Measuring Shear with the Bayesian Fourier Domain Method

Bob Armstrong

Accurate lensing in the era of precision Cosmology

January 15, 2018

Collaborators:
Gary Bernstein, Katie Eckert
HSC Collaboration

LLNL-PRES-765445

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Bayesian Fourier Domain Philosophy

- We want a statistically rigorous shear measurement method
  - Propagate the pixel noise through to shear uncertainty
  - Account for selection effects
  - Avoid model bias
  - Avoid use of external simulations for calibration
  - Treat PSF and other instrumental effects

- This allows us to
  - Handle systematic errors, nuisance parameters appropriately
  - Embed method into forward/hierarchical model
  - Naturally incorporate photo-z into a consistent estimator

- Must be computational feasible

BFD Overview

1. Compress pixel data into feature space *M* with:
   - Known likelihood
   - *M* independent of observing conditions

$$P(\boldsymbol{D}_i|\boldsymbol{g}) = \int d\boldsymbol{M}_{\text{true}} dx_{\text{true}} P(\boldsymbol{M}_{\text{obs}}|\boldsymbol{M}_{\text{true}}) P(\boldsymbol{M}_{\text{true}}|\boldsymbol{g})$$

   - Use low order moments in Fourier space with fixed weight function. Fast, Linear, Gaussian

$$\boldsymbol{M}(\boldsymbol{x}_0) \equiv \begin{pmatrix} M_f \\ M_r \\ M_+ \\ M_\times \end{pmatrix} = \int d^2k \, \frac{\tilde{I}^o(\boldsymbol{k};\boldsymbol{x}_0)}{\tilde{T}(\boldsymbol{k})} W(|\boldsymbol{k}^2|)\boldsymbol{F}; \qquad \boldsymbol{F} \equiv \begin{pmatrix} 1 \\ k_x^2 + k_y^2 \\ k_x^2 - k_y^2 \\ 2k_x k_y \end{pmatrix}$$

BFD Overview

2. Approximate the intrinsic distribution *P*(*M*|*g*) from "noiseless" sample of deep galaxies to create a "template" set. Use kernel density estimator
   – We know how *M* changes under shear and can calculate derivatives
   – Use the measurement error as a smoothing kernel
   – Exploit symmetries to increase sample size.

$$P_i \propto \sum_{\text{deep}} P(\boldsymbol{M}_{\text{obs}}|\boldsymbol{M}_{\text{deep}}) \propto \sum_{\text{deep}} \mathcal{N}(\boldsymbol{M}_{\text{obs}} - \boldsymbol{M}_{\text{deep}}, \sigma_M)$$

BFD Overview

3. Selection function is restricted to subspace of moments
   - For flux only selection this is analytic, but can also make cuts on things like size

4. Weak shear approximation: assume posterior is quadratic and Taylor expand around *g*=0.

$$-\log P(\boldsymbol{g}|\boldsymbol{D}) \approx (\text{const}) - \sum_{i=1}^{N} \log\left(P_i + \boldsymbol{g}\cdot\boldsymbol{Q}_i + \frac{1}{2}\boldsymbol{g}\cdot\mathbf{R}_i\cdot\boldsymbol{g}\right)$$

$$P_i = \int d\boldsymbol{M}_{\text{true}} dx_{\text{true}} P(\boldsymbol{M}_{\text{obs}}|\boldsymbol{M}_{\text{true}}) P(\boldsymbol{M}_{\text{true}}|\boldsymbol{g}=0)$$

