Weak lensing shear estimation and systematics

Rachel Mandelbaum

Dark Energy School, 2/13/17

Outline

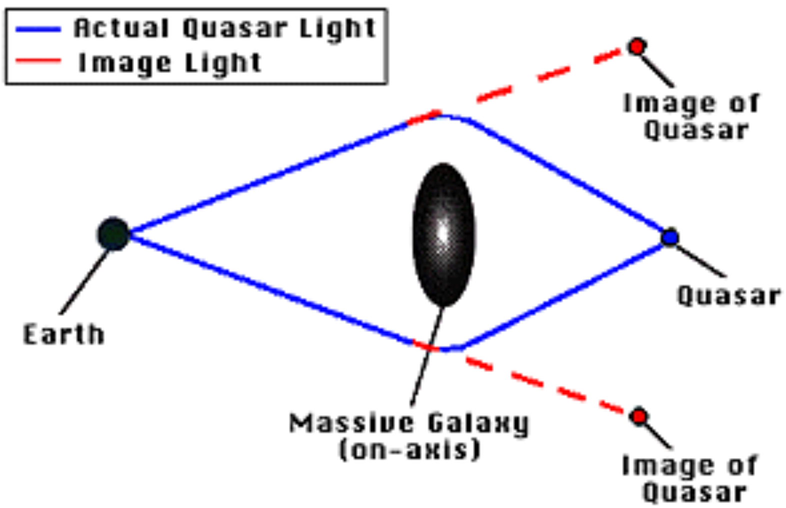
Weak lensing background, followed by 4 short sections:

1. What is the meaning of “galaxy shape”?
2. How do weak lensers measure shapes?
3. One source of systematics: PSF model errors
4. Another source of systematics: selection bias.

Weak lensing background

Gravitational lensing

Deflection of light by all gravitational mass, including dark matter!



Weak gravitational lensing

Depends on:

