+1}$	2.31 ± 0.80	1.09 ± 0.47	$(7.08 \pm 0.70) \times 10^{-1}$

Table 2.3 Final HST power spectra; corrected auto-spectra, $\ell^2 C_\ell / 2\pi$ in units of $(\mathrm{nW\ m^{-2}\ sr^{-1}})^2$, for 5 bands. The quoted errors are the 1 σ uncertainties.

for the cross-correlations are calculated as

$$\delta C_\ell = \sqrt{\frac{1}{f_{\rm sky}(2\ell+1)\Delta\ell}\left[(C_{A\ell}^{\rm auto} + C_{A\ell}^{\rm noise})(C_{B\ell}^{\rm auto} + C_{B\ell}^{\rm noise}) + (C_\ell^{\rm A\times B})^2\right]} \tag{2.6}$$

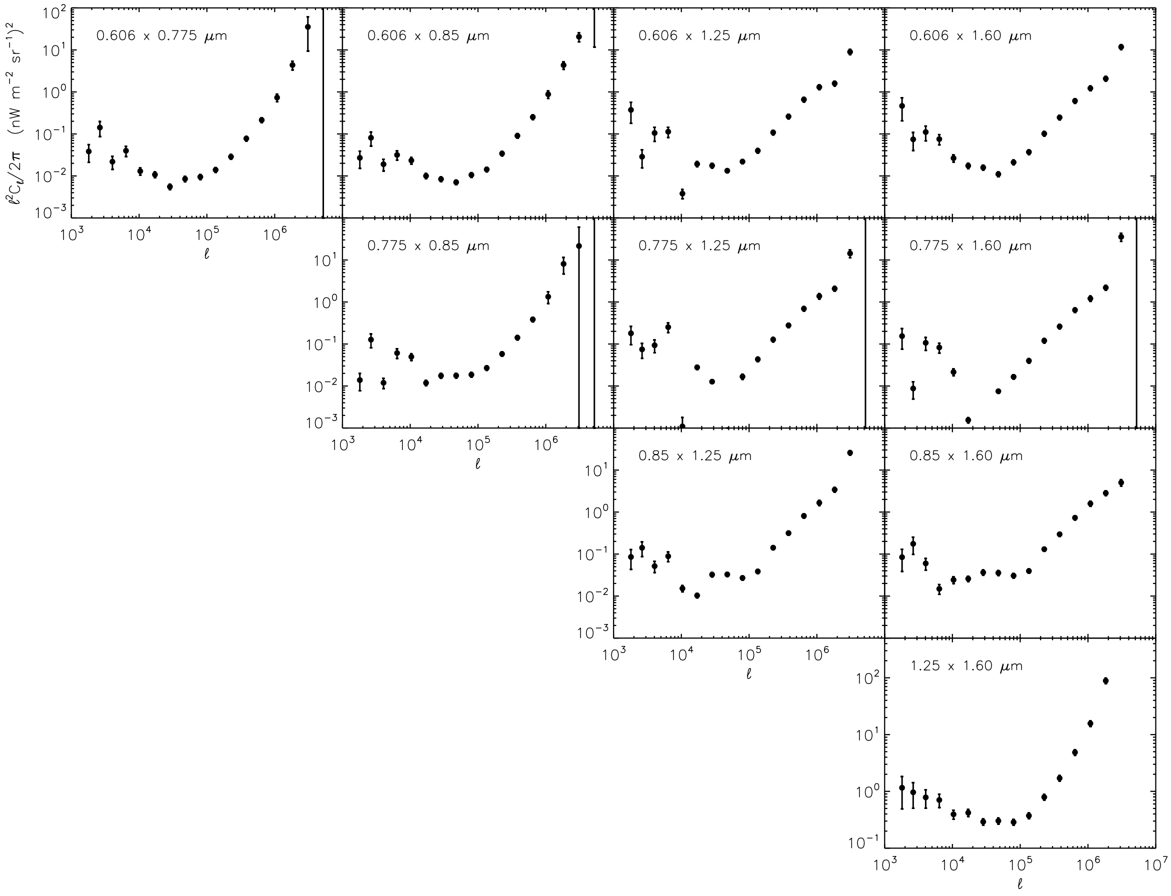

Figure 2.15 Angular cross- power spectra of optical and near-infrared background. Ten cross-correlations between the HST bands. Excess signal is detected in the cross-correlations. The error bars are 1 σ uncertainties which are calculated in a similar way as in Figure 3, which accounts for all sources of noise and error, including map-making, calibration, detector noise, and cosmic variance. However, the noise power spectra for the cross-correlations are calculated slightly differently. For each filter we have two maps, so for each cross-correlation between bands we have four maps (label them A and B for the first filter, and C and D for the second). This enables us to generate a noise power spectrum by computing (A-B) x (C-D), as opposed to taking the auto-spectrum of (A-B) for the auto-correlations. The first row corresponds to all correlations with the 0.606 μm band, the second for all correlations with the 0.775 μm band not found in the first row, the third row corresponds to all correlations with the 0.850 μm band not found in any of the preceding rows, and the last row corresponds to correlations at 1.25 μm. The columns similarly increase in wavelength as you move across the page.

2.7 Theoretical Interpretation – Multi-Component Model

Through models for multiple sources of intensity fluctuations, from diffuse Galactic Light to primordial faint galaxies, we are able to describe the five-band fluctuation measurements in optical and near-IR wavelengths to obtain a constraint on the UV luminosity density of galaxies present during reionization at a redshift above 8. We compare our measurement to existing constraints on the quantity and we also discuss implication of our measurement. First we describe the various model components.

For our theoretical interpretation, we invoke a model which involves four main components: (a) intra-halo light (IHL) following (Cooray et al., 2012b), (b) diffuse galactic light (DGL) due to interstellar dust-scattered light in our Galaxy, (c) low-redshift residual faint galaxiesHelgason et al. (2012); and (d) the high-redshift signal. We assume the flat ΛCDM model with $\Omega_M = 0.27$, $\Omega_b = 0.046$, $\sigma_8 = 0.81$, $n_s = 0.96$ and $h = 0.71$ in our theoretical modeling (Komatsu et al., 2011). We summarize the basic ingredients of our model now while providing references for further details.

For IHL, we follow the model developed in Ref. Cooray et al. (2012b). The mean luminosity of the IHL at rest-frame wavelength λ for a halo with mass M at z is described as

$$\bar{L}_\lambda^{\rm IHL}(M,z) = f_{\rm IHL}(M) L_{\lambda_p}(M)(1+z)^\alpha F_\lambda(\lambda_0/(1+z)), \tag{2.7}$$

where $\lambda_0 = \lambda(1+z)$ is the observed wavelength, α is the power-law index which takes account of the redshift-evolution effect, and $f_{\rm IHL}(M)$ is the IHL luminosity fraction of the total halo, which takes the form

$$f_{\rm IHL}(M) = A_{\rm IHL}\left(\frac{M}{M_0}\right)^\beta . \tag{2.8}$$

Here $A_{\rm IHL}$ is the amplitude factor, $M_0 = 10^{12}\ M_\odot$, and β is the mass power index. In equation (2.7), $L_{\lambda_p}(M) = L_0(M)/\lambda_p$ is the total halo luminosity at 2.2 μm and at $z = 0$, where $\lambda_p = 2.2$ μm and L_0 is given byLin et al. (2004)

$$L_0(M) = 5.64 \times 10^{12} h_{70}^{-2} \left(\frac{M}{2.7 \times 10^{14} h_{70}^{-1} M_\odot} \right)^{0.72} L_\odot. \tag{2.9}$$

Here $H_0 = 70\, h_{70}$ $\mathrm{km\,s^{-1}Mpc^{-1}}$ is the present Hubble constant. The F_λ term is the IHL spectral energy distribution (SED), which can transfer L_{λ_p} to the other wavelengths and is normalized to be 1 at 2.2 μm (see the discussion in Ref. Cooray et al. (2012b) for details). We assume the IHL SED to be the same as the SED of old elliptical galaxies, which are comprised of old red starsKrick & Bernstein (2007).

The angular power spectrum of IHL fluctuations, C_ℓ^{IHL}, is calculated via a halo model approach (Cooray & Sheth, 2002) and involves both a 1-halo term associated with spatial distribution inside a halo and a 2-halo term involving clustering between halos. The details are provided in Cooray et al. (2012b). In the clustering calculation, we assume the IHL density profile follows the Navarro-Frenk-White (NFW) profile (Navarro et al., 1997; Cooray et al., 2012b). We set the maximum IHL redshift at $z_{\max} = 6$. The $M_{\min}$ and $M_{\max}$ are fixed to be 10^9 and 10^{13} $M_\odot/h$, and the power-law index β is fixed to be 0.1 in this workCooray et al. (2012b). The IHL model is then described with two parameters: A_{IHL} in equation (2.8) and the power-law index α in equation (2.7).

A simple test of IHL is to grow the source mask and study how the fluctuation power spectrum varies as a function of the mask size. However, we note that our model for IHL involves clustering at large angular scales. That is, in our description IHL is not restricted to regions near galactic disks only. The dependence of the power spectrum with mask size is studied in Arendt et al. (2010) and Donnerstein (2015). These studies find that the fluctuations do not vary strongly with the mask radius, though such studies ignored the mode-coupling effects associated with the mask as the mask radius is varied. As discussed in Cooray et al. (2012b), in order to test IHL one has to grow the masking area to a factor of 10 larger than the typical mask radius used in the current analyses of fluctuations. In the case of Spitzer data, where we expect fluctuations to be dominated by IHL, the relatively large 2 arcsecond point spread function makes it close to impossible to test IHL directly with a varying mask. However, additional tests of IHL exist in the literature (Arendt et al., 2010). These include correlations between artificial halos and masked sources, and correlations between masked sources and foreground galaxies. Such tests have not ruled out the IHL component. Furthermore, without such a component we are not able to explain the fluctuations measured at wavelengths below 0.8 μm, as residual faint galaxy clustering (Helgason et al., 2012) is not adequate to explain the measurements.

The DGL component involves dust-scattered light and it is likely that the same dust is observed by IRAS at far-infrared wavelengths through thermal emission. The DGL component was considered in Zemcov et al. (2014) and an upper limit on the expected amplitude was included based on the cross-correlation with 100 μm IRAS map of the same fields. The CIBER final model results focussing on IHL to explain the fluctuations did not allow the DGL fluctuations amplitude to vary as a free parameter. With Hubble data, we find stronger evidence for DGL or a DGL-like signal once combined with CIBER, and with an rms amplitude for fluctuations that is at least a factor of 3 larger than the upper limit used in Ref.Zemcov et al. (2014) based on the cross-correlation with CIBER. Moreover, we find that the angular power spectrum is proportional to $\sim \ell^{-3}$ over the degree scales measured from CIBER to tens of arcminute scales of CANDELS measurements. So we model it with an amplitude factor $A_{\rm DGL}$ as

$$C_\ell^{\rm DGL} = A_{\rm DGL}\ell^{-3}. \tag{2.10}$$

To validate this ℓ^{-3} DGL slope dependence, we perform a linear fit in log space at low multipoles of the HST 1.6 μm data simultaneously with the CIBER 1.6 μm data. We measure a slope of -3.05 ± 0.07, so the functional form of our DGL model with ℓ^{-3} is appropriate (Figure 5). The power-law behavior of the DGL signal is consistent with Galactic dust emission power spectra in far-infrared and sub-mm surveys (Amblard et al., 2011), and dust polarization measurements in all-sky experiments like Planck (Planck Collaboration et al., 2014). We summarize a comparison of our DGL intensity measurements to those of existing measurements as a function of wavelength in Figure 5.

The clustering of low-redshift faint galaxies at $z < 5$, where reliable luminosity functions exist in the literature, is based on the detailed models in Helgason et al. (2012). We follow the calculations presented there to establish the expected level of low-redshift clustering, and the uncertainty of that expectation at the depth to which we have masked foreground galaxies. Because the low-redshift luminosity functions are not steep, unless there is a break or steepening in the luminosity function $-$ which is not supported by the halo model $-$ these low-redshift populations do not dominate the clustering we have measured. Given that our fluctuation measurements reach the deepest depths provided by both Hubble and Spitzer/IRAC, it is also unlikely that populations such as extreme red galaxies at $z < 2$ are responsible for the measured fluctuations. If there are populations at low redshift responsible for the SED of the fluctuations we measure, they would need to have individual SEDs that are consistent with a sharp break redshifted between 0.8 and 1.25 μm.While fluctuation measurements in just two bands cannot separate galaxies that have redshifted 4000 Å break, or galaxies that have redshifted Lyman-α break between those two bands, with five bands we have adequate knowledge on the SED of fluctuations, and the shape of the clustering over two decades in angular scales to separate high-z galaxies from low-z faint interlopers. The low-z interlopers are also likely captured by our IHL model as we cannot distinguish between diffuse stars and faint, dwarf galaxies that happen to be a satellite of a large dark matter halo in our modeling description.

2.7.1 Galaxies During Reionization

The final and critical component in our model is the signal from $z > 6$. We break this signal into two redshift intervals given the placement of the five ACS and WFC3 bands, based on the Lyman-dropout signal that moves across these bands. In particular, we consider $8 < z < 13$ and $6 < z < 8$ as the two windows. As we discuss later, given the availability of SFR density measurements in the $6 < z < 8$ interval, we mostly allow the signal in that redshift interval to be constrained by the existing data, and model-fit independently the SFR density in the higher-redshift interval. We do not have a strong independent constraint on the $6 < z < 8$ signal since it is only a Lyman-dropout in our shortest-wavelength band at 0.6 μm. This allows a better separation of the $8 < z < 13$ from the rest of the signals discussed above. To measure $6 < z < 8$ independently, we would need at least one more band below 0.6 μm. The signal from $8 < z < 15$ disappears from the three optical bands (0.6 to 0.85 μm) and is present in the two IR bands at 1.25 and 1.6 μm.

To model the high-redshift signals, we adopt an analytic model (Cooray et al., 2012a; Yue et al., 2013a) based on the work of Fernandez et al. (2012). It involves a combination of two separate classifications of stars − moderate-metallicity, second-generation or later stars (PopII), and the first generation of stars ever formed in the Universe, hence zero metallicity (PopIII). These are modeled with Salpeter (Salpeter, 1955) and Larson (Larson, 1999) initial mass functions (IMFs) for PopII and PopIII stars, respectively. The calculation related to direct stellar emission and the associated nebular lines, including especially Lyman-α emission, follows the work of Fernandez & Komatsu Fernandez et al. (2012). The total integrated intensity from $z_{\rm min} < z < z_{\rm max}$ is

$$\nu I_\nu = \int_{z_{\rm min}}^{z_{\rm max}} dz \frac{c}{H(z)} \frac{\nu(z)\bar{j}_\nu(z)}{(1+z)^2} \,, \tag{2.11}$$

where $\nu(z) = (1+z)\nu$. The comoving specific emissivity, as a function of the frequency is composed of both PopII and PopIII stars with an assumed z-dependent fraction as discussed in Cooray et al. (2012a) with the form given by

$$f_p(z) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{z-10}{\sigma_p}\right)\right] , \tag{2.12}$$

with $\sigma_p = 0.5$. The model thus assumes most of the halos have PopIII stars at $z > 10$ while PopII stars dominate at redshifts lower than that.

There are a number of theoretical parameters related to this model, especially the escape fraction of the Lyman-α photons f_{esc}, the star-formation efficiency denoting the fraction of the baryons converted to stars in high-redshift dark matter halos, or f_*, and the minimum halo mass to host galaxies, or $\mathrm{M}_{\mathrm{min}}$. The overall quality of the data is such that we are not able to independently constrain all of the parameters related to the high-redshift intensity fluctuation signal. Moreover these parameters are degenerate with each other (i.e., changing $f_\star$ can be compensated by a change in $\mathrm{M}_{\mathrm{min}}$ for example). Thus given that we do not have the ability to constrain multiple parameters, we simply model-fit a single parameter $A_{\mathrm{high-z}}$ that scales the overall amplitude from the default model, interpret that scaling through a variation in f_*, and subsequently convert that to a constraint on the SFRD. We fix our default model to a basic set of parameters, and assume $f_{\mathrm{esc}} = 0.2$, $\mathrm{M}_{\mathrm{min}} = 5 \times 10^7\ \mathrm{M}_\odot$, and $f_* = 0.03$. The resulting optical depth to reionization of this default model is 0.07, consistent with the optical depth measured by Planck (Planck Collaboration et al., 2015). Among all these parameters, the most significant change (over the angular scales on which we measure the fluctuations) comes effectively from f_*, or the overall normalization of $\bar{j}_\nu(z)$, given that it is directly proportional to f_*. This can in turn be translated to a direct constraint on the SFRD, $\psi(z)$, since with $f_\star$ we are measuring the integral of the halo mass function such that

$$\psi(z) = f_* \frac{\Omega_b}{\Omega_m} \frac{d}{dt} \int_{M_{\mathrm{min}}}^{\infty} dM M \frac{dn}{dM}(M, z) \tag{2.13}$$

where dn/dM is the halo mass function (Tinker et al., 2008).

Finally, to calculate the angular power spectrum of fluctuations, we also need to assign galaxies and satellites to dark matter halos. For that we make use of the halo model (Cooray & Sheth, 2002). We make use of the same occupation number distribution as in Cooray et al. (2012a) where the central and satellite galaxies are defined following Zheng et al. (2005). However, departing from the low-redshift galaxy models, we take a steep slope for the satellite counts in galaxies with $\alpha_s = 1.5$. The low-redshift galaxy clustering and luminosity functions are consistent with $\alpha_s \sim 1$ (Ref. Zheng et al. (2005)), but such a value does not reproduce the steep faint-end slopes of the LBG luminosity functions (Oesch et al., 2014; Zheng et al., 2012; Coe et al., 2013; Bouwens et al., 2014; Finkelstein et al., 2014). Such a high slope for the satellites also boost the non-linear clustering or the 1-halo term of the fluctuations. We do not have the ability to independently constrain the slope of satellites from our fluctuation measurements. In the future a joint analysis of fluctuations and LBG luminosity functions may provide additional information on the parameters of the galaxy distribution that is responsible for fluctuations. It may also be that the models can be improved with additional external information, such as the optical depth to reionization. We also note that other sources at high redshift include direct collapse black holes (DCBHs; Yue et al. 2013a), but we do not explicitly account for them here as the existing DCBH model is finely tuned to match Spitzer fluctuations, and the low signal-to-noise ratio of the Chandra-Spitzer cross-correlation results in them residing primarily at $z > 12$. DCBHs at such high redshifts will not contribute to Hubble fluctuations.

Finally, at smaller angular scales, the shot noise dominates the optical and IR background intensity fluctuation. Since it is scale-independent, we set it as a free variable parameterized as

$$C_\ell^{\rm shot} = A_{\rm shot}\,. \tag{2.14}$$

This noise term in the fluctuation power spectrum arises because of the Poisson behavior of the galaxies at small angular scales, a product of the finite number of galaxies. Our measured shot noise comes from a combination of the unmasked, faint low-redshift dwarf galaxies, and the high-redshift population. We do not use the information related to the shot noise in our models but instead treat it as a free independent parameter, since we cannot separate the high-redshift shot noise from the shot noise produced by faint, low-redshift dwarf galaxies. Here we focus mainly on the clustering at tens of arcseconds and larger angular scales to constrain SFRD during reionization. In the future, with either a precise model for the low-redshift galaxies or a model for high-redshift galaxies that determines their expected number counts as a function of the free parameters such as $\mathrm{M}_{\mathrm{min}}$ and f_*, it may be possible to separate the overall shot noise associated with reionization sources from that of the low-redshift faint galaxies. If this is the case then it might also be possible to improve the overall constraints on the high-redshift population. It may also be that, under an improved model, shot noise may end up providing complementary information to galaxy clustering to break certain degeneracies in model parameters.

Our overall model for the optical and infrared background fluctuations is

$$C_\ell = \begin{cases} C_\ell^{\mathrm{IHL}} + C_\ell^{\mathrm{DGL}} + C_\ell^{\mathrm{low-z}} + C_\ell^{\mathrm{shot}} + C_\ell^{6<z<8} + C_\ell^{8<z<13} & \text{F125W and above} \\ C_\ell^{\mathrm{IHL}} + C_\ell^{\mathrm{DGL}} + C_\ell^{\mathrm{low-z}} + C_\ell^{\mathrm{shot}} + C_\ell^{6<z<8} & \text{F775W and F850LP} \\ C_\ell^{\mathrm{IHL}} + C_\ell^{\mathrm{DGL}} + C_\ell^{\mathrm{low-z}} + C_\ell^{\mathrm{shot}} & \text{F606W} \end{cases} \quad (2.15)$$

Given that we are not able to constrain the the amplitude of $C_\ell^{6<z<8}$ given the degeneracies with the parameters involving the IHL model, and the fact that we only have a single band below it, we set $C_\ell^{6<z<8}$ based on the default prediction of our model, but allow the overall amplitude $A_{6<z<8}$ to vary such that it uniformly samples the SFRD between [0.003,0.2] $\mathrm{M}_\odot$ yr^{-1} Mpc^{-3}. The range is fully consistent with the existing measurements on the SFRD between $z = 6$ and 8 (Bouwens et al., 2014; Zheng et al., 2012; Coe et al., 2013; Oesch et al., 2014). Our constraint on $A_{8<z<13}$ is mostly independent of this parameter since we can safely constrain the Lyman-dropout signal between 0.8 and 1.25 μm with our existing data.

We also included the CIBER (Zemcov et al., 2014) data at 1.1 and 1.6 μm and Spitzer (Cooray et al., 2012b) data at 3.6 μm in our fitting process. When compared to the Hubble data at 1.25 and 1.6 μm, we find the CIBER data are likely dominated by the emission from a DGL-like signal at large angular scales, and low-z faint galaxies at $z < 5$ at small angular scales (Figure 5). For the fluctuations from faint, low-z galaxies, we adopt a model of residual galaxies which is derived from the observations of the luminosity function for different near-IR bands (Helgason et al., 2012). This model already includes the shot noise term, and we add a scale factor $f_{\mathrm{low}-z}$ to vary the low-z angular power spectrum, $C_\ell^{\mathrm{low}-z}$, in 1 σ uncertainty. For the DGL component, we use the C_ℓ^{DGL} of Hubble data at 1.25 and 1.6 μm to fit the CIBER data at 1.1 and 1.6 μm.

We perform joint fits for Hubble, CIBER and Spitzer data with the Markov Chain Monte Carlo (MCMC) method. The Metropolis-Hastings algorithm is used to find the probability of acceptance of a new MCMC chain point (Metropolis et al., 1953; Hastings, 1970). We estimate the likelihood function as $\mathcal{L} \propto \exp(-\chi^2/2)$, where χ^2 is given by

$$\chi^2 = \sum_{i=1}^{N_d} \frac{(C_\ell^{\mathrm{obs}} - C_\ell^{\mathrm{th}})^2}{\sigma_\ell^2}. \tag{2.16}$$

Here C_ℓ^{obs} and C_ℓ^{th} are the observed and theoretical angular power spectra for HST, Spitzer or CIBER data, respectively. σ_ℓ is the error for each data point at ℓ, and N_d is the number of data points. The total χ^2 of HST, Spitzer and CIBER is $\chi^2_{\mathrm{tot}} = \chi^2_{\mathrm{HST}} + \chi^2_{\mathrm{CIBER}} + \chi^2_{\mathrm{Spitzer}}$.

We assume a flat prior probability distribution for the free parameters; see Table 1 for prior information. Both A_{DGL}, C_ℓ^{shot} vary as independent parameters for each band. Both the A_{DGL} and C_ℓ^{shot} parameters are six-fold, with one for each HST band and for Spitzer/IRAC

3.6 μm. (we combined the two CIBER bands with two of the HST bands). We have two parameters for IHL and one parameter for the normalization of the reionization galaxies with $A_{8<z<13}$. We have two more parameters that we vary, $A_{6<z<8}$ and $f_{\mathrm{low}-z}$. We set a uniform prior on $A_{6<z<8}$ in the SFRD following the existing measurements to be between 0.003 and 0.2 $\mathrm{M}_\odot$ yr^{-1} Mpc^{-3}. We also set a uniform prior on $f_{\mathrm{low}-z}$ over a reasonable range of models to account for the overall uncertainty in the models of Ref. Helgason et al. (2012) to describe the $z < 5$ faint galaxy clustering at the same masking depth as our measurements. We marginalize over both $A_{6<z<8}$ and $f_{\mathrm{low}-z}$ as well as all other parameters when quoting results for an individual parameter. We have a total of 14 free parameters in our MCMC fitting procedure that we extract from the data. Among these parameters, 12 of them simply describe the small and large angular scale fluctuations in each of the bands we have performed the measurements. These parameters are summarized in Table 1. We generate twenty MCMC chains, where each chain contains about 100,000 points after convergence. After thinning the chains, we merge all chains and collect about 10,000 points for illustrating the probability distributions of the parameters. Contour maps for each of the fitted model parameters are shown in Figure 8. Our best-fit model with 14 free parameters have a minimum χ^2 value of 278 for a total degrees of freedom of $N_{\mathrm{dof}} = 104$.

