A probabilistic cosmic shear pipeline

Berkeley Weak Lensing Workshop

January 15, 2019

Michael D. Schneider
Physics Division, LLNL

Collaborators:
Will Dawson, Josh Meyers, Brian Bauman,
D. Bard, N. Golovich, D. Hogg, D. Lang, P. Marshall, K. Ng, L. Poyneer, K. Shah

LLNL-PRES-733055

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

Summary: New cosmological parameter pipeline for cosmic shear without two-point functions

- Avoid the need to compute / estimate **two-point function covariances**
- **Propagate all uncertainties** in the shear measurement and systematics
- Use **all the information** in the cosmological LSS (in principle)

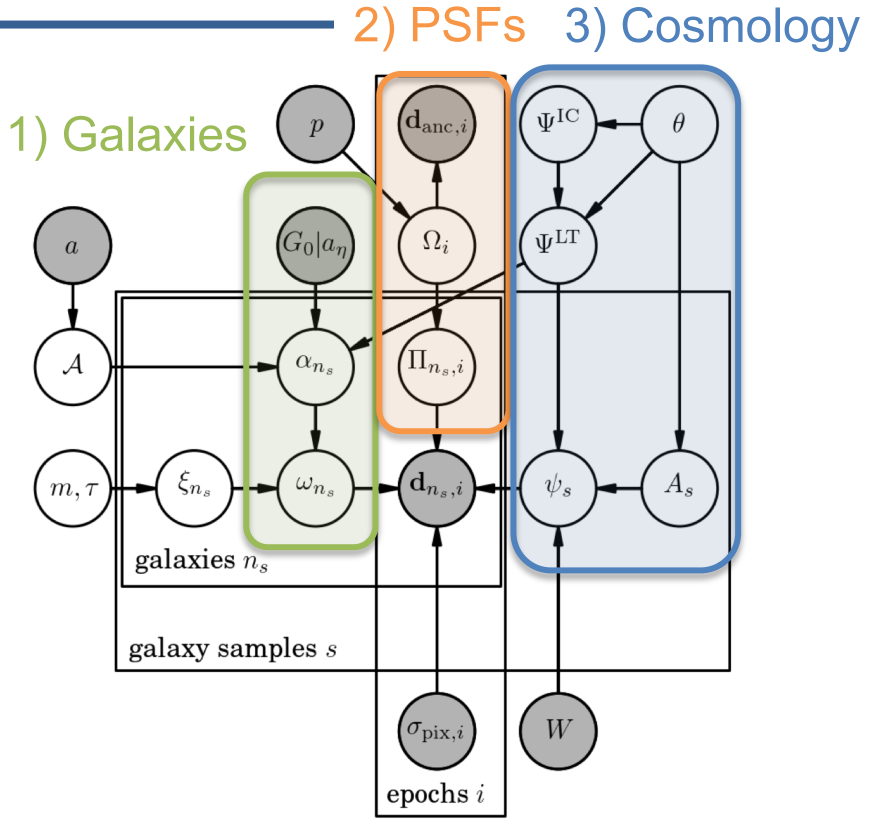
The "data" are pixels not maps or two-point functions.
Exact likelihood is tractable with distributed Monte Carlo computing approaches

