Title: Inference of weak gravitational lensing signals from sky images

Date: Apr 14, 2015 11:00 AM

URL: http://pirsa.org/15040051

Abstract: <p>Weak gravitational lensing is a highly valued tool for inferring the structure of the spacetime metric between an observer and a cosmologically distant â€œwallpaper,â€• most commonly either the CMB or faint background galaxies. The best-measured quantities are the second derivatives of the projected scalar potential(s), which are manifested as apparent shearing and magnification of the wallpaper. Given a collection of faint-galaxy images, what information can we extract about the shear and magnification that these images have undergone? I will describe a new method of lensing inference that, unlike predecessors, is rigorously correct in the presence of noise and other observational realities, nearly optimal, and computationally feasible at the scale of current/future surveys like the Dark Energy Survey and the Large Synoptic Survey Telescope.</p>

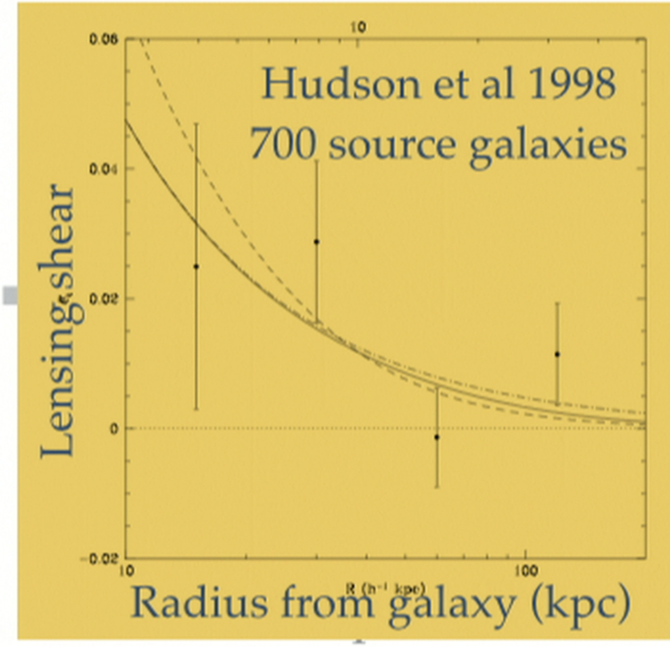
INFERRING WEAK GRAVITATIONAL LENSING FROM SKY IMAGES

G. Bernstein
University of Pennsylvania
14 April 2015
w/Bob Armstrong, Marisa March, Christina Krawiec

1

$$\psi(\vec{\theta}) = \int_0^{\text{source}} dr\, \Psi(r\vec{\theta})\, \frac{2D_{ls}}{c^2 D_l D_s}$$

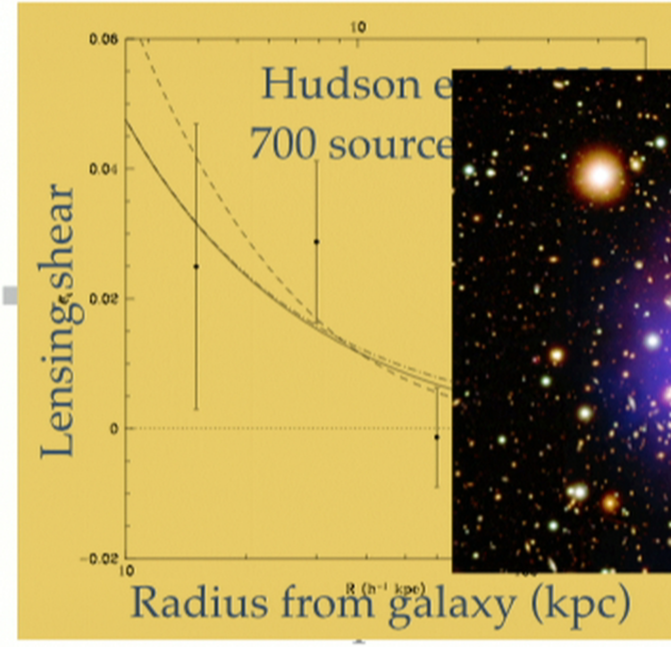
- The projected scalar gravitational potential(s) between us and a source "wallpaper" is a great thing to measure:
  - Relation between total and visible matter
  - Growth of potentials since recombination
  - Ratios of distances vs redshift



$$\Psi(r\vec{\theta})\,\frac{2D_{ls}}{c^2 D_l D_s}$$

- tional potential(s)
- allpaper" is a great thing
- d visible matter
- e recombination
- Ratios of distances vs redshift