2.7.2 Discussion

Our results are summarized in Figure 3, where we show the best-fit model curves. While the dominant contribution to the excess fluctuations comes from DGL at $\ell < 10^4$, at intermediate scales we find the IHL and reionization contributions to be roughly comparable. In Figure 6 we show the rms fluctuation amplitude at ~ 5 arcminute angular scales over the interval $10000 < \ell < 30000$. We find a spectral energy distribution that is consistent with Rayleigh-Jeans (RJ) from 4.5 to 2.4 μm, but diverges between 2.4 and 1.6 μm, and even more rapidly between 1.25 and 0.85 μm. The fluctuations can be explained with a combination of IHL

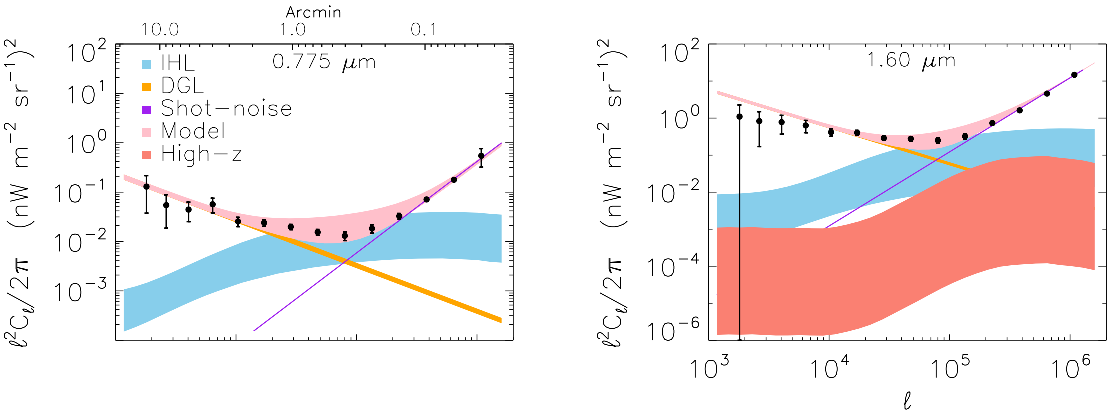

Figure 2.16 Fluctuation power spectra and the best-fit models with 1σ error bounds for the model components are shown at 0.775 μm in **b** and 1.60 μm in **c**. The dominant model contributors to the total power spectrum are DGL at low multipoles, or angular scales greater than a few arcminutes, IHL at intermediate multipoles corresponding to angular scales of about an arcminute, and shot noise associated with faint low-redshift dwarf galaxies dominating the high multipoles or sub-arcminute angular scales. The clustering signal of low-z galaxies is more than an order of magnitude below the lower limit plotted here, thus we did not include a low-z component in our modeling.

and high-redshift galaxies. The residual low-z galaxy signal is small but non-negligible. We find that it is mostly degenerate with IHL, especially if we allow its amplitude to vary more freely than the range allowed by the existing models based on $z < 5$ galaxy luminosity functions (Helgason et al., 2012). Thus modeling uncertainities related to the low-z galaxy confusion do not contaminate our statements about reionization. Assuming the existing low-z galaxy model (Helgason et al., 2012), the best-fit model is such that the IHL intensity peaks at lower redshifts with decreasing wavelength (Figure 7). At 3.6 μm, the IHL signal is associated with galaxies at $z \sim 1$, while at 0.6 μm over 80% of the signal is associated with galaxies at $z < 0.5$. The total intensities are $0.13^{+0.08}_{-0.05}$, $0.23^{+0.17}_{-0.11}$, $0.27^{+0.21}_{-0.13}$, $0.45^{+0.43}_{-0.24}$, and $0.54^{+0.58}_{-0.31}$ nW m^{-2} sr^{-1} at 0.60, 0.77, 0.85, 1.25 and 1.6 μm, respectively. We find that the implied IHL intensities at 1.25 and 1.6 μm are a factor of 10 lower than the implied IHL intensities for a model of CIBER fluctuations with IHL alone. The difference is due to the CIBER model that only included IHL and ignored the presence of DGL.

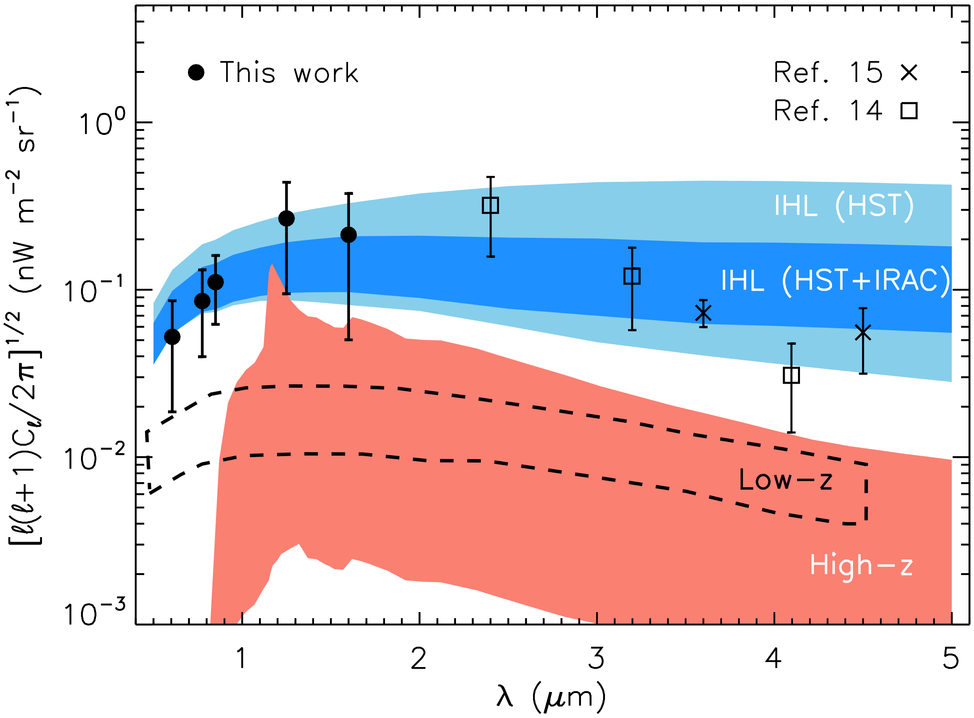

Figure 2.17 Spectral energy distribution of optical and infrared fluctuations at arcminute angular scale. The Hubble/CANDELS points are averaged over $10^4 \leq \ell \leq 3\times10^4$, with the best-fit shot noise and DGL components subtracted. Our model fits for the high-redshift and IHL components, with their 1 σ bounds, are shown as the filled regions. The errors here are propagated from the errors on the auto spectrum at the same ℓ range. The light blue region shows the 1 σ confidence bound for the IHL component when we use only the HST data in our model fitting; the dark blue region shows the 1 σ confidence bound for the IHL component when use both the HST and Spitzer IRAC data in our model fitting. The light red colored region signifies the 1 σ error bound for the high-redshift model component. The dashed line corresponds to the 1 σ bound for the low-redshift component. The Spitzer (Cooray et al., 2012b) and AKARI (Matsumoto et al., 2011) data are taken from previous measurements at $\ell = 3000$. Note the spectral dependence difference between the high-redshift signal and IHL. Below 0.8 μm we do not expect any signal from $z > 8$ galaxies. The presence of fluctuations at optical wavelengths requires a low-redshift signal in addition to high-redshift sources to explain combined optical and IR background intensity fluctuations.

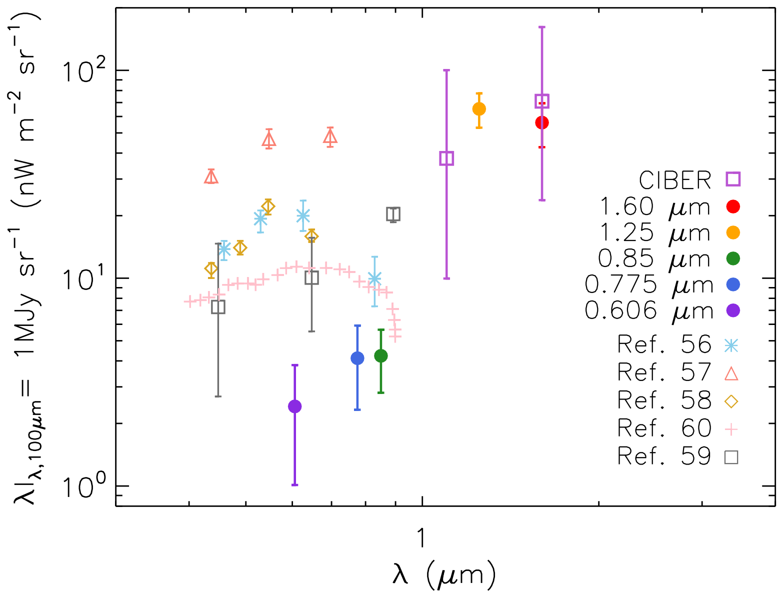

Figure 2.18 Optical and infrared diffuse galactic light (DGL) spectrum. The CANDELS points are taken from the DGL model components at $10^4 \leq \ell \leq 3 \times 10^4$, and the CIBER points are taken directly from Fig 2. of Ref.Zemcov et al. (2014) where they subtract off the shot noise component from their data. The galactic latitude for the optical points are $|b| \simeq 39°, 32°, 41°, 40°$ for the points labeled WittWitt et al. (2008), PaleyPaley et al. (1991), IenakaIenaka et al. (2013) and GuhathakurtaGuhathakurta & Tyson (1989). The BrandtBrandt & Draine (2012) points are modeled over the full sky. GOODS-S is at a galactic latitude of $|b| = 54°$.

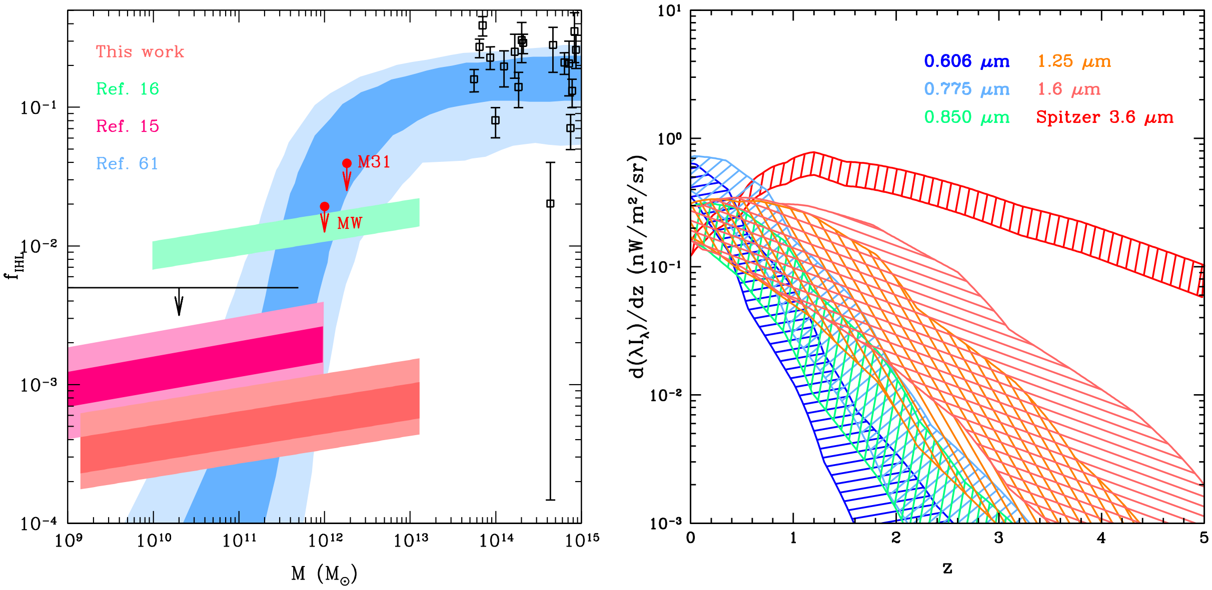

Figure 2.19 Intrahalo light fraction and model intensities. The left panel shows $f_{\rm IHL}$, the intrahalo light fraction, as a function of halo mass. The dark and light shaded regions show the 95% and 68% ranges of $f_{\rm IHL}$ from anisotropy measurements, and from an analytical predictionPurcell et al. (2007) (blue). Intracluster measurements are shown as boxesGonzalez et al. (2005), with 1 σ errors. The red downward arrows denote the 95% confidence upper limit on $f_{\rm IHL}$ estimated for Andromeda (M31) and our Milky Way (MW), following Figure 2 of Ref. Cooray et al. (2012b). The right panel shows $d(\lambda\, I_\lambda)/dz$ from the model, as a function of redshift. We show the 68% confidence uncertainties derived from MCMC fitting of the data at 0.606, 0.775, 0.850, 1.25 and 1.6 μm. The total IHL intensity is $0.13^{+0.08}_{-0.05}$, $0.24^{+0.17}_{-0.11}$, $0.28^{+0.21}_{-0.13}$, $0.45^{+0.43}_{-0.24}$, and $0.54^{+0.58}_{-0.31}$ nW m^{-2} sr^{-1} for 0.606, 0.775, 0.850, 1.25 and 1.6 μm, respectively.

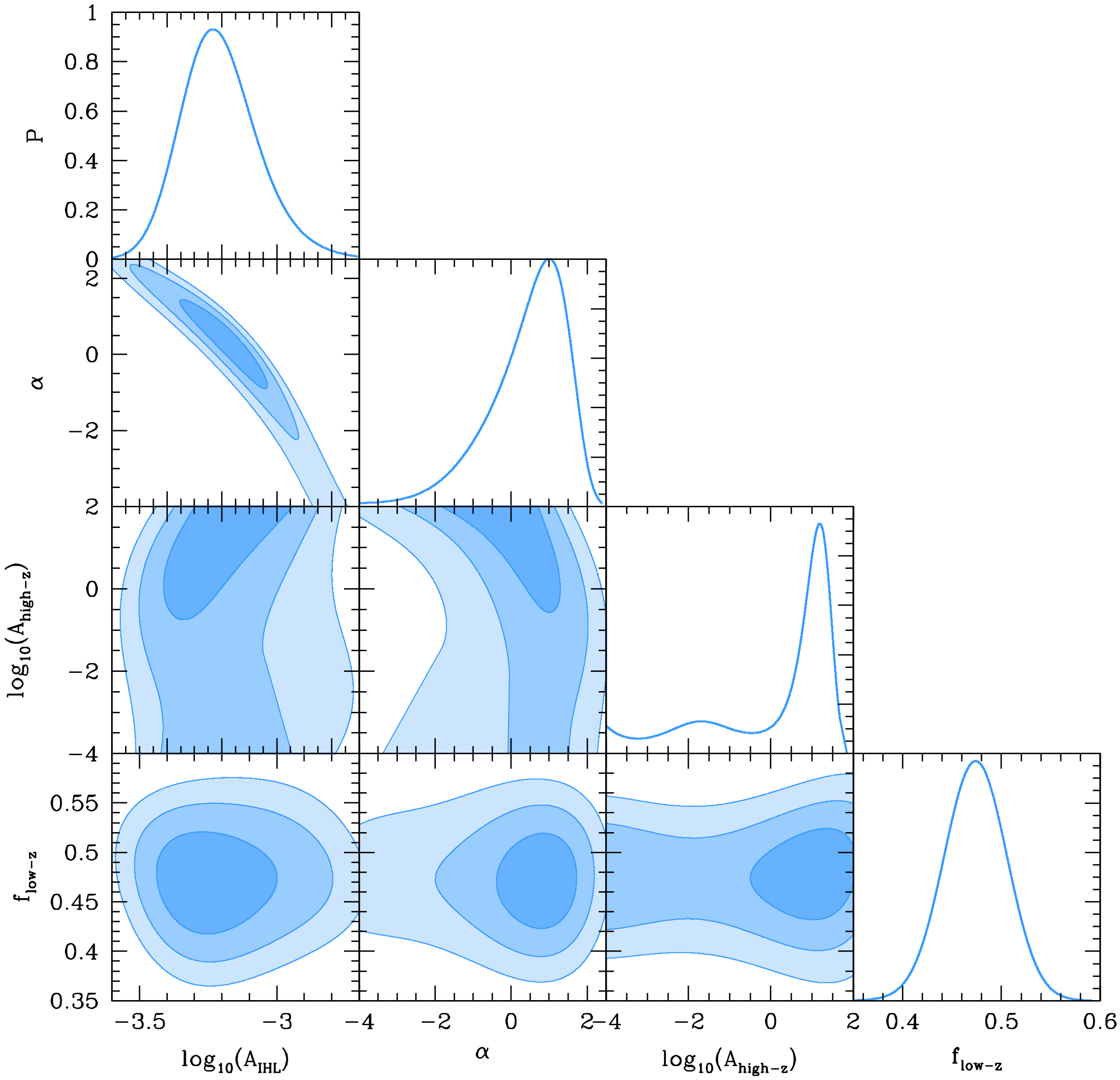

Figure 2.20 Probability distributions of fitted model parameters. Here we show the probablity density distributions for our fitted model parameters $\log_{10}(\mathrm{A}_{\mathrm{IHL}})$, α, $f_{\mathrm{low-z}}$, and $\log_{10}(\mathrm{A}_{\mathrm{high-z}})$ corresponding to the distribution from $8 \leq z \leq 13$. The single curves on the outermost column of each row, labeled with a "P", show the marginalized probability distribution for each parameter labeled on the bottom of the figure. Contour regions to the left of these probability distributions show how the parameters scale with one another. Each of the shaded regions in the contours correspond to the 1, 2 and 3 σ uncertainty ranges.

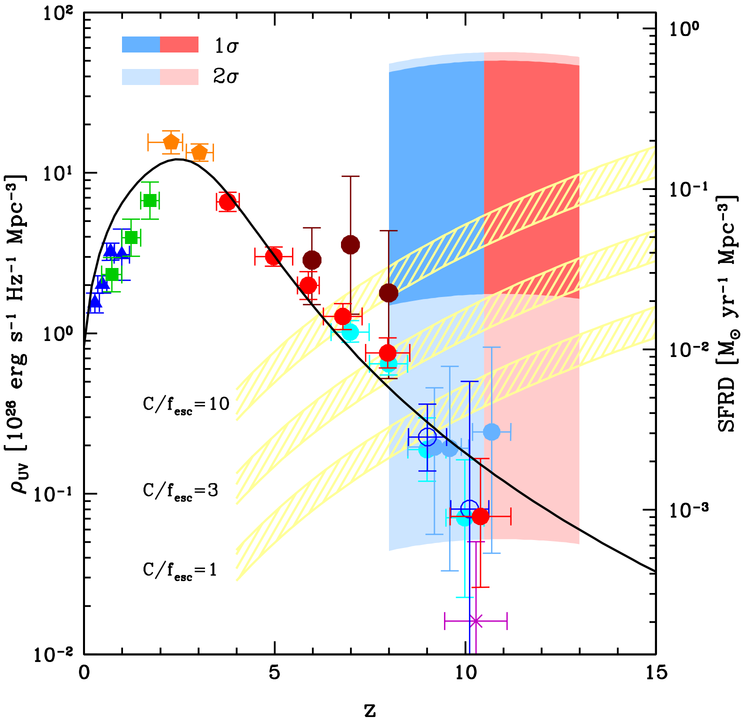

Figure 2.21 Plotted here is the specific UV luminosity density (left axis), with the equivalent star formation rate density (SFRD, right axis), as a function of the redshift z. We show the 1 σ and 2 σ error bounds in our redshift bin as the light and dark blue regions. Results from low-redshift surveys are shown as blue trianglesSchiminovich et al. (2005), bright green squaresOesch et al. (2010), and orange pentagonsReddy & Steidel (2009). At $z \sim 4$ to 10 the star formation rate density is shown to decrease with increasing redshift as measured by previous works, plotted as filled cyan circlesMcLure et al. (2013), filled red circlesBouwens et al. (2014), open red circlesFinkelstein et al. (2014) filled green circlesZheng et al. (2012); Coe et al. (2013) and open blue circlesOesch et al. (2014). Gamma ray burst (GRB) studies are shown as gray trianglesKistler et al. (2009), squaresRobertson & Ellis (2012) and dark gray circlesTanvir et al. (2012). Except for Ref. McLure et al. (2013), other estimates are luminosity function extrapolations and integrations down to $M_{\rm UV} = -13$. Our measured star formation rate densities are consistent with previous works between $z \sim 8$ to 10, however only extremely bright galaxies are directly detected in the aforementioned works with extrapolations down to $M_{\rm UV} = -13$ involves the measured faint-end slope of the luminosity function. For reference we plot the theoretically expected relationMadau et al. (1999) between UV luminosity density and redshift to reionize the universe and/or to maintain reionization using an optical depth to reionization of $\tau = 0.066 \pm 0.012$Planck Collaboration et al. (2015). We take a gas clumping factor of $C = 3$ and show two cases where the escape fraction of galaxies $f_{\rm esc}$ is 6% and 20% (see also Ref. Robertson et al. (2013)).

Parameter	Best fit	Best fit (no high-z)	Prior min, max
$\log_{10}(\mathrm{A}_{8\leq \mathrm{z}\leq 13})$	$1.19^{+0.27}_{-2.62}$	-	-5, 7
$\log_{10}(\mathrm{A}_{\mathrm{IHL}})$	$-3.23^{+0.14}_{-0.12}$	$-3.32^{+0.25}_{-0.09}$	-6, 10
α	$1.00^{+0.61}_{-0.99}$	$1.35^{+0.39}_{-0.73}$	-5, 5
$f_{\mathrm{low-z}}$	0.47 ± 0.03	0.47 ± 0.03	0.1, 10
$\mathrm{A}^{1.6}_{\mathrm{DGL}}$	$(3.74^{+0.30}_{-0.45}) \times 10^4$	$(3.72^{+0.35}_{-0.38}) \times 10^4$	10^3, 10^5
$\mathrm{A}^{1.1\&1.25}_{\mathrm{DGL}}$	$(4.35^{+0.54}_{-0.79}) \times 10^4$	$(4.72^{+0.42}_{-0.48}) \times 10^4$	10^3,10^5
$\mathrm{A}^{0.850}_{\mathrm{DGL}}$	$(2.83^{+0.40}_{-0.42}) \times 10^3$	$(2.77^{+0.32}_{-0.34}) \times 10^3$	$10^2, 10^4$
$\mathrm{A}^{0.775}_{\mathrm{DGL}}$	$(2.74^{-0.36}_{-0.38}) \times 10^3$	$(2.65^{+0.38}_{-0.48}) \times 10^3$	$10^2, 10^4$
$\mathrm{A}^{0.606}_{\mathrm{DGL}}$	$(1.61^{+0.20}_{-0.40}) \times 10^3$	$(1.43^{+0.23}_{-0.22}) \times 10^3$	10^2, 10^4
$\mathrm{C}^{1.6}_{\ell,\mathrm{shot}}$	$(7.54 \pm 0.13) \times 10^{-11}$	$(7.54 \pm 0.13) \times 10^{-11}$	10^{-11},10^{-10}
$\mathrm{C}^{1.25}_{\ell,\mathrm{shot}}$	$(7.77^{+0.21}_{-0.28}) \times 10^{-11}$	$(7.77 \pm 0.14) \times 10^{-11}$	10^{-11}, 10^{-10}
$\mathrm{C}^{0.850}_{\ell,\mathrm{shot}}$	$(7.73^{+0.75}_{-0.45}) \times 10^{-12}$	$(8.10 \pm 0.45) \times 10^{-12}$	10^{-12},10^{-11}
$\mathrm{C}^{0.775}_{\ell,\mathrm{shot}}$	$(4.60^{+0.50}_{-0.30}) \times 10^{-12}$	$(4.65 \pm 0.30) \times 10^{-12}$	10^{-12}, 10^{-11}
$\mathrm{C}^{0.606}_{\ell,\mathrm{shot}}$	$(3.27^{+0.24}_{-0.21}) \times 10^{-12}$	$(3.39 \pm 0.15) \times 10^{-12}$	10^{-13}, 10^{-11}

Table 2.4 Summary of free model parameters. The best-fit values are quoted with 1σ errors. $\log_{10}(\mathrm{A}_{8\leq \mathrm{z}\leq 13})$ is the high-redshift component used to constrain the SFRD during the reionization epoch, which is fit to the 1.25 and 1.60 μm bands. $\log_{10}(\mathrm{A}_{\mathrm{IHL}})$ and α are the two parameters necessary to describe the IHL component (to wit: C_ℓ^{IHL} in Equation 9). $f_{\mathrm{low-z}}$ is the low redshift scaling factor which varies the low redshift power spectrum within a 1 σ uncertainty. $\mathrm{A}^i_{\mathrm{DGL}}$ and $C^i_{\ell,\mathrm{shot}}$ are respectively the DGL amplitude scaling factor and shot noise at wavelength i. All parameter values have units of (nW m^{-2} sr^{-1})^2.

The drop in the fluctuation amplitude from 1.25/1.6 μm to 0.85 μm allows for a signal from reionization, but the presence of fluctuations at shorter wavelengths, such as 0.6 μm, rules out a scenario in which reionization sources are the sole explanation for the fluctuations at wavelengths at 1 μm and above. The 3.6 μm and X-ray cross-correlation (Cappelluti et al., 2013) was explained with primordial direct collapse blackholes at $z > 12$ (Yue et al., 2013a). In our multi-component model we are able to account for the presence of fluctuations at short wavelengths with IHL, DGL and faint low-redshift galaxies, while a combination of those components and high-redshift galaxies are preferred to account for fluctuations at 1.25 and 1.6 μm. The high-z signal is modeled following the calculations in Cooray et al. (2012a). The signal has an overall amplitude scaling that is related to the star-formation rate during reionization. The bright end of the counts are normalized to existing luminosity function measurements, and the faint-end of the luminosity functions to have a steeper slope than measured with counts extending down to arbitrarily low luminosities. In order to test whether a component at high redshift is required to explain the measurements, we also re-ran the MCMC model fits but with $A_{\rm high-z}$ fixed at 0. In this case our best-fit model with 13 free parameters has a minimum χ^2 value of 283 for a total degrees of freedom of $N_{\rm dof} = 105$. The difference in the best-fit χ^2 values with and without a model for high-redshift galaxies suggests a p-value of 0.025. This is consistent with the 2σ to 3σ detection of $8 < z < 13$ signal in the fluctuations (Figure 6).