Introduction

The complete statistical model for cosmic shear

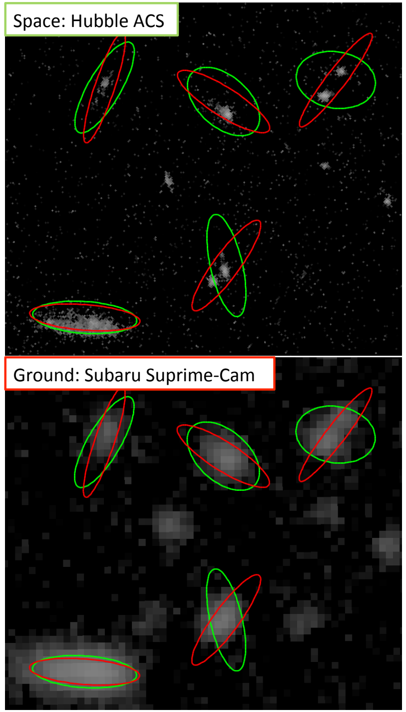
arXiv:1411.2608

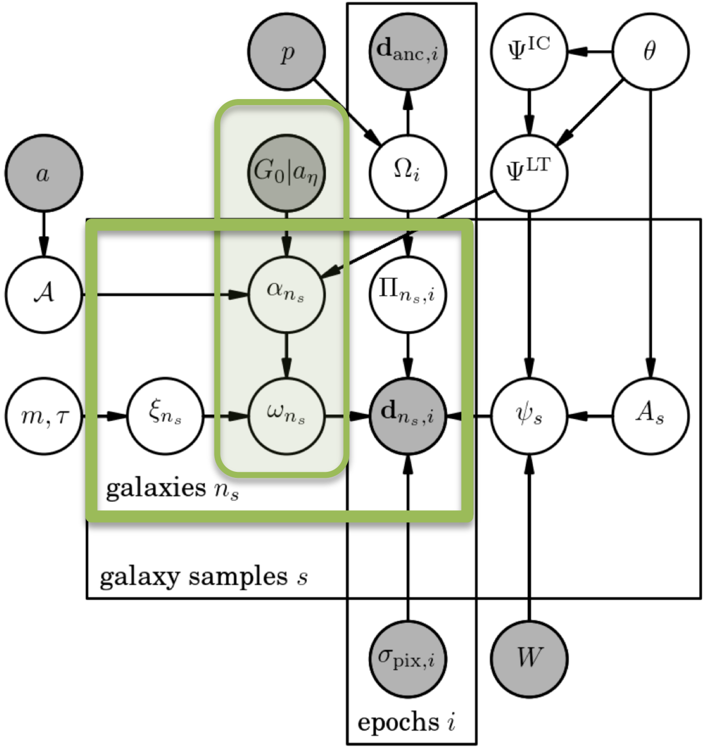
Parameter	Description
$\boldsymbol{\theta}$	Cosmological parameters
$\Psi^{\rm IC}$	Initial conditons for the 3D gravitational potential
$\Psi^{\rm LT}$	Late-time 3D gravitational potential
ψ_s	2D lens potential (given source photo-z bin s)
A_s	Parameters for the line-of-sight source distribution
$\Pi_{n_s,i}$	PSF for galaxy n_s observed in epoch i
Ω_i	Observing conditions in epoch i
$\{\omega_{n_s}\}$	Galaxy model parameters; $n_s = 1, \ldots, n_{{\rm gal},s}$
$\{\alpha_{n_s}\}$	Parameters for the distribution of $\{\omega_{n_s}\}$
$\{\xi_{n_s}\}$	Scaling parameters for $\{\omega_{n_s}\}$
m, τ	Hyperprior parameters for $\{\xi_{n_s}\}$
$\mathcal{A}$	Hyperparameter for $\{\alpha_{n_s}\}$ classifications
$\{\mathbf{d}_{n_s,i}\}$	Pixel data for galaxies $n_s = 1, \ldots, n_{{\rm gal},s}$ in epoch i
$G_0 \vert a_\eta$	Prior specification for $\{\alpha_{n_s}\}$
s	Source sample (e.g., photo-z bin)
W	Survey window function
$\mathbf{d}_{{\rm anc},i}$	Ancillary data for PSF in epoch i
p	Prior params. for observing conditions
a	Prior params. for $\mathcal{A}$
$\sigma_{{\rm pix},i}$	Pixel noise r.m.s. in epoch i
I	Model selection assumptions