- Projected mass
- Separation on sky
- Separation between us, lens, source

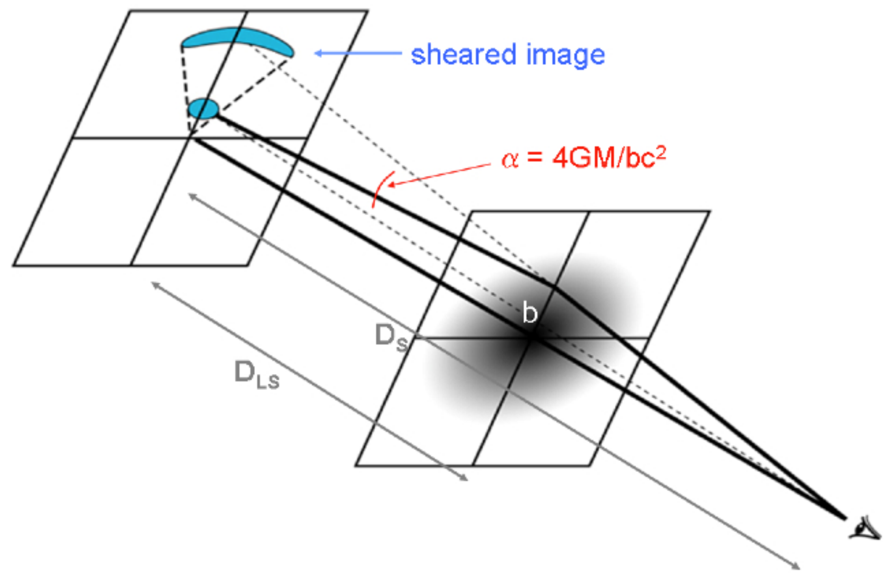
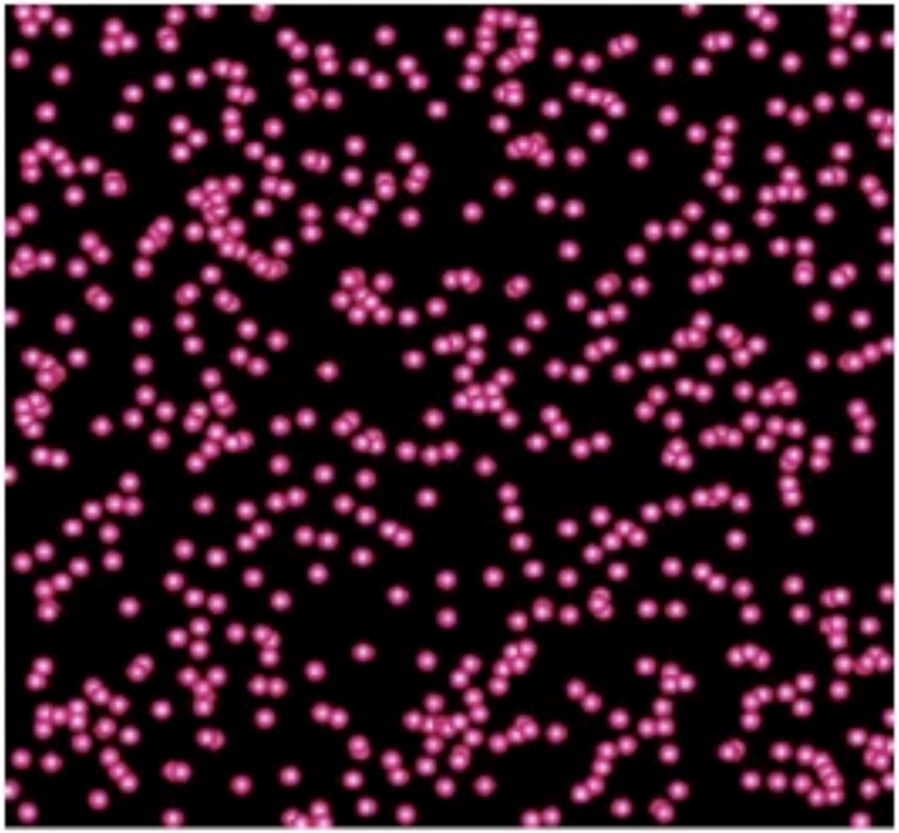
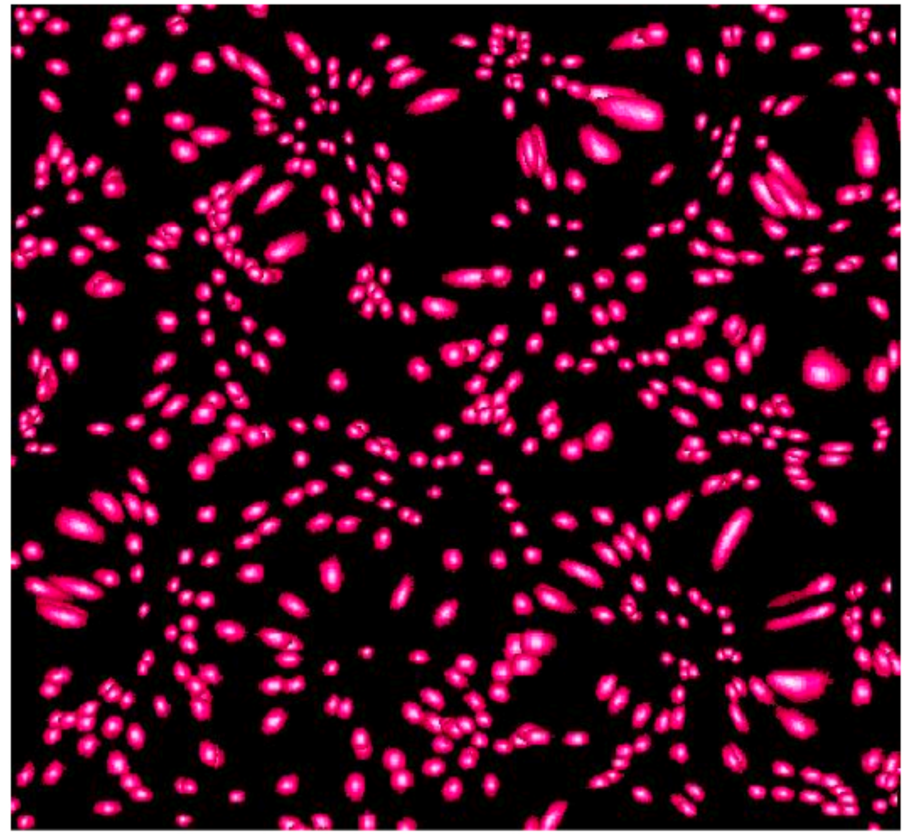