With multi-wavelength measurements extending down to the optical, we are now able to constrain the amplitude of that signal with a model that also accounts for low-redshift sources in a consistent manner. This improves over previous qualitative arguments that have been made, or models involving high-redshift sources alone that have been presented, for the presence of a signal from reionization in the IR background fluctuations (Kashlinsky et al., 2012; Matsumoto et al., 2011; Yue et al., 2013a). In our models, the total intensity arising from all galaxies at $z > 6$ is $\log \nu I_\nu = -0.32 \pm 0.12$ in units of nW m^{-2} sr^{-1} at 1.6 μm. At 1.6 μm the intensity from high-redshift sources is dominated by $z > 8$ galaxies, while at 0.85 μm we find an intensity $\log \nu I_\nu = -0.75 \pm 0.05$ in units of nW m^{-2} sr^{-1} for $6 < z < 8$ galaxies. The total intensity from $z > 8$ galaxies in the 1.6 μm band is comparable to the IHL intensity at the same wavelength (Figure 7). However, at 3.6 μm, the IHL signal is a factor of about 5 times brighter than the $z > 8$ galaxies. At 1.6 μm the total of the IHL, high-z, and integrated galaxy light (Franceschini et al., 2008) of $10.0^{+2.7}_{-1.8}$ nW m^{-2} sr^{-1}is comparable to the EBL intensity inferred by gamma-ray absorption data (H.E.S.S. Collaboration et al., 2013) of $15 \pm 2(\rm stat) \pm 3(\rm sys)$.

Using the best-fit model and uncertainties as determined by MCMC model fits, we also convert the $A_{8<z<13}$ constraint to a measure of the luminosity density of the universe at $z > 8$ (Figure 9). The resulting constraint is $\log \rho_{\rm UV} = \mathbf{27.4^{+0.2}_{-1.2}}$ in units of erg s^{-1} Hz^{-1} Mpc^{-3} at ($1\,\sigma$). As shown in Figure 9, the 68% confidence level constraint on $\rho_{\rm UV}$ is higher

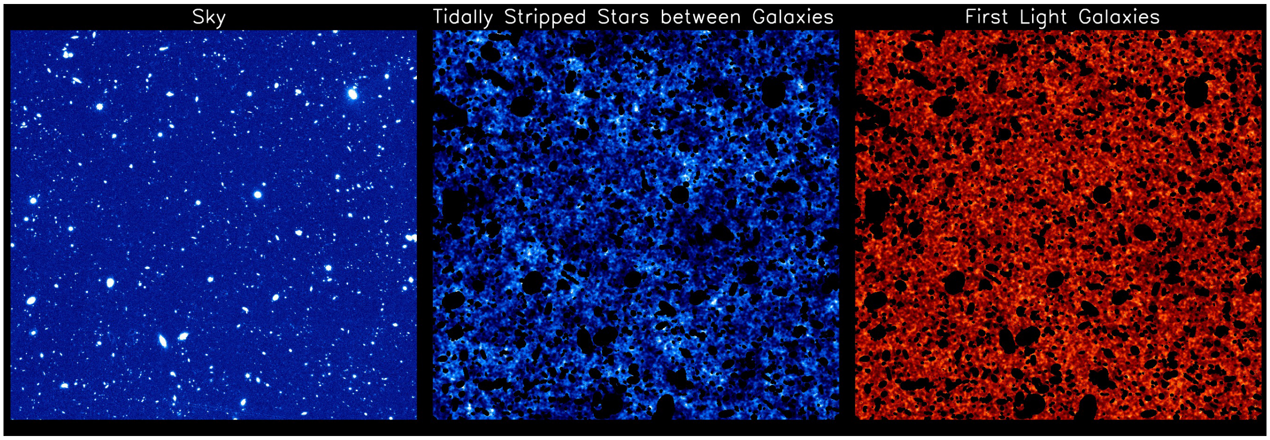

Figure 2.22 Using the model-fit power spectrum parameters, we can re-project the statistical distribution of different model components to real space. In the left panel is the unmasked part of the sky that is reprojected (although the statistics will be the same over the full sky). In the center panel is the IHL model-fit power spectrum reconstructed, with the foreground source mask overlaid. On the right panel is what the first-light galaxies look like according to our models. The relative intensities of these reconstructions are scaled arbitrarily for viewing.

than the existing results from Lyman drop-out galaxy surveys during reionization at $z > 8$ (Atek et al., 2015; Robertson et al., 2013), and especially at $z \sim 10$ (Zheng et al., 2012). At the 95% confidence level, our measurement is fully consistent with the existing results at $z \sim 10$ (Bouwens et al., 2015). Our constraint allows for the possibility that a substantial fraction of the UV photons from the reionization era is coming from fainter sources at depths well below the detection threshold of existing Lyman dropout surveys, as is indeed anticipated from the steep measured slopes of the UV luminosity functions from detected galaxies. Despite their lack of detections in the deepest surveys with HST, the majority of the faint sources responsible for both fluctuations and reionization should be detectable in deep surveys with JWST centered at 1 μm.

Chapter 3

NIR/Optical cross-correlations with X-Ray

3.1 Introduction

In this Chapter we show results of CIB × CXB cross-correlations in the *Chandra* Deep Field-South (CDF-S/GOODS-S; Giavalisco et al. 2004), using data from one of the deepest *Chandra* X-ray surveys, deep *Hubble* exposures from CANDELS and archival HST data, and a combination of a large collection of surveys in IRAC channels 1 and 2. In Section 3.2 we describe our initial data reduction and map-making techinques, and in Section 3.3.1 we present our cross-correlation measurements, which are discussed in context in Section 3.5.

3.2 Map making

We have assembled maps from three different space telescopes to perform the cross-correlations. For the power spectrum analyses we perform jack-knife tests, so we generate two maps for

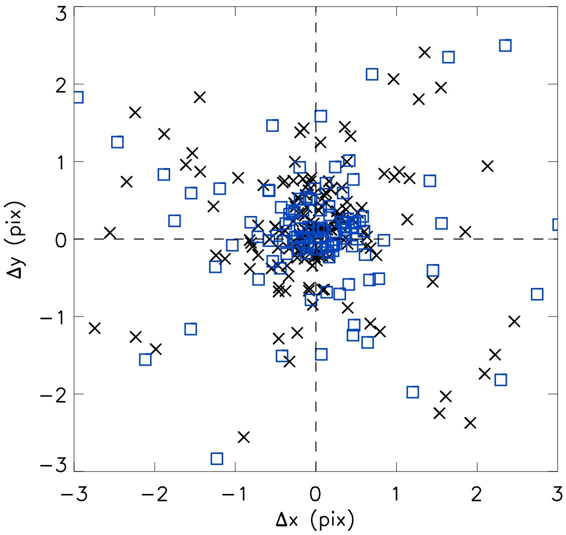

Figure 3.1 Astrometric alignment between the X-ray and optical/NIR maps after performing the astrometric shift with `tweakreg`. IRAC detected sources matched with X-ray are denoted as black crosses, and HST sources matched with X-ray are blue squares. Prior to the astrometric shift with `tweakreg`, the offset was ∼ 0.1−1 pixel at the IRAC pixel scale.

each filter with the exact same sky coverage by subdiving the exposure time into two subsets. Each of the IRAC and *Chandra* maps have large area coverage, with a limiting area defined by the CANDELS observations; both the CANDELS and *Chandra* maps have small pixel scales with $0.''14$ and $0.''5$ pix^{-1} respectively. Thus for all the aligned maps which are used in this work, we are limited by the native IRAC pixel scale at $1.2''$ pix^{-1}, and a ∼ 110 arcmin^2 survey area, defined by the CANDELS observations. Subsequent to generating mosaics for each instrument, they are regridded and cropped to these specifications. Since the mosaics from different instruments are each using different astrometric references, the astrometric alignment between the three was slightly shifted, by 0.1 to 1 pixel. We removed this offset by aligning the X-ray and *Spitzer* maps to the HST astrometry using the `tweakreg` PyRAF module (see Fig 3.1).

3.2.1 *Chandra* X-ray maps

We used *Chandra* data of ECDF-S from Cycle 8 to Cycle 12 as listed in Table 3.1. The data span 4 Ms of observations and we make two maps with an effective exposure time of ~ 1.5 Ms each, generated from a total of 43 observations. The first mosaic is generated from 19 observations, and the second from 24 observations. The details of these observations are listed in Table 3.1, which are a subset of those listed in Table 1 of Luo et al. (2008) and Table 1 of Xue et al. (2011). All observations were done in VFAINT mode. The collection of *Chandra* frames we use here represents the VFAINT subset of the total 4 Ms observations, as it omits all observations done in FAINT mode. Basic processing of the data was done using the *Chandra* X-ray Center (CXC) pipeline. To ensure uniformity, mosaics were produced with the *Chandra* Interactive Analysis of Observations (`CIAO`), following the details described in Giacconi et al. (2002), Alexander et al. (2003) and Xue et al. (2011). However for clarity, we will briefly review the data reduction and analysis procedure here.

Background light curves were inspected using `EVENT_BROWSER` in the Tools for ACIS Real-time Analysis (TARA; Broos 2000). All data were corrected for the radiation damage of the CCDs during the first few months of *Chandra* operations using the charge transfer inefficienty correction procedure of Townsley et al. (2000, 2002). All bad columns, bad pixels and cosmic ray afterglows were removed using the "status" information in the event files, and only observations taken within the CXC-generated good-time intervals was used. `ACIS_PROCESS_EVENTS` was used to identify potential background events and to remove the standard pixel randomization.

The observations are then split into three standard broad bands: 0.5-2 keV ("soft"), 2-8 keV ("hard") and 0.5-8 keV ("full"). The mosaics cover a total area of ~ 450 arcmin^2 (before alignment to the CANDELS field of view), and reach an on-axis sensitivity of 5.1×10^{-18}, 3.7×10^{-17} and 2.4×10^{-17} erg s^{-1} cm^{-2} respectively for the soft, hard and full bands (Lehmer et al., 2012). Our final merged event file includes only 0.3-10 keV events from which make our mosaics.

The background in the X-ray mosaics contains a high energy particle component in addition to the cosmic signal in which we are interested. The spectral shape of the particle background is fairly flat in the soft and hard bands, and above ∼9 keV is solely a product of non-astrophysical particle events. Per the prescription in Hickox & Markevitch (2006), we can isolate the CXB in our X-ray maps by subtracting off the particle background measured in the 9-10 keV exposures. The particle background in the ACIS detectors is a product of interactions of the CCDs and surrounding materials with high-energy particles. Taking ACIS exposures while the instrument is stowed within the detector housing blocks celestial X-rays and thus accurately represents the particle background of the detector.

A stowed image is produced from CALDB version 4.7. A particle background map is then approximated by simply normalizing the stowed image as $C_{\rm data}[9\text{-}10]/C_{\rm stow}[9\text{-}10]$, as is discussed in Hickox & Markevitch (2006), where the counts for $C_{\rm data}$ and $C_{\rm stow}$ are respectively measured from the total counts in the real images and stowed images. We then subtract this particle background from each of our pointings, the same as was done in C13. This will effectively remove the particle background from our analysis, and is an additional step to the reduction procedures detailed in Xue et al. (2011). We exclude energies above 10 keV because all events above that threshold were filtered out in the initial reduction.

In Hickox & Markevitch (2006), they state that data collected in VFAINT mode are the most reliable for background sensitive measurements, however their two other datasets in FAINT mode provide consistent results. In C13 they only use VFAINT data; for the sake of consistency with C13, we have omitted all FAINT observations, which account for ~ 1 Ms of exposure time in CDF-S. In addition to the results presented here, we also performed our analysis with the full 4 Ms data set and obtained very similar results, which is consistent with the results of Hickox & Markevitch (2006).

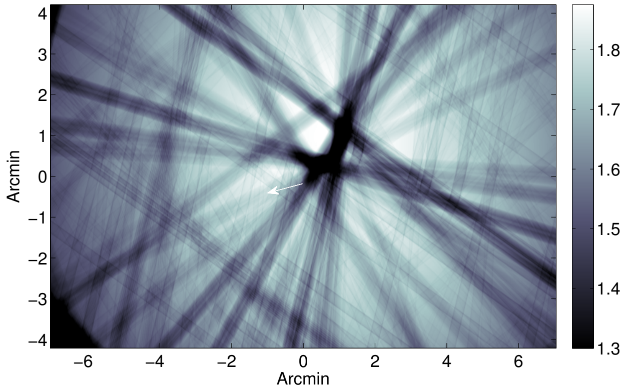

Figure 3.2 Effective exposure time map for the soft band, in the first half CDF-S mosaic, in units of Ms. The white arrow points North from the center of the map.

We convert *Chandra* event maps into flux maps in units of erg s^{-1} cm^{-2} by dividing by the effective exposure time maps (see Fig. 3.2), and assuming a power-law index of 1.4 and a Galactic neutral hydrogen column density of 8.8×10^{19} cm^{-2}. With these criteria, we used the NASA's HEASARC online tool (WebPIMMS; https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3pimms/w3pimms.pl) to obtain conversion factors (ecf) for the soft, hard, and full bands of 2.15×10^{10}, 1.90×10^{10} and 3.53×10^{10} erg^{-1} cm^2, respectively. These mosaics are then aligned to the CANDELS field of view, repixelized to $1.2''$ pix^{-1}, and smoothed by a $3.6''$ Gaussian kernel. The full mosaic, along with the HST CDF-S footprint can be seen in Figure 1 of Xue et al. (2011).

3.2.2 Self-calibration

In order to measure fluctuations over large angular scales, we need to mosaic many individual exposures with smaller fields of view. This becomes especially challenging at NIR wavelengths because of the relatively high intensities of foreground emissions at these wavelengths. Sunlight scattered off of dust in the solar system, or "Zodiacal light", is a temporal foreground that will affect different frames exposed at different times of the year, and leads

Table 3.1. Journal of *Chandra* Deep Field-South Observations.

	Obs. Start	Exposure	Aim Point		Roll Angle	
Obs. ID	(UT)	Time (ks)	α (J2000.0)	δ (J2000.0)	(deg)	Half
8591.......	2007 Sep 20, 05:26	45.4	03 32 28.20	−27 48 06.9	72.7	1
9593.......	2007 Sep 22, 20:34	46.4	03 32 28.20	−27 48 06.9	72.7	1
9718.......	2007 Oct 03, 13:56	49.4	03 32 28.61	−27 48 07.4	62.0	1
8593.......	2007 Oct 06, 02:04	49.5	03 32 28.61	−27 48 07.4	62.0	1
8597.......	2007 Oct 17, 07:07	59.3	03 32 29.25	−27 48 10.4	44.2	1
8595.......	2007 Oct 19, 14:16	115.4	03 32 29.35	−27 48 11.2	41.2	1
8592.......	2007 Oct 22, 12:14	86.6	03 32 29.62	−27 48 13.8	32.4	1
8596.......	2007 Oct 24, 13:20	115.1	03 32 29.62	−27 48 13.8	32.4	1
9575.......	2007 Oct 27, 05:43	108.7	03 32 29.62	−27 48 13.8	32.4	1
9578.......	2007 Oct 30, 22:35	38.6	03 32 29.84	−27 48 16.7	24.2	1
8594.......	2007 Nov 01, 11:51	141.4	03 32 29.84	−27 48 16.7	24.2	1
9596.......	2007 Nov 04, 04:11	111.9	03 32 29.95	−27 48 18.5	19.8	1
12043......	2010 Mar 18, 01:39	129.6	03 32 28.78	−27 48 52.1	252.2	1
12123......	2010 Mar 21, 08:08	24.8	03 32 28.78	−27 48 52.1	252.2	1
12044......	2010 Mar 23, 11:31	99.5	03 32 28.55	−27 48 51.9	246.2	1
12128......	2010 Mar 27, 13:08	22.8	03 32 28.55	−27 48 51.9	246.2	1
12045......	2010 Mar 28, 16:38	99.7	03 32 28.32	−27 48 51.4	240.2	1
12129......	2010 Apr 03, 15:21	77.1	03 32 28.33	−27 48 51.4	240.2	1
12135......	2010 Apr 06, 09:36	62.5	03 32 28.01	−27 48 50.2	231.7	1
12046......	2010 Apr 08, 08:17	78.0	03 32 28.01	−27 48 50.2	231.7	2
12047......	2010 Apr 12, 13:21	10.1	03 32 27.80	−27 48 48.9	225.2	2
12137......	2010 Apr 16, 08:53	92.8	03 32 27.59	−27 48 47.2	219.2	2
12138......	2010 Apr 18, 12:40	38.5	03 32 27.59	−27 48 47.3	219.2	2
12055......	2010 May 15, 17:15	80.7	03 32 26.72	−27 48 32.3	181.4	2
12213......	2010 May 17, 14:22	61.3	03 32 26.69	−27 48 31.1	178.9	2
12048......	2010 May 23, 07:09	138.1	03 32 26.64	−27 48 27.6	171.9	2
12049......	2010 May 28, 18:58	86.9	03 32 26.61	−27 48 24.4	165.5	2
12050......	2010 Jun 03, 06:47	29.7	03 32 26.61	−27 48 21.7	160.2	2
12222......	2010 Jun 05, 02:47	30.6	03 32 26.61	−27 48 21.7	160.2	2
12219......	2010 Jun 06, 16:30	33.7	03 32 26.61	−27 48 21.7	160.2	2
12051......	2010 Jun 10, 11:30	57.3	03 32 26.63	−27 48 19.2	155.2	2
12218......	2010 Jun 11, 10:18	88.0	03 32 26.63	−27 48 19.2	155.2	2
12223......	2010 Jun 13, 00:57	100.7	03 32 26.63	−27 48 19.2	155.2	2
12052......	2010 Jun 15, 16:02	110.4	03 32 26.70	−27 48 14.5	145.7	2
12220......	2010 Jun 18, 12:55	48.1	03 32 26.70	−27 48 14.5	145.7	2
12053......	2010 Jul 05, 03:12	68.1	03 32 27.02	−27 48 06.0	127.0	2
12054......	2010 Jul 09, 11:35	61.0	03 32 27.02	−27 48 06.1	127.0	2
12230......	2010 Jul 11, 03:52	33.8	03 32 27.02	−27 48 06.0	127.0	2
12231......	2010 Jul 12, 03:22	24.7	03 32 27.16	−27 48 03.6	121.2	2
12227......	2010 Jul 14, 21:04	54.3	03 32 27.16	−27 48 03.7	121.2	2
12233......	2010 Jul 16, 10:25	35.6	03 32 27.16	−27 48 03.7	121.2	2
12232......	2010 Jul 18, 19:53	32.9	03 32 27.16	−27 48 03.7	121.2	2
12234......	2010 Jul 22, 19:58	49.1	03 32 27.19	−27 48 03.3	120.2	2

Table 3.2 X-ray map counts. A and B denote each half map; asterisks denote masked maps.

Band	N_A	N_B	N_A^*	N_B^*	$\langle N \rangle$ pix^{-1}
soft	25848	26799	9396	9917	0.18
hard	72997	77424	32065	34116	0.51
full	98845	104223	41462	44034	0.69

to large overall offsets between frames of the same or similar sky area. If these offsets are not properly modeled and removed, they will lead to a fictious anisotropy signal. Atmospheric airglow from the earth can also produce large offsets for HST observations that push its orbital limits and observe close to Earth's limb. The self-calibration algorithm (Fixsen et al., 2000; Arendt et al., 2002b, 2010) is a least-squares calibration algorithm that was explicitly designed to model and remove these offsets. So we utilize this code to generate our own mosaics, instead of the publicly available mosaics genarated by astrodrizzle (Fruchter & Hook, 2002; Hack et al., 2012).

As is described in Arendt et al. (2002b), our data, D^i, are modeled as

$$D^i \approx G^p \, S^\alpha + F^p + F^q \tag{3.1}$$

where i indexes each of the pixels in the entire data set, G^p is the gain of each detector pixel, and S^α is the sky intenisy at each sky position α. F^p and F^q are offset terms per detector pixel (index p) and per frame (index q). If one can model their data in such a way, these last two interloper terms can be efficiently minimized or removed. The details of solving such an equation are nontrivial. We leave it to the reader to refer to Arendt et al. (2002b), Arendt et al. (2000) and Fixsen et al. (2000) for the solution to eq. (3.1). It can be seen that for each additional frame included in the mosaics, the index i increases by the number of pixels in each frame, which becomes exceedingly large for large area mosaics or small pixel scale frames (e.g. ACS). In these cases the self-calibration can take a considerable amount of time and computer memory to run.

3.2.3 *Hubble* optical and NIR maps

All of our HST data are publicly available, and were downloaded from the Barbara A. Mikulski Archive for Space Telescopes (MAST; located at https://archive.stsci.edu/hst/search.php). We assembled our own collection of calibrated, flat-fielded frames (FLT) from the MAST archive from ten different HST proposals (Beckwith et al., 2006; Giavalisco et al., 2004; Grogin et al., 2010b; Koekemoer et al., 2011; Windhorst et al., 2011). These data are collected in five different filters, collected with both the Advanced Camera for Surveys (ACS) and the infrared Wide Field Camera 3 (WFC3/IR). We did not necessarily include all of the frames from any one proposal. As the self-calibration is largely dependent on the dither pattern of the observations and requires significant pixel overlap (Arendt et al., 2000), additional frames of the same pointing, from e.g. the Hubble Ultra Deep Field (Beckwith et al., 2006), will not improve the self-calibration solution, but will increase the variance of the signal to noise of the mosaic (which will propogate into the power spectrum). We thus chose somewhat randomly from these observations, and included the full set of frames from the more wide and contiguous proposals (Giavalisco et al., 2004). We also preferentially selected ACS frames that were exposed before the fourth servicing mission (SM4), as those data are not plagued with the horizontal bias striping that was introduced after the replacement of the readout box during SM4 (Grogin et al., 2010b). This striping is a spatially correlated noise and corrupts any spatial correlation measurements. We preformed simulations (MW15) to show that a mosaic produced with $\lesssim 30\%$ of post-SM4 frames will be adequate to perform reliable angular correlation measurements.

All FLTs downloaded from MAST are reprocessed “on-the-fly”, meaning they use the most recent calibration files to subtract bias and dark frames, and perform flat-field corrections. With our collection of FLTs in hand, we continue with the basic data reduction as prescribed by the Space Telescope Science Institue, using PyRAF version 2.1.1. First, all post-SM4 ACS frames are destriped; all FLTs are then charge transfer efficiency corrected. Cosmic rays are flagged using the CRCLEAN PyRAF module, and sub-arcsecond astrometric alignment against the publicly available CANDELS multidrizzle mosaics[1] is performed with the PyRAF module TWEAKREG.

Finally, we feed these reduced FLT frames to our self-calibration algorithm to generate mosaics. We use the same self-calibration model as in Arendt et al. (2002b), another example of where HST data have been successfully self-calibrated. We de-weight bad pixels and cosmic arrays and iterate three times in order to find a self-calibration solution. Our input FLT frames are geometrically distorted with a pixel size of $0.''1354 \times 0.''120$ for WFC3, and $0.''0498 \times 0.''0502$ for ACS. We remove the distortion in the map making procedure and produce mosaics with a slightly larger geometrically square pixel size of $0.''14$. As was previously mentioned, we perform jack-knife measurements, so each filter has two mosaics of the same sky area generated from different exposures. This way we can perform the cross-correlation of two maps and any uncorrelated noise will drop out of the measurements. For HST observations, each pointing has at least two exposures, so separating the data into two sets is straightfoward. We end up with two mosaics for each of the 0.6 μm (F606W), 0.7 μm (F775W), 0.85 μm (F850LP), 1.25 μm (F125W) and 1.6 μm (F160W) bands. The HST exposure times range between 180 and 1469 seconds; tile patterns for each filter can be seen in Fig. 1 of MW15.

3.2.4 *Spitzer* NIR maps

Spitzer/IRAC Channel 1 (3.6 μm) and Channel 2 (4.5 μm) frames are also susceptible to the same foreground contaminats as the HST data. Therefore we use the self-calibration algorithm to generate the channel 1 and channel 2 mosaics. We assemble basic calibrated data (CBCD) from the Spitzer Heritage Archive. These data are generally of excellent quality and do not require significant reduction prior to mosaicing. All CBCDs are astrometrically

---

[1] http://candels.ucolick.org/data access/GOODS-S.html

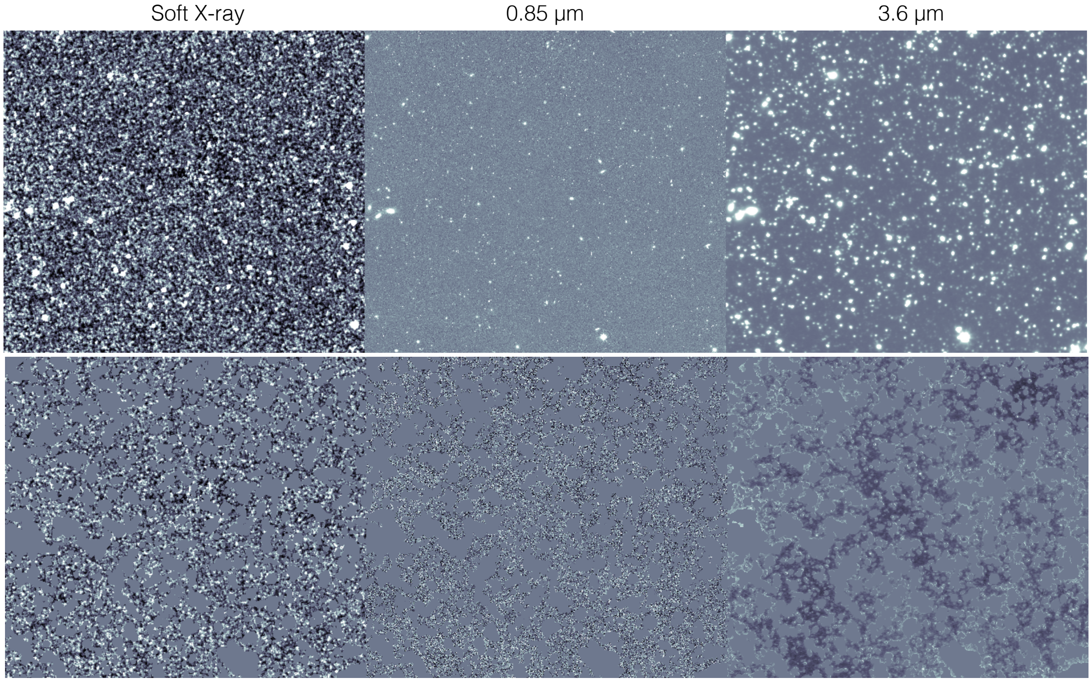

Figure 3.3 A cropped section of the mosaics in each band, both unmasked (top) and median-subtracted masked (bottom). The arrays have been scaled arbitrarily for viewing purposes, with common mean values in the unmasked regions, and common standard deviations in the masked images.

aligned, and those from the cryogenic mission do not, in general, contain arfifacts. The only issue one needs to mitigate before performing the self-calibration relates to the warm mission CBCDs. The Spitzer Science Center (SSC) does not have lab darks to correct for different delay times in the warm mission, which introduces a "stripiness" into the frames. In addition, the sky dark frames that are subtracted from the data during the standard CBCD reduction pipeline usually over-subtracts the background, resulting in overall negative pixel values. The straightforward solution to both of these issues is to generate a median image from all the warm CBCDs in our set, and subtract it from each warm mission CBCD before running the self-calibration.