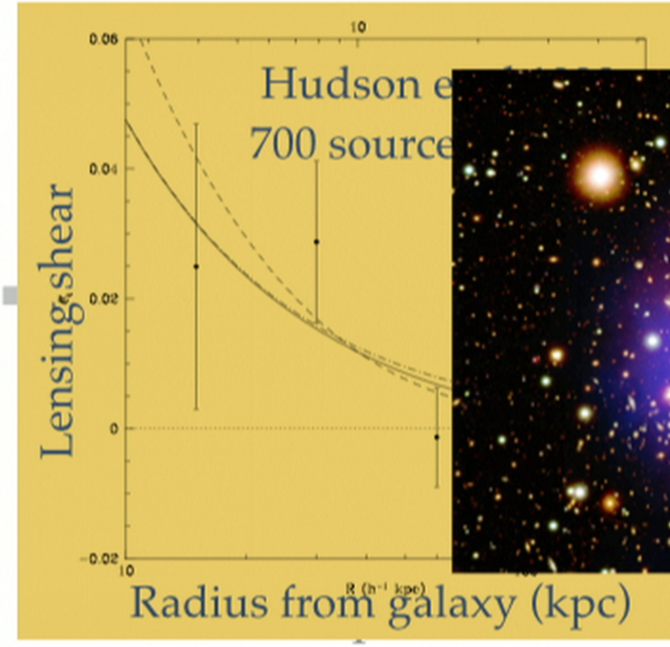
2



ning

e recombination

- Ratios of distances vs redshift

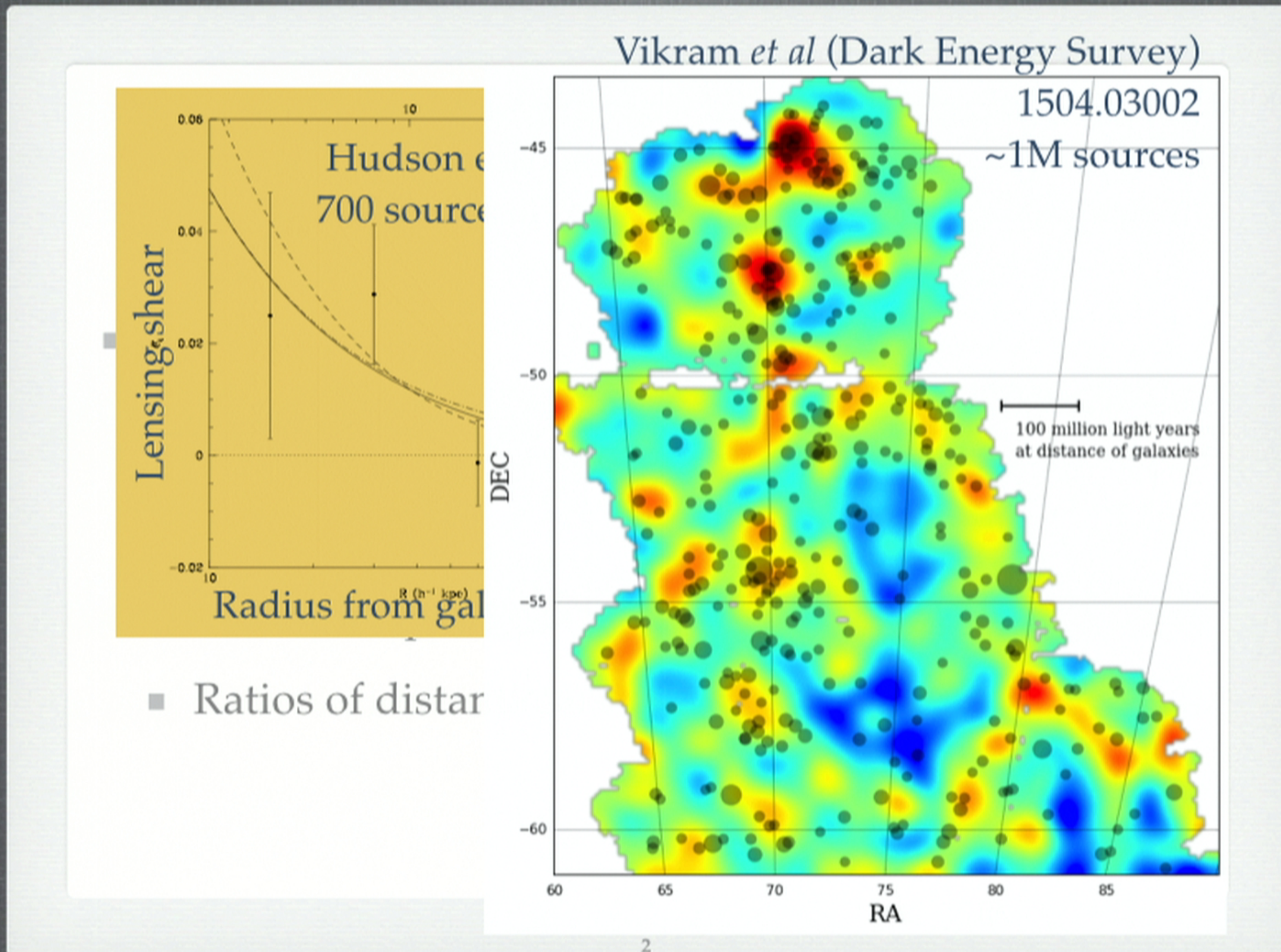


ning

e recombination

- Ratios of distances vs redshift

2


Vikram et al (Dark Energy Survey)
1504.03002
~1M sources
Hudson e
700 source
Lensing shear
Radius from gal
Ratios of distar
100 million light years
at distance of galaxies
DEC
RA
2

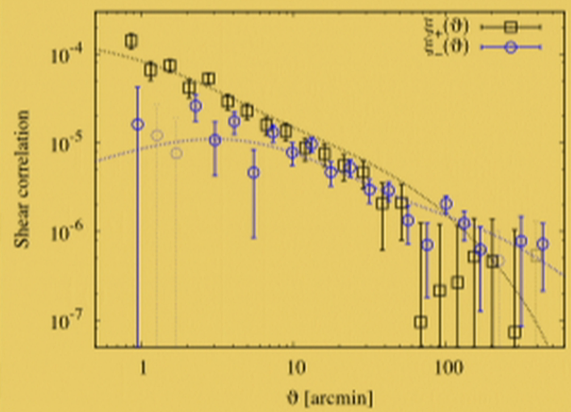
$$\psi(\vec{\theta}) = \int_0^{\text{source}} dr\, \Psi(r\vec{\theta})\, \frac{2D_{ls}}{c^2 D_l D_s}$$

- The projected scalar gravitational potential(s) between us and a source "wallpaper" is a great thing to measure:
  - Relation between total and visible matter
  - Growth of potentials since recombination
  - Ratios of distances vs redshift

Th

bet

to



CFH Lens survey

Kilbinger *et al* (2013)

154 deg²

~5M sources

$$\sigma_8 \left(\frac{\Omega_m}{0.27}\right)^{0.59} = 0.79 \pm 0.03$$

- Growth of potentials since recombination
- Ratios of distances vs redshift

What are the measurable effects of potential?