Image credit:
The LSST science book

Weak lensing



Unlensed



Lensed

Galaxies aren’t round

We measure lensing statistically: look for coherent galaxy shape distortions (<~1%) underneath the ~30%-level ellipticities

Image credits: NASA, ESA, S. Beckwith (STScI), the HUDF Team

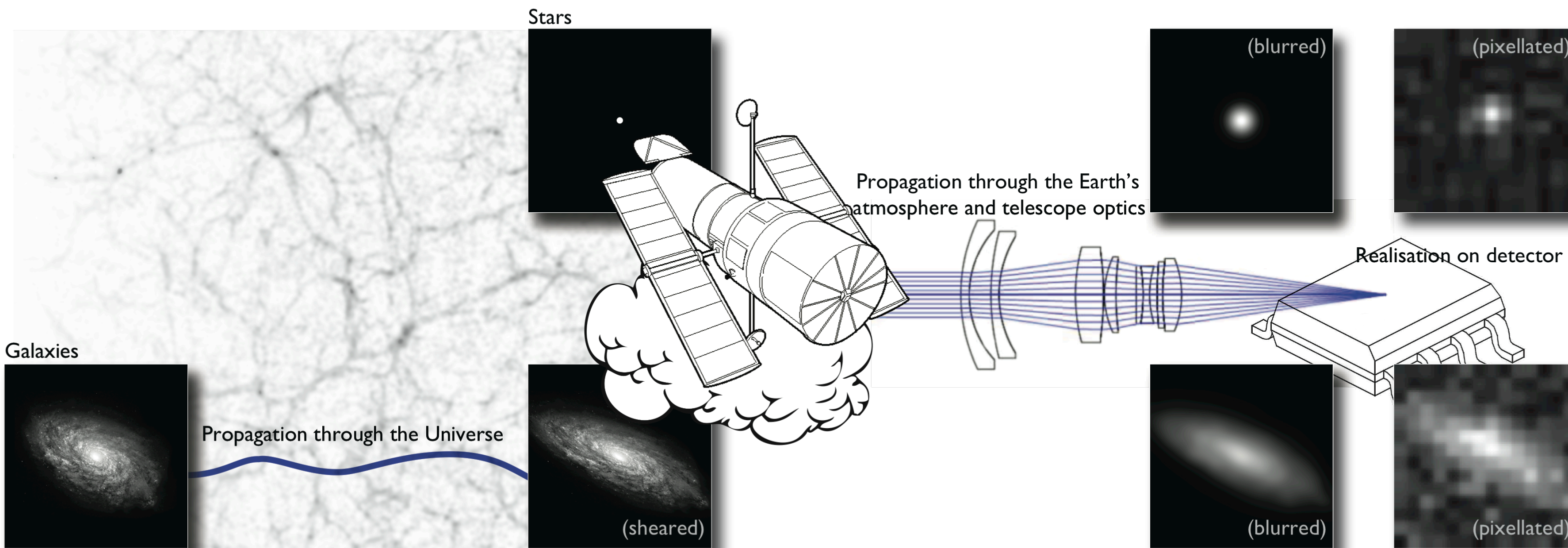
Ansatz

Assume galaxy shapes have a random intrinsic component e_{int}. If measured $e \simeq e_{int} + g$, then the shape correlation is:

$$
\begin{aligned}
\langle e(x)\, e(x+\theta)\rangle = {} & \langle e_{int}(x)\, e_{int}(x+\theta)\rangle \\
& + \langle e_{int}(x)\, g(x+\theta)\rangle + \langle g(x)\, e_{int}(x+\theta)\rangle \\
& + \boxed{\langle g(x)\, g(x+\theta)\rangle}
\end{aligned}
$$

For isotropically oriented galaxy shapes that do not correlate with the shear, only the last (cosmological) term contributes.

