Title
Seeing in the Dark: Weak Lensing from the Sloan Digital Sky Survey

Author
Huff, Eric Michael

Publication Date
2012

Peer reviewed|Thesis/dissertation

Seeing in the Dark: Weak Lensing from the Sloan Digital Sky Survey

by

Eric Michael Huff

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Astrophysics

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:
Professor Uros Seljak, Chair
Doctor David Schlegel
Professor Saul Perlmutter
Professor Martin White

Fall 2012

Seeing in the Dark: Weak Lensing from the Sloan Digital Sky Survey

Abstract

Seeing in the Dark: Weak Lensing from the Sloan Digital Sky Survey

by

Eric Michael Huff
Doctor of Philosophy in Astrophysics

University of California, Berkeley

Professor Uros Seljak, Chair

Statistical weak lensing by large-scale structure – cosmic shear – is a promising cosmological tool, which has motivated the design of several large upcoming astronomical surveys. This Thesis presents a measurement of cosmic shear using coadded Sloan Digital Sky Survey (SDSS) imaging in 168 square degrees of the equatorial region, with $r < 23.5$ and $i < 22.5$, a source number density of 2.2 per arcmin^2 and median redshift of $z_{\rm med} = 0.52$. These coadds were generated using a new rounding kernel method that was intended to minimize systematic errors in the lensing measurement due to coherent PSF anisotropies that are otherwise prevalent in the SDSS imaging data. Measurements of cosmic shear out to angular separations of 2 degrees are presented, along with systematics tests of the catalog generation and shear measurement steps that demonstrate that these results are dominated by statistical rather than systematic errors. Assuming a cosmological model corresponding to WMAP7 (Komatsu et al., 2011) and allowing only the amplitude of matter fluctuations σ_8 to vary, the best-fit value of the amplitude of matter fluctuations is $\sigma_8 = 0.636^{+0.109}_{-0.154}$ (1σ); without systematic errors this would be $\sigma_8 = 0.636^{+0.099}_{-0.137}$ (1σ). Assuming a flat ΛCDM model, the combined constraints with WMAP7 are $\sigma_8 = 0.784^{+0.028}_{-0.026}(1\sigma)^{+0.055}_{-0.054}(2\sigma)$ and $\Omega_m h^2 = 0.1303^{+0.0047}_{-0.0048}(1\sigma)^{+0.009}_{-0.009}(2\sigma)$; the 2σ error ranges are respectively 14 and 17 per cent smaller than WMAP7 alone. Aside from the intrinsic value of such cosmological constraints from the growth of structure, some important lessons are identified for upcoming surveys that may face similar issues when combining multi-epoch data to measure cosmic shear.

Motivated by the challenges faced in the cosmic shear measurement, two new lensing probes are suggested for increasing the available weak lensing signal. Both use galaxy scaling relations to control for scatter in lensing observables.

The first employs a version of the well-known fundamental plane relation for early type galaxies. This modified "photometric fundamental plane" replaces velocity dispersions with photometric galaxy properties, thus obviating the need for spectroscopic data. We present the first detection of magnification using this method by applying it to photometric catalogs from the Sloan Digital Sky Survey. This analysis shows that the derived magnification signal is comparable to that available from conventional methods using gravitational shear. We suppress the dominant sources of systematic error and discuss modest improvements that

may allow this method to equal or even surpass the signal-to-noise achievable with shear. Moreover, some of the dominant sources of systematic error are substantially different from those of shear-based techniques.

The second outlines an idea for using the optical Tully-Fisher relation to dramatically improve the signal-to-noise and systematic error control for shear measurements. The expected error properties and potential advantages of such a measurement are proposed, and a pilot study is suggested in order to test the viability of Tully-Fisher weak lensing in the context of the forthcoming generation of large spectroscopic surveys.

Contents

List of Figures

List of Tables

Acknowledgments

"Academia," it is aptly claimed,[1] "is a pie-eating contest where the prize at the end is more pie." It is difficult to explain to family, friends and acquaintances outside of the ivory tower why anyone would voluntarily spend eight years in graduate school. That there are difficult and exotic problems to solve is certainly part of the explanation. The real reason, however, is the other denizens. The open and freewheeling exchange of ideas, tempered by serious and honest critical attention, makes the whole academic enterprise worth joining. I've been lucky enough to be in the company of people at Berkeley who make this kind of environment possible.

David Schlegel has been my *de facto* advisor for most of the last six years, and gave me the time and space and other non-metric resources necessary to take on a thesis project of this magnitude. I do not know anyone better at turning ambition into science on such a large scale. Chris Hirata and Rachel Mandelbaum have been close collaborators over the same period; Chris's approachable brilliance and irrepressible joy in doing science and Rachel's patience and methodical approach make even the most difficult problems surmountable. All three took the responsibility of mentorship very seriously, to my incalculable benefit. If am very fortunate, I will be able to pass at least some of what they showed me on to my own students.

When the tunnel is very long and dark, it's good to have fellow-travelers with wit, empathy, and their own healthy spirit of intellectual curiosity. The years spent working, kayaking, whiteboarding, and model-building at the Derby Street Institute with Alex and Genevieve will always be counted among my best, despite the recurrent crises and state of "beige alert" which pervaded so much of that time. Without their incomparable company, support, and intellectual companionship things may have turned out very differently.

I am also very grateful to Andrew Wetzel for the six years over which we shared a dining room, whiteboard, many delightful ruminations, and numerous memorable culinary experiments. He remains an invaluable sanity check and source of insight.

Steve Dawson first showed me around Berkeley, and did more than anyone else to persuade me to make my home here.

Mike and Maile Urbancic have been great friends and immensely supportive throughout; I'm especially grateful to Mike and to Brad Delong for the chance to teach economics during my final year.

[1] Genevieve Graves, private communication

Kristen Shapiro and Dan Perley taught this child of the desert how to ski, and along with Charles Hansen and Victor Acosta helped me to discover much of what is great about the northern California outdoors.

Finally, Dexter Stewart's patient and cheerful guidance in navigating the maze of University of California bureaucracy has been invaluable.

Chapter 1

Motivation

Modern cosmologists can simulate the invisible implications of modern cosmological models (e.g., those that can explain the cosmic microwave background, including Komatsu et al. 2011) to what is generally agreed to be a high level of precision (and probably accuracy, c.f. Lawrence et al. 2010). The easily observable consequences of these models for observations of galaxies are not so easy to calculate (e.g., Rudd et al., 2008; Conroy & Wechsler, 2009; Simha et al., 2010), involving as they do the physics of the familiar but nevertheless stubbornly complicated baryons. Most of the precisely calculable components of these models – namely, the properties of the distribution of dark matter on large scales in relatively linear structures – are not readily observable.

For the foreseeable future, the least indirect observation of these dark components is the measurement of the gravitational effects of dark structures on the images of distant background galaxies. These measurements are made almost exclusively via statistical estimation of the distortions in the ellipticities of background galaxies. This takes advantage of the fact that galaxies have no preferred orientation in a homogenous, isotropic universe [1].

Lensing measurements have played a significant role in observational astrophysics in the last two decades, over a range of scales and physical regimes. Studies of galaxy evolution benefit from the ability to understand the dark matter halos that host galaxies (e.g. Hoekstra et al., 2004, 2005; Heymans et al., 2006; Mandelbaum et al., 2006; Mandelbaum et al., 2006c, 2009; Leauthaud et al., 2012). Cosmologists have no other way to directly map the large-scale matter distribution, which is crucial for constraining models of dark energy and modified gravity (Zhang et al., 2007; Reyes et al., 2010). On small scales, maps of the matter distribution can be tied directly to tests of the cold dark matter paradigm and simulations of the formation and evolution of dark matter halos.

Much has been made of the scientific potential of this technique. Six years ago, weak lensing was identified by the Dark Energy Task Force (Albrecht et al., 2006) as the most promising tool for constraining cosmological models. Several large ground-based and space-

[1] For a finite field, this assumption is not strictly true – physically proximate galaxies 'know' about nearby large-scale matter fluctuations, and as a result the covariance between galaxy ellipticities even at large scales does not vanish.

based survey proposals place a weak lensing measurement among their primary science drivers, including the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS)[2], the Dark Energy Survey (DES)[3], the Hyper Suprime-Cam (HSC, Miyazaki et al. 2006) survey, the Large Synoptic Survey Telescope (LSST)[4], Euclid[5], and the Wide-Field Infrared Survey Telescope (WFIRST)[6].

For all the promise, the technical challenges for these future experiments remain formidable. An order-unity distortion to background galaxy images is produced by a physical, projected matter overdensity of

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{d_S}{d_L d_{LS}}, \tag{1.1}$$

where d_L, d_S, and d_{LS} are the angular-diameter distance from the observer to the lens and source, and from the lens to the source, respectively. For characteristic distances of approximately a Gpc, the critical surface density is $0.1\,{\rm g\ cm^{-2}}$. Typical fluctuations in the matter density field projected over cosmological distances are a thousand times smaller than this, so order 10 Mpc-scale density fluctuations in the universe will typically produce changes in galaxy ellipticities of order $e \approx 10^{-3}$ to 10^{-2} in magnitude. In the shot-noise dominated regime, the leading-order contribution to the variance in the correlation function of the ellipticity distortions is

$${\rm Var}\,(\xi_\epsilon) = \frac{\sigma_\epsilon^4}{N_{\rm pair}^2}. \tag{1.2}$$

For a shallow ($\langle z\rangle = 0.5$) galaxy survey with shape noise $\sigma_\epsilon \approx 0.3$ and 100 deg^2 of sky coverage, reducing the shot noise contribution below the expected cosmological signal requires a surface density of usable source galaxies of at least $\sim 4/{\rm arcmin}^2$.

Worse, for ground-based imaging surveys, the observed shape distortions arising from atmospheric turbulence and optical distortions from the telescope are typically of order several percent, with coherence over angular scales comparable to that of the lensing shape distortions. A competitive measurement of the amplitude of matter fluctuations requires suppressing or modeling these coherent spurious distortions to better than one part in 10^3.

Achieving both the statistical precision and control of systematic errors that is required for such a measurement has proved to be a challenge.

Cosmic shear measurements were attempted as early as 1967 (Kristian, 1967), but until the turn of the millennium (Bacon et al., 2000; Kaiser et al., 2000; Van Waerbeke et al., 2000; Wittman et al., 2000), no astronomical survey had the statistical power to detect it. The successful early detections (Bacon et al., 2000; Van Waerbeke et al., 2000; Rhodes et al., 2001; Hoekstra et al., 2002; Brown et al., 2003; Jarvis et al., 2003) showed the promise of the method and confirmed the existence of lensing by large-scale structure at roughly the

[2] http://pan-starrs.ifa.hawaii.edu/public/
[3] http://www.darkenergysurvey.org/
[4] http://www.lsst.org/
[5] http://sci.esa.int/euclid/
[6] http://wfirst.gsfc.nasa.gov/

expected level. However, they also highlighted some of the systematic errors: in particular, B-mode shear (which cannot be produced by lensing at linear order and is thus indicative of systematic effects) was present at a subdominant but non-negligible level. Since then, the weak lensing community has moved in the direction of both deep/narrow surveys with the *Hubble Space Telescope* (*HST*) and wide/shallow surveys on the ground. The Cosmological Evolution Survey (COSMOS) is the premier example of the former: in addition to 2-point statistics (Massey et al., 2007b; Schrabback et al., 2010), it has also produced three-dimensional maps of the matter distribution (Massey et al., 2007b) and the lensing 3-point correlation function (Semboloni et al., 2011). Excellent control of lensing systematics in COSMOS was also achieved thanks to the small number of degrees of freedom controlling the PSF (mostly focus variation; Rhodes et al. 2007) and detailed modeling of charge transfer inefficiency (Massey et al., 2010). However, COSMOS covers only 1.6 deg^2, and the small field of view of *HST* instruments makes significantly larger surveys impractical. The principal recent ground-based cosmic shear program has been the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS). There are now several cosmic shear results from different subsets of the CFHTLS data (Semboloni et al., 2006; Hoekstra et al., 2006; Benjamin et al., 2007; Fu et al., 2008), and the CFHT lensing team is completing a reanalysis using recent advances in PSF determination and galaxy shape measurement.

In light of the efforts shortly to be made by large, expensive surveys to measure cosmic shear, we consider it imperative to show that such a measurement can be performed accurately, without significant contaminating systematic errors, from a ground-based observatory. This goal includes doing a cosmic shear measurement with each of the wide-angle optical surveys that presently exist.

This work combines several methods discussed in the literature as viable techniques for measuring cosmic shear while removing common systematic errors. It begins with the PSF model generated by the Sloan Digital Sky Survey (SDSS) pipeline over a ~ 250 deg^2 stripe of sky that had been imaged many times, and employs a rounding kernel method similar to that proposed in Bernstein & Jarvis (2002). The result, after appropriate masking of problematic regions, is 168 square degrees of deep coadded imaging with a well controlled, homogeneous PSF and sufficient galaxy surface density to measure a cosmic shear signal.

The catalog constructed from this imaging is then used to produce a cosmic shear measurement that is dominated by statistical errors. Chapter 7 enumerates the primary sources of systematic error when measuring cosmic shear using our catalog (the properties of which are summarized briefly in Chap. 5), and describes the approaches to constraining each of them. Chapter 8 outlines the correlation function estimator and several transformations of it that are used for systematics tests. Our methods for estimating covariance matrices for the observable quantities (both due to statistical and systematic errors) are described in Chap. 7. Finally, chapter 9 presents the constraining power of this measurement alone for a fiducial cosmology, and in combination with the 7-year *Wilkinson Microwave Anisotropy Probe* (WMAP7, Komatsu et al., 2011) parameter constraints to produce a posterior probability distribution over $\Omega_m h^2$, $\Omega_b h^2$, σ_8, n_s, and w. In addition to its value as an independent measurement of the late-time matter power spectrum, this measurement provides some ad-

ditional constraining power over WMAP7 within the context of ΛCDM.

While this work was underway, my collaborators and I learned of a parallel effort by Lin et al. (2011). These two efforts use different methods of coaddition, different shape measurement codes, different sets of cuts for the selection of input images and galaxies, and analyze their final results in different ways; what they have in common is their use of SDSS data (not necessarily the same sets of input imaging) and their use of the SDSS PHOTO pipeline for the initial reduction of the single epoch data and the final reduction of the coadded data (however, they use different versions of PHOTO). Using these different methods, both groups have extracted the cosmic shear signal and its cosmological interpretations. We have coordinated submission with them but did not consulted their results prior to conducting our own analysis, so these two efforts are independent, representing versions of two independent pipelines.

Chapter 2

Shape measurement and Shear estimation.

The term of art for shear estimation from galaxy surface-brightness profiles is shape measurement, where 'shape' is defined as any one of several ellipticity definitions, each of which is a linear combination of weighted second moments of the galaxy surface brightness profile. This follows from a physical intuition – shear primarily affects the ellipticity of a galaxy image, so for analysis the focus should be on making unbiased measurements of galaxy ellipticities.

It is well understood that the relationship between measured ellipticity and the gravitational lensing shear is not trivial. Galaxies have a large diversity of surface-brightness profiles, and designing an estimator that has an expectation value of the shear requires using information about the full distribution function of galaxy image properties. Standard shape measurement methods apply corrections for PSF[1] ellipticity and dilution as well as higher-frequency structure in the galaxy and PSF surface-brightness profiles. Accuracy at the demanding standard (shear errors of order 10^{-3}) required for the next generation of experiments has not yet been conclusively demonstrated in realistic simulations or actual experiments.

The approach taken here is to use the maximum-likelihood procedure, recommended (at least as an initial approach) by every statistician who is confronted by a scientist in need of an estimator. What follows does not pretend to be a solution to the shear estimation problem, but rather a useful schema for making sense of the various existing techniques.

[1]Note here the effects of the finite size of CCD pixels is incorporated into the definition of PSF.

2.1 The Maximum Likelihood Estimator: Measuring Shear in a Perfect World

The weak lensing formalism is, by definition, only valid in the regime where the lensing distortion is a linear tranformation[2] of the unlensed image. At this order, the image distortion can be represented by a linear coordinate mapping between the unlensed and observed source planes $\mathbf{x} \mapsto A\mathbf{x}$ where the linear mapping:

$$A = \begin{pmatrix} 1+\kappa+\gamma_1 & \gamma_2 \\ \gamma_2 & 1+\kappa-\gamma_1 \end{pmatrix} \tag{2.1}$$

The shear (γ_1, γ_2) and convergence κ components are related to the matter distribution projected along the line of sight to the lensed galaxy as:

$$(\gamma_1, \gamma_2) = \partial^{-2} \int_0^\infty W(\chi, \chi_s) \left(\partial_x^2 - \partial_y^2, 2\partial_x\partial_y\right) \delta(\chi \hat{\mathbf{n}}_s)\, \mathrm{d}\chi. \tag{2.2}$$

$$\kappa = \int_0^\infty W(\chi, \chi_s)\, \delta(\chi \hat{\mathbf{n}}_s)\, \mathrm{d}\chi. \tag{2.3}$$

where $\mathbf{x} = (x, y)$ is the coordinate vector in the observed image plane.

Here the integral is along the comoving line-of-sight distance χ (where χ_s is the distance to the source), and the matter over-density $\delta = (\rho - \overline{\rho})/\rho$. The lensing kernel in a flat universe is

$$W(\chi, \chi_s) = \frac{3}{2}\Omega_m H_0^2 (1+z)\chi^2 \left(\frac{1}{\chi} - \frac{1}{\chi_s}\right). \tag{2.4}$$

It is conventional to describe the shear components as mapping circular isophotes to ellipses, and the convergence as producing a change in radius of the same. This is a true but incomplete description of the effects of these distortion components; what follows is an elaboration of the effects of the shear distortion on a general galaxy surface-brightness profile.

The unlensed $I(\mathbf{x})$ and lensed $\tilde{I}(\mathbf{x})$ galaxy images can be related to leading order by expanding the distorted image profile to linear order in the shear around the unsheared image.

$$\tilde{I}(\mathbf{x}) = I(A\mathbf{x}) = I(\mathbf{x}) + \gamma_1 \partial_{\gamma_1} I(\mathbf{x}) + \gamma_2 \partial_{\gamma_2} I(\mathbf{x}) + \kappa \partial_\kappa I(\mathbf{x}) + O\left(\gamma^2\right) \tag{2.5}$$

where A here is the lensing distortion transform from equation 2.1 above and $O(\gamma^2)$ is shorthand for terms that are second order in any product of the lensing distortion components.

The derivatives can be written in terms of two of the Pauli matrices, σ_1 and σ_3:

$$\begin{aligned} \partial_{\gamma_1} I(\mathbf{x}) &= -\mathbf{x}^T \sigma_3 \cdot \nabla I(\mathbf{x}) \\ \partial_{\gamma_2} I(\mathbf{x}) &= \mathbf{x}^T \sigma_1 \cdot \nabla I(\mathbf{x}). \end{aligned} \tag{2.6}$$

[2] Higher-order corrections can be significant in some cases, but for the wide-angle cosmological applications being considered here, the linear treatment is sufficient

Next, I suppose the existence of a pixelized image of a galaxy with surface-brightness profile $I^{\rm obs}(\mathbf{x})$ and a fully accurate model of that profile $I^m(\mathbf{x}|\gamma)$ which is a function only of the lensing distortion. For simplicity, I assume that the observed image is detected as a vector of pixel intensities $I_i^{\rm obs}$, and that the pixel intensity noise covariance is $C_{i,j} = \sigma_n^2 \delta_{i,j}$. In this case, the log-likelihood as a function of shear is:

$$\log L = \sum_i^{N_{\rm pix}} \left(\frac{I_i^{\rm objs} - I_i^m}{\sigma_n} \right)^2 - \frac{N_{\rm pix}}{2} \log\left(2\pi\sigma_n^2\right) \tag{2.7}$$

Assuming for the sake of simplicity only a single component of nonzero shear γ_1 (and no convergence), the derivative is:

$$\partial_{\gamma_1} \log L = \sum_i^{N_{\rm pix}} \frac{2}{\sigma_n^2} \left[I_i^{\rm obs} - \left(I_i^m - \gamma_1 \mathbf{x}_i^T \sigma_3 \cdot \nabla I_i^m \right) \right] \mathbf{x}_i^T \sigma_3 \cdot \nabla I_i^m \tag{2.8}$$

Optimizing for γ_1 produces the linear Maximum Likelihood estimator (MLE)[3]:

$$\begin{aligned} \hat{\gamma_1} =& \frac{\sum_{i=0}^{N_{\rm pix}} \left(I_i^{\rm obs} - I_i^m \right) \mathbf{x}_i^T \sigma_3 \cdot \nabla I_i^m}{\sum_{i=0}^{N_{\rm pix}} \left(\mathbf{x}_i^T \sigma_3 \cdot \nabla I_i^m \right)^2} \\ =& \frac{\sum_{i=0}^{N_{\rm pix}} \left(I_i^{\rm obs} - I_i^m \right) w\left(\mathbf{x}_i\right)}{\sum_{i=0}^{N_{\rm pix}} w\left(\mathbf{x}_i\right)^2} \end{aligned} \tag{2.9}$$

I will show below that this expression is equivalent in certain limits to shape measurement methods commonly used or advocated for use in modern shear analyses. Thus far, the MLE is valid only for a single galaxy surface-brightness profile. Applying the same MLE to the great diversity of real galaxies necessitates the use of a model I^m which is in some sense the 'average' galaxy surface brightness; the correct choice of 'average' model depends on on the exact weighting scheme used in the lensing analysis, as well as the precise mixture of galaxy types represented in the survey[4]. For what follows, I will deal only with the MLE for a single galaxy, but the reader should bear in mind the fact that this choice elides one of the major sources of bias in shear estimation.

2.1.1 Multipole moments

If all unlensed galaxy images were perfectly round (that is, made up of cocentric circular isophotes), then equation (2.9) simplifies dramatically. The derivative of the light profile I

[3] As a practical matter, '∇' is meant to be applied to the continuous model before its representation as a vector of pixel intensities.

[4] This can be a strong function of redshift.

with respect to γ_1 becomes:

$$\partial_{\gamma_1} I(\mathbf{x}) = -x\partial_x I + y\partial_y I = r\cos(2\theta)\,\partial_r I \tag{2.10}$$

A small shear applied to a round galaxy induces a small quadrupole proportional to the shear, and so for initially round objects the optimal estimator is only sensitive to the ellipticity of the galaxy. Even in this case, however, the amount of ellipticity generated by a given shear depends explicitly on the radial profile of the galaxy surface brightness profile.

Real galaxies are not perfectly round in the absence of shear, and so the MLE depends on more than just the quadrupole of the light profile. If I and I^m are expanded in multipole moments:

$$\begin{aligned} I &= \sum_{n=-\infty}^{n=\infty} I_n e^{in\theta} \\ I_n &= \int_0^{2\pi} d\theta I(r,\theta)\, e^{-in\theta} \end{aligned} \tag{2.11}$$

then it can be shown (Bernstein & Jarvis, 2002) that the effect of a small shear γ on the n^{th} moment depends on the $n+2$ and $n-2$ moments as:

$$\Delta I_n = -\frac{\gamma}{2}\left[(n-2)\,I_{n-2}(r) - r\partial_r I_{n-2}(r) - (n+2)\,I_{n+2}(r) - r\partial_r I_{n+2}(r)\right] \tag{2.12}$$

When this expression is inserted into equation (2.9), the $n-2$ and $n+2$ moments do not in general vanish or cancel; naive use of a pure $n=2$ shape estimator would produce a bias equal to the sum of the remaining MLE multipole terms. It is possible in principle to design a scheme that nulls all contribution from the higher multipole moments if the full unsheared surface-brightness profile is known in advance. Several current-generation shape measurement methods adopt such a strategy, which I discuss below.

2.1.2 Second Moments in ensemble

Matters simplify when the estimator is averaged over a large ensemble of galaxies, where the real symmetries of the universe allow the neglect of certain terms. Rotational symmetries ensure that the true ensemble average of underlying galaxy shapes, in the absence of coherent shear, should be zero; for the ensemble, then, the second term in the numerator

of equation (2.9) vanishes[5], leaving

$$\left\langle \gamma_1 = \frac{\sum_{i=0}^{N_{\rm pix}} \left(I_i^m\right) w(\mathbf{x}_i)}{\sum_{i=0}^{N_{\rm pix}} w(\mathbf{x}_i)^2} \right\rangle = 0 \tag{2.13}$$

$$\hat{\gamma}_1 = \frac{\sum_{i=0}^{N_{\rm pix}} I_i^{\rm obs} w(\mathbf{x}_i)}{\sum_{i=0}^{N_{\rm pix}} w\left(\mathbf{x}_i\right)^2} \tag{2.14}$$

What remains is a linear operator on the observed galaxy profile, incorporating a 'representative model' I^m for how the observed galaxy should respond to an applied shear.

Expanding $-\log I^m$ in a power series yields:

$$-\log I^m = \sum_{m,n} \frac{a_{m,n}}{m!n!} x^m y^n \tag{2.15}$$

Inserting this expression into equation (2.14), keeping only terms up to second order ($m+n = 2$), and for the moment assuming that the I^m and $I^{\rm obs}$ are both centered (i.e., setting $a_{10} = a_{01} = 0$) and normalized ($a_{00} = 0$), the ensemble-averaged estimator becomes:

$$\hat{\gamma}_1 = \frac{\sum_{i=0}^{N_{\rm pix}} I_i^{\rm obs} \left(a_{2,0}x^2 - a_{0,2}y^2\right) w\left(x,y\right)}{\sum_{i=0}^{N_{\rm pix}} \left[\left(a_{2,0}x^2 - a_{0,2}y^2\right) w\left(x,y\right)\right]^2} \tag{2.16}$$

where the weight function is the gaussian that best fits the unsheared model galaxy:

$$w(x,y) = e^{-\frac{1}{2}\left(a_{2,0}x^2 + a_{0,2}y^2 + 2a_{1,1}xy\right)} \tag{2.17}$$

When averaged over the entire galaxy population, symmetry requires $a_{i,j} = a_{j,i}$, and so setting the coefficients $a_{2,0} = a_{0,2} = a_2$ in the prefactor (though not in the weight function) should not change the expectation value of the shear estimator. The MLE is then an adaptive gaussian-weighted second moment of the galaxy light profile:

$$\hat{\gamma}_1 = \frac{\sum_{i=0}^{N_{\rm pix}} I_i^{\rm obs} \left(x^2 - y^2\right) w\left(x,y\right)}{\sum_{i=0}^{N_{\rm pix}} a_2^2 \left[\left(x^2 - y^2\right) w\left(x,y\right)\right]^2} \tag{2.18}$$

[5]Though this term has nonzero variance, and is the source of the shape noise.

This is identical to the second-moment shape estimators proposed by several authors; see for example Bernstein & Jarvis (2002).

For axisymmetric galaxies, the third moments should vanish, and so the next-to-leading order contribution to the optimal shear measurement depends on fourth moments of the light profile with a more complex weight function. As has been pointed out by Hirata & Seljak (2003) and Zhang & Komatsu (2011), neglecting the higher-order moments here can be a substantial source of bias. Worse, for actual galaxies there is no guarantee that this approximation scheme will converge in a manageably small number of terms.

2.2 The Maximum Likelihood Estimator for Shear with PSF Smearing

Real galaxies are smeared by the effects of atmospheric turbulence and telescope optics, and so the estimator must needs be modified to deal with PSF convolution.

The form of the MLE is modified somewhat in the presence of an anisotropic PSF. The model light profile I^m must be convolved with the PSF, and equation 2.13 no longer holds. Defining $I^P = I^m \otimes K$, where K is the functional form of the PSF and making use of the fact that $\partial_{\gamma_1} K \otimes I^m = K \otimes \partial_{\gamma_1} I^m$ produces the following expression for the shear estimator:

$$\hat{\gamma}_1 = \frac{\sum_{i=0}^{N_{\rm pix}} \left(I_i^{\rm obs}\right) w\left(\mathbf{x}_i\right)}{\sum_{i=0}^{N_{\rm pix}} w\left(\mathbf{x}_i\right)^2} - \frac{\sum_{i=0}^{N_{\rm pix}} \left(I_i^P - I_i^P\right) w\left(\mathbf{x}_i\right)}{\sum_{i=0}^{N_{\rm pix}} w\left(\mathbf{x}_i\right)^2} \qquad = W^{\rm obs} - W^P \tag{2.19}$$

which defines the shear moments $W^{\rm obs}$ and W^P of the observed light profile and the PSF-convolved model galaxy, respectively. A similar approach was taken in Kuijken (1999), the main difference here being the explicit linearity of the shear estimator.

As was independently pointed out by Bernstein & Jarvis (2002) and Refregier & Bacon (2003), much insight can be gained from using Hermite-Laguerre polynomials to analyze the effects of shear on psf-convolved galaxy light profiles. The Hermite-Laguerre polynomials are eigenfunctions of the Fourier transform, so a convolution of two such polynomials is a product of the same.

The Hermite-Laguerre functions are:

$$\psi_{p,q}^{\sigma} = \frac{-1^q}{\sqrt{\pi}\sigma^2}\sqrt{\frac{q!}{p!}}\left(\frac{r}{\sigma}\right)^m e^{im\theta} e^{-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2} L_q^{(m)}\left[\left(\frac{r}{\sigma}\right)\right] \tag{2.20}$$

These are recognizable to the connoiseur as eigenfunctions of the quantized simple harmonic oscillator. The Laguerre polynomials $L_q^{(m)}(x)$ can be calculated from the following recurrence

relation:

$$\begin{aligned} L_0^{(m)} &= 1 \\ L_1^{(m)} &= (m+1) - x \\ L_q^{(m)} &= \frac{1}{q+1}\left[(2q+m+1) - x\right] L_q^{(m)}(x) - (q+m)\, L_{q-1}^{(m)}(x) \end{aligned} \tag{2.21}$$

The orthogonality of the $\psi_{p,q}^{\sigma}$ at fixed σ:

$$\int_{\infty} d^2\mathbf{x}\, \psi_{p,q}^{\sigma}(\mathbf{x})\, \bar{\psi}_{p',q'}^{\sigma}(\mathbf{x}) = \frac{1}{\sigma^2}\delta_{p,p'}\delta_{q,q'} \tag{2.22}$$

allows them to be used as a series approximation scheme for galaxy surface-brightness profiles:

$$I(r,\theta) = \sum_{p,q>0}^{\infty} b_{p,q} \tag{2.23}$$

$$b_{pq} = \sigma^2 \int_{\infty} d^2\mathbf{x}\, I(\mathbf{x})\, \bar{\psi}_{pq}^{\sigma}(\mathbf{x}) \tag{2.24}$$

This is a useful series approximation to for two reasons. First, the Laguerre coefficients are manifestly combinations of gaussian-weighted moments of the surface-brightness profile, and a gaussian seems at first not a bad approximation to a galaxy light profile in many cases. Second, the Laguerre functions are their own Fourier transforms, so in this basis a convolution can be written as a product. If $\mathbf{k} = \left\{k_{pq}^{\sigma_P}\right\}$ and $\mathbf{g} = \left\{g_{pq}^{\sigma}\right\}$ are the vectors of Laguerre function coefficients for the PSF $K(\mathbf{x})$ and the pre-smeared galaxy image $G(\mathbf{x})$, respectively, then the result of the convolution $I = K \otimes G$ is:

$$\begin{aligned} b_{p,q} &= \mathbf{k}^T C \mathbf{g} \\ b_{p,q} &= \sum_{p'q'} \sum_{p''q''} C_{p,q}^{p'q',p''q''} k_{p'q'} g_{p''q''} \end{aligned} \tag{2.25}$$

where the coefficients $C_{p,q}^{p'q',p''q''}$ can be made from the recursion relations set out in Bernstein & Jarvis (2002):

$$\begin{aligned} C_{p,q}^{p'q',p''q''} =& 2\sqrt{\pi}\left[\sqrt{\frac{p''!p'!}{p!\Delta!}} G(p,p',p'')\right]\left[\sqrt{\frac{q''!q'!}{q!\Delta!}} G(q,q',q'')\right] \\ \Delta \equiv& p'' + p' - p = q'' + q' - q \geq 0 \end{aligned} \tag{2.26}$$

Next, it is important to examine the effects of the shear on the coefficient vectors $\mathbf{k}$ and $\mathbf{g}$. Here one can make use of the textbook quantum p- and q-raising and -lowering operator relations on these functions to show that the effect of a shear γ_1 is:

$$G(A\mathbf{x}) = (1 + \frac{1}{2}\left[\left(a_p^{\dagger}\right)^2 + \left(a_q^{\dagger}\right)^2 - (a_p)^2 - (a_q)^2\right])G(\mathbf{x}) \tag{2.27}$$

This can be written as a transformation on the vector $\mathbf{g}$. The effect of shear on the Laguerre coefficients is:

$$\frac{\partial g_{pq}}{\partial \gamma_1} = \frac{1}{2}\left(\sqrt{p(p-1)}b_{(p-2)q} + \sqrt{q(q-1)}b_{p(q-2)} - \sqrt{(p+1)(p+2)}b_{(p+2)q} - \sqrt{(q+1)(q+2)}b_{p(q+2)}\right) \tag{2.28}$$

This is the essence of the Shapelets formalism (Refregier & Bacon, 2003). A direct comparison with the MLE can be got by choosing modestly realistic surface-brightness profiles for the galaxy and PSF, calculating the optimal weight function $w(\mathbf{x})$, and finding the Laguerre-Hermite coefficients of the latter.

Real ground-based PSFs have substantial wings due to atmospheric turbulence, for which a Moffat function :

$$K(r) = \left[1 + \left(\frac{r}{\sigma_{\rm PSF}}\right)\right]^{-\beta} \tag{2.29}$$

with $\beta = 3$ is a common choice (Kitching et al., 2010). Sérsic profiles:

$$I^m(r) = I_0 e^{-\left(\frac{r}{\sigma_G}\right)^{\frac{1}{n}}} \tag{2.30}$$

with $n = 4$ and $n = 1$ are good parameterizatons of elliptical and spiral galaxy light profiles, respectively, in magnitude-limited surveys. The weight function for for a round PSF convolved with a Sérsic $n = 4$ profile is shown in Figure 2.1. The weight function for an elliptical PSF is shown in Figure 2.2. A matrix of the Laguerre-Hermite polynomial coefficients for these weight functions is shown in Figures 2.3 and 2.4; these coefficient matrices show the multipole moments $m = p - q$ and Laguerre polynomial orders $p + q$ that contribute to the MLE in each case.

It is evident that the Shapelets expansion, while formally complete, converges far too slowly to be useful for galaxies and PSFs with extended wings, especially as the higher-order moments are very difficult to measure in the presence of noise.

The non-gaussian structure in the PSF mixes significant amounts of non-quadrupole $m > 2$ power into the shear estimator. Ignoring this dependence and relying only on shape measurement will tend to introduce additional bias that becomes more significant with increasingly complex PSF structure.

This observation generalizes beyond the simple example shown here; in more careful studies, Hirata & Seljak (2003) find substantial shear measurement bias for realistic profiles even after including Hermite-Laguerre terms up to order $p + q = 24$. This mixing is reduced in the presence of a well-behaved PSF; in an experiment that relies on shape measurement as a proxy for shear measurement, a round, gaussian PSF will suffer much less from ignoring the $m > 2$ moments.

The tactic which will be pursued in the cosmic shear study described in this Thesisis to suppress the parts of the galaxy and PSF surface-brightness profile which lead to these

higher moment contributions, so that the gaussian-weighted second moments are a good appoximation to the ensemble-averaged shear moments W defined above [6] .

[6]The suppression of undesirable terms in the light profile is pursued elsewhere as well; see Bernstein (2010) for a different approach with a similar strategy.

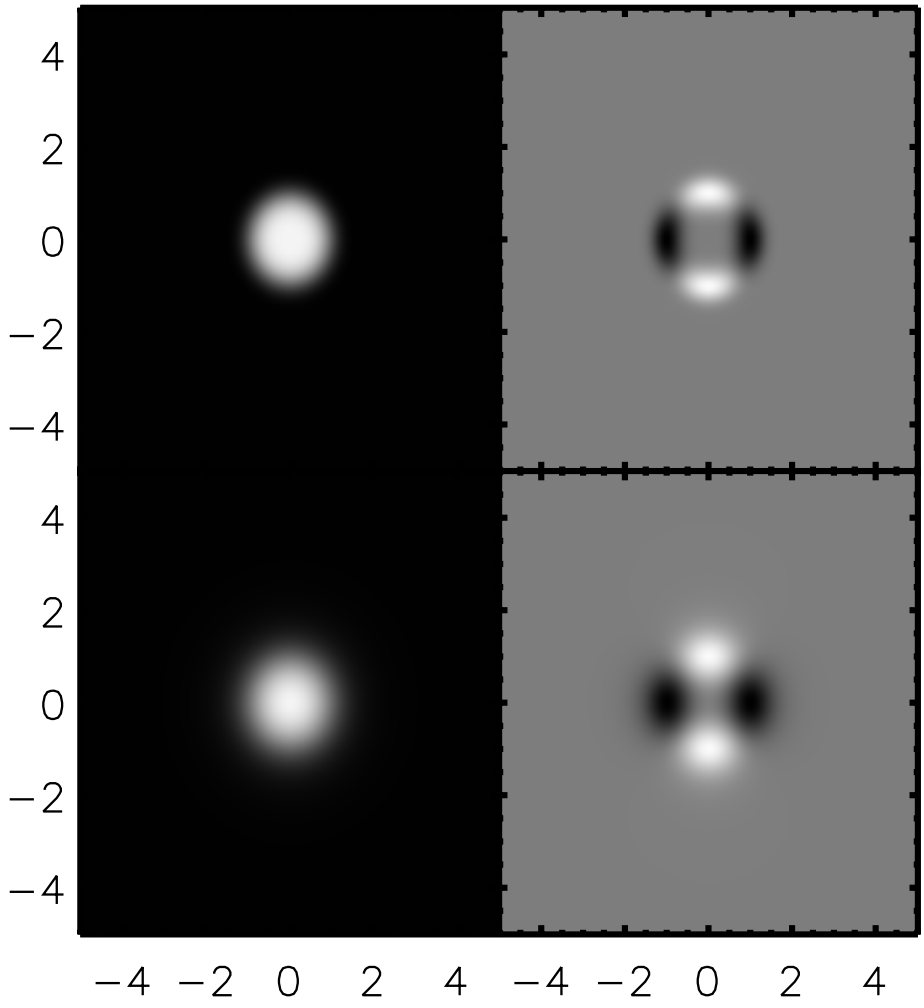

Figure 2.1 Left panels show an $n = 4$ Sérsic profile before (top) and after (bottom) convolution with a round Moffat PSF. Right panels show the corresponding weight functions $w(\mathbf{x})$ used in the linear $\hat{\gamma}_1$ shear estimator.

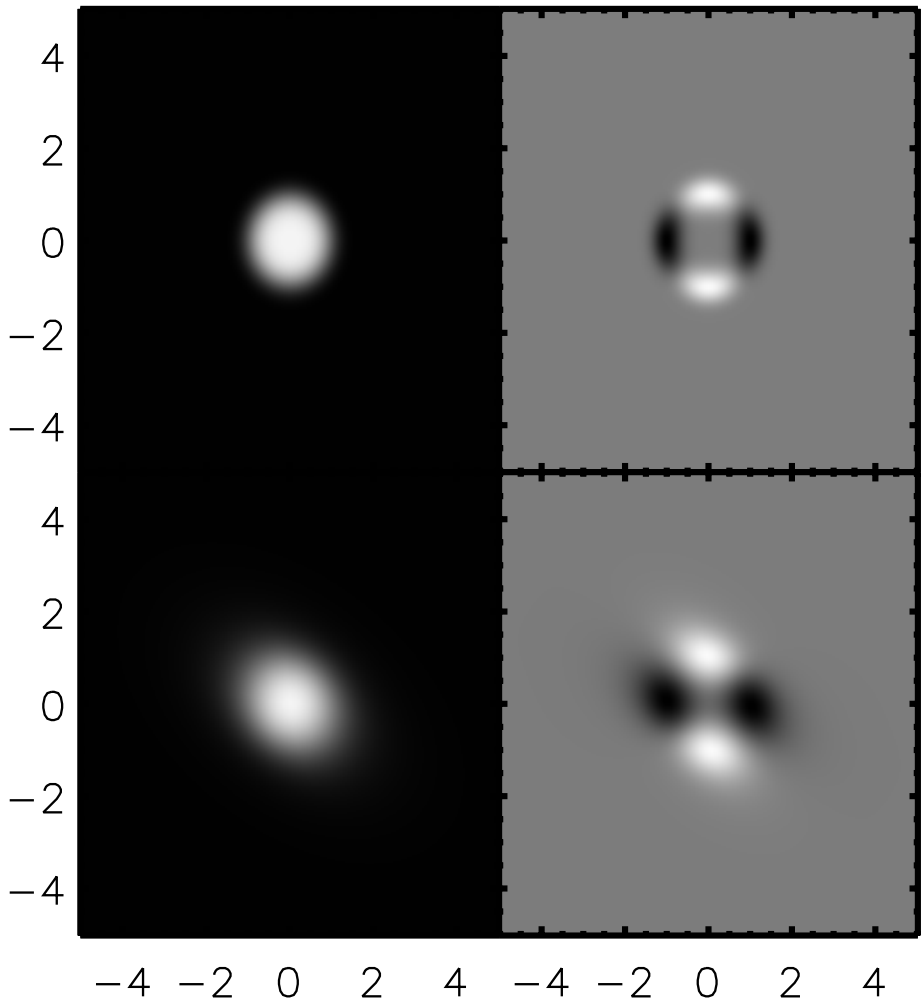

Figure 2.2 Left panels show an $n = 4$ Sérsic profile before (top) and after (bottom) convolution with an elliptical Moffat PSF. Right panels show the corresponding weight functions $w(\mathbf{x})$ used in the linear $\hat{\gamma}_1$ shear estimator.

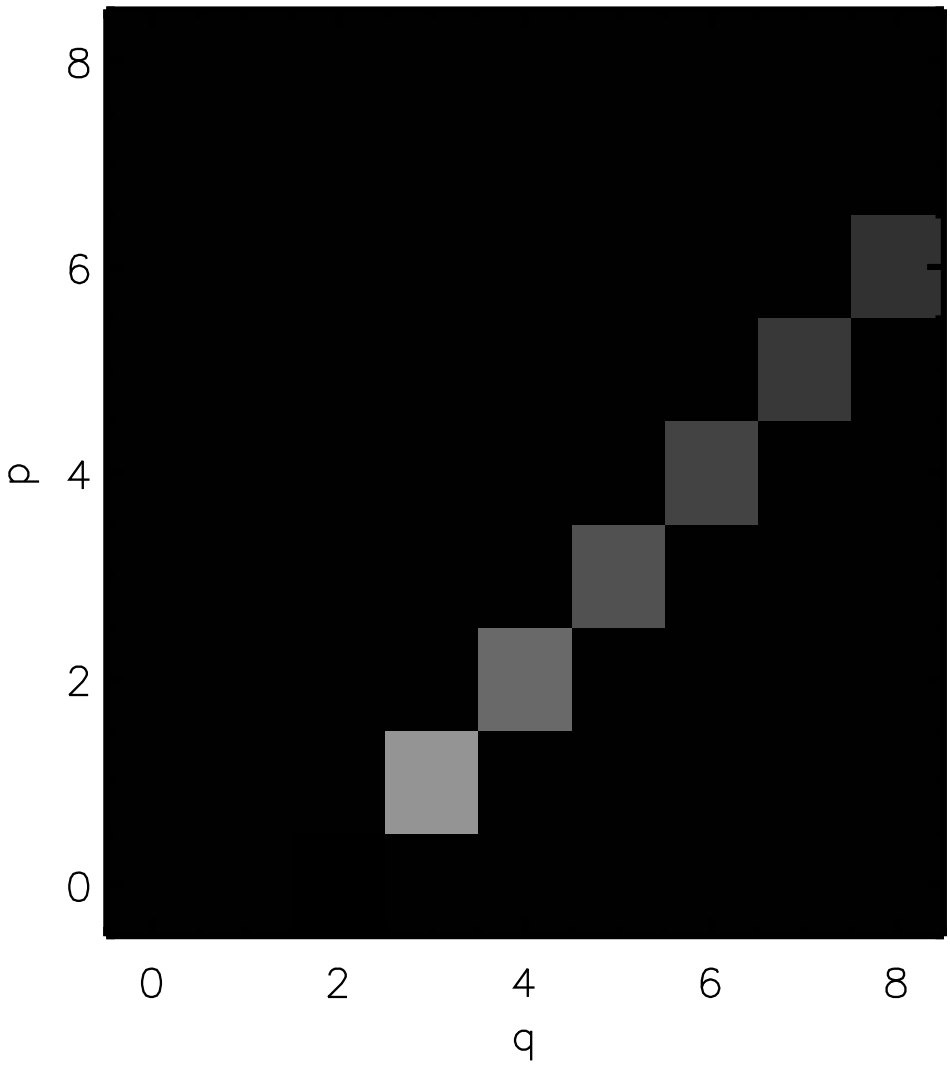

Figure 2.3 The matrix of the first nine Gauss-Laguerre coefficients of the weight function shown in figure 2.1. Note the the corresponding multipole moment that each coefficient contributes to is $m = p - q$. The intensity is linear, and even without scaling it is clear that much power resides in the $m + n > 2$ coefficients.

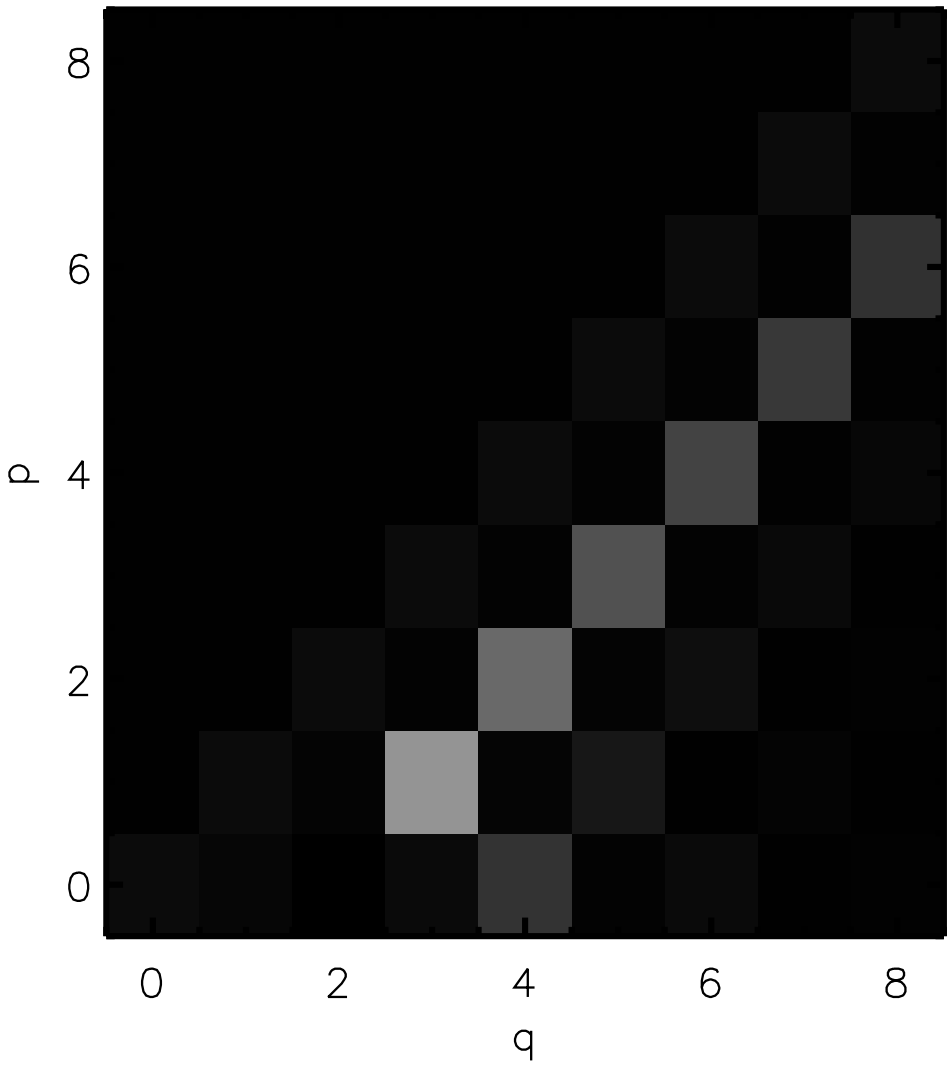

Figure 2.4 The matrix of the first nine Gauss-Laguerre coefficients for the weight function shown in figure 2.2. Note the the corresponding multipole moment that each coefficient contributes to is $m = p - q$. The intensity is linear, and even without scaling it is clear that much power resides in the $m + n > 2$ and $m - n > 2$ coefficients.