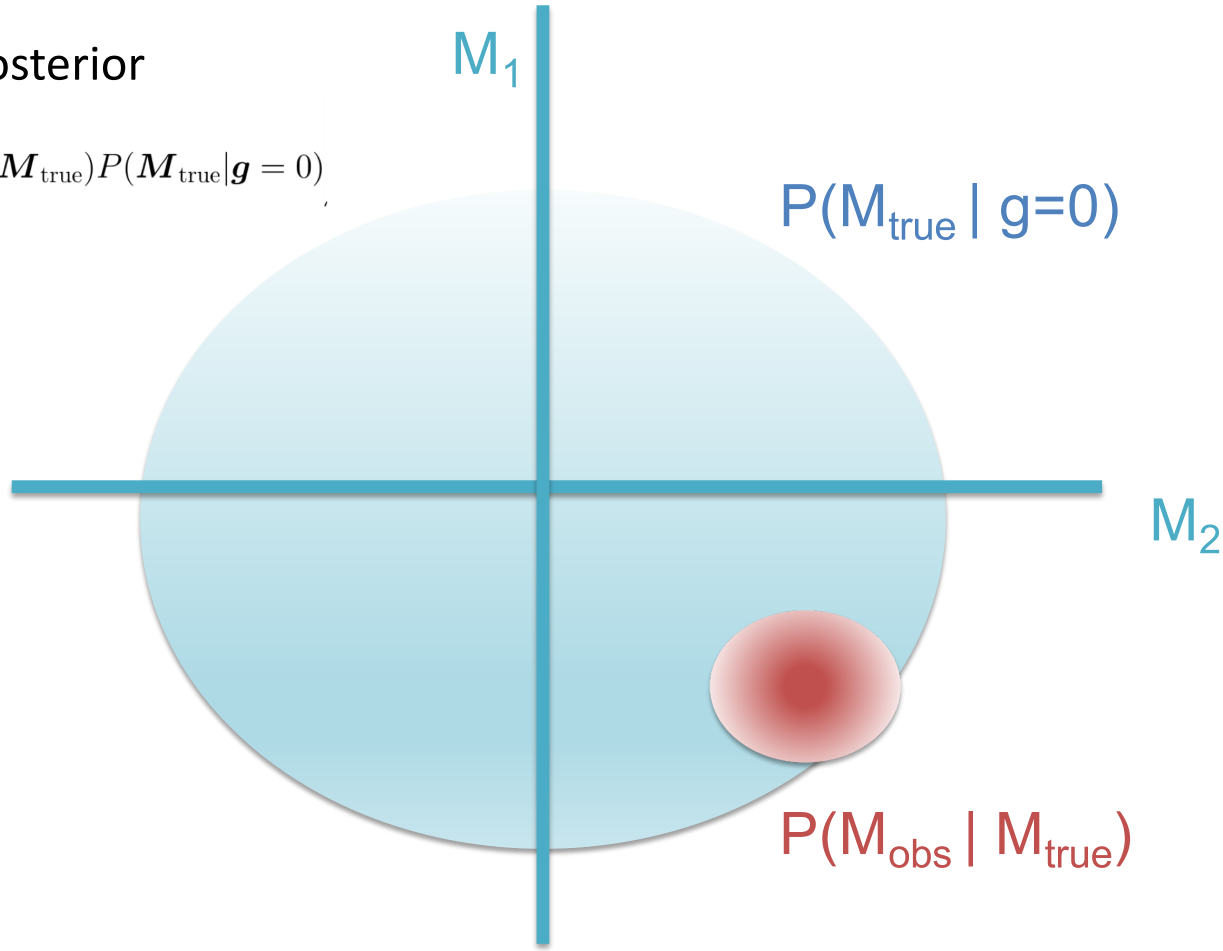
$$\boldsymbol{Q}_i = \int d\boldsymbol{M}_{\text{true}} dx_{\text{true}} P(\boldsymbol{M}_{\text{obs}}|\boldsymbol{M}_{\text{true}}) \nabla_{\boldsymbol{g}} P(\boldsymbol{M}_{\text{true}}|\boldsymbol{g}=0)$$

$$\mathbf{R}_i = \int d\boldsymbol{M}_{\text{true}} dx_{\text{true}} P(\boldsymbol{M}_{\text{obs}}|\boldsymbol{M}_{\text{true}}) \nabla_{\boldsymbol{g}} \nabla_{\boldsymbol{g}} P(\boldsymbol{M}_{\text{true}}|\boldsymbol{g}=0)$$

