SUPPLEMENTARY INFORMATION

In these Supplemental Notes to the main paper, we outline key details related to how we estimated the angular power spectrum of *Spitzer*/IRAC data and its interpretation. The data used for this study are publicly available from the Spitzer Heritage Archive[1] under the program number GO 40839 (PI. D. Stern).

1 The Spitzer Deep Wide Field Survey

We use the IRAC data from the Spitzer Deep Wide Field Survey (SDWFS)[16,31]. Of the NDWFS Boötes fields, the observations were obtained in 4 epochs, with depth per pixel of 90 seconds in each epoch. Each epoch took observations over 7 to 10 days to complete the full mosaic. We started with the basic calibrated data (BCDs). At each epoch, imaging data was obtained in all four IRAC wavelengths or channels (3.6, 4.5, 5.8, and 8 μm); however, for this study we focus on the 3.6 and 4.5 μm data. Each epoch consists of 4300 to 4900 BCDs per channel that were mosaiced to form the final images used in the fluctuation analysis.

Instead of using the public SDWFS mosaics (Available at http://irsa.ipac.caltech.edu), made with the standard *Spitzer* data analysis pipeline MOPEX[32], we produced our own mosaics in order to better control systematic errors. To mosaic the BCDs for each channel and epoch we used a self-calibration algorithm[17,33] to properly match the sky background level from one adjacent frame to the other in the overlapping region using an optimized least squares fitting technique. The SDWFS mapping strategy incorporates several elements to facilitate self-calibration of the data by maximizing inter-pixel correlations. We dithered the observations on small scales and offset by one-third of an IRAC field-of-view between successive passes through each group. This provides inter-pixel correlation information on both small and large scales, so the self-calibrated mosaic has background levels that are stable across the wide area of the SDWFS mosaic. Finally, for the larger, rectangular groups, we cadence the observations such that revisits cover the same area but with a different step size. With a $< 10\%$ penalty in mapping efficiency, cadencing significantly enhanced the inter-pixel correlations across all scales.

These mapping strategies were designed to significantly enhance the self-calibration of the data. Finally, by reobserving the field multiple times at different roll angles, our observing strategy was designed to be robust against bad rows/columns, large scale cosmetic defects on the array, after-images resulting from saturation due to bright stars, variations in the bias level, and the color dependence of the IRAC flat-field across the array[34]. In particular, the challenging diffuse background measurements we report here are vastly aided by the redundant coverage: independent data sets of the same region are the best way to assess and control systematic errors. With this mapping strategy we were able to construct independent sky realizations for carrying out jackknife testing. Furthermore, the power spectrum is estimated by cross-correlating the maps from different epochs, eliminating bias from uncorrelated signals such as instrumental noise and mosaicing artifacts.

---

[1]http://sha.ipac.caltech.edu/applications/Spitzer/SHA/

For each IRAC channel and epoch we passed the cleaned BCD into our self-calibration code, a slightly modified version of an existing code[17], as inputs the cleaned BCDs (cBCDs) with the final, output array size and astrometry defined to correspond to the mosaic of first 3.6 μm epoch. The cBCDs were first cleaned of asteroid trails, hot pixels, and other image artifacts. Since the astrometry is the same for each mosaic, they can be properly coadded and jack-knifed as described below. Each of the cBCDs have an angular pixel scale of 1.2 arcseconds, which was preserved in the final mosaics. The portion of the maps used for analysis are $\sim 3.5 \times 3$ degrees on a side for a total area of ~ 10.5 square degrees. The final mosaics generated from the self-calibration algorithm are shown in Fig. S1.
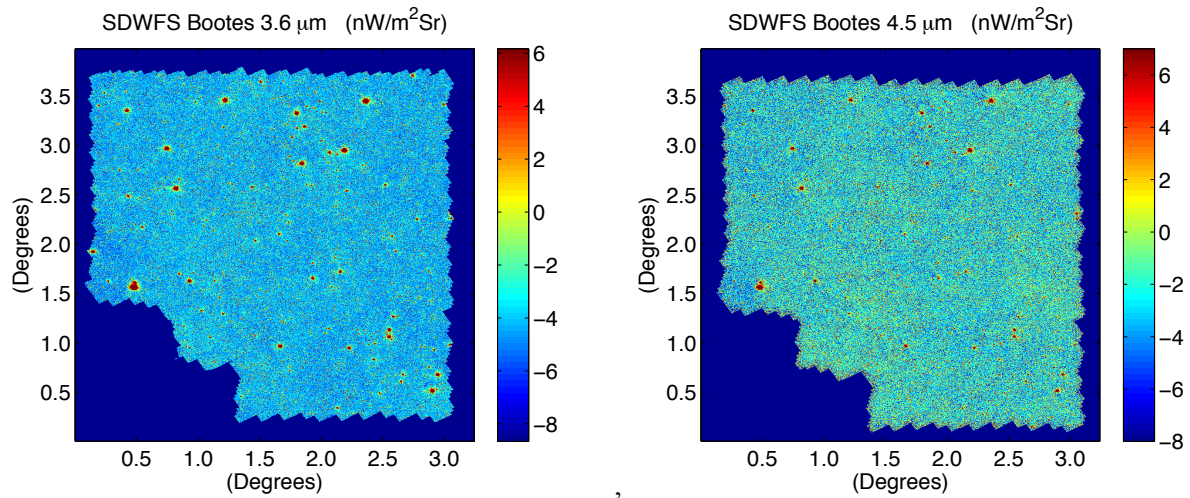

Figure S 1: **The SDWFS Maps.** The final 3.6 μm (left) and 4.5 μm (right) self-calibrated mosaics used for the Boötes SDWFS analysis.

2   Generation of the Detected Source Mask

As the analysis concentrates on the background intensity fluctuations, with the aim of identifying the nature of faint sources below the individual detection level, we must remove the contamination from individual detected sources in our power spectrum measurement. We created source masks based both on the objects detected in the IRAC data at 3.6 and 4.5 μm and the NOAO Deep Wide Field Survey (NDWFS) catalogs for the B_w, R and I bands. The masked sources include stellar point sources, galaxies that are extended, and galaxies that are unresolved but detected as point sources in SDWFS. The mask also accounts for the Spitzer-IRAC point spread function (PSF). For this study we make use of the publicly available "extended" IRAC PSF to properly account for flux wings.

Next, we summarize our recipe to generate the source list and discuss more details below:

1. We create a catalog consisting of all objects detected by SEXTRACTOR at a 3σ detection

threshold on the coadded and individual epoch maps at both 3.6 and 4.5 μm.

2. We add any additional objects detected in the NDWFS optical catalogs, but missing from the IRAC catalogs.

3. We create a map initially of zeros where we place each source detected from our combined catalogs with the proper flux values. The sources that are flagged as extended in the SExtractor catalogs are placed as extended sources with a size that is comparable to their estimated size.

4. We convolve this final map with the IRAC PSF to ensure that the flux wings for each source are properly masked.

5. We make a cut at a certain flux level so that all pixels with intensity down to that threshold are masked. We then histogram the remaining pixels from the data map and cut all pixels that are $\pm 5\,\sigma$ away in the histogram. The latter step is similar to prior approaches to measure the *Spitzer*-IRAC background anisotropy power spectrum[36].

6. To produce the final mask, we set the pixels with intensity values above the final flux and outliers from histogram to a value of zero, and the remaining pixels, the ones used for anisotropy measurements, a value of one.

Before the mosaics were created, source extractor was run iteratively on each epoch and waveband individually as well as on the coadded maps for both wavebands in order to find detected sources. The parameters used for SEXTRACTOR are the same as the ones used for the original SDWFS catalogs[16]. The combined catalog obtained from this iterative source extractor analysis, as well as the objects detected in the NDWFS catalog are merged into a final catalog.
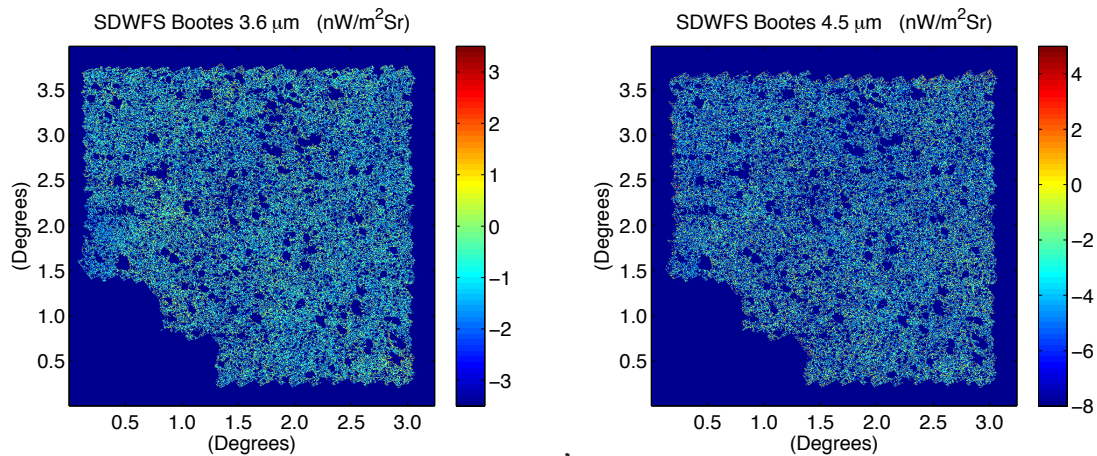

Figure S 2: **The Masked SDWFS maps.** The final 3.6 μm (left) and 4.5 μm (right) masked maps used for the Boötes SDWFS analysis. The mask used for the analysis removes 56% of the pixels in the mosaic.