Chapter 3

Data

3.1 The Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS; York et al. 2000) and its successor SDSS-II (Frieman et al., 2008) mapped 10000 square degrees across the north galactic cap using a dedicated wide-field 2.5 m telescope at Apache Point Observatory in Sunspot, New Mexico (Gunn et al., 2006). The SDSS camera, described in Gunn et al. (1998), images the sky in five optical bands (*u,g,r,i,z*; Fukugita et al. 1996; Smith et al. 2002) with the charge-coupled device (CCD) detectors reading out at the sidereal rate. Each patch of sky passes in sequence through the five filters (in the order *r,i,u,g,z*) along one of the six columns of mosaicked CCDs, and is exposed once in each filter for 54.1 s. The site is monitored for photometricity (Hogg et al., 2001; Tucker et al., 2006). Data undergo quality assessment (Ivezić et al., 2004), and final calibration is done using the "ubercalibration" procedure based on photometry of stars in run overlap regions (Padmanabhan et al., 2008). We use the data from the seventh SDSS data release (Abazajian et al., 2009), with an updated calibration from the subsequent data release.

The footprint of one night's observing is six columns of imaging the width of one CCD (13.52 arcmin) separated by slightly less than one CCD width (11.65 arcmin). Each night's imaging is collectively termed a *run*; each separate column of imaging is, sensibly, a *camera column* (or "camcol"), and the imaging along each camera column is chopped for processing purposes into 8.98 arcmin long *frames* or *fields*. Successive runs are interleaved, in order to fill in the gaps between camera columns. Pairs of interleaved runs along the same great circle are *stripes* (each of which has a north and a south *strip*).

3.2 Stripe 82

Most of the SDSS imaging data were acquired in the northern galactic cap, with galactic latitude $|b| < 30$. For commissioning, and during sidereal times when the primary survey region was unavailable, the telescope frequently imaged a 2.5 degree wide strip of sky along

the celestial equator with right ascension in the interval $[-50, 50]$. In the survey nomenclature, this region is Stripe 82. At any given location on the Stripe, there are on average 120 contributing interleaved imaging runs, comprising in aggregate almost as much imaging data as exist in the remainder of the combined SDSS-I and SDSS-II footprint. It is here that significant gains can be made from image coaddition.

3.3 Single-epoch data processing

The raw imaging data is processed by the automated SDSS photometric pipeline, PHOTO (Lupton et al., 2001). This pipeline has components to handle astrometric and photometric calibration as well as catalog construction; it also generates an array of data quality measurements describing the telescope point-spread function (PSF), the locations of unreliable pixels, and measurements of the photometric quality of individual frames. Many of these data quality indicators are used during the construction of the coadd imaging and its associated catalog. Their use is described below. A detailed description of the image processing pipeline and its outputs can be found in Stoughton et al. (2002). Outputs can be found in locations specified by the SDSS data release papers (Abazajian et al., 2003, 2004, 2005; Adelman-McCarthy et al., 2006, 2007, 2008; Abazajian et al., 2009).

PHOTO produces a number of intermediate outputs for the single-epoch SDSS imaging that we use in the coaddition process. Corrected Frames, or `fpC` files, are produced by the pipeline from the raw CCD images of single frames; these are bias-subtracted, sky-subtracted and flat-fielded, and a nonlinearity correction is applied where appropriate. These are the images that are combined during the coaddition process below.

PHOTO also generates a bitmask (an `fpM` file) for each frame describing pixels that are known to be defective. Pixels are marked in this bitmask as saturated, cosmic-ray contaminated, interpolated (if a column or pixel is known to be saturated, or is *a priori* marked as unreliable, PHOTO interpolates over that region). We use these bitmasks to exclude bad pixels from the image coaddition.

Astrometric solutions (`asTrans` files) are produced by PHOTO for each SDSS frame. Systematic errors in the astrometric positioning are reported to less than 50 mas, and the relative astrometry between successive overlapping frames is approximately 10 mas (Pier et al., 2003). The astrometric solution for each run (Pier et al., 2003) is determined by matching against astrometric standard stars from the USNO CCD Astrograph Catalogue (UCAC Zacharias et al., 2000) and Tycho-2 (Høg et al., 2000) catalogs. Since stars bright enough to appear in Tycho-2 are saturated, a two-step calibration ladder is necessary: the astrometric solution from the science CCDs is tied to a set of 22 "astrometric" CCDs observing through neutral density filters, and these can observe the Tycho-2 stars without saturating. The coaddition algorithm relies on the astrometric solutions provided; we have found it unnecessary to re-solve the astrometry.

For photometric calibration, we use the "ubercal" solutions derived by Padmanabhan et al. (2008).

The SDSS pipeline uses bright, isolated stars with apparent magnitudes brighter than 19.5 to construct a model of the PSF and its variation across each frame. For each frame, the stellar images for the three neighboring frames along the scan in both directions are used to produce a set of Karhunen-Loève (KL) eigenimages describing the PSF variation (Lupton et al., 2001). A global PSF model for the frame is constructed by allowing the first few KL components to vary up to second order in the image coordinates across the frame, with the coefficients of the variation being tied to the aforementioned bright stars. The KL eigenimages and coefficients of their spatial variation are stored by PHOTO for each band in the `psField` files. These are taken as inputs to the coaddition process and used for PSF correction. We will test the fidelity of this PSF model in the coadded images on stars that were too faint to perform a reliable PSF determination in the single-epoch data.

Chapter 4

Methods

4.1 Algorithms

Our general strategy for correcting for the effects of seeing is similar to that suggested in Bernstein & Jarvis (2002). We will apply a rounding kernel to each single-epoch image prior to stacking the ensemble. The large variation in SDSS PSF sizes (see Fig. 4.1) will require a trade-off between rejection of a large fraction of the available imaging, and significant dilution of the signal due to the rounding convolution. Stacking the images without a kernel, however, will produce a PSF with large variations that will be difficult to model accurately.

4.1.1 Field smoothing

This section describes the operation of smoothing the map so as to make the effective PSF equal to some target PSF. Here we will denote the intrinsic PSF of the telescope by $G(\mathbf{x})$, so that if the intrinsic intensity of an object on the sky is $f(\mathbf{x})$, the actual image observed is

$$I(\mathbf{x}) = \int G(\mathbf{y}) f(\mathbf{x}-\mathbf{y}) \mathrm{d}^2\mathbf{y} \equiv [\mathrm{G} \otimes \mathrm{f}](\mathbf{x}). \tag{4.1}$$

Of course this image is only sampled at values of $\mathbf{x}$ corresponding to pixel centres. Our principal objective here is to construct the kernel K such that

$$[I \otimes K](\mathbf{x}) = [\Gamma \otimes f](\mathbf{x}) \quad \text{or} \quad [G \otimes K](\mathbf{x}) = \Gamma(\mathbf{x}), \tag{4.2}$$

where Γ is the target PSF. In order to do this, we need to first choose a target PSF Γ and then determine the appropriate convolution kernel K, which will differ for every imaging run contributing to the coadds at a given position depending on the full position-dependent PSF model in each run. These are the subjects of Secs. 4.1.1 and 4.1.1 respectively.

The target PSF

Here we consider the target PSF Γ. It must be constant across different runs in order for the co-add procedure to make sense, although it need not be the same in different filters.

There is a large advantage in having Γ be circularly symmetric. Gaussians are convenient since most galaxy shape measurement codes are based on Gaussian moments, but this is not a requirement. In fact the PSF G delivered by most telescopes, including the SDSS, has "tails" at large radius that are far above what one could expect from a Gaussian. These can in principle be removed by a convolution kernel K that has negative tails at large radius, but there are problems when these tails extend to the field boundaries or across bad columns in the CCD. Therefore we have chosen the double-Gaussian form for Γ:

$$\Gamma(\mathbf{x}) = \frac{1 - f_{\rm w}}{2\pi\sigma_1^2} \mathrm{e}^{-\mathbf{x}^2/2\sigma_1^2} + \frac{f_{\rm w}}{2\pi\sigma_2^2} \mathrm{e}^{-\mathbf{x}^2/2\sigma_2^2} \tag{4.3}$$

with $\sigma_2 > \sigma_1$. This functional form manifestly integrates to unity, and has a fraction $f_{\rm w}$ of the light in the "large" Gaussian. The two Gaussians have widths σ_1 and σ_2, respectively, with $\sigma_1 < \sigma_2$.

The parameters of the double-Gaussian were adjusted by trial and error so that a compactly supported kernel K (13×13 pixels) can achieve $G \otimes K \approx \Gamma$ for a wide range of real PSFs G delivered by the SDSS. The most critical parameter is the width of the central Gaussian, σ_1. This is the main parameter controlling the seeing of the final co-added image: if it is set too high then many galaxies become unresolved, whereas if it is set too low then a large number of fields with moderate seeing will have to be rejected because it will be impossible to find a kernel K that achieves the target PSF without dramatically amplifying the noise.

The PSF size distribution in the r band is shown in Fig. 4.1.

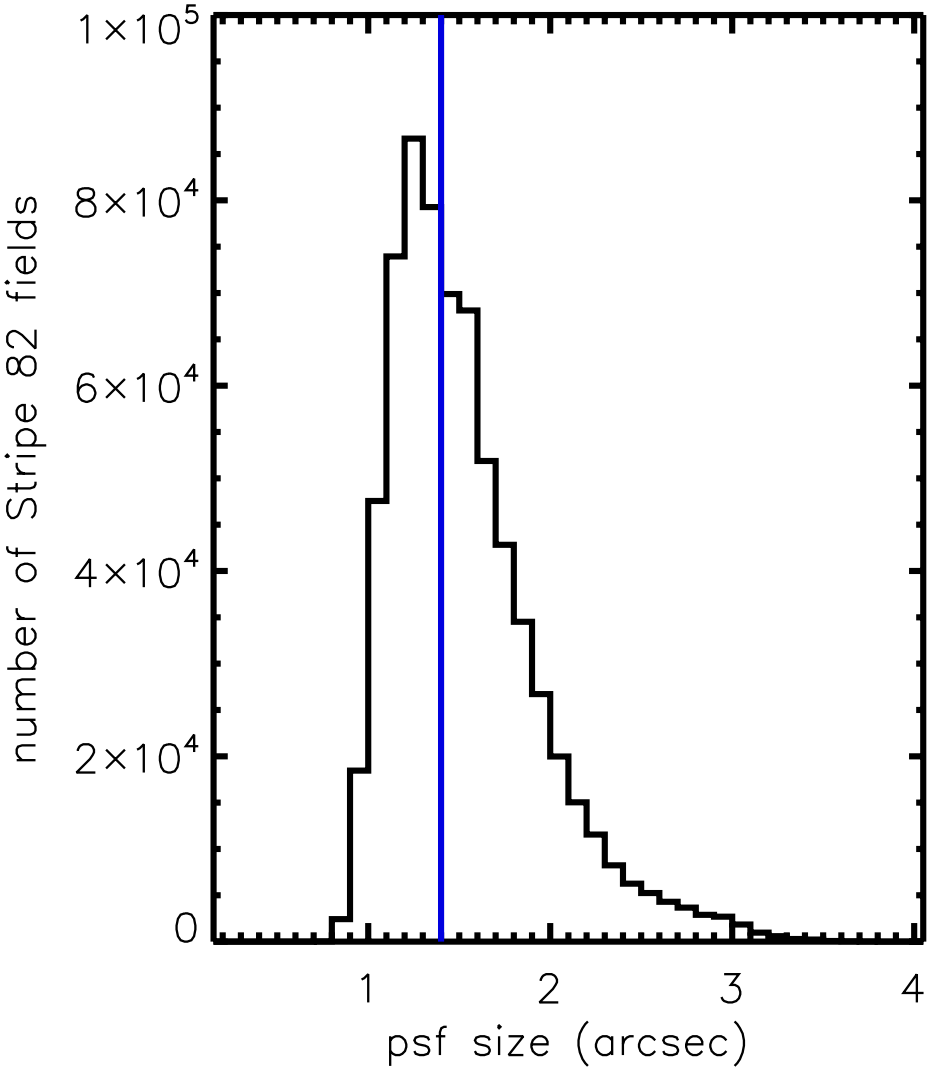

Figure 4.1 The distribution of PSF FWHM in the r band for all frames on Stripe 82. The half-width of the target PSF after rounding is indicated by the solid vertical line.

Table 4.1 Parameters for the PSF repair in different filters.

Parameter	u	g	r	i	z	Units
Target PSF parameters						
σ_1 (PSF_SIZE)	1.80	1.40	1.40	1.40	1.40	pixels
σ_2 (PSF_SIZE_WING)	5.10	5.10	5.10	5.10	5.10	pixels
$f_{\rm w}$ (FRACWING)	0.035	0.035	0.035	0.035	0.035	pixels
FWHM of target PSF Γ	1.68	1.31	1.31	1.31	1.31	arcsec
50 per cent Encircled Energy Radius	0.86	0.67	0.67	0.67	0.67	arcsec
Kernel acceptance parameters						
CUT_L2	0.001	0.0025	0.0025	0.0025	0.0025	
CUT_OFFSET	0.04	0.01	0.01	0.01	0.01	
CUT_ELLIP	0.002	0.0005	0.0005	0.0005	0.0005	
CUT_SIZE	0.01	0.0025	0.0025	0.0025	0.0025	
CUT_PROF4	0.04	0.01	0.01	0.01	0.01	
Co-addition parameters						
DELTA_SKY_MAX1	0.5	0.25	0.25	0.25	0.25	nmgy arcsec^{-2}
DELTA_SKY_MAX2	0.04	0.02	0.02	0.02	0.02	nmgy arcsec^{-2}

The convolution kernel and its application

Equation (4.2) can formally be solved in Fourier space by taking the ratio, $\tilde{K}(\mathbf{k}) = \tilde{\Gamma}(\mathbf{k})/\tilde{G}(\mathbf{k})$, where the tilde denotes the Fourier transform and $\mathbf{k}$ the wave vector. Unfortunately, this simple idea comes with two well-known problems. One is that if the PSF has power only up to a certain wave number $k_{\max}$, then it is impossible to divide by $\tilde{G}(\mathbf{k}) = 0$. The other is that the PSF varies slowly across the field, i.e. G in Eq. (4.2) formally depends on $\mathbf{x}$ as well as $\mathbf{y}$.

The solution to the first problem is that instead of taking a simple ratio in Fourier space, we minimize the L^2 norm of the error,

$$\mathcal{E}_1 = \int |\Gamma(\mathbf{x}) - [G \otimes K](\mathbf{x})|^2 \, \mathrm{d}^2\mathbf{x} \equiv \|\Gamma - G \otimes K\|^2, \tag{4.4}$$

subject to a constraint on the L^2 norm of the kernel:

$$\mathcal{E}_2 = \int |K(\mathbf{x})|^2 \mathrm{d}^2\mathbf{x} \equiv \|K\|^2. \tag{4.5}$$

If the input noise is white (which is a good approximation), then the noise variance on an individual pixel in the convolved image is "$\mathcal{E}_2$ times the noise variance in the input image. Roughly speaking, for kernels that attempt to "deconvolve" the original PSF, and consequently have large positive and negative contributions, $\mathcal{E}_2$ will come out to be very large. We adopt a requirement that $\mathcal{E}_2 \leq 1$. For kernels that poorly approximate the target PSF Γ, $\mathcal{E}_1$ will be very large. The problem of minimizing $\mathcal{E}_1$ subject to a constraint on $\mathcal{E}_2$ is most easily solved by transforming to the Fourier domain and then using the method of Lagrange multipliers:

$$\tilde{K}(\mathbf{k}) = \frac{\tilde{G}(\mathbf{k})\tilde{\Gamma}^*(\mathbf{k})}{|\tilde{G}(\mathbf{k})|^2 + \Lambda}. \tag{4.6}$$

Here the positive real number Λ is the Lagrange multiplier and its value is adjusted until $\mathcal{E}_2 = 1$. Λ plays the role of regulating the deconvolution; indeed one can see that for Fourier modes present in the image, $\tilde{G}(\mathbf{k}) \neq 0$, we have $\lim_{\Lambda \to 0^+} \tilde{K}(\mathbf{k}) = \tilde{\Gamma}(\mathbf{k})/\tilde{G}(\mathbf{k})$.

To summarize, Eq. (4.6) finds the convolving kernel K that makes the final PSF $G \otimes K$ as close as possible (in the least-squares sense) to Γ without amplifying the noise. The kernel is truncated into a 13×13 pixel region centred at the origin in order to avoid boundary effects and to prevent problems such as bad columns, saturated stars, or cosmic rays from "leaking" all over the field. We also re-scale the resulting kernel to integrate to unity ($\tilde{K}(\mathbf{0}) = 1$) but since Λ is small, typically of order 10^{-5}, this has no practical effect. Note that since $G(\mathbf{x})$ and $\Gamma(\mathbf{x})$ are both real functions, it follows that in Fourier space they satisfy the conditions $\tilde{G}(\mathbf{k}) = \tilde{G}^*(-\mathbf{k})$ and $\tilde{\Gamma}(\mathbf{k}) = \tilde{\Gamma}^*(-\mathbf{k})$, and then Eq. (4.6) guarantees that a similar condition holds for K: the convolution kernel $K(\mathbf{x})$ is real.

The second problem – the variation of the PSF across the field – is handled by taking the reconstructed PSF on a grid of 8×6 points separated by 298 pixels (2 arcminutes) in

each direction, and constructing a grid of 48 kernels K. The kernels are then interpolated between neighboring points, and then the final image $F(\mathbf{x})$ is constructed according to

$$F(\mathbf{x}) = \int K_{\mathbf{x}}(\mathbf{y}) I(\mathbf{x}-\mathbf{y})\, \mathrm{d}^2\mathbf{x}, \tag{4.7}$$

where $K_{\mathbf{x}}$ is the kernel reconstructed at position $\mathbf{x}$ in the field.

The convolution kernel will not capture PSF model fluctuations on scales below 2 arcminutes. We show below that, even at 1 arcminute, the remaining PSF variations not captured by the kernel are very small compared to the expected shot-noise errors in the two-point statistics at those scales.

Obviously there will be instances, in which the kernel reconstruction is not good enough. Therefore a set of cuts must be applied to the resulting kernels[1]. In order to construct these cuts, we consider the Gaussian-weighted moments of the residual $\Gamma - G \otimes K$, i.e.

$$M_{\alpha\beta} = \frac{1}{\pi\sigma_1^2} \int [\Gamma - G \otimes K](\mathbf{x})\, \frac{x_1^\alpha x_2^\beta}{\sigma_1^{\alpha+\beta}} \mathrm{e}^{-\mathbf{x}^2/2\sigma_1^2}\, \mathrm{d}^2\mathbf{x}. \tag{4.8}$$

The cuts are then:

1. We reject an entire field if PSP failed to determine a good PSF model in the single-epoch imaging, or was forced to use a low-order fit to the PSF (`PSP_STATUS!=0`).

2. We reject cases where the PSF residual is too large regardless of the moments, i.e.

$$\frac{\|\Gamma - G \otimes K\|^2}{\|\Gamma\|^2} > \texttt{CUT\_L2}. \tag{4.9}$$

   Our shape measurement method relies on the presumption of the chosen form of the PSF, and large deviations from this form may have unpredictable effects on catalog construction or shape measurement; for example, the relative photometric calibration in SDSS is known to be accurate to approximately 1%. The SDSS photometric measurements used here for catalog definition rely on PSF-convolved model fits, so large deviations between the outcome of the rounding kernel and the target model should be kept much smaller than the desired photometric precision.

3. We reject cases where the Gaussian-weighted offset is more than `CUT_OFFSET` σ_1, i.e.

$$\sqrt{M_{01}^2 + M_{10}^2} > \texttt{CUT\_OFFSET}. \tag{4.10}$$

[1]It should be noted that residual anisotropies from the Lagrange multiplier Λ are affected by the quality cuts described below.

4. We reject cases where the ellipticity of the final PSF exceeds `CUT_ELLIP`, i.e.

$$\sqrt{(M_{02} - M_{20})^2 + (2M_{11})^2} > \texttt{CUT\_ELLIP}. \tag{4.11}$$

   In the limit of small PSF ellipticities, this corresponds to half the ellipticity (by the measure we use for galaxy shapes) of the core of the PSF. If the PSF model used to construct the kernel is correct, this cut operates as an effective ceiling on the spurious ellipticity resulting from the PSF. The value is chosen so as to reduce the PSF ellipticity for nearly unresolved objects just below the anticipated shear signal.

5. We reject cases where the PSF size error exceeds `CUT_SIZE`, i.e.

$$|M_{02} + M_{20} - M_{00}| > \texttt{CUT\_SIZE}. \tag{4.12}$$

6. We reject cases where the radial profile of the PSF is severely in error as determined by the fourth moment, i.e.

$$|M_{40} + 2M_{22} + M_{04} - 2M_{00}| > \texttt{CUT\_PROF4}. \tag{4.13}$$

The specific values of the parameters chosen for the cuts depend on the band and are shown in Table 4.1. The tightest constraints on the quality of the PSF are in g, r, i, and z bands (r and i are used to measure galaxy shapes). In the u band, where the average image quality is much lower than in the other bands, more liberal cuts can be applied because we are interested primarily in the total flux, not the shape. Also there is more to gain from liberal cuts because the signal-to-noise ratio in u band is lower. Nevertheless, a serious error in the size of the PSF will result in erroneous photometry, and spurious ellipticity could introduce colour/photo-z or selection biases that depend on galaxy orientation, so some cuts must be applied.

4.1.2 Noise symmetrization

It is a well-known fact in weak lensing that even if the PSF in an image has been corrected to have perfectly circular concentric isophotes, it is possible to produce spurious ellipticity if there is anisotropic correlated noise. For example, if the PSF is elongated in the x_1 direction and is "fixed" by smoothing in the x_2-direction, the resulting map has more correlations in the x_2 direction than x_1. This can lead to (1) centroiding biases, in which the error on the galaxy centroid is larger in the x_2 than the x_1 direction, thus yielding more galaxies that appear aligned in the x_2 than x_1 direction; and (2) biases in which noise fluctuations tend to be elongated in the x_2 direction, so that positive noise fluctuations on top of a galaxy (which increase its likelihood of detection) tend to make it aligned in the x_2 direction whereas negative fluctuations (which decrease the likelihood of detection) make the galaxy aligned in the x_1 direction. For a detailed description of noise-induced ellipticity biases, see Kaiser (2000) or Bernstein & Jarvis (2002). These phenomena can

all mimic lensing signals and hence should be eliminated from the data. Our method of doing this is to add synthetic noise to each field so as to give the noise properties 4-fold rotational symmetry. To be precise, we want the power spectrum of the total noise (actual plus synthetic) to satisfy:

$$P_N(\mathbf{k}) = P_N(\mathbf{e}_3 \times \mathbf{k}), \tag{4.14}$$

where $\mathbf{e}_3$ is a vector normal to the plane of the image; the cross product operation $\mathbf{e}_3\times$ rotates a vector by 90 degrees. Even though it is not circularly symmetric, this is sufficient to guarantee zero mean ellipticity for a population of randomly oriented galaxies because ellipticity reverses sign under 90 degree rotations.[2] In principle m-fold symmetry for any integer $m \geq 3$ would suffice, however 4-fold symmetry is the only practical possibility for a camera with square pixels. For obvious reasons, we would like to achieve this by adding the minimal amount of synthetic noise possible.

The simplest way to achieve Eq. (4.14) is to decompose the power spectrum into its actual ("act") and synthetic ("syn") components:

$$P_N(\mathbf{k}) = P_N^{(\mathrm{act})}(\mathbf{k}) + P_N^{(\mathrm{syn})}(\mathbf{k}). \tag{4.15}$$

The actual component is the white noise variance v in the input image, smoothed by the convolving kernel:

$$P_N^{(\mathrm{act})}(\mathbf{k}) = v|\tilde{K}(\mathbf{k})|^2. \tag{4.16}$$

Since K is real, this power spectrum has 2-fold rotational symmetry: $P_N^{(\mathrm{act})}(\mathbf{k}) = P_N^{(\mathrm{act})}(-\mathbf{k})$. The minimal synthetic noise power spectrum that satisfies Eq. (4.14) is then

$$P_N^{(\mathrm{syn})}(\mathbf{k}) = \max\left[P_N^{(\mathrm{act})}(\mathbf{e}_3 \times \mathbf{k}) - P_N^{(\mathrm{act})}(\mathbf{k}), 0\right]. \tag{4.17}$$

Gaussian noise with this spectrum can be obtained by taking its square root,

$$\tilde{T}(\mathbf{k}) = \sqrt{P_N^{(\mathrm{syn})}(\mathbf{k})}, \tag{4.18}$$

and transforming to configuration space $T(\mathbf{x})$. Then one generates white noise with unit variance and convolves it with T. Since the PSF and hence K varies across the field, T must also vary; its value is interpolated from the same 8×6 grid of reference points as used for K.

The Gaussian white noise was generated using Numerical Recipes `gasdev` modified to use the `ran2` uniform deviate generator (Press et al., 1992). The seed was chosen by a formula based on the run, camcol, field number, and filter to guarantee that the same seed was never used twice in the reductions, and that the same sequence will be generated if the software is re-run. For each field, a sequence of 2048×1361 Gaussian deviates is generated;

[2] In group theory language, the noise properties are symmetric under the 4-fold rotation group $\mathcal{C}_4$, which is a subgroup of the full rotations $SO(2)$. The condition for zero mean ellipticity due to noise is that ellipticity fall into one of the non-trivial represenations of the noise symmetry group.

since there are 128 rows of overlap between successive fields, we fill in the last 128 rows of each field with the first 2048×128 deviates from the next field. It is also essential that the period of the generator be longer than the total number of pixels in the survey (of order a few $\times 10^{12}$), a requirement which is not fulfilled by many generators, since otherwise the same synthetic "noise" pattern will repeat itself throughout the survey.

The image $F(\mathbf{x})$ after addition of the synthetic noise is a `kImage`.

4.1.3 Single-image masking

Once the kernel-convolved, noise-added image (`kImage`) is constructed for each run that will contribute to the coadds at a given position, it is necessary to construct a mask before co-addition. The mask must remove the usual image defects as well as diffraction spikes. It is constructed as described in this section, and is termed the `kMask`.

We begin by masking out all pixels in $F(\mathbf{x})$ for which the convolution (Eq. 4.7) integrates over a bad pixel. Since K has compact support – it is nonzero only in a 13×13 pixel region – this means that for each bad pixel in $I(\mathbf{x})$ we mask out a 13×13 block in $F(\mathbf{x})$. Our definition of a "bad pixel" is one that is out of the field; was interpolated by PHOTO (usually due to being in a bad column); is saturated; is potentially affected by ghosting; was not checked for objects by PHOTO; is determined by PHOTO to be affected by a cosmic ray; or had a model subtracted from it. Note that the first cut means that a 6-pixel region is rejected around the edge of the field.

The second and more sophisticated mask is applied to remove diffraction spikes from stars. The secondary support structure responsible for the diffraction spikes is on an altitude-azimuth mount, so that the diffraction spikes appear at position angles of 45, 135, 225, and 315 degrees in the altitude-azimuth coordinate system. Therefore in the equatorial runs, the orientation of the diffraction spikes relative to equatorial coordinates changes depending on the hour angle of observation. If no correction for this is made, then after co-addition of many runs, even moderately bright stars have a hedgehog-like pattern of diffraction spikes at many position angles that can affect a significant fraction of the area.

Our procedure for removing diffraction spikes is as follows. We first identify objects with a PSF flux (i.e. flux defined by a fit to the PSF) exceeding some threshold (corresponding to 9.7×10^5, 8.5×10^5, 2.2×10^5, 1.7×10^5, and 1.1×10^6 electrons in filters r, i, u, z, and g respectively). Around these objects, we mask a circle of radius 20 pixels (8 arcsec) and four rectangles of width 8 pixels (3 arcsec) and length 60 pixels (24 arcsec). The rectangles have the object centroid at the center of their short side, and their long axis extends radially from the centroid in the direction of the expected diffraction spike.

4.1.4 Resampling

In order to co-add images, we must first resample them into a common pixelization. Ideally we would like this pixelization to be both conformal (no local shape distortion) and equal-area (convenient for total flux measurements). Unfortunately because the sky is

curved, it is impossible to achieve both of these conditions. However since our analysis uses a narrow range of declinations around the equator ($|\delta| \leq 1.3°$) we can come very close by choosing a cylindrical projection; the obvious choices are Mercator (perfectly conformal) or Lambert (perfectly equal-area). In our case the Mercator projection would result in the pixel scale being different by $\Delta\theta/\theta = 2.6 \times 10^{-4}$ at the Equator versus at $\delta = \pm 1.3°$. (The area error is twice this, or 5.2×10^{-4}.) The Lambert projection would preserve shapes at the Equator but the coordinate system would have a shear of $\gamma = 2.6 \times 10^{-4}$ at $\delta = \pm 1.3°$. Neither of these problems is particularly serious, since either could if necessary be corrected in the flux or shape measurements. We have chosen the Mercator projection because the small cosmic shear signal means that we are much more sensitive to a given percentage error in shear than in flux. Also, a flux error of 5.2×10^{-4} is insignificant compared to the error in the flatfields, so there is no point in eliminating it at the expense of complicating the shear analysis.

The scale of the resampled pixels must be smaller than the native pixel scale on the CCD (~ 0.396 arcsec) in order to preserve information. However it is desirable for it not to be too small, since this increases the data volume with no increase in information content. It should also not be nearly equal to the CCD scale in order to avoid production of a moiré pattern with large-scale power. We have used 0.36 arcsec.

The actual resampling step requires us to interpolate the image from the native pixelization onto the target Mercator pixelization. This is done by 36-point second-order polynomial interpolation on the 6×6 grid of native pixels surrounding the target pixel[3]. A target pixel is considered masked if any of the 36 surrounding pixels are masked.

4.1.5 Addition of images

After resampling the images, the next step is to combine them to produce the co-add. The combination proceeds in three steps: comparison of images to reject "bad" regions that were not masked in earlier stages of the analysis; relative sky estimation; and stacking. Note that bad regions must be explicitly rejected: "robust" algorithms such as the median are nonlinear and result in a final co-added PSF that depends on object flux and morphology, which is not acceptable for lensing studies.

Rejection of bad regions is critical because it is possible for some serious defects such as satellite trails to "leak through" earlier stages of the analysis and not be `kMask`ed. Rejection at this stage is also the best way to eliminate solar system objects, which are not easily identified in the single-epoch `fpC`s but of course will not show up at the same coordinates in successive runs. We first bin each input image into 4×4 resampled pixels. We then compare the binned images and reject the brightest or faintest image if it differs by more than `DELTA_SKY_MAX1` from the mean. When this rejection is done, we actually mask a 20×20 resampled pixel region around the affected area. (We found that without this padding region,

[3]Polynomial interpolation on an equally-spaced grid of points converges to sinc interpolation in the limit that the number of gridpoints is taken to infinity.

satellite trails were often incompletely masked because they passed through the corners of some 4×4 regions and did not sufficiently affect the mean flux.)

Next we compute the difference in sky level among all of the N images[4]. This difference must be determined and removed because otherwise a masked pixel in an image with below-average sky will appear as a bright spot in the co-added image. We compute the relative sky level as follows. For each pair (i, j) of co-added images, we compute the difference map $F_i - F_j$ and take the median in 125×125 resampled pixel blocks. This is taken as an estimate of the sky difference $S_i - S_j$. From these differences we obtain the unweighted least-squares solution for the sky levels $\{S_i\}$, up to an additive offset (the absolute sky level cannot be determined by this procedure). The mean of these levels is denoted by $\bar{S} = \sum_{i=1}^{N} S_i/N$. We add to the ith image the quantity $\bar{S} - S_i$ interpolated to a particular point $\mathbf{x}$ by 4-point interpolation from the nearest block centres. An entire block is masked out if $|\bar{S} - S_i| >$`DELTA_SKY_MAX2` and if it is an extremal value (either the highest or lowest sky value).

The stacking of the images works by the usual formula

$$F_{\rm tot}(\mathbf{x}) = \frac{\sum_{i=1}^{N} w_i(\mathbf{x}) F_i(\mathbf{x})}{w_{\rm tot}(\mathbf{x})}, \tag{4.19}$$

where $w_{\rm tot}(\mathbf{x}) = \sum_{i=1}^{N} w_i(\mathbf{x})$ and w_i are the weights. Because the noise is correlated, the optimal weights are scale-dependent; we have chosen the optimal weights in the limit of small $\mathbf{k}$, i.e. large scales. That is, $w_i = v^{-1}$ where v is the noise variance in image i. For photometry of large objects, $w_{\rm tot}$ can be thought of as an inverse white noise variance, i.e. the mean square noise flux in a region of area Ω is $1/w_{\rm tot}\Omega$. However for small objects (which are always our concern) this is not the case and the error bars must be computed from the measured noise properties of the co-add.

4.1.6 Additional masking

Before constructing the photometric catalogs, we zero all pixels contaminated by bright stars in the Tycho catalogs (Høg et al., 2000), replacing them with random noise of appropriate amplitude. Pixels masked in this manner have the 'INTERP' bit set in the input fpM files, so that the downstream analysis can exclude objects that incorporate pixels from a masked region. Regions of the coadd images that incorporate fewer than seven single-epoch images, or which are otherwise `kMask`ed, also have 'INTERP' bits set, as the coadd image quality in these regions (which often include very bright stars or galaxies which saturate in some, but not all, of the single-epoch exposures) is typically poor. This final step results in a catalog with a complex geometry, which will be demonstrated explicitly in Sec. 5.3.

[4]While the fpC images generated from single-epoch data by PHOTO are sky-subtracted, in practice this initial sky subtraction was not sufficiently smooth to avoid the appearance of large background brightness variation in the coadd images. This should not be surprising, as PHOTO has known sky-subtraction problems (Aihara et al., 2011a)

Chapter 5

Catalogs

5.1 Photometric catalogs

Once each coadded image is constructed, we detect objects using the catalog-construction portion of the SDSS photometric pipeline, PHOTO-FRAMES. The details of FRAMES's catalog construction and object measurement process are described more fully elsewhere (Stoughton et al. 2002; Lupton et al., in prep.). It is nevertheless useful to review the important parts of the FRAMES algorithms.

PHOTO-FRAMES requires as input a set of long integer images, and a considerable array of inputs describing the properties of the telescope and the observing conditions. Principal among these is a description of the telescope point-spread function. For single-epoch data, FRAMES uses a principal-component decomposition of the variation of the PSF across five adjacent fields. The components of this decomposition are allowed to vary as a polynomial (typically quadratic) in the image coordinates across each frame. As the coadded images have the same target PSF in every image, this target PSF is stored as the first principal component. For fast computation of object properties, the pipeline uses a double-Gaussian fit to the PSF; as this is the exact form of the target PSF resulting from the rounding kernel applied above, we simply use the target PSF parameters.

FRAMES first smooths the image with the narrower of the two gaussian widths describing the PSF. Collections of connected pixels greater than seven times the standard deviation of the sky noise are marked as objects. Each object is grown by six pixels in each direction. For each object, the list of connected pixels is then culled of peaks less than three times the local standard deviation of the sky.

In order to avoid including objects that represent random noise fluctuations, catalog galaxies are required to have statistically significant ($> 7\sigma$) detections in both the r and i bands. Note that this is a higher threshold than the $> 5\sigma$ cut used at this stage in the standard single-epoch SDSS processing. This was necessitated by the fact that the pixel noise in the `kImage`s is correlated by the convolution process.

In the standard SDSS pipeline, FRAMES then re-bins the image and repeats the search.

We choose not to use objects found in this manner, as the shape measurements of these very low surface brightness galaxies would not be reliable.

This detection algorithm is repeated in each filter separately. Objects detected in multiple bands are merged to contain the union of the pixels in each band if they overlap on the sky. The list of peak positions in each band is preserved. The center of the resulting single object is determined by the location of the highest peak in the r-band. Objects with multiple peaks are deblended: the deblending algorithm assigns image flux to each peak in the parent object.[1]

Once a complete list of deblended peaks (hereafter objects) has been constructed, the properties of each peak are measured. For the purposes of this paper, the most important outputs are the `MODELFLUX` and `MODELFLUX_IVAR` parameters[2], which are determined from the total flux in the best-fit (PSF-convolved) galaxy profile in the r band (comparing the likelihoods for an exponential and a de Vaucouleurs model), with the amplitude re-fit separately to each of the other bands. This flux measure approximates the true, total flux in the r-band, and provides a robust colour measurement, which is crucial for photometric estimates of the object redshift distribution.

The final crucial output of PHOTO-FRAMES, for lensing purposes, is a postage stamp image for every unique object detected in the catalog, except for those objects for which the deblender algorithm failed.

[1]Short descriptions of the SDSS deblending can be found in Stoughton et al. (2002, Sec. 4.4.3) and on the SDSS website at `http://www.sdss.org/dr7/algorithms/deblend.html`. A detailed paper describing the deblender is forthcoming (Lupton et al., *in prep*).

[2]`http://www.sdss3.org/dr8/algorithms/magnitudes.php`

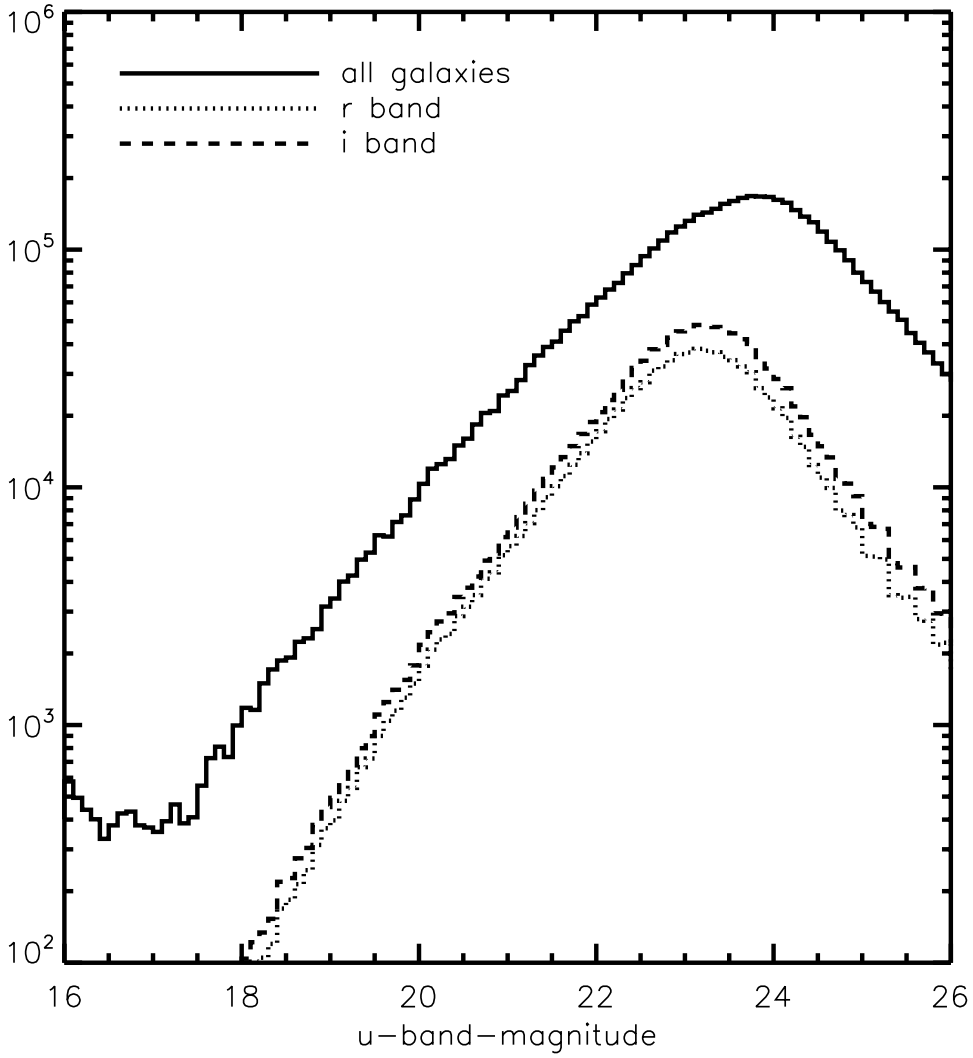

Figure 5.1 The distribution of u-band galaxy apparent magnitudes. The solid line shows the apparent magnitudes for all unique extended objects; dotted and dashed show the r- and i-band lensing catalogs, respectively.

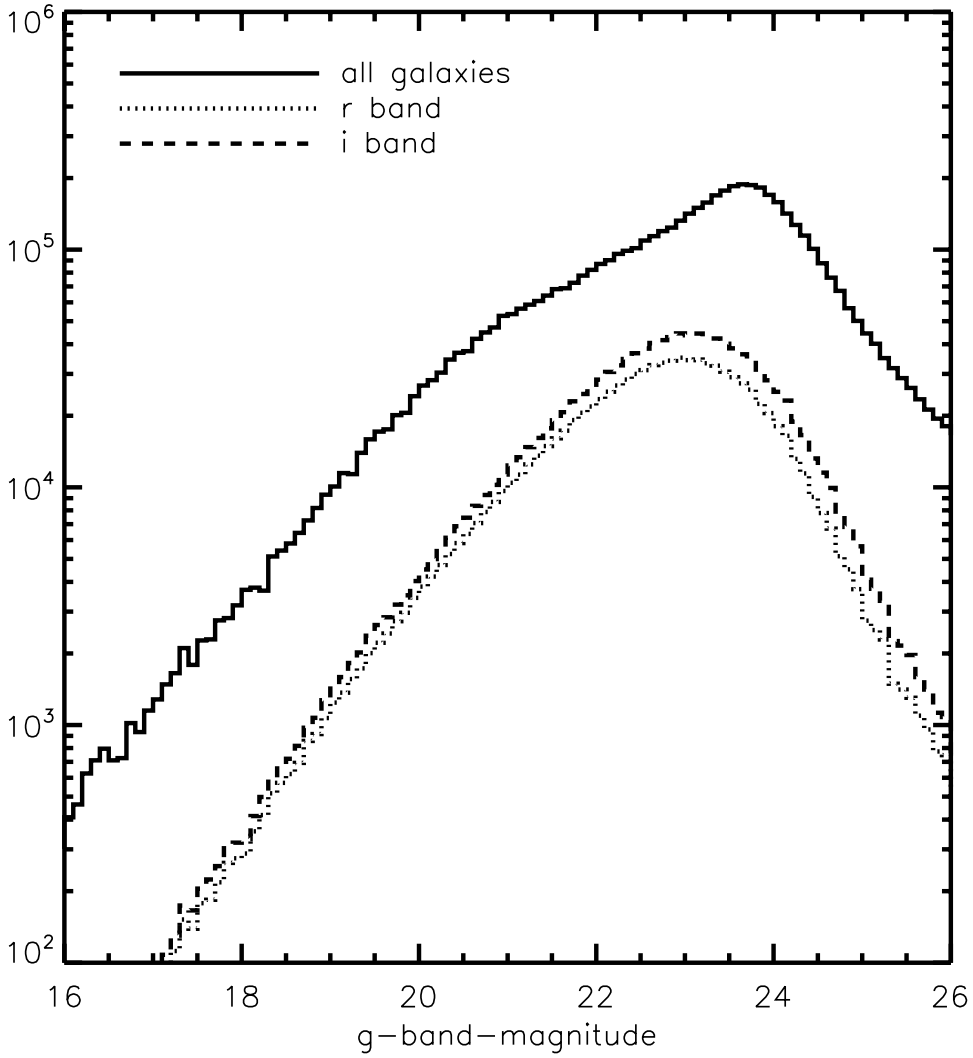

Figure 5.2 The distribution of g-band galaxy apparent magnitudes. The solid line shows the apparent magnitudes for all unique extended objects; dotted and dashed show the r- and i-band lensing catalogs, respectively.

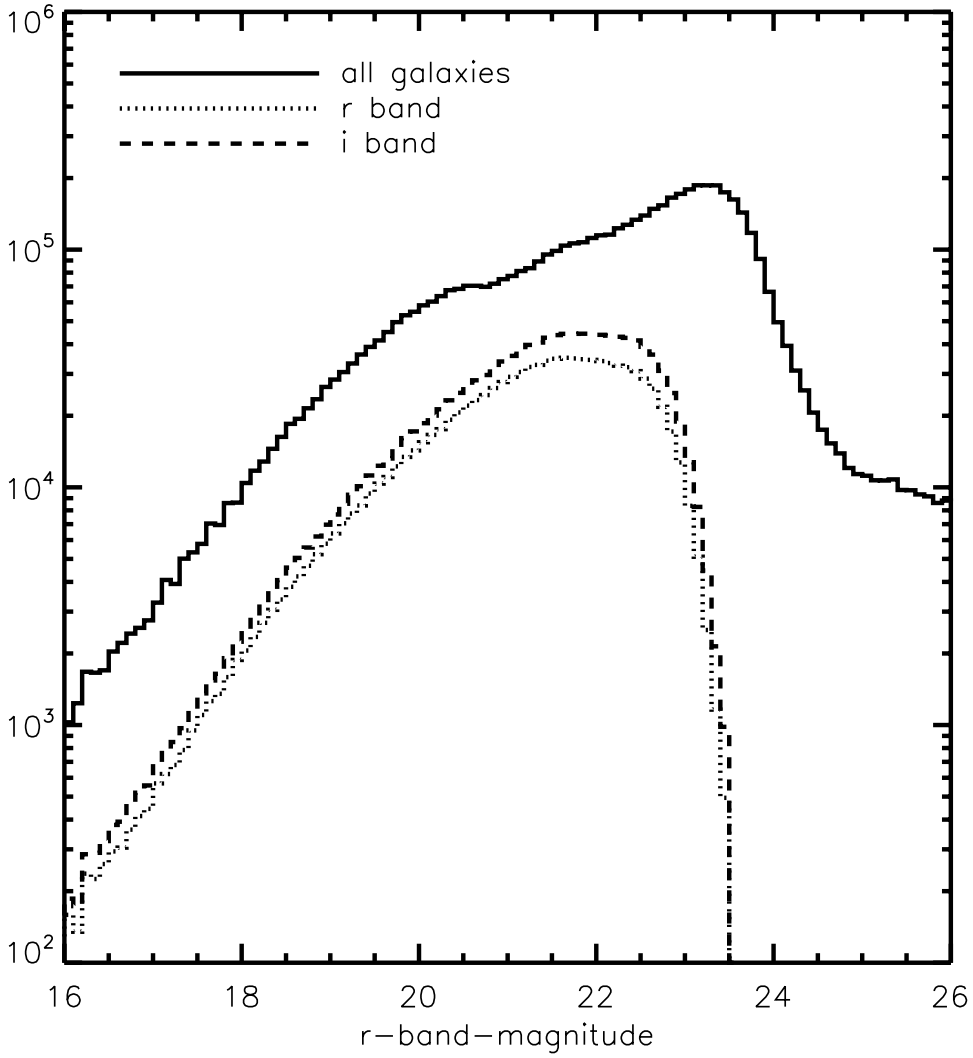

Figure 5.3 The distribution of r-band galaxy apparent magnitudes. The solid line shows the apparent magnitudes for all unique extended objects; dotted and dashed show the r- and i-band lensing catalogs, respectively.