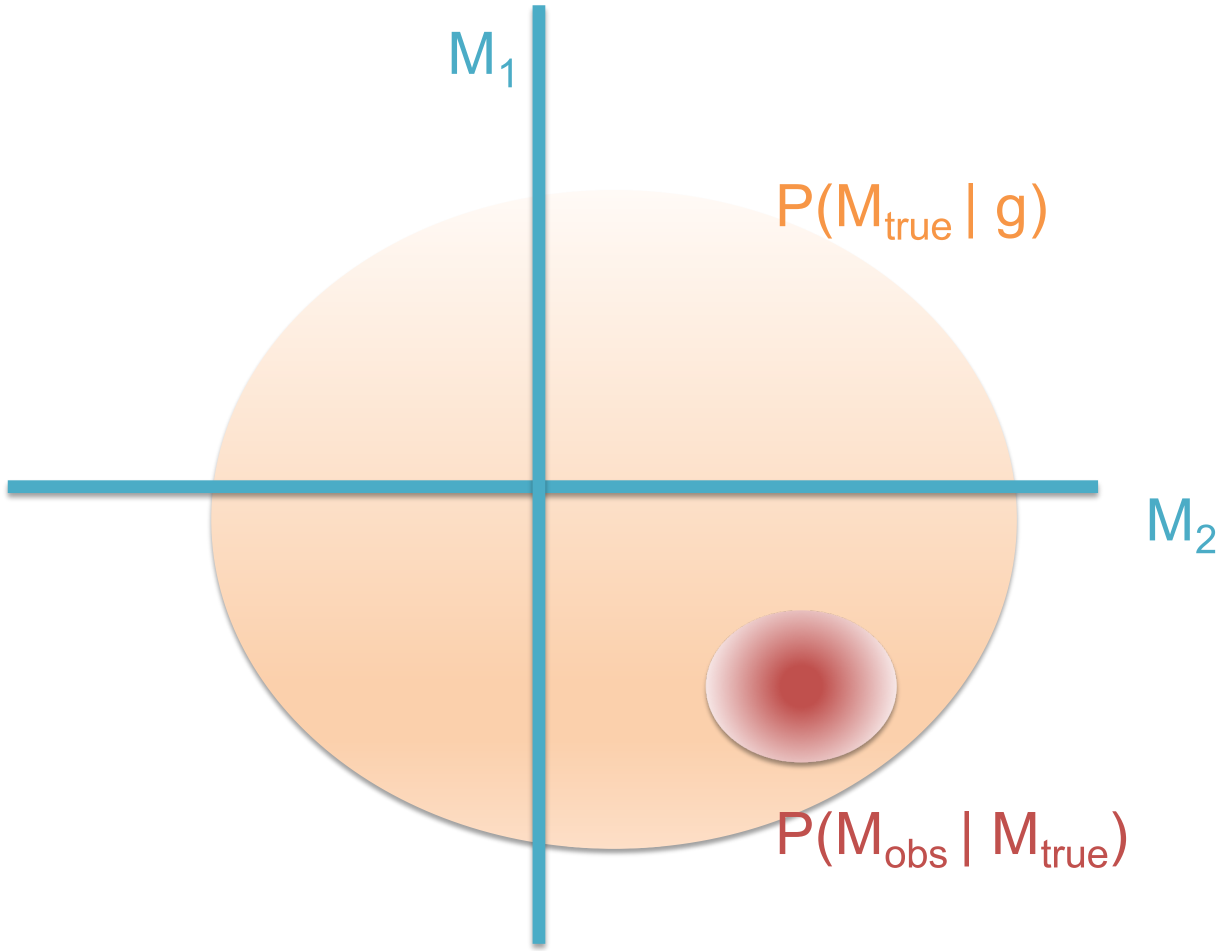
   - Need to keep 6 numbers per galaxy
   - Assume *g* is constant or that you have a model for *g* at the location of each galaxy

Example computing posterior

$$P_i = \int d\boldsymbol{M}_{\rm true} dx_{\rm true} P(\boldsymbol{M}_{\rm obs}|\boldsymbol{M}_{\rm true}) P(\boldsymbol{M}_{\rm true}|\boldsymbol{g}=0)$$