Signal	Effect	Information?
ψ	Time delay	**None** (unless multiply imaged)
ψ'	Image shift	**None** (unless multiply imaged)
ψ''	Shear, magnification	**YES!**
ψ'''…	"flexions,"…	Yes, but weaker and more difficult than

4

What are the measurable effects of potential?

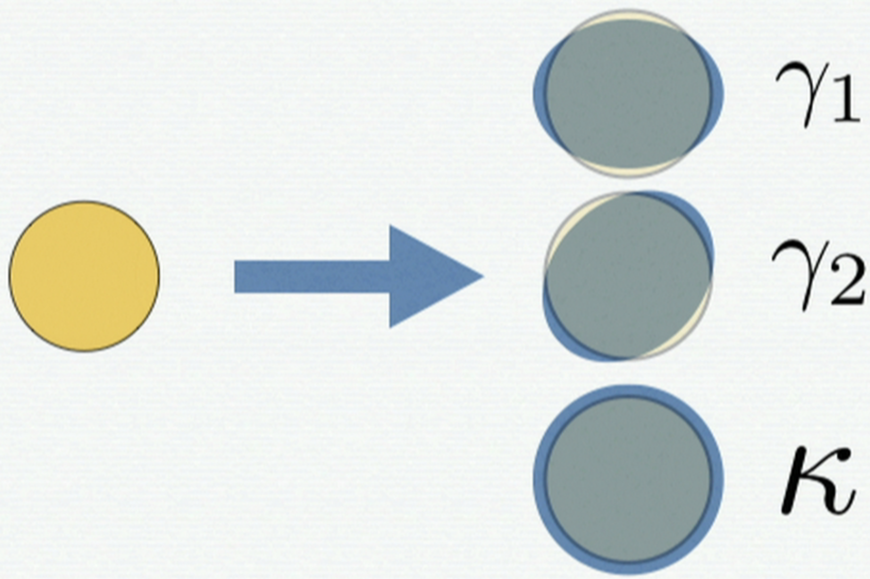
Signal	Effect	Information?
ψ	Time delay	**None** (unless multiply imaged)
ψ'	Image shift	**None** (unless multiply imaged)
ψ''	Shear, magnification	**YES!**
ψ'''…	"flexions,"…	Yes, but weaker and more difficult than

What are the measurable effects of potential?

Signal	Effect	Information?
ψ	Time delay	**None** (unless multiply imaged)
ψ'	Image shift	**None** (unless multiply imaged)
ψ''	Shear, magnification	**YES!**
ψ'''…	"flexions,"…	Yes, but weaker and more difficult than

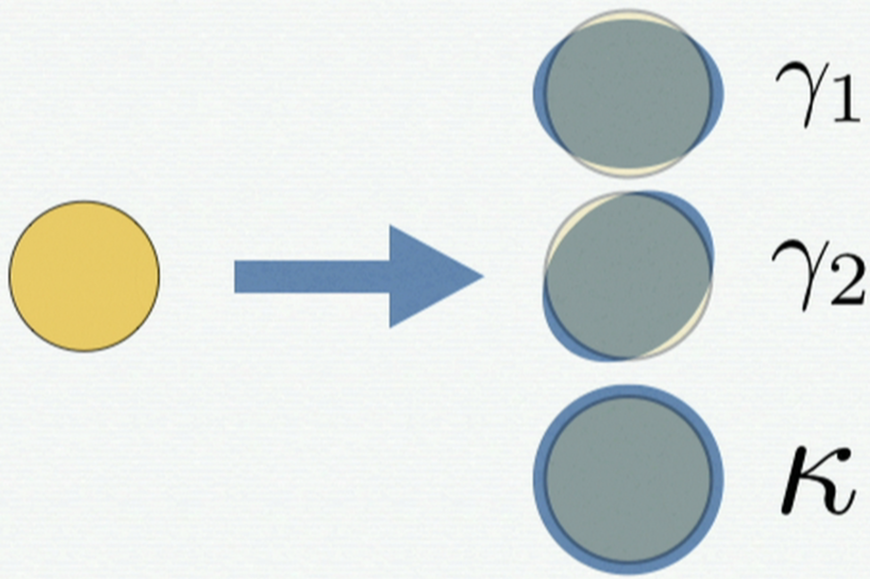
Weak lensing observables

$$\frac{d\theta_{\rm src}}{d\theta_{\rm obs}} = \begin{pmatrix} 1-\kappa+\gamma_1 & \gamma_2 \\ \gamma_2 & 1-\kappa-\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1+g_1 & g_2 \\ g_2 & 1-g_1 \end{pmatrix}$$



Weak lensing observables

$$\frac{d\theta_{\rm src}}{d\theta_{\rm obs}} = \begin{pmatrix} 1-\kappa+\gamma_1 & \gamma_2 \\ \gamma_2 & 1-\kappa-\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1+g_1 & g_2 \\ g_2 & 1-g_1 \end{pmatrix}$$



5

- What is $P(g \mid D)$ given data D that are the pixel values of a set of images of the sky?
- Two successful kinds of wallpaper to date:
  - CMB: unlensed T & E are isotropic Gaussian random fields, B is zero. Optimal g estimation is "easy." Nothing
  - Galaxies: Only *a priori* knowledge of unlensed field is isotropy.