Other stuff happens to galaxy images besides shear

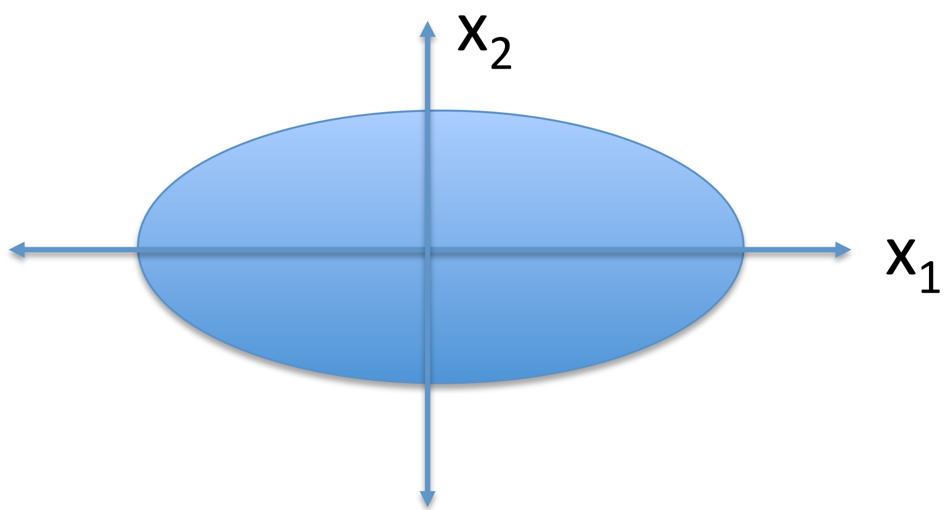


Need to disentangle shear from the PSF convolution, detector effects – for galaxies with finite resolution and SNR

The meaning of “galaxy shapes”

How do you define a galaxy shape?

- Let's start with something really simple:
  - No PSF (and no pixel response function)
  - No detector effects
  - Infinite SNR
  - Galaxy shape $I(x_1,x_2)$ just looks like an ellipse



Consider the moments

- Zeroth moment: flux (1 number)
  - $M_0 = \int I(x_1,x_2)\, dx_1 dx_2$
- 1st moment: centroid (2 numbers)
  - $M_i = \int I(x_1,x_2)\, x_i\, dx_1 dx_2 / M_0$
- 2nd moment: shape and size (3 numbers)
  - $M_{ij} = \int I(x_1,x_2)\, (x_i-M_i)(x_j-M_j)\, dx_1 dx_2 / M_0$
  - Circular gaussian: $M_{11}=M_{22}=\sigma^2$, $M_{12}=0$

Now add noise…

- Consider moments estimation as summation over $I(x_1,x_2) + N(x_1,x_2)$ where $N(x_1,x_2)$ values are random deviates drawn from Gaussians with some standard deviation
- Result: moments diverge due to noise!
- *Solution?*

Weighted moments

- Choose a weight function that is somewhat matched to the galaxy light profile
  - Put W(x,y) into all integrals defining the moments
- How to choose weight function? Some examples:
  - Use crude size estimate for circular Gaussian
  - Use iterative “adaptive moments”: effectively an elliptical Gaussian fit

Now consider this (perfectly reasonable) galaxy

What does the shape/size defined by the moments mean? How does it depend on the weight function?
How will it depend on the SNR?

Shape is not a well-defined property for realistic galaxies!

- But does this problem matter?
- Weak lensers use shapes as a means for getting at the shear
- As long as the ensemble statistics of the shape (however we arbitrarily define it) has a well-defined response to shear, it’s OK
- One consequence: should never compare shape measurements from two estimators – only compare ensemble shears!