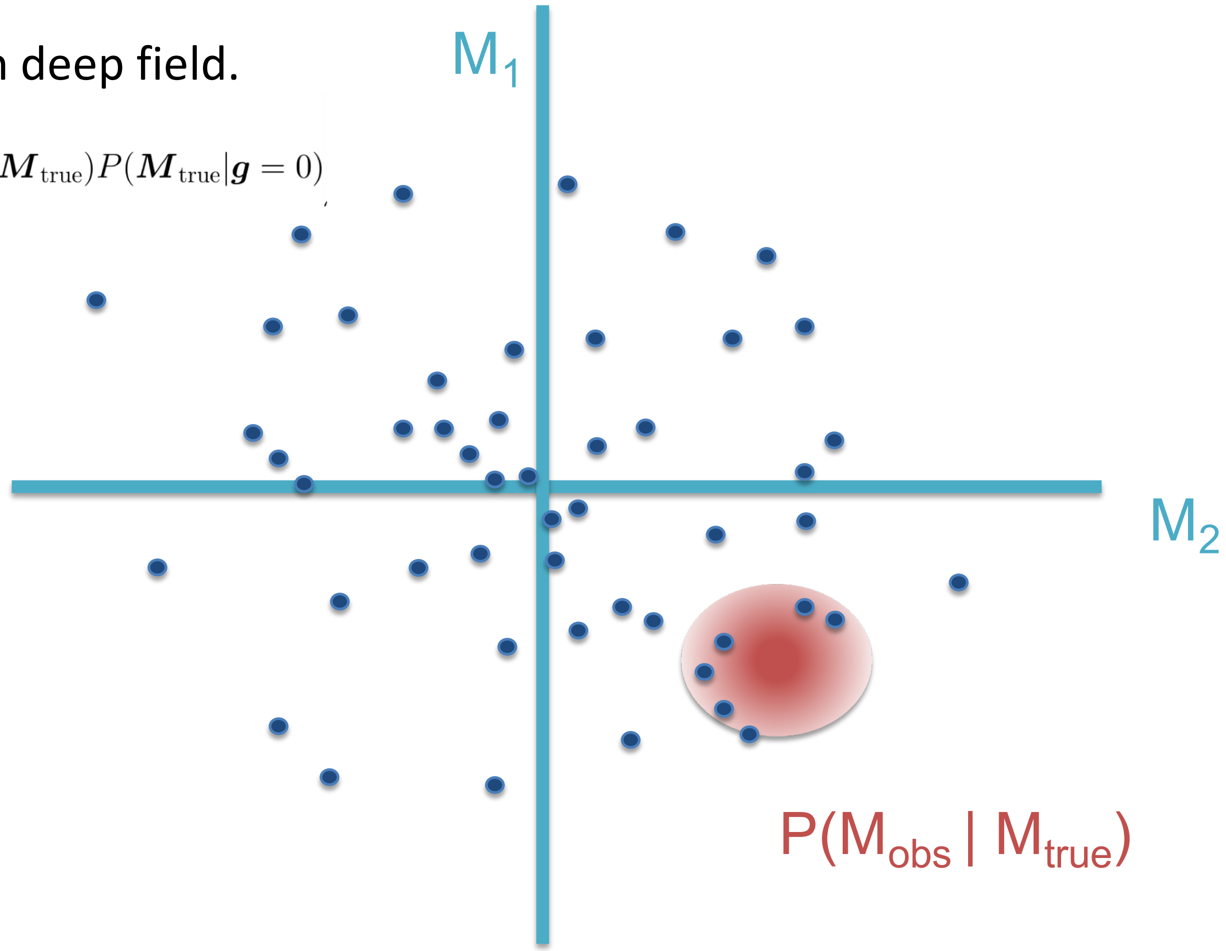


Sheared distribution



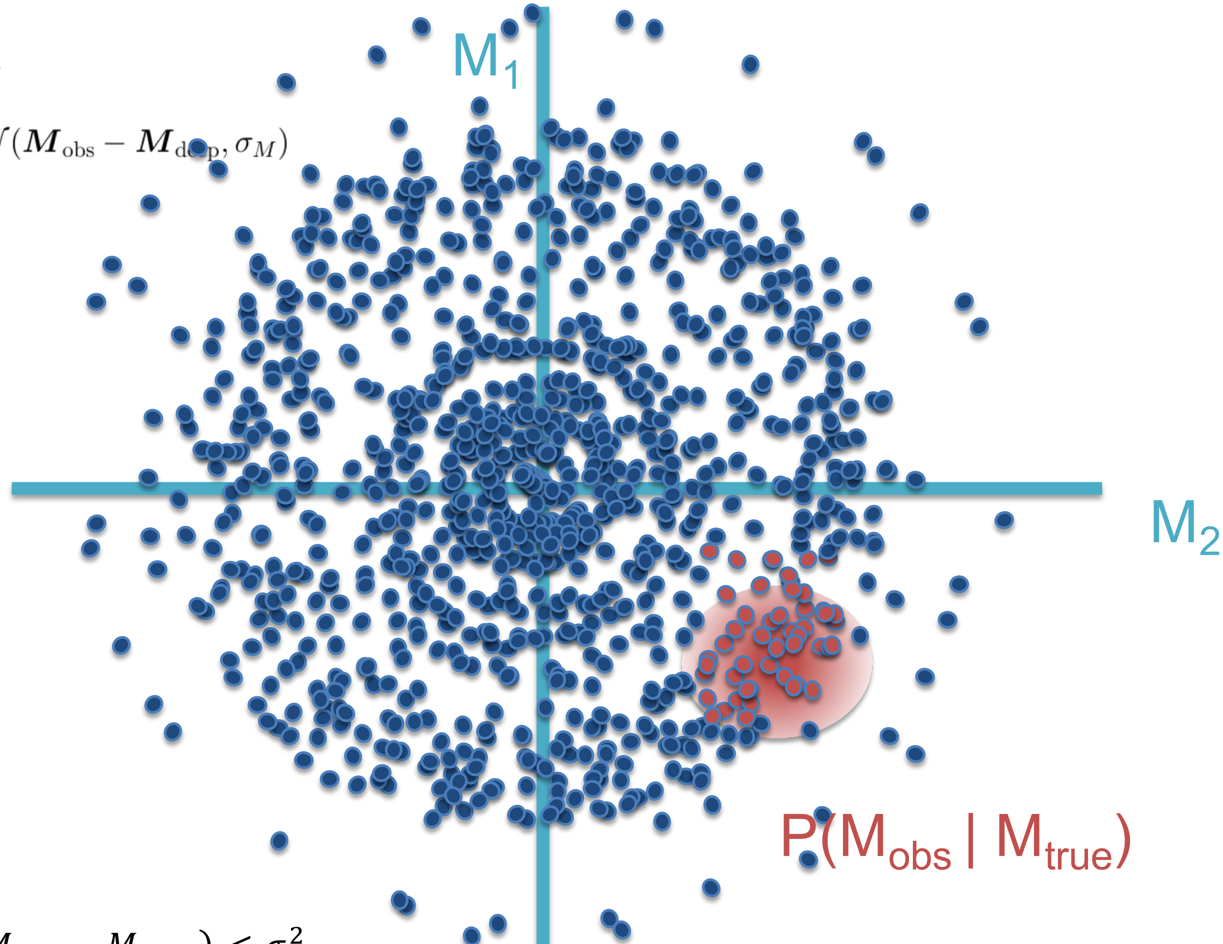
Sample moments from deep field.

$$P_i = \int d\boldsymbol{M}_{\mathrm{true}} dx_{\mathrm{true}} P(\boldsymbol{M}_{\mathrm{obs}}|\boldsymbol{M}_{\mathrm{true}}) P(\boldsymbol{M}_{\mathrm{true}}|\boldsymbol{g}=0)$$



Replicate deep sample