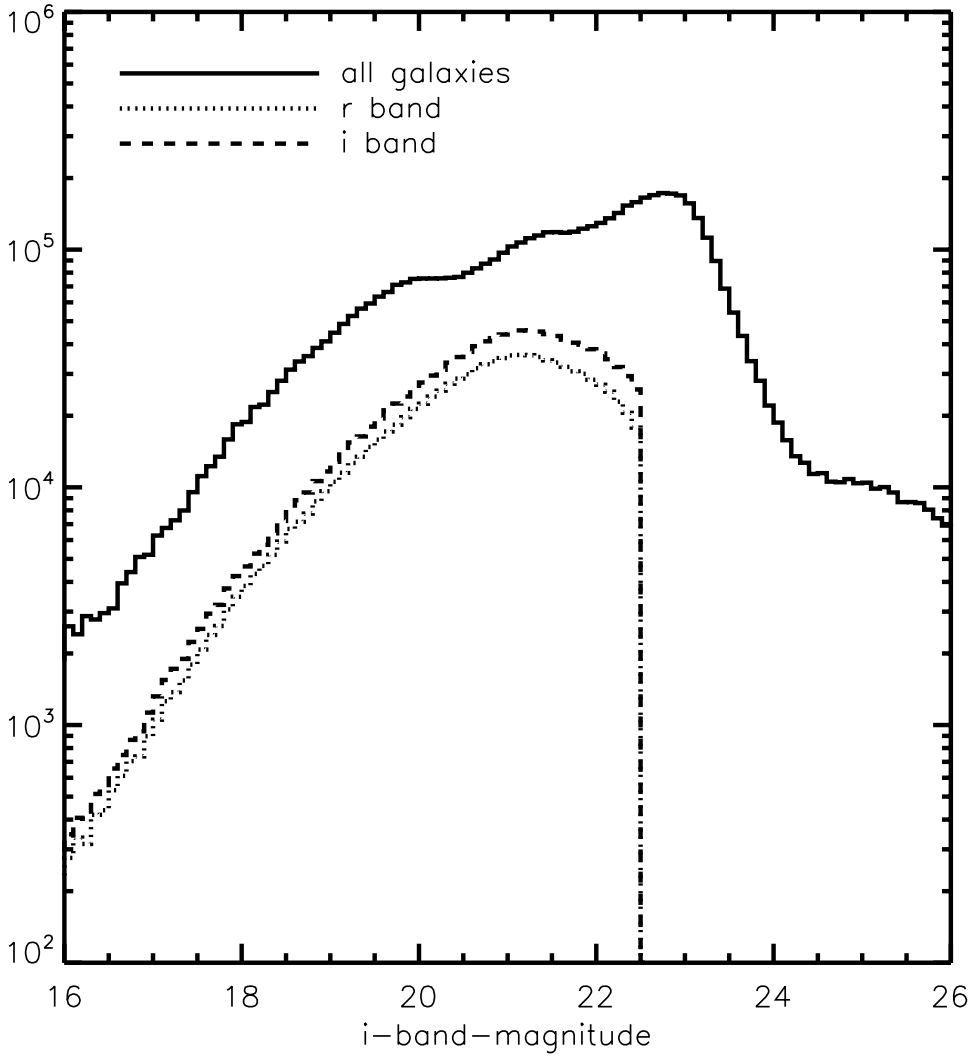

Figure 5.4 The distribution of i-band galaxy apparent magnitudes. The solid line shows the apparent magnitudes for all unique extended objects; dotted and dashed show the r- and i-band lensing catalogs, respectively.

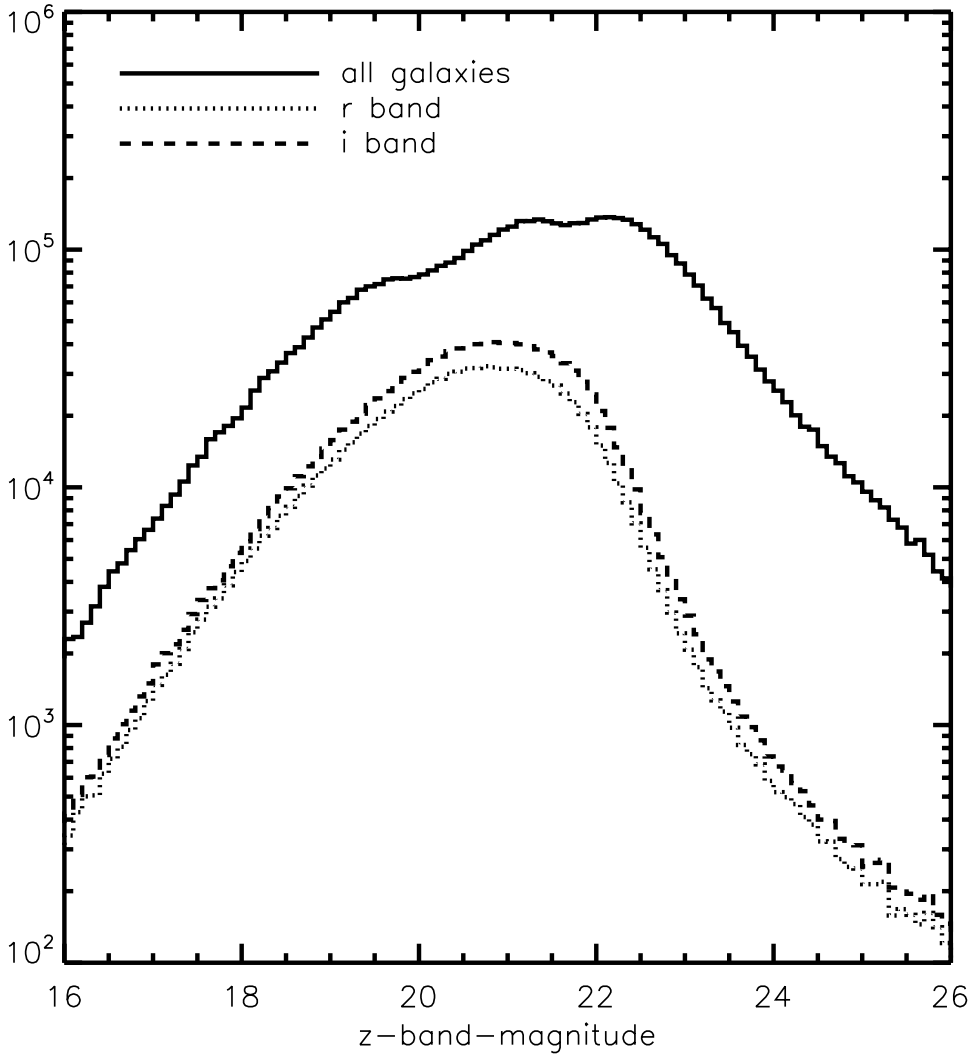

Figure 5.5 The distribution of z-band galaxy apparent magnitudes. The solid line shows the apparent magnitudes for all unique extended objects; dotted and dashed show the r- and i-band lensing catalogs, respectively.

5.2 Lensing Catalog Construction

After PHOTO-FRAMES has constructed an object catalog from the coadded images, we attempt to eliminate spurious detections, stars, and galaxies that are unsuitable for shape measurement. Information from the input mask (fpM) files is propagated through to the catalog, so that objects that incorporate bad pixels identified earlier in the pipeline can be excluded as needed. Due to the nature of the `kImage`s produced by the image coaddition, many of the standard SDSS flags will not be used (e.g, by construction, there are no saturated pixels). As we describe above, masked regions of the `kImage`s are marked as interpolated; objects in the photometric catalog outputs with these bits set are removed from the catalog at this stage. Any galaxies on which the deblending algorithm failed are also excluded, as PHOTO-FRAMES will not generate unique postage stamps for these objects.

PHOTO-FRAMES also attempts to classify objects as "stars" or "galaxies" on the basis of the relative fluxes in the point spread function and galaxy model fits (Lupton et al., *in prep*). Objects that are well-described by a PSF are classified as stars; we do not include these objects in the shape catalog, but set them aside as aids for detecting systematic errors.

To minimize these effects, we also match against a list of all objects classified as stars in the single-epoch SDSS catalogs[3] with apparent magnitudes in i or r band brighter than 15. We remove objects from the catalog within an angular separation of these bright stars that depends on the latter's apparent magnitude as described in Table (5.1).

[3]As our sky coverage is less complete than the single-epoch data, we use the single-epoch catalogs in masking so as to remove objects that are in close proximity to a star that is in one of our masked regions.

Table 5.1 Masking radius as a function of apparent stellar magnitude.

Magnitude range	Masking radius (arcsec)
$r, i < 12$	100
$12 < r, i < 13$	70
$13 < r, i < 14$	50
$14 < r, i < 15$	40
$15 < r, i < 16$	30

In addition to these basic cuts, we cull objects from the lensing catalog if any of the following criteria are satisfied:

1. All objects where the model flux or ellipticity moment measurement failed;
2. All objects within 62 pixels of the beginning or end of a frame;
3. All objects detected only in the rebinned images (BINNED2 or BINNED4);
4. All objects where a bad pixel was was close to the object center (INTERP_CENTER) in either of the r or i bands;
5. All objects that are parents of blends (i.e., not unique)
6. All objects for which the apparent r-band magnitude is fainter than 23.5, and for which the apparent i-band magnitude is fainter than 22.5.

Many of these cuts are applied in only one band. The result of this process is to produce two separate shape catalogs, one for each of the two shape-measurement bands, so there are a small number of galaxies which appear in only one of the two catalogs.

The SDSS photometric pipeline is known to produce significant sky proximity effects, wherein the photometric properties of objects detected near a bright star are systematically biased. The effect of bright stars on the measured tangential shear of nearby galaxies in single-epoch SDSS data is shown in figure 5.8. Motivated by the scales of the effects seen there, we mask the regions around bright stars with a masking radius that depends on the apparent r-band (model) magnitude of the stars as given in table 5.1.

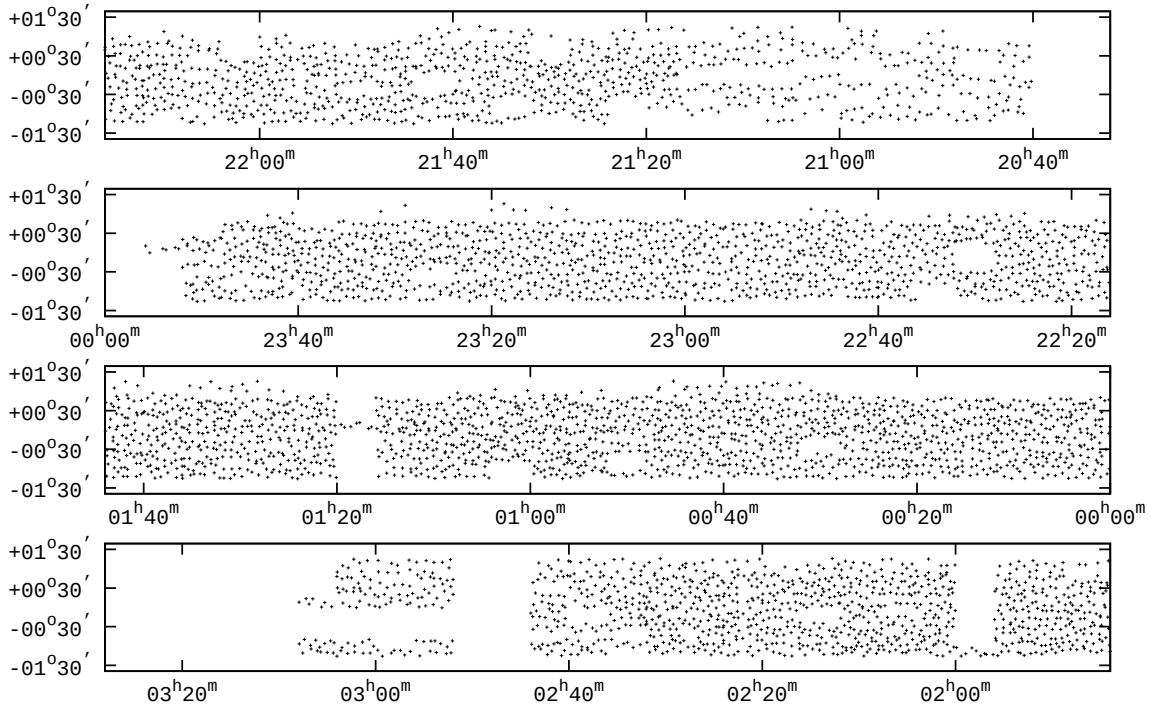

Figure 5.6 The angular distribution (in J2000 right ascension and declination) of the i-band galaxy catalog. A subsample of every 250th galaxy is shown. The r-band sample is identical except for the missing range of $-00°48' <\text{Dec}< -00°24'$. Note the complex survey geometry. Coverage gaps at Dec > 0.8 are primarily due to the severe PSF quality cuts made during the image coaddition step.

5.3 Shape measurement

Once the final shape catalog has been constructed, we use the re-Gaussianization shape measurement method of Hirata & Seljak (2003) to generate an ellipticity measure for each object. The processing code and script are a modification of those used in Mandelbaum et al. (2005). Re-Gaussianization is not an especially modern shape measurement technique, but we have used it previously on SDSS data, it meets our requirements for shear calibration given the expected statistical power, and we had well-tested code that interfaced to PHOTOFRAMES outputs at the time of initiating the cosmic shear project. Therefore we chose to continue using it for this analysis.

5.3.1 Overview of re-Gaussianization

The re-Gaussianization method was an outgrowth of previous work by Bernstein & Jarvis (2002). They defined the adaptive moments M_I of an image I by finding the Gaussian $\mathcal{G}[I]$ that minimizes the L^2 norm $\|I - \mathcal{G}[I]\|$. A Gaussian has 6 parameters – an amplitude, 2 centroids $\bar{\mathbf{x}}_I$, and 3 components of the symmetric covariance matrix – and the last of these is by definition the 2×2 adaptive moment matrix. The ellipticity of the galaxy is defined via its components

$$e_+^{(f)} = \frac{M_{f,xx} - M_{f,yy}}{M_{f,xx} + M_{f,yy}} \tag{5.1}$$

and

$$e_\times^{(f)} = \frac{2M_{f,xy}}{M_{f,xx} + M_{f,yy}}. \tag{5.2}$$

For Gaussian PSFs and galaxies, it is easily seen that the adaptive moment of the intrinsic galaxy image f can be extracted from that of the observed image via $M_f = M_I - M_G$. If the PSF is both circular and Gaussian (a situation that does not arise in practice) then one can relate the ellipticity of the observed image to that of the true galaxy image via the resolution factor R_2:

$$\mathbf{e}^{(f)} = \frac{\mathbf{e}^{(I)}}{R_2} \quad \text{and} \quad R_2 = 1 - \frac{T_G}{T_I}, \tag{5.3}$$

where we have used T to denote the trace of the adaptive moment matrix: e.g., $T_G \equiv M_{G,xx} + M_{G,yy}$. Re-Gaussianization seeks to apply corrections to Eq. (5.3) to correct for the non-Gaussianity of the PSF and the galaxy.[4]

5.3.2 Non-Gaussian galaxies

First is the non-Gaussian galaxy correction – i.e. we consider the case of a Gaussian PSF and non-Gaussian galaxy. Appendix C of Bernstein & Jarvis (2002) used a Taylor

[4]There are also steps in the Hirata & Seljak (2003) code that correct for non-circularity of the PSF. However since the co-add code has already circularized the PSF, these portions of the code are vestigial and we do not describe them here.

expansion method to show that if the galaxy is well-resolved, then in this case Eq. (5.3) could be corrected by using a different formula for the resolution factor,

$$R_2 = 1 - \frac{(1+\beta_{22}^{(I)})T_G}{(1-\beta_{22}^{(I)})T_I}, \tag{5.4}$$

where $\beta_{22}^{(I)}$ is the radial fourth moment,

$$\beta_{22}^{(I)} = \frac{\int (\rho^4 - 4\rho^2 + 2) I(\mathbf{x}) \mathcal{G}[I](\mathbf{x})\, \mathrm{d}^2\mathbf{x}}{2\int I(\mathbf{x})\mathcal{G}[I](\mathbf{x})\, \mathrm{d}^2\mathbf{x}}, \tag{5.5}$$

where $\mathcal{G}[I]$ is the adaptive Gaussian and the rescaled radius ρ is given by

$$\rho \equiv \sqrt{(\mathbf{x}-\bar{\mathbf{x}}_I)\cdot M_I^{-1}(\mathbf{x}-\bar{\mathbf{x}}_I)}. \tag{5.6}$$

This is equivalent to an elliptical version of the $n = 4, m = 0$ polar shapelet (Refregier, 2003; Refregier & Bacon, 2003), and we have $\beta_{22}^{(I)} = 0$ for a Gaussian galaxy (in practice usually $\beta_{22}^{(I)} > 0$).

5.3.3 Non-Gaussian PSF

The shape measurement procedure described in the previous section is only valid where the PSF itself is gaussian. One additional step is required to account for the fact that our rounding kernel was chosen so as to produce a PSF that is the sum of two gaussians. We start by constructing a Gaussian approximation G_1 to the PSF G,

$$G(\mathbf{x}) \approx G_1(\mathbf{x}) = \frac{1}{2\pi\sqrt{\det M_{G_1}}} \exp\left(-\frac{1}{2}\mathbf{x}^T M_{G_1}^{-1}\mathbf{x}\right). \tag{5.7}$$

The choice G_1 is chosen according to the adaptive moments of G. The function G_1 is determined from the centroid and covariance, but the amplitude in Eq. (5.7) is chosen to normalize the Gaussian G_1 to integrate to unity.[5]

We may then define the residual function $\epsilon(\mathbf{x}) = G(\mathbf{x}) - G_1(\mathbf{x})$. It follows that the measured image intensity will satisfy $I = G \otimes f = G_1 \otimes f + \epsilon \otimes f$, where $\otimes$ represents convolution. This can be rearranged to yield:

$$G_1 \otimes f = I - \epsilon \otimes f. \tag{5.8}$$

This equation thus allows us to determine the Gaussian-convolved intrinsic galaxy image I', if we know f. At first glance this does not appear helpful, since if we knew f it would be

[5]The reason for doing this is that while this increases the overall power $\int(\epsilon^2)$ of the residual function, it yields $\int \epsilon = 0$, which ensures that for well-resolved objects (i.e. objects for which the PSF is essentially a δ-function), the "correction" $\epsilon \otimes f_0$ applied by equation (5.10) does not corrupt the image I.

trivial to determine $G \otimes f$. However, f appears in this equation multiplied by (technically, convolved with) a small correction ϵ, so equation (5.8) may be reasonably accurate even if we use an approximate form for f. The simplest approach is to approximate f as a Gaussian with covariance:

$$\begin{aligned} f_0 &= \frac{1}{2\pi\sqrt{\det M_f^{(0)}}} \exp\left(-\frac{1}{2}\mathbf{x}^T[M_f^{(0)}]^{-1}\mathbf{x}\right), \text{ with} \\ M_f^{(0)} &= M_I - M_G, \end{aligned} \tag{5.9}$$

where M_I and M_G are the adaptive covariances of the measured object and PSF, respectively. Then we can define:

$$I' \equiv I - \epsilon \otimes f_0 (\approx G \otimes f). \tag{5.10}$$

The adaptive moments of I' can then be computed, and the PSF correction of Eq. (5.6) can then be applied to recover the intrinsic ellipticity $\mathbf{e}^{(f)}$.

Simple simulations with (noise-free) toy galaxy profiles indicate that this method has shear calibration errors at the level of a few percent depending on the galaxy profile, with the worst performance for de Vaucouleurs profiles at low resolution and high ellipticity (Mandelbaum et al., 2005, fig. 5). Moreover, simulations of SDSS data based on real galaxy profiles from COSMOS, single-epoch SDSS PSFs, and realistic noise levels show that the shear calibration biases are not markedly different under more realistic conditions (Mandelbaum et al., 2012). An investigation of the shear calibration bias for the SDSS cosmic shear sample is presented in Chapter 7.

To select the galaxy sample used for the final analysis, we impose a resolution factor cut at $R_2 > 0.333$ in both r and i (we will justify this choice in Sec. 6.2.3 based on our desire to minimize additive PSF systematics). The parameters of the final shape catalog are shown in Table 5.2, and the survey geometry can be found in Fig. 5.6.

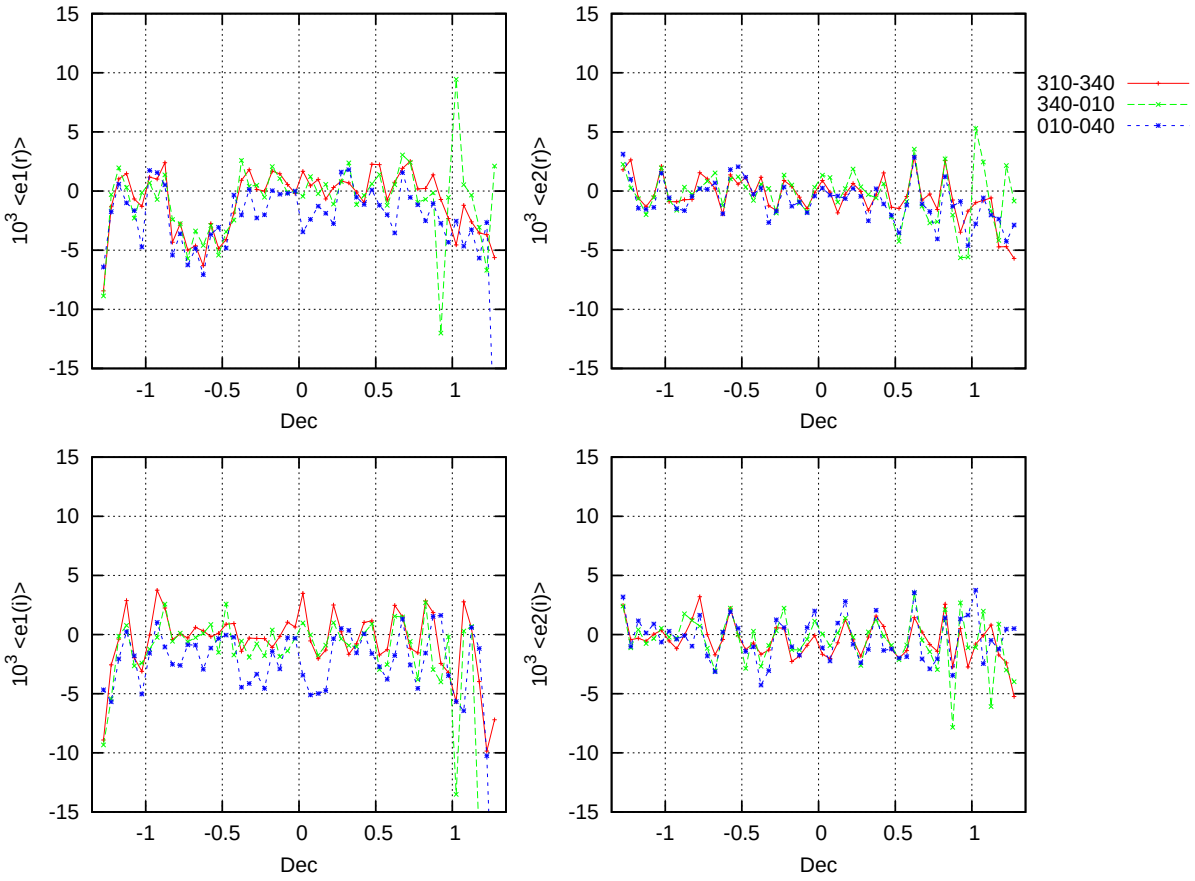

Figure 5.7 The mean ellipticities of stars in the r band as a function of declination for different ranges of right ascension. The top panels show the r band and the bottom panels show the i band, while the left and right panels show different ellipticity components. This was computed using a version of the star catalog prior to final cuts. Note the spurious effect in camcol 2 r band in the e_1 component (declinations -0.8 to $-0.4°$). The apparent magnitude range for this plot was $19.5 < r < 21.5$.

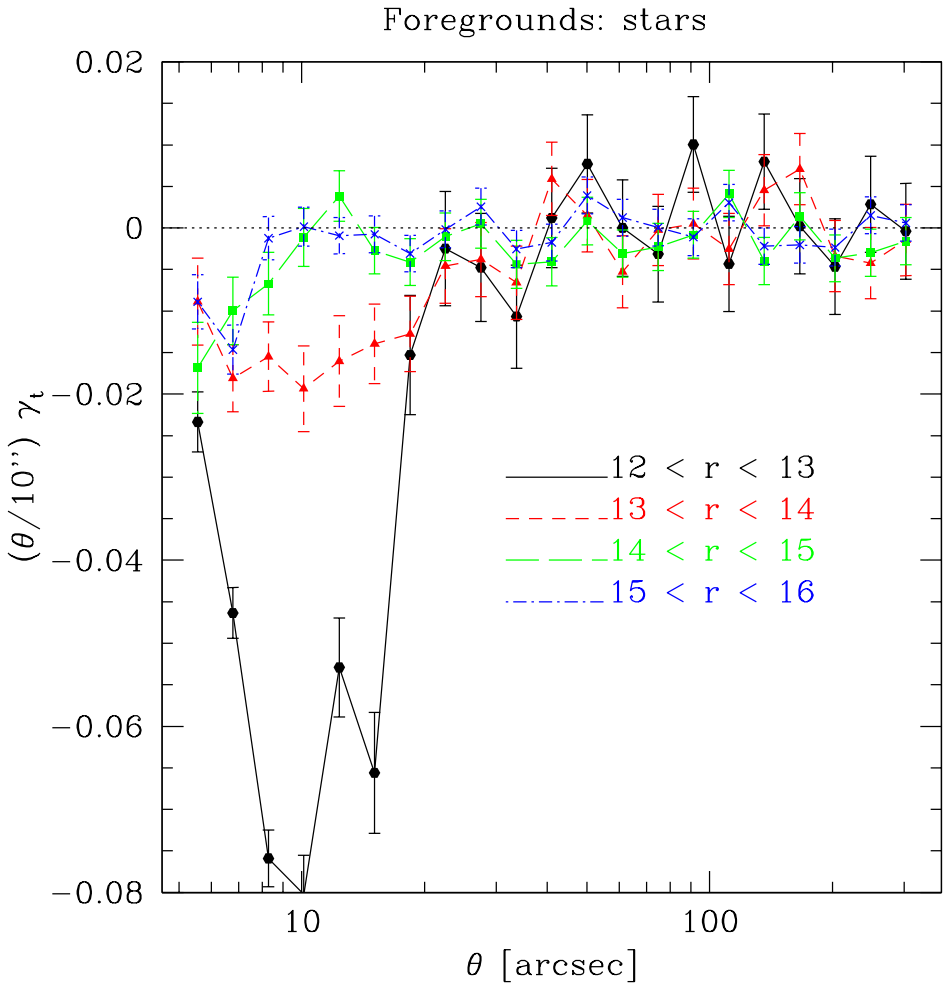

Figure 5.8 Tangential shear γ_t as a function of separation θ from stars, as measured in the single epoch SDSS imaging using the shape catalog from Reyes et al. (2011a). The different lines with points show different bins in r-band stellar apparent magnitude, as labeled on the plot. The ideal expected value of zero is shown as a dotted horizontal line.

Table 5.2 Parameters of the final shape catalog.

Parameter	Value		Units
	r-band	i-band	
Total number of source galaxies	1 328 885		
Number of sources per band	1 067 031	1 251 285	
Effective number of sources downweighted by noise, $N_{\rm eff} = \sum_i \varpi_i$	882 345	1 065 807	
Median magnitude	21.5	20.9	mag AB
Median resolution factor R_2	0.55	0.53	
RMS measured ellipticity per component (noise not subtracted)	0.48	0.47	

Chapter 6

Catalog Systematics Tests

6.1 Correlation function estimation

As stated previously, the primary systematic error of concern in this paper are additive shear systematics, due to PSF ellipticity leaking into the galaxy shapes even after the PSF correction is carried out. This concern will drive our choice of diagnostics to use on the shape catalogs. There are several possible choices for diagnostics that we could use:

1. 1-point statistics of the star and galaxy shapes: For example, we calculate the mean stellar and galaxy ellipticities in bins of some chosen size and look for deviations from zero, including coherent patterns. We use this diagnostic in Sec. 6.2.1.

2. The tangential shear as a function of scale around random points (e.g., Mandelbaum et al., 2005): If there is some additive systematic shear, then on scales that are such that we start losing lens-source pairs off the survey edge, it will show up as a nonzero tangential shear. However, this test alone does not tell us much about the correlations between systematic shears at different points, and therefore we ignore it in favour of more informative tests.

3. Cross-correlations between the stellar shapes and galaxy shapes, as a function of separation θ: These correlation functions tell us not only about the amplitude of any systematic shear, but also about the characteristic scales that are affected by it. This section will describe our methodology for calculating these correlation functions.

4. The B mode shear, which should be zero due to gravitational lensing: While this test is an important one as it can signal a variety of problems with PSF correction, it is not strictly a measure of additive shear systematics. Thus, we neglect this test for this (which focuses on additive shear systematics) until Chapter 8.

6.1.1 The estimator and weighting

In order to compute the star-galaxy cross-correlations, we employ a direct pair-count correlation function code. It is slow (~ 3 hours for 2×10^6 galaxies on a modern laptop) but robust and well-adapted to the Stripe 82 survey geometry. The code sorts the galaxies in order of increasing right ascension α; stars and galaxies galaxies are assigned to the range $-60^\circ < \alpha < +60^\circ$ to avoid unphysical edge effects near $\alpha = 0$. It then loops over all pairs with $|\alpha_1 - \alpha_2| < \theta_{\rm max}$. The usual ellipticity correlation functions can be computed via summation over galaxies i and stars j, e.g.,

$$\xi_{11,\rm psf}(\theta) = \frac{\sum_{\alpha\beta} w_i e_{\alpha 1} \texttt{M\_E1}_\beta}{\sum_{\alpha\beta} w_\alpha} \tag{6.1}$$

and similarly for $\xi_{22,\rm psf}$. Here e_{i1} is the PSF-corrected galaxy e_1 for galaxy index i, and $\texttt{M\_E1}_j$ is the stellar e_1 derived from the adaptive moments described in Sec. 5.3. The sum is over pairs with separation in the relevant θ bin, and we weight each pair according only to the weight associated with the galaxy in each pair:

$$w_i = \frac{1}{0.37^2 + \sigma_e^2}, \tag{6.2}$$

Following Reyes et al. (2011a), we have for weighting purposes adopted an intrinsic shape noise $e_{\rm rms}$ per component of 0.37. The weight of a galaxy relative to a galaxy with perfectly measured shape is

$$\varpi_i = \frac{w_i}{w(\sigma_e = 0)} = \frac{1}{1 + \sigma_e^2/0.37^2}. \tag{6.3}$$

Since the imaging is taken in drift-scan mode, which introduces a potential preferred direction for PSF distortions, we compute our diagnostic correlations between the components aligned along ($-e_1$ and $-\texttt{M\_E1}$) and at 45 degrees to (e_2 and $\texttt{M\_e2}$) the scan direction.

The code works on a flat sky, i.e. equatorial coordinates (α, δ) are approximated as Cartesian coordinates. This is appropriate in the range considered, $|\delta| < 1.274^\circ$, where the maximum distance distortions are $\frac{1}{2}\delta^2_{\rm max} = 2.5 \times 10^{-4}$.

All of our shape correlations are computed over the range $1 < \theta < 120$ arcminutes, evenly spaced in $\log\theta$.

6.1.2 Statistical errors

The direct pair-count correlation function code can directly compute the Poisson error bars, i.e. the error bars neglecting the correlations in $e_{i\alpha}\texttt{M\_E}\alpha_j$ between different pairs. This estimate of the error bar is

$$\sigma^2[\xi_{++}(\theta)] = \frac{\sum_i w_i^2 |\mathbf{e}_i|^2 |\texttt{M\_Ej}|^2}{2\left(\sum_i w_i\right)^2}. \tag{6.4}$$

Equivalently this is the variance in the correlation function that one would estimate if one randomly re-oriented all of the galaxies. As the star-galaxy correlations described here are approximate indicators of the amplitude of the additive PSF shear, and not precision estimates for use in a cosmic shear analysis, we will not attempt to infer the covariance matrix for the full diagonal star-galaxy cross-correlation functions.

6.2 Diagnostics

Here we present our two main systematics tests described in Sec. 6.1, namely the 1-point statistics of the stellar and galaxy ellipticities, and the star-galaxy shape cross-correlations.

6.2.1 Average shapes

We first estimate the influence of residual PSF ellipticities on the galaxy shapes by mapping the stellar shape field.

We computed a set of star shapes binned by right ascension and declination. The stars were chosen to be moderately faint, $19.5 < r < 21.5$, such that they were not used to estimate the PSF model in the single-epoch images that was used to construct the rounding kernel applied to each single epoch image. Figure 5.7 shows the results: the mean stellar ellipticities are usually small, of order 10^{-3}, but in the r band in a particular declination range covered by camcol 2, the shapes are systematically elongated in the scan direction by $-e_1 = 0.005$. We find no significant changes in the amplitude of this artifact when splitting the stellar populations by colour ($r - i <$ or > 0.3) or by apparent magnitude ($r <$ or > 20.5). We did not definitively determine the source of this elongation, but we have confirmed that it appears in the single-epoch SDSS imaging (including the galaxy shape catalogs from Mandelbaum et al. 2005 and Reyes et al. 2011a), so is not merely an artifact of the coaddition and catalog-making process of this work[1]. There is no counterpart feature in the i-band. We exclude all r-band galaxy data in camera column 2 from the r-band shape autocorrelation used for science.

6.2.2 Star-galaxy cross-correlation

Our primary tasks in producing a shear measurement are to demonstrate that the additive systematic shear is below the target threshold set above (Chap. 1), and that our shape measurement method allows us to correctly translate the measured shapes into shears with sufficient accuracy.

[1]One possible explanation (Robert Lupton and Jim Gunn, priv. comm.) is incorrect nonlinearity corrections for the r-band camcol 2 CCD. The stars used to construct the PSF model are sufficiently bright that they require non-linearity corrections, but the stars used for our tests here do not. Therefore if the non-linearity correction is wrong for that CCD, it could affect the PSF model for that CCD alone.

In order to test for residual additive shear systematics, we calculate the cross-correlation between the measured shapes of the stars and those of the galaxies in our sample. Any remaining contribution to the inferred shear field of the galaxies that is sourced by the point-spread function will produce a non-zero cross-correlation. It is important to note that this measurement is performed using measurements of the images of stars *not* used to construct the model PSF; the shape measurements of these objects are not in any sense corrected, and do not incorporate knowledge of the PSF model. As a result, the star-galaxy shape correlations are diagnostic of any spatially-varying PSF modeling errors. Constant multiplicative errors due, for example, to finite-pixel effects, noise bias, or similar problems will be handled in the shear calibration step.

We estimate the star-star and star-galaxy cross correlations as in Eq. (6.1) for all star-galaxy pairs within and between the r and i bands. The results are shown in figures 6.3-6.6.

For the systematic error diagnostics considered here, we are primarily interested in computing the cross-correlation between resolved galaxies and unresolved point sources.

6.2.3 Resolution cuts

Due to the PSF dilution correction applied to all galaxy shapes in Sec. 5.3, noisy measurements of poorly resolved galaxies can significantly amplify any residual additive shear systematics not corrected for in the rounding kernel process. To assess the effects of a resolution cut, we compute the star-galaxy cross-correlations in each band for $R_2 > 0.25$, > 0.333, and > 0.4. Adopting the first of these cuts leads to unacceptably large star-galaxy correlations, indicative of residual systematics. Adopting the second of these of these thresholds appears to be sufficient to minimize the amplitude of the star-galaxy shape correlation signal. As a result, we adopt a cut of $R_2 > 0.333$ for both the i and r-band galaxy catalogs. This resolution cut corresponds to galaxies with a typical half-light radius (as determined from the SDSS model fits) of 0.7 arcseconds. Any potential selection bias resulting from resolution cuts will be dealt with when we derive the empirical shear calibration.

6.2.4 Star-galaxy separation

A nonzero amplitude of ξ_{sg} can also be produced by imperfect star-galaxy separation. Poorly-resolved galaxies masquerading as stars sample both the PSF- and cosmic shear-sourced shape fields. If the fraction of stars that are actually mistakenly classified as galaxies is $f_{\rm gal}$, then the measured ξ_{sg} will include a contribution proportional to $f_{gal}\xi_{\gamma}$. As the ellipticity of nearly-unresolved galaxies will be diluted by PSF convolution, this represents an upper limit to the level of star-galaxy correlation that can be introduced via imperfect star-galaxy separation.

The PHOTO-FRAMES pipeline classifies an object as a star or a galaxy on the basis of the relative fluxes of PSF and galaxy model fits to the object's surface brightness profile. As an independent check on this scheme, we have defined a sample of stars for which aperture-matched UKIRT Infrared Deep Sky Survey (UKIDDS) colours are available. The UKIDSS

project is defined in Lawrence et al. (2007). UKIDSS uses the UKIRT Wide Field Camera (WFCAM; Casali et al. 2007). The photometric system is described in Hewett et al. (2006), and the calibration is described in Hodgkin et al. (2009). The pipeline processing and science archive are described in Hambly et al. (2008). Stars and galaxies separate fairly cleanly in $J-K, r-i$ colour space, so we attempt to use this matched catalog to put some limits on galactic contamination of the stellar sample. This constraint on $f_{\rm gal}$ will give us our upper limit $f_{\rm gal}\xi_\gamma$ on the ξ_{sg} due to contamination of the star sample by galaxies.

We match the objects classified as stars in both bands from our coadd to UKIDDS objects with valid $J-K$ colours; objects with angular separations between the two catalogs less than one arcsecond are identified. We find 93 753 such stars. Of these, 11 331, or 12 per cent, have $J-K, r-i$ colours inconsistent with the stellar population. The UKIDDS matches are shallower than the rest of the catalog in the i band, but of comparable depth in the r band. Only 16 per cent of our stars have UKIDDS matches in either band, however, so the contamination fraction is not well-constrained in the entire star sample.

If, however, this fraction is representative of the galaxy contamination in the entire stellar catalog, then (for an unresolved population with a typical resolution just below our resolution cut), then it would explain a substantial fraction of the residual PSF systematic amplitude that we see.

As a test for this, we compute the star-galaxy shape correlation using only those objects identified as stars in the manner described above. For this population, the amplitude of the star-galaxy correlation is not significantly reduced below the star-galaxy correlations for the full sample, so we conclude that imperfect star-galaxy separation is not a primary source of the systematics signal present in the catalog.

After all of the above cuts have been applied, the final shape catalog consists of 1 067 031 r-band and 1 251 285 i-band shape measurements, over an effective area of 140 and 168 square degrees, respectively.

6.3 Discussion

We have constructed deep, lensing-optimized coadd imaging of the SDSS equatorial stripe. The procedure is designed to enable the construction of a catalog suitable for weak lensing measurements by suppressing the effects of PSF anisotropy on the measured galaxy shapes below the level of statistical error achievable with a cosmic shear survey on this Stripe.

We have measured the star-galaxy and star-star cross-correlations in order to constrain the amplitude of spurious, PSF-induced shape correlations in the catalogs we have obtained from coadding Stripe 82 imaging.

This procedure is successful if and only if it renders the PSF shape distortions sufficiently small that they are negligible compared to the statistical errors expected for a cosmic shear signal in this survey. To estimate the amplitude and scale-dependence of the residual PSF systematics, we fit a power law of the form:

$$\xi_{sg} = A\theta^{-p} \tag{6.5}$$

to the average of the four measured star-galaxy cross-correlations, using the Poisson errors output by the correlation function code. The best-fit power law and average star-galaxy correlations are shown in figure 6.11.

We compare the ratio of this best fit to the shot-noise errors expected for a shape-shape correlation function for this survey. To estimate the shot noise, we follow Schneider et al. (2002a) to calculate the statistical errors expected due to shot noise for a 168 square degree lensing survey with an effective source surface density of 2 galaxies per arcmin^2:

$$\mathrm{Var}\,(\xi) = \left(3.979 \times 10^{-9}\right) \left(\frac{\sigma_e}{0.3}\right)^4 \left(\frac{\mathrm{Area}}{1\,\mathrm{deg}^2}\right)^{-1} \times \left(\frac{n_{\mathrm{eff}}}{30\,\mathrm{arcmin}^{-2}}\right)^{-2} \left(\frac{\theta}{1\,\mathrm{arcmin}}\right)^{-2} \tag{6.6}$$

The ratio of the systematics amplitude to the shot noise is shown as a function of scale in figure 6.12. From this, we can see that PSF systematics for these data should be, on average, 50 per cent of the size of the statistical error budget for a cosmic shear measurement with this catalog; on degree scales, this becomes comparable to the shot-noise errors.

As discussed above, this is an upper limit; imperfect star-galaxy separation at the level of a few per cent can produce a star-galaxy correlation signal in the absence of uncorrected PSF effect, the response of a galaxy shape to a PSF anisotropy is typically less than unity, and the Poisson error estimate will underestimate the true variance on larger scales.

In addition, masks defined as sets of pixels can introduce a shape selection bias. We tested the effects of masking on the spurious shear statistics during the catalog-making step by applying a strict cut to eliminate those regions of the coadd imaging with fewer than seven contributing single-epoch images. Introducing this cut actually increased the spurious shear amplitude; the star-galaxy correlations in the presence of this more aggressive masking step reach an amplitude at degree scales of 10^{-5}. In later chapters, we use detailed image simulations to estimate the fraction of galaxies that are rejected *only* due to proximity to mask boundaries at ~ 1 per cent, and we will discuss in detail whether that level of rejection due to mask boundaries is enough to explain the observed star-galaxy correlations.

The relative contributions of mask selection and PSF anisotropy biases can be ascertained from the relative amplitudes of the star-star and star-galaxy correlation functions. A PSF anisotropy will produce a similar signal in both metrics. The stellar shape dispersion is much smaller than that of the galaxies, so a selection bias will produce a much larger systematics signal in the star-galaxy correlation function than in the star-star correlation functions. This is indeed the case, as shown in Figs.6.13-6.16 – substantial evidence that mask selection bias will be a significant fraction of the systematic error budget. Excluding objects near the boundaries of masked regions on the basis of their centroid positions could remove this effect; however, as Fig. 6.12 shows, the statistical errors should dominate for this catalog, so reducing the catalog further at this stage would not improve the quality of a final cosmic shear measurement.

These results suggest that a cosmic shear analysis that is statistics-limited is possible with these data. We have shown that the effects of the point-spread function are small compared to the statistical errors. The mask selection bias is larger, but still on average significantly smaller than the expected statistical errors.

The systematics floor for the rounding kernel method we have employed here is set by the SDSS PSF model. Inaccuracies in this PSF model are documented both here (fig. 5.7) and in other work (Reyes et al., 2011a). Coherent variations in the PSF model errors in both components across the camera columns are visible with a characteristic amplitude of 2×10^{-3}. Aside from the very striking and atypical effect seen in the r band in camcol 2, it is likely that the shortcomings of the polynomial interpolation method employed in PHOTO play an important role here, as documented in Bergé et al. (2012) for more general simulated ground-based data. As this is close to the level of residual PSF systematics seen in our final lensing catalog, it is very likely that an improvement in the underlying model construction would allow the rounding kernel method deployed here to achieve a much greater level of sytematics control.

The masking problem is not extensively treated in the literature; to the knowledge of the authors, it has not been taken into account in existing studies. It is standard in modern photometric pipelines to define the survey mask and object rejection algorithms in terms of sets of pixels, rather than (for example) galaxy centroids, which is the ultimate source of the masking bias we see here. This effect will be important to take into account in the photometric pipeline construction in the next generation of lensing measurements.

The PSF correction method described here is suitable for deployment in the next generation of weak lensing surveys. All of these surveys will include multi-epoch data over their full footprint; where the image quality (as measured by PSF isotropy and size) distribution is fairly narrow, the rounding kernel method will result in little loss of information.

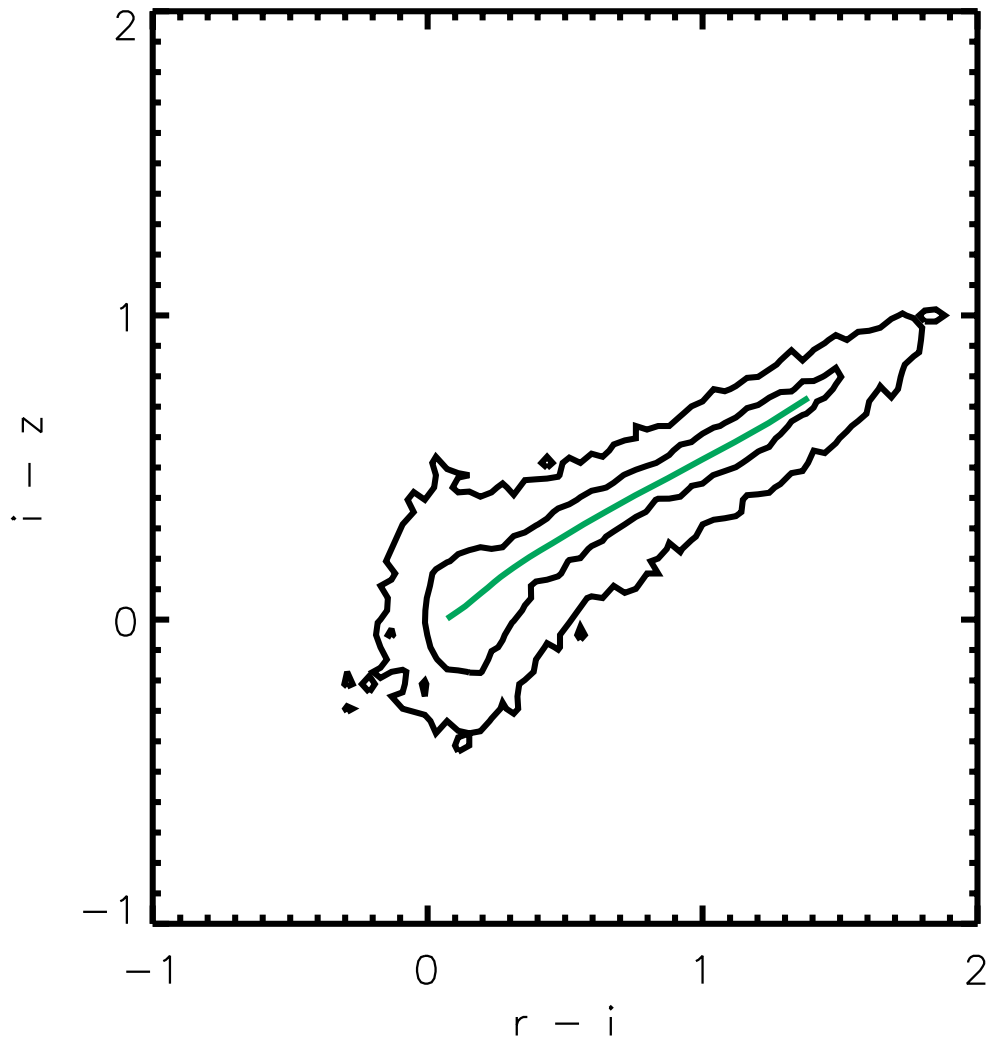

Figure 6.1 $r - i$ vs $i - z$ colours for objects identified as stars using PHOTO's star-galaxy separation based on the concentration of the light profile; the contours containing 68 and 95 per cent of the density are shown. The stellar locus is shown as a solid line through the center of the contours.