CBCDs are are collected from 6 different proposal IDs: 10076 (IGOODS; Labbe et al. 2015), 60022 (SEDS; Ashby et al. 2013), 61052 (SERVS; Mauduit et al. 2012), 70145 (IUDF; Labbe et al. 2015), 70204 (ERS; Fazio et al. 2011), 80217 (S-CANDELS; Ashby et al. 2015). We again do not necessarily use all the frames from each proposal; instead we limit those to a region just slightly larger than the CANDELS area (so the self-calibration solution converges more quickly). Each of the *Spitzer*/IRAC mosaics are generated from 8089 individual CBCD frames.

Spitzer/IRAC observations do not necessarily have multiple exposures for each pointing. To split the CBCDs into two halves, in order to make jack-knife maps, we simply sort all CBCDs in each filter according to their right ascension and take every other frame for each half. Apart from this sorting, and subtracting the warm median frame from warm mission CBCDs, we do not perform any additional reduction prior to self-calibrating.

3.3 Power Spectra

3.3.1 Outline

Standard fast fourier transform (FFT) techniques are used to measure the angular power spectrum on these small area maps. We're looking for correlations in the diffuse background light, so this requires generating a resolved source mask in all the bands, and quantifying the mode-mode coupling introduced by the mask. We also need to quantify the map-making and tiling patterns (transfer function), and correct for the finite resolution of the telescope (beam transfer function). Although these procedures are discussed in detail in other works (Amblard et al., 2011; Hivon et al., 2002; Mitchell-Wynne et al., 2015; Thacker et al., 2013, 2015; Zemcov et al., 2014), we will discuss them here for clarity.

3.3.2 Resolved source mask

Resolved sources are masked in the HST bands down to $\sim 27\ m_{\mathrm{AB}}$; the mask generated for the work in MW15 is regridded to the larger pixel scale, which removes 47% of the pixels. We use the public X-ray main catalog from Xue et al. (2011) to generate an X-ray source mask, which in our field area corresponds to ~ 200 sources. We mask within circular radii proportional to the logarithm of the source counts; this increases the masking percentage by $\sim 5\%$. We do not include X-ray clusters in our source mask, as it would omit too many pixels. We mask IRAC sources down to $\sim 24\ m_{\mathrm{AB}}$ by convolving a 3σ detected `SExtractor` mask with the IRAC beam. After taking the union of these three masks, we sigma clip each masked map at a 5σ level, which removes an addtional $\sim 1\%$ of the pixels, to obtain our final mask. These 1% residual outliers are products of imperfect masking. In the end we remove $\sim 53\%$ of the pixels with the common source mask, which is applied to all bands and instruments. A cropped section of the source mask applied to the mosaics can be seen in the bottom panel of Fig 3.

When we apply the mask to our maps, it breaks some large scale modes into smaller scale modes, which needs to be corrected for. This correction is described in Hivon et al. (2002). We apply a mode-mode coupling correction with an $M_{\ell\ell'}$matrix (see Technical Supplement of Cooray et al. 2012a), which has been demonstrated with simulations (Cooray et al., 2012a; Mitchell-Wynne et al., 2015; Thacker et al., 2015) to recover the true power spectrum of a masked map. We generate the $M_{\ell\ell'}$matrix with the same methods described in Cooray et al. (2012a), using 100 simulations.

3.3.3 $T(\ell)$ correction

The next correction that is performed is the map-making transfer function, $T(\ell)$. For HST and *Spitzer*/IRAC we do this in a six step process:

1. Generate simulated image with known power spectrum which spans the entire survey area.
2. Fill each of the FLTs (for HST) or CBCDs (for *Spitzer*/IRAC) with the signal from the appropriate position on simulated map from the previous step.
3. Add Gaussian noise proportional to the inverse square root exposure time to each FLT or CBCD.
4. Add an overall offset equal to the median value of the science frame.
5. Feed this set of FLTs or CBCDs to selfcal and generate a mosaic.
6. Compare the input power spectrum to the output power spectrum to quantify $T(\ell)$.

Step 4 ensures that any residual offsets in the mosaics are corrected for. This six-step procedure is done 100 times for each band; the error bars for our transfer functions are just the standard deviation at each ℓ bin. The $T(\ell)$ for the HST mosaics can be seen in the Supplementary Section of MW15. The Spitzer $T(\ell)$ is plotted in Fig. 3.4.

Since the X-ray maps aren't generated from frames plagued with offsets, we just need to check how the varying effective exposure time across the mosaic affects the power spectrum. This is done simply by generating 100 Gausian maps (constant C_ℓ's), multiplying each map with the effective exposure time map (shown in Fig. 3.2), and measuring how far from constancy the resulting power spectra are. Since the variance of the effective exposure times is small, this results in a transfer function that is roughly constant at all scales, as is shown in Fig. 3.4. Error bars are again taken as the standard deviation at each bin.

3.3.4 B(ℓ) correction

Because of the finite resolutions of the telescopes used in this analysis, the power spectra will drop off at small angular scales. We correct this with a beam transfer function, $B(\ell)$. We measure the *Hubble*/WFC3 and ACS $B(\ell)$ using theoretical PSFs generated with TINY TIM (Krist et al., 2011). The $B(\ell)$ is simply the FFT of these PSFs. The beam transfer functions measured in this way are consistent with those in MW15, and can be seen in the Supplementary Section of MW15. We chose to re-measure the $B(\ell)$'s using TINY TIM as a kind of sanity check that all was unchanged at the larger pixel scale used in this work. This larger pixel scale also made it more difficult to perform stacking, hence our use of the theoretical PSFs. For *Spitzer*/IRAC, we us the $B(\ell)$ that was measured in Cooray et al. (2012a), and simply linearly interpolate those values at our ℓ bins. Those points are plotted in Fig. 3.4

The point spread function (PSF) of *Chandra* varies across the mosaic as a function of off axis

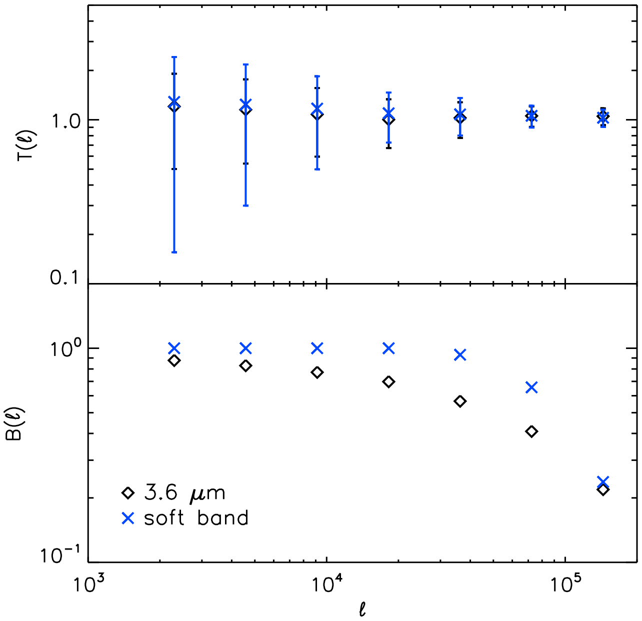

Figure 3.4 Map-making transfer functions (top) and beam transfer functions (bottom) for IRAC and the soft X-ray band. The different X-ray bands have slightly different effective exposure time maps, so the $T(\ell)$'s for each are unique but very similar. Beams for each X-ray band are roughly constant, with a 3 pixel-wide Gaussian PSF. Similar plots for the HST bands can be seen in Supplementary Figure 2 of MW15.

angle. However, since we are only concentrating on a relatively small area and convolving by a relatively large beam, we do not need to quantify and correct for this PSF variation. The 50% encircled energy average radius of the *Chandra* PSF increases from $\sim 0''$ at the center of the image to $\sim 3''$ at an off axis angle of $8'$ [2]. The maximum off axis position of a *Chandra* source in the CANDELS field is $\sim 8.5'$. Therefore we use the $3.6''$ Gaussian convolution kernel as an approximation to the *Chandra* beam. This approximation is further justified as we are not particularly interested in these small-scale fluctuations, which correspond to angular scales at $\ell > 3 \times 10^5$. The X-ray $B(\ell)$ is then computed with a Gaussian of 3.6 arcseconds, which corresponds to the three-pixel Gaussian kernel that was used to convolve the X-ray maps. This curve can also be seen in Fig.!3.4.

---

[2] `http://cxc.cfa.harvard.edu/proposer/POG/html/chap4.html#fg:hrma_ee_offaxis_hrci`

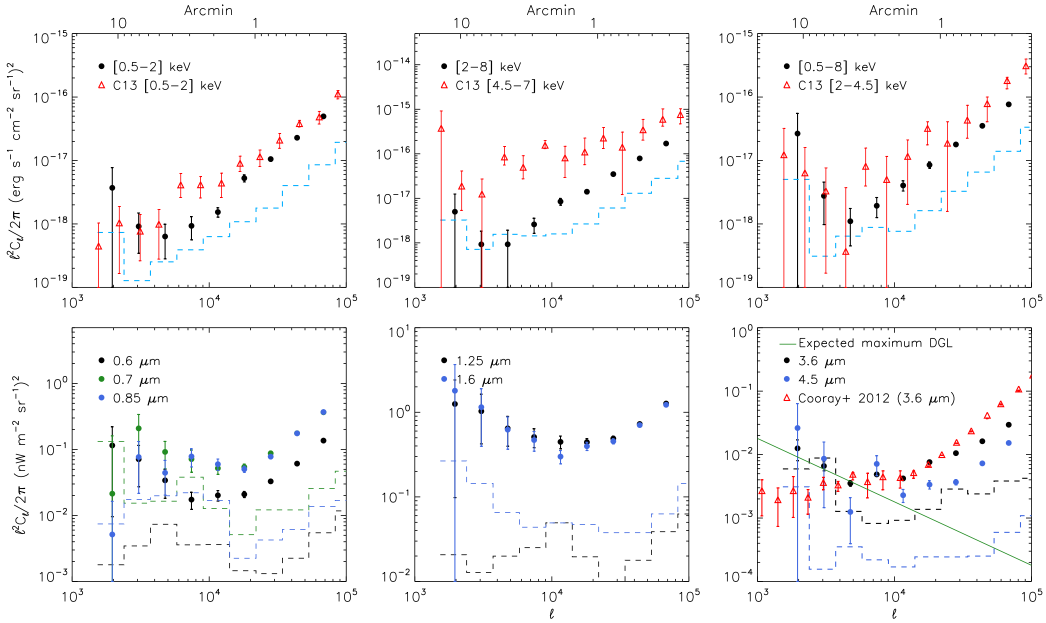

Figure 3.5 Auto power spectra in each band. Red triangles show measurements from previous works. Dashed lines show the noise power spectra (e.g. the auto spectra of (A-B)/2). The green line in the final plot shows the upper limit of the DGL contribution to the 3.6 μm auto-spectrum, modeled as $C_\ell^{\mathrm{DGL}} = A_{\mathrm{DGL}}^{3.6}\, \ell^{-3}$.

3.3.5 Auto Spectra

For jack-knife maps A and B (see MW15 for details of jack-knife tests), we compute the auto-power spectrum as the cross-spectrum of $A \times B$, and quantify the noise-power spectrum as the auto spectrum of $(A - B)/2$. Taking the cross-spectrum will omit any uncorrelated noise. Cosmic variance is defined as

$$\delta C_\ell = \sqrt{\frac{2}{f_{\mathrm{sky}}(2\ell+1)\Delta\ell}}(C_\ell^{\mathrm{auto}} + N_\ell), \tag{3.2}$$

where N_ℓ is the noise-power spectrum and f_{sky} is the fraction of the total sky unmasked in our maps. We compute our auto-spectra error bars by adding in quadrature the errors from the $T(\ell)$ and cosmic variance. The auto spectra for each of our 9 bands, including noise spectra, are plotted in Fig 3.5.

C13 performed auto-correlation measurements between similar X-ray bands, and *Spitzer*/IRAC Channels 1 and 2. We compare our X-ray auto-spectra to theirs, and find a statistically significant disparity at angular scales of $\ell \sim 10^4$. The C13 auto-spectra have a bump at $\ell \sim 10^4$, where ours are dominated by the shot noise from unresolved AGNs at this scale. This discrepancy can likely be attributed to the difference in the two X-ray datasets. The CDF-S data used in this work are at a minimum 4 times deeper than the EGS/AEGIS data used in C13 (Goulding et al., 2012). The broad X-ray band passes chosen in C13 are also slightly different than those that are used in this work.

3.3.6 Cross-correlations: Measurements

Cross-power spectra measurements are made by computing the cross-spectrum of $(A_m + B_m)/2 \times (A_n + B_n)/2$ for filters m and n. The errors for a cross-correlation are computed as

$$\delta C_\ell = \sqrt{\frac{1}{f_{\rm sky}(2\ell+1)\Delta\ell}\left[(C_{\ell,m}^{\rm auto} + N_{\ell,m})(C_{\ell,n}^{\rm auto} + N_{\ell,n}) + \left(C_\ell^{m\times n}\right)^2\right]},$$

where $C_\ell^{m\times n}$ is the cross-correlation power spectrum between bands m and n. Our total error budget is then the quadratic sum of the δC_ℓ's and the geometric mean of the transfer function between the two bands, e.g. $T_{\ell,mn} = \sqrt{T_{\ell,m}\,T_{\ell,n}}$. The beam transfer functions for the cross spectra are approximated in a similar way as $T_{\ell,mn}$.

Since the measurements in C13 were presented with none of the corrections we consider here, we first compare our raw CIB × CXB cross-correlations with IRAC 3.6 and 4.5 μm with theirs in Fig 3.6. Our corrected CIB × CXB cross-correlations can be seen in Figs 3.7. To be consistent with C13, we define the significance of the cross-correlation detections by concentrating only on angular scales larger than $20''$ ($\ell \lesssim 6.5 \times 10^4$), however we compute these values from our corrected cross spectra, not the raw spectra. The maximum angular scale probed in our study is about $10'$ and the significance is computed between $1.1 \times 10^3 \lesssim \ell \lesssim 6.5 \times 10^4$. We compute p-values from χ^2 estimates of all the C_ℓ's within that range of angular scales. Our measured cross-correlations are generally in good agreement with those measured in C13. In that work they claimed a high significance with their soft band cross-correlations (3.8σ at 3.6 μm and 5.6σ for 4.5 μm), and lower significances in their cross-correlations with harder X-ray bands. We detect a similar result, with 3.7σ and 2.7σ significances for the soft and hard bands respectively at 3.6 μm, and 4.2σ and 3.7σ respectively at 4.5 μm(see Table 3.3). In C13 they had a total of $\sim 1.3 \times 10^5$ photons, where in our work we have $\sim 8 \times 10^4$ (see Table 3.2), which corresponds to a signal-to-noise ratio of about 1.3 times lower. The 3-4σ significances that we measure here is more or less what one expects if we are seeing the same population as C13.

All 15 cross-correlations of the three X-ray bands with the five *Hubble* bands can be seen in Figure 3.8. These are computed in the same manner as discussed above for *Spitzer* fluctuations. The statistical significances for the five *Hubble* bands (as well as those from the 3.6 and 4.5 μm correlations) are shown in the top-right corners of their respective plots, and listed in Table 3.3. Interestingly, we find the HST cross-correlations are largely anticorrelated, with the exception of two products at 0.775 μm, with anticorrelation significances ranging from 1.4 - 3.5σ. The two positive correlations with 0.775 μm have significances of 2.5 and 2.8σ. Statistically, all these (anti) cross-correlations are less significant, but we note that 10 out of the 15 cross-correlations below 1.6 μm are negative. The presence of anticorrelations suggest that the fluctuation regions that are bright in the optical are faint in the X-rays, for example. One possibility for this is the Galactic absorption of soft X-rays, as discussed in Wang & Yu (1995) using ROSAT and IRAS, and (Snowden et al., 2000). While the extragalactic soft X-ray background is absorbed by the Galactic clouds, the same clouds (and associated cirrus) contributes positively to infrared fluctuations through DGL, Galactic dust-scattered interstellar light. While our detections of the anti cross-correlations are statistically less significant, the general behavior is consistent with suggestions in the

	0.6 μm	0.7 μm	0.85 μm	1.25 μm	1.60 μm	3.6 μm	4.5 μm
soft	2.2*	2.8	2.4*	2.2*	2.6*	3.7	4.2
hard	1.4*	2.8*	2.6*	2.6*	3.5*	2.7	3.7
full	1.4*	2.5	3.3*	1.7*	3.4*	3.2	4.1

Table 3.3 Cross-correlation statistical significances (σ) for $\ell \gtrsim 20''$. Values with an asterisk denote anticorrelations.

literature that optical and IR background fluctuations can be impacted by Galactic signals such as DGL (MW15, Yue et al. 2016).

In addition to these cross-correlations, we show the correlation coefficient for each of the 21 cross-correlations. For bands n and m, this coefficient is defined as $C_\ell^{n\times m}/\sqrt{C_{\ell,n}^{\rm auto}\,C_{\ell,m}^{\rm auto}}$. These are shown in Fig 3.9. The correlation coefficient can be interpreted as the fraction of total emission that is common between the two populations at filters m and n, which means the physical source of emission is the same between the two bands, or that separate emitters at band m and n (e.g. IR and X-ray) are spatially separated by an angle less than the beam size.

C13 and Helgason et al. (2014) used their measured values of the correlation coefficient at large angular scales of the soft X-ray band and the 4.5 μm band to infer a lower limit fraction of BH emission in the CIB at that 4.5 μm band. Their values are quoted without errors that $\simeq$ 15%-25% of the emission in that IR channel is a product of BH emission. As can be seen in Fig 3.9, the correlation coefficients we measure come with sufficiently large errors that we will not make any definitive statements about what fraction of the CIB emission comes from low-z or high-z sources and/or BHs. For example, the correlation coefficient we measure at large angular scales for the soft band and 4.5 μm ranges from -0.56 ± 1.55 to 2.62 ± 6.86 with an average value of 0.32 ± 0.66.

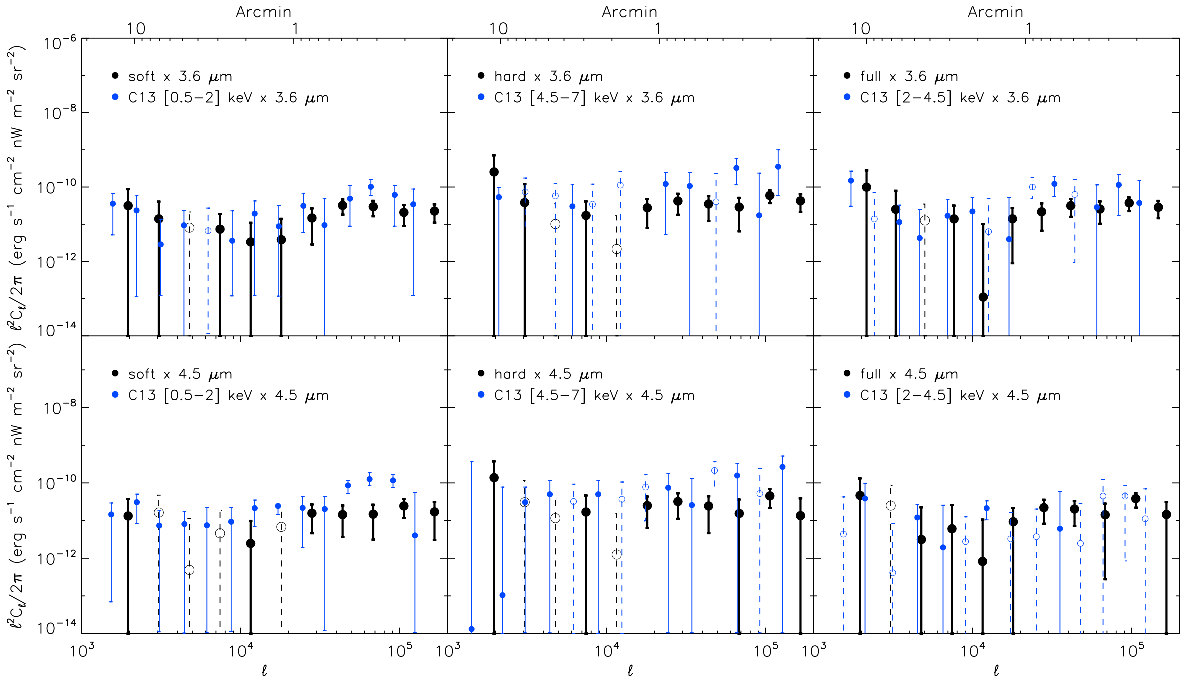

Figure 3.6 Raw cross-correlations of X-ray and IRAC 3.6 μm (top) and 4.5 μm (bottom). Open circles with dashed error bars denote negative values. Blue points are taken from Fig. 5 and 6 of C13. The spectra published in C13 are raw, so for consistency, here we compare our uncorrected spectra with theirs. There is good agreement between the two works.

3.3.7 Cross-correlations: Model Predictions

We show in Fig. 3.8 the expected contribution of unresolved AGN and galaxies (from their X-ray binary populations) based on the population models of Helgason et al. (2014). In this study, source emission was modeled separately in the near-IR and X-rays for both galaxies and AGN at $z < 6$, to derive the unresolved background for a common NIR/X-ray masking threshold. The angular cross power spectra were calculated using a halo model and validated by an independent semi-analytical model mapped onto N-body simulations (Henriques et al., 2012). For this paper, we extend these models to include optical wavelengths as well as hard X-ray energies using the same methodology as described in Helgason et al. (2014).

The models are displayed as solid lines alongside the measurements in Figures 3.7 and 3.8. Whereas the galaxy contribution shows some increased power from clustering toward large

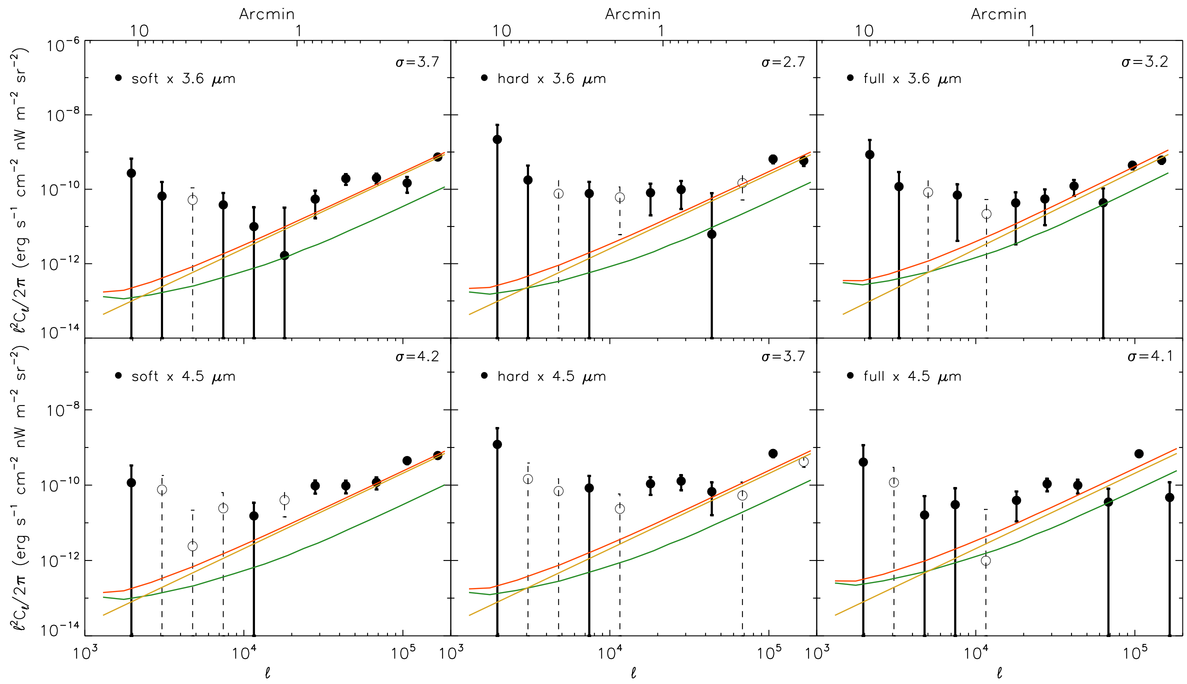

Figure 3.7 Corrected cross-correlations of X-ray and IRAC 3.6 μm (top) and 4.5 μm (bottom). Open circles with dashed error bars denote negative values. Blue points are taken from Fig. 5 and 6 of C13. The statistical significance of each cross-correlation, computed at angular scales $\gtrsim 20''$, is shown in the top-right of each plot. Theoretical models are shown as solid lines: the AGN component as a yellow line, unresolved galaxies as a green line, with a slight excess at low-ℓ, and the total contribution as a red line.

angular scales, the AGN contribution is almost purely shot noise dominated. This is because many AGN that are relatively X-ray bright will still remain unresolved if they are optically dim. However, the two contributions are similar in amplitude because the abundance of the two source classes is comparable at the depth of the ECDFS. The models are somewhat sensitive to the extent to which the IR mask removes sources that remain undetected in X-rays.