The final flux cut for the mask was found iteratively by lowering the flux limit until further expansions of the masked regions no longer affect the final power spectrum. Since the PSF is slightly different for the $3.6\mu m$ and $4.5\mu m$ wavebands the two masks are independently constructed. Fig. S2 shows the final masked epoch 1 maps. The final mask is such that 56% of the original pixels are removed from the subsequent analysis.

3 Power Spectrum Estimation

With the final mosaics and masks in hand, initial, raw auto and cross-correlations may be computed to measure the level of clustering in the maps at each scale. To calculate the cross-correlation between masked maps M_1 and M_2 in real space, we first take the 2D Fourier transform of each map which we call $\tilde{M}_1$ and $\tilde{M}_2$, as

$$\tilde{M}[l_x, l_y] = \Delta \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} M[m, n] e^{-2\pi i(l_x m/M + l_y n/N)} \tag{1}$$

where M and N are the number of discrete points in the two dimensions of the map and Δ is the sampling interval in radians. The power spectra C_l formed from the cross-correlation of M_1 and M_2 for a specific l_i bin between between l-modes l_1 and l_2 is equal to the weighted mean of the squared Fourier modes $\tilde{M}_1 \tilde{M}_2^*$ between l_1 and l_2,

$$C_{l_i} = \frac{\sum_{l_x^2+l_y^2 \geq l_1}^{l_x^2+l_y^2 \leq l_2} w[l_x, l_y] \tilde{M}_1[l_x, l_y] \tilde{M}_2^*[l_x, l_y]}{\sum_{l_x^2+l_y^2 \geq l_1}^{l_x^2+l_y^2 \leq l_2} w[l_x, l_y]}, \tag{2}$$

where $w[l_x, l_y]$ is a window function in Fourier space that is non-zero for each mode of the analysis and zero for modes that are discarded. To compute the raw auto-correlation we have the special case where $M_1 = M_2$.

Being able to form a power spectrum from a cross-correlation rather than an auto-correlation is highly advantageous, as the noise bias and other contaminants that can dominate an auto-correlation calculation are minimized in a cross-correlation. This is because the pixel of each map $M_i = S_i + N_i$ is really a sum of the signal S_i plus noise N_i. The noise contributes to the auto-correlation such that $M_1 \times M_1 = S_1^2 + N_1^2$, but is minimized in a cross-correlation $M_1 \times M_2 = (S_1 + N_1) \times (S_2 + N_2) = S_1^2$, since $S_1 = S_2$. Another advantage to a cross-correlation study in this analysis is the availability of multi-epoch data over a four-year period. In such data any time varying signals that are not correlated between epochs cannot contribute to the background anisotropy power spectrum. One such possibility is the zodiacal light associated with scattered Sunlight off of dust particles. This is due to dynamical dust particles in near-Earth orbits. Furthermore with varying *Spitzer* orbit the lines of sights to the SDWFS Boötes field during the four epochs will also be different. Thus, we expect the zodiacal light contamination to have a time varying component that is not correlated between epochs. While autocorrelations in single epochs may be contaminated by zodiacal light, we expect cross-correlations to reduce the contamination.
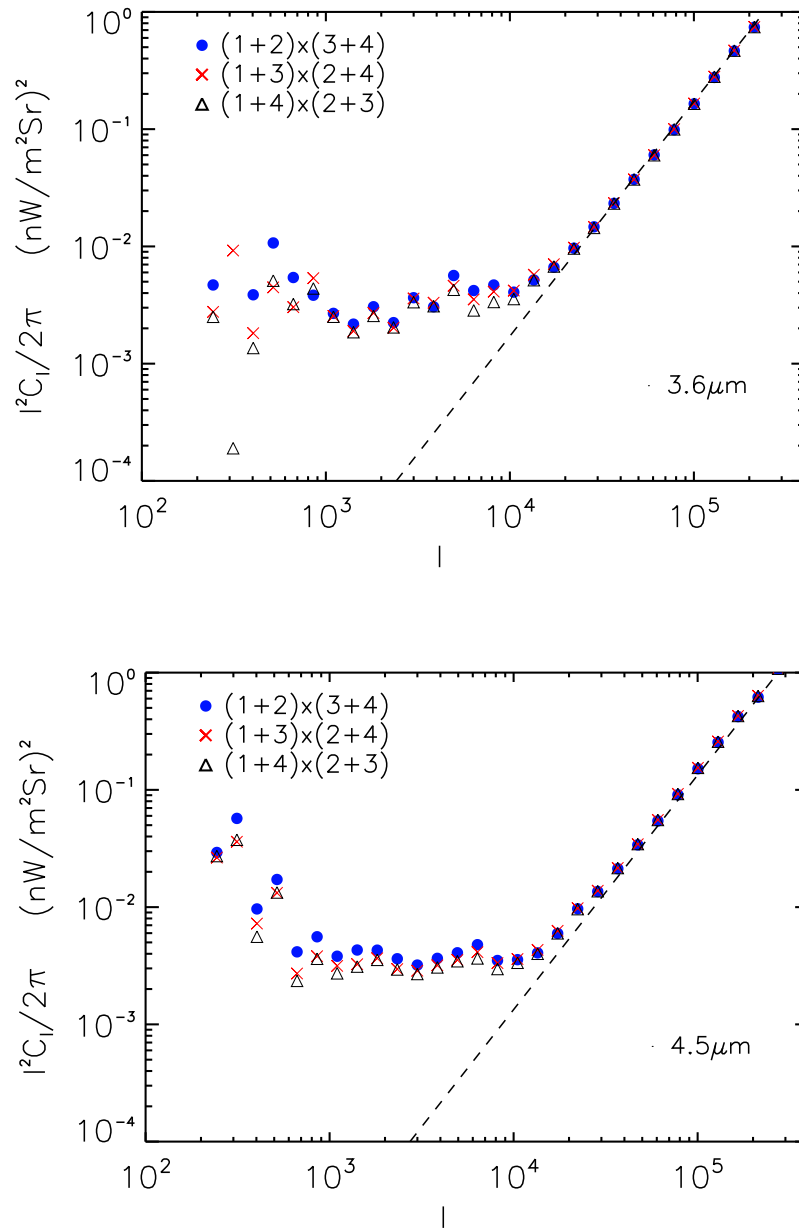
Figure S 3: **The cross power spectra of the sum of multi-epoch maps**. The cross-correlation power spectra of different epoch summed maps with 3.6 μm (top) and 4.5 μm (bottom) shown separately. The average of the summed maps are taken to be the power spectrum. The notation $(a+b) \times (c+d)$ indicates a cross correlation between the average of the $a+b$ and the $c+d$ epochs.