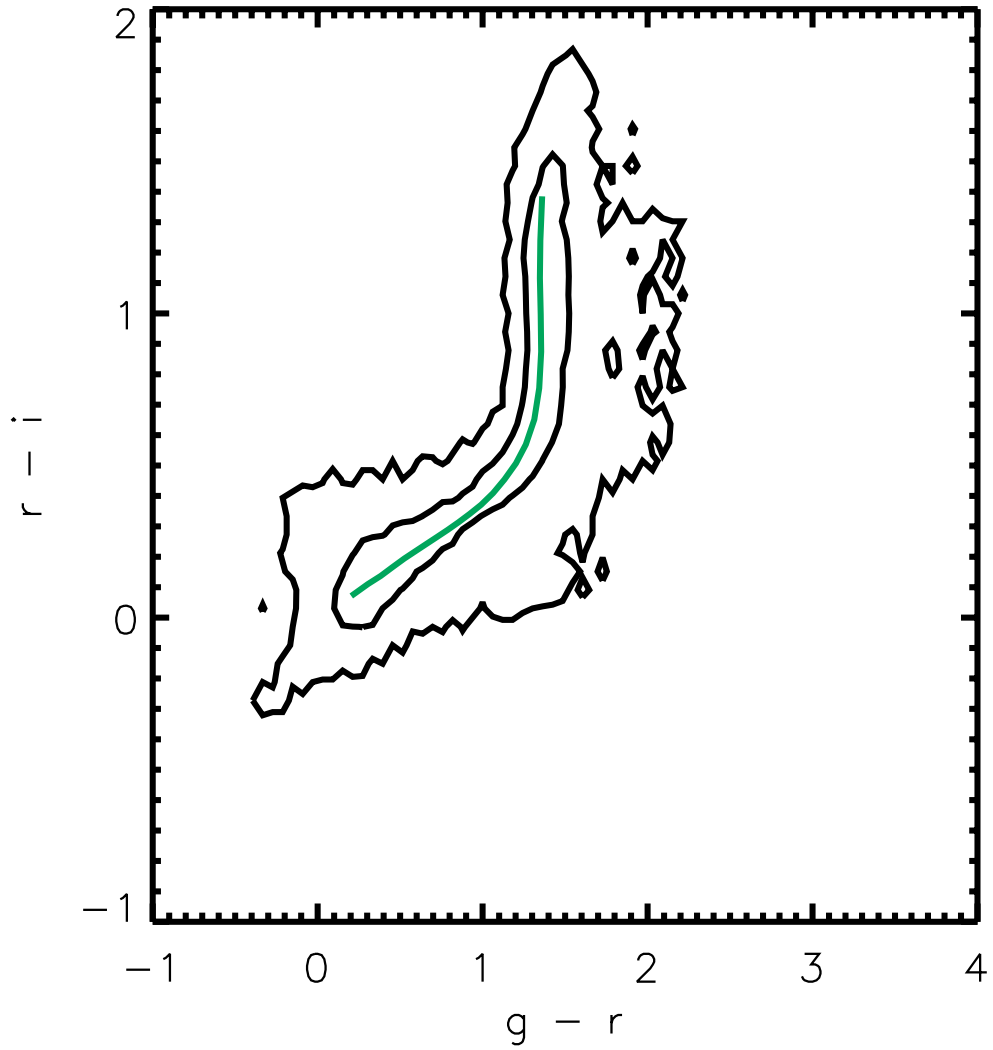

Figure 6.2 $g - r$ vs $r - i$ colours for objects identified as stars using PHOTO's star-galaxy separation based on the concentration of the light profile; the contours containing 68 and 95 per cent of the density are shown. The stellar locus is shown as a solid line through the center of the contours.

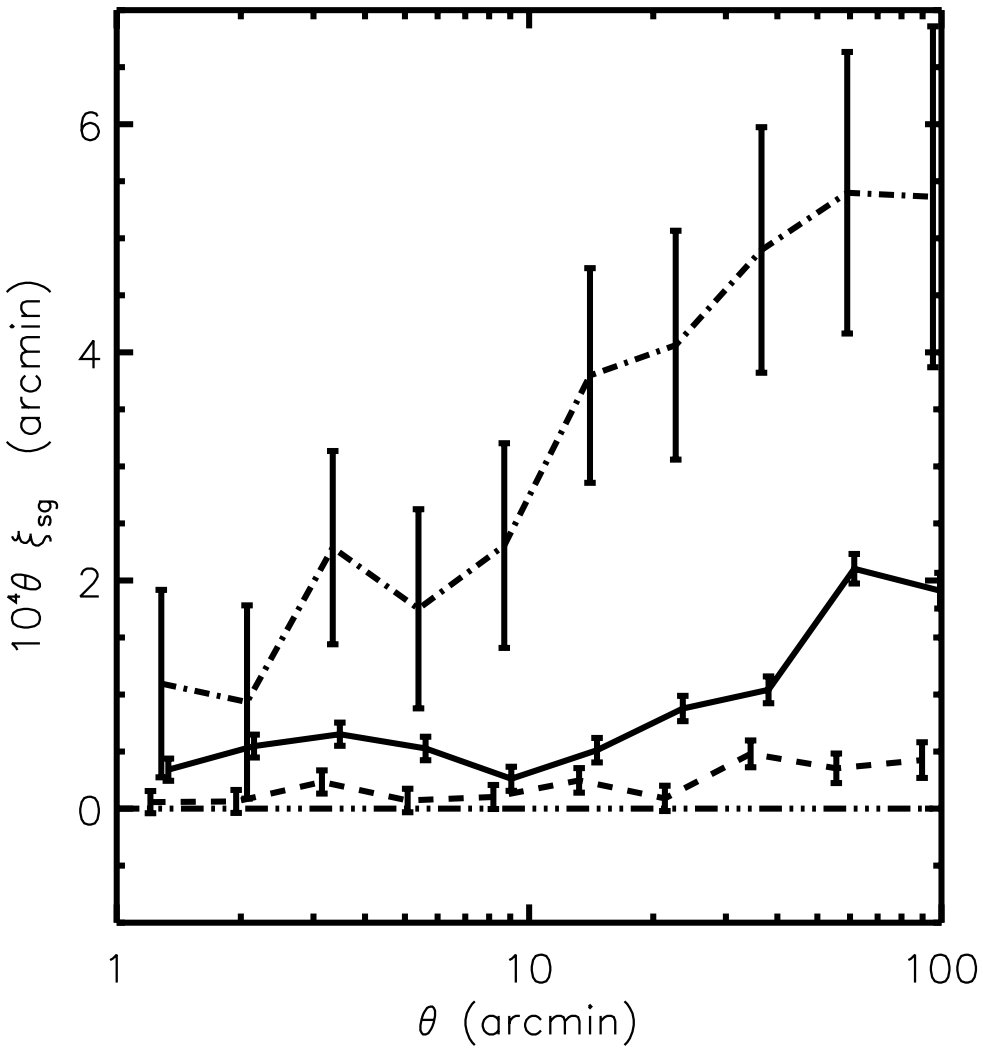

Figure 6.3 The cross-correlation of i-band star shape moments and i-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

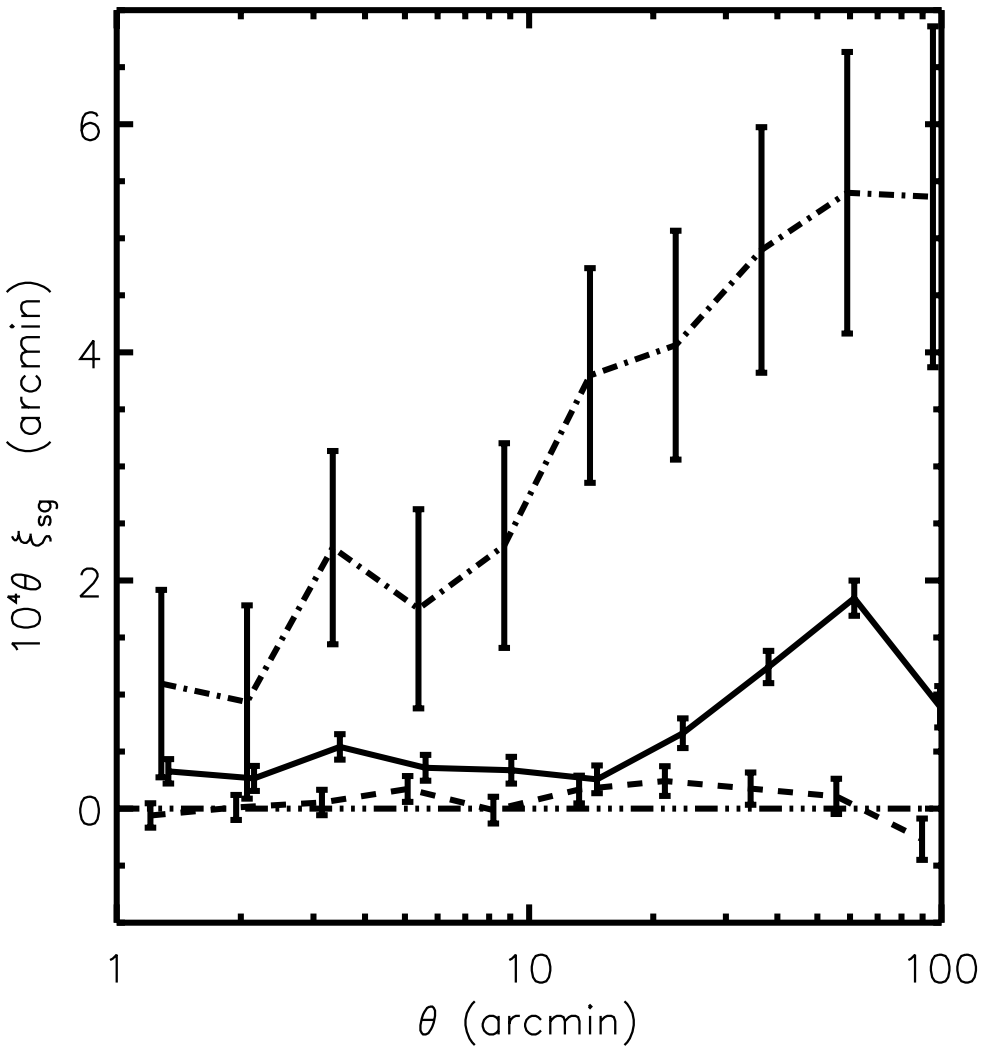

Figure 6.4 The cross-correlation of r-band star shape moments and r-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

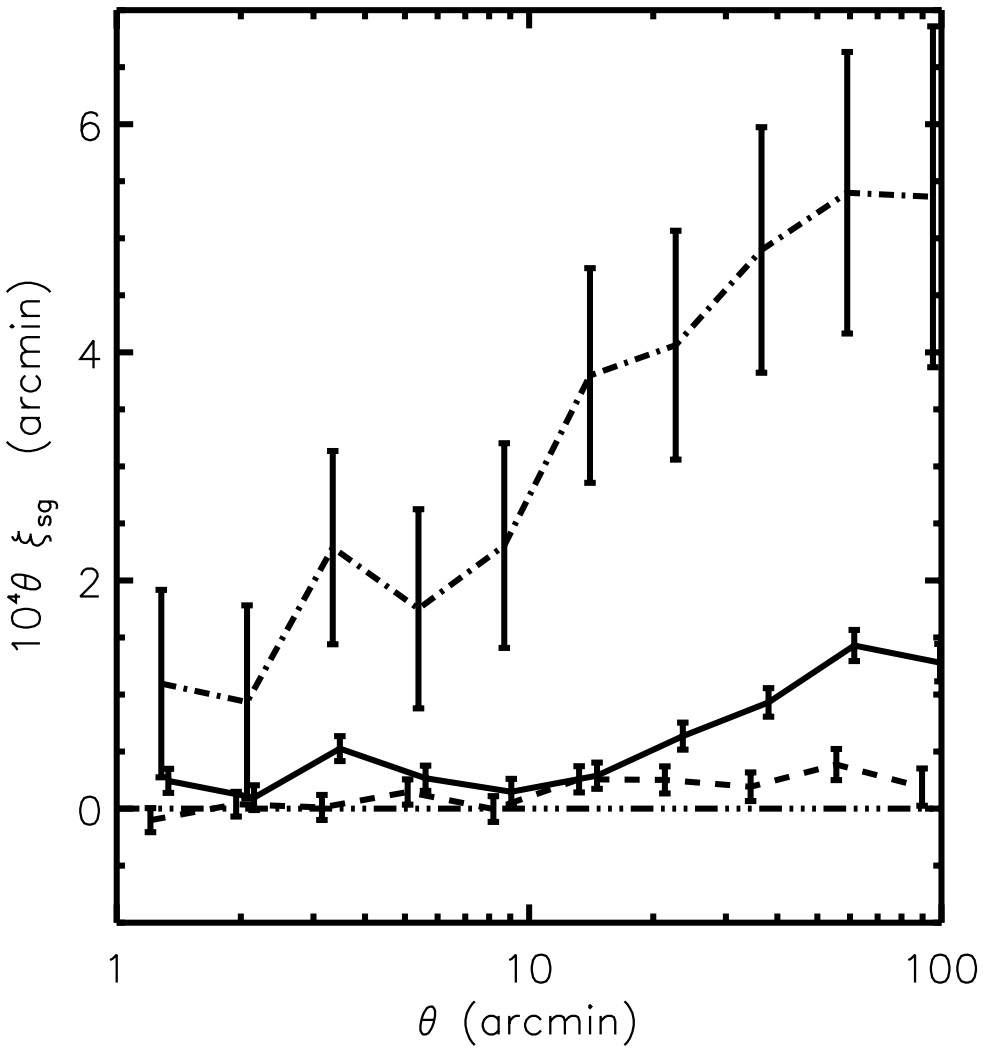

Figure 6.5 The cross-correlation of r-band star shape moments and i-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

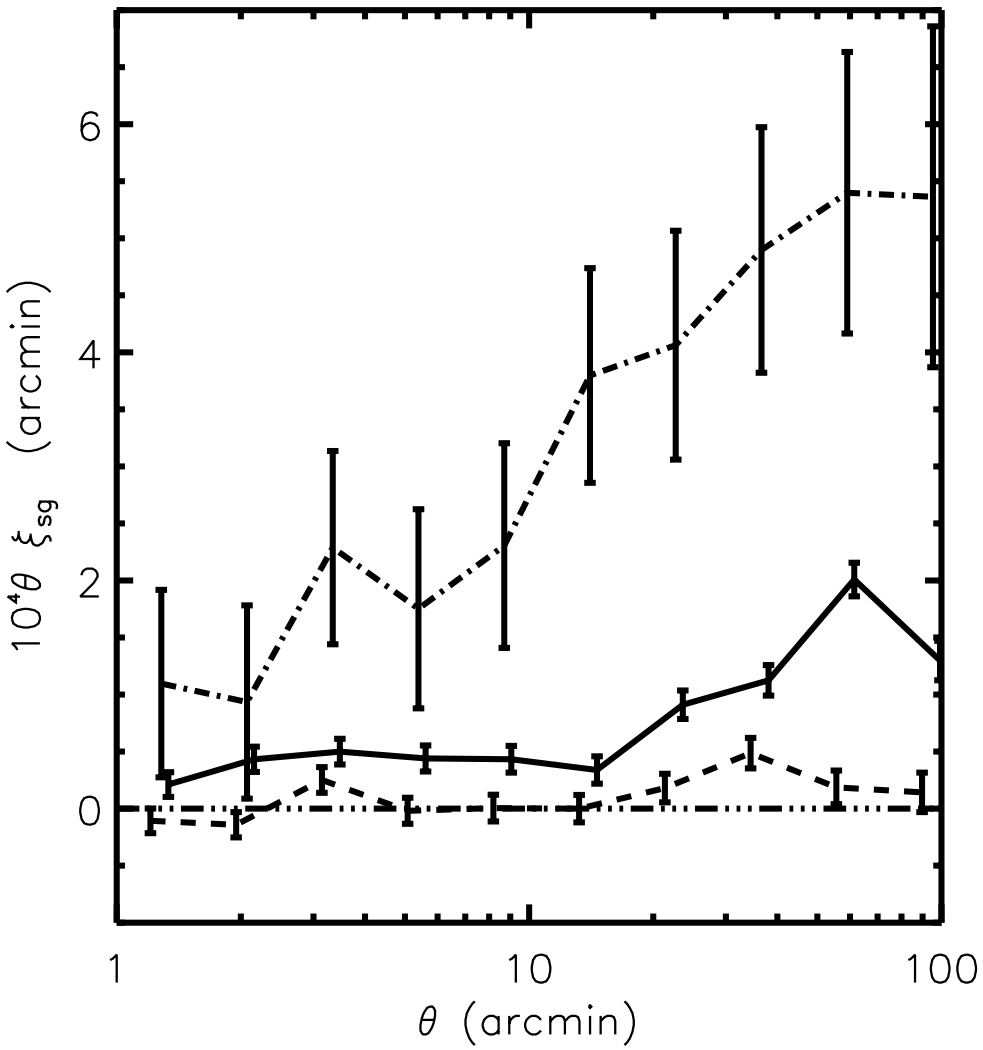

Figure 6.6 The cross-correlation of i-band star shape moments and r-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

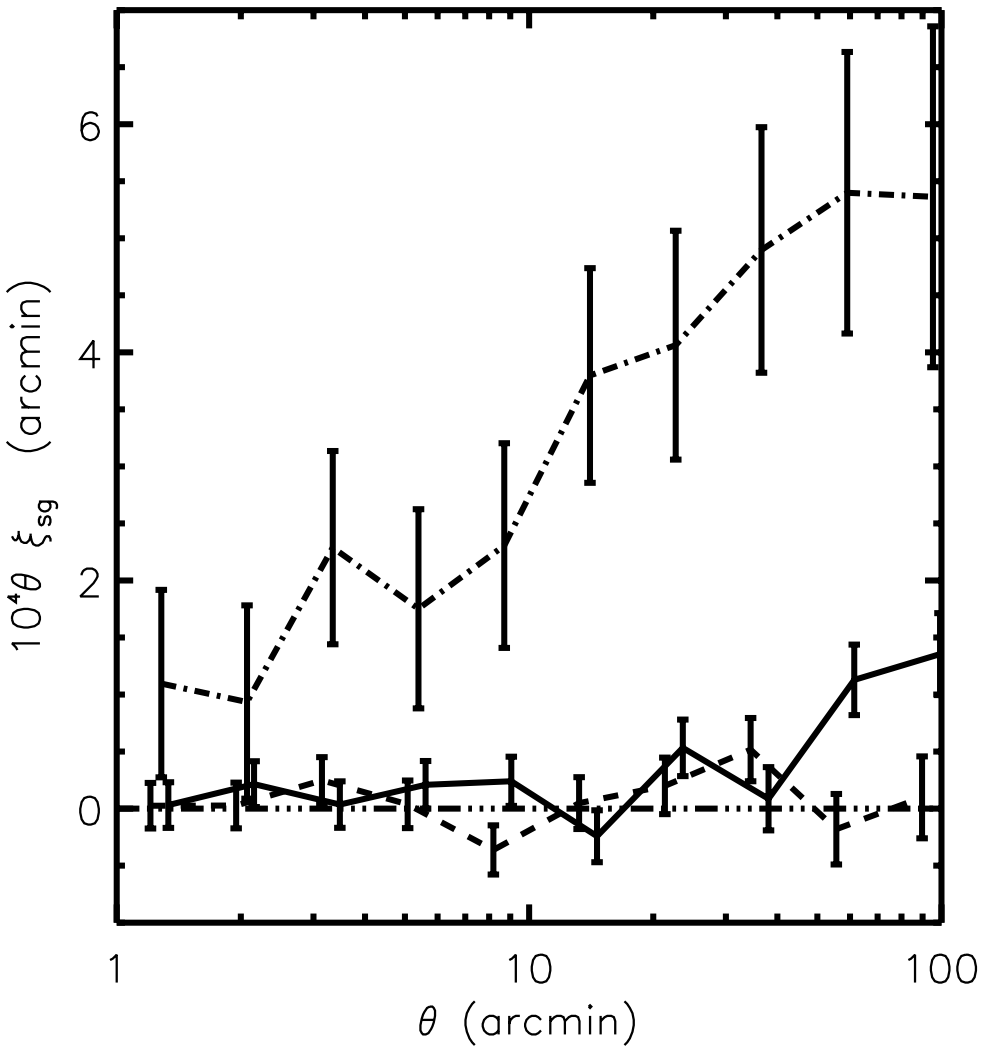

Figure 6.7 The cross-correlation of r-band, UKIDDS-selected star shape moments and r-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

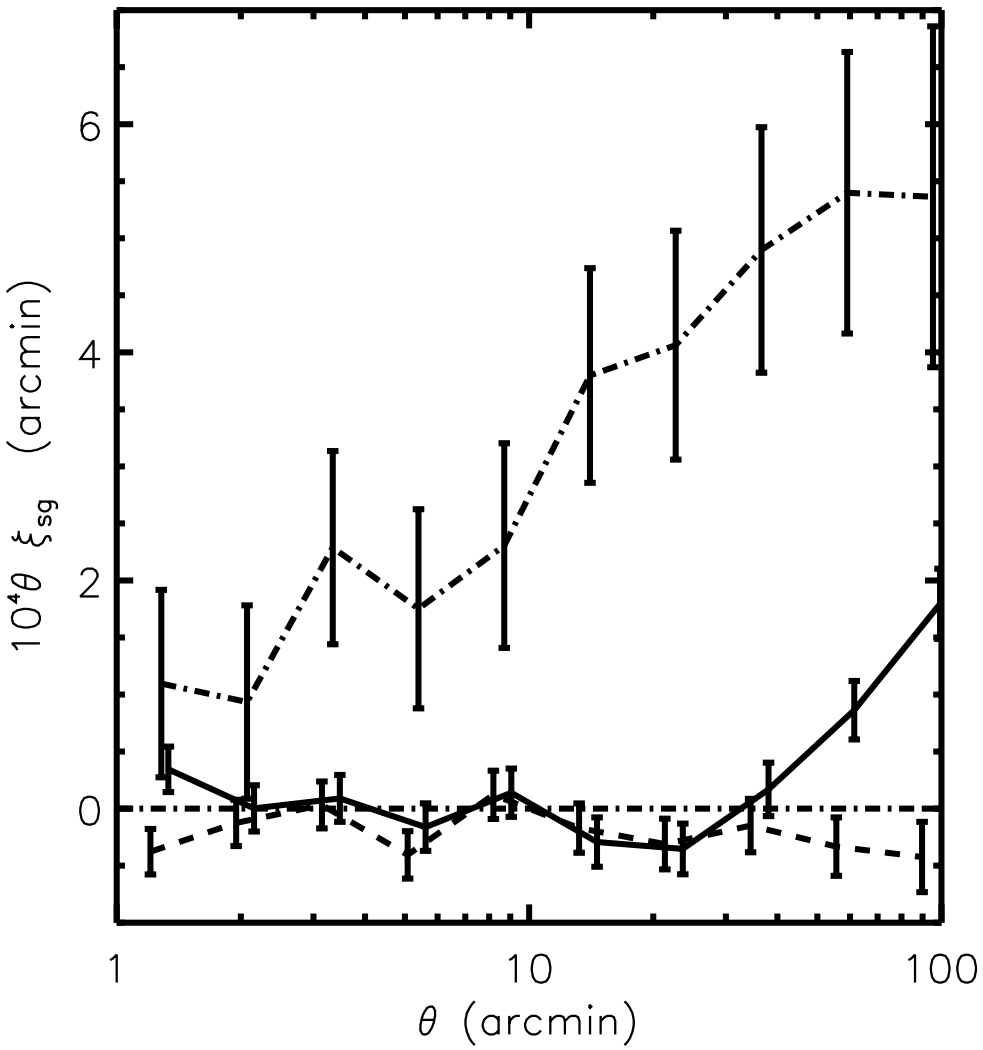

Figure 6.8 The cross-correlation of i-band, UKIDDS-selected star shape moments and i-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

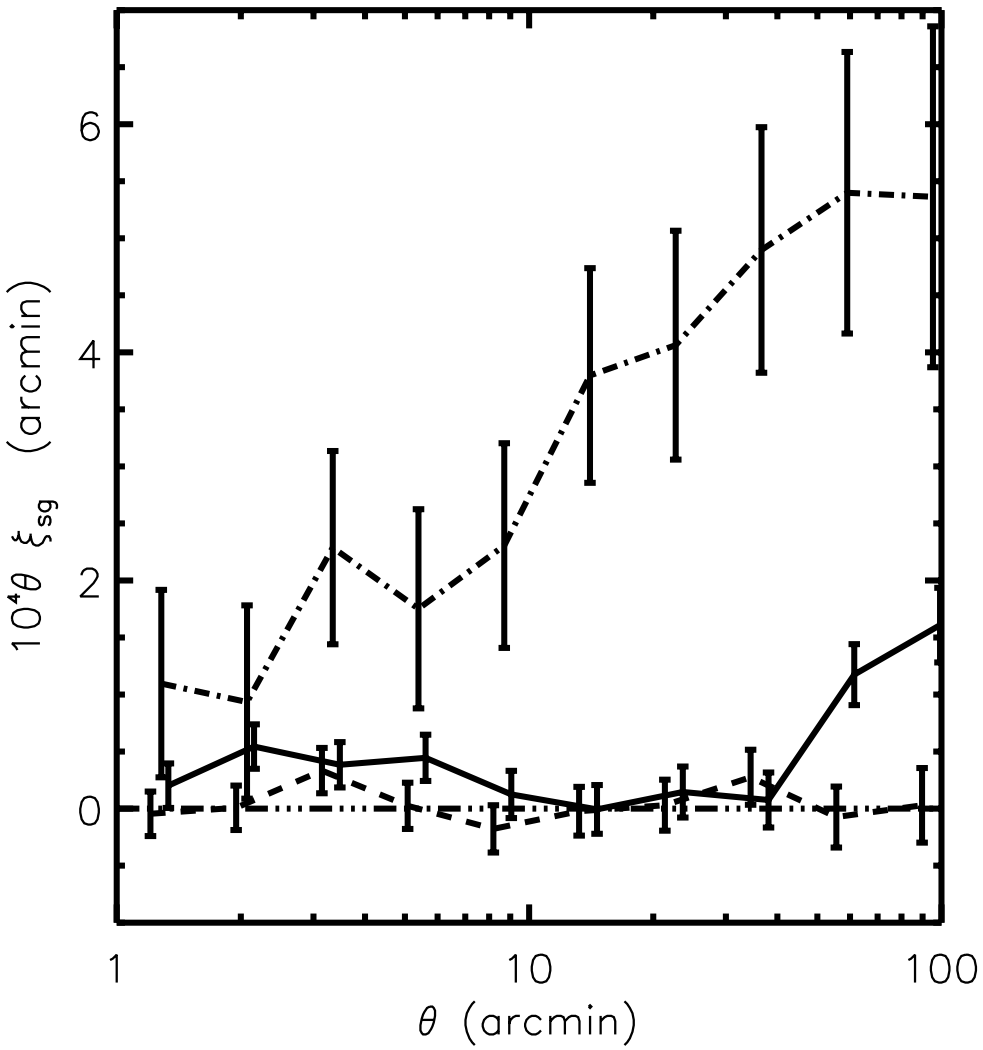

Figure 6.9 The cross-correlation of r-band, UKIDDS-selected star shape moments and i-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

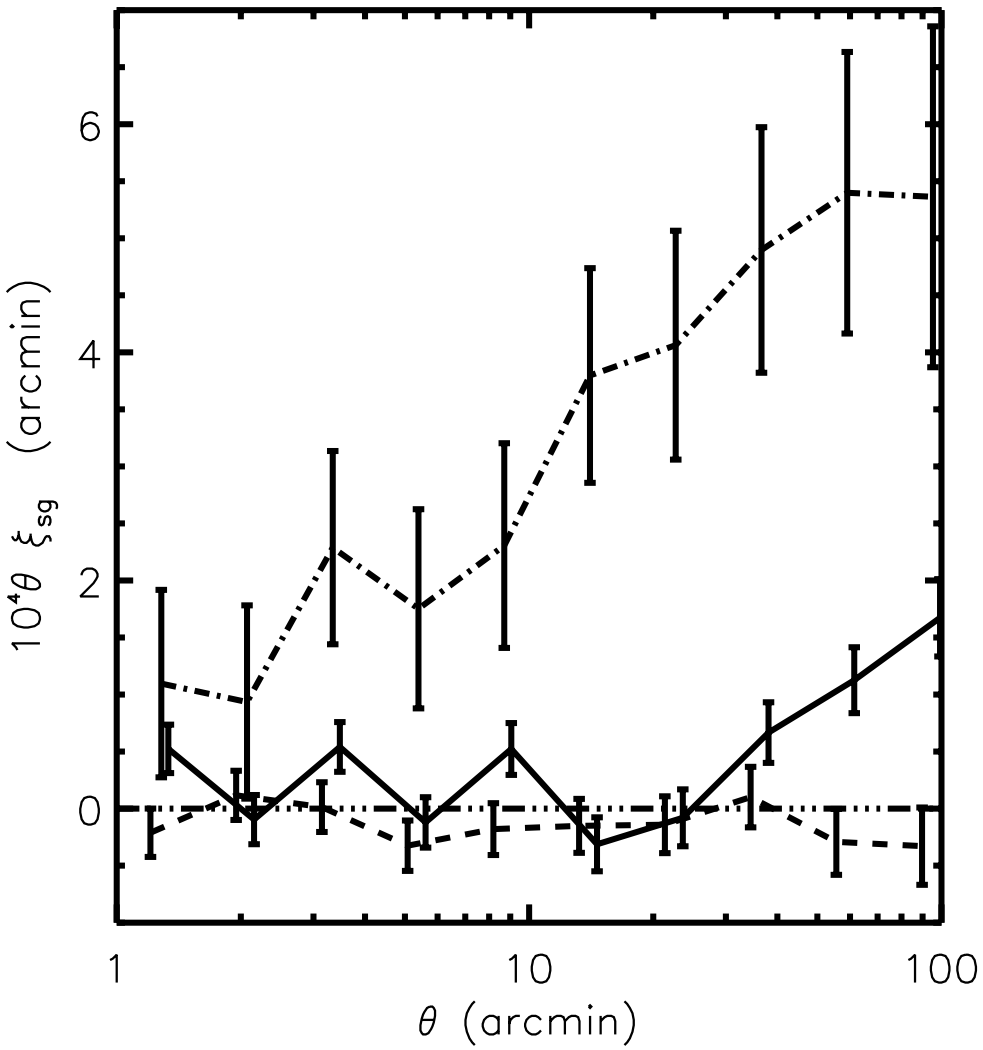

Figure 6.10 The cross-correlation of i-band, UKIDDS-selected star shape moments and r-band galaxy shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

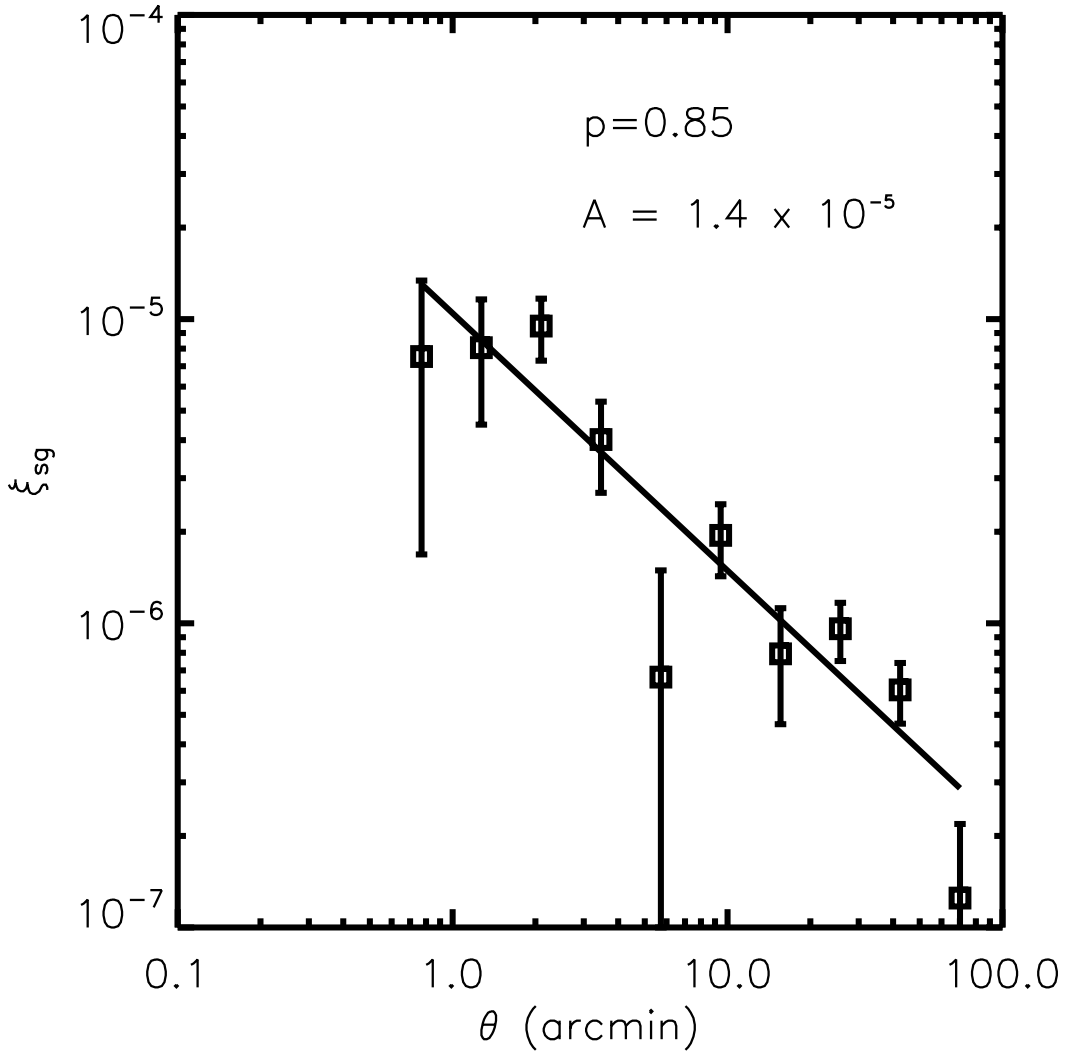

Figure 6.11 Average of the $r-i$, $i-r$, $r-r$, and $i-i$ star-galaxy cross-correlation functions.

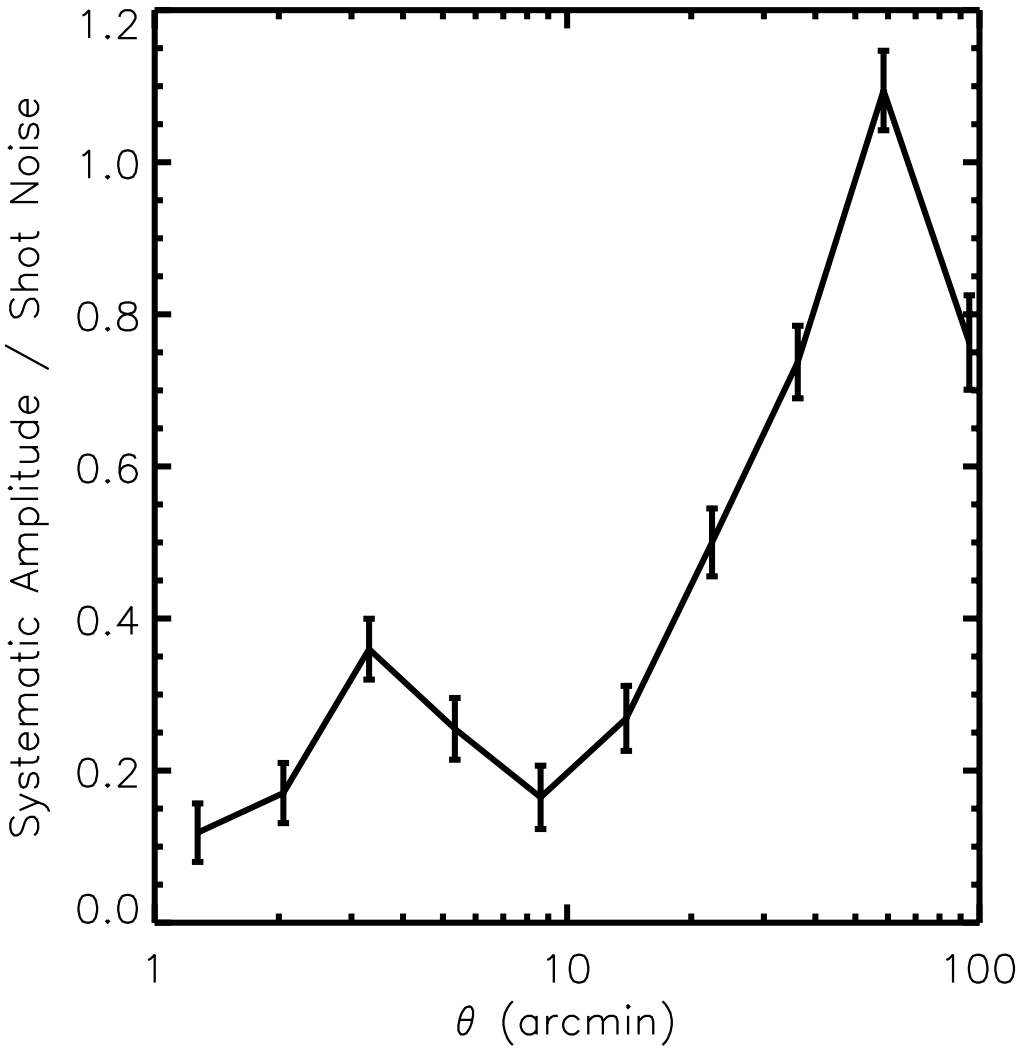

Figure 6.12 Ratio of the best-fit star-galaxy cross-correlation power law to the expected shot-noise errors for a cosmic shear measurement using the catalogs described here. As the star-galaxy amplitude is only poorly constrained, this should be taken as a rough indication of the level of significance of the systematics.

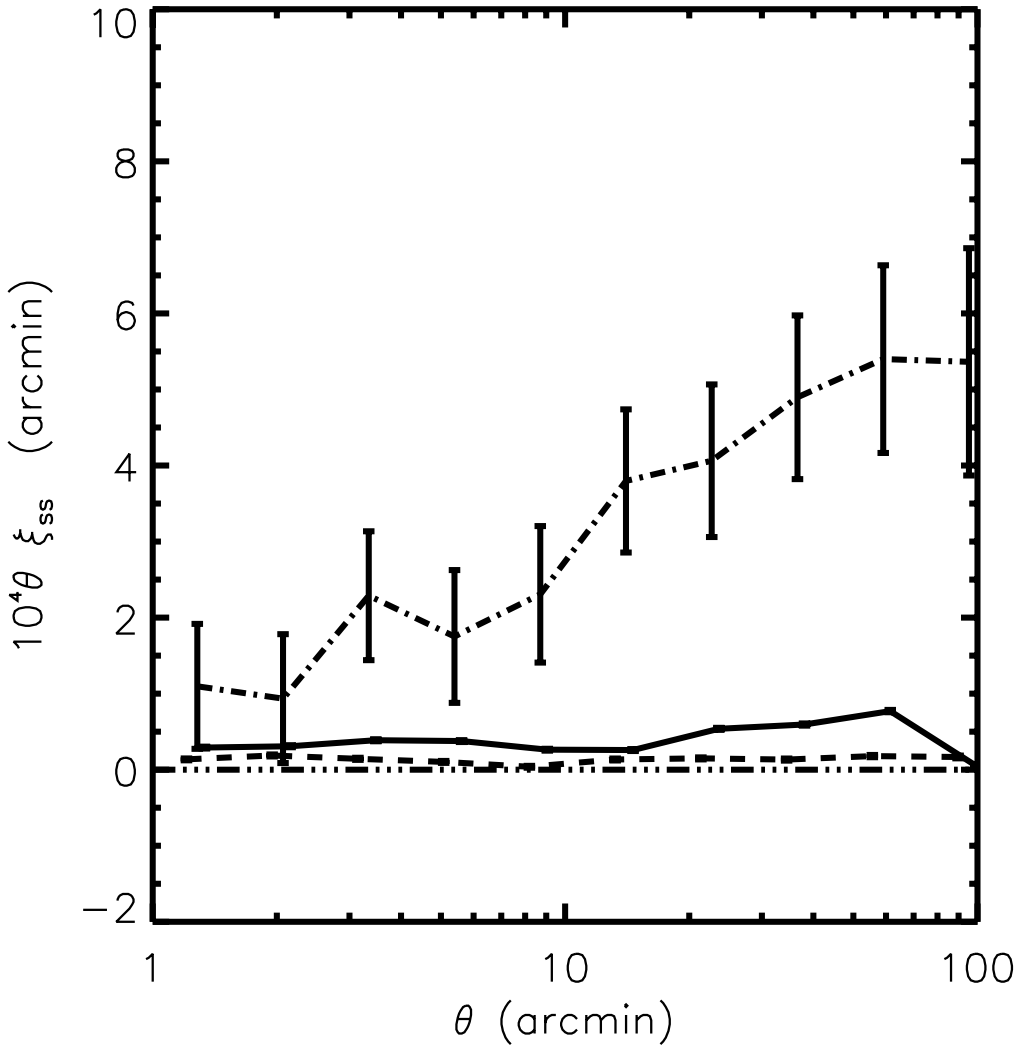

Figure 6.13 The autocorrelation of r-band star shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

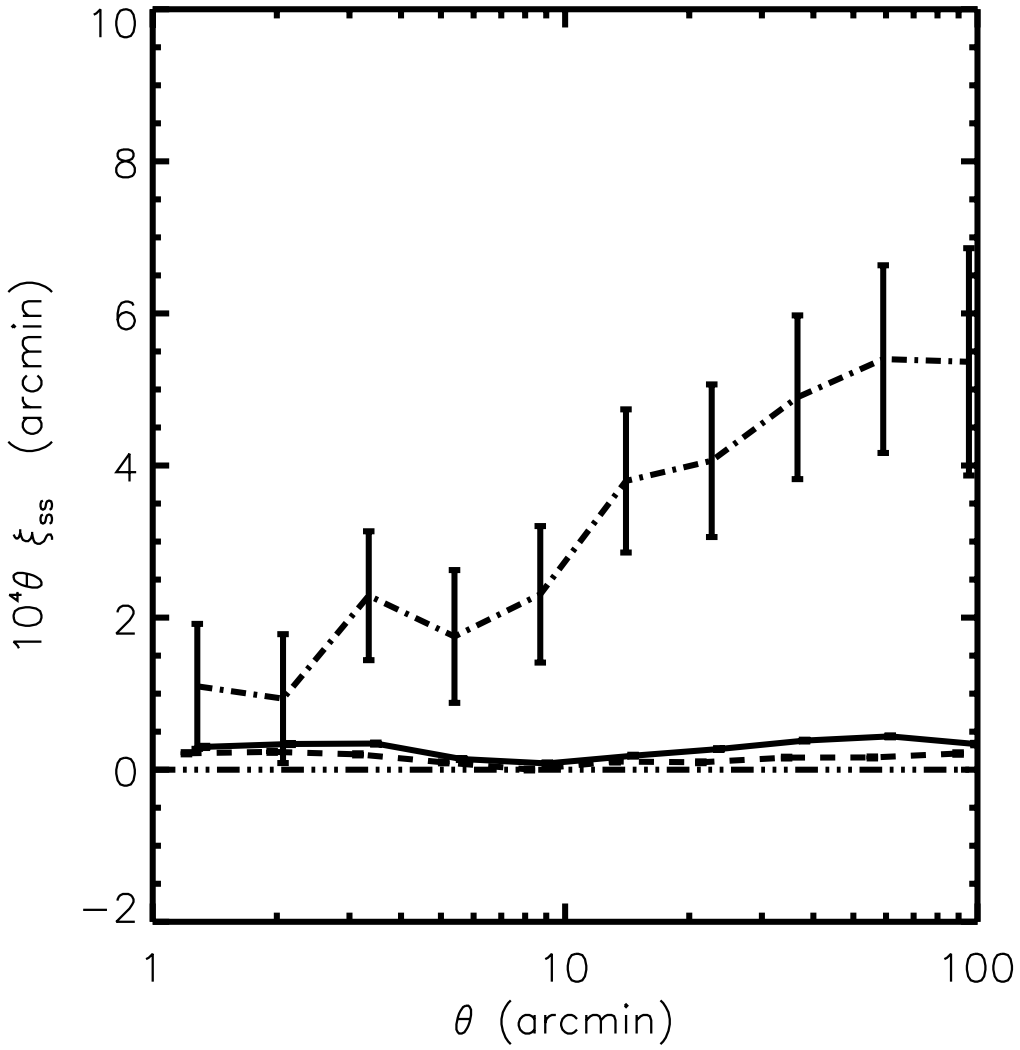

Figure 6.14 The autocorrelation of i-band star shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

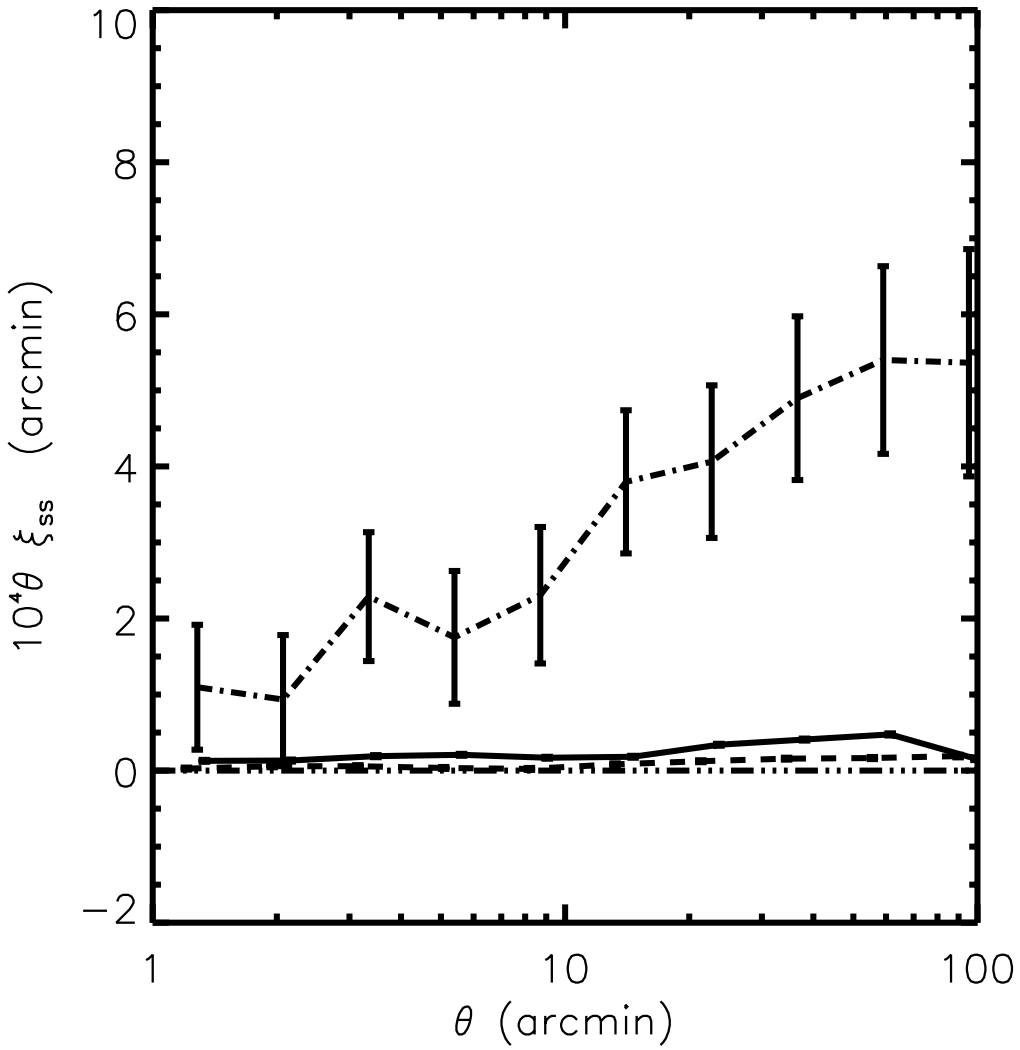

Figure 6.15 The cross-correlation of i-band star shapes with i-band star shapes. The $\langle e_1\, e_1\rangle$ correlation is the solid line, while the $\langle e_2\, e_2\rangle$ correlation is the dashed line. The dot-dashed line shows the expected cosmic shear $\langle e_+\, e_+\rangle$ shape-shape correlation for a survey of this depth and size, with shot-noise errors.