$$P_i \propto \sum_{\text{deep}} P(\boldsymbol{M}_{\text{obs}} | \boldsymbol{M}_{\text{deep}}) \propto \sum_{\text{deep}} \mathcal{N}(\boldsymbol{M}_{\text{obs}} - \boldsymbol{M}_{\text{deep}}, \sigma_M)$$



Keep all galaxies with

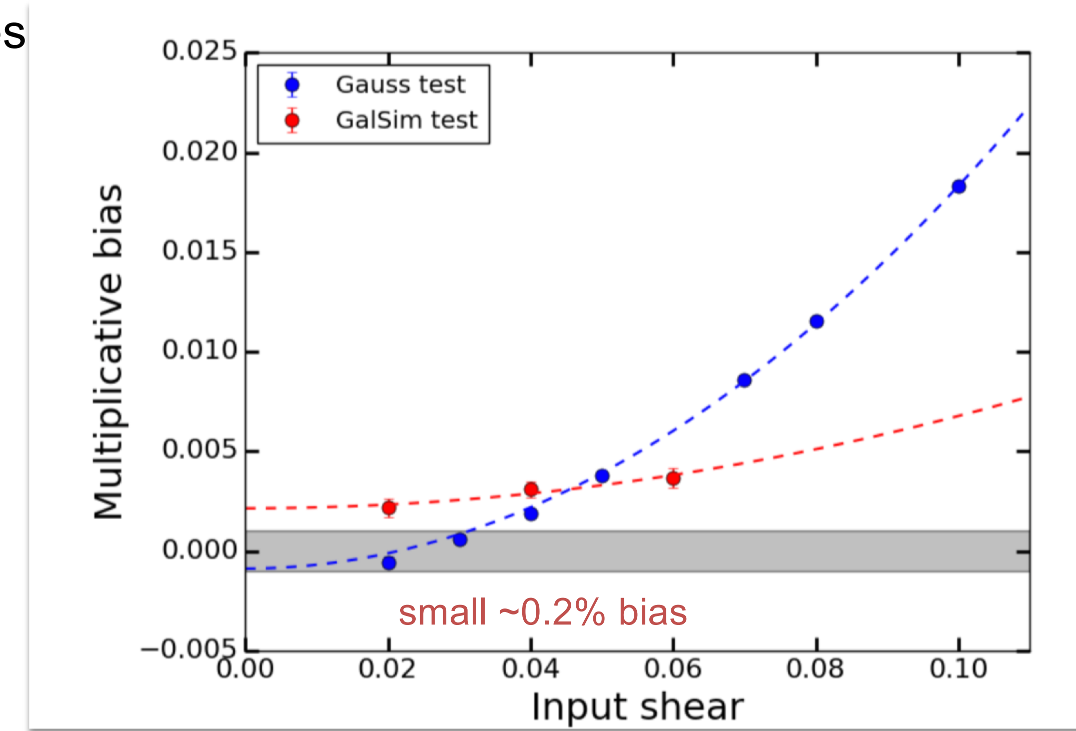
$$\chi^2 = (M_{obs} - M_{true}) C_M^{-1} (M_{obs} - M_{true}) < \sigma_{max}^2$$