A previous analysis[36] showed that the contamination from zodiacal light spatial fluctuations is at least an order of magnitude below the background anisotropy level at 3.6 μm. Thus, zodiacal light
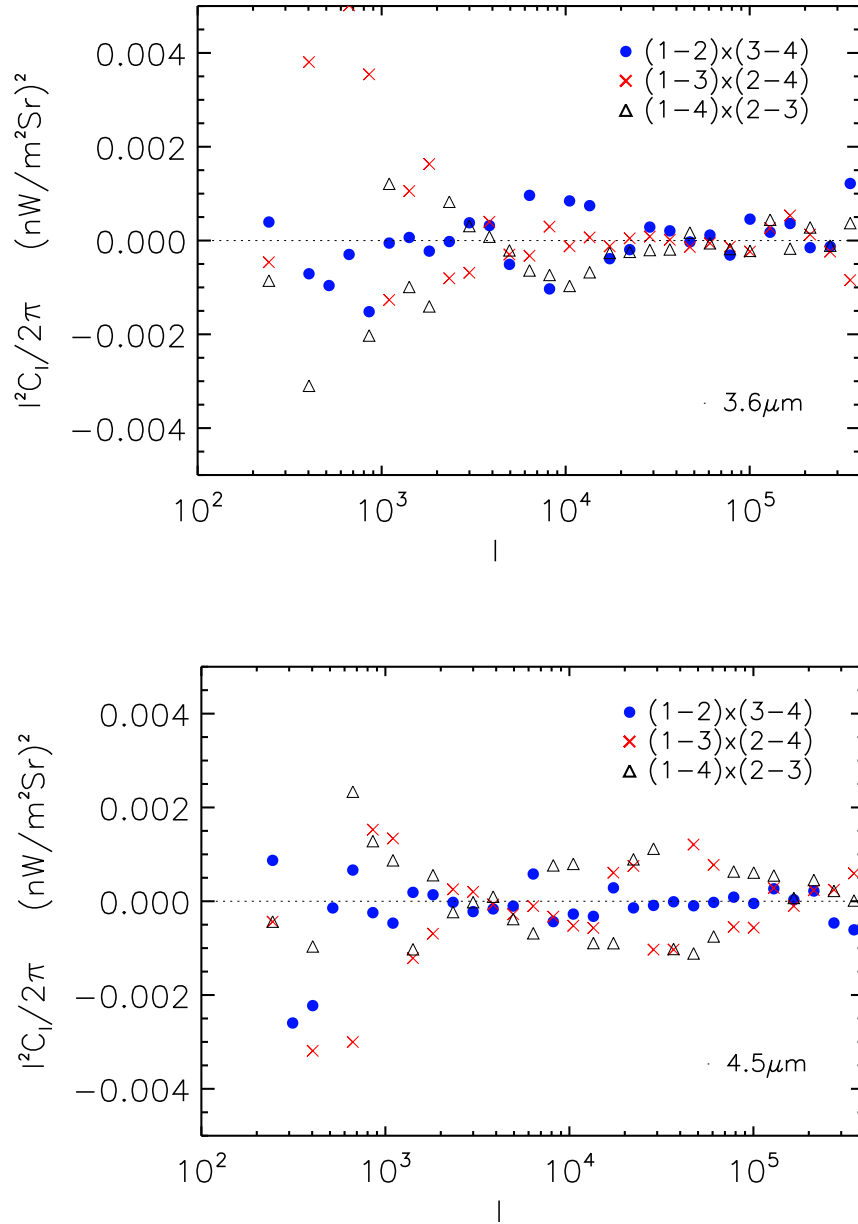
Figure S 4: **The cross power spectra of the difference of multi-epoch maps**. The cross-correlation power spectra of the difference of multi-epoch maps between epochs 1 to 4 with 3.6 μm (top) and 4.5 μm (bottom) shown separately. The cross-correlations are consistent with zero and the variance between the different cross-correlations provide one part of the final error budget associated with the power spectrum measurement.
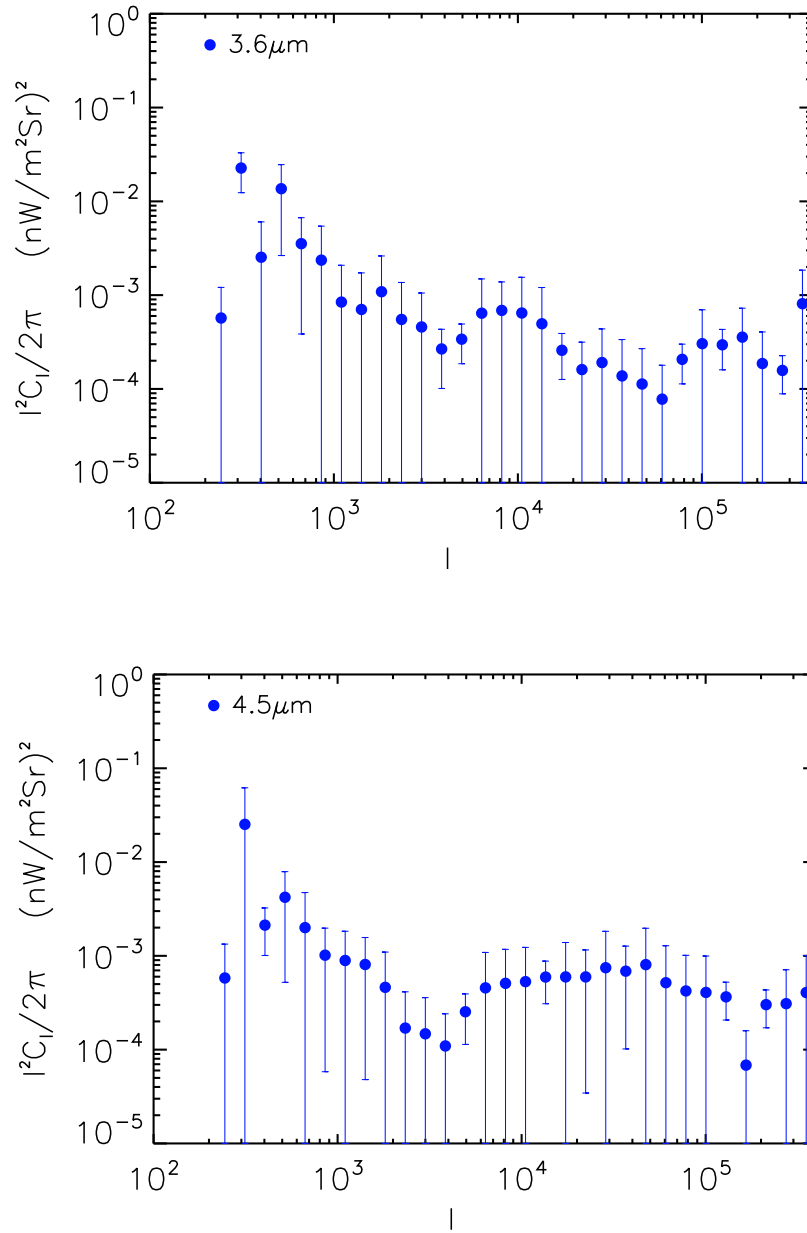
,

Figure S 5: **The noise floor to detect near-IR background anisotropies**. The mean and the variance (1σ) of the different-epoch difference maps cross-correlations. We show the absolute value of the mean and the error on the mean from the variance of the three independent cross-correlations.

should not be the dominant systematic effect in the present analysis. The cross-correlations using sum maps of epochs 1 to 4 are shown in Fig. S3.
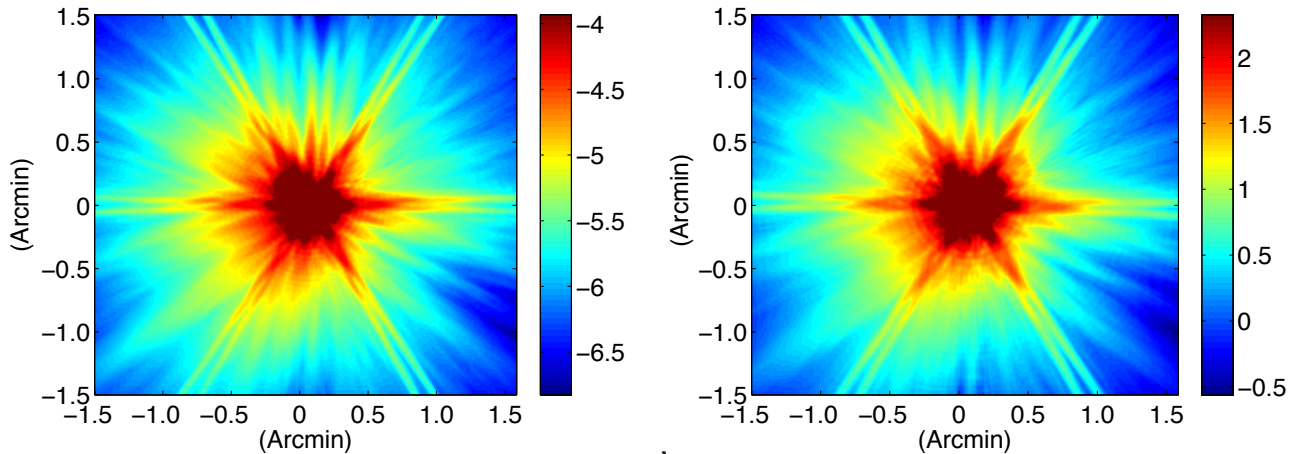
Figure S 6: **The IRAC PSF.** The 3.6 μm (left) and 4.5 μm (right) IRAC PSFs on a logarithmic intensity scale.

We can check the assumption that $N_1 \times N_2$ terms cancel by examining the cross-correlation between, say, $M_1 - M_2 = N_1 - N_2$ and $M_3 - M_4 = N_3 - N_4$. For differences of the same region the signal terms cancel and the amplitude of the cross-correlation $(M_1 - M_2) \times (M_3 - M_4)$ provides an estimate for the floor level at which noise contributions cancel. In Fig. S4 we show these differences, where we find that the cross-correlation power spectra between different epoch differences are consistent with zero. The variance of these difference cross-correlations provide a part of the error budget associated with noise correlations between different epochs including any systematic effects that are not canceled out in difference maps and are correlated between epochs. Since these spectra represent the noise floor the absolute value, we add them to the final error budget in quadrature with other uncertainties (Fig. S5).

Even after cross-correlation, the raw spectra are contaminated by several different sources and require additional corrections. These issues include resolution damping from the beam, the fictitious correlations introduced by the mask and shot noise. All three of these contaminants are dealt with as described below.

4   The Beam Correction

The first correction that needs to be applied to the raw power spectra is the correction for the beam. Realistic detectors have limits to their resolving power which causes a fictitious drop in power at high multipoles. For a known beam structure, the resolution limits of the instrument can be modeled in harmonic space with a function that encodes the full-width-half-max (FWHM) of the telescope. The resulting scale dependent function is known as the beam transfer function, b_l. For a
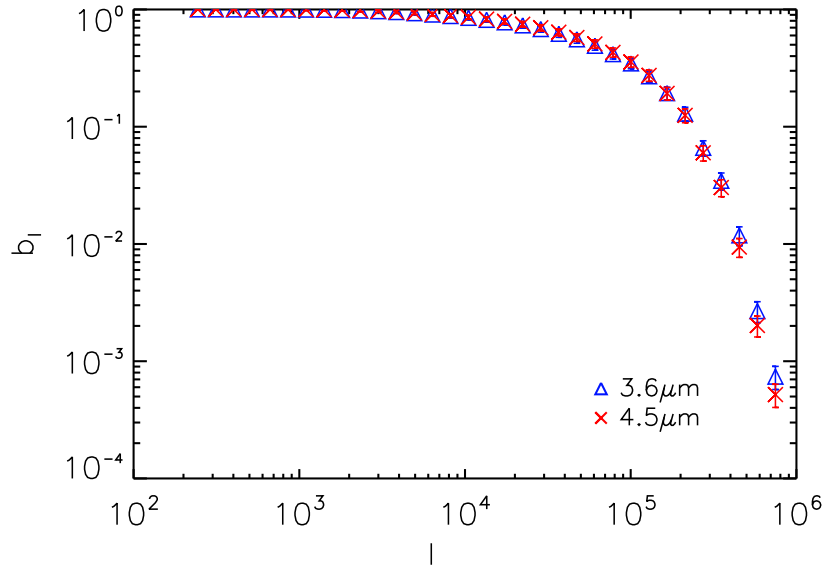
Figure S 7: **The IRAC beam function.** The beam transfer functions for the SDWFS Boötes field in the multipole space ℓ. We show the functions at 3.6 and 4.5 μm with triangles and crosses, respectively. The plotted error bars are the 1 σ uncertainties.