Methods for estimating galaxy shapes

References for different methods

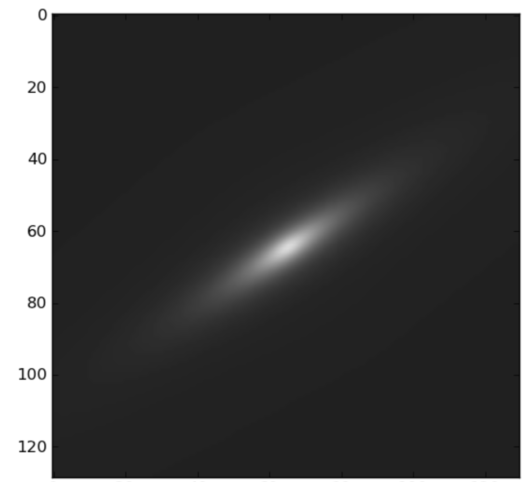
- Mandelbaum+ (2015) – GREAT3 results
  - Moments-based (8)
  - Model fitting (7): mostly maximum-likelihood (e.g., im3shape), one partly Bayesian (LensFit)
  - Other: 3 ML-based methods
- Different philosophies:
  - Inferring shear without per-galaxy shape estimates: Bernstein & Armstrong (2014)
  - Hierarchical inference: Schneider+ (2015)
  - Meta-calibration (direct calibration on images): Huff & RM (2017), Sheldon & Huff (2017)

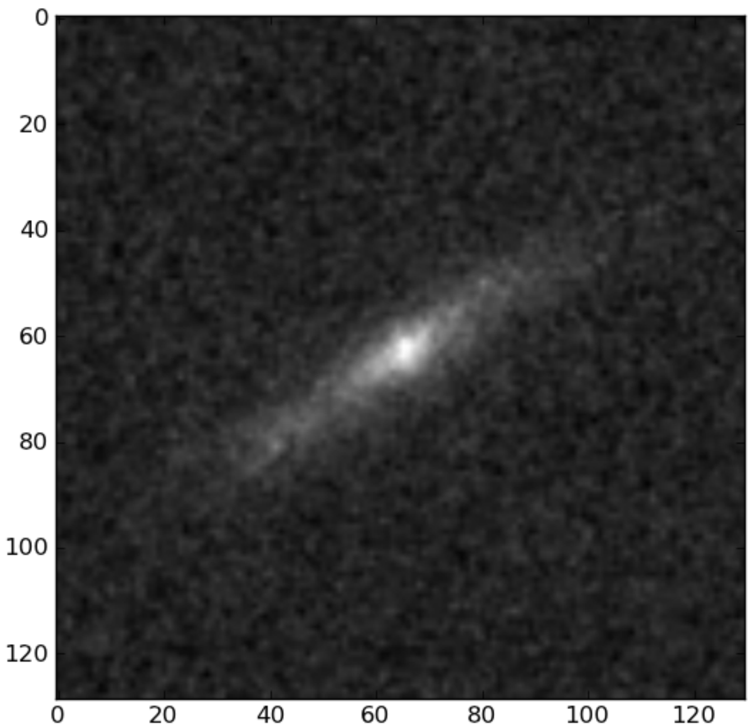
The idea behind model fitting

- Build up some idea for what galaxy light profiles look like $g(x_1,x_2)$:
  - Family of models like Sersic
  - Decomposition into basis functions
- Take the PSF models from the data $P(x_1,x_2)$
- Use maximum-likelihood fitting to find the best-fitting galaxy light profile parameters after comparing the observed image with $g(x_1,x_2) \otimes P(x_1,x_2)$