- What is $P(g \mid D)$ given data D that are the pixel values of a set of images of the sky?
- Two successful kinds of wallpaper to date:
  - CMB: unlensed T & E are isotropic Gaussian random fields, B is zero. Optimal g estimation is "easy." Nothing
  - Galaxies: Only *a priori* knowledge of unlensed field is isotropy.

How good should it be?

$$\mathrm{Var}\,\hat{g}_1 = \frac{\mathrm{Var}\,e_1}{N_g} \approx \frac{0.3^2}{10^9} \approx 10^{-10}$$

$$g_{\mathrm{rms}} \approx 0.02 \quad \Rightarrow \quad \frac{\text{signal}}{\text{noise}} \approx \frac{0.02}{10^{-5}} = 2 \times 10^{-3}$$

In order that lensing measurement systematic errors not degrade the statistical power of a large-scale WL survey, the shear inference should be accurate to 1 part in 1000, and any spurious shear signal should be $<10^{-5}$ RMS over cosmologically interesting *l* values.

How good should it be?

$$\mathrm{Var}\,\hat{g}_1 = \frac{\mathrm{Var}\,e_1}{N_g} \approx \frac{0.3^2}{10^9} \approx 10^{-10}$$

$$g_{\mathrm{rms}} \approx 0.02 \quad \Rightarrow \quad \frac{\text{signal}}{\text{noise}} \approx \frac{0.02}{10^{-5}} = 2 \times 10^{-3}$$

In order that lensing measurement systematic errors not degrade the statistical power of a large-scale WL survey, the shear inference should be accurate to 1 part in 1000, and any spurious shear signal should be $<10^{-5}$ RMS over cosmologically interesting *l* values.

Standard approach 2: Model fits

Create a galaxy model that predicts the galaxy signal at each pixel, maximize the likelihood of the model:

$$\mathcal{L}(D_i|D(G)), \quad G = \{e_1, e_2, x, y, \mathit{flux}, \mathit{size}, B/D, \ldots\}$$

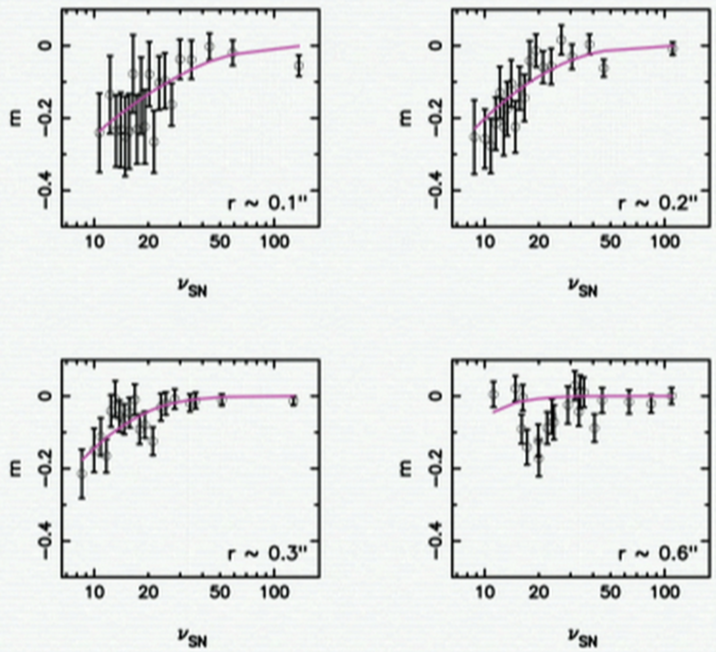
Standard approach 2: Model fits

Create a galaxy model that predicts the galaxy signal at each pixel, maximize the likelihood of the model:

$$\mathcal{L}(D_i|D(G)), \quad G = \{e_1, e_2, x, y, \mathit{flux}, \mathit{size}, B/D, \ldots\}$$

How good is it?

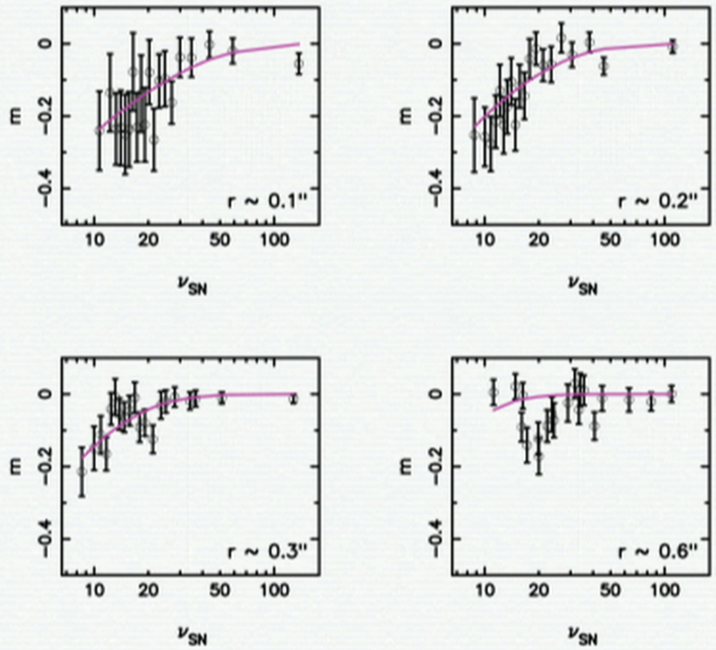
Largest WL survey to date:
CFHLS, 150 deg^2
Miller *et al* (2013) show that *shear* measurement requires ~10% ad hoc correction based on measurements of simulated sheared sky images as S/N<20.