As shown in Figs 6. and 7 existing models for the X-ray background, with simple populations involving faint galaxies and AGNs, are adequate to explain the measured cross-correlations with *Spitzer*. These models, however, do not predict the anti-correlation behavior seen at wavelengths below 1.6 μm since they do not involve Galactic absorption of X-rays. We do

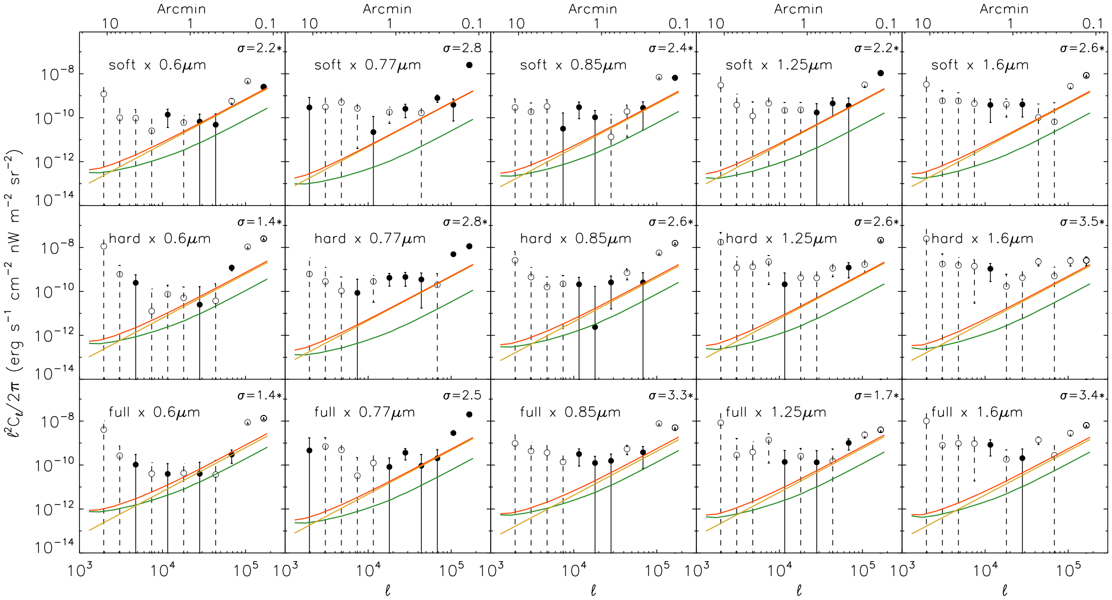

Figure 3.8 Cross-correlations of X-ray and optical/NIR HST maps. Open circles with dashed error bars denote negative correlations. The color schemes for the theoretical models curves are the same as in Fig. 3.7. The significance of each cross-correlations at angular scales $> 20''$ is shown in the top right of each plot. Sigma values with an asterisk denote an overall anticorrelation.

not pursue such detailed modeling here as we do not have adequate statistics to constrain additional model parameters. In the future with wide area surveys both in the X-rays (e.g., with Athena) and optical/infrared (such as with Euclid and WFIRST) the study can be expanded to 10-100 deg^2 areas for sufficient statistics for studies we have alluded here. It should also be noted that the evolution of X-ray binaries may be complex (Lehmer et al., 2014), which is not considered in the modeling.

3.4 Null Tests

In this section we briefly address a kind of "safety test", in that we cross-correlate NIR/optical null maps with X-ray signal maps and vice versa. For jack-knfe maps A_1 and B_1 for band 1, and A_2 and B_2 for band 2, we recompute the cross-correlations twice for each combination of

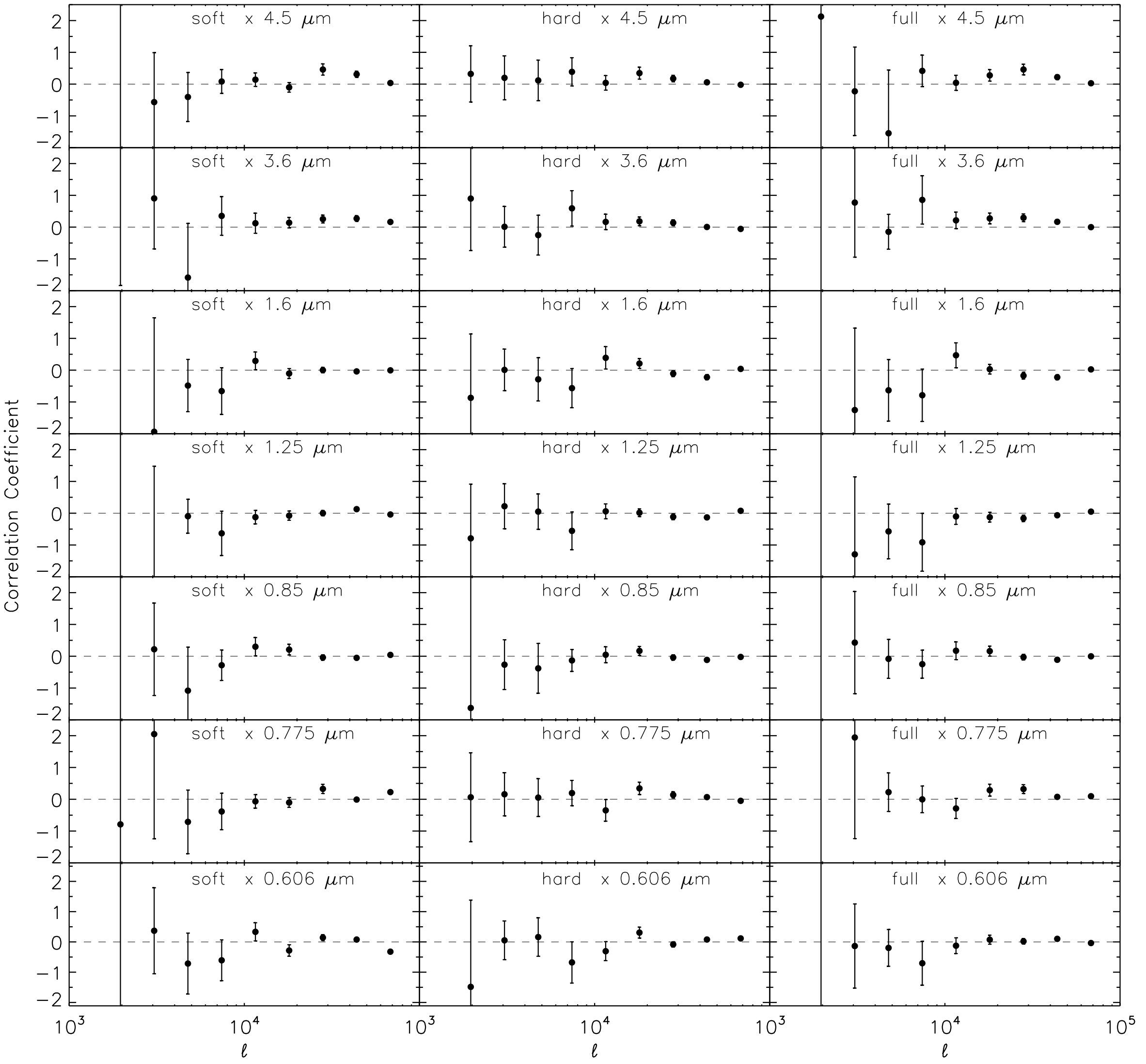

Figure 3.9 Correlation coefficients of X-ray and optical/NIR HST cross-correlations. All the points are statistically between -1 and 1.

null and sky map signals: $(A1-B1)/2 \times (A2+B2)/2$ and $(A1+B1)/2 \times (A2-B2)/2$. We do this for each cross-correlation considered in this work, shown in Fig 3.10. We show in that plot the null tests, and for comparison the sky map cross-correlations presented throughout this work. The sky map raw cross spectra are generally an order of magnitude or more above the null test cross-correlations, not dissimilar to what was found in C13. This is a good test that spurious instrumental features are not dominating our sky map cross-correlations. Note that these null tests are encapsulated in our error bars throughout this work, as they do have contributions from the noise power spectra, discussed above.

3.5 Discussion

A previous study focusing on the *Spitzer*/Extended Groth Strip (EGS; Davis et al. 2007) found $\sim 3.8\sigma$ and 5.6σ cross-correlation detections at large angular scales between unresolved fluctuations at *Spitzer*/IRAC 3.6 and 4.5 μm respectively, and the *Chandra* soft X-ray band of 0.5 to 2 keV (C13). This cross-correlation has been explained as due to a population of primordial DCBHs at $z > 12$ (Yue et al., 2013b). If DCBHs explain all of the IRAC fluctuations at 3.6 μm, due to the Lyman-break cut-off redshifted to infrared wavelengths today, we do not expect to see any fluctuations generated by DCBHs at wavelengths less than 1.6 μm. Despite this general expectation, the recent study by MW15 found a significant detection of the intensity fluctuations from optical to infrared bands (0.6 to 1.6 μm) with *Hubble*/ACS and WFC3 imaging data in the GOODS and CANDELS surveys. These fluctuations, as well as previous detections of fluctuations at 3.6 μm, have been explained as mostly due to IHL (Cooray et al. 2012), with some evidence for fluctuations generated by galaxies present during reionization (MW15). Since IHL is the signal from tidally stripped stars, it will not have any X-ray component directly associated with it.

In the present study we return to the claimed detection between X-rays and *Spitzer*/IRAC.

We extend the analysis to consider multi-wavelength optical and infrared data between 0.6 and 1.6 μm, in addition to 3.6 and 4.5 μm data in the CDF-S. Here we have complied and produced sky maps in the CDF-S in 10 passbands, from 3 different space telescopes, corresponding to 3 bands at X-ray wavelengths and 7 bands between 0.6 to 4.5 μm with both *Hubble* and *Spitzer*. After masking the resolved sources to isolate the background light with a common mask, which itself involves the product of individual source-detection masks in each of the wavelengths, we performed cross-correlations of the X-ray maps with the optical and NIR data.

The X-ray maps are sensitive to emission from black hole accretion and hot ionized gas, such as that found in galaxy clusters. The optical and infrared maps trace the faint galaxies throughout the cosmic history. The redshift dependence of the galaxy poupulation is captured by the Lyman-break signal that is moving across the bands from 0.5 to 1.6 μm as a function of the redshift. The optical and infrared fluctuations also contain a signature of DGL, arising from dust-scattered light in our Galaxy. We found a 3.7-4.2σ correlation between the soft X-ray band with both 3.6 and 4.5 μm, confirming the general result of C13. As the measurement is in an independent field, and utilizes independent analysis techniques and methods, we can generally confirm that there is indeed a significant cross-correlation between unresolved infrared fluctuations at 3.6, 4.5 μm and soft X-rays. Similar to C13, we also find a lower statistical significance in cross-correlation between the hard X-ray bands and IRAC.

Extending the analysis in C13, we also present results related to 0.6 and 1.6 μm fluctuations and find that neither soft nor hard X-ray bands are correlated with the optical or infrared fluctuations at a statistically significant level ($\gtrsim 3.5\sigma$). Those cross-correlations are mostly anti-correlated. Wang & Yu (1995) and Snowden et al. (2000) found shadows in the CXB introduced by the galaxy. We can then attribute, at least in part, these anti-correlations between HST and *Chandra* to X-ray absorption by dust in the Milky Way (diffuse Galactic light; DGL), which is present in the NIR HST maps. However if this is the mechanism responsible for our measurements, it would also affect the the cross-correlation between the soft band and 3.6-4.5 μm, which would imply our measured cross-correlation significances at those wavelengths are lower limits. Furthermore it is unlikely this interpretation can be extended to the hard X-ray band since the galaxy is largely transparent at those wavelengths. The Galactic signal is present between 1.6 and 3.6 μm, complicating an easy interpretation of the auto and cross power spectra. Due to the lack self-calibrated maps at wavelengts between 1.6 and 3.6 μm, we do not have a complete understanding of the exact spectral energy distribution of the component that is correlated with the X-ray background.

Giallongo et al. (2015) claimed a population of AGNs at $z > 4$ could be part of the population responsible for reionization. From their work, 7 of those sources are outside our cropped field of view and the remaining 15 sources are behind our source mask. So those purported signals do not contribute to the measurements reported here. Treister et al. (2013), Weigel et al. (2015) and Cappelluti et al. (2016) all find weak or no evidence of a significant population of AGNs at $z \gtrsim 4-5$, which would be consistent with our results assuming a high-redshift population of X-ray sources exist. In that case, the X-ray cross-correlation below 1.6 μm should disappear, due to the Lyman-break cut-off. In addition, (Kashlinsky et al., 2007) found no correlation with IRAC and faint ACS galaxies, which also aligns with our measurements. If there was a significant cross-correlation between those bands, it would follow that they would both be correlated, to some degree, with the same X-ray maps. The cross-correlation between the CIBER wavelengths and *Spitzer*/IRAC (Zemcov et al., 2014) can be wholly attributed to DGL Yue et al. (2016) and thus does not add contradiction to this interpretation or previous findings.

The signal-to-noise of our cross-correlation measurements are ultimately limited by the number of X-ray photons, so larger area fields will improve the cross-correlation statistics. However large area, self-calibrated HST maps, generated from e.g. the 1.7 deg^2 F814W COSMOS observations (Koekemoer et al. 2007; Scoville et al. 2007), have proven difficult if not impossible to generate with the current number of exposures. Although the weak correlation coefficient between HST and *Spitzer*/IRAC (compounded with a weak X-ray/HST cross-correlation) does suggest the background sources of *Spitzer*/IRAC and HST are dissimilar, it is plausible that our HST/X-ray cross-correlations are just not sensitive to a signal that is actually there, because of the limited number of X-ray photons. It is also plausible that the X-ray background sources reside at $z > 13$, which the longest HST band (1.6 μm) would not be sensitive to.

Our result is consistent with the existing framework related to our understanding of the nature and origin of optical and infrared intensity fluctuations, but with some modifications. Part of the component that is correlated between X-rays and 3.6, 4.5 μm could be in the form of DCBHs or some other source population. For example, apart from DCBHs, the soft X-rays can also come from hot gas associated with either star-forming regions or halos around galaxies, both of which may be expected to correlate with 3.6 and 4.5 μm. Due to the presence of optical fluctuations, we also require a signal from low redshifts and, in current models, such an origin involves IHL or tidally-stripped diffuse stars that populate the extended dark matter halos of galaxies. A combination of this IHL component and the component that is correlated with X-rays, such as DCBHs, likely contribute to fluctuations at 3.6 and 4.5 μm. Despite significant progress and overall improvement in the statistical accuracy in recent years, the cross-correlation between *Spitzer*/IRAC and X-rays is at the level of 4σ significance and we are not able to model-fit to establish accurately the relative fraction between IHL and DCBH amplitudes at 3.6 and 4.5 μm. The lack of a significant cross-correlation between 1.6 and 3.6 μm could be interpreted as DCBHs primarily dominating the fluctuations at 3.6 and 4.5 μm, while IHL is dominating the fluctuations at 1.6 μm. Given the overall statistical uncertainties in the cross-correlation between 1.6 and 3.6 μm, even a scenario in which 100% of the fluctuations are from IHL in 3.6 and 4.5 μm cannot be ruled out by the present measurements. To separate IHL from DCBHs, we need to improve not only the sensitivity to large-scale fluctuations in infrared and X-ray data, but also improve methods to separate reliably the DGL signal in the fluctuation power spectra. Naturally, additional deeper and wider fields in X-rays, and in the optical between 0.5 and 4 μm, will be one way to improve the current situation. In the future, appropriate data will likely come from survey telescopes such as *Athena* for X-rays and Euclid/WFIRST for optical and infrared.

3.6 Summary

In this paper we have measured the cross-correlations between the CXB and CIB and found a correlation between the soft X-ray background and the *Spitzer*/IRAC 3.6 and 4.5 μm backgrounds only, which is consistent with the previous findings of C13. Extending the results of C13, we find X-ray cross-correlations between the shorter wavelengths, ranging from 1.6 μm down to 0.6 μm, are statistically insignificant or anticorrelated. All correlations with the hard X-ray band were also found to be less significant or anticorrelated. This result implies that a significant portion of the CIB at 3.6 and 4.5 μm could be a product of DCBHs; however it does not simultaneously rule out a partial or full contribution from IHL in the 3.6 and 4.5 μm intensity fluctuations. Furthermore if the DCBHs do reside at $z > 13$, we would expect no correlations in bands shortward of 1.6 μm with X-ray, which is consistent with our measurements. This lack of correlations in the shorter bands may be evidence that the sources responsible for the 3.6 and 4.5 μm backgrounds are fundamentally different from those at 1.6 μm and below.

Given the overall low significance of the detections in the near-IR and lack of detections in the optical, it is hard to draw strong conclusions from this study. We recommend further studies with wider X-ray fields of areas greater than 4 sq. degrees, since these studies benefit from the wide area coverage probing the linear clustering and not the narrow, deep fields more sensitive to the shot-noise. We recommend Chandra conduct a sufficient survey for the purposes of a cross-correlation study in a field overlapping with sufficient ancillary data in the near-infrared.

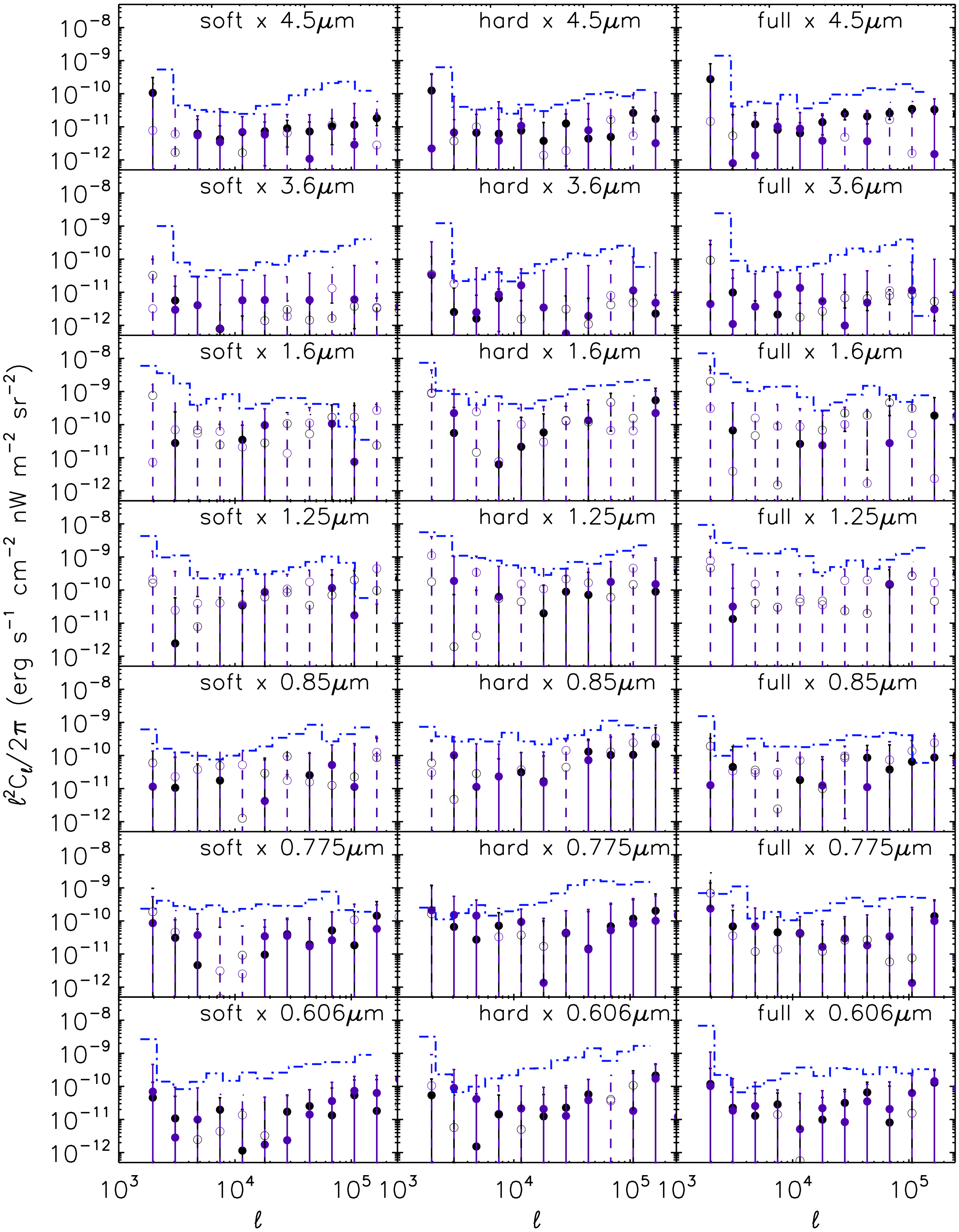

Figure 3.10 Null and signal map cross-correlations of X-ray and optical/NIR maps. Open circles with dashed error bars denote negative correlations. The black points correspond to the null (A-B)/2 NIR or optical map cross-correlated with the X-ray signal map. The purple points correspond to the null X-ray map cross-correlated with the signal (A+B)/2 NIR or optical map. The dot-dashed blue line is the absolute value of the signal-signal cross-correlations presented throughout this work.

Chapter 4

Herschel-SPIRE N(z)

4.1 Introduction

For this study we make use of data from the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012), which mapped a large number of well-known fields with existing multi-wavelength ancillary data using SPIRE. To cross-correlate against SPIRE-selected galaxies, we make use of near-IR selected galaxy samples from *Spitzer* observations based on the 1.6 μm "bump", which has long been established as a redshift indicator (Sawicki 2002; Simpson & Eisenhardt 1999; Wright & Fazio 1994). The bump results from the fact that the H^- absorption in stellar atmospheres is minimally opaque at 1.6 μm. This leads to a bump in the spectral energy distributions of cool stars at 1.6 μm (John, 1988) that is nearly ubiquitous in galaxy spectra. For $z > 0$, the wavelength at which the bump in the SED peaks allows for a redshift determination based on the colors in IRAC channels between 3.6 and 8 μm, and covering the redshift range of 1 to 2.5. We complement these "bump" galaxy samples with a 24 μm and an R-band based sample of dust obscured galaxies, which has a redshift distribution that peaks around $z \sim 2.3$ (Dey et al. 2008). We also make use of optical-selected galaxy samples with SDSS spectroscopic and photometric redshifts out to about 0.7.

This Chapter is organized as follows. In Section 2 we describe source selection in SPIRE and galaxy sample selection with IRAC and MIPS bands, complemented with optical data to remove outliers. In Section 3 we describe the redshift distribution of the galaxy samples used for the cross-correlation analysis, and in Section 4 we describe the cross-clustering measurements. Fitting results are presented in Section 5. In Section 6 we present $N(z)$ and $b(z)$, and discuss our results. We assume a flat-ΛCDM cosmological model and fix the cosmological parameters to the best-fit values of $\Omega_{\rm m} = 0.27$, $\Omega_b = 0.046$, $\sigma_8 = 0.81$, $n_s = 0.96$ and $h = 0.71$ (Komatsu et al., 2011) when performing MCMC model fits.

4.2 SPIRE Source and Galaxy Sample Selection

4.2.1 *Herschel*-SPIRE sample

The HerMES SPIRE source catalogs used for this paper come from a combined analysis involving both a direct source extraction and an attempt to account for blending at 350 and 500 μm wavelengths given the positions of 250 μm detections (Wang et al. in prep). The method updates the source extraction pursued by HerMES at each of the three SPIRE bands independently that ignored issues associated with blending at longer wavelengths (Smith et al. 2011).

In order to maximize the overlap with multi-wavelength data, we concentrate our study on the Boötes field with HerMES SPIRE data covering 12.5 deg^2. The field has been imaged with *Spitzer* IRAC as part of the *Spitzer* Deep Wide Field Survey (SDWFS; Ashby et al. 2009) and from the ground with optical to near-IR observations as part of the NOAO Deep Wide

Field Survey (NDWFS; Jannuzi and Dey 1999), with coverage also provided by the Sloan Digital Sky Survey (SDSS; Abazajian et al. 2009).

For this study we selected SPIRE sources with a flux density greater than 20 mJy in the Böotes field. We find in excess of 22 000 sources in each of the SPIRE bands covering the entire 12.5 deg^2 of SPIRE observations, while 3775, 3243 and 958 galaxies at 250 μm, 350 μm and 500 μm, respectively, were used in the cross-correlation study – an area covering 6.7 deg^2, where various ancillary data best overlap.

At 20 mJy, the SPIRE catalogs are $\sim$ 30% complete at each of the three wavelengths. The 90% completeness of the catalogs is at a flux density of about 55 mJy (Wang et al. in prep). At such a high flux density level, the number of SPIRE sources in the area overlapping with ancillary data is down by at least factor of 8 and the resulting low surface density does not allow useful constraints on the redshift distribution. We note some caution in interpreting our results with models due to the incompleteness. We are unable to correct for it through simulations due to the lack of a priori information on the redshift distribution of missing sources. It is unlikely, however, that the redshift distributions presented here are biased due to catalog incompleteness since the source detection algorithm is primarily based on the flux density and not the individual redshifts of SPIRE sources.