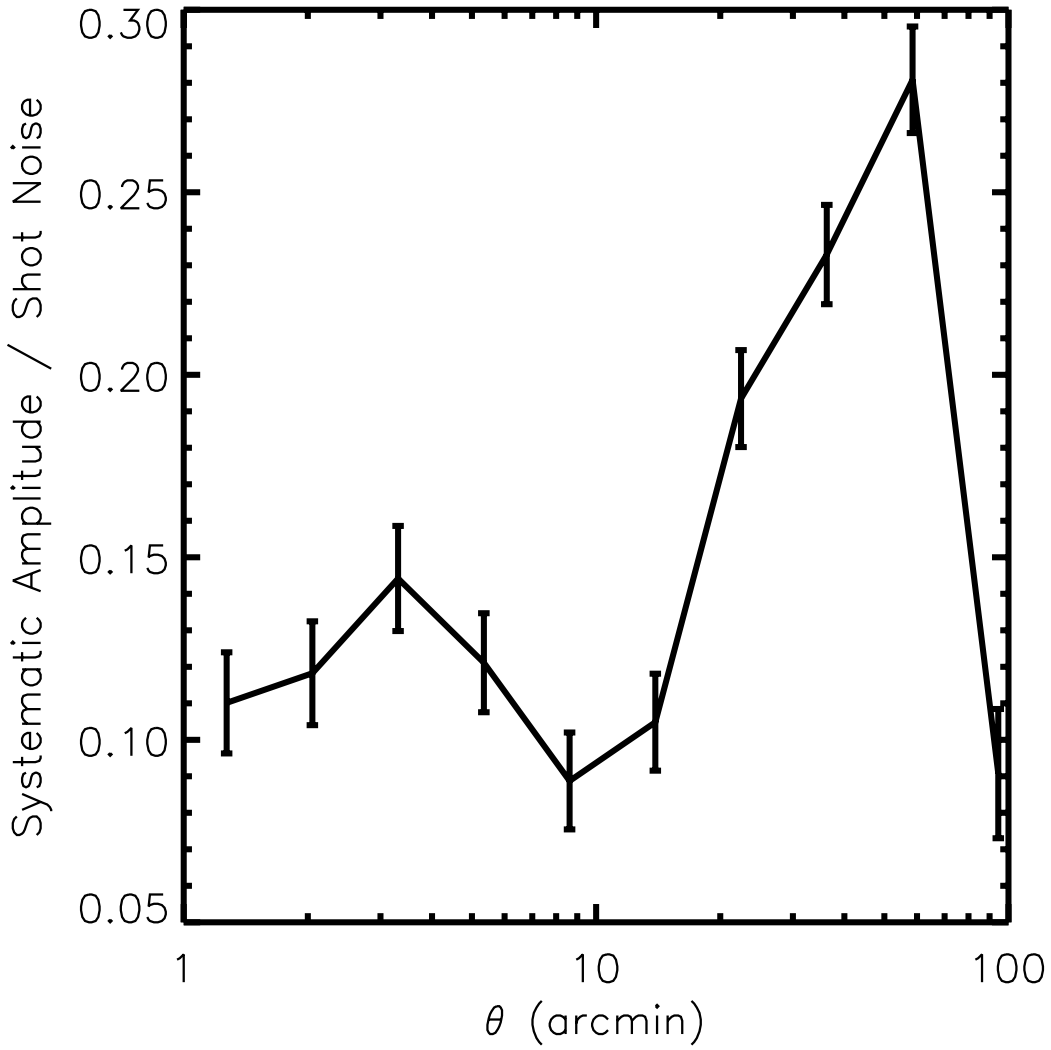

Figure 6.16 The mean stellar cross-correlation signal as a fraction of the expected Poisson error.

Chapter 7

The Model

We model the observed galaxy shape field as the sum of a cosmic shear component, an independent systematics field produced by anisotropies in the effective PSF $e_{\rm psf}$, and a systematics field produced by the intrinsic spatial correlations of galaxy shapes $e_{\rm int}$ (intrinsic alignments; e.g., Hirata & Seljak 2004). For this work, we follow (Bernstein & Jarvis, 2002) and define shapes as 'distortions', which are related to the axis ratio q of an ellipse as:

$$|e| = \frac{1-q^2}{1+q^2} \tag{7.1}$$

and to the adaptive second moment matrix of a surface brightness profile $I(\mathbf{x})$ as:

$$\begin{aligned} e_1 &= \frac{M_{xx} - M_{yy}}{M_{xx} + M_{yy}} \\ e_2 &= \frac{2M_{xy}}{M_{xx} + M_{yy}} \end{aligned} \tag{7.2}$$

where the adaptive moments themselves are:

$$M_{x_i,x_j} = \int_\infty d^2\mathbf{x}\, x_i x_j w(\mathbf{x})\, I(\mathbf{x}) \tag{7.3}$$

and w is an elliptical Gaussian weight function that has been matched in shape to the galaxy light profile.

We allow for a shear calibration factor that depends on the shear responsivity $\mathcal{R}$ (Bernstein & Jarvis, 2002) of the ensemble of galaxy surface brightness profiles to the underlying gravitationally-induced shear γ. We consider $\mathcal{R}$ to be a general factor that includes the standard response (see below) as well as any biases due to effects such as uncorrected PSF dilution, noise-related biases, or selection biases. We assume that the galaxy shape response to PSF anisotropies $\mathcal{R}_{\rm psf}$ is not a priori known, but rather suffers from a similar set of 'calibration' uncertainties as the response of the ensemble of galaxy images to gravitational lensing shear. Thus we define our model for the two ellipticity components $\mathbf{e} = (e_1, e_2)$ as

$$\mathbf{e} = \mathcal{R}\gamma + \mathcal{R}_{\rm psf}\mathbf{e}_{\rm psf} + \mathbf{e}_{\rm int}. \tag{7.4}$$

We assume that the two-point statistics of the underlying (cosmological) shear field $\langle\gamma\gamma\rangle$ consist entirely of E-modes, $e_{\gamma,E}$ (which is a good enough approximation given the size of our errors; Crittenden et al. 2002; Schneider et al. 2002b), and are statistically independent of the PSF when averaged over large regions. We also assume that the PSF and the intrinsic alignments are independent – but not that the lensing shear and intrinsic alignments are independent (Hirata & Seljak, 2004). The two-point correlation of the galaxy shapes contains terms resulting from gravitational lensing and from systematic errors:

$$\langle \mathbf{e}e\rangle = \mathcal{R}^2\xi_{\gamma,E} + \mathcal{R}^2_{\rm psf}\xi_{\rm psf} + \xi_{\rm int} + \langle \mathbf{\Gamma}\mathbf{e}_{\rm int}\rangle. \tag{7.5}$$

Here, $\xi_{\rm psf}$ is the auto-correlation of the PSF ellipticity field. Errors in the determination of the galaxy redshift distribution will enter as a bias in the predicted $\xi_{\gamma,E}$.

Our goal is to carry out a statistics-limited measurement of $\xi_{\gamma,E}$. This will entail showing that the combined amplitudes of $\mathcal{R}^2_{\rm psf}\xi_{\rm psf}$, $\xi_{\rm int}$, $\langle \mathbf{\Gamma}\,\mathbf{e}_{\rm int}\rangle$, the uncertainty in the theoretically-predicted $\xi_{\gamma,E}$ arising from redshift errors, and the uncertainty in the shear calibration (via the responsivity $\mathcal{R}$) contribute less than 20 per cent to the statistical errors in $\langle \mathbf{e}e\rangle$.

Our approach to handling of systematic error is as follows: we attempt to reduce each systematic to a term that can be robustly and believably estimated from real data (either the data here or in other, related work), and we then explicitly correct for it. These corrections naturally have some uncertainty associated with them, which we use to derive a systematic error component to the covariance matrix. The exception to the rule given here is if there is a systematic error for which there is no clear path to estimating its magnitude, then we do not attempt any correction, and simply marginalize over it by include an associated uncertainty in the covariance matrix.

7.1 Cosmic shear

Foreground anisotropies in the matter distribution along the line of sight to a galaxy will generically distort the galaxy image. For weak lensing, the leading order lensing contribution to galaxy shapes can be thought of as arising from a linear transformation of the image coordinates $A\mathbf{x}_{\rm true} = \mathbf{x}_{\rm obs}$, where

$$A = \begin{pmatrix} 1+\kappa+\gamma_1 & \gamma_2 \\ \gamma_2 & 1+\kappa-\gamma_1 \end{pmatrix}. \tag{7.6}$$

The convergence κ causes magnification, whereas the shear components γ_1 and γ_2 map circles to ellipses. The shear is related to the projected line-of-sight matter distribution, weighted by the lensing efficiency:

$$(\gamma_1,\gamma_2) = \partial^{-2}\int_0^\infty W(\chi,\chi_i)\left(\partial_x^2-\partial_y^2, 2\partial_x\partial_y\right)\delta(\chi\hat{\mathbf{n}}_i)\,{\rm d}\chi. \tag{7.7}$$

Here we integrate along the comoving line-of-sight distance χ (where χ_i is the distance to the source), and the matter over-density $\delta = (\rho-\overline{\rho})/\rho$. The window function in a flat universe

is

$$W(\chi, \chi_i) = \frac{3}{2}\Omega_m H_0^2 (1+z)\chi^2 \left(\frac{1}{\chi} - \frac{1}{\chi_i}\right). \tag{7.8}$$

The two-point correlation function of the shear can be calculated by identifying pairs of source galaxies, and defining shear components (γ_t, γ_x) for each one to be the shear in the coordinate system defined by the vector connecting them, and in the $\pi/4$ rotated system. This two-point correlation function can be expressed as a linear transformation of the matter power spectrum P_δ averaged over the line of sight to the sheared galaxies:

$$\begin{aligned} \xi_\pm &= \langle \gamma_t \gamma_t \rangle \pm \langle \gamma_\times \gamma_\times \rangle \\ &= \frac{1}{2\pi} \int_0^\infty \mathrm{d}\ell\, \ell\, P_\kappa(\ell)\, J_{0,4}(\ell\theta) \end{aligned} \tag{7.9}$$

and

$$\begin{aligned} P_\kappa &= \left(\frac{3\Omega_m}{2d_H^2}\right) \int_0^\infty \frac{\mathrm{d}\chi}{a(\chi)^2} P_\delta\left(\frac{\ell}{d(\chi)}\right) \\ &\quad \times \left[\int_\chi^\infty \mathrm{d}\chi' n(\chi') \frac{d(\chi'-\chi)}{d(\chi')}\right]^2, \end{aligned} \tag{7.10}$$

where the last expression makes use of Limber's approximation and $d(\chi)$ is the distance function, i.e. χ in a flat universe, $K^{-1/2}\sin K^{1/2}\chi$ in a closed universe, and $(-K)^{-1/2}\sinh(-K)^{1/2}\chi$ in an open universe. In the expression in brackets, $n(\chi')$ represents the source distribution as a function of line-of-sight distance (normalised to integrate to 1). This statistic (P_κ) is sensitive both to the distribution of matter δ and to the background cosmology, via both the explicit Ω_m dependence and the distance-redshift relations.

7.2 Intrinsic alignments

Many studies have discussed intrinsic alignments of galaxy shapes due to effects such as angular momentum alignments or tidal torque due to the large-scale density field (for pioneering studies, see Croft & Metzler, 2000; Heavens et al., 2000; Catelan et al., 2001; Crittenden et al., 2001; Jing, 2002; Hopkins et al., 2005). While these effects can generate coherent intrinsic alignment 2-point functions, Hirata & Seljak (2004) pointed out that the large-scale tidal fields that can cause intrinsic alignments are sourced by the same large-scale structure that is responsible for producing a cosmic shear signal. Thus, in this model, the intrinsic alignments do not just have a nonzero auto-correlation, they also have a significant anti-correlation with the lensing shear which can persist to very large transverse scales and line-of-sight separations. If left uncorrected, this coherent alignment of intrinsic galaxy shapes suppresses the lensing signal, since the response of the intrinsic shape to an applied tidal field has the opposite sign from the response of the galaxy image to a shear with the same magnitude and direction. We generally refer to the intrinsic alignment auto-correlation as the

"II" contamination and its correlation with gravitational lensing as the "GI" contamination. This can be compared to the pure gravitational lensing auto-correlation ("GG").

To address the uncertainty related to intrinsic alignments, we rely on empirical measurements that constrain the degree to which they might affect our measurement. Several studies using SDSS imaging and spectroscopic data (e.g., Mandelbaum et al., 2006b; Hirata et al., 2007; Okumura et al., 2009; Joachimi et al., 2011; Mandelbaum et al., 2011) have demonstrated the existence of intrinsic alignments of galaxy shapes on cosmological distance scales. Hirata et al. (2007) used the luminosity and colour-dependence of intrinsic alignments for several SDSS galaxy samples to estimate the contamination of the cosmic shear signal due intrinsic alignments for lensing surveys as a function of their depth. These estimates were a function of the assumptions that were made, for example about evolution with redshift. The "central" model given in that paper leads to a fractional contamination of

$$\frac{C_{\ell=500,GI}}{C_{\ell=500,GG}} \approx -0.08 \tag{7.11}$$

for a limiting magnitude of $m_{R,\mathrm{lim}} = 23.5$, which is close to the limiting magnitude of our sample. Subsequent work (Joachimi et al., 2011; Mandelbaum et al., 2011) provided more information about redshift evolution; primarily those results were in broad agreement with the previous ones, and were sufficient to rule out both the "optimistic" and the "pessimistic" models in Hirata et al. (2007).

We thus adopt the "central" model, and apply the correction given in Eq. (7.11) to our theory predictions for the C_ℓ due to cosmic shear, multiplying the predicted cosmic shear power spectrum by 0.92 before transforming into the statistics that are used for the actual cosmological constraints[1]. We also assume this correction has a conservative systematic uncertainty of 50 per cent, which amounts to an overall 4 per cent uncertainty in the theory prediction (see Sec. 8.2 for a quantitative description of how we incorporate this and other systematic uncertainties into the covariance matrix).

Since the GI correlation is first order in the intrinsic alignment amplitude, while the II power is second order, we expect the first to be the dominant systematic. In principle, the GI effect could be smaller than II if the correct alignment model is quadratic in the tidal field rather than linear (Hirata & Seljak, 2004). However, in the aforementioned cases in which intrinsic alignment signals are detected at high significance (i.e. for bright ellipticals) the linear model for intrinsic alignments appears to be valid (Blazek et al., 2011). Therefore we attempt no correction for II.

[1]While the intrinsic alignment contamination is in principle scale-dependent, the plots in Hirata et al. (2007) suggest that this scale dependence is in fact quite weak for the scales used for our analysis, so we ignore it here.

7.3 Shear calibration

Another source of systematic error for weak lensing measurements is uncertainty in the shear calibration factor. The galaxy ellipticity $(e_+, e_\times)$ observed after isotropizing the PSF need not have unit response to shear: in general, averaged over a population of sheared galaxies, we should have

$$\langle (e_+, e_\times) \rangle = \mathcal{R}(\gamma_+, \gamma_\times), \tag{7.12}$$

where $\mathcal{R}$ is the shear responsivity. It depends on both the shape measurement method and the galaxy population (e.g. Massey et al., 2007a; Bernstein, 2010; Zhang & Komatsu, 2011).

For this work, we used the re-Gaussianization method (Hirata & Seljak, 2003), which is based on second moments from fits to elliptical Gaussians, and has been previously applied to SDSS single-epoch imaging (Mandelbaum et al., 2005; Reyes et al., 2011a). For this class of methods, in the absence of selection biases and weighting of the galaxies, perfectly homologous isophotes, and no noise, there is an analytic expectation (Bernstein & Jarvis, 2002):

$$\mathcal{R} = 2(1 - e_{\rm rms}^2), \tag{7.13}$$

where $e_{\rm rms}$ is the root-mean-square ellipticity per component (+ or ×).

The calibration errors for re-Gaussianization and other adaptive-weighting methods are well-studied in the literature (e.g., Hirata et al., 2004b; Mandelbaum et al., 2005, 2012; Bergé et al., 2012). They arise from all of the deviations from the assumptions of Eq. (7.13). Higher-order[2] departures from non-Gaussianity in the galaxy light profile cause errors in the PSF dilution correction. Errors in the measurement of the PSF model will cause a similar error in the dilution correction. The resolution factor of an individual galaxy depends on its ellipticity, so any resolution cut on the galaxy sample will introduce a shear bias in the galaxy selection function. Due to the non-linearity of the shear inference procedure, noise in the galaxy images causes a bias in the shears (rather than just making them noisier). The estimation of the shear responsivity, or even of $e_{\rm rms}$, is another potential source of error, as the response of the galaxies to the shear depends on the true, intrinsic shapes, rather than the gravitationally sheared, smeared (by the PSF), noisy ones that we observe.

Past approaches to this problem have included detailed accounting for these effects one by one. In this paper, we instead use detailed simulations of the image processing and shape measurement pipelines, including real galaxy images, to estimate both the shear calibration and the redshift distribution of our catalogue. The advantage is that this includes all of the above effects and avoids uncertainties associated with analytic estimates of errors. The SHERA (SHEar Reconvolution Analysis) simulation package[3] has been previously described (Mandelbaum et al., 2012) and applied to single-epoch SDSS data for galaxy-galaxy lensing (Reyes et al., 2011a), but this is its first application to cosmic shear data.

[2]Nonzero higher-order terms in the elliptical Gauss-Laguerre expansion of the galaxy light profile; see (Hirata & Seljak, 2003) for details.‘

[3]`http://www.astro.princeton.edu/~rmandelb/shera/shera.html`

To simulate our images, we require a fair, flux-limited sample of any galaxies that could plausibly be resolved in our coadd imaging, including high-resolution images with realistic morphologies[4]. For this purpose we use a sample of 56 662 galaxy images drawn from the COSMOlogical Evolution Survey (COSMOS: Koekemoer et al., 2007; Scoville et al., 2007b,a) imaging catalogues. The deep *Hubble Space Telescope* (HST) Advanced Camera for Surveys/Wide Field Camera (ACS/WFC) imaging in $F814W$ ("broad I") in this 1.6 deg^2 field is an ideal source of a fairly-selected galaxy sample with high resolution, deep images[5]. These images consist of two samples – a "bright" sample of 26 116 galaxies in the magnitude range $I < 22.5$, and a "faint" sample consisting of the $22.5 < I < 23.5$ galaxies. The charge transfer inefficiency-corrected (Massey et al., 2010) and multi-drizzled (Koekemoer et al., 2003; Rhodes et al., 2007, to a pixel scale of $0.03''$) galaxy postage stamp images have been selected to avoid CCD edges and diffraction spikes from bright stars, and have been cleaned of any other nearby galaxies, so they contain only single galaxy images without image defects. The bright sample is used for ground-based image simulations in Mandelbaum et al. (2012); the faint sample is selected and processed in an identical way[6]. Each postage stamp is assigned a weight to account for the relative likelihoods of generating postage stamps passing all cuts (avoidance of CCD edges and bright stars) for galaxies of different sizes in the COSMOS field; this weight is calculated empirically, by comparing the size distribution of galaxies with postage stamps to the size distribution of a purely flux-limited sample of galaxies.

Each of these postage-stamp images has several properties associated with it that are of interest for this analysis. The COSMOS photometric catalogues (Ilbert et al., 2009) contain HST $F814W$ magnitudes as well as photometric redshifts and Subaru $r-i$ colours based on PSF-matched aperture magnitudes.

In order to simulate our observations, we first select a coadd 'run' consisting of five adjacent frames in the scan direction at random from the list of completed runs. We draw 1250 galaxies (exactly 250 per frame) at random from the list of COSMOS postage stamps according to the weights described above, up-weighting the probability of drawing the faint galaxies by a factor of 1.106 to account for the fact that we have sampled the faint population at a lower rate than the bright one in constructing the image sample.

Once a list of postage-stamp images is selected, we assign r- and i-band magnitudes by re-scaling each image; each galaxy image is inserted into the coadded imaging with the flux it would have been observed to have in SDSS before the addition of pixel noise. The i-band is chosen to be 0.03 magnitudes brighter than the COSMOS $F814W$ (I) band `MAG_AUTO` values; this small offset is based on empirical comparison with SDSS magnitudes for brighter galaxies, to account for slight differences in the $F814W$ and i passbands (Mandelbaum et al., 2012). The r-band is chosen so as to match the Subaru PSF-matched aperture colours for

[4]Simple models with analytic radial profiles and elliptical isophotes are not adequate to measure all sources of systematic error such as under-fitting biases or those due to non-elliptical isophotes (Bernstein, 2010).

[5]Admittedly there may be some sampling variance that affects the morphological galaxy mix.

[6]We thank Alexie Leauthaud for kindly providing these processed images.

each object. Each postage stamp is assigned a random, uniformly-sampled position in the coadd run, with the postage stamps distributed equally among the frames.

We use the SHERA code to pseudo-deconvolve the HST point-spread function, apply (if necessary; see below) a shear to each galaxy, reconvolve each image with the known coadd point-spread function, renormalise the flux appropriately, and resample from the COSMOS pixel scale to the coadd pixel scale before adding that postage stamp to the coadd image. This procedure, suggested by Kaiser (2000) and implemented to high precision in Mandelbaum et al. (2012), can be used to simulate ground-based images with a shear appropriately applied, despite the space-based PSF in the original COSMOS images, and with a user-defined PSF.

The normal coadd masking algorithm is then applied, and shear catalogues are generated by running the SDSS object detection and measurement pipeline, PHOTO-FRAMES, followed by the shape measurement code described in chapter 5. The output catalogues are matched against the known input object positions, and a simulation catalogue of the matches is created. We employ these simulations below to determine the shear calibration and as an independent validation of our inferred redshift distribution.

For each suite of simulation realisations, we use the same random seed (i.e., we select the same galaxies from our catalogue and place them at identical locations in the coadded image) but with different applied shears per component ranging from -0.05 to $+0.05$. We measure the mean weighted shape of the detected simulation galaxies produced by our pipeline, and fit a line to the results. Since the same galaxies are used without rotation, only the slope and not the intercept is meaningful. The shear response in each component for each applied shear is shown in Fig. 7.1. The responsivities in the two components are consistent, which is expected on oversampled data with a rounded PSF. (The unequal size of the error bars reflects the number of runs that we were able to process by the time the shear calibration solution was frozen.) The total number of galaxies in the final simulated catalogues was 130 063. The response appears to be linear for small applied shears. Based on these results, we adopt a shear responsivity for this galaxy population of 1.776 ± 0.043. For the galaxy population used in this measurement, the shape dispersion $e_{\rm rms}$ is 0.37; the corresponding responsivity for an unbiased shape measurement method, by Eq. (7.13), is 1.72. Even in the absence of any correction from the simulations above, this measurement would only suffer a 2.8% shear calibration bias, which is already an unusually small bias given that it includes many realistic effects such as selection bias, noise rectification bias, and effects due to realistic galaxy morphologies. This bias is well below the statistical errors of our measurement, but we correct for it in any case by using the simulation-based responsivity rather than the "ideal" one based on the RMS ellipticity.

7.4 Redshift distribution

The explicit dependence of the shear signal in Eqs. (7.7) and (7.10) on the distribution of lensed galaxy redshifts, combined with the practical impossibility of acquiring a spectroscopic redshift for the millions of faint galaxies statistically necessary for a cosmic shear

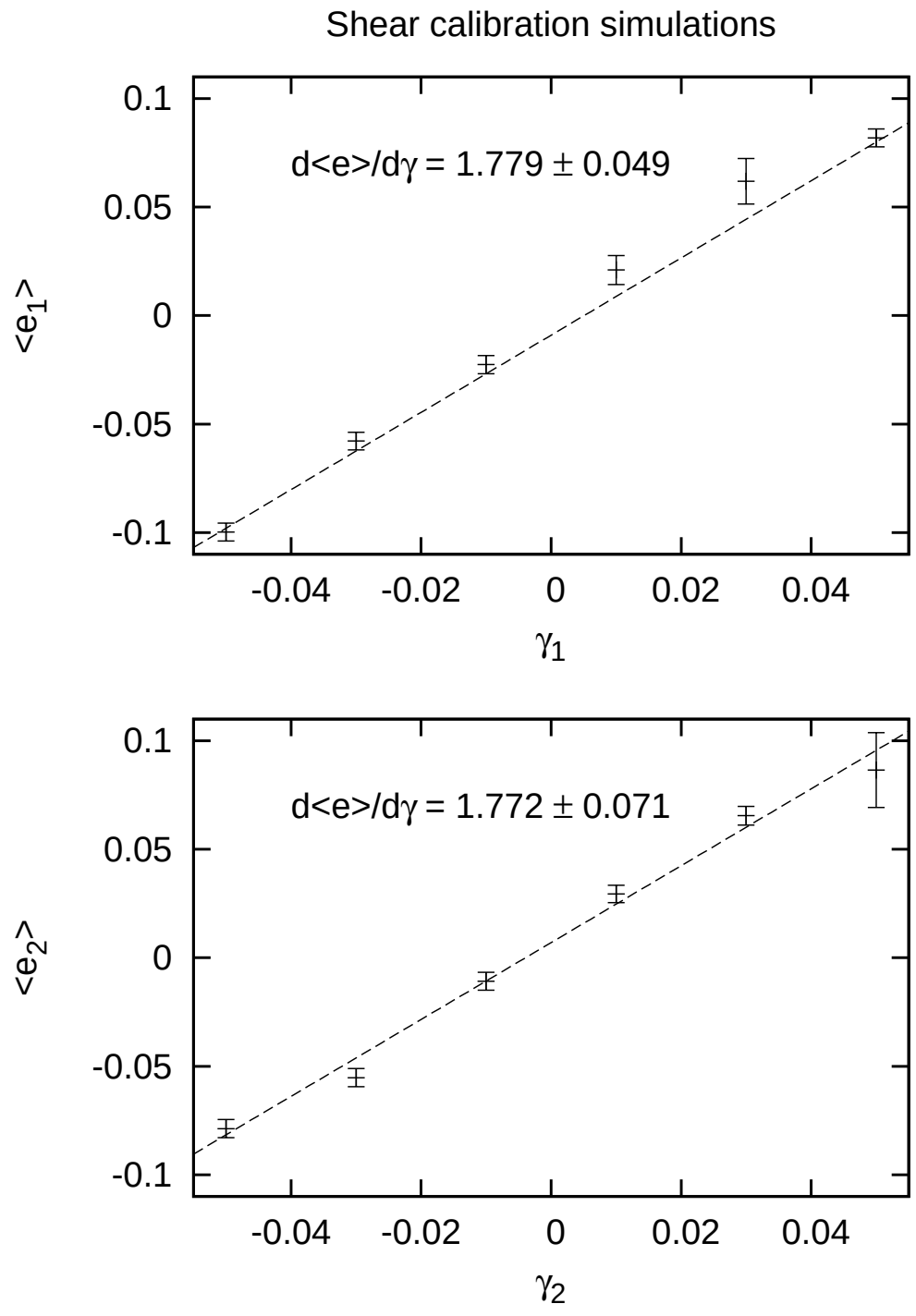

Figure 7.1 The response of the mean ellipticities $\langle e_1 \rangle$ and $\langle e_2 \rangle$ to applied shear, as determined in the SHERA-based simulations. Poisson error bars are shown. The additive offset to the response curve is not shown in the fit; these simulations do not accurately measure an additive shear bias.

measurement, can be a troublesome source of bias and systematic uncertainty for cosmic shear measurements.

An error in the estimated redshift distribution leads to an incorrect prediction for the amplitude of the shear signal at a given cosmology. This is similar in principle to the bias arising in the amplitude of the galaxy-galaxy lensing signal due to photometric redshift biases explored in Nakajima et al. (2012); uncorrected, standard photometric redshift estimation techniques can lead to biases in the predicted lensing signal at the ~ 10 per cent level. For cosmic shear measurements, an imperfect estimate of the redshift distribution leads to biases in σ_8 and Ω_m that are comparable in amplitude to the errors in the estimated mean of the redshift distribution (van Waerbeke et al., 2006).

As a fiducial reference, the redshift distribution of the single-epoch SDSS imaging catalogue is established to approximately 1 per cent (Sheldon et al., 2011); for deeper surveys over a smaller area, this becomes a more difficult problem, as the spectroscopic calibration samples available for inferring the redshift distribution are limited in their redshift coverage and widely dispersed across the sky. We employ a colour-matching technique similar to that employed by Sheldon et al. (2011); in what follows, we describe the technique, our estimate of its uncertainty, and several cross-checks on the results.

7.4.1 Fiducial redshift distribution

The source redshift distribution used in our analysis is derived following Lima et al. (2008) and Cunha et al. (2009), and is similar in spirit to Sheldon et al. (2011); the principle is that, for two galaxy samples that span broadly similar ranges in redshift, colour, and limiting magnitude, matched colour samples correspond to matched redshift distributions.

Our spectroscopic calibration sample is composed of 12 360 galaxies, from the union of the VIMOS VLT Deep Survey (Le Fèvre et al., 2005, VVDS) 22h field, the DEEP2 Galaxy Redshift Survey (Davis et al., 2003; Madgwick et al., 2003), and portions of the PRism MUlti-object Survey (PRIMUS; Coil et al. 2011, Cool et al. 2011 in prep.). We follow the procedures outlined in Nakajima et al. (2012) for selecting good quality spectroscopic redshifts, and avoiding duplicate galaxies in samples that overlap (such as DEEP2 and PRIMUS). Each of these samples has a redshift distribution that is likely to differ substantially from the redshift distribution of our lensing catalogue: the DEEP2 catalogue in the fields we use at 23^h30^m and 02^h30^m is heavily colour-selected (in non-SDSS bands) towards objects at $z > 0.7$; the PRIMUS catalogue includes several fields, some of which are selected from imaging with a shallower limiting magnitude; and the VVDS catalogue is selected in the I band ($I < 22.5$) with a relatively high redshift failure rate that exhibits some colour dependence.

We assign a redshift from a galaxy in the union calibration sample to the closest galaxy in the lensing catalogue within 3 arcsec, finding 12 360 matches. To generate a representative training sample of galaxies from the lens catalogue, we draw 4×10^5 galaxies with replacement from the full area (not just in these regions), with sampling probability proportional to the mean of the weights assigned in the r and i bands to that galaxy for the correlation analysis (Eq. 8.1). Note that this procedure does not incorporate those galaxies in the excluded

camcol 2 region.

We use the Lima et al. (2008) code[7] to solve for a set of weights over the calibration sample, such that the re-weighted 5-dimensional magnitude distributions of the calibration sample match those of the representative random subset of the lensing catalogue.

All photometric redshift estimation methods assume (at least implicitly) that two galaxy populations with similar distributions in color and magnitude have similar distributions in redshift – otherwise. If that is the case, and if the spectroscopic sample spans the full range of properties of the photometric sample, then the photometric distribution $n_p(z)$ can be written as a product of the true spectroscopic redshift distribution and a redshift-dependent function:

$$n_p(z) = ns(z)w(z). \tag{7.14}$$

The algorithm attempts to find a weight $w(z_i)$ for the i^{th} galaxy such that the re-weighted spectroscopic calibration sample has the same properties as a fair sample of the true redshift distribution of the photometric sample. It uses a nearest-neighbor method to define volume elements in observable space such that for any given volume element, the galaxies in that element can be assigned a weight $w(z) = np(z)/ns(z)$ without the ratio introducing unmanageable amounts of noise.

Because the COSMOS tests described below agree perfectly (within statistical errors, of course) with the redshift histogram, major biases are extremely unlikely – such biases would require a significant population of galaxies at z¡1 for which no spectroscopic redshifts in PRIMUS, VVDS, or DEEP2 are successful, and which are also invisible to any checks on the COSMOS photo-z's. While not impossible, the existence of such a population in this sample seems very unlikely.

The histogram of the calibration sample redshifts reweighted in this manner is shown as a solid line in Fig. 7.2. The inferred mean redshift is 0.51; in contrast to the redshift distribution for single-epoch imaging, there is a non-negligible fraction of the galaxy sample above $z > 0.7$. We use the solid curve based on the colour-matching techniques to calculate the shear covariance matrix, and to predict the shear correlation function for any given cosmology.

7.4.2 Uncertainty

We expect that the primary source of error in the redshift distribution as estimated from the combined calibration sample is sample variance, resulting from the finite volume of the calibration sample. To estimate its magnitude, we use the public code of Moster et al. (2011) for estimating the cosmic variance of number counts in small fields.

Our redshift binning scheme has 19 bins between $0 < z < 1.5$. For a collection of disparate calibration fields, we use the Moster et al. (2011) code to produce a fractional error in the number counts $\sigma_{\text{gg},i,j}$ for the jth redshift bin in the i field (where fields are distinguished by their coverage area) in bins of stellar mass.

[7] `http://kobayashi.physics.lsa.umich.edu/~ccunha/nearest/`

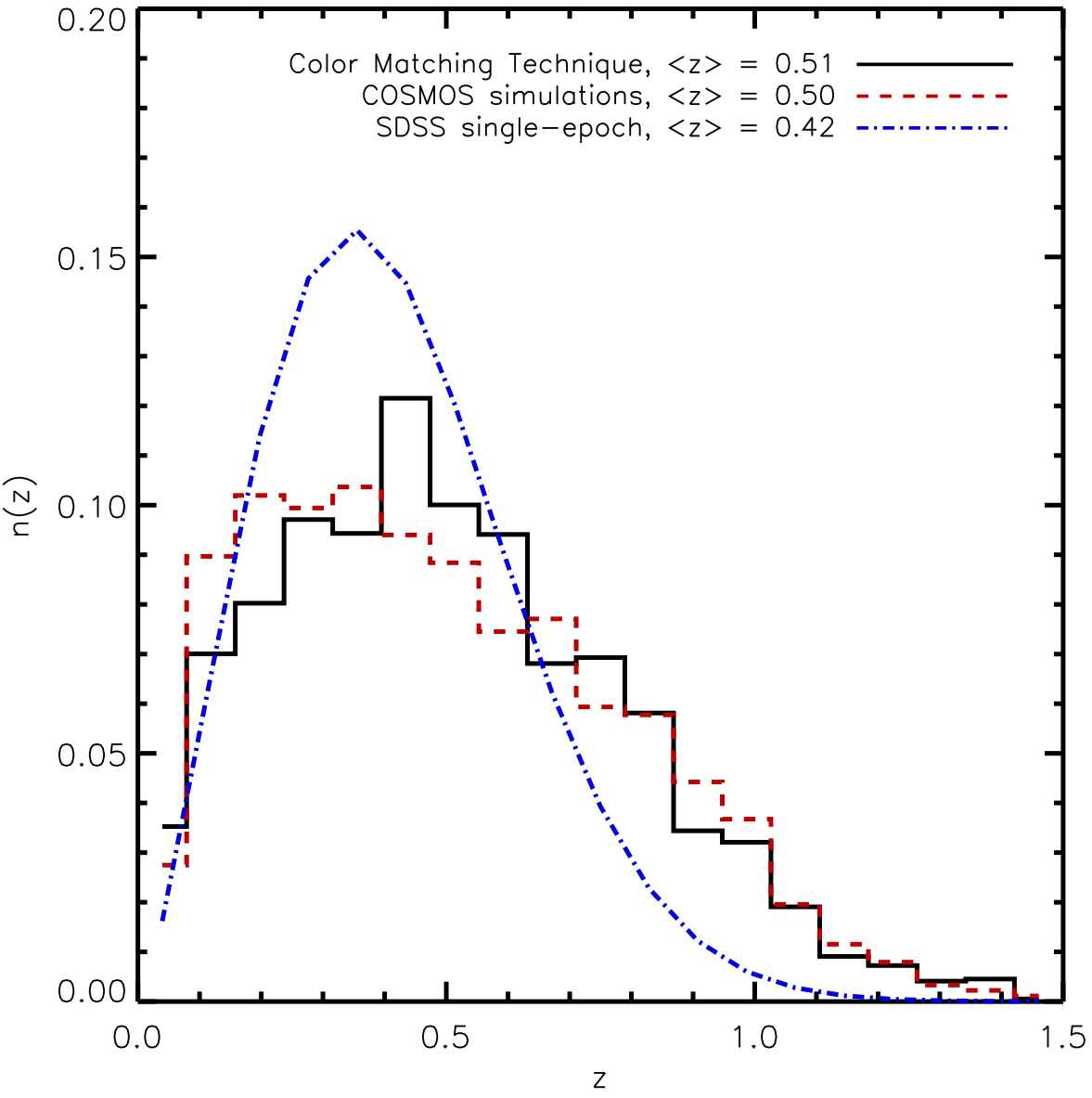

Figure 7.2 The redshift distribution inferred from matching the colours of the spectroscopic calibration sample to those of the lensing catalogue (solid black line, Sec. 7.4.1) shown alongside the noisier redshift distribution inferred from the shear calibration simulations (dashed red line, Sec. 7.4.3). The best-fit distribution for the single-epoch SDSS lensing catalogue from Nakajima et al. (2012) is shown for reference as the blue dot-dashed line.

The redshift sampling rate of each distinct survey in the calibration sample differs, and so the balance of contributions to the final redshift distribution will change as well. To account for this, we sum over every calibration field's contribution to the reweighted redshift distribution in the j bin to estimate an absolute (not relative) overall error:

$$\sigma_j^2 = \sum_i \left(\sigma_{\mathrm{gg},i,j} n_{\mathrm{eff},i,j}\right)^2 \tag{7.15}$$

where the effective number of galaxies contributed in the j bin by the i survey is just the sum over the nearest-neighbour derived weights assigned to calibration sample galaxies k in that field i and bin j:

$$n_{\mathrm{eff},i,j} = \sum_k w_{nn,i,j,k} \tag{7.16}$$

To propagate these errors into the covariance matrix for ξ_E, we first fit a smooth function of the form

$$n_z\left(z\right) \propto z^a \mathrm{e}^{-(z/z_0)^b} \tag{7.17}$$

to the nearest neighbour weighting-derived redshift distribution shown in Figure 7.2; the best fit parameters are $a = 0.5548$, $z_0 = 0.7456$, and $b = 2.5374$. We perturb this smooth distribution by adding a random number drawn from a normal distribution with mean $n_z\left(z_j\right)$ (normalised to the weighted number of calibration galaxies in that bin) and standard deviation σ_j at the location of the jth redshift bin. We then renormalise the perturbed distribution to unity, and compute the predicted cosmic shear signal. The covariance matrix of 402 realisations of this procedure is added to the statistical covariance matrix.

7.4.3 Other tests

As an independent check on the redshift distribution, we also use the shear calibration simulations (Sec. 7.3) to constrain the redshift distribution of our sources. The COSMOS photometric redshifts, inferred as they are from many more imaging bands (typically with deeper imaging) than for the SDSS data discussed here, are very accurate. For example, Ilbert et al. (2009) finds a photo-z scatter of $\sigma_z/(1+z) \sim 0.01$ for a galaxy sample with the flux limit of the SDSS coadds. In contrast, Nakajima et al. (2012) found that in the SDSS single-epoch imaging, the scatter defined in the same way was ~ 0.1 despite the brighter flux limit of the single-epoch imaging (due in part to the more limited number of bands, but primarily to the far lower signal-to-noise ratio). If we treat the COSMOS photometric redshifts as we would spectroscopic data, then the redshift distribution of COSMOS galaxies that pass successfully into the shear catalogue is the same as that of our source catalogue – assuming, of course, that the COSMOS field is representative of the whole of Stripe 82. It is not, of course; large-scale structure in the COSMOS field (which can be significant, as COSMOS covers only 1.7 square degrees; Kovač et al., 2010) can bias a determination of the redshift distribution in this manner. The $n(z)$ inferred from the COSMOS-based simulations

is also shown in Fig. 7.2, and agrees extremely well with the fiducial $n(z)$ derived from colour-matching.

A final (but obviously not independent) sanity check is to compare to the COSMOS Mock Catalogue (Jouvel et al., 2009), which is being used extensively to plan future dark energy programs, using the cuts $r_{\rm eff} > 0.47$ arcsec, limiting magnitudes $r < 23.5$, and $i < 22.5$. This test predicts $\langle z \rangle = 0.51$, identical to that obtained via the re-weighting procedure. Given the crudeness of the procedure for comparing the results, this is an excellent validation of the COSMOS Mock Catalogue as a forecasting tool.

7.5 Stellar contamination

Stellar contamination of the galaxy catalogue reduces the apparent shear by diluting the signal with round objects that are not sheared by gravitational lensing. Because the image simulations described in Sec. 7.3 only included galaxies, the resulting shear responsivities do not include signal dilution due to accidental inclusion of stars in the galaxy sample. In Chapter 6, we estimated the stellar contamination by comparison with the DEEP2 target selection photometry (which is deeper and was acquired at the Canada-France-Hawaii Telescope under much better seeing conditions than typical for SDSS), and found a contamination fraction of 0.017. We also argued that the mean stellar density in the stripe must be larger than in the high-latitude DEEP2 fields, by a factor as large as 2.8. We therefore conservatively take the stellar contamination fraction $f_{\rm star}$ to be

$$f_{\rm star} = 0.017(1.9 \pm 0.9) = 0.032 \pm 0.015. \tag{7.18}$$

The resulting suppression of the cosmic shear signal is treated in much the same way as for intrinsic alignments: we reduce the theory signal by a factor of $(1 - 0.032)^2 = 0.936$, and add a contribution to the covariance of 0.030 times the theory signal.

7.6 Additive systematics

7.6.1 Shear Selection Bias from Masking

Among the most worrying systematics in the early detections of cosmic shear was additive power. This comes from any non-cosmological source of fluctuations in shapes such as PSF anisotropy that add to the ellipticity correlation function of the galaxies. Such power is clearly detected in the form of systematic variation of both star and galaxy e_1 as a function of declination. The sense of the effect – a negative contribution to e_1 (in r band we have[8] $\langle e_1 \rangle = -0.0018$ and $\langle e_2 \rangle = +0.0004$, while in i band $\langle e_1 \rangle = -0.0022$ and $\langle e_2 \rangle = -0.0002$) – is suggestive of masking bias, in which the selection of a galaxy depends on its orientation, with galaxies aligned in the along-scan direction ($e_1 < 0$) being favoured, and with no effect

[8]The 1σ Poisson uncertainty in these numbers is 0.0005 (0.0004) per component in r (i) band.

on e_2 (consistent with zero mean over the whole survey). In this appendix, we argue that the selection bias due to pixel masking described above is a quantitatively plausible explanation for the remaining spurious shear correlations. This calculation is motivated by the observation that the spurious star-galaxy correlations increase in amplitude when more galaxies are rejected from the catalog using the pixel mask.

Consider an elliptical galaxy isophote centered at the origin, with an ellipticity e and a major axis at some arbitrary position angle θ with respect to the x-axis. The ellipticity e is defined in terms of the axis ratio q such that:

$$\begin{aligned} e &= \frac{1-q^2}{1+q^2} \\ q &= \frac{b}{a},\ b \leq a \end{aligned} \tag{7.19}$$

where b and a are the semi-minor and -major axes of the ellipse, respectively.

Define the two ellipticity components e_+ and $e_\times$ as:

$$\begin{aligned} e_1 &= e\cos(2\theta) \\ e_2 &= e\sin(2\theta) \end{aligned} \tag{7.20}$$

Naturally, the average ellipticity over all position angles θ is zero.

Now place a vertical barrier at position $x = d, d > 0$ and compute the expected value of e over all θ again, this time removing the contribution from all position angles where the ellipse crosses the barrier.

Because of the $\theta \mapsto \theta + \pi$ symmetry of ellipticity, the average value of e_2 will be unchanged. When the ellipse is far enough from the barrier such that $d > a$, the expectation value will be zero; when the minor axis of the ellipse meets the barrier, $d < b$, isophotes at any position angle will be rejected from the catalog.

The extremal point on the ellipse in the direction of the barrier is:

$$x_{max} = \sqrt{\frac{T}{2}(1+e_1)} \tag{7.21}$$

where $T = a^2 + b^2$, and the galaxy is masked if $x_{max} \geq d$. To compute the mean shape over the survey geometry, average over ellipticity weighting by the total survey area where it is possible to measure each ellipticity:

$$\langle e_1 \rangle = \frac{\int_{-1}^{1} de_1\, A(e_1)\, e_1}{\int_{-1}^{1} de_1\, A(e_1)}. \tag{7.22}$$

and $A(e_1)$, the total survey area where it is possible to measure an ellipticity e_1, is:

$$A(e_1) = (A_{\rm tot} - x_{\rm max} P) \tag{7.23}$$

where $A_{\rm tot}$ is the total unmasked survey area and P is the total length of the mask perimeter. For $|e| << 1$, the mean e_1 evaluates to:

$$< e_1 > \simeq -\frac{P}{A_{\rm tot}}\sqrt{\frac{T}{2}}\left\langle\frac{e_1^2}{2}\right\rangle \tag{7.24}$$

The prefactor $\frac{P}{A_{\rm tot}}\sqrt{\frac{T}{2}}$ is the fraction of the total survey area that lies nearer than the characteristic radius of a galaxy to a mask boundary.

For the SDSS Stripe 82 data presented here, the variance in ellipticity-per-component $< e_1^2 >$ is 0.1 and the characteristic limiting galaxy isophotal radius [9] is twice the median half-light radius, or 2.5″. Roughly one percent of the survey area lies within this distance of a mask boundary. The shape measurement procedure described in section 5.3 will amplify the masking selection bias by a factor of $\frac{1}{R_2}$, a characteristic value of which is $\frac{1}{R_2} = 1.8$. This yields a masking bias of order-of-magnitude 10^{-3}, which is potentially large enough to dominate the systematic error budget for this measurement.

It is important to note that, while a square or round masking pattern will generally eliminate a bias in the *average* catalog ellipticity, this selection bias produces a coherent shape pattern along the boundary of a masked region of any size or shape. The exact effect on the shear statistics can be calculated by convolving the mask with the amplitude of the mask selection bias in each shape component, and calculating the two-point ellipticity correlation function (or other shear statistic of interest) of the resulting map; for this measurement, however, the bias is small enough that it does not contribute significantly to the shear correlation function, so we content ourselves with the crude treatment (projection) described below.

The reason for the prominent masking effect on the mean statistics can be seen in Figure 5.7; as shown, bad columns along the scan direction tend to be repeated at the same location in multiple images, resulting in significant (non-isotropic) masks with that directionality. Direct evidence for masking bias comes from the change in mean ellipticity due to increased masking: when we removed from the coadded image pixels that were observed in fewer than 7 input runs and reran PHOTO-FRAMES, the $\langle e_1\rangle$ signal became worse: -0.0051 in r band and -0.0044 in i band, whereas $\langle e_2\rangle$ was essentially unchanged. This increase is difficult to explain in terms of spurious PSF effects. Further, the large difference between the amplitude of the star-star and star-galaxy cross-correlation signals is consistent with a masking effect; the stars are nearly round, so $\left\langle\frac{e_1^2}{2}\right\rangle$ is much smaller than for the galaxies. We conclude that our galaxy catalogue likely contains a mixture of masking bias as well as possible additive systematics from PSF ellipticity in the coadded image, though it is very likely in light of the above that the dominant additive systematic signal is due to the masking.

The mean e_1 signal as a function of declination is shown in Fig. 7.3 in bins of width 0.05 degrees. We take this as a template for mask-related selection biases (combined with

[9] For a 10σ circular gaussian galaxy light profile, the outermost detectable isophote is a factor of 1.78 larger than the half-light radius.

any systematic uncorrected PSF variation as a function of declination, which in west-to-east drift-scan observations is a highly plausible type of position dependence). Before computing the correlation function, we subtracted this mean signal from the galaxy ellipticity catalogue.[10]

One danger in this procedure to remove spurious $\langle e_1 \rangle$ is that some real power could be removed – that is, even in the absence of any systematic error, some of the actual galaxy shape correlation function signal could be suppressed since the method determines the mean e_1 of the real galaxies and by subtracting it introduces a slight artificial anti-correlation. The best way to guard against this is with simulations. Using the Monte Carlo simulation tool of Sec. 8.2.1, we generated simulated realisations of our ellipticity catalogue and either implemented the $\langle e_1 \rangle$ projection or not. The difference in the correlation functions is a measure of how much power was removed. The result is shown in Fig. 7.4, and shows that the loss of real power is insignificant compared to our error bars.