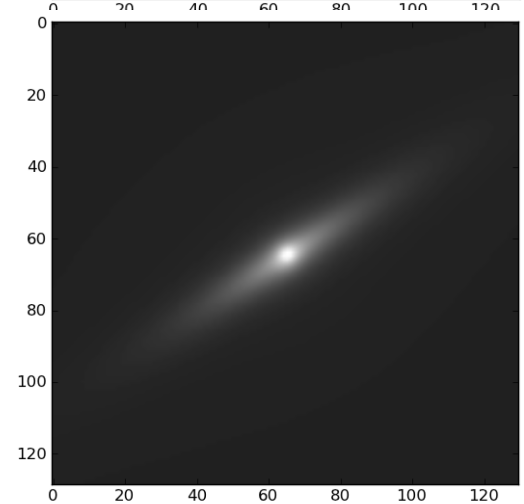
A worked example

Sersic fit: flux, n, size, shape (2 params), centroid (2 params) = 7 parameters



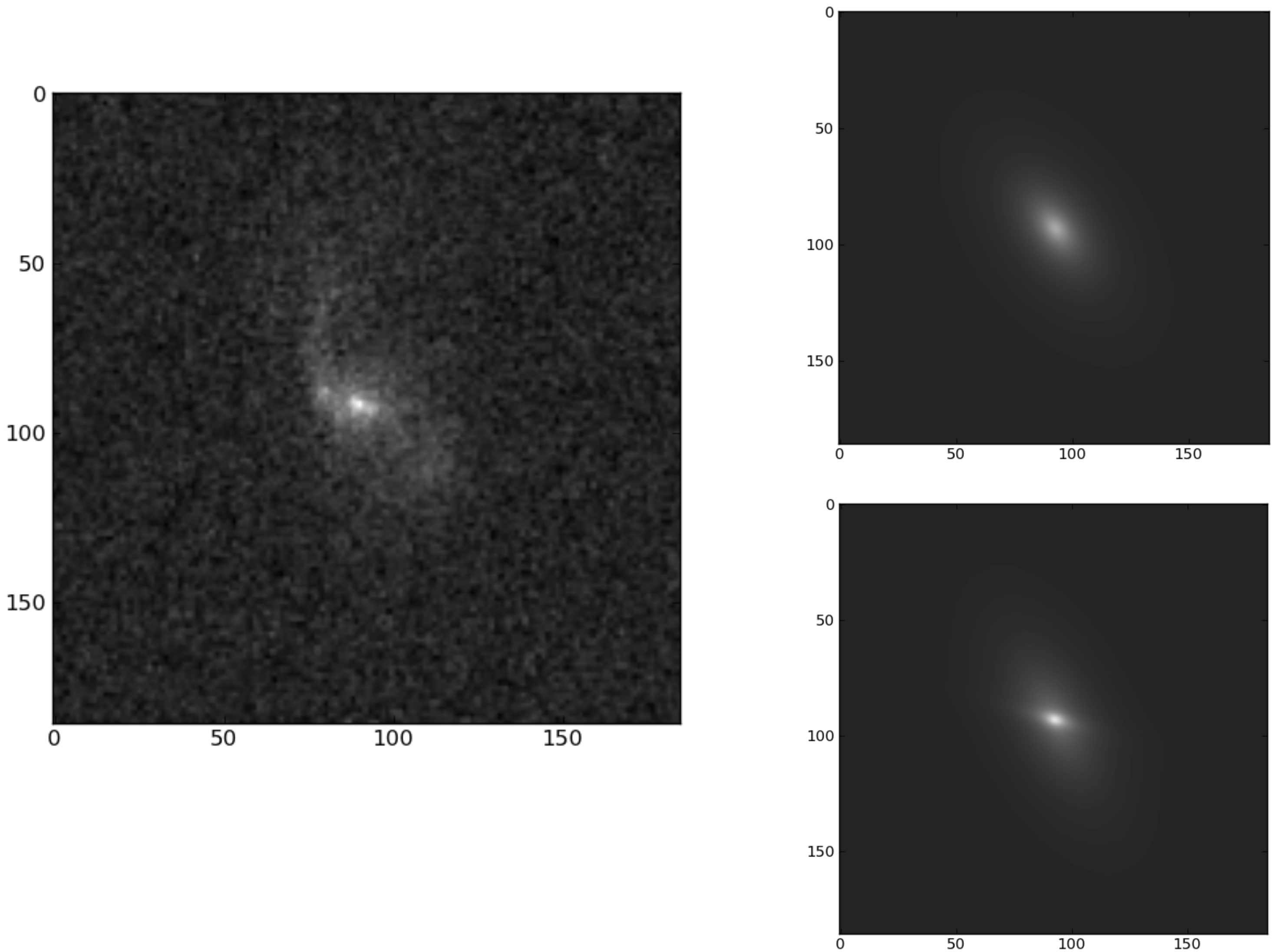


HST i-band image



Bulge+disk fit: 2 flux, 2 size, 4 shape, 4 centroid = 12 parameters

A worked example



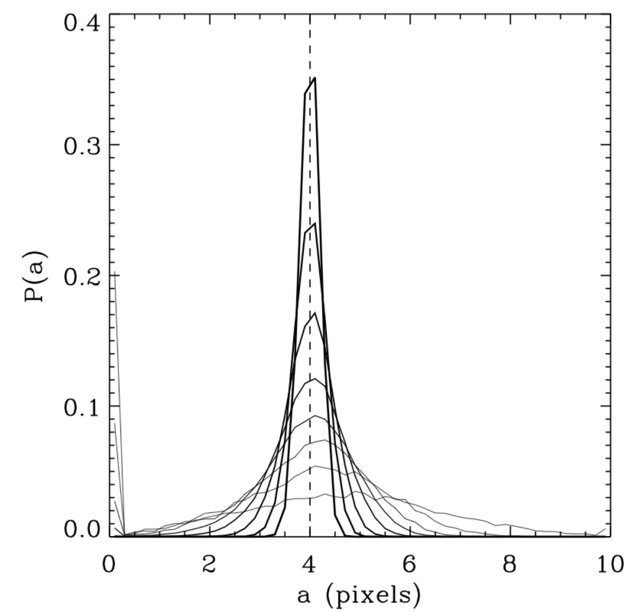
What could possibly go wrong?

(Answer: nearly everything)

- Errors due to inputs
  - PSF model errors
  - Detector effects
- Errors due to failure of ansatz (uncorrelated intrinsic shapes and shear): selection biases
- Errors due to limitations of ML fitting
  - Noise bias, model bias
- And more… (see additional references on DE School webpage)

Example: noise bias