4.2.2 IRAC Sample Selection and Star-Galaxy Separation

Using the SDWFS data combined with ground-based *K*-band data from NDWFS, we generated three different catalogs of 1.6 μm-bump sources based on the IRAC channel where the SED peaks. These three samples are as follows: bump-1 with a peak in the 3.6 μm channel ($0.5 \lesssim z \lesssim 1.5$); bump-2 peaking in the 4.5 μm channel ($0.8 \lesssim z \lesssim 2.2$); and bump-3 with a peak at 5.8 μm ($1.5 \lesssim z \lesssim 3.0$). Using the photometric redshifts computed via a template fitting method (Csabai et al. 2003) in the SDSS DR7 catalog, we also constructed two separate redshift distributions with peaks at $z \sim 0.3$ and 0.7.

In order to establish catalogs of bump-1 to bump-3 galaxy populations we first had to remove stars and other contaminants from our optical and IR catalogs. This was done using a combination of infrared and optical data. We used the SDWFS four-epoch stacked catalog (Ashby et al. 2009) which contains all sources detected in the first channel of IRAC at or above 5σ. This catalog was matched with the NDWFS third data release catalog and the SDSS catalog, using a $2.5''$ matching radius. For sources with multiple matches (<3%), magnitudes from the NDWFS catalog were then compared with the $3.6\,\mu$m magnitude and entries with the most similar values were kept.

Stars and spurious sources were removed from the resulting merged catalog using various techniques. Vega magnitudes and $6''$ diameter aperture photometry are used throughout the star-galaxy separation unless otherwise noted. We employed a three stage process to remove stars from our catalog. An initial selection of sources with $[3.6] < 16$ were identified as stars, where [3.6] represents the vega magnitude at 3.6 μm. Using the combination of optical and IRAC photometry we further classify sources as stars that either satisfy $(B_{\rm w} - I) > 2(I - [3.6]) - 1.65$, or $-1.65 > (B_{\rm w} - I) - 2(I - [3.6]) > -3.35$ (Eisenhardt et al., 2004). The former criterion defines a sequence of BIK stars (Huang et al., 1997), and the latter a sequence of giant stars (Johnson, 1966). Lastly, for IRAC sources without optical counterparts, we used a binning method that involved only the IRAC bands (Waddington et al., 2007). Three flux density bins were defined with the criteria $[3.6] \leq 19.5$, $19.5 < [4.5] \leq 20.0$ and $[4.5] \leq 23$, with color cuts $[3.6]$ - $[4.5] < -0.35$, -0.30 and -0.25, respectively; all sources satisfying these criteria were assumed to be stars. The results of these extractions are shown in Fig. 4.1.

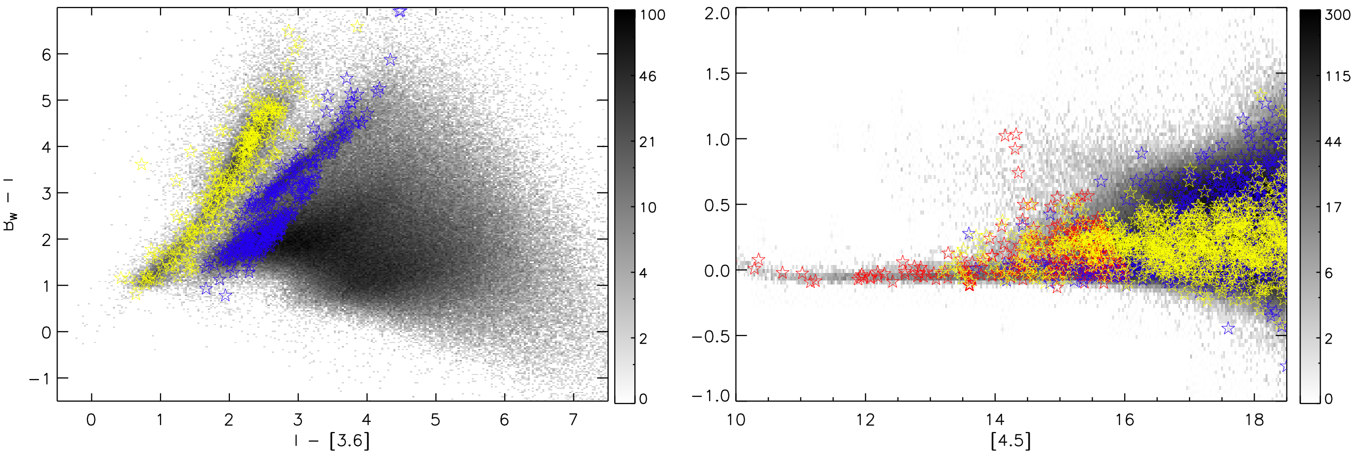

Figure 4.1 Color-color (top; $B_\mathrm{w} - I$ vs. $I - [3.6]$) and color-magnitude (bottom; $[3.6] - [4.5]$ vs. [4.5]) density plots of sources (both stars and galaxies) in our Boötes field catalogs. Red stars indicate the 3.6 μm magnitude selection for stars, while giants are depicted as yellow stars, and the *BIK* sequence as blue stars (see Section 2.2 for details). Only a fraction of the sources identified as stars are plotted to avoid over-crowding in the plots. Black points are the galaxies.

4.2.3 Bump, DOG and SDSS Selections

From our resulting merged and star-subtracted catalog, we invoked simple color constraints to classify three different types of bump sources and dust-obscured galaxies (DOGs; Dey et al. 2008), using 4” aperture diameter photometry. Bump-1, bump-2 and bump-3 sources each display excess emission in IRAC channels 1, 2 and 3, respectively. Bump-1 sources were selected using the criteria $K - [3.6] > 0.1$ and $[3.6] - [4.5] < 0$; bump-2 with $K - [3.6] > 0$, $[3.6] - [4.5] > 0$ and $[4.5] - [5.8] < 0$; bump-3 with $[3.6] - [4.5] > 0$, $[4.5] - [5.8] > 0$ and $[5.8] - [8] < 0$. The number of bump-1 sources in the SDWFS catalogs were found to be $\sim 1.3 \times 10^4$ at the 5σ detection limit of the 3.6 μm channel of the IRAC instrument. Bump-2 source identification yielded 6.5×10^3 galaxies, while the bump-3 catalog contains 4×10^3 galaxies.

We also make use of a sample of dust obscured galaxies, selected with 24 μm *Spitzer*-MIPS and optical *R*-band data to have extreme red colors from dust obstruction, with $S_{24}/S_R >$

1000 (where S_{24} is the 24 μm flux density), or equivalently $R - [24] \geq 14$ and $S_{24} \geq 0.3$ mJy ($\approx 6\sigma$; Dey 2009). We found that a total of 2838 galaxies satisfied the selection criteria. Based on spectroscopic follow-up, they are now known to have a mean redshift around $z \sim 2.3$ (Dey & Ndwfs/MIPS Collaboration, 2009). We make use of the full, broad redshift distribution for this sample, spanning the range of $0.5 < z < 3.5$, with a peak around $z \sim 2$, found from a similar identification of DOGs in the COSMOS field (see Section 3 and Fig. 2) for the present analysis. These dust-obscured galaxies have been suggested to be an intermediary phase of the evolution of quasi-stellar objects from gas-rich mergers (Dey & Ndwfs/MIPS Collaboration, 2009). They have also been shown to be strongly clustered and are believed to be progenitors of massive ($3 - 6L_*$) galaxies at low redshift (Brodwin et al., 2008).

Finally, to cover the redshift range of $0 < z < 0.7$ efficiently we also selected optical galaxies from SDSS. These sources have photometric redshifts, individually determined with SED fits to SDSS photometry, in the above range. We make use of 8 000 SDSS galaxies and we consider two sub-samples peaking at $z \sim 0.2$ and 0.5 with roughly equal numbers. The first of these two sub-samples is obtained by selecting sources which obey $2.6 < B - I < 3$ and $-0.8 < I - R < 0.1$, while the second selection obeys $B - I > 4$ and $-0.9 < I - R < 0$. These six galaxy samples (3 bump catalogs, DOGs, and two SDSS samples) provide adequate redshift coverage over the range of $0 < z < 3$.

4.3 COSMOS photo-z and spec-z

While we are able to generate large samples of galaxies to cross-correlate against the SPIRE catalogs of the Boötes field, the existing spectroscopic and photometric redshift information in the Boötes field is not adequate to establish the redshift distributions of the *Spitzer* galaxy samples. For that we turn to data in the Cosmological Evolution Survey (COSMOS; Scoville

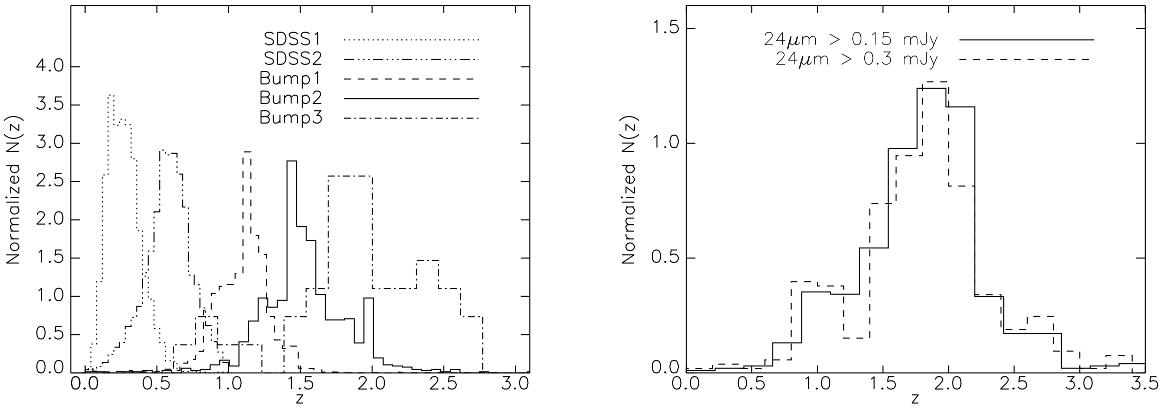

Figure 4.2 $N(z)$ distribution obtained from SDSS photometric redshifts in the Boötes field as well as bump-1 to -3 (top plot) and DOGs (bottom plot) redshift distributions from the COSMOS field. We assume the latter four redshift distributions measured directly in COSMOS are also applicable for galaxy samples under the same color selection criteria in the Boötes field.

et al. 2006; Capak et al. 2007) where we can select similar galaxy samples as in the Boötes field, using the same depth and color criteria. For those galaxies we are able to use either the existing spectroscopic or the photometric redshifts from the public COSMOS catalog (Ilbert et al., 2009). We assume that the redshift distributions for galaxy samples in COSMOS is the same as for Boötes when interpreting clustering measurements from the wider Boötes area that overlaps with SPIRE.

The COSMOS field has extensive photometric redshift measurements over 2 deg^2 using 30 broad, intermediate and narrow band filters from space-based telescopes (*Hubble, Spitzer, GALEX, XMM, Chandra*) and ground-based telescopes (Subaru, VLA, ESO-VLT, UKIRT, NOAO, CFHT, and others). We used a public COSMOS source catalog containing $\sim 10^4$ spectroscopic redshifts, and $\sim 3 \times 10^5$ photometric redshifts that were computed using a χ^2 galaxy template-fitting technique (Ilbert et al., 2009). In both the Böotes and COSMOS fields, we imposed the same brightness thresholds in the selection bands: for IRAC channel 1 and K-band we imposed a 5σ detection limit. We require that each of the galaxies detected at above 5σ is also detected in IRAC Channcel 2 to 4. In those cases, however, we considered a source that has a flux density above 50% completeness level to be considered as detected, while sources with flux densities below 50% completeness level of each of the three channels were dropped from the final catalog. This selection process was chosen to ensure that we are probing equal depths between the two fields, COSMOS and Böotes.

Using the same selection methods described above, we were able to obtain $N(z)$ measurements for each of our different source classifications from the COSMOS source catalog (see Fig. 4.2). Note that the galaxy type selections are mutually exclusive, so there are no overlapping sources between different samples. All of the bump sources have well defined redshift distributions, and the DOG distribution agrees well with those in the literature (see figure 1 in Dey 2009). We identified 384 sources as DOGs in COSMOS, with $S_{24} > 0.3\,\mu$Jy, and 683 with $S_{24} > 0.15\,\mu$Jy, a number density consistent with statistics of the DOG population in other fields (Brodwin et al., 2008).

4.4 Angular Cross-correlation and Covariance Matrix

To obtain the redshift distributions of SPIRE sources, we first cross-correlate the SDSS-selected sample and bumps and DOGs from the Boötes field, against sub-mm sources in each SPIRE band, from arcminute to degree angular scales. We also measure the auto-correlation functions of the galaxy and SPIRE samples, as these are needed to model the clustering strength and to extract the unknown redshift distribution.

We use a bootstrap method to establish the covariance matrix for each of the cross- or auto-correlation functions, as an accounting of the covariance is needed to properly model the clustering measurements. We do this by selecting 200 separate catalogs from the original SPIRE data by removing about 5% of the sources randomly. We measure the auto and cross-correlations with each of those catalogs and build the mean auto and cross-correlation functions, the variance from the scatter, and the covariance from the correlations between the measured auto and cross-correlation functions.

The angular cross-correlation function is modeled analytically using the COSMOS redshift distribution of the bump-1, bump-2, bump-3 and DOGs, while for the SDSS galaxy samples we make use of the public photometric redshifts from SDSS DR-7. For simplicity we bin the unknown SPIRE redshift distribution from $z = 0$ to 4 in 5 bins in redshift. To extract the best-fit values and uncertainties in the redshift distribution bins, and the other parameters in the analytical model, we make use of a likelihood fitting technique based on the Markov Chain Monte Carlo (MCMC) method.

In this section, we first discuss the method of modeling the angular cross-correlation $w_{\rm cross}$ using the redshift distribution of galaxies and the linear matter power spectrum. Then we describe the measurement of the $w_{\rm cross}$ functions as well as the covariance matrix from the galaxy samples.

4.4.1 Modeling the Angular Cross-Correlation

The angular cross-correlation function $w_{\rm cross}$ for two galaxy samples is defined by

$$w_{\rm cross}(\theta) = \langle \delta n_1(\hat{\phi}) \delta n_2(\hat{\phi}') \rangle, \tag{4.1}$$

where $\delta n_i(\hat{\phi}) = (n_i(\hat{\phi}) - \bar{n}_i)/\bar{n}_i$, and $n_i(\hat{\phi})$ is the number density of galaxies observed in direction $\hat{\phi}$ in the sky ($\theta = \phi - \phi'$), and $\bar{n}_i$ is the mean number density of the galaxy sample i. δn_i can be decomposed into two terms–one term from the real clustering of galaxies and a second term caused by lensing magnification. Here we ignore the few percent contribution from lensing (Wang et al., 2011) and only consider the clustering term, which is

$$
\begin{aligned}
w_{\rm gg}(\theta) = & \; b_1 \;\; b_2 \int_0^{\chi_{\rm H}} d\chi N_1(\chi) N_2(\chi) \\
\times & \int_0^{\infty} \frac{k}{2\pi} P(\chi, k) J_0[k r(\chi)\theta],
\end{aligned}
\tag{4.2}
$$

where b_1 and b_2, $N_1(\chi)$ and $N_2(\chi)$ are the galaxy bias and the normalized radial distribution for the two galaxy samples, respectively. $P(\chi, k)$ is the power spectrum of the dark matter, $J_0(x) = \sin(x)/x$ is the zero-order Bessel function, and χ and $r(\chi)$ are the radial comoving distance and the comoving angular diameter distance respectively ($r(\chi) = \chi$ in flat space). $\chi_{\rm H}$ denotes the radial distance to the horizon, or Hubble length. Note that $w_{\rm gg}(\theta)$ will be zero if the positions of the two galaxy samples do not overlap with eachother.

When modeling the measurements, we make use of the linear theory power spectrum to describe $P(\chi, k)$ and focus only on modeling the measurements over the angular scales of $6'$, and above where the clustering is in the linear regime (Cooray et al. 2000). At these large angular scales, the 1-halo term makes less than a 1% correction to the correlation function and can be safely ignored.

4.4.2 The Measurement and Covariance Matrix of the Cross-correlation

The angular cross-correlation function $w_{\rm cross}(\theta)$ is defined as the fractional excess of the probability relative to a random distribution (Peebles, 1980), and can be measured from galaxy samples by the *pair counts* method. There are several kinds of estimators that are proposed to measure the cross-correlation (e.g. Blake et al. 2006); the one we adopt here is the modified Landy-Szalay estimator which is derived from the auto-correlation (Landy &

Szalay, 1993),

$$w_{\rm cross}(\theta) = \frac{D_1D_2 - D_1R_2 - D_2R_1 + R_1R_2}{R_1R_2}, \tag{4.3}$$

where $D_1D_2(\theta)$, $D_1R_2(\theta)$, $D_2R_1(\theta)$ and $R_1R_2(\theta)$ are the normalized pair counts for data (D_i) and random (R_i) catalogs with separation θ.

We generate random un-clustered catalogs with varying catalog sizes that contain 5 to 10 times more sources than the observed samples, with a larger number of sources than in data catalogs to avoid biases coming from Poisson fluctuations. The angular cross-correlation extracted from the observational data, and the theoretical estimation using the best-fit value (see next section) of the SPIRE distribution in the Boötes field are shown in Fig.4.3. The auto- and cross-correlation of the SPIRE surveys for $250\,\mu$m, $350\,\mu$m and $500\,\mu$m are also shown in Fig. 4.4.

As was mentioned in the previous Section, to avoid biases coming from non-linear clustering we only use $w(\theta)$ data from 0.1 to 1° to fit the model, since adding the 1-halo term with three or four extra parameters for the halo occupation number will result in extra degeneracies, degrading the $N(z)$ estimates, consistent with theoretical suggestions in the literature (e.g., Neyrinck et al. 2006). Also keeping to scales larger than 0.1°, we avoid the need to introduce a transfer function for $w(\theta)$ for SPIRE sources and their cross-correlations since at smallest scales close to the SPIRE point response function, clustering is expected to be affected by source blending and issues related to map making. As studied in Cooray et al. (2010), at $\theta > 0.05°$, there are no corrections to the measured $w(\theta)$.

To evaluate the covariance matrix of the angular correlation $w(\theta)$, we use a bootstrap method to generate 200 realizations for the galaxy samples. Then the covariance matrix of $w_{\rm cross}$ is

$$C_{ij} = \frac{1}{N-1}\sum_k^N [w_k(\theta_i) - \bar{w}(\theta_i)][w_k(\theta_j) - \bar{w}(\theta_j)], \tag{4.4}$$

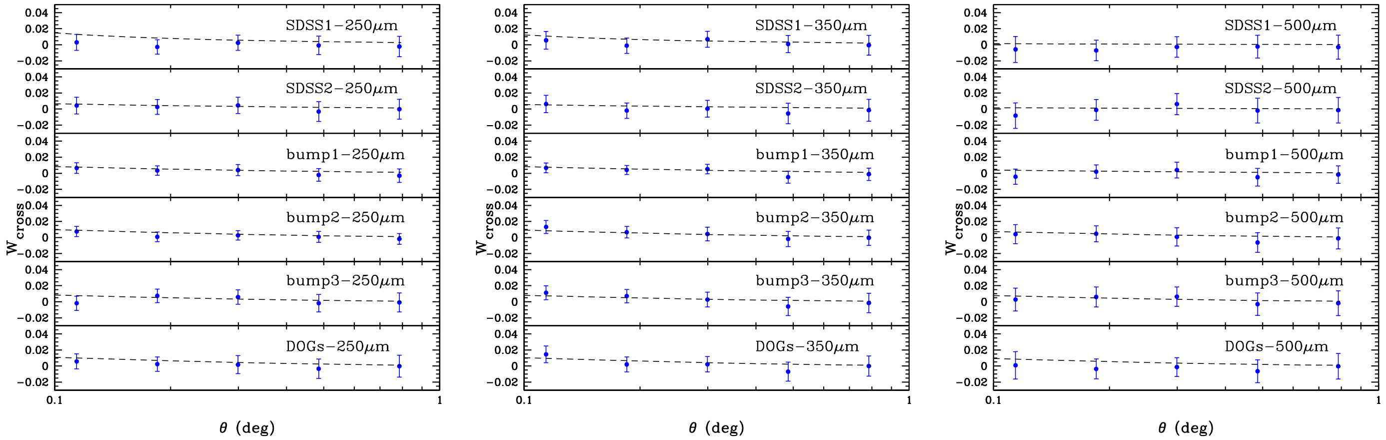

Figure 4.3 Angular cross-correlation between the 250, 350, and 500 μm objects of the SPIRE surveys and the SDSS1, SDSS2, bump1, bump2, bump3 and DOGs in the Boötes field. The 1σ error bars are derived from 200 bootstrap realizations. The black dashed lines are theoretical estimates of the angular cross-correlation using the best-fit value of the SPIRE redshift distribution.

where $N = 200$ is the number of the bootstrap realization, and $\bar{w}(\theta)$ is the average angular correlation for all bootstrap realizations at θ. The error of the angular correlation thus takes the form of $\sigma_w(\theta_i) = \sqrt{C_{ii}}$.

We use nine logarithmic bins from 0.01 to 1° to calculate the angular auto- and cross-correlation and their covariance matrix. The model correlation and cross-correlation functions, $w^{\rm th}$, are calculated for a given $N(z)$ and clustering bias factors (described in the next Section), and are compared with measurements, $w^{\rm data}$, using the covariance matrix from the data. In calculating $w^{\rm th}$, we make use of the measured $N(z)$ of the SDSS, bumps and DOGs, derived in the last section.

4.5 Estimating the SPIRE galaxy redshift distribution

We employ a Markov Chain Monte Carlo (MCMC) technique, using the Metropolis-Hastings algorithm (Hastings, 1970), to fit the SPIRE redshift distribution $N(z)$ of sources with flux

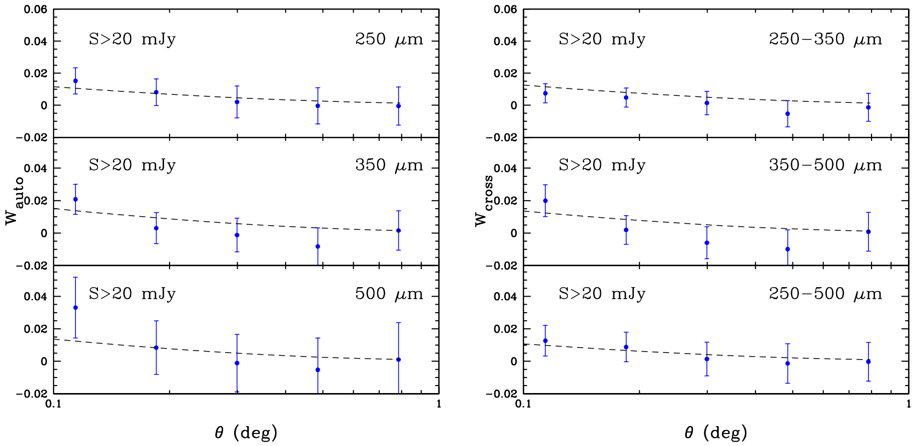

Figure 4.4 Angular auto- and cross-correlation for the 250, 350, and 500 μm SPIRE sources in the Boötes field. The 1σ error bars are derived from 200 bootstrap realizations. The black dashed lines are the theoretical estimation using the best-fit value of the SPIRE redshift distribution.

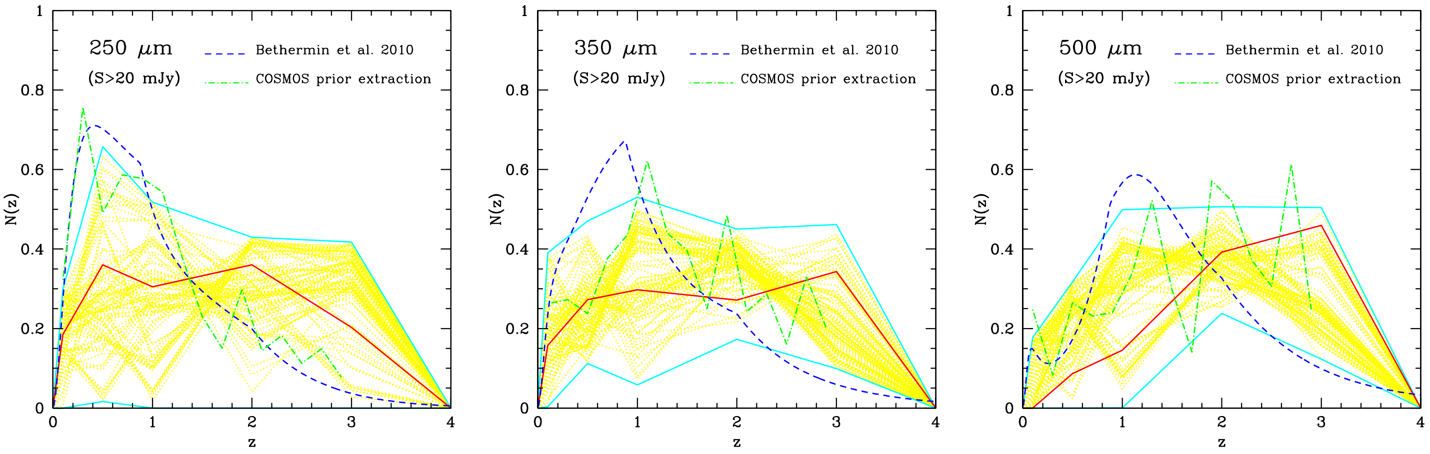

Figure 4.5 Best-fit normalized redshift distributions (red solid line) and 1σ error regions (cyan line) for sources with flux densities greater than 20 mJy for 250, 350 and 500 μm SPIRE bands using SDSS, bump-1 to bump-3, and DOGs catalogs in the Boötes field. As examples 100 best-fit N(z) from our MCMC results are also shown in yellow dotted lines. The analytical model predictions on $N(z)$ from the literature (blue dashed line; Béthermine et al. 2010) for galaxies in the three SPIRE bands are also shown for comparison. The green line is a direct estimate of $N(z)$ using the combination of a stacking and a cross-identification analysis involving 24 μm MIPS and SPIRE sources from Béthermin et al. (2012).

densities greater than 20 mJy at each of the three wavelengths. We follow established standard procedures in fitting the data, including thinning of the chains and separation of steps that are part of the initial burn-in period.