Gaussian beam it can be computed analytically as

$$b_l = \exp(l^2\sigma_{\mathrm{beam}}^2/2) \qquad and \qquad (3)$$

$$\sigma_{\mathrm{beam}} = \frac{\theta_{\mathrm{FWHM}}}{\sqrt{8\ln 2}}, \qquad (4)$$

where θ_{FWHM} encodes the full-width-half-max of the instrument's resolving power. The Spitzer team has measured $\theta_{\mathrm{FWHM}} = 1.9$ arcsec. The beam transfer function can also be computed directly from the data by measuring the point spread function. Each bright point source in the sky really shines a thin point like beam of light that should only illuminate a single pixel of the detector. However, due to the finite resolution of the telescope, the source is spread out over many pixels and often has a complex shape very different from a Gaussian. Fig. S6 shows this is true of the SDWFS PSF, and for this reason the beam transfer function needs to be calculated directly from the PSF.

In general, the beam transfer function b_l in Fourier or multipolar space is

$$b_l^2 = \frac{C_l^{M_{\mathrm{psf}}}}{C_l^{M_{\mathrm{point}}}} \qquad (5)$$
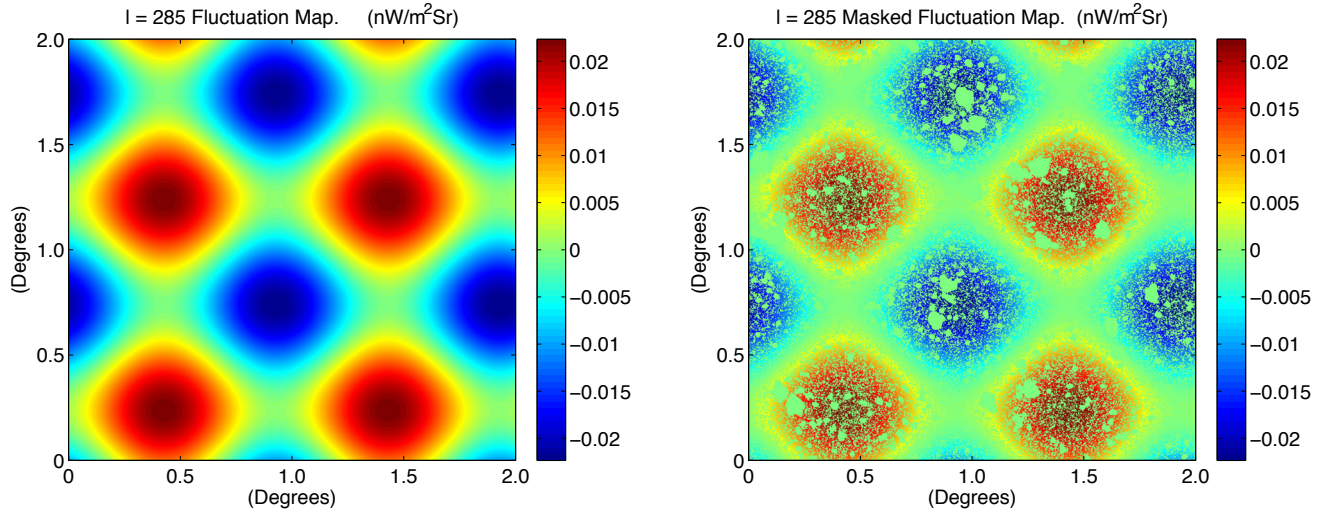
,

Figure S 8: **The masking effects in a map.** The left shows a map of a $l = 285$ fluctuation. On the right is the same map masked by one of the SDWFS masks. Note how what was only a large scale fluctuation gets broken up into smaller modes by the mask, contaminating the true power.

where $C_l^{M_{\mathrm{psf}}}$ is the power spectrum of M_{psf}, the observed image of a point source including the effects of the telescope, and $C_l^{M_{\mathrm{point}}}$ is the power spectrum of a true point source where all the light lies in one pixel of the map. Fig. S7 shows the beam calculated using Eq. 5 and the publicly available models of the *Spitzer*-IRAC PSF. The PSFs differ for each epoch and waveband and the appropriate beam transfer function for a cross-correlation between maps M_1 and M_2 becomes:

$$b_l^{1 \times 2} = \sqrt{b_l^1 b_l^2} \tag{6}$$

where b_l^1 and b_l^2 are the beam transfer functions for maps M_1 and M_2 respectively.

The beam corrected spectra C_l are then computed from the raw power spectrum C_l^{raw} by dividing by the beam transfer function b_l,

$$C_l = C_l^{raw}/b_l^2. \tag{7}$$

To measure the uncertainty in the beam transfer function we must understand the uncertainties in the PSF. The SDWFS team provides several measured PSFs taken across the Boötes images from which the uncertainty in b_l can be measured from Eq. 5 as $\delta b_l = \delta C_l^{M_{\mathrm{psf}}}/C_l^{M_{\mathrm{point}}}$ with $\delta C_l^{M_{\mathrm{psf}}}$ estimated from the variance of the differences from using the various PSF models.
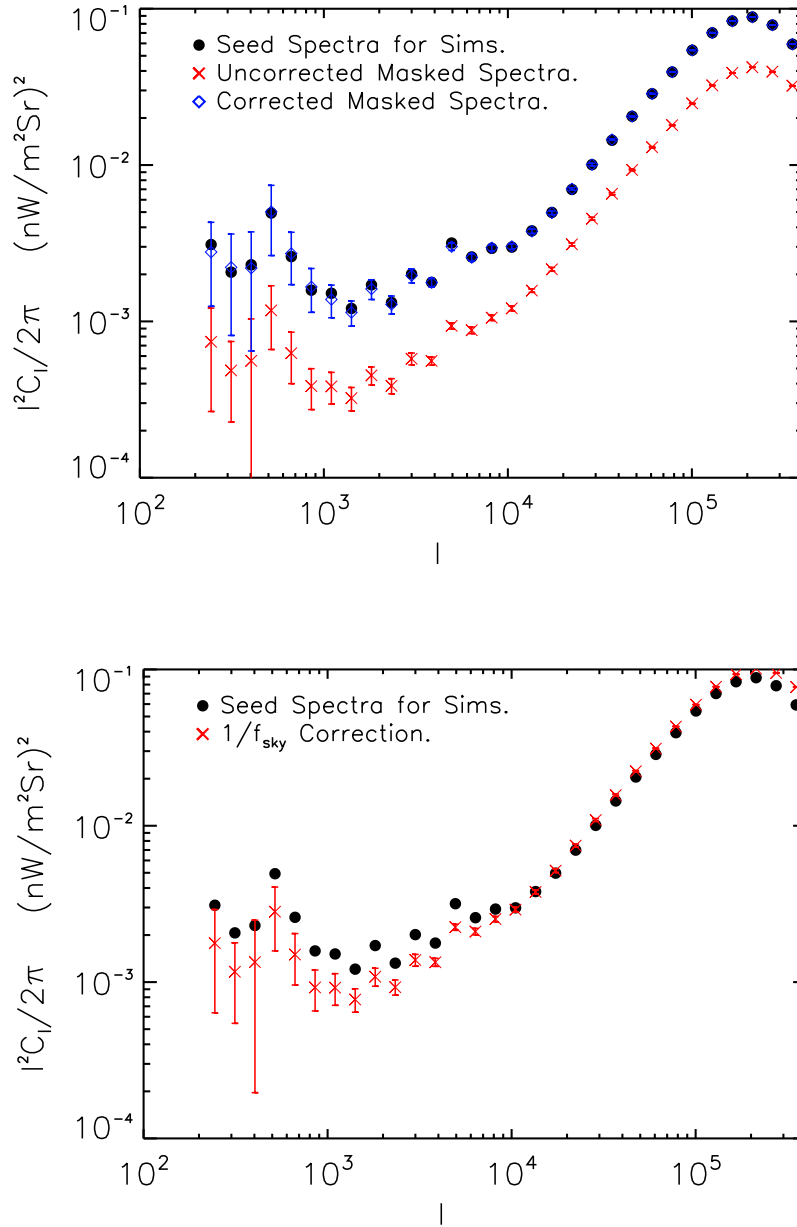
Figure S 9: **Masking effects on the power spectrum.** The top panel shows how well we can recover the true spectrum of a masked sky using the mode coupling matrix. The black points show the seeded spectra we used to create 100 simulations. The red shows what happens to the spectra after masking using the 3.6 μm mask. The blue points show what happens after correcting with the mode-coupling matrix. The bottom panel shows the result if corrected with only the masked sky fraction given by $f_{\rm sky}$. The mask breaks many large modes into smaller modes, so that after the f_{sky} correction the large-scale modes are under-represented and the small-scale modes are over represented. This illustrates the need to use the full coupling matrix to correct for the mask. The plotted error bars are the 1 σ uncertainties.