Magnification measurements are substantially cruder.

How good is it?

Largest WL survey to date: CFHLS, 150 deg^2
Miller *et al* (2013) show that *shear* measurement requires ~10% ad hoc correction based on measurements of simulated sheared sky images as S/N<20.



Magnification measurements are substantially cruder.


TEST
LOCAL REMOTE
EMISSION
VOICE CW
MCW
CHANNEL
MANUAL
BATTERY VOLTAGE
P.A GRID
P.A PLATE
CALIBRATE
TUNE
OPERATE
ANTENNA LOADING
E
COARSE
C
ANTENNA TUNING
FINE
D
COARSE
A
HIGH FREQUENCY TUNING
FINE
B
ATTENTION
LACHER
MANIP. OU
MICRO SUR
SECTEUR NON
GRAVE DU
BOUTON
INSTRUCTIONS FOR PRESETTING AUTOTUNE
SIDE TONE
1 2
MICROPHONE
KEY
SET KNOBS ACCURATELY
SET KNOBS ACCURATELY

Why so difficult?

- Point spread function (PSF): observed image is
$$I_{\text{obs}} = I_{\text{src}} * T$$
- Sampling
- Noise - ML estimates are *not* unbiased!
- No dictionary: galaxy images have no finite basis
- Selection biases
- Missing data (cosmic rays, etc.)
- Crowding

14

	Model fitting	Moments
PSF	✔ Forward model	✔ Fourier-domain moments
Sampling	✔ Forward model	✔ (Nyquist sample)
Noise	✘ Correction from simulation	✘ Correction from simulation
No model	✘ Correction from simulation	✔ Non-parametric
Selection	✘ Correction from simulation	✘ Correction from simulation
Missing data	✔ Fit extant data	✘
Crowding	✘ (multi-galaxy fit?)	✘

15

- *Can we produce a lensing estimator from images that does not have poor assumptions or approximations that need to be corrected with ad hoc factors derived from simulations?*

- *Can we do so with sufficient speed?*

(1 core-sec per galaxy x 10^9 sources / 1000 cores = 2 weeks)

Start over again…

- The classical Bayesian inference formula:

$$P(\boldsymbol{g}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{g})P(\boldsymbol{g})}{P(\boldsymbol{D})}$$

- Chop the (image) data into N_g disjoint selected galaxy images D_i:

$$P(g|D) \propto P(g)P(N_g|g)\prod_{i=1}^{N_g} P(D_i|g)$$

17

Start over again…

- The classical Bayesian inference formula:

$$P(\boldsymbol{g}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{g})P(\boldsymbol{g})}{P(\boldsymbol{D})}$$

- Chop the (image) data into N_g disjoint selected galaxy images D_i:

$$P(g|D) \propto P(g)P(N_g|g)\prod_{i=1}^{N_g} P(D_i|g)$$

17

$$\ln P(g|D) = \text{const} + \ln P(g) + \ln P(N_g|g) + \sum_i \ln P(D_i|g)$$

Introduce a descriptive vector G for a galaxy.
Simple case: G fully predicts each pixel value.

$$P(D_i|g) = \int dG\, P(D_i|G) P(G|g)$$
$$= \int dG\, \mathcal{L}\left(D_i - \hat{D}(G)\right) P(G|g)$$

Where do we get P(G | g)? It ***must*** *be empirical; there is no accurate theory for galaxy shapes!*

$$\ln P(g|D) = \text{const} + \ln P(g) + \ln P(N_g|g) + \sum_i \ln P(D_i|g)$$

Introduce a descriptive vector G for a galaxy.
Simple case: G fully predicts each pixel value.

$$P(D_i|g) = \int dG\, P(D_i|G) P(G|g)$$
$$= \int dG\, \mathcal{L}\left(D_i - \hat{D}(G)\right) P(G|g)$$

Where do we get P(G | g)? It ***must*** *be empirical; there is no accurate theory for galaxy shapes!*

Building P(G | g)

- Approximate the 7+ dimensional function by sampling galaxies from the actual sky.
  - *Replicate each real galaxy with rotated copies to closely approximate unsheared, unmagnified distribution.*
  - *Observe at high S/N so we know G(g) well for lensed version of each "template" galaxy.*

$$P(D_i|g) = \sum_{\mu} \mathcal{L}\left(D_i - \hat{D}[G_\mu(g)]\right)$$

19

Building P(G | g)

- Approximate the 7+ dimensional function by sampling galaxies from the actual sky.
  - *Replicate each real galaxy with rotated copies to closely approximate unsheared, unmagnified distribution.*
  - *Observe at high S/N so we know G(g) well for lensed version of each "template" galaxy.*