We describe the unknown redshift distribution $N(z)$ at five values, using five "pivot" redshifts $z_{\rm p} = 0.1, 0.5, 1, 2$ and 3, and set $N(z_{\rm p} = 0) = 0$ and $N(z_{\rm p} = 4) = 0$ to describe the two end points of the redshift distribution. To describe $N(z)$ when $0 < z < 4$, we linearly interpolate the fitted $N(z)$ distribution at each of the pivot redshifts $z_{\rm p}$ and use those linearly interpolated values between two pivots in our model fitting algorithm. The assumption that $N(z > 4) = 0$ does not bias our results since we only expect at most a few percent of the sub-mm galaxies to be located at $z > 4$ (e.g., Pope & Chary 2010). Furthermore, we do not have sensitivity to such high redshifts, given that the optical and near-IR galaxy samples we have used for the cross-correlation study are restricted to $z < 4$.

Before deciding on this description, we also considered a description of $N(z)$ that involved five bins in redshift, with $N(z)$ taking the same value in each of the bins. However, we failed to obtain fits to the binned case since in the first bin $0 < z < z_1$, $N(z)$ prefers a value that is non-zero at z_1, but zero at $z = 0$. The use of pivot redshifts and linear interpolation between pivots avoids the discontinuities that were present with the binned case, leading to issues with the numerical integrations of the clustering in equation 2.

As discussed earlier (related to equation 2), we also need to account for the clustering bias factor of galaxies and SPIRE sources relative to the linear matter power spectrum. Instead of keeping the bias in each of the bins as a free parameter, which leads to a large number of model parameters to be determined from the data, we assume a model for the galaxy bias, as a function of redshift, to be of the form

$$b(z) = b_0(1+z)^c, \tag{4.5}$$

where b_0 and c are free parameters to be determined from data using the MCMC analysis. In addition to this model we also consider two other approaches with: (i) $b(z) = b_0 + b_1 z$, a simple linear interpolation with redshift; and (ii) $b(z) = b_0$ when $z < 2$ and $b(z) = b_1$ when $z > 2$. We found results consistent within 1σ uncertainties in both $N(z)$ and $b(z)$ with the power-law form when using the linear relation.

For optical and IR galaxy samples we assume that each has an average bias factor, and we do not account for the redshift evolution of the bias factor in each of the galaxy samples. This is a fair assumption since each of the samples we have created has a narrow redshift distribution compared to the distribution expected for the SPIRE galaxies.

Altogether we have thirteen free parameters in our MCMC fitting, which contains five parameters for the SPIRE redshift distribution and six bias parameters for SDSS-1, SDSS-2, bump-1, bump-2, bump-3, and DOGs, plus two parameters to describe the SPIRE galaxy bias and its evolution with redshift. While the redshift distribution and bias factor and evolution for SPIRE sources are different at each of the three SPIRE wavelengths, the bias factors for optical and IR-selected galaxies remain the same. Thus, the six bias parameters for the galaxy samples, with assumed or known redshifts, can be determined jointly from cross-correlation data at the three SPIRE wavelengths together with their auto-correlation functions. We fix all the other cosmological parameters and assume the flat ΛCDM model as mentioned in Section 1.

We fit the data following the χ^2 distribution estimated as

$$\chi^2 = \sum_{\text{datasets}} \mathbf{\Delta^T C^{-1} \Delta}, \tag{4.6}$$

where $\mathbf{\Delta} = [w^{\text{data}}(\theta_1) - w^{\text{th}}(\theta_1),\ ...,\ w^{\text{data}}(\theta_9) - w^{\text{th}}(\theta_9)]$, $\mathbf{C}$ is the covariance matrix of $w(\theta)$, w^{th} is obtained directly from the $N(z)$, and "data" here are the full angular cross-correlations for the SPIRE, SDSS-1, SDSS-2, bump-1, bump-2, bump-3 and DOG samples (21 cross-correlations for each SPIRE band), and their auto-correlations in the Boötes field. The angular auto-correlation and the cross-correlation between the SPIRE and the SDSS1, SDSS2, bump1, bump2, bump3 and DOG sub-samples extracted from the observational data in the Boötes field are shown in Figs. 4.3 and 4.4 as examples.

We adopt an adaptive step-size Gaussian sampler given by Doran & Müller (2004) for the MCMC fitting process. We generate six chains with about 10^5 points after the convergence process. At the end we resample the chains to get about 10 000 points to illustrate the probability distribution of the parameters.

4.6 Results and Discussion

In Fig. 4.5 we show the best-fit results and the 1σ errors of the redshift distribution, $N(z)$, for the three SPIRE bands (see also Table 1 for the values). The redshift distributions are normalized such that $\int dz N(z) = 1$. The 1σ error bars (cyan lines) are derived from the Markov chains, which are statistically estimated via the values of $N(z)$ calculated using each chain point at different redshifts. As an example, the 100 best-fit N(z) are also shown in yellow dotted lines. As shown by the errors of $N(z)$ at high redshift ($z > 3$), the galaxy distribution could be larger when going from the $250\,\mu m$ to $500\,\mu m$ bands, which implies there may be more high-redshift galaxies for the $500\,\mu m$ band than the $250\,\mu m$ and $350\,\mu m$ bands. In Table 1 we also tabulate the average redshift of the SPIRE sources by calculating $\int dz\, zN(z)$, and these values range from 1.8 ± 0.2 at $250\,\mu$m to 1.9 ± 0.2 for $500\,\mu$m. We also derive the correlation coefficient for the $N(z)$ at five pivot redshifts from our Marcov chains (see Appendix). We find the correlation is weak (∼10%) between adjacent $N_{\rm i}$ pivots for each SPIRE band.

Two additional $N(z)$ predictions from the literature are also shown in the plot for comparison. The dashed line is a direct estimate of $N(z)$ from PSF-fitted extraction using 24 μm positions as a prior (Béthermin et al., 2012b) and the green curve is a model prediction for the SPIRE

Table 4.1 The best-fit SPIRE redshift distribution and bias parameters

$N(z)$	z-pivot	250 μm	350 μm	500 μm
N_1	0.1	$0.19^{+0.11}_{-0.19}$	$0.16^{+0.23}_{-0.15}$	$0.00^{+0.18}_{-0.00}$
N_2	0.5	$0.36^{+0.30}_{-0.34}$	$0.27^{+0.20}_{-0.16}$	$0.09^{+0.23}_{-0.09}$
N_3	1.0	$0.31^{+0.21}_{-0.31}$	$0.30^{+0.23}_{-0.24}$	$0.15^{+0.35}_{-0.15}$
N_4	2.0	$0.36^{+0.07}_{-0.36}$	$0.27^{+0.18}_{-0.10}$	$0.39^{+0.11}_{-0.15}$
N_5	3.0	$0.20^{+0.22}_{-0.20}$	$0.34^{+0.12}_{-0.24}$	$0.46^{+0.04}_{-0.34}$
Average Redshift				
$\langle z \rangle$		1.8 ±0.2	1.9 ±0.2	1.9 ±0.2
Sub-mm Bias				
b_0		$1.0^{+0.8}_{-0.5}$	$1.0^{+1.0}_{-0.5}$	$0.9^{+0.6}_{-0.5}$
c		$1.1^{+0.4}_{-0.6}$	$1.2^{+0.3}_{-0.7}$	$1.1^{+0.5}_{-0.8}$

redshift distribution (Béthermin et al., 2011). Our estimation for $N(z)$ for the 250 μm band agree well with both the direct extraction based on 24 μm identifications and a model prediction, while we find some differences at 350 and 500 μm. However, given the large uncertainties in our binned $N(z)$ estimate these differences are statistically insignificant.

In addition to $N(z)$ and bias factors of SPIRE-selected galaxies, we also measure the bias factors of the optical and IR-selected galaxy samples that we have used for cross-correlations. In Fig. 4.6 we show the two-dimensional error plots, and in Table 4.2 we list the best-fit bias values and their uncertainties. These results are obtained by combining the likelihoods from the MCMC chains of all three SPIRE bands. These values are consistent with values quoted in the literature for the bias of these samples. For example, the dust-obscured galaxies have an estimated bias factor of 3.1 ± 0.5 (Brodwin et al. 2008), which can be compared to our estimate of $2.6^{+1.1}_{-1.9}$. While fully consistent with the Brodwin et al. (2008) estimate, our central value is lower than their value, as we account for the full redshift distribution of

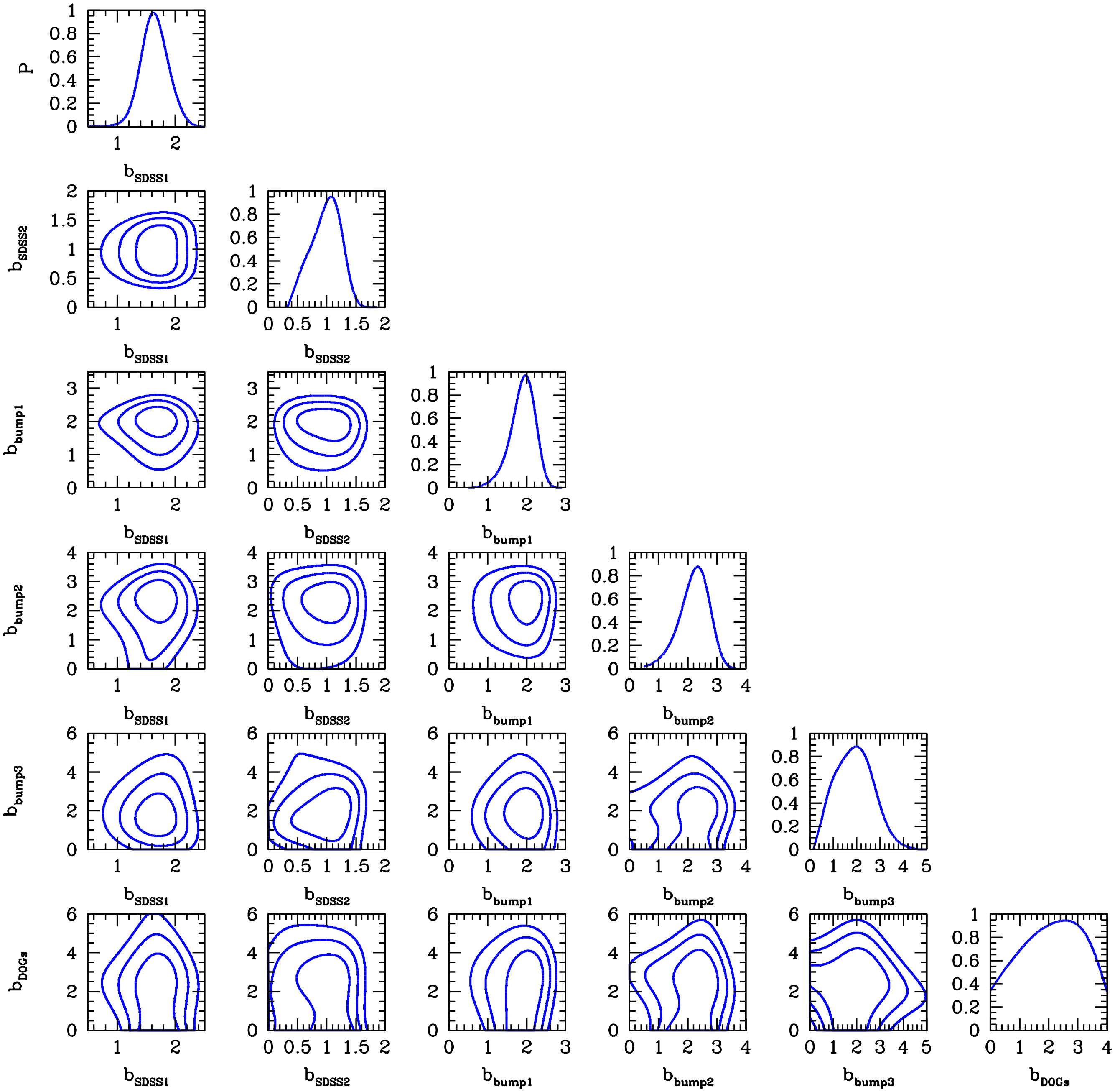

Figure 4.6 The one and two-dimensional probability distribution functions for bias parameters for all of the Boötes samples used throughout this paper. These bias parameters are estimated by combining the likelihoods from the MCMC chains of all three SPIRE bands. The 68.3%, 95.5% and 99.7% uncertainties from the fits are shown in the two-dimensional error plots.

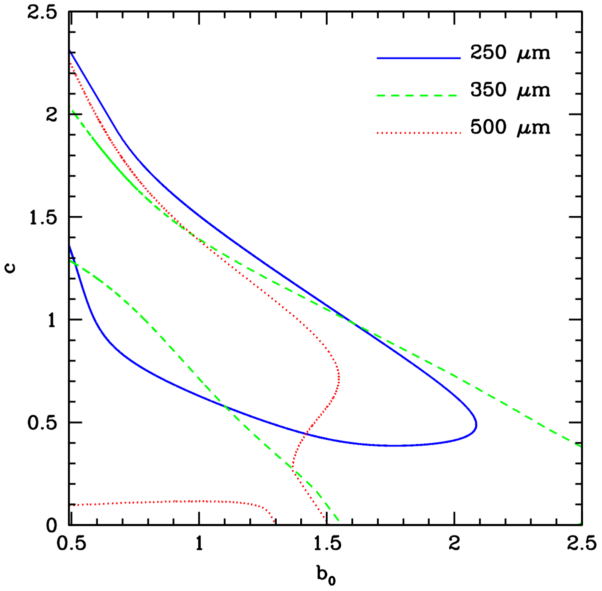

Figure 4.7 The 68% contour maps of the bias parameters b_0 and c, determined from the MCMC analysis, at 250, 350 and 500 μm with $S > 20\,mJy$ in the Boötes field.

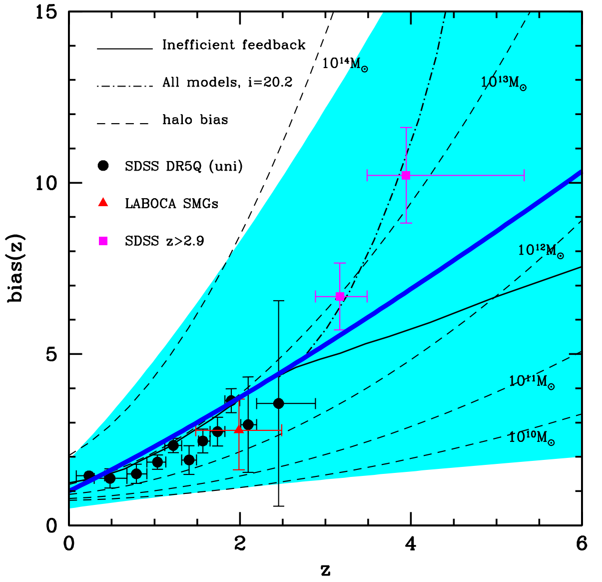

Figure 4.8 Clustering bias of $S_{350} > 20$ mJy SPIRE sources as a function of redshift. The shaded region shows the 68% confidence region allowed, with the blue solid line showing the best-fit $b(z)$ relation. For reference we plot the bias factor of dark matter halos as a function of halo mass. The range allowed by the data over $0 < z < 4$ is occupied by halos with mass $10 < \log M/\mathrm{M}_\odot < 14$. We also show samples of galaxies and quasars from the literature (Shen et al., 2007; Ross et al., 2009; Hickox et al., 2012), and two models for the evolution of the bias factor of merging galaxies from Hopkins et al. (2008) involving all three models at $i = 20.2$ (dash-dotted lines) and inefficient (black solid) feedback.

Table 4.2 Bias factors of optical and IR-selected galaxy samples

Sample	Approximate z-Range	bias
SDSS-1	$0 - 0.4$	$1.6^{+0.2}_{-0.2}$
SDSS-2	$0.3 - 0.7$	$1.1^{+0.2}_{-0.3}$
Bump-1	$0.8 - 1.5$	$2.0^{+0.3}_{-0.3}$
Bump-2	$1.2 - 2.0$	$2.3^{+0.4}_{-0.5}$
Bump-3	$1.6 - 2.5$	$2.0^{+0.8}_{-1.1}$
DOGs	$0.7 - 3.0$	$2.6^{+1.1}_{-1.9}$

these galaxies, while their analysis assumed a redshift of 2 for the whole DOG sample in the Boötes field.

In Fig. 4.7 we show the 68% confidence contour maps of the bias factors of SPIRE sources at the three wavelengths (the values and uncertainties are listed in Table 1). We generally find that the SPIRE galaxy bias factors are consistent with $b(z) \sim 1 + z$ (i.e. $c \approx 1$). To understand further the evolution of the sub-mm galaxy bias factor, we plot the redshift dependence in Fig. 4.8, where we compare with the bias factor of dark matter halos at several halo masses, from dwarf galaxy mass to galaxy cluster scales. The bias factors we find at all three wavelengths indicate a halo mass in the range of few times 10^{10} to few times 10^{13} $M_{\odot}$. The SPIRE clustering analysis in Cooray et al. (2010) found a halo mass for sub-mm galaxies that is about $3{\times}10^{12}$ $M_{\odot}$, under the assumption of a redshift distribution for the sub-mm galaxy population with a peak at $z \sim 2.3$, similar to the DOG redshift distribution in Fig. 2. We now find a slightly lower bias factor, and this is primarily due to the fact that the underlying redshift distribution of the SPIRE galaxies, especially at 250 μm, contains more sources at lower redshifts ($z \lesssim 1$). While the result here is for bright sub-mm sources that are individually detected, the model interpretation of the SPIRE anisotropy power spectrum by Amblard et al. (2011) found a minimum halo mass of 3×10^{11} $M_{\odot}$.

In Fig. 4.8 we also compare the SPIRE sub-mm galaxy bias factors to samples of galaxies and quasars from the literature (Shen et al., 2007; Ross et al., 2009; Hickox et al., 2012). Our results are generally consistent with the possibility that SMGs and quasars trace similar evolutionary paths and that the hosts correspond to dark matter halos that contain a $\sim$ few $L_\star$ ellipticals at $z \sim 0$. The exact mechanism on how the starburst galaxies seen in SPIRE feed the black holes that result in the quasars, and the subsequent feedback that suppresses star-formation, remains uncertain.

In Fig. 4.8 we also plot two models for the evolution of the bias factor of merging galaxies from Hopkins et al. (2007). While these models have similar behavior at $z < 3$, differences exist at higher redshift. A clustering study of SPIRE-selected sub-mm galaxies at $z > 4$ on its own, or as a cross-correlation with high-redshift quasars, could potentially be used to understand the intricate role of starbursts and quasars and to separate the subsequent feedback processes.

4.7 Conclusions

The wide-area sub-mm surveys with the SPIRE instrument aboard the *Herschel* Space Observatory have now led to catalogs of order one hundred thousand dusty, star-forming galaxies at 250, 350, and 500 μm. While some properties of this sub-mm source population are now understood, the redshift distribution of these galaxies, $N(z)$, is not yet well determined observationally. We make a statistical estimate of $N(z)$ using a clustering analysis involving the cross-correlation of sub-mm galaxies detected at each of 250, 350 and 500 μm from the *Herschel* Multi-tiered Extragalactic Survey (HerMES) centered on the Boötes field, against samples of galaxies detected at optical and near-IR wavelengths from the Sloan Digital Sky Survey (SDSS), the NOAO Deep Wide Field Survey (NDWFS), and the *Spitzer* Deep Wide Field Survey (SDWFS).

We create optical and near-IR galaxy samples based on their photometric or spectroscopic redshift distributions and test the accuracy of these redshift distributions with similar galaxy samples defined via catalogs of the Cosmological Evolution Survey (COSMOS). We fit the clustering auto and cross-correlations of SPIRE and optical/IR galaxy samples at angular scales of 0.1 to 1°, where clustering of each of the galaxy samples is expected to be linear, with the amplitude determined by a bias factor together with the redshift distribution of the sources. We make use of a Markov Chain Monte Carlo (MCMC) method to sample $N(z)$ at five nodes in the range $0 < z < 4$, as well as the bias factors. The SPIRE-selected sub-mm galaxy bias factor is found to vary with redshift according to $b(z) = 1.0^{+1.0}_{-0.5}(1+z)^{1.2^{+0.3}_{-0.7}}$. We find clear evidence of evolving redshift distributions as the wavelength increases from $250\,\mu$m to $500\,\mu$m, with the $250\,\mu$m band containing the largest number of low redshift sources. We also compare the measured redshift distribution to model predictions in the literature, and find an excess of sources in the highest redshift bin when compared to the model prediction from Béthermin et al. (2010), although in general our results agree with both predictions from the literature. With subsequent observations in more fields, this analysis could potentially be carried out again – incorporating more data in this analysis would reduce the size of the errors and more fully constrain the $N(z)$ of these sub-mm galaxies.

Chapter 5

Angular clustering of strongly lensed submillimeter galaxies

5.1 Introduction

In this work we consider the angular clustering of *Herschel* lensed sources and the effect their clustering has on previous measurements of the overall *Herschel* SMG population. Cooray et al. (2010) used the clustering of submm galaxies over 10 deg^2 and found what appears to be clear evidence of a 1-halo term at arcmin angular scales. With model fits, they infer the minimum dark matter halo mass a submm galaxie can reside in. However they did not include the (unknown) effects of the underlying lensed distribution in their $w(\theta)$ measurements. Although lenses are identified with $S_{500\,\mu m} > 100$ mJy, there still exists a distribution extending down to lower flux densities, which effectively boosts galaxies with lower intrinsic luminosities above the lower threshold flux cut (typically 3-5σ) performed in clustering studies.

In the next section we briefly discuss the lensing galaxy candidate selection, followed by

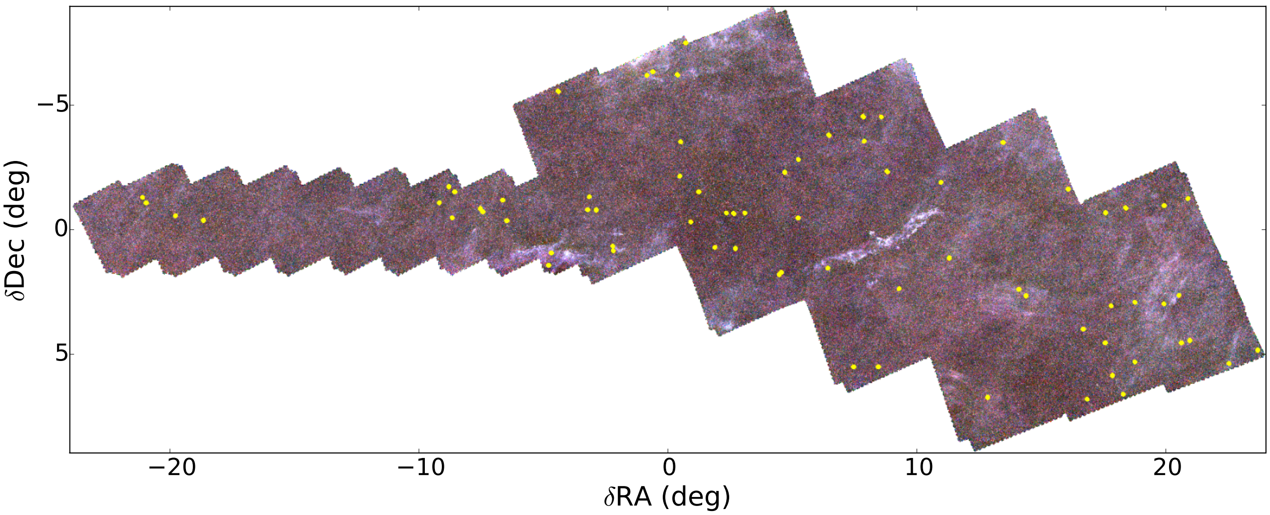

Figure 5.1 HerS (left portion/shorter scan lengths) and HeLMS three color mosaic made from 250 (blue), 350 (green) and 500 μm (red) bands. The yellow circles denote the postions of the lensed candidates.

the $w(\theta)$ measurements. In §5.3, we describe simulations of random, wide-area lensing realizations, which exhibit the variance of the lensing $w(\theta)$ over 350 deg^2 areas, assuming a lensing number density of 0.1-0.2 deg^2. In §5.4, we show the effect faint lenses have on the overall SMG population $w(\theta)$ over ~ 1 arcmin angular scales. The results of our measurements and simulations are discussed in §5.5.

5.2 Catalogs

Cosmic lensing events occur when the line-of-sight of a massive foreground galaxy or galaxy cluster coincides with a high-z background galaxy. In this work we assume luminous red galaxies (LRGs) make up the foreground distribution of our lensing field. In the next section we show the $w(\theta)$ of LRGs with a catalog taken from the 2dF-SDSS LRG and QSO (2SLAQ) Luminous Red Galaxy Survey (Cannon et al., 2006). This $w(\theta)$ is made useful in the simulations discussed in the subsequent sections.