- Noise modifies the likelihood surface, shifting location of the maximum.
- This causes bias in ML fitting
  - More model parameters often = larger biases
- Maximum likelihood estimators are biased in the presence of noise.
- Unavoidable feature of ML fitting



Refregier et al (2012): likelihood for size of 2D Gaussian

Systematic error model

- Multiplicative biases
  - Tend to depend on galaxy, PSF properties
  - Easy to divide samples and identify relative multiplicative bias
  - Very hard to detect overall absolute multiplicative bias
- Additive biases
  - Can be removed / detected via cross-correlations
  - Tend to depend on galaxy, PSF properties

PSF model errors

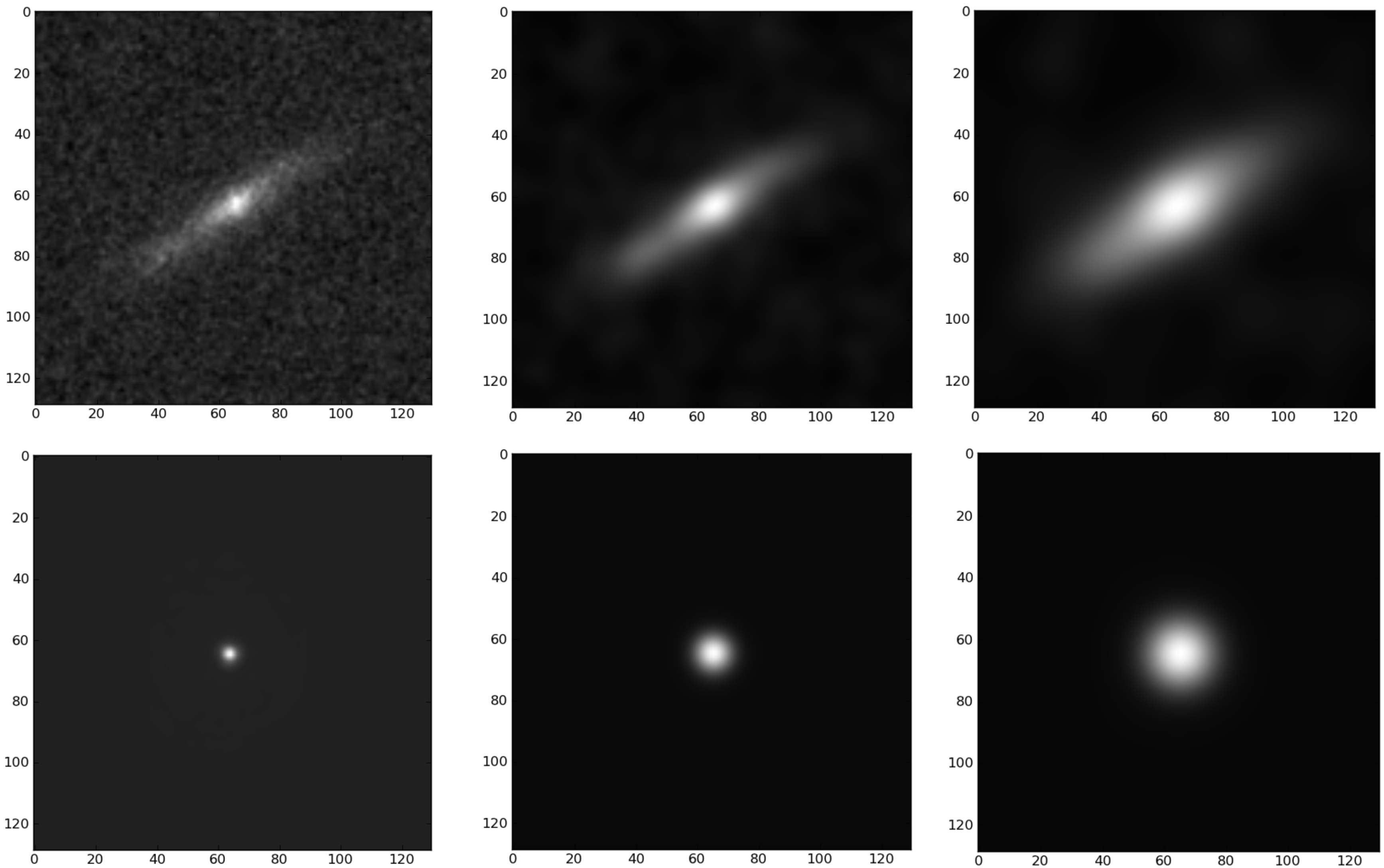
How does the PSF model affect shear estimation?