7.6.2 PSF anisotropy

Convolution with an elliptical PSF will induce a spurious ellipticity in observed galaxy surface-brightness profiles. While the effective PSF for these coadds is a circular double Gaussian to quite high precision, the tests in Chapter 6 indicate a low level of residual anisotropy that we must consider here.

Possible sources of this issue include: (i) inaccuracies in the single-epoch PSF model used to determine the kernel to achieve the desired PSF; (ii) colour-dependence of the PSF that means the single-epoch PSF model from the stars is not exactly the PSF for the galaxies; or (iii) the fact that we determine the rounding kernel on a fixed grid, so that smaller-scale variations in PSF anisotropy might remain uncorrected. All of these must be present at some level, although the last two cannot be the full solution: (ii) does not explain the residual stellar ellipticity[11], and (iii) does not explain why there is structure in the declination direction on the scale of an entire CCD (0.23 degrees).

For a galaxy and a PSF that are both well-approximated by a Gaussian, the PSF-correction given above produces a measured ellipticity of:

$$e^{\rm obs} = \mathcal{R}_{\rm psf} e^{\rm PSF} = \frac{1 - R_2}{R_2} e^{\rm PSF}; \tag{7.25}$$

see e.g. Bernstein & Jarvis (2002). The weighted (by the same weights used for the correlation function; see Eq. 8.1) average of the PSF anisotropy response defined in Eq. (7.25) over the sample of galaxies considered in this work is $\mathcal{R}_{\rm psf} = 0.86$ (r band) or 0.95 (i band); in what follows we take a value of 0.9.

[10] We refer below to this step as projection, as the intent is to map the shape catalog onto a subspace of itself that does not include the spurious masking-induced modes.

[11] We have searched for a $g - i$ dependence in the stellar ellipticities in the coadded image. We only found effects at the ~ 0.002 level, and while they are statistically significant, we have not established whether they correspond to true colour dependence versus e.g. variation of stellar colour distributions along the stripe.

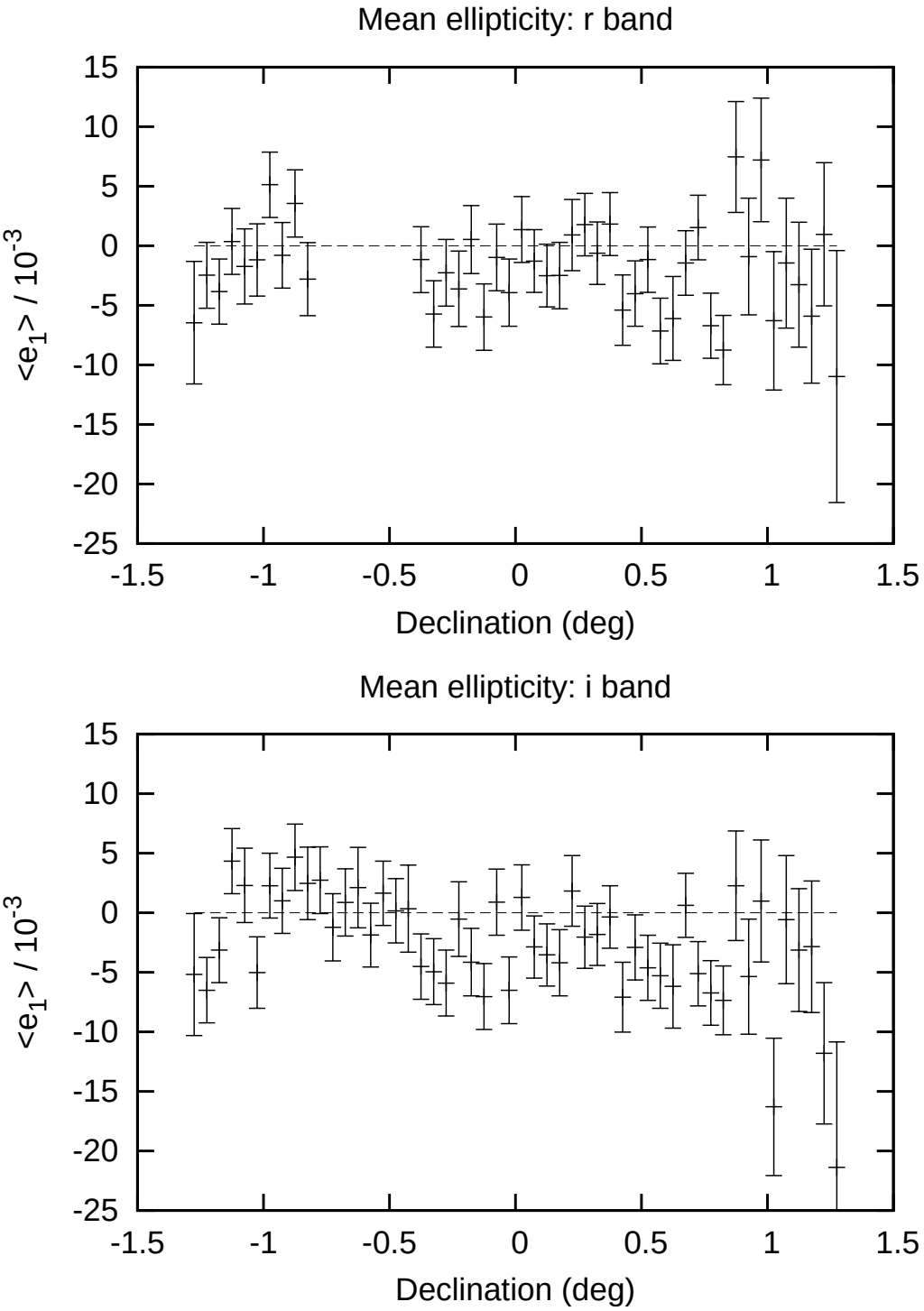

Figure 7.3 The mean ellipticity $\langle e_1 \rangle$ as a function of declination in the r and i bands. This signal was removed from the galaxy catalogue prior to computing the final correlation function. The r band data between declination $-0.8°$ and $-0.4°$ were rejected due to the known problems with camcol 2. The error bars are Poisson errors only.

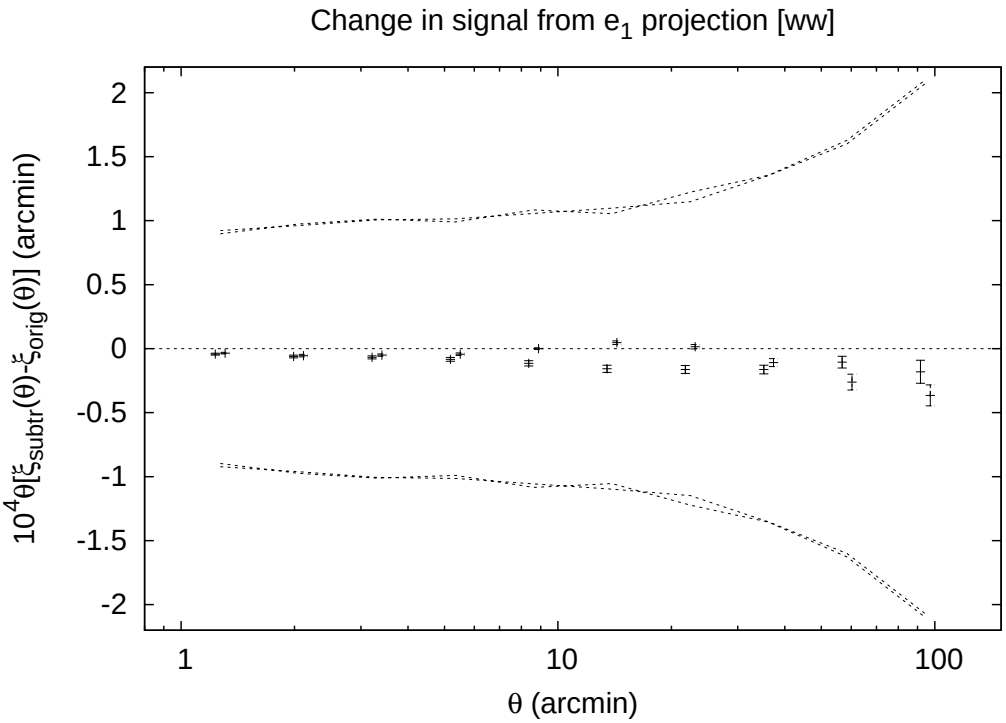

Figure 7.4 The loss of actual power due to e_1 projection. Using 36 realizations from the Monte Carlo simulation, we find the difference in post-projection ellipticity correlation function $\xi(\theta)$ and original $\xi(\theta)$. These are shown as the solid points (ξ_{++}) and dashed points ($\xi_{\times\times}$) in the figure, re-binned to 10 bins in angular separation θ. The dashed lines at top and bottom are the $\pm 1\sigma$ statistical error bars of our measurement. The reduction of actual power is detectable by combining many simulations, but is very small compared to the error bars on the measurement.

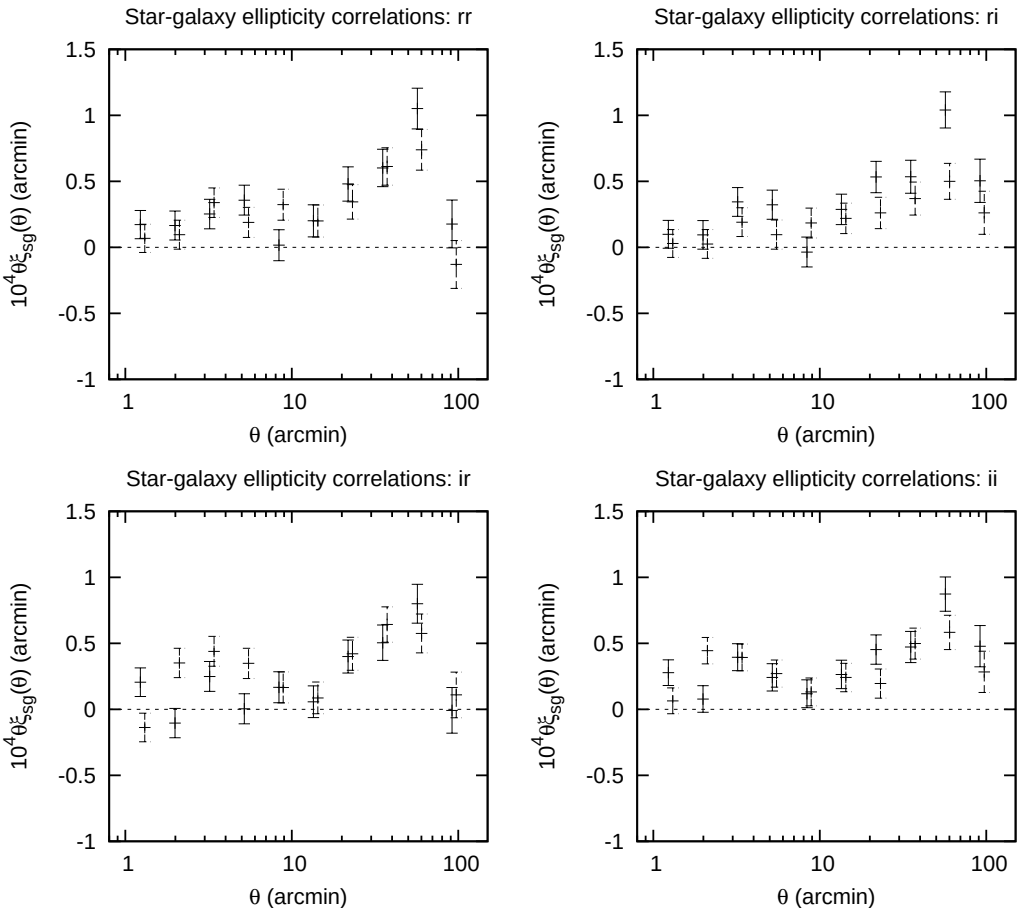

Figure 7.5 The star-galaxy ellipticity correlation functions. Shown are the rr, ri (i.e. star r $\times$ galaxy i), ir, and ii correlation functions, reduced to 10 bins. The solid points, which are offset to slightly lower θ-values for clarity, are the $+\!+$ correlation functions, and the dashed points are the $\times\times$ functions. All error bars are Poisson only.

A nonzero star-galaxy correlation function $\xi_{\rm sg}$ resulting from systematic PSF anisotropy indicates the presence of a spurious contribution to the shear-shear correlation function with amplitude $\approx 0.9\xi_{\rm sg}$. We will not determine this response to high enough accuracy to subtract the effect with small residual error: doing so would not require just a simulation, but a simulation that knows the correct radial profile of the PSF errors.[12] In our case, the star-galaxy correlation function is detectable but below the errors on the galaxy-galaxy ellipticity auto-correlation (although not by very much), so a highly accurate correction is unnecessary.

We constrain the PSF anisotropy contribution by computing the star-galaxy correlation function. This was done above, but some of the star-galaxy signal is due to the systematic variation of PSF ellipticity with declination and is removed by the subtraction procedure above. The star-galaxy ellipticity correlation function with the corrected catalogue is shown in Fig. 7.5. The implied contamination to the galaxy ellipticity correlation function, appropriately averaging the bands and applying the factor of $\mathcal{R}_{\rm psf} = 0.9$, is shown in Fig. 7.6.

These measured star-galaxy correlations can be used to construct a reasonable systematics covariance matrix for this systematic. We take the amplitude of the diagonal elements of the PSF systematic covariance to be equal to the amplitude of the measured contamination. We also assume that the off-diagonal terms are fully-correlated between bins, which is equivalent to fixing the scaling of this systematic with radius, and saying that only the overall amplitude of the systematic is uncertain.

Since there are a number of uncertainties in this procedure, we do not apply any correction for these additive PSF systematics as we do for ones that are previously discussed, such as intrinsic alignments or stellar contamination. Instead, we simply include a term in the systematics covariance matrix to account for it. We also will present a worst-case scenario for the impact of this term on cosmological constraints; in Chap. 9 we will show what happens to the cosmology constraints if we assume that the systematic error is $+2\sigma$ from its mean, i.e. 40 per cent of the statistical errors. This should be taken as a worst-case scenario for this particular systematic.

[12] This might be an option in future space-based surveys if the type of error can be traced to the source of ellipticity (astigmatism×defocus, coma, or jitter). In either space or ground-based data, one could imagine doing cross-correlations of higher-order shapelet modes (Refregier, 2003) to extract the particular form of the errors. None of these options are pursued here.

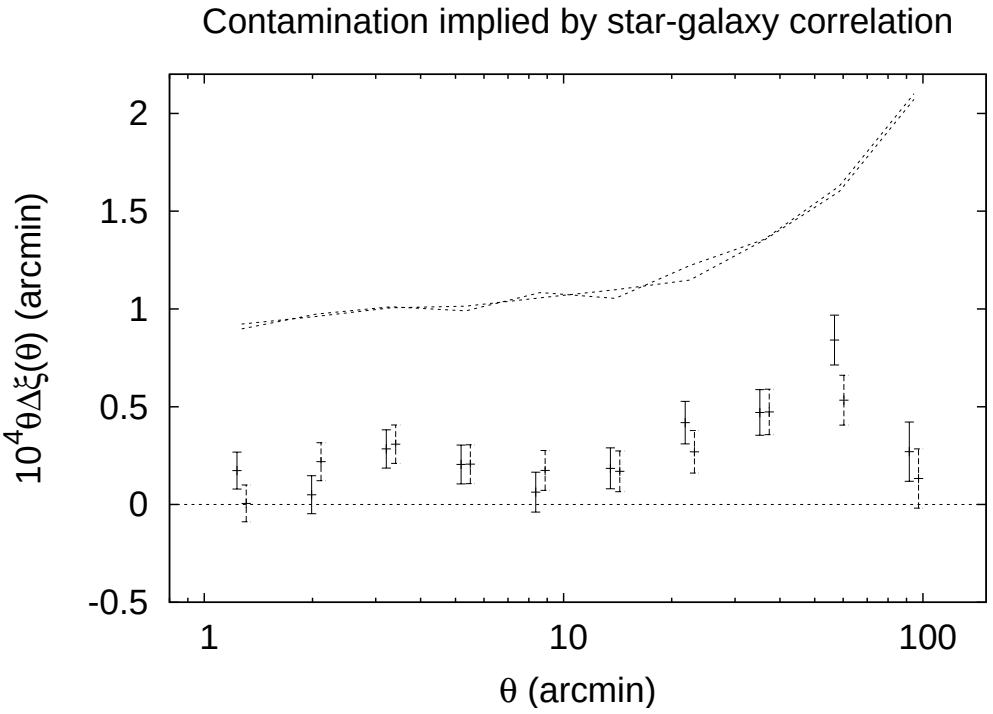

Figure 7.6 The implied contamination to the galaxy ellipticity correlation function if the star-galaxy correlation function is used as a measure of the additive PSF power. The solid points are the ++ correlation functions, and the dashed points are the $\times\times$ functions. All error bars are propagated from the Poisson errors assuming correlation coefficient +1 (a better assumption than independent errors, but likely an overestimate). The dotted curves show the 1σ errors in each radial bin from the Monte Carlo simulations (see Sec. 8.2.1) which include both Poisson and cosmic variance uncertainties. Note also that the shapes and normalisations of the ++ and $\times\times$ signals are nearly identical.

Chapter 8

Analysis Tools

8.1 Analysis tools

8.1.1 Ellipticity correlation function

We compute the ellipticity correlation functions defined in Eq. (7.9) on scales from 1–120 arcminutes. For the cosmological analysis, we start by computing the correlation function in 100 bins logarithmically spaced in separation θ to avoid bin width artifacts. For the cosmological parameter constraints, we project these onto the Complete Orthogonal Sets of E-/B-mode Integrals (COSEBI) basis (Schneider et al., 2010) to avoid the instabilities of inverting a large covariance matrix estimated via Monte Carlo simulations (we will describe our implementation of COSEBIs in Sec. 8.1.3). However, for display purposes, it is more convenient to reduce the θ resolution to only 10 bins so that the real trends are more visually apparent.

Weighting

The correlation functions used here are weighted by the inverse variance of the ellipticities, where the "variance" includes shape noise. Specifically, we define a weight for a galaxy

$$w_i = \frac{1}{\sigma_e^2 + 0.37^2}, \tag{8.1}$$

where σ_e is the ellipticity uncertainty per component defined by our shape measurement pipeline. As demonstrated by Reyes et al. (2011a), these may be significantly underestimated in certain circumstances; however, this will only make our estimator slightly sub-optimal, so we do not attempt to correct for it. The value of 0.37 for the root-mean-square (RMS) intrinsic ellipticity dispersion per component comes from the results of Reyes et al. (2011a), for $r < 22$, and therefore we are implicitly extrapolating it to fainter magnitudes. Given that Leauthaud et al. (2007) found a constant RMS ellipticity to far fainter magnitudes in

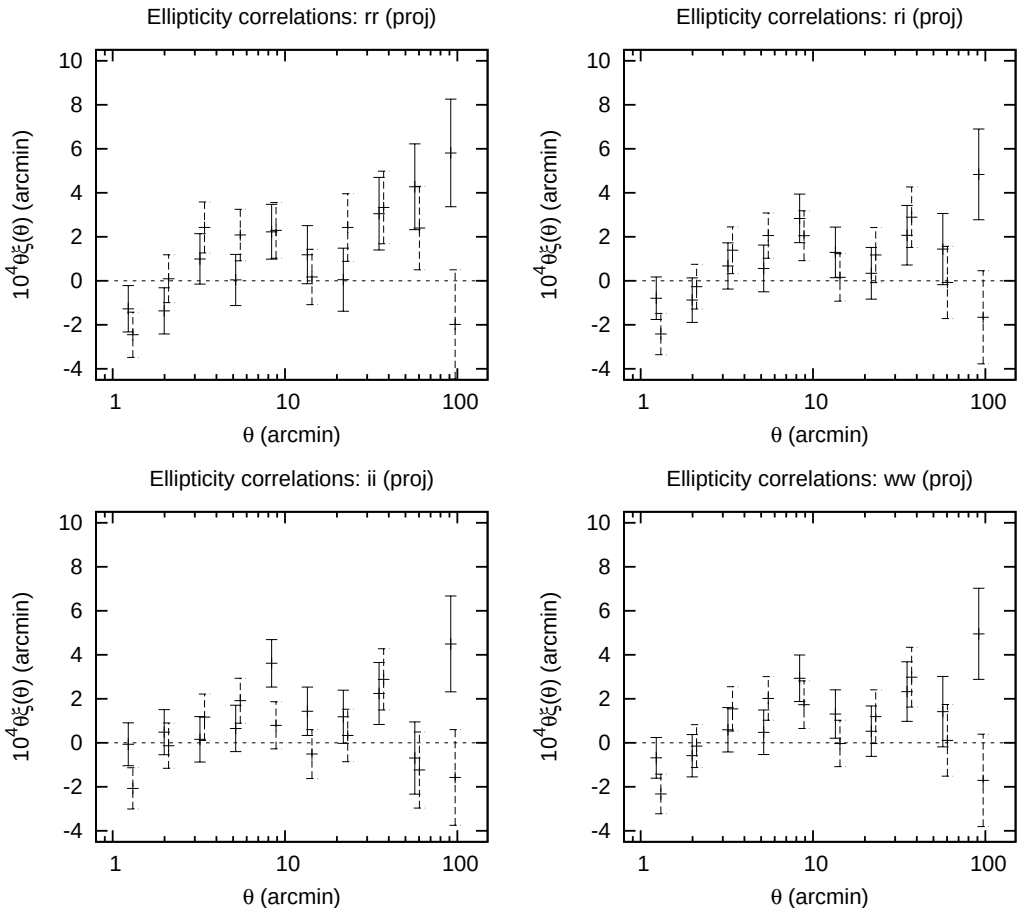

Figure 8.1 The ellipticity correlation functions in the rr, ri, ii and ww (combined) band combinations. The solid points denote the $++$ and the dashed points denote the $\times\times$ components of the correlation function. The points have been slightly displaced horizontally for clarity. The Monte Carlo errors are shown.

the COSMOS data, we consider this extrapolation justified[1].

Direct pair-count code

A direct pair-count correlation function code was used for the cosmological analysis. It is slow ($\sim$ 3 hours for 2×10^6 galaxies on a modern laptop) but robust and well-adapted to the Stripe 82 survey geometry. The code sorts the galaxies in order of increasing right ascension α; the galaxies are assigned to the range $-60° < \alpha < +60°$ to avoid unphysical

[1]Note that we do not use the actual value of RMS ellipticity from Leauthaud et al. (2007) – only the trend with magnitude – because, as demonstrated by Mandelbaum et al. (2012), the RMS ellipticity value in Leauthaud et al. (2007) is not valid for our adaptively-defined moments, which use an elliptical weight function matched to the galaxy light profile.

edge effects near $\alpha = 0$. It then loops over all pairs with $|\alpha_1 - \alpha_2| < \theta_{\rm max}$. The usual ellipticity correlation functions can be computed, e.g.

$$\xi_{++}(\theta) = \frac{\sum_{ij} w_i w_j e_{i+} e_{j+}}{\sum_{ij} w_i w_j}, \tag{8.2}$$

where the sum is over pairs with separations in the relevant θ bin, and the ellipticity components are rotated to the line connecting the galaxies. The direct pair-count code works on a flat sky, i.e. equatorial coordinates (α, δ) are approximated as Cartesian coordinates. This is appropriate in the range considered, $|\delta| < 1.274°$, where the maximum distance distortions are $\frac{1}{2}\delta^2_{\rm max} = 2.5 \times 10^{-4}$. The direct pair-count code is applicable to either auto-correlations of galaxy shapes measured in a single filter (rr, ii) or cross-correlations between filters or between distinct populations of objects (ri and all of the star-galaxy correlations).

Simple post-processing allows one to compute the ξ_+ and ξ_- correlation functions, defined by

$$\xi_+(\theta) \equiv \xi_{++}(\theta) + \xi_{\times\times}(\theta) \tag{8.3}$$

and

$$\xi_-(\theta) \equiv \xi_{++}(\theta) - \xi_{\times\times}(\theta). \tag{8.4}$$

Combining bands

Finally, the different band correlation functions rr, ri, and ii must be combined according to some weighting scheme:

$$\xi^{ww}_{++}(\theta) = w_{rr}\xi^{rr}_{++}(\theta) + w_{ri}\xi^{ri}_{++}(\theta) + w_{ii}\xi^{ii}_{++}(\theta), \tag{8.5}$$

where the label "ww" indicates that the bands were combined. The relative weights were chosen according to the fraction of measured shapes in r- and i-bands, i.e. $w_{rr} = f_r^2$, $w_{ri} = 2f_r f_i$, and $w_{ii} = f_i^2$ where the weights are $f_r = 0.4603$ and $f_i = 0.5397$.

The final ellipticity correlation functions (with the θ resolution reduced to 10 bins) are shown in Fig. 8.1.

8.1.2 Tests of the correlation function

We implement several null tests on the correlation function to search for remaining systematic errors.

The first test, shown in Fig. 8.2, constructs the difference between the cross-correlation function of r and i band galaxy ellipticities versus the rr and ii auto-correlations. The differences in the two types of correlation functions are small compared to the statistical uncertainty in the signal. This is consistent with our expectations, as the true cosmic shear signal should be independent of the filters in which galaxy shapes are measured.

The second test, shown in Fig. 8.3, compares the (band averaged or ww) correlation function computed using galaxy pairs separated in the cross-scan (north-south) direction

versus pairs separated in the along-scan (east-west) direction. This difference should be zero if the signal we measure is due to lensing in a statistically isotropic universe. The error bars shown are Poisson errors, so they may be slight underestimates at the larger scales, where cosmic variance becomes important. Visual inspection shows no obvious offset from zero, but the error bars are larger for this test than in Fig. 8.2 because the null test includes no cancellation of galaxy shape noise.

8.1.3 E/B-mode decomposition

As a final check for systematics, we decompose the 2-point correlation function (2PCF) into E- and B-modes, where, to leading order, gravitational lensing only creates E-modes. The B-modes can arise from the limited validity of the Born approximation (Jain et al., 2000; Hilbert et al., 2009), redshift source clustering (Schneider et al., 2002b), and lensing (magnification) bias (Schmidt et al., 2012; Krause & Hirata, 2010), however the amplitude of B-modes from these sources should be undetectable with our data. At our level of significance, a B-mode detection would indicate remaining systematics, e.g. due to spurious power from an incomplete PSF correction.

Formerly used methods to decompose E-and B-modes, such as the aperture mass dispersion

$$\langle M_{\rm ap}^2\rangle(\theta) = \int_0^{2\theta} \frac{{\rm d}\vartheta\,\vartheta}{2\,\theta^2}\left[\xi_+(\vartheta)T_+\left(\frac{\vartheta}{\theta}\right) \;+\; \xi_-(\vartheta)T_-\left(\frac{\vartheta}{\theta}\right)\right], \tag{8.6}$$

with the filter functions $T_\pm$ as derived in Schneider et al. (2002b), or the shear E-mode correlation function, suffer from E/B-mode mixing (Kilbinger et al., 2006), i.e. B-modes affect the E-mode signal and vice versa. These statistics can be obtained from the measured 2PCF, for an exact E/B-mode decomposition, however they require information on scales outside the interval $[\theta_{\rm min};\theta_{\rm max}]$ for which the 2PCF has been measured.

The ring statistics (Schneider & Kilbinger, 2007; Eifler et al., 2010; Fu & Kilbinger, 2010) and more recently the COSEBIs (Schneider et al., 2010) perform an EB-mode decomposition using a 2PCF measured over a finite angular range. COSEBIs and ring statistics can be expressed as integrals over the 2PCF as

$$EB = \int_{\theta_{\rm min}}^{\theta_{\rm max}} \frac{{\rm d}\theta}{2}\,\theta[T_{+n}^{\log}(\theta)\xi_+(\theta) \pm T_{-n}^{\log}(\theta)\xi_-(\theta)] \tag{8.7}$$

and

$$R_{\rm EB}(\theta) = \int_{\theta_{\rm min}}^{\theta} \frac{{\rm d}\theta'}{2\theta'}[\xi_+(\theta')Z_+(\theta',\theta) \pm \xi_-(\theta')Z_-(\theta',\theta)]. \tag{8.8}$$

For the ring statistics, we use the filter functions $Z_\pm$ specified in Eifler et al. (2010). The derivation of the COSEBI filter functions $T_{\pm n}$ is outlined in Schneider et al. (2010), where the authors provide linear and logarithmic filter functions indicating whether the separation of the roots of the filter function is distributed linearly or logarithmically in θ. Note that whereas the ring statistics are a function of angular scale, the COSEBIs are calculated

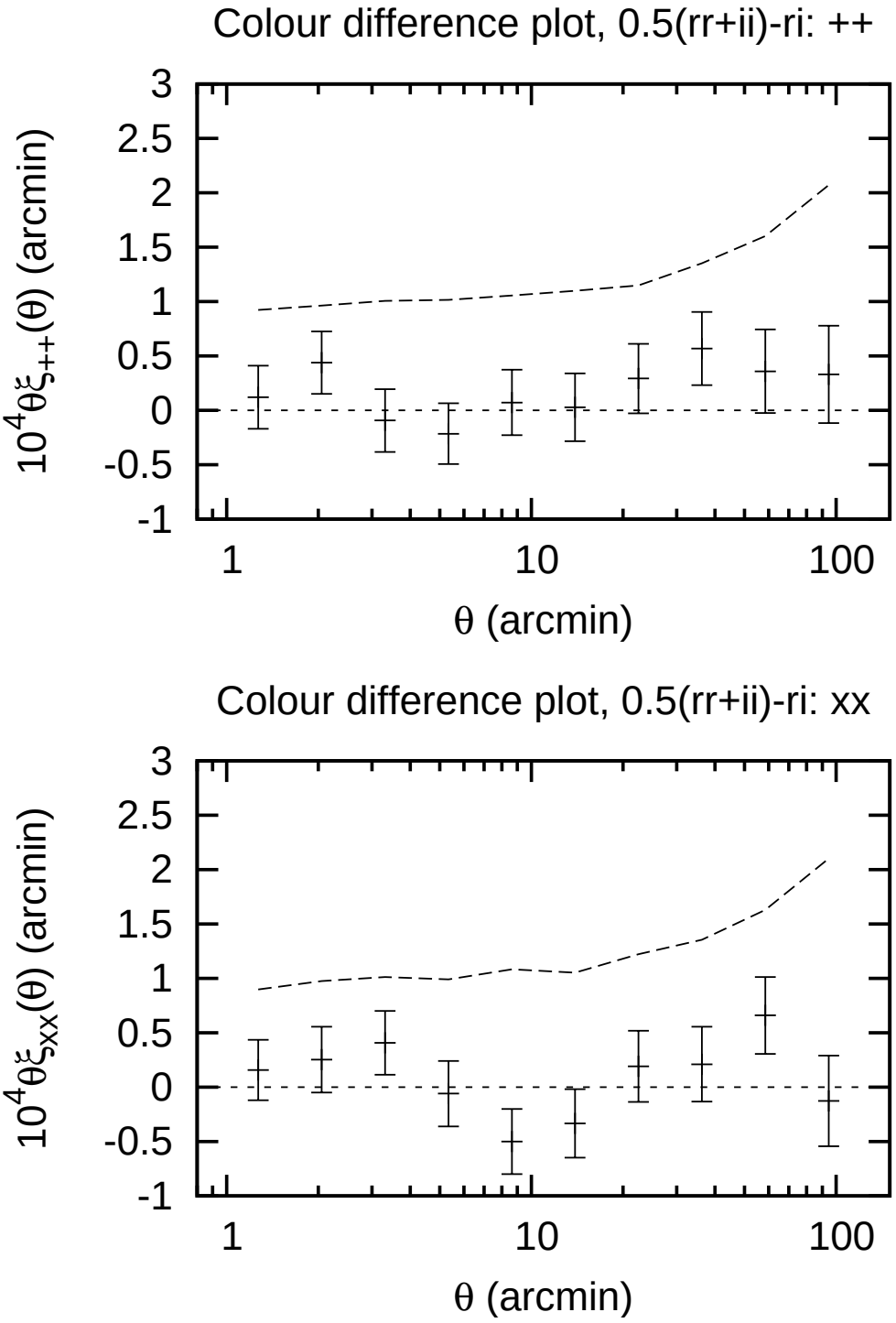

Figure 8.2 The difference between the galaxy ellipticity cross-correlations (ri) and the auto-correlations $(rr + ii)/2$, with error bars determined from the Monte Carlo simulations. The upper panel shows the ++ correlations and the lower panel shows the $\times\times$ correlations. The dashed line is the 1σ statistical error bar on the actual signal.

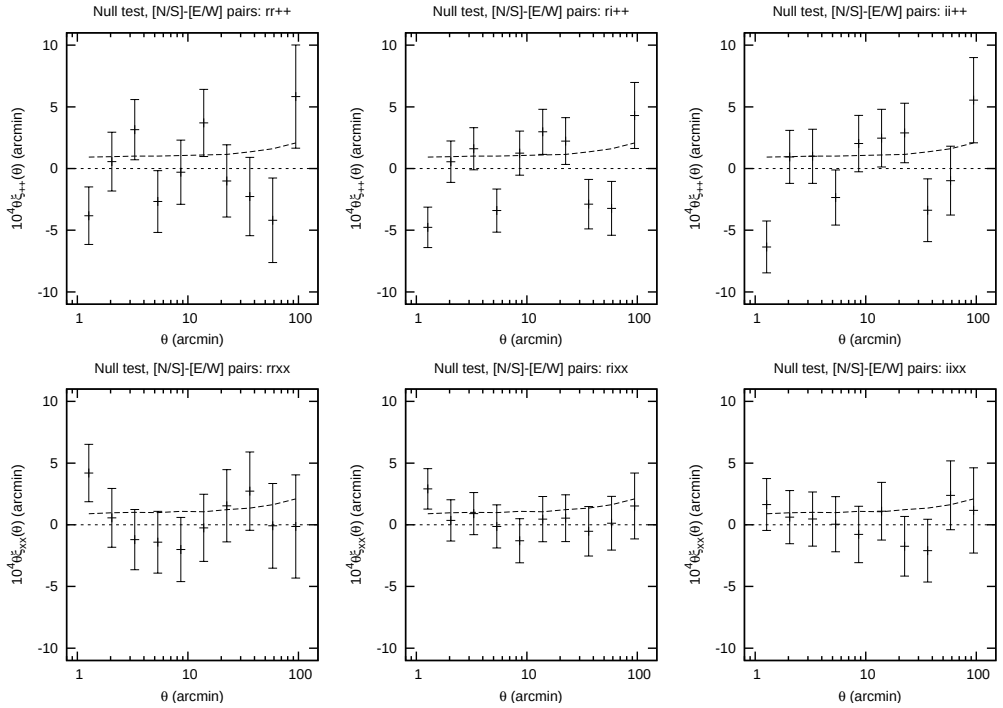

Figure 8.3 The null test of the correlation functions measured using galaxy pairs whose separation vector is within 45° of the north-south direction, minus that measured using galaxy pairs whose separation vector is within 45° of the east-west direction. The error bars shown are the Poisson errors only. The dashed curve shows the 1σ error bars of the actual signal (all colour combinations and separation vectors averaged). The 6 panels show the three colour combinations (rr, ri, and ii) and the 2 components ($++$ or $\times\times$).

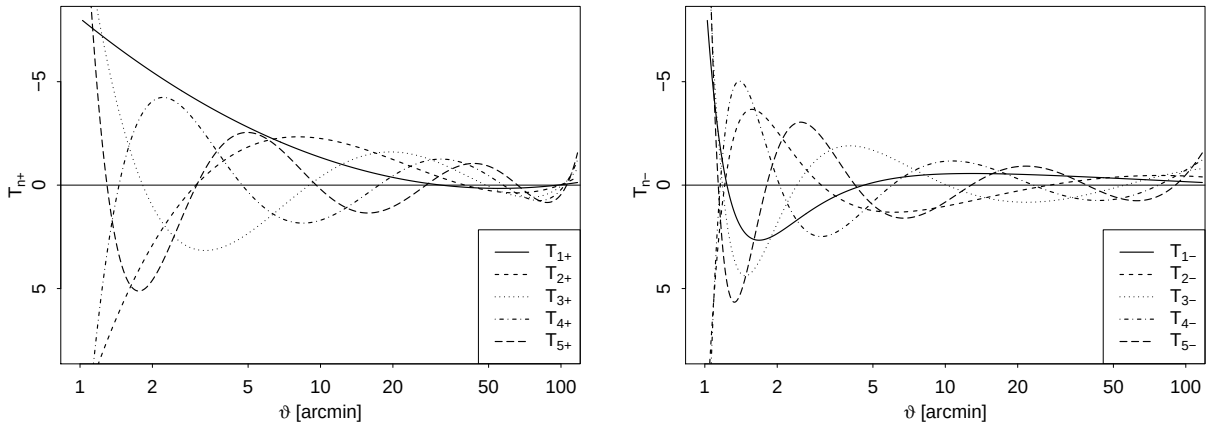

Figure 8.4 The COSEBI filter functions T_{n+} (upper panel) and T_{n-} (lower panel) for the first 5 modes.

over the total angular range of the 2PCF, condensing the information from the 2PCF naturally into a set of discrete modes. The linear T-functions can be expressed conveniently as Legendre polynomials, however $T_{\pm n}^{\log}$ compresses the cosmological information into significantly fewer modes; we therefore choose the logarithmic COSEBIs as our second-order shear statistic in the likelihood analysis in Chapter. 9. The COSEBI filter functions are displayed graphically in Fig. 8.4.

Figure 8.5 shows three different E/B-mode statistics derived from our measured shear-shear correlation function, i.e. the COSEBIs, the ring statistics, and the aperture mass dispersion. The error bars are obtained from the square root of the corresponding covariances' diagonal elements (statistics only). Note that the COSEBIs data points are significantly correlated. Slightly smaller is the correlation for the aperture mass dispersion, and the ring statistics' data points have the smallest correlation.

From the COSEBIs, we find a reduced χ^2 for the E-modes to be consistent with zero of 6.395, versus 1.096 for the B-modes (5 degrees of freedom each). The latter is consistent with purely statistical fluctuations.

8.2 Covariance estimation

8.2.1 Ellipticity correlation function covariance matrix

The covariance matrix of the ellipticity correlation function estimated via Eq. (8.5) was computed in several ways. The preferred method for our analysis is a Monte Carlo method (Sec. 8.2.1) but we compare that covariance matrix with an estimate of the Poisson errors (Sec. 8.2.1) as a consistency check.

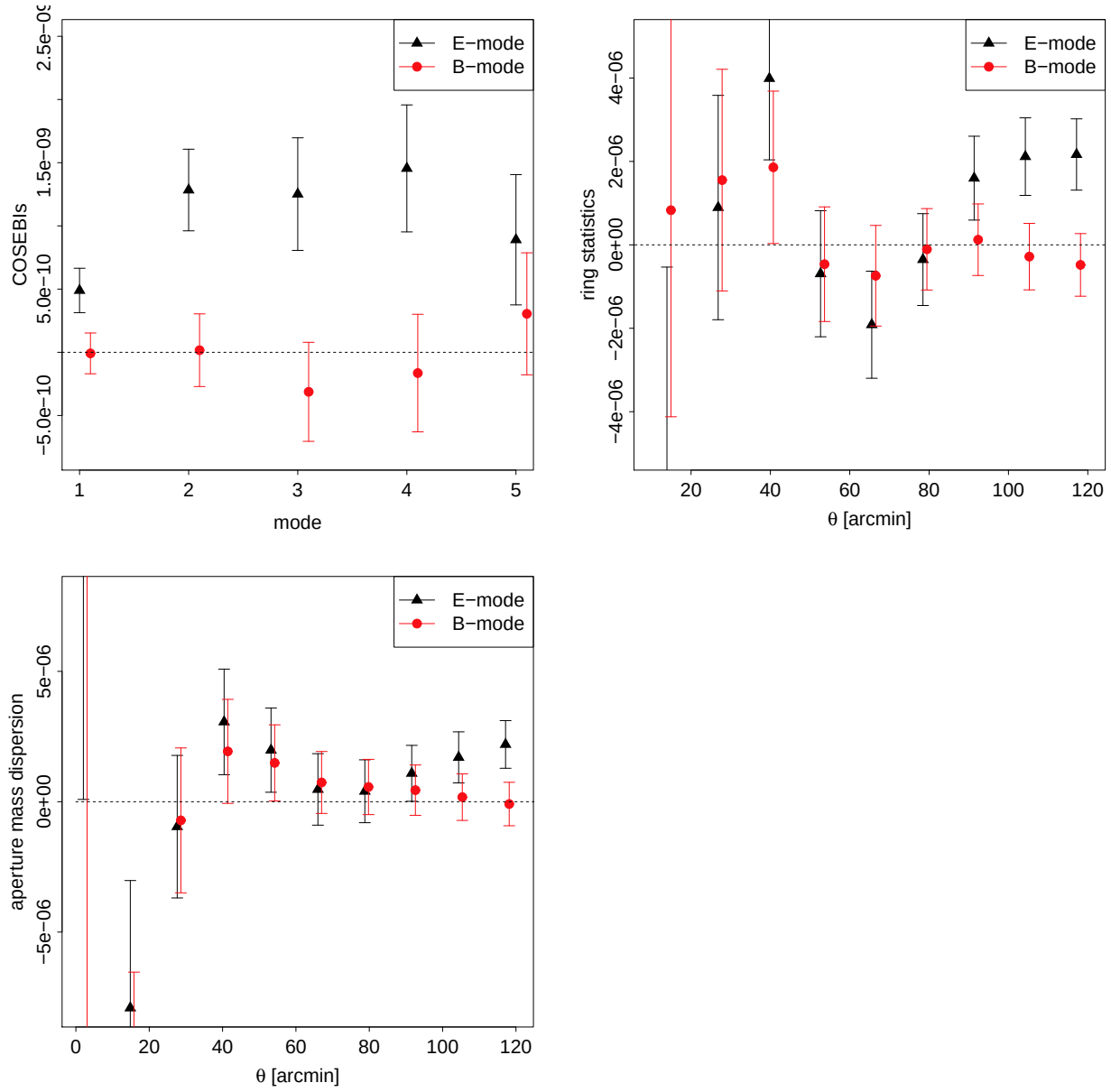

Figure 8.5 The measured COSEBIs, ring statistics, and aperture mass dispersion from the combined cosmic shear signal. The error bars equal the square root of the corresponding covariances' diagonal elements (statistics only). Note that the COSEBIs data points are significantly correlated. Slightly smaller is the correlation for the aperture mass dispersion, and the ring statistics' data points have the smallest correlation.

Poisson method

The direct pair-count correlation function code can compute the Poisson error bars, i.e. the error bars neglecting the correlations in $e_{i+}e_{j+}$ between different pairs. This estimate of the error bar is

$$\sigma^2[\xi_{++}(\theta)] = \frac{\sum_{ij} w_i^2 w_j^2 |\mathbf{e}_i|^2 |\mathbf{e}_j|^2}{2\left[\sum_{ij} w_i w_j\right]^2}. \tag{8.9}$$

Equivalently, this is the variance in the correlation function that one would estimate if one randomly re-oriented all of the galaxies. The Poisson method is simple, however it is not fully appropriate for ri cross-correlations (since the same intrinsic shape noise is recovered twice for pairs that appear in both ri and ir cross-correlations). Moreover, at scales of tens of arcminutes and greater there is an additional contribution because the cosmic shear itself is correlated between pairs. Therefore the Poisson error bars should be used only as a visual guide: they would underestimate the true uncertainties if used in a cosmological parameter analysis.

Monte Carlo method

We used a Monte Carlo method to compute the covariance matrix of $\xi_{++}(\theta)$ and $\xi_{\times\times}(\theta)$. The method is part theoretical and part empirical: it is based on a theoretical shear power spectrum, but randomizes the real galaxies to correctly treat the noise properties of the survey. The advantages of the Monte Carlo method – as implemented here – are that spatially variable noise, intrinsic shape noise including correlations between the r and i band, and the survey window function are correctly represented. The principal disadvantages are that the cosmic shear field is treated as Gaussian and a particular cosmology must be assumed (see Eifler et al., 2009, for alternative approaches). However, so long as this cosmology is not too far from the correct one (an assumption that can itself be tested!), the Monte Carlo approach is likely to yield the best covariance matrix.

The Monte Carlo approach begins with the generation of a suite of 459 realizations of a cosmic shear field in harmonic space according to a theoretical spectrum. For our analysis, the theoretical spectrum was that from the *WMAP* 7-year (Larson et al., 2011) cosmological parameter set (flat ΛCDM; $\Omega_{\rm b}h^2 = 0.02258$; $\Omega_{\rm m}h^2 = 0.1334$; $n_{\rm s} = 0.963$; $H_0 = 71.0\,{\rm km\,s^{-1}\,Mpc^{-1}}$; and $\sigma_8 = 0.801$), and the shear power spectrum code used in Albrecht et al. (2009), itself based on the Eisenstein & Hu (1998) transfer function and the Smith et al. (2003) nonlinear mapping. The redshift distribution discussed in section 7.4.1, based on a calibration sample from DEEP2, VVDS, and PRIMUS, was used as the input to the shear power spectrum calculation.

From this power spectrum we generate a sample set of Gaussian E-mode shear harmonic space coefficients $a^{\rm E}_{lm}$. The full power spectrum is used at $l \leq 1500$; a smooth cutoff is applied from $1500 < l < 2000$ and no power at $l \geq 2000$ is included. This is appropriate for a covariance matrix since the power at smaller scales is shot noise dominated and cannot

be recovered. (The E-mode power spectrum is $C_{1500}^{EE} = 3.6 \times 10^{-11}$, as compared to a shot noise of $\gamma_{\rm int}^2/\bar{n} \sim 1.8 \times 10^{-9}$.) No B-mode shear is included. The particle-mesh spherical harmonic transform code of Hirata et al. (2004a) with a 6144×3072 grid ($L' = 6144$) and a 400-node interpolation kernel ($K = 10$) was used to transform these coefficients into shear components (γ_1, γ_2) at the position $\hat{\mathbf{n}}_j$ of each galaxy j.[2]

A synthetic ellipticity catalogue was then generated as follows. For each galaxy, we generated a random position angle offset $\psi_j \in [0, \pi)$ and rotated the ellipticity in both r and i bands by ψ_j.[3] We then added the synthetic shear weighted by the shear responsivity to the randomized ellipticity to generate a synthetic ellipticity:

$$\mathbf{e}_j^{\rm syn} = {\rm e}^{2{\rm i}\psi_j}\mathbf{e}_j^{\rm true} + 1.73\mathbf{\Gamma}(\hat{\mathbf{n}}_j). \tag{8.10}$$

The 1.73 prefactor was estimated from Eq. (7.13), which we expected to be good enough for use in the Monte Carlo analysis, so that the Monte Carlos could be run in parallel with the shear calibration simulations. The latter gave a final result of 1.78 ± 0.04, which is not significantly different.

The direct pair-count correlation function code, in all versions (rr, ri, and ii) was run on each of the 459 Monte Carlo realisations, before combining the different correlations to get the weighted value via Eq. (8.5).

The Monte Carlo and Poisson error bars are compared in Fig. 8.6. The correlation coefficients of the correlation functions in different bins are plotted graphically in Fig. 8.7.

From each Monte Carlo correlation function we compute the COSEBIs via Eq. (8.7) and use their covariance matrix in our subsequent likelihood analysis. In order to test whether our covariance has converged, meaning that the number of realisations is sufficient to not alter cosmological constraints, we perform 3 likelihood analyses in σ_8 vs. Ω_m space varying the numbers of realisations from which we compute the covariance matrix (see Chap. 9 for detailed methodology; for now we are just establishing convergence of the covariance matrix). In Fig. 8.8 we show the 68 and 95 per cent likelihood contours, i.e. the contours enclose the corresponding fraction of the posterior probability (within the ranges of the parameters shown). We see that the contours hardly change when going from 300 to 400 realizations and show no change at all when going from 400 to 459 realisations, hence the 459 Monte Carlo realizations are sufficient for our likelihood analysis.

8.2.2 Systematic contributions to the covariance matrix

The following additional contributions are added to the Monte Carlo covariance matrix (and if appropriate the theory result) described in Sec. 8.2.1.