l_{eff}	$l^2 C_l / 2\pi$ (3.6 μm) [nW m^{-2} sr^{-1}]^2	$l^2 C_l / 2\pi$ (4.5 μm) [nW m^{-2} sr^{-1}]^2
243	$(0.27 \pm 0.36) \times 10^{-2}$	$(2.04 \pm 2.82) \times 10^{-2}$
313	$(0.52 \pm 0.58) \times 10^{-2}$	$(2.81 \pm 3.33) \times 10^{-2}$
402	$(0.18 \pm 0.24) \times 10^{-2}$	$(0.54 \pm 0.49) \times 10^{-2}$
517	$(0.51 \pm 0.50) \times 10^{-2}$	$(0.89 \pm 0.87) \times 10^{-2}$
665	$(0.32 \pm 0.26) \times 10^{-2}$	$(0.23 \pm 0.23) \times 10^{-2}$
854	$(0.43 \pm 0.42) \times 10^{-2}$	$(0.28 \pm 0.18) \times 10^{-2}$
1099	$(0.25 \pm 0.12) \times 10^{-2}$	$(0.23 \pm 0.12) \times 10^{-2}$
1412	$(0.18 \pm 0.11) \times 10^{-2}$	$(0.24 \pm 0.12) \times 10^{-2}$
1815	$(0.26 \pm 0.16) \times 10^{-2}$	$(0.27 \pm 0.08) \times 10^{-2}$
2332	$(0.19 \pm 0.09) \times 10^{-2}$	$(0.21 \pm 0.04) \times 10^{-2}$
2997	$(0.34 \pm 0.08) \times 10^{-2}$	$(0.20 \pm 0.03) \times 10^{-2}$
3851	$(0.29 \pm 0.05) \times 10^{-2}$	$(0.22 \pm 0.04) \times 10^{-2}$
4949	$(0.43 \pm 0.09) \times 10^{-2}$	$(0.24 \pm 0.05) \times 10^{-2}$
6360	$(0.32 \pm 0.13) \times 10^{-2}$	$(0.28 \pm 0.09) \times 10^{-2}$
8173	$(0.36 \pm 0.13) \times 10^{-2}$	$(0.22 \pm 0.08) \times 10^{-2}$
1.05×10^4	$(0.34 \pm 0.10) \times 10^{-2}$	$(0.22 \pm 0.09) \times 10^{-2}$
1.35×10^4	$(0.42 \pm 0.08) \times 10^{-2}$	$(0.25 \pm 0.10) \times 10^{-2}$
1.735×10^4	$(0.53 \pm 0.05) \times 10^{-2}$	$(0.35 \pm 0.11) \times 10^{-2}$
2.229×10^4	$(0.72 \pm 0.03) \times 10^{-2}$	$(0.54 \pm 0.12) \times 10^{-2}$
2.865×10^4	$(1.02 \pm 0.04) \times 10^{-2}$	$(0.71 \pm 0.16) \times 10^{-2}$
3.682×10^4	$(1.49 \pm 0.04) \times 10^{-2}$	$(1.03 \pm 0.17) \times 10^{-2}$
4.731×10^4	$(2.14 \pm 0.03) \times 10^{-2}$	$(1.48 \pm 0.19) \times 10^{-2}$
6.081×10^4	$(3.05 \pm 0.03) \times 10^{-2}$	$(2.10 \pm 0.15) \times 10^{-2}$
7.814×10^4	$(4.28 \pm 0.05) \times 10^{-2}$	$(2.97 \pm 0.13) \times 10^{-2}$
1.004×10^5	$(5.87 \pm 0.06) \times 10^{-2}$	$(4.11 \pm 0.14) \times 10^{-2}$
1.291×10^5	$(7.67 \pm 0.09) \times 10^{-2}$	$(5.27 \pm 0.12) \times 10^{-2}$
1.658×10^5	$(8.99 \pm 0.08) \times 10^{-2}$	$(6.15 \pm 0.06) \times 10^{-2}$
2.131×10^5	$(9.28 \pm 0.04) \times 10^{-2}$	$(5.92 \pm 0.12) \times 10^{-2}$
2.739×10^5	$(7.67 \pm 0.02) \times 10^{-2}$	$(4.84 \pm 0.08) \times 10^{-2}$
3.52×10^5	$(5.21 \pm 0.14) \times 10^{-2}$	$(3.06 \pm 0.07) \times 10^{-2}$
4.523×10^5	$(3.25 \pm 0.16) \times 10^{-2}$	$(1.70 \pm 0.15) \times 10^{-2}$

Table S 1: The final SDWFS power spectrum values $l^2 C_l / 2\pi$ for both the 3.6 μm and 4.5 μm bands. The quoted error is the 1 σ uncertainty of the final power spectrum.

5 The Mode Coupling Correction.

Fictitious correlations introduced by the mask must be corrected. When an image is masked, the sources are replaced by the value zero in the image. When the power spectrum in computed, these zeros in real space make fictitious contributions to the two-dimensional Fourier transform that are then added to the final power spectrum. This can be easily seen in Fig. S9. On the top we see an unmasked fluctuation pattern for a specific l-mode. After the mask is applied, this fluctuation gets broken up into fluctuations of different sizes causing both a diminishing and a reshuffling of power in Fourier space.

A matrix correction method exists[37] to model the effects of the mask on the power spectrum by using a matrix $M_{ll'}$ whose inverse removes the effects of the mask from the measurement by matrix multiplication. If $\tilde{C}_l$ is the masked sky power spectrum and C_l be the true power spectrum, the relation between the true and masked sky spectrum is

$$\tilde{C}_l = M_{ll'} C_{l'} \,, \tag{8}$$

where Einstein summation notation is being used. Since this relationship is matrix multiplication, the masking effects can be removed from the masked power spectrum by simply using a matrix multiplication $M_{ll'}^{-1} \tilde{C}_l$ to recover the true power spectrum C_l. In the limit of no $l-$mode coupling, $M_{ll'} = f_{\text{sky}}$ where f_{sky} is the fraction of the masked map that is non-zero.

Calculating the mode coupling matrix $M_{ll'}$ analytically[37] is computationally expensive for large maps. For this reason, we developed a new way to generate the mode-coupling matrix as follows:

1. For each ℓ in the power spectrum create many realizations of maps consisting of a pure tone where $C_m = 1$ if $\ell = m$ and $C_m = 0$ otherwise (an example case is shown in Fig. S9).

2. For each of these trial maps, mask the maps and calculate an observed power spectrum $\tilde{C}_m(\ell)$.

3. The mode coupling matrix $M_{\ell m} = \left\langle \tilde{C}_m(\ell) \right\rangle$ is the average of the masked power spectra found for the random realizations of model ℓ. The inverse of the mode coupling matrix gives the sky power spectrum corrected for the masking effect, $C_l = M_{lm}^{-1} \tilde{C}_m$.

To see how well this this works, consider Fig. S9. The black line shows the exact power spectrum from which we drew 100 simulated images. The red points show the observed power spectra of the masked sky. The blue points show those same 100 power spectra after correcting with the mode coupling matrix described above. This mode-coupling transformation does an excellent job recovering the input power spectra of an unmasked sky. As the lower panel of Fig. S9 shows using the simplified model based only on the masked sky fraction f_{sky} is not a good approximation.
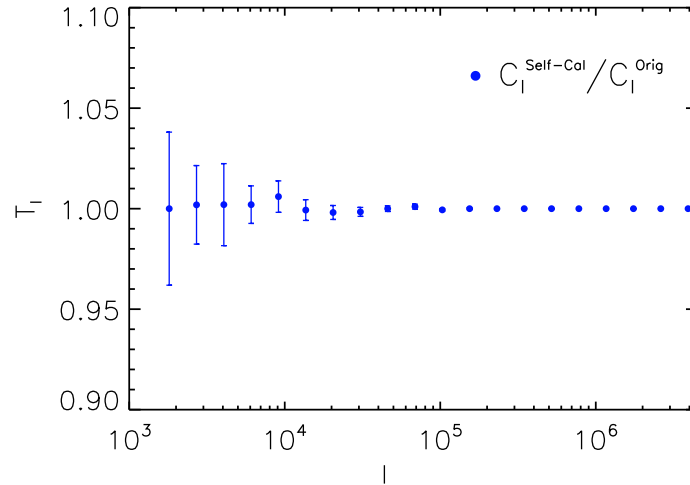
Figure S 10: **The map-making transfer function.** *Spitzer*-IRAC mosaic transfer function based on the self-calibration algorithm used to make data maps for this study. We show the transfer function at 3.6 μm. The 4.5 μm transfer function is similar to the result at 3.6 μm. The plotted error bars are the 1 σ uncertainties.

To estimate the effects of cosmic variance on the power spectrum one must make multiple Gaussian simulations with the exact power spectrum measured in the data. Without applying a mask, these simulations will have a cosmic variance:

$$\delta C_l^{\mathrm{CV}} = \sqrt{\frac{2}{(2l+1)\delta l f_{\mathrm{sky}}}} C_l \tag{9}$$

where δC_l^{CV} is the cosmic variance for multipole l, δl is the width of that l bin, f_{sky} is the fraction of the total sky covered by the unmasked region, and l in the denominator is the mid-point of the l bin. It should be clear that masks should increase cosmic variance because they further reduce the coverage of the unmasked sky. The mode-coupling matrix does not restore errors from cosmic variance and therefore, masking and correcting with $(M_{ll'})^{-1}$ will combine errors from the mode-coupling matrix as well as the reduced coverage from the mask and thus have a greater variance then cosmic variance alone.