Naïve approach is intractable

1. Galaxy models be joint fitted to all available epochs i

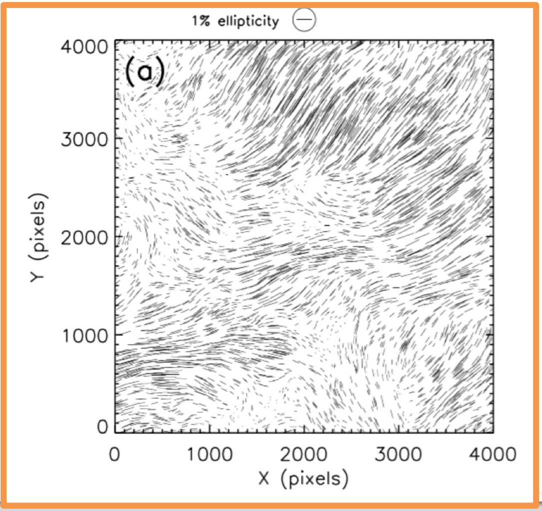


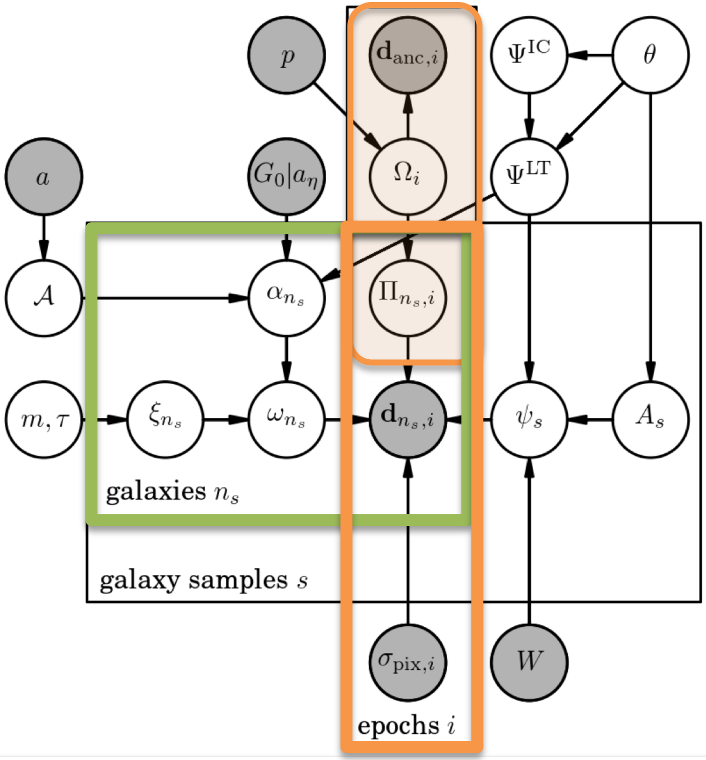


Naïve approach is intractable

1. Galaxy models be joint fitted to all available epochs i

2. PSF models must be joint fitted to all galaxies in an exposure n_s

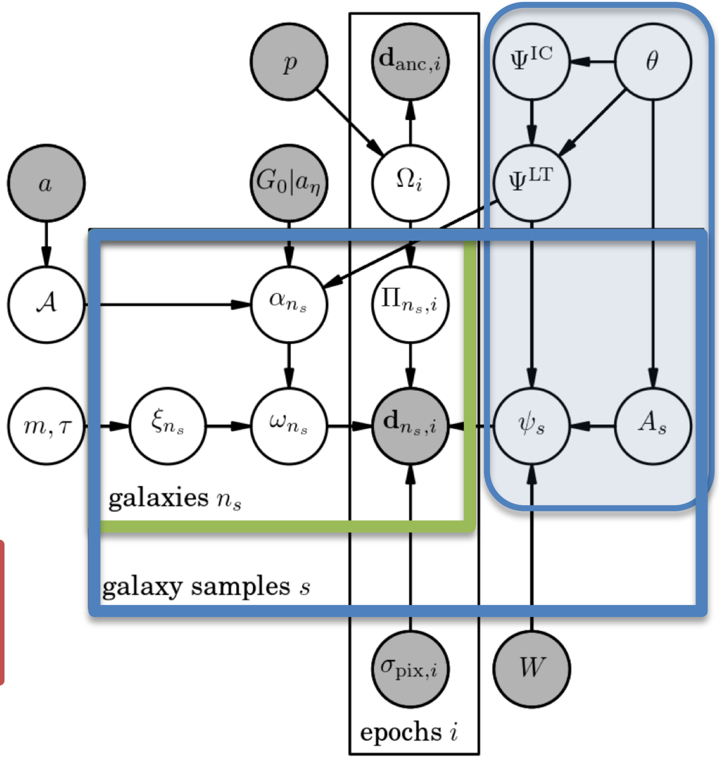




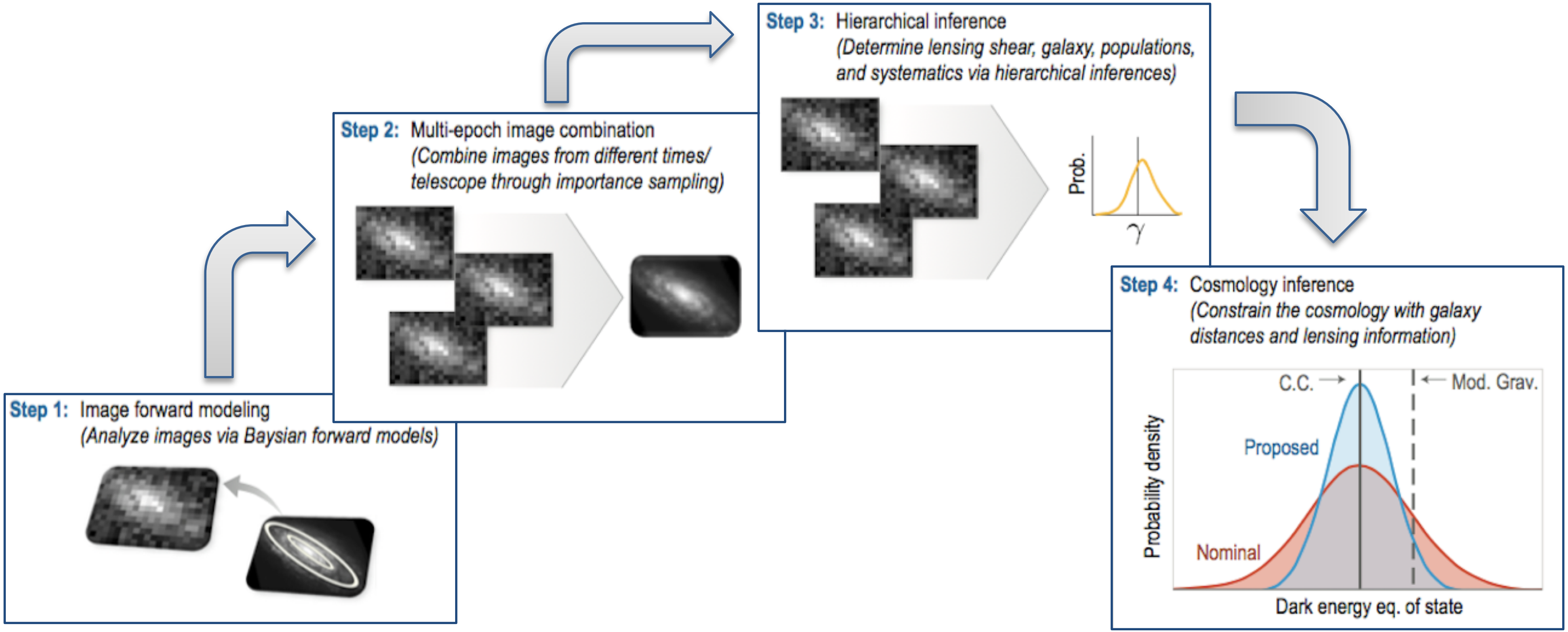
Naïve approach is intractable

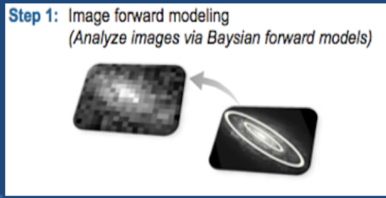
1. Galaxy models be joint fitted to all available epochs i
2. PSF models must be joint fitted to all galaxies in an exposure n_s
3. Cosmology must be joint fitted to all galaxy samples & epochs

The principled inference requires fitting all pixels of all surveys simultaneously



We have developed a tractable ‘divide & conquer’ computational approach for the complete statistical model for multiple imaging surveys