[2]The use of a full-sky approach for the Monte Carlo realisations was not necessary for the SDSS Stripe 82 project, but was the simplest choice given legacy codes available to us.

[3]To simplify bookkeeping, the actual implementation was that a sequence of 10^7 random numbers was generated, and a galaxy was assigned one of these numbers based on its coordinates in a fine grid with 0.36 arcsec cells in (α, δ).

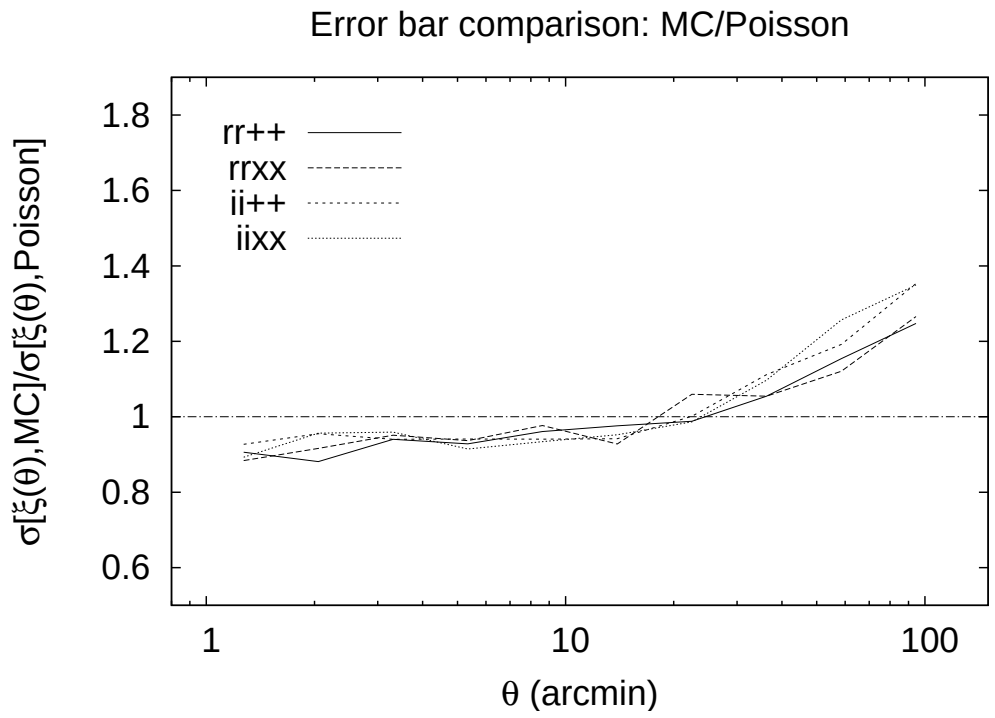

Figure 8.6 The ratio of error bars obtained by the Monte Carlo method to those obtained by the Poisson method, for 10 angular bins. The four curves show either rr or ii band correlation functions, and either the $++$ or $\times\times$ component. Note the rise in the error bars at large values of the angular separation, due to mode sampling variance.

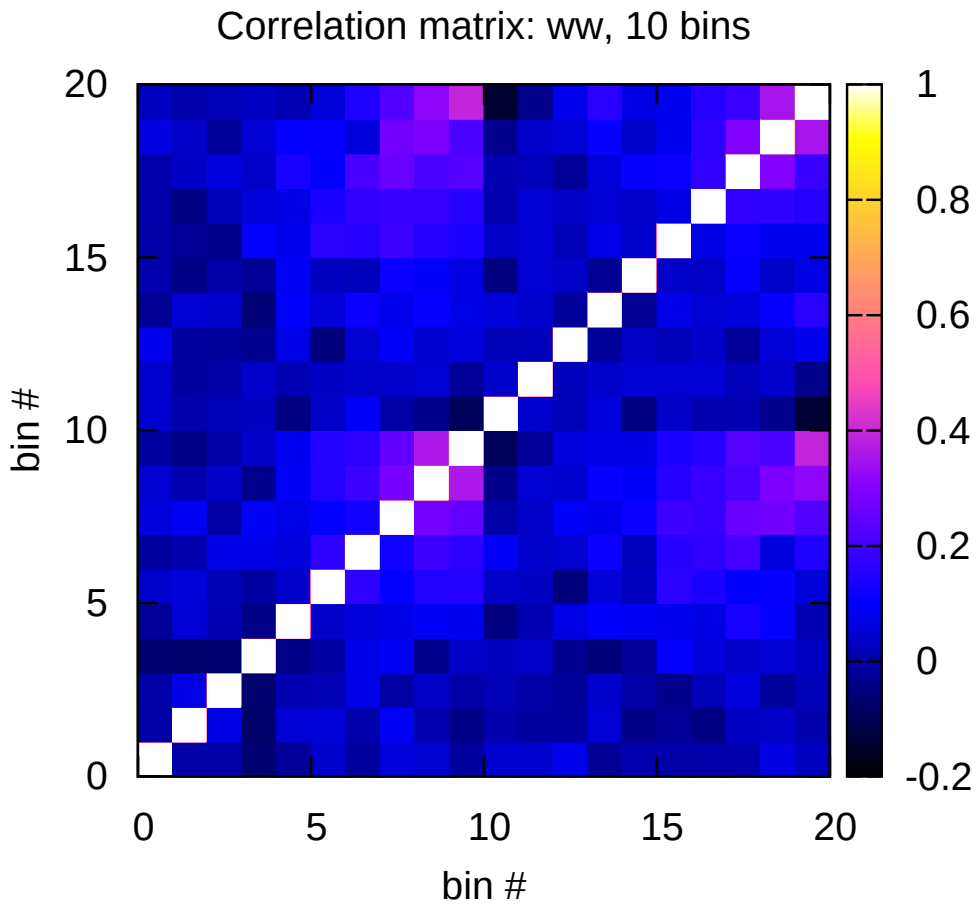

Figure 8.7 The matrix of correlation coefficients for the combined (ww) correlation functions in the 10 angular bins for which the correlation function is plotted in the companion figures. The bin number ranges from 0–9 for $\xi_{++}(\theta)$ and from 10–19 for $\xi_{\times\times}(\theta)$; all diagonal components are by definition equal to unity. Based on 459 Monte Carlo realisations.

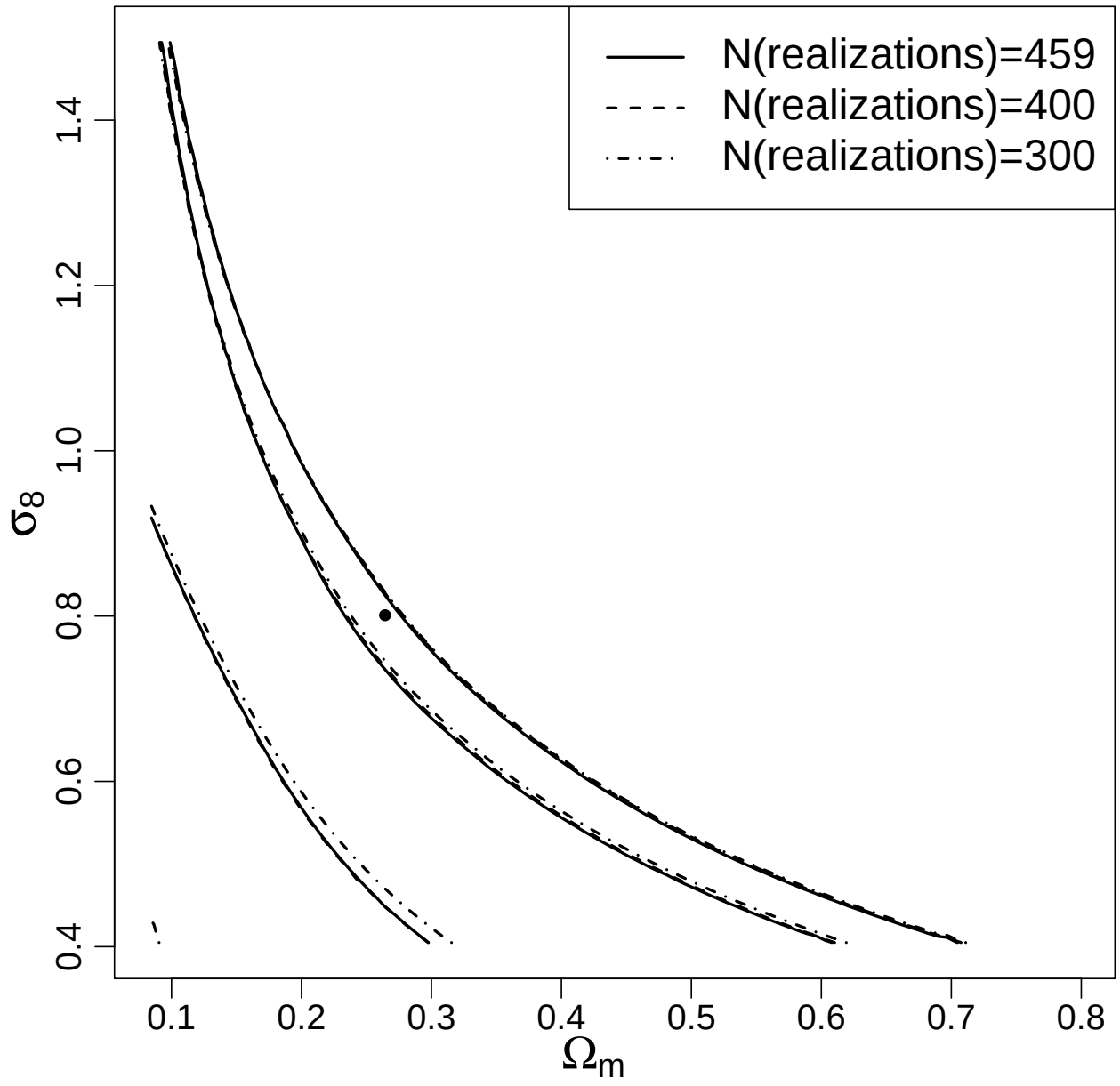

Figure 8.8 Convergence test of the σ_8 vs. Ω_m parameter constraints as a function of the number of Monte Carlo realizations used to compute the covariance. The plot shows the 68 and 95 per cent likelihood contours (however, the lower 95 per cent contours are not visible). The covariance includes statistical errors only.

1. The intrinsic alignment error was included following Sec. 7.2: the theory shear correlation function was reduced by a factor of 0.92, and an uncertainty of 4 per cent of the theory was added to the covariance matrix, i.e. we add an intrinsic alignment contribution
$$\mathrm{Cov}[\xi_i,\xi_j](\text{intrinsic alignment}) = 0.04^2\xi_i^{(\mathrm{th})}\xi_j^{(\mathrm{th})}, \tag{8.11}$$
where the theory curve (th) is obtained at the fiducial WMAP7 point. This covariance matrix includes perfect correlation between radial bins, implying that we treat this systematic as being an effect with a fixed scaling with separation, so the only degree of freedom is its amplitude.

2. The stellar contamination was included following Sec. 7.5: the theory shear correlation function was reduced by a factor of 0.936, and an uncertainty of 3 per cent of the theory was added to the covariance matrix, i.e. we add a stellar contamination contribution
$$\mathrm{Cov}[\xi_i,\xi_j](\text{stellar contamination}) = 0.03^2\xi_i^{(\mathrm{th})}\xi_j^{(\mathrm{th})}, \tag{8.12}$$
where the theory curve (th) is obtained at the fiducial WMAP7 point.

3. The implied error from the redshift distribution uncertainty is derived from 402 realisations of the sampling variance simulations as described in Sec. 7.4.2. We construct the covariance matrix of the predicted E-mode COSEBIs.

4. The shear calibration uncertainty was conservatively estimated in Sec. 7.3 to be ± 2.4 per cent, or equivalently 4.8 per cent in second-order statistics. We thus add another term to the covariance matrix,
$$\mathrm{Cov}[\xi_i,\xi_j](\text{shear calibration}) = 0.048^2\xi_i^{(\mathrm{th})}\xi_j^{(\mathrm{th})}. \tag{8.13}$$

5. In Chap. 7.6, we described a procedure for including uncertainty due to additive PSF contamination. According to this procedure, the relevant systematics covariance matrix is related to the amplitude of the measured contamination signal:
$$\mathrm{Cov}[\xi_i,\xi_j](\text{PSF contamination}) = 0.9^2\xi_{\mathrm{sg},i}\xi_{\mathrm{sg},j}, \tag{8.14}$$
again assuming a fixed scaling with radius for this systematic uncertainty. Since all entries scale together, we do not spuriously "average down" our estimate of the systematic error by combining many bins.

The final data vector and its covariance matrix (including all the statistical and systematic components) are given in Tables 8.1 and 8.2. Note that given our procedure of applying the systematic corrections to the theory, the data vector is the observed one without any such corrections for the stellar contamination and intrinsic alignments contamination. With this in hand, we can estimate the significance of the E- and B-mode signals described in section 8.1.3. The probability that the COSEBI E-mode signal that we observe is due to random chance given the null hypothesis (no cosmic shear) is 6.0×10^{-6}. The probability of measuring our B-mode signal due to random chance given the null hypothesis of zero B modes is .36, evidence that there is no significant B-mode power.

Table 8.1 Our data vector. The first five elements are COSEBI mode amplitudes; the final is the correlation function averaged in the range $29.2296 \leq \theta \leq 44.9730$.

4.89797E-10
1.28335E-09
1.25136E-09
1.45616E-09
8.92333E-10
1.46457E-05

Table 8.2 The covariance matrix for the data vector shown in table 8.1.

Data vector index	Data vector index	Covariance
0	0	3.37161E-20
0	1	4.67637E-20
0	2	4.00484E-20
0	3	2.49916E-20
0	4	9.84257E-21
0	5	3.01770E-17
1	1	1.06383E-19
1	2	1.19226E-19
1	3	8.39508E-20
1	4	3.86519E-20
1	5	1.82344E-16
2	2	1.99923E-19
2	3	1.87469E-19
2	4	1.12196E-19
2	5	5.07790E-16
3	3	2.56568E-19
3	4	2.13363E-19
3	5	8.02118E-16
4	4	2.67774E-19
4	5	5.67797E-16
5	5	3.68112E-11

Chapter 9

Cosmological Constraints

Having described the measured cosmic shear two-point statistics, and shown that the systematic bias in this measurement is small compared with the statistical constraints, we now turn to the cosmological interpretation. We work in the context of the flat ΛCDM parametrisation, taking where necessary the WMAP7 (Komatsu et al., 2011) constraints for our fiducial parameter values.

9.1 The prediction code: modeling second-order shear statistics

To produce a cosmological interpretation of our measured cosmic shear signal from our model framework, we require a method to convert a vector of cosmological parameters into a prediction of the observed cosmic shear signal. Due to projection effects, we expect that a significant fraction of the observed cosmic shear signal is produced by the clustering of matter on nonlinear scales, so a suitably accurate prediction algorithm must ultimately rely on numerical simulations of structure formation.

The prediction code used in our likelihood analysis is a modified version of the code described in Eifler (2011). We combine Halofit (Smith et al., 2003), an analytic approach to modeling nonlinear structure, with the Coyote Universe Emulator (Lawrence et al., 2010), which interpolates the results of a large suite of high-resolution cosmological simulations over a limited parameter space, to obtain the density power spectrum. The derivation is a two-step process: First, we calculate the linear power spectrum from an initial power law spectrum $P_\delta(k) \propto k^{n_\mathrm{s}}$ employing the dewiggled transfer function of Eisenstein & Hu (1998). The non-linear evolution of the density field is incorporated using Halofit. In order to simulate wCDM models we follow the scheme implemented in ICOSMO (Refregier et al., 2011), interpolating between flat and open cosmological models to mimic Quintessence cosmologies (see Schrabback et al. 2010 for more details). In a second step, we match the Halofit power spectrum to the Coyote Universe emulator (version 1.1) power spectrum, which emulates P_δ over the range $0.002 \leq k \leq 3.4h/\mathrm{Mpc}$ within $0 \leq z \leq 1$ to an accuracy of 1 per cent.

Wherever possible, the matched power spectrum exactly corresponds to the Coyote Universe Emulator; of course this is limited by the cosmological parameter space of the Emulator and its limited range in k and z. However, even outside the range of the Emulator, we rescale the Halofit power spectrum with a scale factor $P_\delta(\text{Coyote})/P_\delta(\text{Halofit})$ calculated at the closest point in parameter space (cosmological parameters, k, and z) where the Emulator gives results. Outside the range of the Emulator, the accuracy of this "Hybrid" density power spectrum is of course worse than 1 per cent, however it should be a significant improvement over a density power spectrum from Halofit only. From the so-derived density power spectrum we calculate the shear power spectrum via Eq. (7.10) and the shear-shear correlation function via Eq. (7.9). As a final step, we transform these predicted correlation functions to the COSEBI basis as described above in Sec. 8.1.3.

For our final results in the (Ω_m, σ_8) likelihood analysis, we used both prediction codes; the results are compared in Fig. 9.1, where they are seen to agree to much better than 1σ. We therefore conclude that uncertainty in the theory predictions is sub-dominant to the other sources of systematic error, and to the statistical error.

9.2 Constructing the input data vector

For our primary science results, we use the measured 5 COSEBI modes (see Fig. 8.5, left panel). As a first step we want to determine the number of COSEBI modes that need to be included in our likelihood analysis. In Fig. 9.2 we show a likelihood analysis in the σ_8-Ω_m parameter space varying the number of modes in the data vector. We find that there is hardly a change in the likelihood contours when going from 4 to 5 modes; we therefore conclude that 5 modes is a sufficient number to capture the cosmological information encoded in our data set.

As shown in Eifler et al. (2008), the information content of the aperture mass dispersion can be greatly improved when including 1 data point of the shear-shear correlation function ξ_+ into the data vector; here we adopt this concept for the COSEBIs. The basic idea is that the data point of the correlation function is sensitive to scales of the power spectrum to which the COSEBIs are insensitive. We incorporate only a single data point of the correlation function as this is sufficient to capture the bulk of the additional information while simultaneously minimizing possible B-mode contamination.

In order to determine the optimal scale of the data point that is to be included, we consider 10 bins of ξ_+ ranging from 1.3 to 97.5 arcmin and perform 10 likelihood analyses for a combined data vector consisting of 5 COSEBI modes and one additional data point of ξ_+. We quantify the information content through the so-called q figure of merit (q-FoM)

$$\begin{aligned} q &= \sqrt{|\mathbf{Q}|}, \quad \text{where} \\ Q_{ij} &= \int \mathrm{d}^2\pi \, p(\pi|\,\mathbf{d}) \, (\pi_i - \pi_i^\mathrm{f})(\pi_j - \pi_j^\mathrm{f}) \,, \end{aligned} \tag{9.1}$$

$\pi = (\Omega_m, \sigma_8)$ is the parameter vector, $p(\pi|\mathbf{d})$ is the posterior likelihood at this parameter

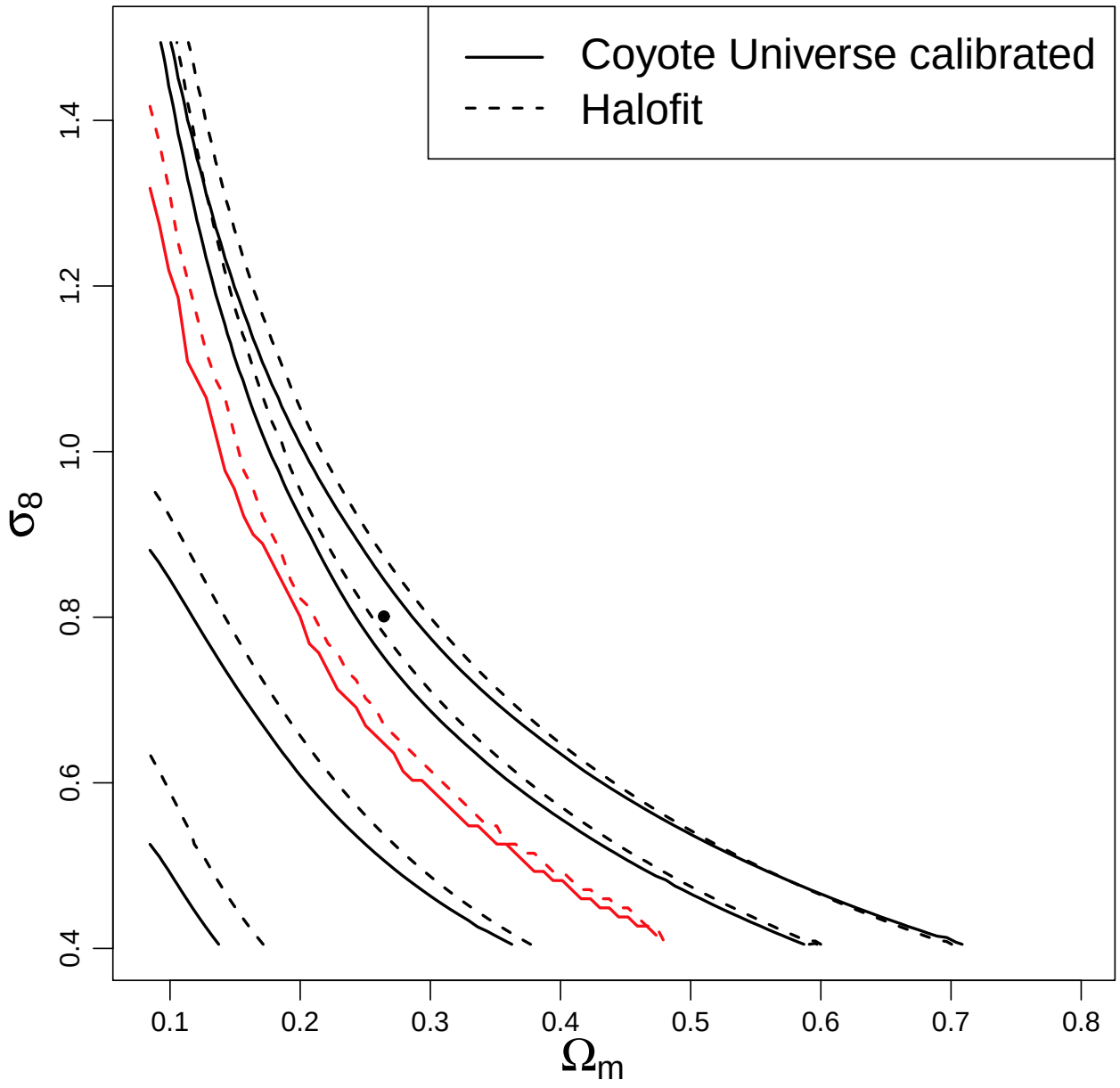

Figure 9.1 The 68 and 95 per cent likelihood contours of the combined data vector including a full treatment of systematics when using the Halofit prediction code (dashed) and when using the Coyote Universe-calibrated prediction code (solid). The red lines correspond to the best-fitting value of σ_8 for a given Ω_m. The dot indicates the WMAP7 best-fitting values.

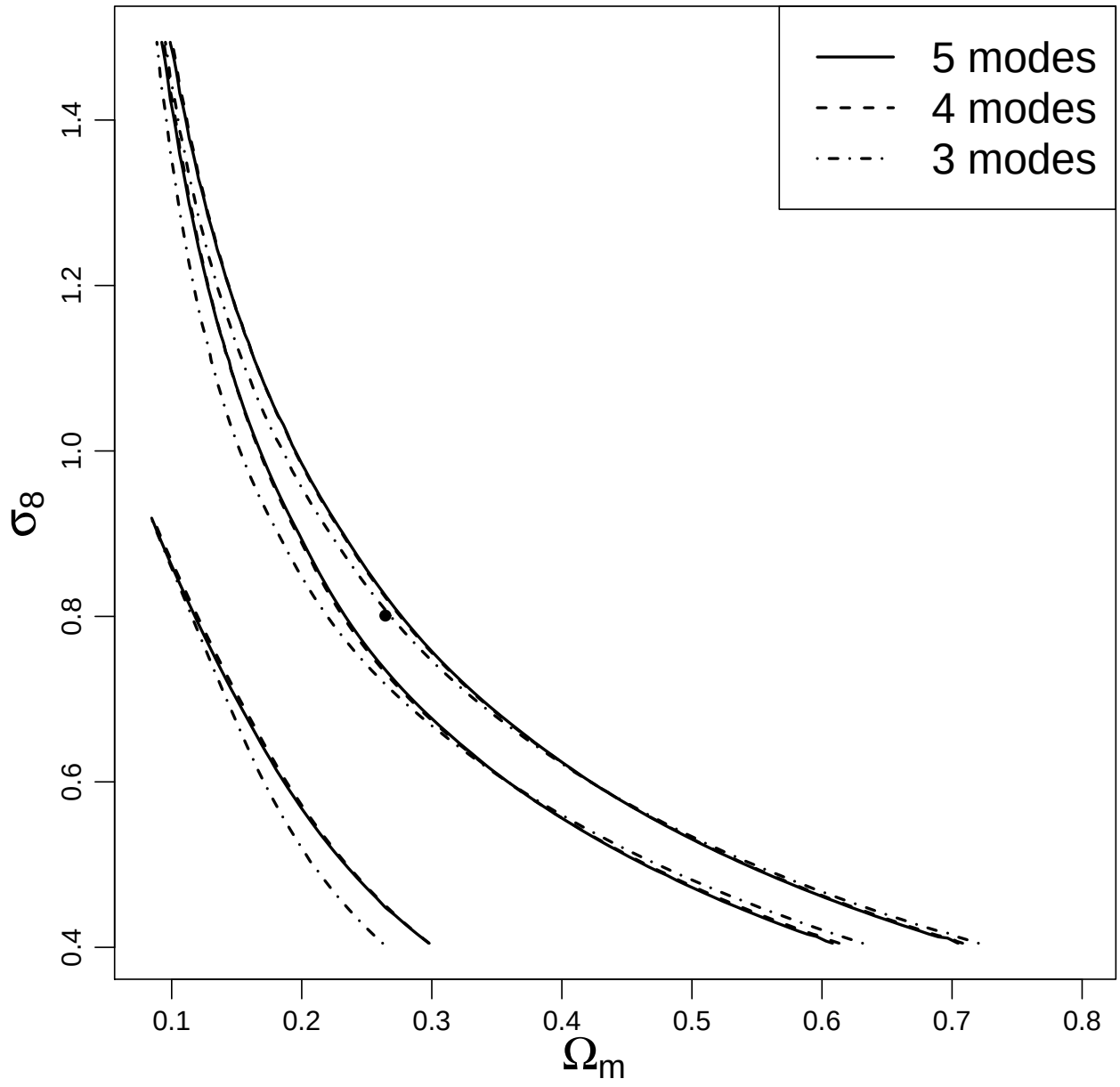

Figure 9.2 Convergence test of the σ_8 vs. Ω_m parameter constraints as a function of number of COSEBI modes in the data vector. The plot shows the likelihood contours enclosing 68 and 95 per cent of the posterior distribution. (The lower bounding curve for the 95 per cent contours is not visible on the plot.) The covariance contains statistical errors only. The dot indicates the WMAP7 best-fitting values.

point, and $\pi_i^{\rm f}$ denotes the fiducial parameter values. If the likelihood in parameter space (i.e. the posterior probability) is Gaussian, the q-FoM corresponds to the more common Fisher matrix based figure of merit $f = 1/|\sqrt{\mathbf{F}}|$. The Fisher matrix $\mathbf{F}$ can be interpreted as the expectation value of the inverse parameter covariance evaluated at the maximum likelihood estimate parameter set, which in our ansatz corresponds to the fiducial parameters. Mathematically we can express this equivalence as

$$f = \frac{1}{\sqrt{|\mathbf{F}|}} = \sqrt{|\mathbf{C}_\pi|} = \sqrt{|\mathbf{Q}|} = q\,. \tag{9.2}$$

Since the assumption of a Gaussian posterior is clearly violated in the σ_8-$\Omega_{\rm m}$ parameter space, we perform a full likelihood analysis and calculate q to quantify the size of the likelihood. Note that smaller q-FoM is "better."

We varied the angular scale (in arcmin) of the added $\xi_+(\theta)$ data point, and found a minimal q-FoM at $\theta = 37.8$ arcmin. We will use this scale for the additional ξ_+ data point henceforth. Note that this analysis uses a simulated input data vector in order to avoid biases from designing a statistical test based on the observed data. The constraints coming from the various possible data vectors – the COSEBIs, the COSEBIs supplemented with a single ξ_+ point, and the full shear correlation function – are compared in Fig. 9.3. They are not identical, which is expected since they weight the data in different ways, but are consistent with each other.

The COSEBI modes are highly correlated with each other, and they are correlated to a lesser extent with ξ_+ at 38 arcmin. The correlation matrix is shown in Fig. 9.4, and the corresponding covariance matrix is tabulated in the Appendix in Table 8.2.

9.3 Parameter Fits

We perform all of our fits to a standard five-parameter ΛCDM model[1]. For the initial likelihood analysis, we fix n_s, $\Omega_b h^2$, $\Omega_m h^2$, and w_0 at their fiducial best-fit WMAP7 values (Komatsu et al., 2011), and vary σ_8. The upper panel of Fig. 9.5 shows the likelihood of σ_8 with all other parameters fixed, with a value at the peak and 68 per cent confidence interval of $0.636^{+0.109}_{-0.154}$. For a survey of this size and depth, the constraints are comparable to the statistically achievable confidence limits.

We also perform a likelihood analysis fixing three parameters, and varying Ω_m and σ_8 simultaneously, as these two parameters are much more sensitive to the measured cosmic shear signal than the others. The resulting two-dimensional constraints are shown in the bottom panel of Fig. 9.5. Our 68 per cent confidence limits on the degenerate product $\sigma_8 \left(\frac{\Omega_m}{0.264}\right)^{0.67}$ are $0.65^{+0.12}_{-0.15}$ for the Coyote Universe prediction code (see Fig. 9.5, solid red line), and $\sigma_8 \left(\frac{\Omega_m}{0.264}\right)^{0.72} = 0.67^{+0.12}_{-0.15}$ for the Halofit prediction code (see Fig. 9.5, dashed red line).

[1]The optical depth to reionization τ is a sixth parameter implicitly included in the WMAP7 chains, but with no effect on the lensing shear correlation function.

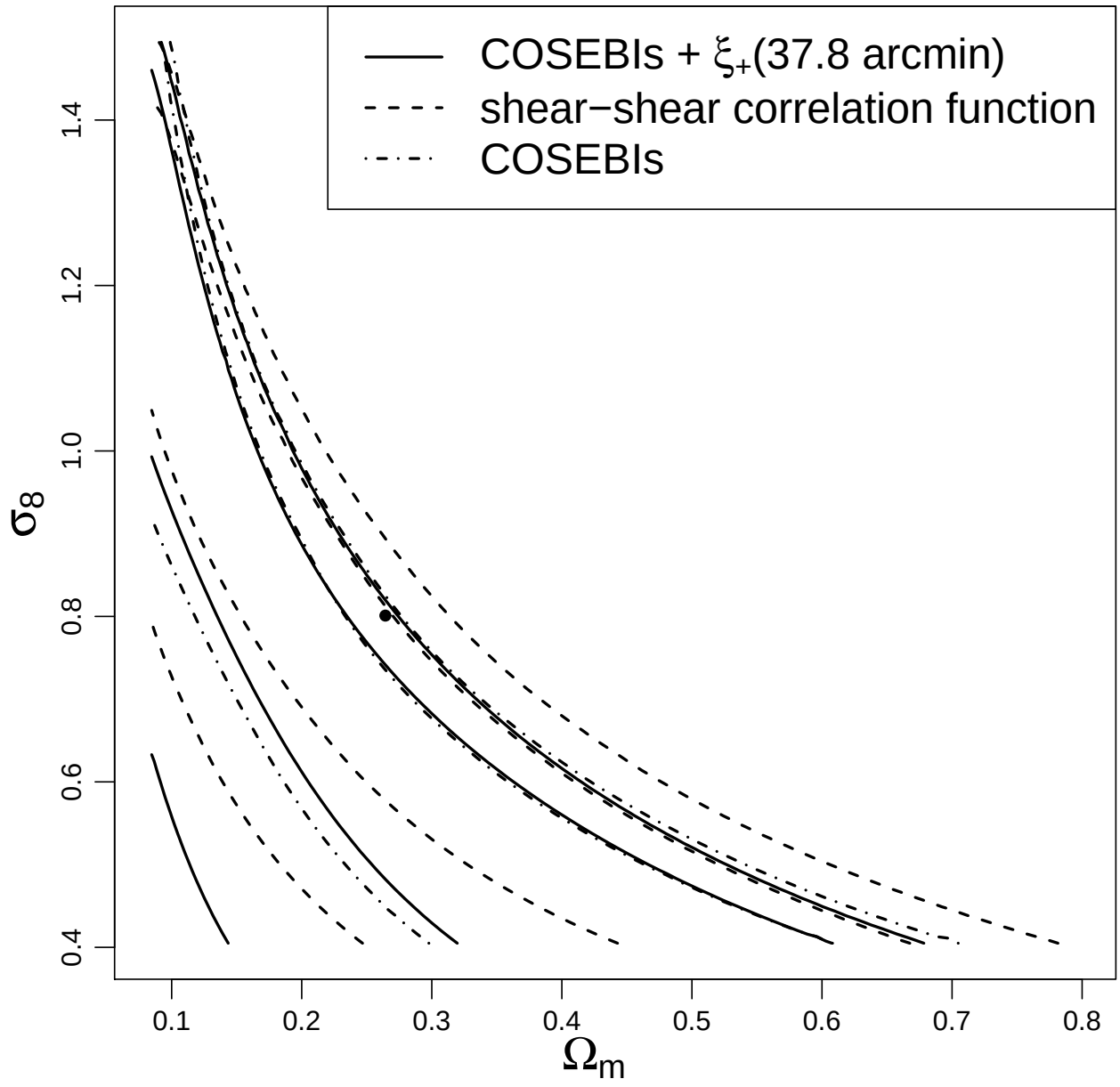

Figure 9.3 The likelihood contours of the combined data vector (solid), the shear-shear correlation function (dashed), and the COSEBIs (dotted) data vector to illustrate how much information is gained when including the additional data point. Note that the COSEBIs' lower 95 per cent contour is outside the considered region. The dot indicates the WMAP7 best-fitting values.

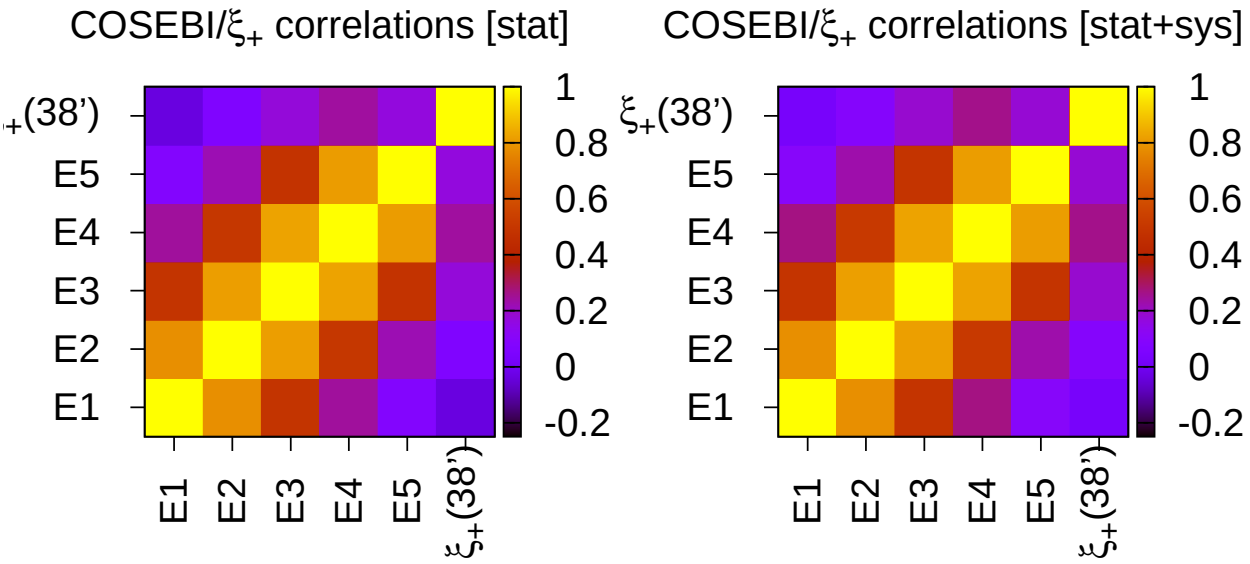

Figure 9.4 The correlation matrix of the COSEBI modes 1–5 ("E1...E5" in the figure) and $\xi_+(38')$. The left panel shows only the statistical (Monte Carlo) errors, and the right panel includes the systematics as well.

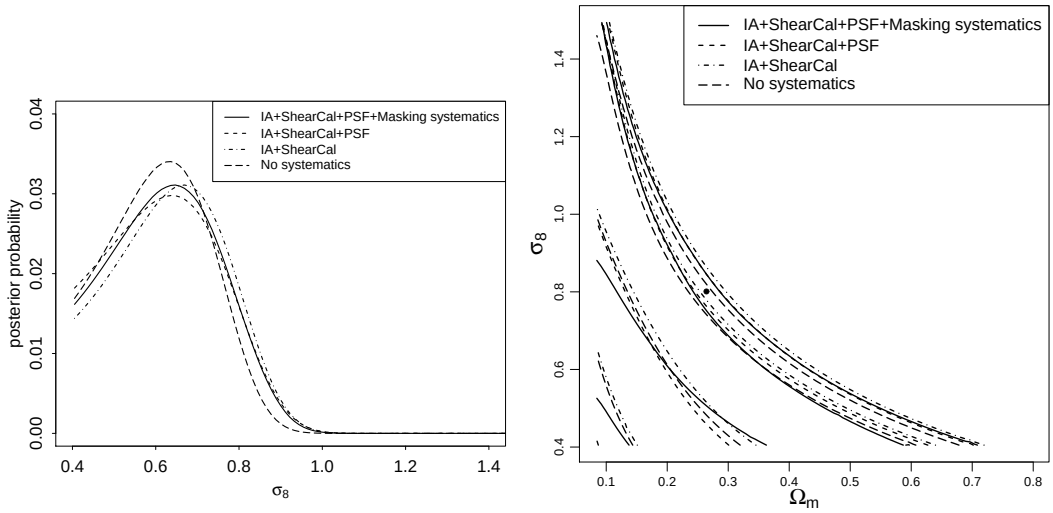

Figure 9.5 The effect of systematic errors in the 1-D likelihood of σ_8 (upper panel) and in the 2-D constraints (68 per cent likelihood contours only) in the $\sigma_8 - \Omega_m$ plane (lower panel). The solid curve shows our final analysis, while the other curves show results including subsets of the systematic errors. The dot-dashed curve labeled "no systematics" shows only the statistical errors, without any systematic error corrections either to the theory or to the covariance matrix. The dot indicates the WMAP7 best-fitting values.

We show the effects of removing each systematic error correction, Fig. 9.5 also shows, for both the one- and two-dimensional analyses, the impact of systematic error corrections. The combined effects of these uncertainties are clearly substantially smaller than the statistical error on the amplitude of the shear signal.

Finally, we adopt the WMAP7 likelihoods as priors, and evaluate our likelihood at each link in the WMAP7 Markov chain. For each chain element, we assign a weight equal to our likelihood function evaluated at the parameter vector for that chain element. For each of the parameter constraint plots shown here, we first assign each Markov Chain Monte Carlo (MCMC) chain element to a point on a regular grid in the parameter space; the value of the marginalised likelihood at each grid-point, $H_{i,j}$ is then the sum of our likelihood weights over the MCMC chain elements at the (i, j) grid-point,

$$H_{i,j} = \sum_k I_k(i,j) L_k, \tag{9.3}$$

where the indicator function $I_k(i, j)$ is equal to unity when the (i, j) grid-point in parameter space is nearest the kth chain element, and zero otherwise. The likelihood L_k for each chain element is evaluated in the usual way as:

$$L_k = \exp\left(-\frac{\bar{\mathbf{d}}_k^T C^{-1} \bar{\mathbf{d}}_k}{2}\right). \tag{9.4}$$

Here C is the full covariance matrix for the measurement, incorporating both the statistical and systematic uncertainties, and the normalization is arbitrary. The data vector $\bar{\mathbf{d}}_k$ is the extended COSEBI vector described above; where shown, the WMAP7 priors are simply this sum with $L_k = 1$ for each point.

We estimate the detection significance for the final signal, the difference $\sqrt{-2\Delta \log L}$ between the highest-likelihood Markov Chain element for both the ΛCDM and wCDM models and the likelihood evaluated with no signal. The 1σ detection significances for these two models are 2.64 and 2.88, respectively. This is not the significance of the detection of cosmic shear (as in Sec. 8.2.2), but rather a measurement of the likelihood of these two models given the combination of WMAP7 priors with this experiment.

In Fig. 9.6, we show marginalized posterior likelihoods in the case of fixed ΛCDM (i.e., $w = -1$) for $\Omega_m h^2$, $\Omega_b h^2$, n_s, and σ_8. The results with a free equation of state of dark energy (wCDM) are in Fig. 9.7. Our measurement provides some additional constraints beyond those from WMAP7 on these parameters. In particular, the low amplitude of the measured shear signal rules out some of the previously allowed volume of $\Omega_{\rm m} h^2$ and σ_8 WMAP7 constraints.

9.4 Conclusions

Using coadded imaging constructed from SDSS Stripe 82 data, we have constructed a weak lensing catalogue of 1 328 885 galaxies covering 168 square degrees, and showed that the additive shear systematics arising from the PSF are negligible compared to the cosmic shear signal. In this paper, we carried out a cosmic shear measurement that resulted in a 20 per cent constraint on σ_8 (with all other cosmological parameters fixed). This adds constraining power beyond that from WMAP7, and serves as an important independent data point on the amplitude of the matter power spectrum at late times. In particular, the primary CMB anisotropies presently provide only a modest constraint on $\Omega_m h^2$, and (due to the effect of matter density on the growth of structure) there is then an elongated allowed region in the $(\Omega_m h^2, \sigma_8)$ plane; see Fig. 9.6. The WMAP7-allowed region is ideally oriented for lensing to play a role: the lensing signal at the high-$\Omega_m h^2$, high-σ_8 end of the ellipse leads to a much higher lensing signal than low $\Omega_m h^2$, low σ_8. The low amplitude of cosmic shear observed in this paper eliminates the high-$\Omega_m h^2$, high-σ_8 solutions, and leads to a WMAP7+SDSS lensing solution of $\sigma_8 = 0.784^{+0.028}_{-0.026}(1\sigma)^{+0.055}_{-0.054}(2\sigma)$ and $\Omega_m h^2 = 0.1303^{+0.0047}_{-0.0048}(1\sigma)^{+0.0091}_{-0.0092}(2\sigma)$; the 2σ error ranges are respectively 14 and 17 per cent smaller than for WMAP7 alone.

We have also carefully evaluated other sources of uncertainty such as the source redshift distribution, intrinsic alignments, and shear calibration, to ensure that our measurement is dominated by statistical errors rather than systematic errors. This achievement is important when considering that (i) the SDSS data were never designed with this application in mind, and indeed includes several features (e.g. the minimal amount of cross-scan dithering) that cause significant difficulty, and (ii) with the multitude of upcoming multi-exposure lensing surveys in the next few years, it is important to cultivate new data analysis techniques (such

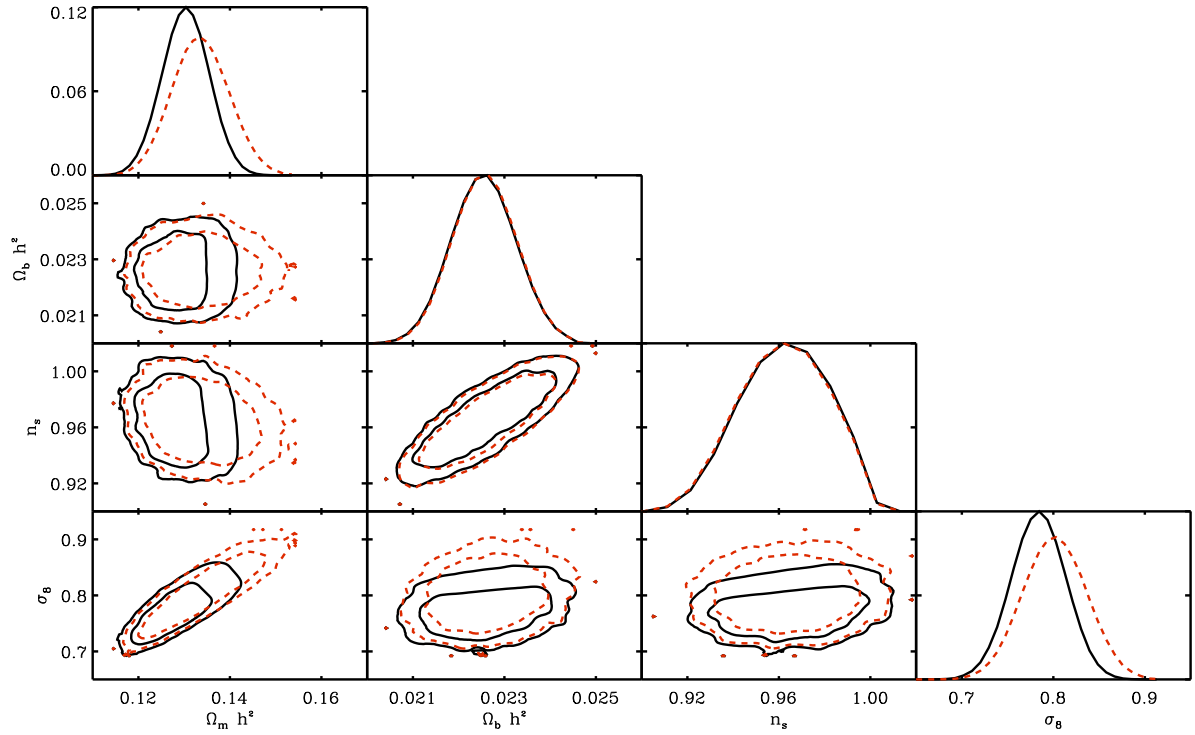

Figure 9.6 The cosmological parameter constraints using the extended COSEBI data vector, fixing the dark energy equation of state w at -1, but allowing all other parameters to vary. Off-diagonal panels show joint two-dimensional constraints after marginalization over all the other parameters, which are shown. For these, the red contours show the WMAP7 priors containing 68.5 and 95.4 per cent of the posterior probability. The black contours are the same but for WMAP7+SDSS lensing. Diagonal panels show the fully-marginalized one-dimensional posterior distribution for each parameter; for these panels, the red (dashed) contours show the marginalized WMAP7 constraints.