6 The Map-making Transfer Function

The last correction that must be accounted for is to correct for the effects of the map-making procedure in producing the final mosaic. The mosaicing procedure is often combines many images and between dealing with overlap regions as well adding more tiles to a mosaic, the addition or subtraction of total power must be quantified and corrected. The best way to uncover the effects

of the mosaicing procedure on the power spectrum is to take a map with a known power spectrum, re-build the map using your mosaicing procedure and compare the output and input power spectra. More specifically to build the map-making transfer function we followed the following procedure:

1. We created a large map similar in size, pixel scale and astrometry to our SDWFS data mosaic, but assuming a shape for the power spectrum (we started with pure white noise for which $C_l = $ constant).

2. We break the map into the tiles similar in size, pixel scale and astrometry to our original data tiles. We added instrumental noise to the tiles consistent with the noise of the different epochs.

3. We ran these simulated signal plus noise tiles through the same self-calibration mosaicing procedure described in Section 1 to produce a new map identical to the original map modulo the mosaicing effects.

4. The map-making transfer function is then $T_l = C_l^{\mathrm{orig}}/C_l^{\mathrm{mosiaced}}$, where C_l^{orig} is the known power spectrum of the original simulation and C_l^{mosiaced} is the power of the final map.

5. We repeated Steps 1 to 4 above for different initial power spectra to test if the transfer function remains the same or is different. We found that the transfer function is independent of the assumption for the input power spectrum shape or the amplitude.

For our analysis we used this process with the self-calibration algorithm for the mosaic transfer function. Simulated maps of pure white noise broken into tiles and remosaiced using the self-calibration algorithm to determine the transfer functions. An example transfer function for SDWFS is given in Fig. S10. This was obtained by simulating 10^3 independent maps and using the mean and the standard deviation of $(T_l)_i$, where i denotes each simulation, to determine the best-determined transfer function and its error.

Given the transfer function T_l, beam correction b_l, the mode-coupling matrix $M_{ll'}$ the final power spectrum is estimated as

$$C_l = M_{ll'}^{-1}T_l'\tilde{C}_{l'}/b_{l'}^2 \tag{10}$$

where $\tilde{C}_{l'}$ is the raw power spectrum after masking out the foreground sources and C_l is the final corrected power spectrum. C_l is the power spectrum that we present in the main paper and compared to previous results.

7 Final Power Spectrum Results

To compute the power spectrum for the infrared background at 3.6 μm we used all four epochs of the SDWFS data as described above. In order to minimize contamination from instrumental noise, zodiacal light and other systematics we used the cross-correlation $1/2(E_1 + E_2) \times 1/2(E_3 + E_4)$

and the two additional permutations by switching the epochs, where $E_1..E_4$ correspond to the four epochs. Masks were generated to remove the foreground and applied as described above before the cross-correlations were taken. We used the same cross-correlation procedure to obtain the 4.5 μm power spectrum. After the raw spectra are obtained from the mosaics, they are corrected for the beam, mode-coupling, mosaicing (or map-making) effects as described above using Equation 10.

Fig. S11 shows the results from the *Spitzer* SDWFS Boötes field (the power spectra values are listed in Table 1). Results from a recent analysis[12] are also given for reference. These spectra are the final spectra after all corrections have been applied. We note the strong agreement between our measurements and the previously published ones. The difference at small angular scales, high ℓ, is due to differences in the depth of the mask. It is captured by a difference of the shot-noise levels in the point source detection level between SDWFS and deeper SEDS data[12].

Fig. S12 shows the angular cross-power spectrum of near-IR anisotropies measured with SDWFS at 3.6 and 4.5 μm. The left panel shows the cross power spectrum ($C_l^{3.6-4.5}$) between the two channels, while the right panel shows the correlation coefficient calculated as

$$r = C_l^{3.6-4.5}/\sqrt{C_l^{3.6}C_l^{4.5}}, \tag{11}$$

where $C_l^{3.6}$ and $C_l^{4.5}$ are the auto power spectra at 3.6 and 4.5 μm, respectively (Fig. S11). The correlation coefficient is consistent with unity. The errors are the 1σ overall uncertainty in the correlation coefficients found by propagating errors on the cross power spectrum and auto power spectra through equation 11.

8 Theoretical Interpretation of Near-IR Anisotropies as Spatial Fluctuations of Integrated Intrahalo Light

Our intrahalo light (IHL) model presented in the *Letter* is described in this Section. In Fig. 1 of the main *Letter*, we also show results from two descriptions related to the near-IR background anisotropies. One involves the faint galaxies that fall below the magnitude cut-off of the masks that are applied to measure fluctuations. The contribution from faint galaxies, primarily dwarf has been studied in detail with latest information on the faint-end of the galaxy luminosity functions[7] and we use their results in Fig. 1. A second model involves the $z > 6$ galaxy contribution. The shaded region for $z > 6$ galaxies in Fig. 1 combines the analytical models[19] with results from numerical simulations[20]. The predictions are normalized to the measured luminosity functions of galaxies at $z > 6$ and uses a reionization history that is consistent with the WMAP 7-year optical depth to electron scattering[19].