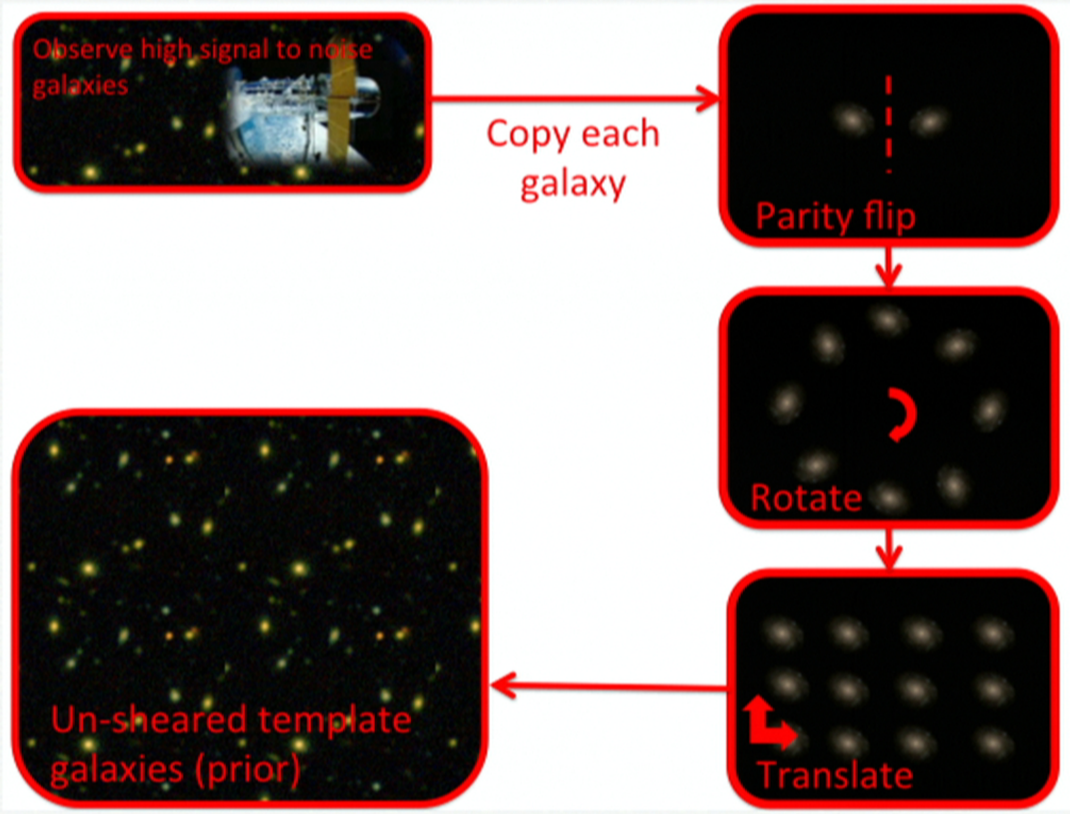
$$P(D_i|g) = \sum_{\mu} \mathcal{L}\left(D_i - \hat{D}[G_\mu(g)]\right)$$

Building P(G|g)

- Approximate the 7+ dimensional function by sampling galaxies from the actual sky.
  - *Replicate each real galaxy with rotated copies to closely approximate unsheared, unmagnified distribution.*
  - *Observe at high S/N so we know G(g) well for lensed version of each "template" galaxy.*

$$P(D_i|g) = \sum_{\mu} \mathcal{L}\left(D_i - \hat{D}[G_\mu(g)]\right)$$

MAKING AN UNLENSED PRIOR



20

Computationally feasible?

$$P(D_i|g) = \sum_{\mu} \mathcal{L}\left(D_i - \hat{D}[G_\mu(g)])\right)$$

- Sum over >10^8 templates for each of 10^9 targets, repeat at every g of interest - ***too slow.***
- It's **weak** lensing: Taylor expand about g=*0*.

$$P(D_i|g) \approx P_i + g \cdot Q_i + \frac{1}{2} g \cdot R_i \cdot g$$

$$P_i = \sum_{\mu} \mathcal{L}\left(D_i - \hat{D}[G_\mu(g=0)]\right)$$

$$Q_i = \nabla_{\boldsymbol{g}} P_i|_{g=0}$$

$$R_i = \nabla_{\boldsymbol{g}} \nabla_{\boldsymbol{g}} P_i|_{g=0}$$

Non-parametric galaxy compression

- We want to describe galaxies by a short vector M having the following properties:
  - Effect of PSF can be removed from M
  - Measurement noise distribution on M is small and is known exactly, rapidly calculable.
  - Does not assume parametric form for galaxies
  - M captures nearly all information on change in galaxy under lensing.

One choice for non-parametric galaxy compression: Fourier-domain moments

$$\begin{Bmatrix} M_I \\ M_x \\ M_y \\ M_r \\ M_1 \\ M_2 \end{Bmatrix} = \int d^2k \, \frac{\tilde{I}_{\rm obs}(\boldsymbol{k})}{\tilde{T}(\boldsymbol{k})} W(|\boldsymbol{k}^2|) \begin{Bmatrix} 1 \\ ik_x \\ ik_y \\ k_x^2 + k_y^2 \\ k_x^2 - k_y^2 \\ 2k_x k_y \end{Bmatrix}.$$

$\mathcal{L}(M_i|M)$ is Gaussian in M_i-M in most imaging modes as M is linear in the pixel values D_i.

$$P(M_i|g) = \int dM \, P(M_i|M) P(M|g) \rightarrow \sum_\mu \mathcal{L}\left[M_i - M_\mu(g)\right]$$