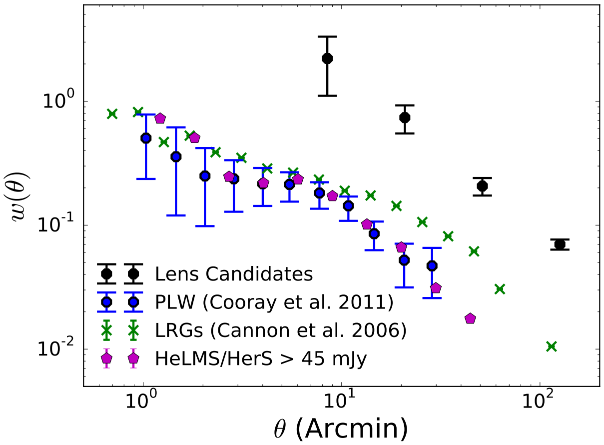

Figure 5.2 Angular correlation function for the lensed candiate population (black circles), the overall population of the PLW band measured with the HerMES observations of Lockman-SWIRE and the Spitzer First Look Survey (FLS) (Cooray et al. 2010; blue circles), the contribution from LRGs (Cannon et al., 2006), and the overall PLW population $>$ 45 mJy in HeLMS/HerS.

For the lensing candidate selection, we use the majority of the catalog from Nayyeri et al. (2016). The lenses we consider in this work are shown overlaid on the HeLMS/HerS mosaic in Fig 5.1.

We use a public two-point angular correlation estimator called TreeCorr (Jarvis et al., 2004; Jarvis, 2015) to estimate $w(\theta)$, which uses ball trees for high computational efficiency. We implement the Landy-Szalay estimator (Landy & Szalay, 1993), which approximates $w(\theta)$ as

$$w(\theta) = \frac{DD - 2DR + RR}{RR}. \tag{5.1}$$

DD is the number of galaxy-galaxy pairs, DR is the number of galaxy-random pairs, and RR is the number of random-random pairs. We generate random uniform catalogs over the same sky area as the data catalogs. Our measured angular correlation function for various catalogs is show in Fig. 5.2.

5.3 Cosmic Lensing Variance

In the HeLMS/HerS field we are working with only 72 sources, which is a statistically insignificant number. It is then reasonable to assume that this relatively small distribution might not be representative of the full sky lensing $w(\theta)$. In order to test this, we construct Monte Carlo simulations to assess the variance of the lensing $w(\theta)$ that can be reasonably expected between different fields – what we call a "cosmic lensing variance". This is done by measuring the $w(\theta)$ of many different realizations of lensing distributions, which are products of chance LRG-SMG line-of-sight alignments, out to some characteristic angle. This characteristic angle, in arcsec, is the Einstein radius:

$$\theta_E = \sqrt{\left(\frac{M}{10^{11.09} M\odot}\right) \Big/ \left(\frac{D_L D_S / D_{LS}}{\text{Gpc}}\right)} \tag{5.2}$$

where M is the mass of the foreground galaxy, D_L is the distance to the lensing galaxy, D_S is the distance to the "source", or background SMG, and D_{LS} is the difference of D_L and D_S. In our simulations we approximate the mass of all the foreground galaxies to be $10^{11} M_{\odot}$. Distances are calculated with $z_L = 0.6$ and $z_S = 1.5$, using the standard cosmological parameters (Planck Collaboration et al., 2015).

The simulations are initialized with a foreground LRG population with an average number density of 70 deg^2. These sources are put down on a field of ~ 350 deg^2 following clustering of the form $w(\theta) \propto \theta^{-1}$. A simulated SMG population is added to the field with similar clustering statistics, and a number density congruent with the cumulative number counts in Fig 6 of Wardlow et al. (2013), down to 50 mJy. Lenses are then flagged as any SMG within some Einstein radius θ_E. For each simulation, this is typically 50-80 sources. The $w(\theta)$ of these resulting lenses is calculated and one iteration of the simulation is complete.

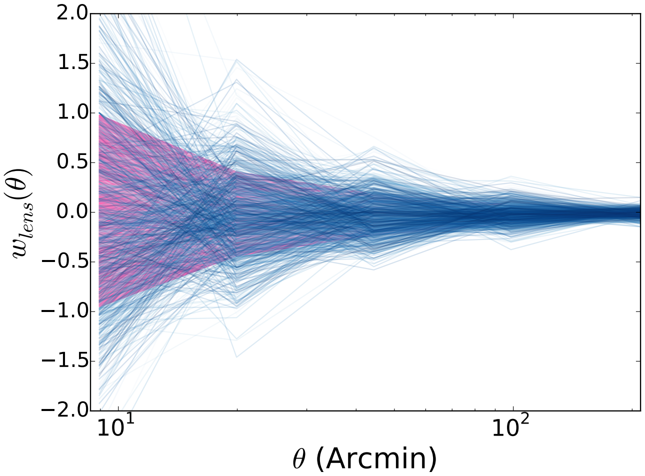

Figure 5.3 Monte Carlo simulations show a cosmic lensing variance effect over 350 deg^2. Each line is a new realization of an SMG distribution. The pink underlying area represents the 1 σ interval.

Table 5.1 Results from 1000 simulations over 350 deg^2.

θ (arcmin)	$< w(\theta) >$	Var[$w(\theta)$]
9.0	0.014	0.96
19.9	-0.016	0.42
44.3	-0.021	0.21
98.7	-0.017	0.094
219.6	-0.015	0.042

Results of this exercise are show in fig 5.3, which was generated from 10^3 different SMG distributions. Each curve from that figure is one realization of the simulation; the pink shaded region shows the 1-σ interval at each bin, which is also found in Table 5.2. The large variance around 10 arcmin essentially shows the range of lensing $w(\theta)$ measurements one can expect from surveys limited to hundreds of square degrees, from e.g. the North and South Galactic Pole fields (NGP, SGP) observed in the H-ATLAS program.

5.4 Lensing bias of the *Herschel* $w(\theta)$

With slight modifications to the simulations discussed in the previous section, we can attempt to quantify what the contribution of the underlying ($S < 100$ mJy) lensing population has on the overall *Herschel* $w(\theta)$. Cooray et al. (2010) used the angular clustering of *Herschel* sources down to 30 mJy to infer the minimum dark matter halo mass that host submm galaxies. They found evidence for a 1-halo term at 1 arcmin. These measurements were made over ~ 10 deg^2 area fields and did not consider the contribution from faint lens clustering.

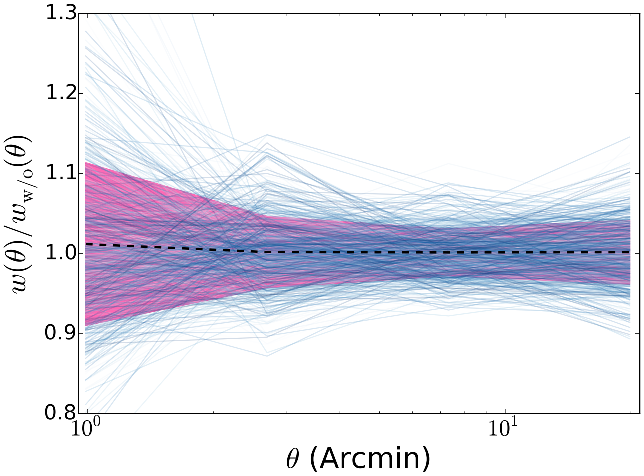

Figure 5.4 Simulations to show a slight bias introduced by the underlying lenses, at $\sim$ 1 arcmin. As in the previous plot, each line is a new realization, the pink underlying area is one standard deviation, and the dashed line is the average. $w(\theta)$ is measured from the total catalog and $w_{\mathrm{w/o}}(\theta)$ is measured with the lenses omitted from that same catalog.

Table 5.2 $w(\theta)$ bias results from 1000 simulations over 350 deg^2.

θ (arcmin)	$< w(\theta) >$	Var$[w(\theta)]$
0.98	1.01	0.10
2.67	1.00	0.045
7.31	1.00	0.029
19.87	1.00	0.040

These simulations are contrived in such a way that the field-of-view is reduced to 10 deg^2 and the number density of *Herschel* sources is consistent with a cumulative number count distribution down to 30 mJy. In this scenario, we obtain 1500-1700 submm sources per realization. We again put this distribution on the same (reduced area) LRG map and flag all the lensed sources – submm galaxies occuring within some Einstein radius of any LRG. We compute $w(\theta)$ of the total distribution, containing both lensed SMGs and unlensed SMGs – this should be statistically equivalent to the measurements in Cooray et al. (2010). The flagged lensed sources are then removed from the simulated catalog and $w(\theta)$ is computed again. We take the ratio of these two $w(\theta)$ measurements, which is shown in Fig 5.4, where $w_{\mathrm{w/o}}(\theta)$ represents the correlation function from the catalog without lenses.

5.5 Discussion

The results of these simulations are two-fold. First, the wide-area field of HeLMS + HerS is insufficient to be representative of the clustering distribution of lensed sources over the full sky. Secondly, the bias simulations show that the current $w(\theta)$ measurements may be biased up to 10%, which would in turn affect the minimum halo mass values inferred from the $w(\theta)$ measurements.

Bibliography

Abraham, P., Leinert, C., & Lemke, D. 1997, 328, 702

Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2003, *Astron. J.*, 126, 539

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

Arendt, R. G., Fixsen, D. J., & Moseley, S. H. 2000, *Astrophys. J.*, 536, 500

Arendt, R. G., Fixsen, D. J., & Moseley, S. H. 2002a, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 217

Arendt, R. G., Fixsen, D. J., & Moseley, S. H. 2002b, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 217

Arendt, R. G., Kashlinsky, A., Moseley, S. H., & Mather, J. 2010, *Astrophys. J. Suppl.*, 186, 10

Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, *Astrophys. J.*, 769, 80

—. 2015, *Astrophys. J. Suppl.*, 218, 33

Atek, H., Richard, J., Kneib, J.-P., et al. 2015, *Astrophys. J.*, 800, 18

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, *Astron. J.*, 132, 1729

Béthermin, M., Dole, H., Lagache, G., Le Borgne, D., & Penin, A. 2011, *Astro. Astrophys.*, 529, A4

Béthermin, M., Daddi, E., Magdis, G., et al. 2012a, *Astrophys. J.*, 757, L23

Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012b, *Astro. Astrophys.*, 542, A58

—. 2012c, *Astro. Astrophys.*, 542, A58

Blain, A. W., & Longair, M. S. 1993, *Mon. Not. Royal Astron. Soc.*, 264, 509

Bouwens, R. J., Bradley, L., Zitrin, A., et al. 2014, *Astrophys. J.*, 795, 126

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *Astrophys. J.*, 803, 34

Brandt, T. D., & Draine, B. T. 2012, *Astrophys. J.*, 744, 129

Brodwin, M., Dey, A., Brown, M. J. I., et al. 2008, *Astrophys. J.*, 687, L65

Broos, P. 2000, in University Park: Pennsylvania State Univ., Vol. 0, "User's Guide for the TARA Package"

Bussmann, R. S., Pérez-Fournon, I., Amber, S., et al. 2013, *Astrophys. J.*, 779, 25

Cannon, R., Drinkwater, M., Edge, A., et al. 2006, *Mon. Not. Royal Astron. Soc.*, 372, 425

Cappelluti, N., Ranalli, P., Roncarelli, M., et al. 2012, *Mon. Not. Royal Astron. Soc.*, 427, 651

Cappelluti, N., Kashlinsky, A., Arendt, R. G., et al. 2013, *Astrophys. J.*, 769, 68

Cappelluti, N., Comastri, A., Fontana, A., et al. 2016, *Astrophys. J.*, 823, 95

Casey, C. M., Narayanan, D., & Cooray, A. 2014, Phys. Rep., 541, 45

Chary, R.-R., Cooray, A., & Sullivan, I. 2008, *Astrophys. J.*, 681, 53

Clements, D. L., Rigby, E., Maddox, S., et al. 2010, *Astro. Astrophys.*, 518, L8

Coe, D., Zitrin, A., Carrasco, M., et al. 2013, *Astrophys. J.*, 762, 32

Cooray, A., Bock, J. J., Keatin, B., Lange, A. E., & Matsumoto, T. 2004, *Astrophys. J.*, 606, 611

Cooray, A., Gong, Y., Smidt, J., & Santos, M. G. 2012a, *Astrophys. J.*, 756, 92

Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1

Cooray, A., Sullivan, I., Chary, R.-R., et al. 2007, *Astrophys. J.*, 659, L91

Cooray, A., Amblard, A., Wang, L., et al. 2010, *Astro. Astrophys.*, 518, L22

Cooray, A., Smidt, J., de Bernardis, F., et al. 2012b, *Nature*, 490, 514

Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, *Astrophys. J.*, 660, L1

de Zotti, G., Ricci, R., Mesa, D., et al. 2005, *Astro. Astrophys.*, 431, 893

Dey, A., & Ndwfs/MIPS Collaboration. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 408, The Starburst-AGN Connection, ed. W. Wang, Z. Yang, Z. Luo, & Z. Chen, 411

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

Donnerstein, R. L. 2015, *Mon. Not. Royal Astron. Soc.*, 449, 1291

Doran, M., & Müller, C. M. 2004, J. Cosmology Astropart. Phys., 9, 003

Eales, S., Dunne, L., Clements, D., et al. 2010, PASP, 122, 499

Eisenhardt, P. R., Stern, D., Brodwin, M., et al. 2004, *Astrophys. J. Suppl.*, 154, 48

Fan, X., Carilli, C. L., & Keating, B. 2006a, *Ann. Rev. of Astron. Astrophys.*, 44, 415

Fan, X., Strauss, M. A., Becker, R. H., et al. 2006b, *Astron. J.*, 132, 117

Fazio, G., Soifer, T., Willner, S., et al. 2011, The Spitzer Ultra-Deep Survey of the HST/ERS Field, Spitzer Proposal

Fernandez, E. R., Iliev, I. T., Komatsu, E., & Shapiro, P. R. 2012, *Astrophys. J.*, 750, 20

Fernandez, E. R., Komatsu, E., Iliev, I. T., & Shapiro, P. R. 2010, *Astrophys. J.*, 710, 1089

Finkelstein, S. L., Ryan, Jr., R. E., Papovich, C., et al. 2014, ArXiv e-prints, arXiv:1410.5439

Fixsen, D. J., Moseley, S. H., & Arendt, R. G. 2000, *Astrophys. J. Suppl.*, 128, 651

Franceschini, A., Rodighiero, G., & Vaccari, M. 2008, *Astro. Astrophys.*, 487, 837

Fruchter, A. S., & Hook, R. N. 2002, PASP, 114, 144

Giacconi, R., Zirm, A., Wang, J., et al. 2002, *Astrophys. J. Suppl.*, 139, 369

Giallongo, E., Grazian, A., Fiore, F., et al. 2015, *Astro. Astrophys.*, 578, A83

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *Astrophys. J.*, 600, L93

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

Goulding, A. D., Forman, W. R., Hickox, R. C., et al. 2012, *Astrophys. J. Suppl.*, 202, 6

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

Grogin, N. A., Lim, P. L., Maybhate, A., Hook, R. N., & Loose, M. 2010a, in Hubble after SM4. Preparing JWST

Grogin, N. A., Lim, P. L., Maybhate, A., Hook, R. N., & Loose, M. 2010b, in Hubble after SM4. Preparing JWST

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *Astrophys. J. Suppl.*, 197, 35

Guhathakurta, P., & Tyson, J. A. 1989, *Astrophys. J.*, 346, 773

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *Astrophys. J. Suppl.*, 207, 24

Hack, W. J., Dencheva, N., Fruchter, A. S., et al. 2012, in American Astronomical Society Meeting Abstracts, Vol. 220, American Astronomical Society Meeting Abstracts #220, 135.15

Harris, A. I., Baker, A. J., Frayer, D. T., et al. 2012, *Astrophys. J.*, 752, 152

Hastings, W. K. 1970, Biometrika, 57, 97

Helgason, K., Cappelluti, N., Hasinger, G., Kashlinsky, A., & Ricotti, M. 2014, *Astrophys. J.*, 785, 38

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

Henriques, B. M. B., White, S. D. M., Lemson, G., et al. 2012, *Mon. Not. Royal Astron. Soc.*, 421, 2904

H.E.S.S. Collaboration, Abramowski, A., Acero, F., et al. 2013, *Astro. Astrophys.*, 550, A4

Hickox, R. C., & Markevitch, M. 2006, *Astrophys. J.*, 645, 95

Hickox, R. C., Wardlow, J. L., Smail, I., et al. 2012, *Mon. Not. Royal Astron. Soc.*, 421, 284

Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, *Astrophys. J.*, 567, 2

Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, *Astrophys. J. Suppl.*, 175, 356

Hu, W. 2000, Phys. Rev. D, 62, 043007

Huang, J.-S., Cowie, L. L., Gardner, J. P., et al. 1997, *Astrophys. J.*, 476, 12

Ienaka, N., Kawara, K., Matsuoka, Y., et al. 2013, *Astrophys. J.*, 767, 80

Ilbert, O., Capak, P., Salvato, M., et al. 2009, *Astrophys. J.*, 690, 1236

Ivison, R. J., Swinbank, A. M., Swinyard, B., et al. 2010, *Astro. Astrophys.*, 518, L35

Jarvis, M. 2015, TreeCorr: Two-point correlation functions, Astrophysics Source Code Library, ascl:1508.007

Jarvis, M., Bernstein, G., & Jain, B. 2004, *Mon. Not. Royal Astron. Soc.*, 352, 338

John, T. L. 1988, *Astro. Astrophys.*, 193, 189

Johnson, H. L. 1966, *Ann. Rev. of Astron. Astrophys.*, 4, 193

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

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

—. 2007, *Astrophys. J.*, 654, L1

Kashlinsky, A., Mather, J. C., Helgason, K., et al. 2014, ArXiv e-prints, arXiv:1412.5566

Kelsall, T., Weiland, J. L., Franz, B. A., et al. 1998, 508, 44

Kistler, M. D., Yüksel, H., Beacom, J. F., Hopkins, A. M., & Wyithe, J. S. B. 2009, *Astrophys. J.*, 705, L104

Koekemoer, A. M., & et al. 2002, HST Dither Handbook

Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, *Astrophys. J. Suppl.*, 172, 196

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *Astrophys. J. Suppl.*, 197, 36

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *Astrophys. J. Suppl.*, 192, 18

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

Krist, J. E., Hook, R. N., & Stoehr, F. 2011, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8127, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 0

Labbe, I., Oesch, P. A., Illingworth, G. D., et al. 2015, ArXiv e-prints, arXiv:1507.08313

Landy, S. D., & Szalay, A. S. 1993, *Astrophys. J.*, 412, 64

Larson, R. B. 1999, in Star Formation 1999, ed. T. Nakamoto, 336–340

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, 379, 1599

Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, *Astrophys. J.*, 752, 46

Lehmer, B. D., Berkeley, M., Zezas, A., et al. 2014, *Astrophys. J.*, 789, 52

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

Loeb, A., & Barkana, R. 2001, *Ann. Rev. of Astron. Astrophys.*, 39, 19

Luo, B., Bauer, F. E., Brandt, W. N., et al. 2008, *Astrophys. J. Suppl.*, 179, 19

Lupu, I., Courbin, P., George, L., & Goossens, J. 2010, ArXiv e-prints, arXiv:1004.3715

Madau, P., Haardt, F., & Rees, M. J. 1999, *Astrophys. J.*, 514, 648

Marsden, G., Ade, P. A. R., Bock, J. J., et al. 2009, *Astrophys. J.*, 707, 1729

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

Mauduit, J.-C., Lacy, M., Farrah, D., et al. 2012, PASP, 124, 714

McLure, R. J., Dunlop, J. S., Bowler, R. A. A., et al. 2013, *Mon. Not. Royal Astron. Soc.*, 432, 2696

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

Mitchell-Wynne, K., Cooray, A., Gong, Y., et al. 2015, Nature Commun, 6, 7945

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

Nayyeri, H., Keele, M., Cooray, A., et al. 2016, *Astrophys. J.*, 823, 17

Negrello, M., Perrotta, F., González-Nuevo, J., et al. 2007, *Mon. Not. Royal Astron. Soc.*, 377, 1557

Negrello, M., Hopwood, R., De Zotti, G., et al. 2010, Science, 330, 800

Oesch, P. A., Bouwens, R. J., Carollo, C. M., et al. 2010, *Astrophys. J.*, 725, L150

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2014, *Astrophys. J.*, 786, 108

Oliver, S. J., Wang, L., Smith, A. J., et al. 2010, *Astro. Astrophys.*, 518, L21

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, *Mon. Not. Royal Astron. Soc.*, 424, 1614

Paley, E. S., Low, F. J., McGraw, J. T., Cutri, R. M., & Rix, H.-W. 1991, *Astrophys. J.*, 376, 335

Peebles, P. J. E. 1968, *Astrophys. J.*, 153, 1

—. 1980, The large-scale structure of the universe

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *Astro. Astrophys.*, 518, L1

Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2014, ArXiv e-prints, arXiv:1409.5738

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, ArXiv e-prints, arXiv:1502.01589

Purcell, C. W., Bullock, J. S., & Zentner, A. R. 2007, *Astrophys. J.*, 666, 20

Reddy, N. A., & Steidel, C. C. 2009, *Astrophys. J.*, 692, 778

Riechers, D. A., Cooray, A., Omont, A., et al. 2011, *Astrophys. J.*, 733, L12

Riess, A. G., Strolger, L.-G., Tonry, J., et al. 2004, *Astrophys. J.*, 607, 665

Rix, H.-W., Barden, M., Beckwith, S. V. W., et al. 2004, *Astrophys. J. Suppl.*, 152, 163

Robertson, B. E., & Ellis, R. S. 2012, *Astrophys. J.*, 744, 95

Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, *Astrophys. J.*, 768, 71

Ross, N. P., Shen, Y., Strauss, M. A., et al. 2009, *Astrophys. J.*, 697, 1634

Salpeter, E. E. 1955, *Astrophys. J.*, 121, 161

Salvaterra, R., & Ferrara, A. 2003, *Mon. Not. Royal Astron. Soc.*, 339, 973

Santos, M. R., Bromm, V., & Kamionkowski, M. 2002, *Mon. Not. Royal Astron. Soc.*, 336, 1082

Schiminovich, D., Ilbert, O., Arnouts, S., et al. 2005, *Astrophys. J.*, 619, L47

Scott, K. S., Lupu, R. E., Aguirre, J. E., et al. 2011, *Astrophys. J.*, 733, 29

Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, *Astrophys. J. Suppl.*, 172, 38

Seo, H. J., Lee, H. M., Matsumoto, T., et al. 2015, ArXiv e-prints, arXiv:1504.05681

Serjeant, S., & Harrison, D. 2005, *Mon. Not. Royal Astron. Soc.*, 356, 192

Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, *Astron. J.*, 133, 2222

Snowden, S. L., Freyberg, M. J., Kuntz, K. D., & Sanders, W. T. 2000, *Astrophys. J. Suppl.*, 128, 171

Sullivan, I., Cooray, A., Chary, R.-R., et al. 2007, *Astrophys. J.*, 657, 37

Tanvir, N. R., Levan, A. J., Fruchter, A. S., et al. 2012, *Astrophys. J.*, 754, 46

Thacker, C., Gong, Y., Cooray, A., et al. 2015, *Astrophys. J.*, 811, 125

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

Thompson, R. I., Eisenstein, D., Fan, X., Rieke, M., & Kennicutt, R. C. 2007, *Astrophys. J.*, 666, 658

Thompson, R. I., Illingworth, G., Bouwens, R., et al. 2005, *Astron. J.*, 130, 1

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, *Astrophys. J.*, 688, 709

Townsley, L. K., Broos, P. S., Garmire, G. P., & Nousek, J. A. 2000, *Astrophys. J.*, 534, L139

Townsley, L. K., Broos, P. S., Nousek, J. A., & Garmire, G. P. 2002, Nuclear Instruments and Methods in Physics Research A, 486, 751

Treister, E., Schawinski, K., Volonteri, M., & Natarajan, P. 2013, *Astrophys. J.*, 778, 130

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

Viero, M. P., Asboth, V., Roseboom, I. G., et al. 2014, *Astrophys. J. Suppl.*, 210, 22

Waddington, I., Oliver, S. J., Babbedge, T. S. R., et al. 2007, *Mon. Not. Royal Astron. Soc.*, 381, 1437

Wang, L., Cooray, A., Farrah, D., et al. 2011, *Mon. Not. Royal Astron. Soc.*, 414, 596

Wang, Q. D., & Yu, K. C. 1995, *Astron. J.*, 109, 698

Wardlow, J. L., Cooray, A., De Bernardis, F., et al. 2013, *Astrophys. J.*, 762, 59

Weigel, A. K., Schawinski, K., Treister, E., et al. 2015, *Mon. Not. Royal Astron. Soc.*, 448, 3167

Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, *Astrophys. J. Suppl.*, 193, 27

Witt, A. N., Mandel, S., Sell, P. H., Dixon, T., & Vijh, U. P. 2008, *Astrophys. J.*, 679, 497

Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *Astrophys. J. Suppl.*, 195, 10

Yue, B., Ferrara, A., & Salvaterra, R. 2016, ArXiv e-prints, arXiv:1601.02514

Yue, B., Ferrara, A., Salvaterra, R., & Chen, X. 2013a, *Mon. Not. Royal Astron. Soc.*, 431, 383

Yue, B., Ferrara, A., Salvaterra, R., Xu, Y., & Chen, X. 2013b, *Mon. Not. Royal Astron. Soc.*, 433, 1556

—. 2014, *Mon. Not. Royal Astron. Soc.*, 440, 1263

Zemcov, M., Smidt, J., Arai, T., et al. 2014, Science, 346, 732

Zheng, W., Postman, M., Zitrin, A., et al. 2012, *Nature*, 489, 406

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, *Astrophys. J.*, 633, 791