For the interpretation presented in the *Letter*, we model the IHL intensity angular power spectrum using the halo model. The IHL model differs from galaxy clustering models in that we assign a profile to the diffuse stars. The standard galaxy clustering models assume a central galaxy at the halo center and satellites that are distributed randomly in the halo with a profile that is tracing the dark matter distribution. We now add a diffuse extended component in addition to the central
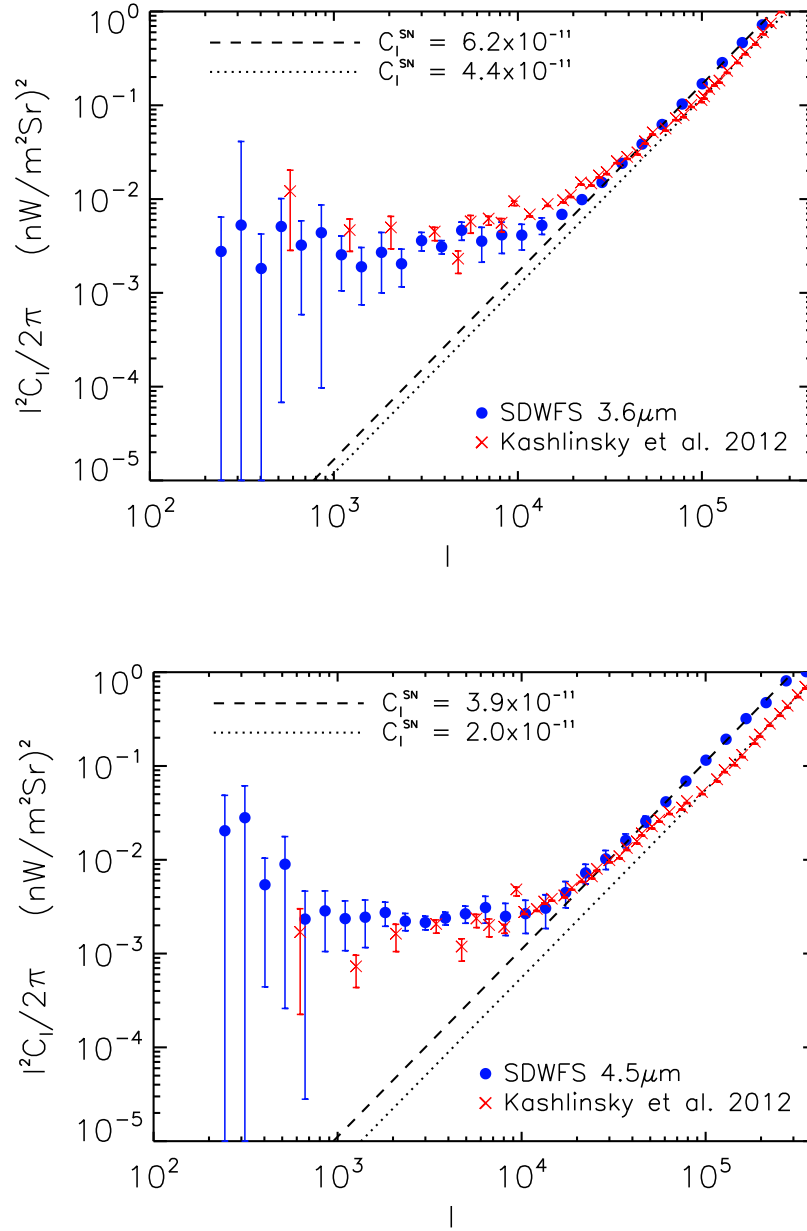
Figure S 11: **The angular power spectrum of near-IR anisotropies.** The angular power spectrum of near-IR anisotropies measured with SDWFS at 3.6 and 4.5 μm. The 1 σ error bars include all uncertainties we have discussed in the Supplement and the measurements are beam corrected. We also compare our measurements to existing results[12] where we find a general agreement on clustering. The large-ℓ difference between the two datasets reflect the depth of the point source identification and removal in the mask.
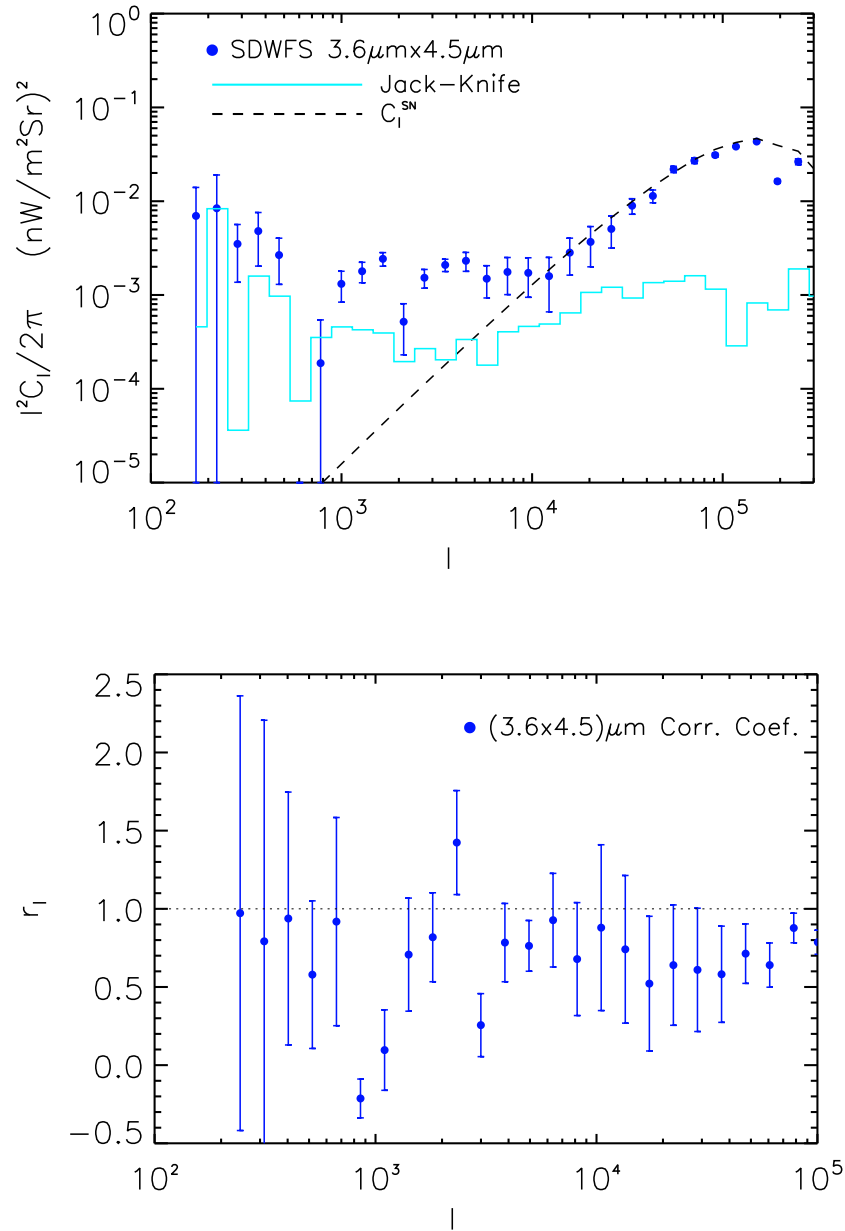
Figure S 12: **The angular cross-power spectrum of near-IR anisotropies.** The angular cross-power spectrum of near-IR anisotropies measured with SDWFS at 3.6 and 4.5 μm. The 1 σ error bars include all uncertainties we have discussed in the Supplement and the measurements are beam corrected. The upper panel shows the cross power spectrum ($C_l^{3.6-4.5}$), while the lower panel shows the correlation coefficient calculated as $r = C_l^{3.6-4.5}/\sqrt{C_l^{3.6}C_l^{4.5}}$, where $C_l^{3.6}$ and $C_l^{4.5}$ are the auto power spectra at 3.6 and 4.5 μm, respectively (Fig. S11)
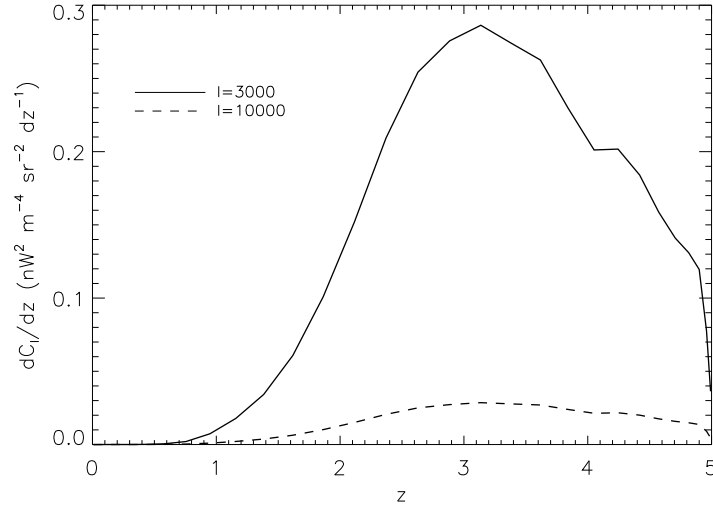
Figure S 13: **The redshift dependence of the IHL anisotropy power spectrum.** dC_l/dz as a function of redshift for $\ell = 3 \times 10^3$ and 10^4. The majority of near-IR anisotropies originate from $1 < z < 4$.

and satellite galaxies. The halo number density[43] as a function of redshift and mass $dn(M,z)/dM$ is

$$\frac{dn}{d\ln M} = \frac{\rho_m}{M} f(\nu) \frac{d\nu}{d\ln M} \,, \tag{12}$$

with

$$\nu f(\nu) = A\sqrt{\frac{2}{\pi} a\nu^2} [1 + (a\nu^2)^{-p}] \exp[-a\nu^2/2] \,. \tag{13}$$

Given a galaxy luminosity-halo mass relation, a certain fraction $f_{ICL}(M)$ of this luminosity will be in the form of IHL. The IHL luminosity-mass relation is then:

$$L_{IHL,\lambda}(M,z) = f_{IHL}(M)L(M,z=0)(1+z)^\alpha f_\lambda(\lambda/(1+z)) \,, \tag{14}$$

where α is the power-law index that accounts for a possible redshift evolution and $f_\lambda(\lambda/(1+z))$ is the spectral energy distribution of the IHL (see also discussion below). We model the fraction of total luminosity in form of IHL as a power-law in halo mass,

$$f_{IHL}(M) = A_f \left(\frac{M}{10^{12} M_\odot}\right)^\beta \,. \tag{15}$$

The total luminosity as a function of halo-mass $L(M, z = 0)$ at $z = 0$ is taken to be the

best-fit relation from[46] at 2.2 μm

$$L(M, z = 0) = 5.64 \cdot 10^{12} h_{70}^{-2} \left(\frac{M}{2.7 \cdot 10^{14} h_{70}^{-1} M_\odot} \right)^{0.72} L_\odot \,. \tag{16}$$

This is measured for galaxy groups and clusters. At smaller mass scales one no longer has the issue of multiple galaxies in a halo and the total luminosity is simply that of the central galaxy[47]. We extend it to lower masses using the same power-law slope since analyses of the total luminosity-halo mass relation using galaxy-galaxy lensing find the slope continues down to mass scales below 10^{11} $M_\odot$ [48,49]. Given that this total luminosity-halo mass relation is measured at 2.2 μm, we scale this to other wavelengths with the SED f_λ, but with the normalization that $f_\lambda = 1$ at 2.2 μm at $z = 0$. We consider the SED of IHL to be consistent with that of old elliptical galaxies comprised of old, red stars[44]. In the main *Letter*, in Fig. 3, we also consider alternatives for the SED using a variety of galaxy SED templates.

Under these assumptions the angular power spectrum of the IHL flux fluctuations can be written as the sum of a 1-halo term, that originates from small-scale fluctuations within individual halos

$$C_\ell^{1h} = \frac{1}{(4\pi)^2} \int dV \frac{1}{(1+z)^2 \chi^4(z)} \int_{M_{min}}^{M_{max}} dM \frac{dn(M, z)}{dM} u_{\mathrm{IHL}}^2(k|M) L_{IHL, \lambda}^2(M, z) \,, \tag{17}$$

and a 2-halo term, related to the large scales dark matter fluctuations, and hence to the linear dark matter power spectrum $P_M(k, z)$ as

$$C_\ell^{2h} = \frac{1}{(4\pi)^2} \int dV \frac{1}{(1+z)^2 \chi^4(z)} \left[\int_{M_{min}}^{M_{max}} dM \frac{dn(M, z)}{dM} u_{\mathrm{IHL}}(k|M) b_h(M, z) L_{IHL, \lambda}(M, z) \right]^2$$
$$\times P_M(k = \ell/\chi(z), z) \,, \tag{18}$$

where $u_{\mathrm{IHL}}(k, z|M)$ is the Fourier transform of the IHL profile in a dark matter halo of mass M at redshift z, $b_h(M)$ is the linear bias, $\chi(z)$ is the comoving radial distance, and dV is the comoving volume element $dV = \chi(z)^2 d\chi/dz$. The redshift integration is performed up to a maximum redshift $z_{\max} = 5$. We found that integrating to a higher redshift did not change our results. The values of M_{min} and M_{max} in Eqs. 17 and 18 determine the relative amplitude of the 1-halo and 2-halo terms and we let them vary freely. The power spectrum also contains a shot-noise contribution from unresolved fluctuations, so that the total power spectrum is

$$C_\ell = C_\ell^{1h} + C_\ell^{2h} + C_\ell^{\mathrm{SN}} \,. \tag{19}$$

We also take C_ℓ^{SN} to be a free parameter that is varied during our model fitting.