Results on Simulations

- Test on analytic Gaussian galaxies and on complicated ones using GalSim down to *S/N*=5.

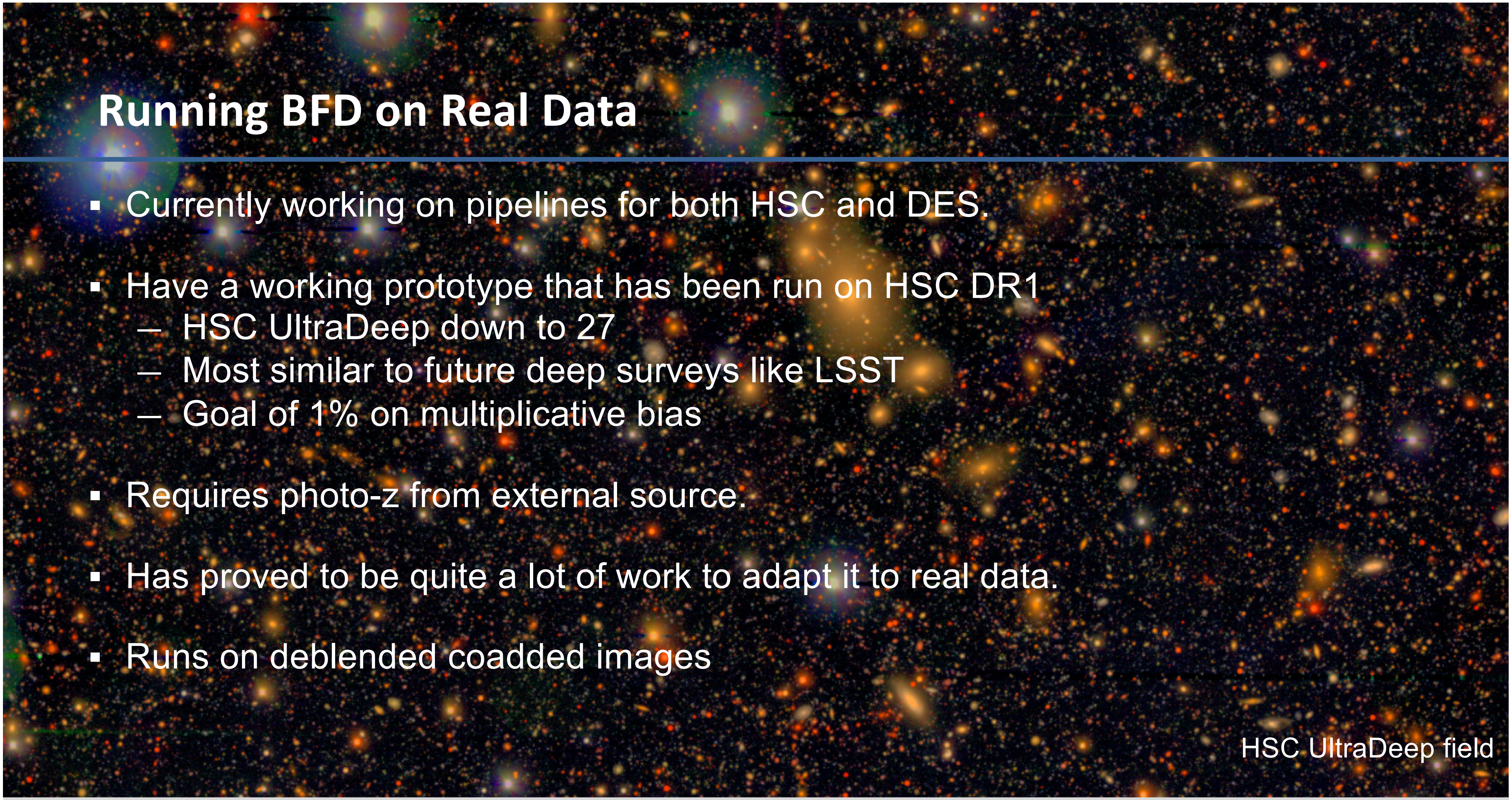
- Predicted/Observed selection:
  – 0.69111 ± 0.00006
  – 0.69105