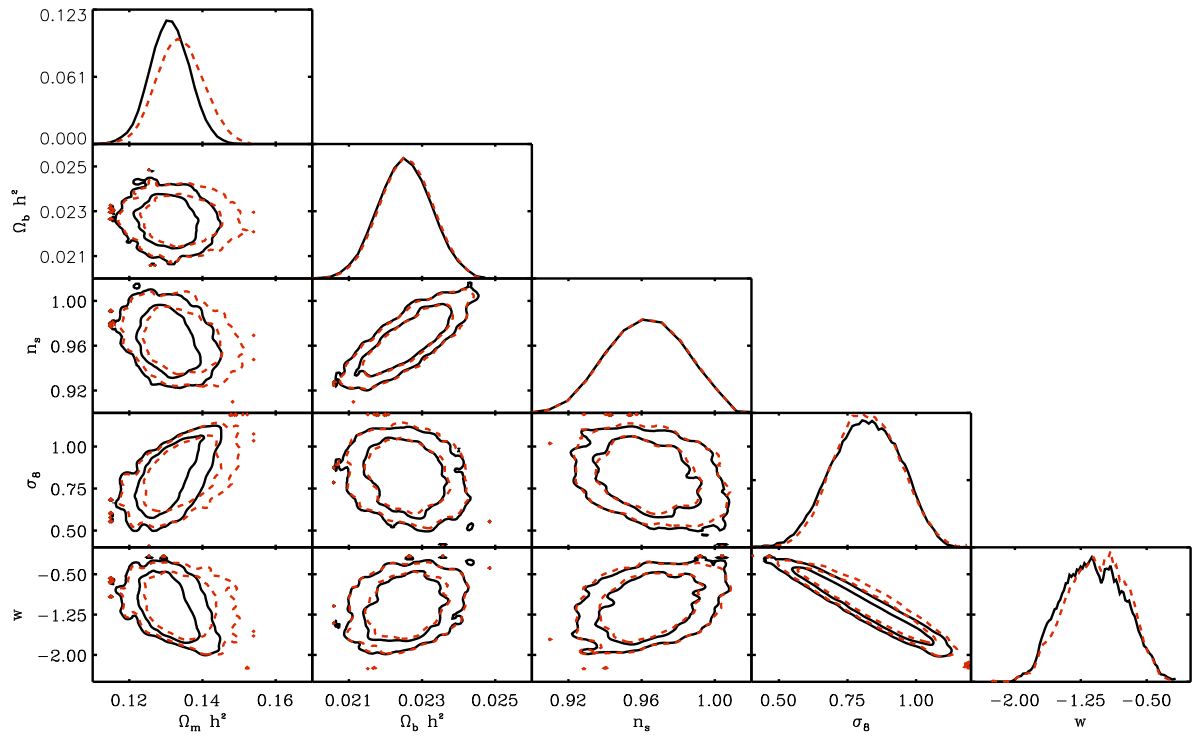

Figure 9.7 The cosmological parameter constraints using the extended COSEBI data vector, varying all five parameters. Off-diagonal panels show joint two-dimensional constraints after marginalization over all the other parameters, which are shown. For these, the red contours show the WMAP7 priors containing 68.5 and 95.4 per cent of the posterior probability. The black contours are the same but for WMAP7+SDSS lensing. Diagonal panels show the fully-marginalized one-dimensional posterior distribution for each parameter; for these panels, the red (dashed) contours show the marginalized WMAP7 constraints.

as the one used here) that are capable of producing homogeneous data with tight control over PSF anisotropies. As a quantitative measure of the extent of PSF correction possible with SDSS data, we take the RMS residual spurious shear at a particular scale estimated from the star-galaxy correlations,

$$\gamma_{\rm rms,eq}(\theta) = \frac{\sqrt{\mathcal{R}_{\rm psf}\,\xi_{+,\rm sg}(\theta)}}{\mathcal{R}}. \tag{9.5}$$

From Fig. 7.6, we see that this is $\sim 2\times10^{-3}$ at the smallest scales (1–6 arcmin), is $< 10^{-3}$ at scales $\theta > 0.1$ degree, and drops to 3.7×10^{-4} in the final bin (1.2–2.0 degrees).[2] There is almost no difference between the ++ and $\times\times$ signals, suggesting that the spurious additive ellipticity signal contains similar amounts of E- and B-modes[3]; something similar was seen in the SDSS single-epoch data via run-by-run comparisons of ellipticity measurements on the same galaxies (Mandelbaum et al., 2006a, Fig. 8). This is good news for the use of the B-mode as a diagnostic of PSF systematics, although an understanding of the generality of this pattern remains elusive.

A major lesson learned from this project is the importance of masking bias, in which the intrinsic orientation of a galaxy affects whether it falls within the survey mask. This is likely the main reason why we had to implement the $\langle e_1\rangle$ projection. While we have clearly not exhausted the range of options for removing this bias at the catalogue level, future surveys should be designed to produce more uniform data quality via an appropriate dithering strategy and suppress the masking bias at the earliest stages of the analysis.

Our major limitation in the end was the source number density, which was driven by the fact that our PSF-matching procedure was limited by the worst seeing in the images that we use, and therefore we had to eliminate the images with seeing worse than the median. This means that the coadds were not as deep as they could have been, and the final effective seeing was 1.31 arcsec (full-width half maximum). In principle this will be an obstacle to applying this technique in the future, but in fact, that statement depends on context. For example, for a survey such as HSC or LSST where we expect typically ~ 0.7 arcsec seeing, and with plans to preferentially use the best-seeing nights for r and i-band imaging that will be used for shape measurement, it is conceivable that nearly all images intended for lensing will have seeing in the 0.6–0.8 arcsec range. In that context, a PSF-matched coadd that has the rounding kernel applied may actually not result in much loss of information about the shapes of most useful galaxies, and will have the advantage of the removal of PSF anisotropies. Moreover, even for surveys for which the loss of information that results from this method may not be suitable for the final cosmological analysis, this method may still serve as a useful diagnostic of the additive PSF systematics.

[2] We used $\mathcal{R}_{\rm psf} = 0.9$ and $\mathcal{R} = 1.776$, as described in the text.

[3] Recall that $\xi_{++}(\theta) - \xi_{\times\times}(\theta)$ and $P_E(\ell) - P_B(\ell)$ are J_4 Hankel transforms of each other.

Chapter 10

The Future

Broadly speaking, weak lensing is best thought of as deeply signal-starved. For a comparable level of effort, the marginal impact of an increase in signal-to-noise is likely to be much larger than methodological improvements or increased statistical sophistication.

The additive shear bias due to improper PSF modeling is a good example. The response of the measured shear field to a small PSF anisotropy is discussed in Chapter 4. This error scales strongly (as R_2^4 in the two-point statistics) with the relative size of the PSF and galaxy. An increase in signal-to-noise that permitted the lensing analysis to rely on slightly better-resolved galaxies could reduce the magnitude of all PSF-dependent systematics by a large factor.

In general a stronger signal should allow more accurate modeling of systematic effects. For this reason, it is likely that the major gains yet to be realized in weak lensing will come from new probes of the underlying signal.

In this Chapter, I discuss two ideas that may offer opportunities to realize such gains. Both are still relatively unproven, but early results are very promising.

10.1 The Photometric Fundamental Plane

The fundamental plane (FP) is in many ways an ideal tool for measuring magnification. It is an observed correlation between galaxy effective radius (R_e), which is magnified by gravitational lensing, and two galaxy properties which are unaltered by lensing: galaxy surface brightness (μ) and the stellar velocity dispersion (σ). The intrinsic scatter in the FP is ~ 0.08 dex (Jørgensen et al., 1996; Bernardi et al., 2003), or 20%. Thus the FP makes it possible to predict the intrinsic value of R_e from observations of μ and σ, which can then be compared with the observed values of R_e to measure magnification.

The FP was in fact proposed as a tool for this purpose by Bertin & Lombardi (2006), but it has never been used as such due to a critical flaw. Placing galaxies on the FP requires σ measurements. Even with the tight scatter in the FP, a statistically viable measurement would require high-resolution spectroscopic measurements for millions of galaxies.

Identifying a purely photometric analog to the FP with comparable scatter would solve this problem. Such a relation has already been identified by Graham (2002), where the concentration of the galaxy light profile fills the role normally served by σ. This works in part because concentration and velocity dispersion are both strongly correlated with galaxy mass, and in part because at fixed mass, galaxies with more concentrated mass profiles have higher velocity dispersions. The relation between the spectroscopic fundamental plane and the photometric relation deployed here will be explored more fully in subsequent work.

10.1.1 Background Sources

To define a photoFP for this work, a sample of galaxies is drawn from the Sloan Digital Sky Survey III (SDSS-III) Eighth Data Release (DR8, Aihara et al. 2011b). The sample is limited to resolved sources that meet basic quality cuts (e.g., are not saturated). For these, we estimate photometric redshifts (photo-z's) based on the SDSS *ugriz* photometry using the public code ZEBRA (Feldmann et al., 2006) run with the default templates, allowing interpolation between the standard templates without template optimization. To select a sample of early type background sources that should lie on the photoFP, we exclude the $\sim 2/3$ of the galaxies with best-fitting templates inconsistent with that of a passive stellar population. The sample selection for background sources will be described in greater detail in Paper II.

The SDSS photometric pipeline does not measure Sérsic index. Here, the SDSS petrosian concentration $C = R\,90/R\,50$, defined as the ratio of the radii containing 90% and 50% of the Petrosian flux (e.g., Shimasaku et al. 2001), is substituted for n. All reported quantities are measured in the r band.

A photoFP of the form

$$\log R_e = \alpha\mu + \beta \log C + \gamma, \tag{10.1}$$

is fit, where R_e is the half-light radius of the best-fit de Vaucouleurs light profile converted into physical units using the ZEBRA photo-z, μ is the mean de Vaucouleurs surface brightness within R_e, and α, β, and γ are free parameters. To avoid errors resulting from a redshift-dependent selection function, evolution in the photoFP, and K-corrections to the radii due to the fact that the morphological measurements are all made in the observed-frame r band, the galaxy sample is divided into redshift bins with width $\Delta z = 0.01$ and fit the photoFP separately in each bin. The best-fit coefficients are chosen to minimize the dispersion in effective radius at fixed μ and $\log C$, taking into account only the errors in R_e.

Figure 10.1 shows an edge-on view of the photoFP for our source sample. The dispersion around the photoFP in the direction of effective radius is 0.15 dex, or 35%.

10.1.2 Magnification using the photoFP

A line-of-sight matter overdensity at lens redshift z_l will produce an image convergence κ of amplitude:

$$\kappa = \frac{\Sigma(d_l\vec{\theta})}{\Sigma_{\rm crit}}, \tag{10.2}$$

where Σ is the projected surface density on the sky at z_l and $\Sigma_{\rm crit}$ is the characteristic surface density of matter required for lensing. $\Sigma_{\rm crit}$ is defined by the lensing geometry, such that

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{d_s}{d_l d_{ls}\,(1+z_l)^2}, \tag{10.3}$$

where d_l, d_s, and d_{ls} are the angular diameter distances from the observer to the lens, from the observer to the source, and from the lens to the source, respectively. The factor of $(1+z_l)^2$ arises from the use of comoving coordinates.

The lensing convergence re-scales the light profile, in the limit of very weak lensing, by a factor of $(1+\kappa)$. The radius and luminosity increase, but as the light profile is simply rescaled, the concentration is left unchanged. In the presence of the scaling relation described above, this implies an estimator $\hat{\kappa}$ of:

$$\begin{aligned}\log(1+\hat{\kappa}) &= \Delta \log R_e \\ &\equiv \log R_e - (\alpha\mu + \beta \log C + \gamma).\end{aligned} \tag{10.4}$$

If the errors in the observables are uncorrelated, the variance in our estimator $\hat{\kappa}$ is just the variance in the photoFP in the direction of R_e. A galaxy-galaxy lensing signal is extracted by cross-correlating this estimator with a population of foreground lenses.

10.2 A Magnification Measurement

10.2.1 Lens Sample

The lens sample is selected from the NYU Value-Added Catalog (Blanton et al., 2005) version of the SDSS Data Release 7 (DR7) spectroscopic survey (Abazajian et al., 2009), using only Luminous Red Galaxy Sample targets (LRGs, Eisenstein et al. 2003). The sample is limited to massive galaxies with absolute r-band magnitudes $-21.5 > M_{0.0_r} > -22.6$ and redshifts $0.15 < z < 0.35$. In order to compare with the results of Mandelbaum et al. (2008). The magnitudes are k-corrected and evolution corrected to $z = 0.0$ as in Mandelbaum et al. (2006c, hereafter M+06). Finally, to exclude satellite galaxies that are not at the centers of their dark matter haloes, galaxies with brighter nearby LRGs are removed, again following M+06 . This gives a sample of $\sim 55,000$ lenses that have comparable properties to the combined LRG sample of M+06.

10.2.2 Correcting Biases due to Photometric Redshift Errors

In the presence of photo-z errors, the overdensity of sources clustered near a lens will produce an excess of galaxies with incorrect photo-z (z_p) along the line of sight to the lens. As a result, averaging $\Delta \log R_e$ over the foreground or background source galaxies systematically mis-estimates the residuals from the plane associated with a lens due to the 'shadow' cast by photo-z errors.

This bias is dealt with by calculating the magnitude of this spurious signal directly from the data, and subtract it from our measured signal. It is important to first estimate the error in $\Delta \log R_e$ induced by a galaxy being assigned the wrong z_p ($\Delta \log R_e^{\rm err}$), and then to calculate what fraction f_l of the galaxies at each z_p have been scattered in from z_l. In these terms, the observed mean photoFP residual is:

$$\Delta \log R_e^{\rm obs} = (1 - f_l) \log (1 + \kappa) + f_l \Delta \log R_e^{\rm err}, \tag{10.5}$$

where κ is the true convergence.

$\Delta \log R_e^{\rm err}$ can be estimated by assuming that the galaxy lies on the photoFP at z_l but is incorrectly assigned to z_p. The inferred effective radius of a galaxy with true redshift z_l that is mistakenly assigned to z_p will be off by a factor of $d_s(z_p)/d_s(z_l)$. The surface brightness dimming correction will be similarly incorrect, with $\mu_p = \mu_l - 10 \log[(1+z_p)/(1+z_l)]$. Finally, the photoFP fits differ between redshift bins. A galaxy with an incorrect photo-z will therefore lie off the photoFP in its assigned redshift bin by

$$\Delta \log R_e^{\rm err} = \log \left(\frac{d_s(z_p)}{d_s(z_l)} \frac{R_e^p(\mu_p, C)}{R_e^l(\mu_l, C)} \right). \tag{10.6}$$

The expressions $R_e^p(\mu_p, C)$ and $R_e^l(\mu_l, C)$ are the radii that would be predicted by the photoFP for that galaxy's surface brightness and concentration in the bins corresponding to z_p and z_l, respectively.

The quantity f_l can be estimated by cross-correlating the positions of sources at z_p with lenses at z_l. The positions of galaxies in widely separated redshift bins are assumed to be uncorrelated and that any observed excess of sources far behind a lens is due to scattering from z_l. This means that

$$f_l = \frac{w_{il}(\theta)}{1 + w_{il}(\theta)}, \tag{10.7}$$

where $w_{il}(\theta)$ is the angular cross-correlation between the positions of sources at z_i and lenses at z_l. A cross-correlation signal of this form can also be produced by the boosted number counts of magnified background sources (e.g., Jain & Lima 2011) but that effect is too weak to detect with a lens sample of this size.

The cross-correlations for $0.20 < z_l < 0.25$ with a range of z_s bins are shown in figure 10.2. An angular correlation function of the form

$$1 + w_{il}(\theta) = \frac{A_{il}}{\theta^{0.8}} + B_{il} \tag{10.8}$$

is fit, where A_{il} and B_{il} are free parameters. The choice of power law index is motivated by the angular correlation function measurements of Wake et al. (2011), which are in agreement with the observed w_{ll}. Incorrectly estimating the true mean density of galaxies at z_p will cause B_{il} to deviate from unity, as is observed. The effects of this uncertainty are removed when calculating f_l by setting $B_{il} = 1$. Sources with $f_l > 0.20$ (above the black horizontal line in Figure 2) are excluded from the lensing measurement, while sources with $f_l < 0.20$ are corrected using equation 10.5.

In addition to the effects of galaxy clustering on photometric redshift errors, a mean offset between the true and photometric redshifts in a z_p bin will cause an incorrect estimation of the critical density $\Sigma_{\rm crit}$ for all of the galaxies in that bin. This error depends on the distribution of foreground lens redshifts. Using the method of Mandelbaum et al. (2008), the effect of a mean shift in the photo-z's on the signal is estimated to be no more than 10%. This uncertainty is small relative to the other corrections discussed here, so this calculation is deferred to subsequent work.

10.2.3 Sky Proximity Bias Correction

The SDSS photometric pipeline produces known sky subtraction proximity effects, where the photometry of objects near bright stars or galaxies is systematically biased (c.f. Aihara et al. 2011b). This may induce a systematic bias in the estimated radii, surface brightnesses, and concentrations that contaminates the lensing signal. Sky subtraction effects cannot distinguish between foreground and background galaxies (with respect to the bright lens), so this proximity bias can in principle be estimated from the photoFP residuals for galaxies in the *foreground* of the lenses, which are unaffected by lensing.

Figure 10.3 shows the average deviation from the photoFP as a function of source-lens angular separation for both foreground and background sources. The sky proximity bias systematically induces a reduction in effective radius relative to the photoFP trend. The lensing signal is thus the difference between the background and foreground photoFP deviations at each angular separation. Of note is the fact that the empirical sky correction extends beyond the size of the SDSS sky subtraction box, which is $\sim 100''$; this is a result of galaxy clustering. Each of the bright objects used as a lens will tend to be associated with a galaxy overdensity on the sky. This excess will also impact the sky correction, even in neighboring sky subtraction cells, so the angular scale of the resulting correction will be set by the galaxy correlation function.

The low redshift of the lens sample and the poor quality of the photo-z's (which preferentially scatter higher-z sources to lower z) result in a large fraction of source galaxies near lenses with photometric redshifts $z_p < z_l$ that are actually at $z > z_l$. This means that a foreground sample of sources with photo-z's will be contaminated by objects from higher z. The cut on f_L described above removes many such contaminating galaxies, at the cost of dramatically reducing the signal-to-noise ratio of the sky proximity bias estimate. This remains the major source of uncertainty in this measurement.

As a check against this effect, the deviation from the photoFP trend of those foreground

sources with spectroscopic redshifts is also shown. Any large bias to this sky subtraction estimate resulting from imperfect photo-z's should produce a substantial difference between the spectroscopic and photometric foreground estimates; this is not observed.

10.2.4 Halo Mass Profile

After controlling for the systematic errors described above, the line-of-sight surface matter density Σ is calculated by weighting each lens–background source pair by the critical density for lensing, $\Sigma_{\rm crit}\,(z_s, z_l)$. This density is binned by physical separation in the lens plane. The results are shown in figure 10.4, along with existing measurements from M+06 for a similar lens population.

10.3 Discussion: The Way Forward

The magnification signal demonstrated above, while many times stronger than previous magnification measurements, is still somewhat noisier than the shear signal for a comparable sample. This is because the convergence dispersion resulting from the measured photoFP width is 35% (1.8 times larger than the intrinsic shear dispersion of 20%) and because only the third of the source sample consistent with early-type SEDs has been used.

If the fundamental achievable limit for this technique is the intrinsic scatter in the *spectroscopic* fundamental plane, then the average magnification S/N for an early-type galaxy is the same as in shear; a comparable photometric Tully-Fisher relation for late-type galaxies would bring us to the point where magnification and shear provide comparable information. And any improvement in our understanding of galaxy evolution and dynamics that further diminishes the scatter in these scaling relations will boost the magnification signal beyond that available for shear measurement.

Perhaps just as valuable, magnification by this method is not sensitive to the same systematic biases that challenge upcoming shear measurements. For instance, the intrinsic galaxy alignment signal on large scales should not affect galaxy sizes, concentrations, and mean surface brightnesses in same the manner in which it affects shapes. We expect that this technique will also prove useful in extracting and removing instrumental sytematics, such as those arising from variations in the telescope point-spread function, and will investigate this prospect in a subsequent paper.

10.4 Shear using the Optical Tully-Fisher Relation

The low signal-to-noise of shear measurements has always necessitated the use of photometric data, as weak lensing measurements must average over many galaxy images in order to achieve a usably strong signal. This necessity introduces a set of major observational challenges. The statistical requirements entail measuring shapes for faint, poorly-resolved galaxies, for which the effects of errors in the telescope point-spread function model are large.

Precise, unbiased photometric redshifts, which are necessary to interpret a lensing signal, are very difficult to obtain.

In the shot-noise limit, the signal-to-noise in a lensing map will scale as $\sigma_e/\sqrt{n}$, where n is the number density of usable galaxy shape measurements on the sky. A deep imaging survey such as LSST [REF] may achieve a target density of around 40 usable galaxies per square arcminute.

Existing or planned large spectroscopic surveys such as SDSS-III/BOSS Schlegel et al. (2009) and BigBOSS Schlegel et al. (2011) achieve a target density on the sky in excess of many hundreds of targets per square degree; for present purposes, a modest fiducial value of 0.3 galaxies per square arcminute will serve as an achievable spectroscopic target density using existing technology.

The optical Tully-Fisher relation is a well-known scaling relation between the circular velocity of disc galaxies and their intrinsic luminosity (or stellar mass) Tully & Fisher (1977). It was originally presented as a tool for distance measurement, though it has in recent years been replaced in this by Type Ia supernovae.

In standard Tully-Fisher measurements, the inclination of the galaxy with respect to the line of sight, $\sin i$, is estimated from the ellipticity of the galaxy image. This is important, as the circular velocity actually measured in the spectroscopic rotation curve is $v_c \sin i$. Failing to correct for inclination angle scatters galaxies off of the Tully-Fisher in the direction of smaller v_{circ}. As i is uniformly distributed, this scatter is quite large compared with the intrinsic dispersion about the mean relation, as shown in figure 10.5.

If the true Tully-Fisher relation is known, however, then the amount by which a galaxy's circular velocity lies off of the relation is a very accurate estimator for $\sin i$, and thus its intrinsic unlensed ellipticity

$$\log_{10}\left(\sin\left(i\right)\right) = \log_{10}\left(v_{\mathrm{circ,\,obs}}\right) - \left[a + b\log_{10}\left(M_\star\right)\right] \tag{10.9}$$

The sine of the inclination angle is related to the observed axis ratio of the galaxy q as:

$$\sin\theta = \left(\frac{1-q^2}{1-q_z^2}\right)^{1/2} \tag{10.10}$$

where q_z is the ratio of the vertical and radial disk scale lengths; this is typically taken to be 0.19 Reyes et al. (2011b). The axis ratio and position angle in the image plane together determine the two ellipticity components that are normally used as a (noisy) estimator of the weak lensing shear. The scatter off of the Tully-Fisher relation due to uncorrected line-of-sight inclination is much larger than the intrinsic scatter in the relation itself. This means that the intrinsic ellipticity of the galaxy can be predicted to high precision using only the stellar mass and the rotation curve. Lensing will distort the galaxy ellipticity without modifying the rotation curve. Comparing the observed galaxy ellipticty to that estimated from the circular velocity offset from the mean Tully-Fisher relation can dramatically reduce the galaxy shape noise.

Figure 10.5 shows a simple simulation of the achievable reduction in shape noise which adopts the Tully-Fisher fit parameters and measured scatter from Reyes et al. (2011b) and assumes a uniform distribution in i. Each galaxy's intrinsic axis ratio is inferred from its distance off of the mean TF relation, using equation 10.10. The relation between the true and estimated ellipticities is shown in the right panel.

Estimating the inclination angle as above yields a correct estimate of the galaxy's shape to within 8%. Notably, the axis ratios of galaxies with small inclinations (i.e., those that are more nearly face-on) have a much smaller dispersion – only 4% for the less inclined half of the sample. For comparison, the dispersion in shapes for typical lensing surveys is around 35%.

This reduction in the shape noise would allow a spectroscopic weak lensing survey with signal strength approximately as great as that available from LSST, while avoiding the challenging systematic errors (e.g., from photometric redshift biases) associated with purely photometric lensing measurements. This should be possible with a target density comparable to that achieved by current spectroscopic galaxy surveys such as SDSS-III/BOSS or BigBOSS, at significantly lesser cost, and without any of the systmatic errors associated with photometric redshifts. It should be noted that this technique is beyond the reach of the aforementioned spectroscopic surveys. Boss and BigBOSS use or are planning to use fibers, so they cannot produce rotation curves.

A practical pilot study is necessary in order to understand whether such a measurement is feasible. The most straightforward approach is to use a wide-field multi-slit spectrograph to obtain rotation curves for a sufficiently dense sample of disk galaxies behind a massive low-redshift cluster with existing shear maps. The weak lensing maps derived from the Tully-Fisher shape measurements can then be compared with conventional lensing maps for the same field.

A successful demonstration of Tully-Fisher spectroscopic weak lensing would be very timely, as there are several large spectroscopic surveys (PFS, BigBOSS) planned for the near future.

10.5 Discussion

Much additional work will be required before either photometric or spectroscopic galaxy scaling relations can be used for high-precision cosmology. What is not in doubt, however, is the existence of a substantial amount of unexploited weak lensing signal. The systematic errors will be easier to model as the available signal grows; what is more difficult to account for, but much more valuable, is that the scope for asking new questions will grow as well.

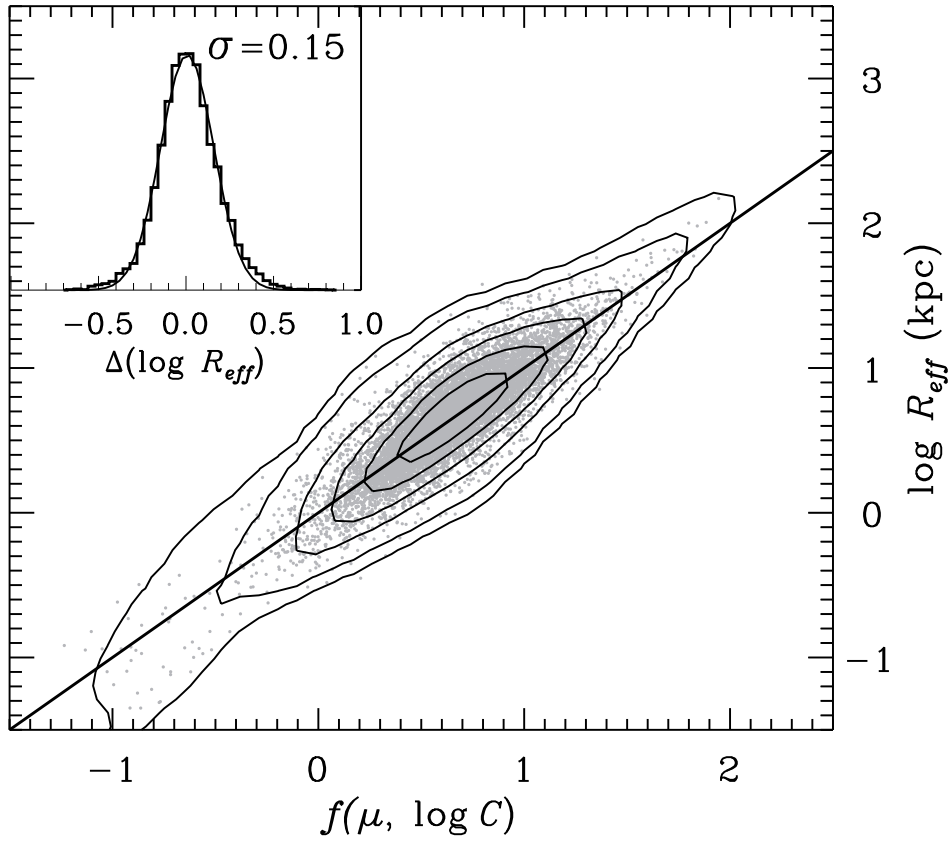

Figure 10.1 The photometric fundamental plane for the source sample of 8.4 million galaxies, shown edge-on. $\log R_e$ is fit as a function of effective surface brightness (μ) and concentration ($\log C$) separately in redshift bins of width $\Delta z = 0.01$. Gray points show a random subset of 100,000 galaxies from the source catalog, while the solid line shows the one-to-one relation. Contours enclose the 0.5σ, 1σ, 1.5σ, 2σ, 2.5σ, and 3σ boundaries of the 2D distribution for the full source catalog. The inset shows the distribution of residuals in $\log R_e$ from the photoFP fits, which has width $\sigma = 0.153$ dex.

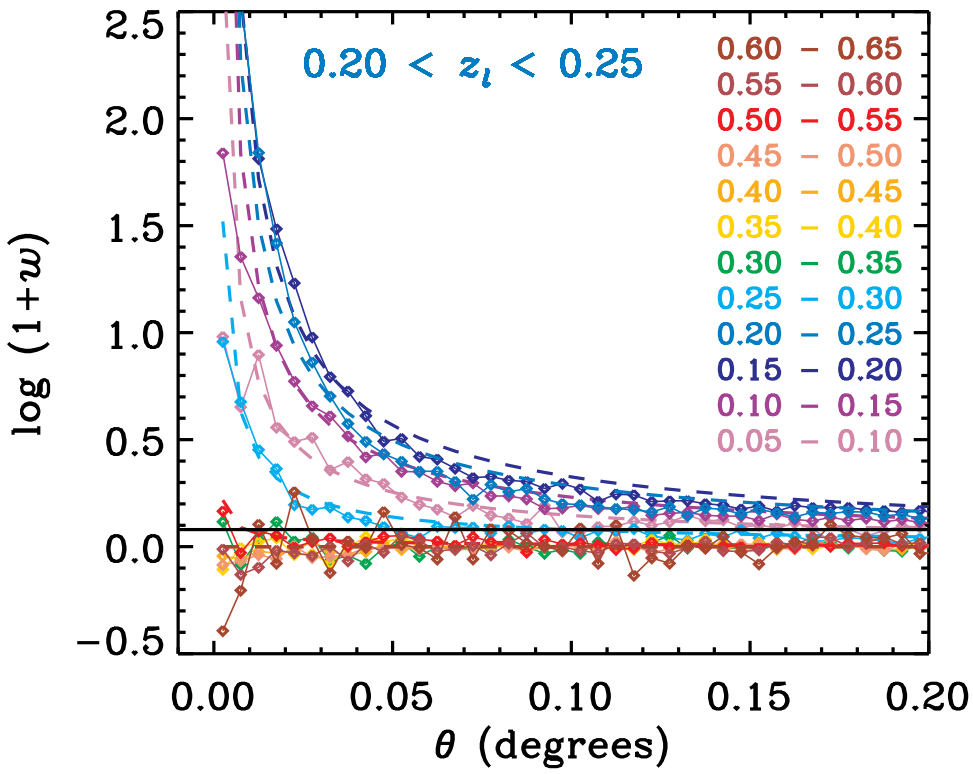

Figure 10.2 The projected correlation function for sources around lenses with $0.20 < z_l < 0.25$. Greyscale (colors in online version) indicate different bins in z_s. Solid lines show fits to the data. At small separations, a large fraction of the galaxies in nearby z bins are likely scattered in from z_l through photo-z errors. See text for details.

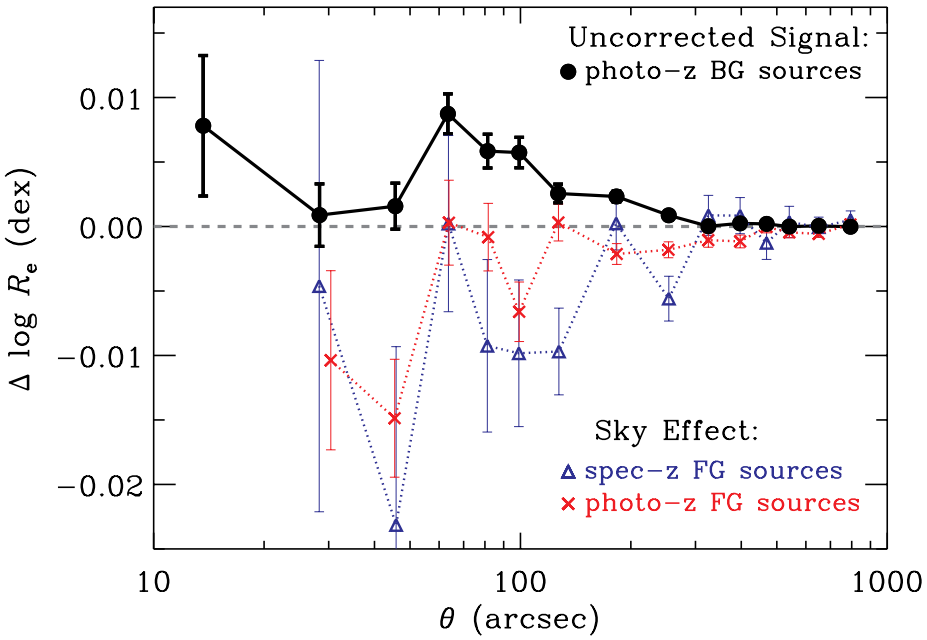

Figure 10.3 The raw magnification signal around the galaxy lenses (filled black circles) compared with the sky proximity bias measured from foreground sources. The red crosses show the estimated sky subtraction effect using sources with photo-z's; the blue triangles show the same estimate, but using those foreground galaxies with spectroscopic redshifts.

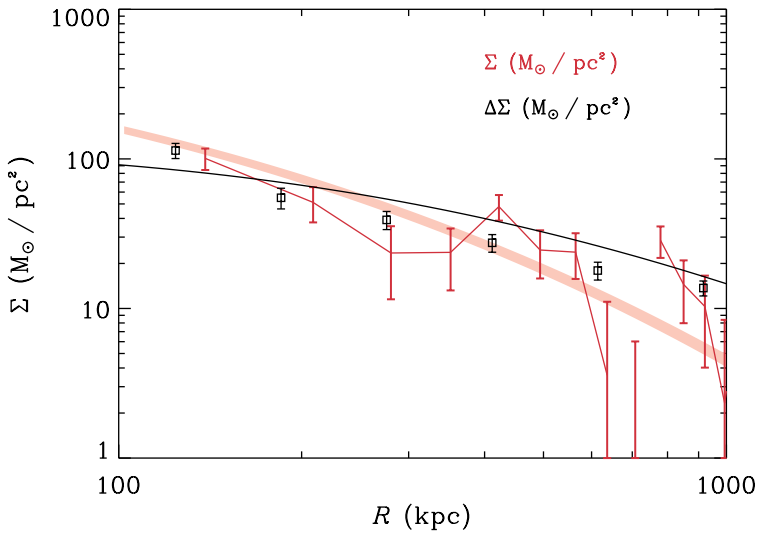

Figure 10.4 Solid line with error bars (red in online version): Σ from this work. Open squares with error bars: $\Delta\Sigma$ from M+06 measured using shear. That measurement used a smaller lens sample than considered here, so those error bars have been reduced in order to allow for a fair comparison of the statistical power of the two samples. The solid line is the best-fit $\Delta\Sigma$ profile from M+06. The shaded region (red in the online version) shows the corresponding Σ profile (with 68% confidence interval) derived from the M+06 data.

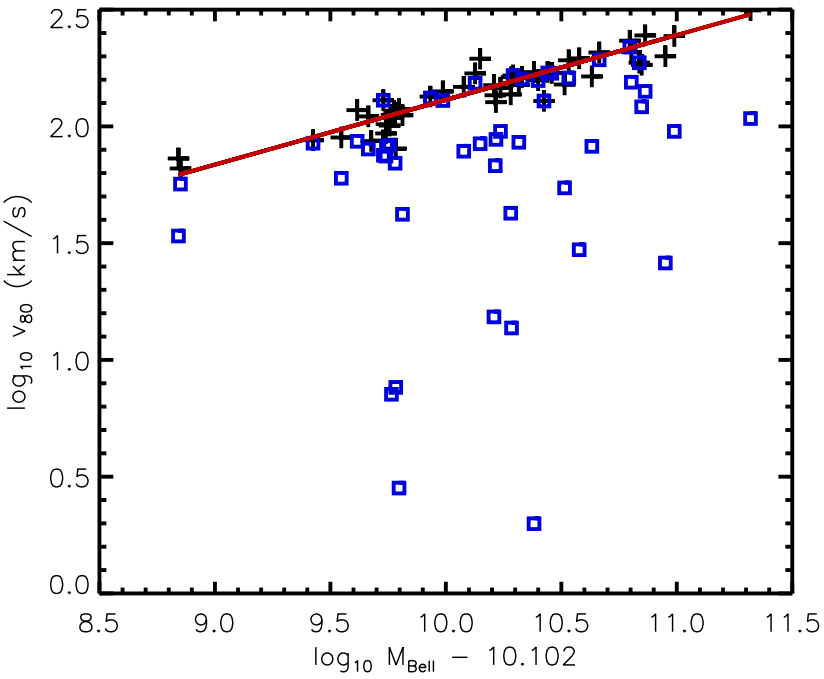

Figure 10.5 The Tully-Fisher relation, using the best-fit trend (red line) and scatter reported in Reyes et al. (2011b). Blue squares show the positions of galaxies before inclination correction, and black crosses show the typical results from standard inclination corrections. The difference between these two clusters of points is the magnitude of the signal. Errors in the inclination estimate generate part of the known scatter around the trend.

Bibliography

Abazajian, K., et al. 2003, AJ, 126, 2081
—. 2004, AJ, 128, 502
—. 2005, AJ, 129, 1755
Abazajian, K. N., et al. 2009, ApJS, 182, 543
Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38
—. 2007, ApJS, 172, 634
—. 2008, ApJS, 175, 297
Aihara, H., et al. 2011a, ApJS, 193, 29
—. 2011b, ApJS, 193, 29
Albrecht, A., et al. 2006, ArXiv Astrophysics e-prints
—. 2009, ArXiv e-prints
Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625
Benjamin, J., et al. 2007, MNRAS, 381, 702
Bergé, J., Price, S., Amara, A., & Rhodes, J. 2012, MNRAS, 419, 2356
Bernardi, M., et al. 2003, AJ, 125, 1866
Bernstein, G. M. 2010, MNRAS, 406, 2793
Bernstein, G. M., & Jarvis, M. 2002, The Astronomical Journal, 123, 583
Bertin, G., & Lombardi, M. 2006, ApJ, 648, L17
Blanton, M. R., et al. 2005, AJ, 129, 2562
Blazek, J., McQuinn, M., & Seljak, U. 2011, J. Cosmology Astropart. Phys., 5, 10
Brown, M. L., Taylor, A. N., Bacon, D. J., Gray, M. E., Dye, S., Meisenheimer, K., & Wolf, C. 2003, MNRAS, 341, 100
Casali, M., et al. 2007, A&A, 467, 777
Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, MNRAS, 320, L7
Coil, A. L., et al. 2011, ApJ, 741, 8
Conroy, C., & Wechsler, R. H. 2009, ApJ, 696, 620
Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2001, ApJ, 559, 552
—. 2002, ApJ, 568, 20
Croft, R. A. C., & Metzler, C. A. 2000, ApJ, 545, 561
Cunha, C. E., Lima, M., Oyaizu, H., Frieman, J., & Lin, H. 2009, MNRAS, 396, 2379
Davis, M., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4834, Society of Photo-Optical Instrumentation Engineers (SPIE) Con-

ference Series, ed. P. Guhathakurta, 161–172
Eifler, T. 2011, MNRAS, 418, 536
Eifler, T., Kilbinger, M., & Schneider, P. 2008, A&A, 482, 9
Eifler, T., Schneider, P., & Hartlap, J. 2009, A&A, 502, 721
Eifler, T., Schneider, P., & Krause, E. 2010, A&A, 510, A7
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Eisenstein, D. J., et al. 2003, ApJ, 585, 694
Feldmann, R., et al. 2006, MNRAS, 372, 565
Frieman, J. A., et al. 2008, AJ, 135, 338
Fu, L., & Kilbinger, M. 2010, MNRAS, 401, 1264
Fu, L., et al. 2008, A&A, 479, 9
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
Graham, A. W. 2002, MNRAS, 334, 859
Gunn, J. E., et al. 1998, AJ, 116, 3040
—. 2006, AJ, 131, 2332
Hambly, N. C., et al. 2008, MNRAS, 384, 637
Heavens, A., Refregier, A., & Heymans, C. 2000, MNRAS, 319, 649
Hewett, P. C., Warren, S. J., Leggett, S. K., & Hodgkin, S. T. 2006, MNRAS, 367, 454
Heymans, C., et al. 2006, MNRAS, 371, L60
Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31
Hirata, C. M., Mandelbaum, R., Ishak, M., Seljak, U., Nichol, R., Pimbblet, K. A., Ross, N. P., & Wake, D. 2007, MNRAS, 381, 1197
Hirata, C. M., Padmanabhan, N., Seljak, U., Schlegel, D., & Brinkmann, J. 2004a, Phys. Rev. D, 70, 103501
Hirata, C. M., & Seljak, U. 2003, Mon.Not.Roy.Astron.Soc., 343, 459
Hirata, C. M., & Seljak, U. 2004, Phys. Rev. D, 70, 063526
Hirata, C. M., et al. 2004b, MNRAS, 353, 529
Hodgkin, S. T., Irwin, M. J., Hewett, P. C., & Warren, S. J. 2009, MNRAS, 394, 675
Hoekstra, H., Hsieh, B. C., Yee, H. K. C., Lin, H., & Gladders, M. D. 2005, ApJ, 635, 73
Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, ApJ, 606, 67
Hoekstra, H., Yee, H. K. C., Gladders, M. D., Barrientos, L. F., Hall, P. B., & Infante, L. 2002, ApJ, 572, 55
Hoekstra, H., et al. 2006, ApJ, 647, 116
Høg, E., et al. 2000, A&A, 355, L27
Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129
Hopkins, P. F., Bahcall, N. A., & Bode, P. 2005, ApJ, 618, 1
Ilbert, O., et al. 2009, ApJ, 690, 1236
Ivezić, Ž., et al. 2004, Astronomische Nachrichten, 325, 583
Jain, B., & Lima, M. 2011, MNRAS, 411, 2113
Jain, B., Seljak, U., & White, S. 2000, ApJ, 530, 547
Jarvis, M., Bernstein, G. M., Fischer, P., Smith, D., Jain, B., Tyson, J. A., & Wittman, D.

2003, AJ, 125, 1014
Jing, Y. P. 2002, MNRAS, 335, L89
Joachimi, B., Mandelbaum, R., Abdalla, F. B., & Bridle, S. L. 2011, A&A, 527, A26
Jørgensen, I., Franx, M., & Kjaergaard, P. 1996, MNRAS, 280, 167
Jouvel, S., et al. 2009, A&A, 504, 359
Kaiser, N. 2000, ApJ, 537, 555
Kaiser, N., Wilson, G., & Luppino, G. A. 2000, ArXiv Astrophysics e-prints
Kilbinger, M., Schneider, P., & Eifler, T. 2006, A&A, 457, 15
Kitching, T., et al. 2010
Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2003, in The 2002 HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System, Proceedings of a Workshop held at the Space Telescope Science Institute, Baltimore, Maryland, October 17 and 18, 2002. Edited by Santiago Arribas, Anton Koekemoer, and Brad Whitmore. Baltimore, MD: Space Telescope Science Institute, 2003., p.337, ed. S. Arribas, A. Koekemoer, & B. Whitmore, 337
Koekemoer, A. M., et al. 2007, ApJS, 172, 196
Komatsu, E., et al. 2011, ApJS, 192, 18
Kovač, K., et al. 2010, ApJ, 708, 505
Krause, E., & Hirata, C. M. 2010, A&A, 523, A28
Kristian, J. 1967, ApJ, 147, 864
Kuijken, K. 1999, A&A, 352, 355
Larson, D., et al. 2011, ApJS, 192, 16
Lawrence, A., et al. 2007, MNRAS, 379, 1599
Lawrence, E., Heitmann, K., White, M., Higdon, D., Wagner, C., Habib, S., & Williams, B. 2010, ApJ, 713, 1322
Le Fèvre, O., et al. 2005, A&A, 439, 877
Leauthaud, A., et al. 2007, ApJS, 172, 219
—. 2012, ApJ, 744, 159
Lima, M., Cunha, C. E., Oyaizu, H., Frieman, J., Lin, H., & Sheldon, E. S. 2008, MNRAS, 390, 118
Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., Kent, S., & Yasuda, N. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne, 269
Madgwick, D. S., et al. 2003, ApJ, 599, 997
Mandelbaum, R., Hirata, C. M., Broderick, T., Seljak, U., & Brinkmann, J. 2006a, MNRAS, 370, 1008
Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006b, MNRAS, 367, 611
Mandelbaum, R., Hirata, C. M., Leauthaud, A., Massey, R. J., & Rhodes, J. 2012, MNRAS, 420, 1518
Mandelbaum, R., Li, C., Kauffmann, G., & White, S. D. M. 2009, MNRAS, 393, 377
Mandelbaum, R., Seljak, U., Cool, R. J., Blanton, M., Hirata, C. M., & Brinkmann, J.

2006c, MNRAS, 372, 758
Mandelbaum, R., Seljak, U., & Hirata, C. M. 2008, JCAP, 8, 6
Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, Mon.Not.Roy.Astron.Soc., 368, 715
Mandelbaum, R., et al. 2005, MNRAS, 361, 1287
—. 2008, MNRAS, 386, 781
—. 2011, MNRAS, 410, 844
Massey, R., Rowe, B., Refregier, A., Bacon, D. J., & Bergé, J. 2007a, MNRAS, 380, 229
Massey, R., Stoughton, C., Leauthaud, A., Rhodes, J., Koekemoer, A., Ellis, R., & Shaghoulian, E. 2010, MNRAS, 401, 371
Massey, R., et al. 2007b, ApJS, 172, 239
Miyazaki, S., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113
Nakajima, R., Mandelbaum, R., Seljak, U., Cohn, J. D., Reyes, R., & Cool, R. 2012, MNRAS, 420, 3240
Okumura, T., Jing, Y. P., & Li, C. 2009, ApJ, 694, 214
Padmanabhan, N., et al. 2008, ApJ, 674, 1217
Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, AJ, 125, 1559
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in C. The art of scientific computing
Refregier, A. 2003, MNRAS, 338, 35
Refregier, A., Amara, A., Kitching, T. D., & Rassat, A. 2011, A&A, 528, A33
Refregier, A., & Bacon, D. 2003, MNRAS, 338, 48
Reyes, R., Mandelbaum, R., Gunn, J. E., Nakajima, R., Seljak, U., & Hirata, C. M. 2011a, ArXiv e-prints
Reyes, R., Mandelbaum, R., Gunn, J. E., Pizagno, J., & Lackner, C. N. 2011b, MNRAS, 417, 2347
Reyes, R., Mandelbaum, R., Seljak, U., Baldauf, T., Gunn, J. E., Lombriser, L., & Smith, R. E. 2010, Nature, 464, 256
Rhodes, J., Refregier, A., & Groth, E. J. 2001, ApJ, 552, L85
Rhodes, J. D., et al. 2007, ApJS, 172, 203
Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, ApJ, 672, 19
Schlegel, D., White, M., & Eisenstein, D. 2009, in ArXiv Astrophysics e-prints, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey, 314
Schlegel, D., et al. 2011, ArXiv e-prints
Schmidt, F., Leauthaud, A., Massey, R., Rhodes, J., George, M. R., Koekemoer, A. M., Finoguenov, A., & Tanaka, M. 2012, ApJ, 744, L22
Schneider, P., Eifler, T., & Krause, E. 2010, A&A, 520, A116
Schneider, P., & Kilbinger, M. 2007, A&A, 462, 841

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002a, A&A, 396, 1
Schneider, P., van Waerbeke, L., & Mellier, Y. 2002b, A&A, 389, 729
Schrabback, T., et al. 2010, A&A, 516, A63
Scoville, N., et al. 2007a, ApJS, 172, 38
—. 2007b, ApJS, 172, 1
Semboloni, E., Schrabback, T., van Waerbeke, L., Vafaei, S., Hartlap, J., & Hilbert, S. 2011, MNRAS, 410, 143
Semboloni, E., et al. 2006, A&A, 452, 51
Sheldon, E. S., Cunha, C., Mandelbaum, R., Brinkmann, J., & Weaver, B. A. 2011, ArXiv e-prints
Shimasaku, K., et al. 2001, AJ, 122, 1238
Simha, V., Weinberg, D., Dave, R., Fardal, M., Katz, N., & Oppenheimer, B. D. 2010, ArXiv e-prints
Smith, J. A., et al. 2002, AJ, 123, 2121
Smith, R. E., et al. 2003, MNRAS, 341, 1311
Stoughton, C., et al. 2002, AJ, 123, 485
Tucker, D. L., et al. 2006, Astronomische Nachrichten, 327, 821
Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661
van Waerbeke, L., White, M., Hoekstra, H., & Heymans, C. 2006, Astroparticle Physics, 26, 91
Van Waerbeke, L., et al. 2000, A&A, 358, 30
Wake, D. A., et al. 2011, ApJ, 728, 46
Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143
York, D. G., et al. 2000, AJ, 120, 1579
Zacharias, N., et al. 2000, AJ, 120, 2131
Zhang, J., & Komatsu, E. 2011, MNRAS, 414, 1047
Zhang, P., Liguori, M., Bean, R., & Dodelson, S. 2007, Physical Review Letters, 99, 141302