Since the IHL profile for small mass halos has yet to be determined precisely, we consider two model descriptions under the assumption that IHL (i) traces the Navarro-Frenk-White (NFW) profile of dark matter halos[50] and (ii) falls as r^{-2} with an exponential cut-off[51] such that $\rho_{\mathrm{IHL}} \propto 1/r^2 \exp(-r/2R_{\mathrm{vir}})$. There is limited information on the light profile from the stacking analysis

of SDSS galaxies[52]. However, we are unable to use those measurements for the current study as we do not have information on how the profile changes with the halo mass. Moreover given the limited information both in terms of the angular scale of fluctuations and the large uncertainties we find that we are not able to statistically distinguish one IHL profile over another.

In Fig. 1 of the main *Letter*, the best-fit anisotropy power spectrum makes use of the description involving the NFW profile with

$$\rho(r) = \frac{\rho_s}{(cr/r_{\mathrm{vir}})(1 + cr/r_{\mathrm{vir}})^2} \, , \tag{20}$$

where r_{vir} is the virial radius and c the halo concentration parameter. We define the concentration using the result from numerical simulations[39] that find

$$c(M, z) \;=\; \frac{9}{1+z} \left(\frac{M}{M_*} \right)^{-0.13} , \tag{21}$$

where M_* is the mass scale at which the critical density contrast δ_c required for spherical collapse is equal to the square root of the variance in the initial density field $\sigma(M_*) = \delta_c$. While we make use of this particular fitting function[39], an alternative fitting function[40] led to the same results. Thus, our best-fit parameter values are independent of the assumption on the halo mass-concentration relation.

To analyze the data we conducted a Monte Carlo-Markov Chain (MCMC) analysis using a modified version of the publicly available package cosmoMC[45] with convergence diagnostic based on the Gelman-Rubin statistic[41]. We fit a total of six free parameters: the minimum and maximum masses (M_{min}, M_{max}), the power-law index of the redshift dependence in the luminosity-mass relation α, the amplitude and the power-law index with halo mass of the IHL fraction A_f and β, respectively, and the shot noise contribution C_ℓ^{SN}. We fix the cosmological parameters to the best-fit for the ΛCDM concordance model from WMAP 7-year analysis[42].

The best-fit parameters for the fit to the data at 3.6 μm from SDWFS are shown in Table 6. The minimum and maximum halo masses of the halo mass range contributing to measured near-IR anisotropies is $10^{9.03\pm0.05}$ M$_\odot$ and $10^{11.91\pm0.05}$ M$_\odot$, respectively. Our model-fitting suggests a small dependence of the IHL fraction on halo mass with a power-law slope of $\beta = 0.09\pm0.01$. We found values consistent with the 1σ uncertainties when using the alternate IHL profile described above and fitting the model to the 4.5 μm power spectrum measurements. Our results for $f_{\mathrm{IHL}}(M)$ are summarized in Fig. 2 of the main *Letter*. There we also compare our determination to an analytical prediction of the IHL fraction relative to the total luminosity, in the literature[14]. These analytical models exist only down to a halo mass scale of about 5×10^{10} M$_\odot$ and $f_{\mathrm{IHL}} > 5 \times 10^{-4}$. In addition to the power-law behavior it was found also found in previous analytical work[13] that the IHL fraction is constant at a value of about 0.005 when $M < 5 \times 10^{11}$ M$_\odot$ in some scenarios to generate IHL. This flattening behavior could be related to our observation that the IHL fraction is not strongly halo mass dependent over the mass ranges that we are probing with near-IR background anisotropy power spectrum. More detailed studies are necessary to properly

understand how our results can be used to understand the merger rate and generation of IHL in low mass halos at redshifts of 1 to 4.

In Fig. S13 we summarize the redshift dependence of the IHL power spectrum by calculating dC_l/dz as a function of redshift. The contributions peak at a redshift of 3, but has a broad distribution ranging from $1 < z < 4$. The measured shot-noise level of 120 ± 10 nJy nW m^{-2} sr^{-1} is about a factor of 2 higher than the shot-noise in the deeper SEDS data[12] with a value of ~ 57 nJy nW m^{-2} sr^{-1}.

In Fig. 3 of the main *Letter*, we summarize the SED of IHL. Here, we make use of a variety of stellar templates from B to K-type stars for this purpose. The prediction converges at the longer wavelengths due to existing measurements but we find large deviations at 1 μm and shorter wavelengths. A measurement of the background anisotropy at optical wavelengths is clearly desirable and could be used to both identify the stars that are primarily contributing to IHL at $z \sim$1 to 4.

9   References

31. Eisenhardt, P. R. et al. The Infrared Array Camera (IRAC) Shallow Survey, Astrophys. J. Suppl. 154, 48-53 (2004).

32. Makovoz, D. & Khan, I. 2005, Astronomical Data Analysis Software and Systems, Astronomical Society of Pacific Conference Series 132, Astronomical Data Analysis Software and Systems VI, edited by P. L. Shopbell, M. C. Britton, & R. Ebert (San Francisco: ASP).

33. Fixsen, D. J., Moseley, S. H. & Arendt, R. G. Calibrating Array Detectors, Astrophys. J. Suppl. 128, 651 (2000).

34. Quijada, M. A., Marx, C. T., Arendt, R. G. & Moseley, S. H. 2004, Angle-of-incidence effects in the spectral performance of the infrared array camera of the Spitzer Space Telescope, Optical, Infrared, and Millimeter Space Telescopes. Edited by Mather, John C. Proceedings of the SPIE, 5487, 244-252 (2004).

35. Bertin, E. & Arnouts, S. SExtractor: Software for source extraction. Astron. and Astroph. Supp. 317, 393 (1996).

36. Arendt, R. G., Kashlinsky, A., Moseley, S. H. & Mather, J. Cosmic Infrared Background Fluctuations in Deep Spitzer IRAC Images: Data Processing and Analysis, Astrophys. J. Suppl. 186, 10-47 (2010).

37. Hivon, E. et al. MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets, Astrophys. J. 567, 2-17 (2002).

38. Sullivan, I. et al. Clustering of the IR Background Light with Spitzer: Contribution from Resolved Sources, Astrophys. J. 657, 37-50 (2007).

39. Bullock J. S. et al. Profiles of dark haloes: evolution, scatter and environment, Mon. Not. R. Astron. Soc., 321, 559-575 (2001).

40. Duffy, A. R., Schaye, J., Kay, S. T. & Della Vecchia, C. 2008, Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology, Mon. Not. of the Roy. Astron. Soc., 390, L64-L68 (2008).

41. Gelman A. & Rubin. D. B. Inference from Iterative Simulation Using Multiple Sequences, Statist. Sci. 7, 457-472 (1992).

42. Komatsu E. et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation, Astrophys. J. Suppl. 192, 18 (2011).

43. Sheth R. K. & Tormen G. Large-scale bias and the peak background split, Mon. Not. R. Astron. Soc., 308, 119-126 (1999).

44. Krick J. E. & Bernstein R. A. Diffuse Optical Light in Galaxy Clusters. II. Correlations with Cluster Properties, Astrophys. J. 134, 466-493 (2007).

45. Lewis A. & Bridle S. Cosmological parameters from CMB and other data: A Monte Carlo approach, Phys. Rev. D 66, 103511 (2002).

46. Lin Y. -T. , Mohr J. J. & Stanford S. A. K-Band Properties of Galaxy Clusters and Groups: Luminosity Function, Radial Distribution, and Halo Occupation Number, Astrophys. J. 610, 745-761 (2004).

47. Vale, A. & Ostriker, J. P. Linking halo mass to galaxy luminosity, Mon. Not. of the Roy. Astron. Soc., 353, 189-200 (2004).

48. Cooray, A. & Milosavljević, M. Dissipationless Merging and the Assembly of Central Galaxies, Astrophys. J. 627, L85-L88 (2005).

49. Yang, X., Mo, H. J., van den Bosch, F. C. Constraining galaxy formation and cosmology with the conditional luminosity function of galaxies, Mon. Not. of the Roy. Astron. Soc., 339, 1057-1080 (2003).

50. Navarro J. F., Frenk C. S. & White S. D. M. A Universal Density Profile from Hierarchical Clustering, Astrophys. J., 490, 493 (1997)

51. Masaki, S. & Yoshida, N. Distribution of Dust around Galaxies: An Analytic Model, arXiv.org:1203.6414 (2012).

52. Tal, T. & Dokkum, P. The faint stellar halos of massive red galaxies from stacks of more than 42000 SDSS LRG images, Astrophys. J. 731, 89 (2011).

A_f	0.0015 ± 0.0002
$\log(M_{\min}/M_\odot)$	9.03 ± 0.05
$\log(M_{\max}/M_\odot)$	11.91 ± 0.05
β	0.094 ± 0.005
α	1.23 ± 0.09
C_ℓ^{SN} (nW^2 m^{-4} sr^{-1})	$(9.8 \pm 0.5) \times 10^{-11}$

Table S 2: The best-fit parameter values of the IHL anisotropy power spectrum model to the 3.6 μm data using MCMC model fits. The quoted error bars are the 1σ uncertainties for each of the parameter likelihoods marginalizing over other parameters.