- Predicted/Observed shear error per 50,000 galaxies
  – 3.58 e^{-4}
  – 3.56 ± 0.02 e^{-4}



Running BFD on Real Data

- Currently working on pipelines for both HSC and DES.
- Have a working prototype that has been run on HSC DR1
  - HSC UltraDeep down to 27
  - Most similar to future deep surveys like LSST
  - Goal of 1% on multiplicative bias
- Requires photo-z from external source.
- Has proved to be quite a lot of work to adapt it to real data.
- Runs on deblended coadded images



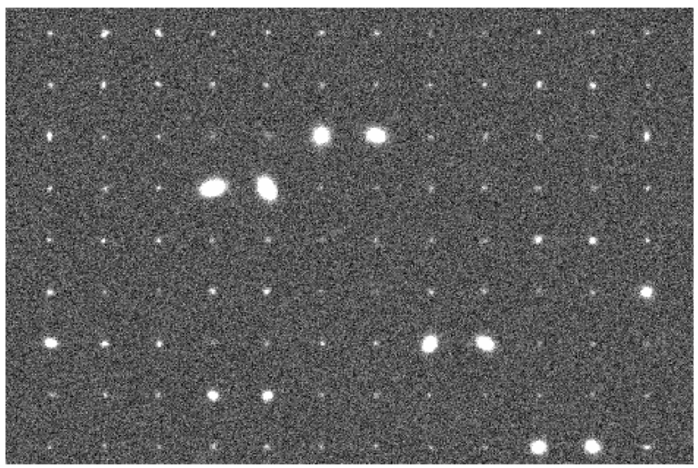
HSC UltraDeep field

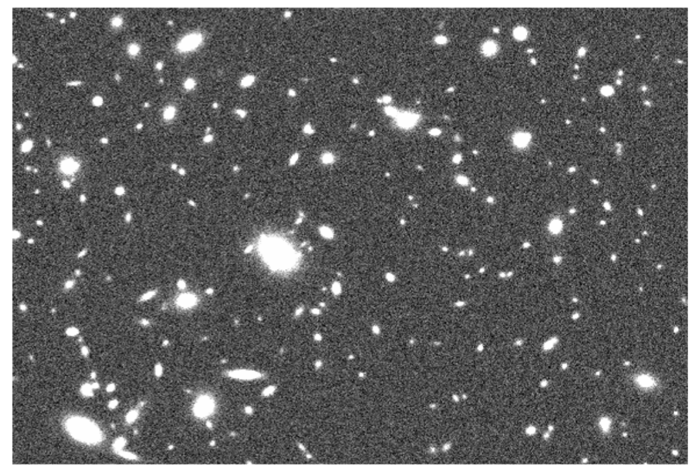
Blending Constant Shear

Keep deep sample as isolated postage stamps, but change noisy galaxy density.

If you match the density between deep and noisy bias is removed.