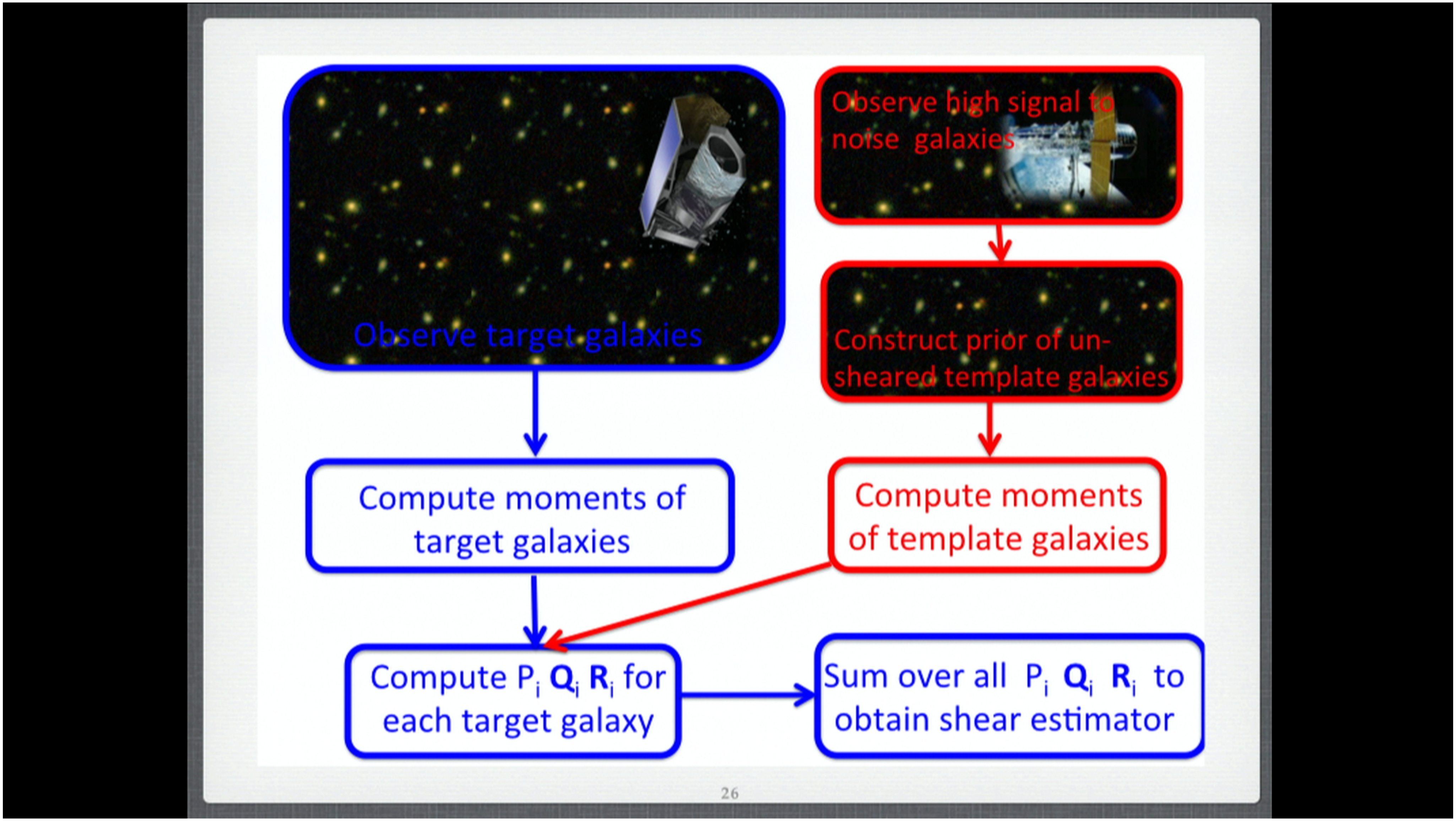
One choice for non-parametric galaxy compression: Fourier-domain moments

$$\begin{Bmatrix} M_I \\ M_x \\ M_y \\ M_r \\ M_1 \\ M_2 \end{Bmatrix} = \int d^2k \, \frac{\tilde{I}_{\rm obs}(\boldsymbol{k})}{\tilde{T}(\boldsymbol{k})} W(|\boldsymbol{k}^2|) \begin{Bmatrix} 1 \\ ik_x \\ ik_y \\ k_x^2 + k_y^2 \\ k_x^2 - k_y^2 \\ 2k_x k_y \end{Bmatrix}.$$

$\mathcal{L}(M_i|M)$ is Gaussian in M_i-M in most imaging modes as M is linear in the pixel values D_i.

$$P(M_i|g) = \int dM \, P(M_i|M)P(M|g) \rightarrow \sum_{\mu} \mathcal{L}\left[M_i - M_\mu(g)\right]$$

25



Observe target galaxies
Observe high signal to noise galaxies
Construct prior of un-sheared template galaxies
Compute moments of target galaxies
Compute moments of template galaxies
Compute P_i $\mathbf{Q}_i$ $\mathbf{R}_i$ for each target galaxy
Sum over all P_i $\mathbf{Q}_i$ $\mathbf{R}_i$ to obtain shear estimator
26

- PSF
- Sampling
- Noise
- Model bias
- Selection bias
- Missing data
- Crowding
- Speed?

Selection effects

- Our starting point should have accounted for the fact that only galaxies passing selection s make it into our data vector:

$$P(g|D) \propto P(g)P(N_g|g)\prod_{i=1}^{N_g} P(D_i|s,g)$$

- Compressing the data into the non-parametric moments:

$$P(M_i|s,g) = \frac{P(M_i,s|g)}{P(s|g)}$$

$$= \frac{\int dM\, P(M_i,s|M)P(M|g)}{\int dM\, P(s|M)P(M|g)}$$

Selection effects

- Our starting point should have accounted for the fact that only galaxies passing selection s make it into our data vector:

$$P(g|D) \propto P(g)P(N_g|g) \prod_{i=1}^{N_g} P(D_i|s, g)$$

- Compressing the data into the non-parametric moments:

$$P(M_i|s, g) = \frac{P(M_i, s|g)}{P(s|g)}$$

$$= \frac{\int dM \, P(M_i, s|M)P(M|g)}{\int dM \, P(s|M)P(M|g)}$$

29

Own your selection!