Model-fitting: Use maximum-likelihood fitting to find the best-fitting galaxy light profile parameters after comparing the observed image with $g(x_1,x_2) \otimes P(x_1,x_2)$

Think of PSF as having two effects:

- blurring/rounding out the light profile
- Imprinting a coherent shear corresponding to the PSF anisotropy

A demonstration: PSF rounding effect



Activity: pairwise discussion

Come up with ideas for how shear estimation can be systematically biased due to PSF modeling errors.

Some answers

- Coherent size errors (PSF model too large / small): shear coherently over/under-estimated = multiplicative bias
- Coherent shape errors: additive biases
- Spatial patterns in size/shape: scale-dependent issues in shear correlations
- Higher-order moments wrong: hard to predict exact nature of error on shear, but something will undoubtedly go wrong

Reference

See Jarvis et al (2016) for clear explanation of how PSF model errors propagate into shear correlations + relevant null tests

Selection bias

What is selection bias?

Assume we select galaxies based on property Y.

- Does the observed value of Y depends on the shear magnitude/direction (non-zero dY/dg)?
- Does the observed value of Y depend on the PSF ellipticity magnitude/direction?

Either condition would result in a shear/PSF-dependent selection. Violates the assumption that intrinsic shapes are isotropically distributed.

Activity: pairwise discussion

Come up with a list of common galaxy selection criteria, and discuss which could cause selection bias for lensing.

My answer

- Cuts on SNR/magnitude
  - Depends on how photometry is measured.
  - Note intrinsic vs. artificial selection biases.
- Apparent size (depending on how it's defined)
- Photo-z-related cuts
- Blending / local density
- Masking
- …?

How to estimate

- Simulations: full estimate of <e_{int} g> etc.
- Simulations: estimate of dY/dg, then use dn/dY to estimate bias (Hirata et al. 2004)
- Typical magnitude: a few % - far above statistical floor for LSST (important even for ongoing surveys)

Take-away lessons

1. The concept of "galaxy shape" is ill-defined, and that's OK – "shape" is just an intermediate goal towards estimating shears from the galaxy ensemble.
2. Most methods of shear inference via averaging galaxy shapes are subject to multiplicative and additive systematics that naturally arise for basic mathematical reasons.
3. PSF model errors cause all kinds of subtle shear estimation biases.
4. Selection bias is an under-appreciated but important weak lensing systematic (of order few %).