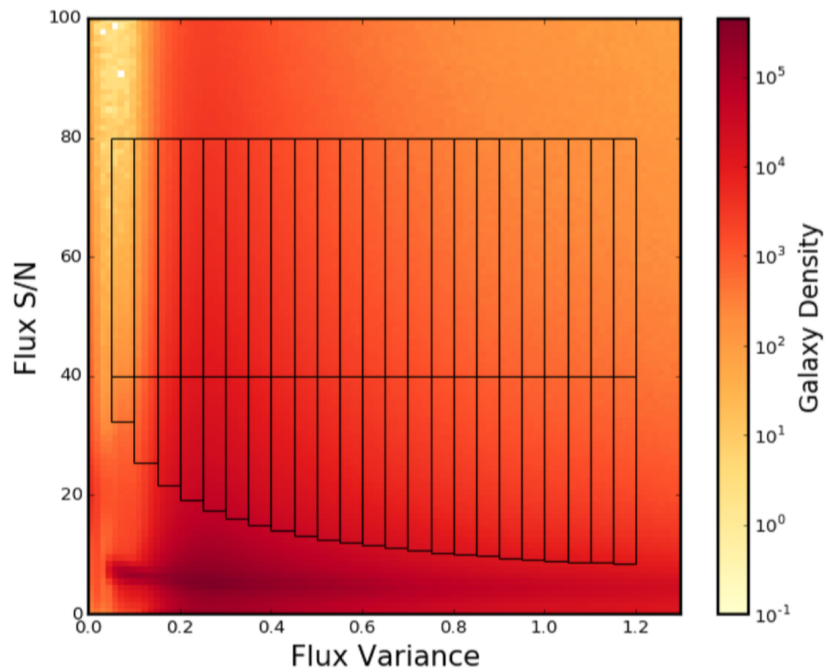
Sample	Shear bias low S/N	Shear bias high S/N
Isolated	<1%	<1%
Random 0.5x density	<1%	<1%
Random 1x density	-3%	<1%
Random 2x density	-8%	-5%



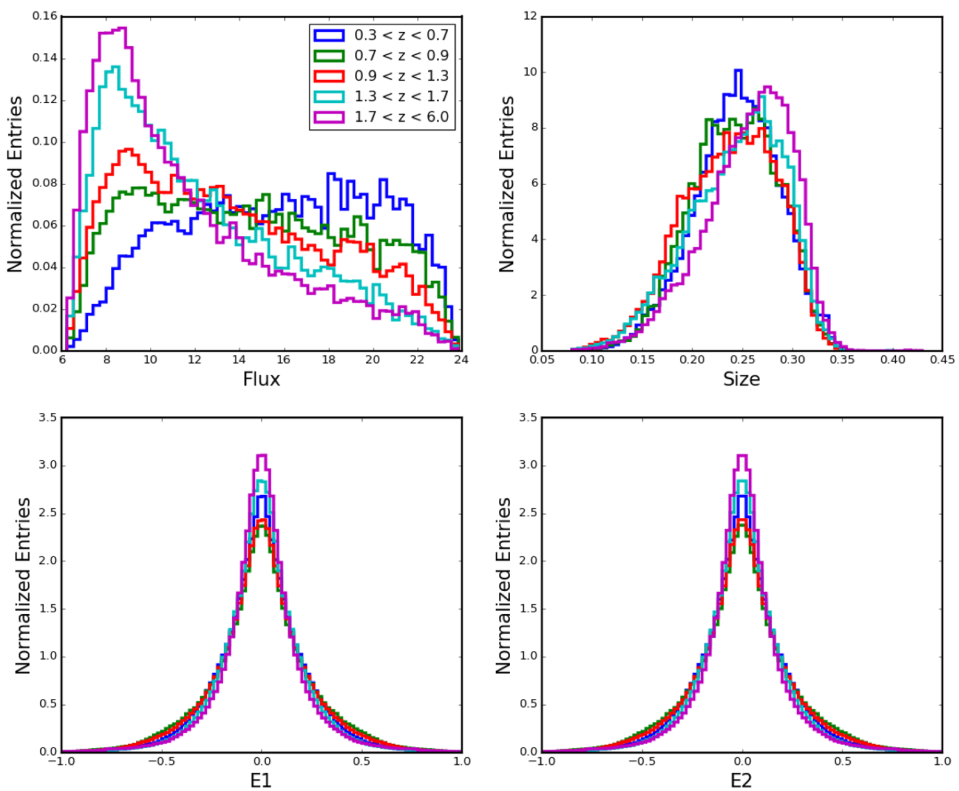


Noise complications on real data

- The noise properties vary significantly over the survey.
  – Currently use the moment covariance to reject galaxies far from our selection. Must divide the data into groups with similar noise.
- For high S/N measurement error is too small to have sufficient deep galaxies. We artificially inflate the noise.
- Use conservative cut at ~24.5 because photozs are not reliable beyond that.



Variation with redshift



- Moments also vary with redshift. Must construct separate deep galaxies for different redshift bins.
- Combined with noise bins, we construct over 200 separate deep galaxy samples!

Future

- Future developments
  - Magnification measurement
  - Joint shear + photoz to infer tomographic shear posterior
  - Integrating into hierarchical model being developed at LLNL

- Active software development
  - HSC and DES
    - Application to HSC data is almost finished
  - Euclid and LSST