- In most surveys the selection s is difficult to even define, or at best is a complex function of the D_i images.
- We can make our life hugely easier by making our selection *based on the observed values* M_i, in which case $P(M_i \mid s, M) = P(M_i \mid M)$ for all selected galaxies! Now

$$P(M_i|s,g) = \frac{\sum_\mu \mathcal{L}\left[M_i - M_\mu(g)\right]}{\sum_\mu \int_{\hat{M}\in s} \mathcal{L}\left[\hat{M} - M_\mu(g)\right]}$$

Missing data

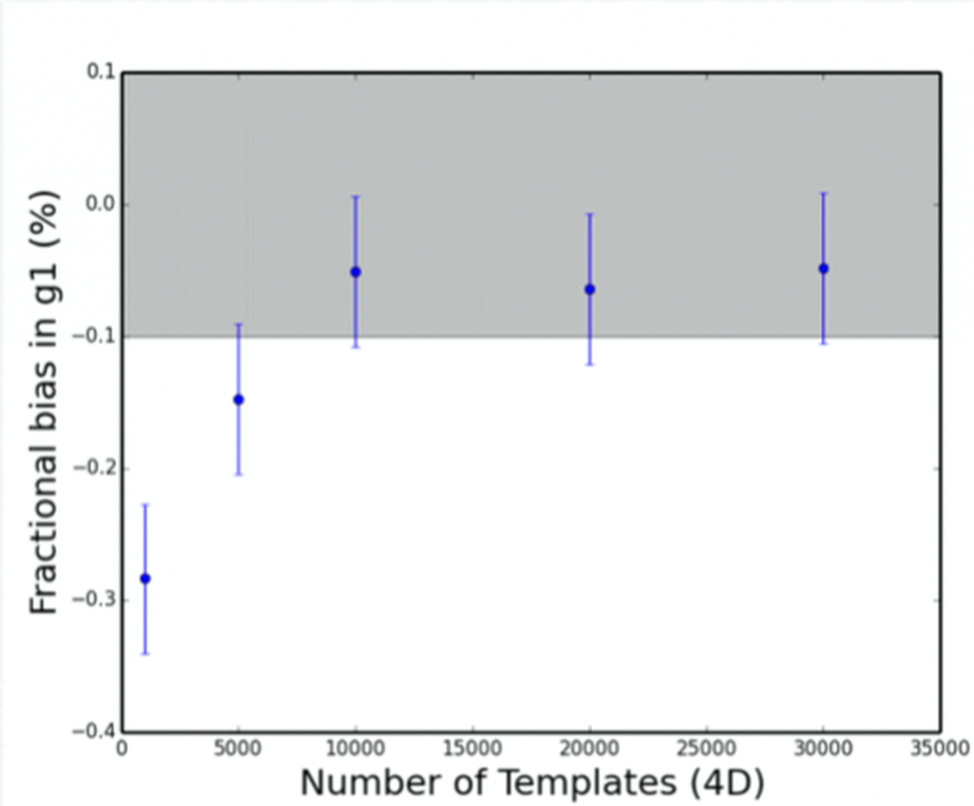
- With multiple Nyquist-sampled exposures, missing data in 1 exposure can be filled in using information from other exposures.
- By definition a model-free measurement cannot be conducted on under-determined image data!
- If the amount of missing data is small, we could inpaint using prior assumptions about galaxies: either conduct a model fit, or use compressed sensing techniques (sparsity prior) to replace the data. Model biases are suppressed.
- If the missing data contain a large fraction of the galaxy's flux, just throw it away! Will be rare in ground-based images.

Crowding

- Overlapping galaxies at same z: treat as one galaxy!
- Overlaps at different z: tough since g is z-dependent.
- Division of image flux into two galaxies is ill-posed without strong priors on galaxy morphologies. No rigorous approach.
- Model-fitting methods can proceed with joint fits to pixel data *if the foreground galaxy is transparent*. It is generally not.
- This is a tough problem. For DES I will choose to discard the <10% of galaxies that have a detectably bright neighbor significantly impinging.
- LSST aims to measure fainter galaxies at higher surface density, so a larger overlap rate that must be dealt with. Euclid, WFIRST have better resolution so can use smaller windows to reduce overlaps.

"Bayesian Fourier Domain" (BFD) development

- Formalism produces lensing estimators that are unbiased by any of the known difficulties, save missing data (not bad) and crowding (could be bad).
- Method has been numerically verified to part-per-thousand accuracy on simple model systems.
- Efficient algorithm for doing the 6d integrals over large template sets.
- Currently verifying the method and code on simulated pixelized images of sheared sky, first application to the *Dark Energy Survey* this year.
- ***Nothing in the method limits it to shear: magnification can be derived with the same method.*** Work of Huff *et al* suggests that kurtosis moment will enhance information content for magnification
- Need to develop smart inpainting method to deal with real data.
- We will probably just dodge the crowding challenge for DES. Open problem how to get the fainter sources from LSST data in 2020's.



Example numerical test: this case, testing number of template galaxies needed in the restricted case of known galaxy center (by Bob Armstrong)

BFD method virtues

- First rigorous treatment of noise, PSF, selection, centering for galaxy lensing
- Each target galaxy contributes what it “knows” about the shear without explicit weighting, e.g. zero weight at S/N=0, fixed weight at high S/N.
- The galaxy template can be constructed with same instrument as the survey is done.
- Selection criteria are included naturally as a cut on the weighted moments.
- Computationally simple and fast.
- Trivial extension to include magnification along with shear.
- Multiple exposures included in an obvious way.
- (Almost) no iterative solving: very difficult for the measurement to fail
- *No assumptions about the nature of galaxies on the sky!*