Biased shear estimate due to minimal apparent size cut

(Dated: August 10, 2008)

Some first studies on the size bias.

I. SIZE BIAS

In the following, I will use a very naive approach of measuring the shear; it might be wrong to use such an oversimplified ansatz, hence this should be checked... Furthermore, I will use shear and convergence interchangeably. Basically, the ansatz says that we wish to measure the average convergence in some solid angle bin on the sky. The shear κ in this bin follows a distribution $f(\kappa)$, so that the average shear in the solid angle bin is given by:

$$\langle\kappa\rangle = \int_{-\infty}^{\infty} d\kappa\, \kappa\, f(\kappa) \tag{1}$$

We assume that somehow an estimate of $\langle\kappa\rangle$ is calculated by observing several galaxies in that patch of sky, estimating κ for each one, and thus sampling from the κ distribution. If we sample fairly from this distribution, our estimate of $\langle\kappa\rangle$ should be unbiased. However, a cut in apparent size results in a distortion of this distribution.

For the shear to be measurable, we assume that the observed apparent angular size of the galaxy, s, has to be greater than some minimal value s_{min}. However, the *observed* size of the galaxy is not the actual one, since lensing magnifies objects:

$$s_{\mathrm{obs}} = s_0\sqrt{\mu} = \frac{s_0}{\sqrt{(1-\kappa^2)+|\gamma|^2}} = s_0[1+\kappa+O(\kappa^2,\gamma^2,\kappa\gamma)], \tag{2}$$

where s_0 is the intrinsic apparent angular size of the object and μ the magnification, i.e. the ratio of apparent areas of the lensed and unlensed objects. Note that if the cut is not on overall "size" but, e.g., on area, the coefficient of the linear κ term will change. For a pure number count, this effect leads to a bias analogous to the magnification bias for flux-limited surveys:

$$N_{\mathrm{obs}} = N_0(1+\frac{dN}{d\ln s}|_{s_{\mathrm{min}}}\kappa). \tag{3}$$

However, the number count of galaxies in a bin does *not* influence the measured average shear (of course, if that is oversimplified, and there is an effect of the number count in a bin on the shear measurement, this should be studied as well). Size bias enters the shear measurement in a more subtle way (assuming the derivation here is correct).

Consider a galaxy with an intrinsic size s_0 and a convergence κ in the direction of it. In order to be counted in the $\langle\kappa\rangle$ measurement, it has to satisfy:

$$s_0(1+\kappa) > s_{\mathrm{min}} \quad \Rightarrow \quad s_0 > \frac{s_{\mathrm{min}}}{1+\kappa}. \tag{4}$$

Let $f_0(\kappa)$ denote the distribution of κ in the absence of size bias, i.e. if we were able to cut on the actual intrinsic size of galaxies. Since intrinsic size and shear are uncorrelated, $f_0(\kappa)$ should be equal to the actual shear distribution. Assuming this, the observed shear distribution $f_{\mathrm{obs}}(\kappa)$ is given by:

$$f_{\mathrm{obs}}(\kappa) = f_0(\kappa) + \int_{s_{\mathrm{min}}/(1+\kappa)}^{s_{\mathrm{min}}} ds\, \frac{dN}{ds} f_0(\kappa) \tag{5}$$

$$\approx f_0(\kappa)\left(1 - \frac{dN}{d\ln s}|_{s_{\mathrm{min}}}\kappa\right), \tag{6}$$

where dN/ds is the distribution of galaxy apparent sizes normalized to unity, $dN/d\ln s = s\, dN/ds$, and the second line is the result to lowest order in κ, consistent with equation (2). This is the crucial point, and I hope it makes sense. The fraction $f(\kappa)$ of galaxies between κ and $\kappa + d\kappa$ is enhanced by the fraction of galaxies below s_{min} which can be raised above s_{min} by the magnification κ (or vice versa for $\kappa < 0$). Basically, at high convergences κ, we preferentially pick lensed galaxies close to the size threshold, and hence bias the distribution towards larger κ (assuming $dN/d\ln s$ is negative, larger galaxies being rarer). The bias in the estimated shear, again to lowest order in κ, is then given by:

$$\langle\kappa\rangle = \int_{-\infty}^{\infty} d\kappa\, \kappa\, f_{\mathrm{obs}}(\kappa) = \langle\kappa\rangle_0 - \frac{dN}{d\ln s}|_{s_{\mathrm{min}}} \int_{-\infty}^{\infty} d\kappa\, \kappa^2\, f_0(\kappa) \tag{7}$$

$$= \langle\kappa\rangle_0 - \frac{dN}{d\ln s}|_{s_{\mathrm{min}}} \left(\langle\kappa^2\rangle_0 + \langle\kappa\rangle_0^2\right) \tag{8}$$

Here $\langle \cdot \rangle_0$ denotes averages according to the undistorted distribution $f_0(\kappa)$. Some interesting features are worth noting: the effect depends on the distribution of κ, not only on its average value (like e.g. for the magnification bias). It is quadratic in κ, but not simply in $\langle \kappa \rangle_0$, so that a nonzero scatter in κ will lead to a bias even for very small $\langle \kappa \rangle_0$. In fact, even if there was zero average shear in the solid angle bin considered, size bias could lead to a nonzero observed shear. Also, the effect is always positive (for a negative size count slope), whether $\langle \kappa \rangle_0$ is positive or negative. The *relative* size bias effect is then:

$$\frac{\langle \kappa \rangle - \langle \kappa \rangle_0}{\langle \kappa \rangle_0} = -\frac{dN}{d \ln s}|_{s_{\min}} \left(\frac{\langle \kappa^2 \rangle_0}{\langle \kappa \rangle_0} + \langle \kappa \rangle_0 \right) \tag{9}$$

Assuming $dN/d \ln s$ is not terribly small, the effect could thus well be above the percent level, if there is significant scatter in κ. Large scatter in κ means even very small (and abundant) galaxies can be magnified over the threshold size.

II. NEXT STEPS

If we have convinced ourselves that the derivation above is correct, the next steps toward an actual calculation of the effect would be:

- What is the distribution of κ in a solid angle bin ? (first step: Gaussian)
- What is $dN/d \ln s$? This also depends on the redshift bin considered, since $s = L/d_A(z)$, where L is the physical size of the galaxy.
- This leads to another interesting question: both κ and s depend on redshift, so there *is* a correlation between apparent size and κ. Will that change the simple ansatz above ? To zeroth order, this effect is calibrated out by using the observed redshift distribution of galaxies in the shear calculation.
- Generalize to smooth rather than sharp size cutoffs. Is the observational cut actually on size or on minor axis (which should result in the same effect, to first order) ? Or in area ($\leftrightarrow$ photon count) ?

Another issue which might be worth investigating: other systematic lensing effects could be linked to the apparent size; these will then also influence the measured κ via size bias.