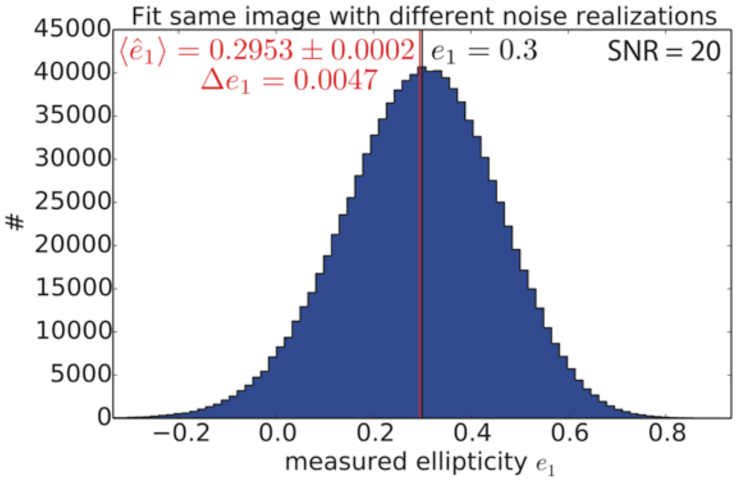


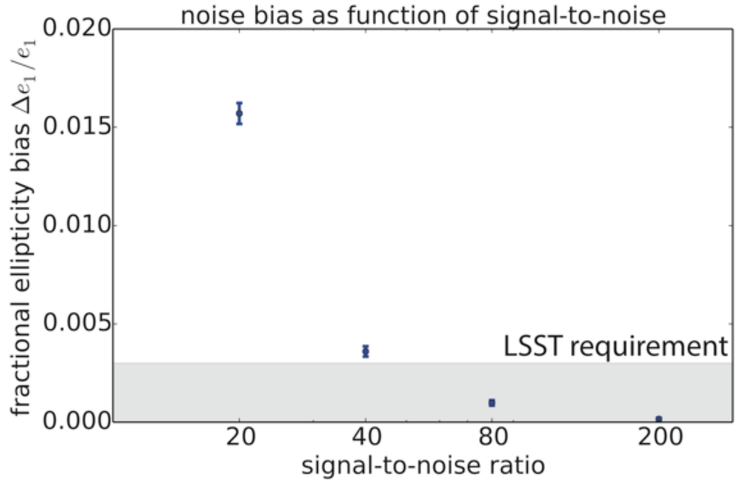


Step 1: Image forward modeling
Shear bias

Shape to Shear: Noise Bias

- Ellipticity: $e = \dfrac{a-b}{a+b}\exp(2i\theta)$
- Ensemble average ellipticity is an unbiased estimator of shear.
- However, maximum likelihood ellipticity in a model fit is **not** unbiased.
- Ellipticity is a non-linear function of pixel values.





Mitigating Noise Bias – at least 2 strategies

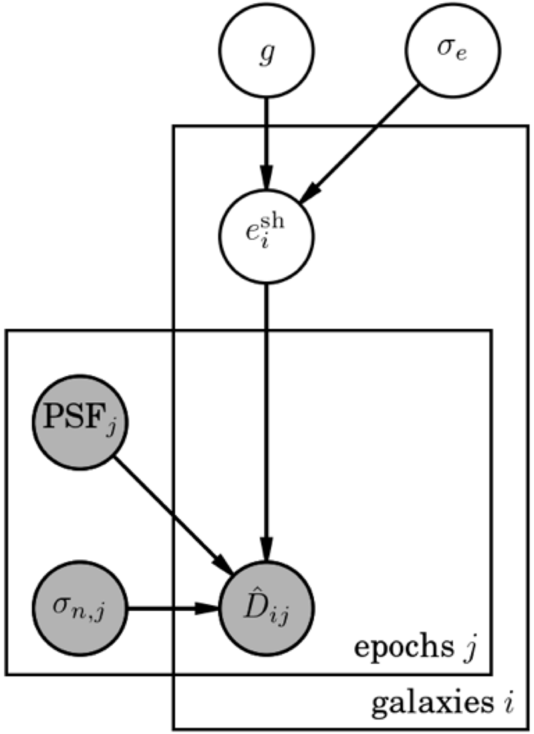
1. Calibrate using simulations. (im3shape, sfit)
   - But corrections are up to 50x larger than expected sensitivity!

2. Propagate entire ellipticity distribution function P(ellip | data)
   - Use Bayes' theorem: P(ellip | data) ∝ P(data | ellip) P(ellip)
   - Measure P(ellip) in deep fields. (lensfit, ngmix, BFD).
   - Infer simultaneously with shear in a hierarchical model. (MBI).

Metacalibration

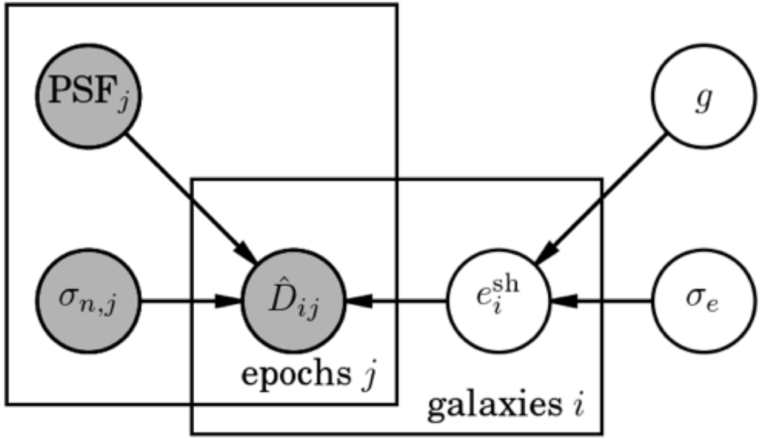
A hierarchical model for the galaxy distribution

- σ_e = intrinsic ellipticity dispersion
- e^{int} = galaxy intrinsic ellipticity
- g = shear
- e^{sh} = galaxy sheared ellipticity
- PSF = point spread function
- D = model image
- σ_n = pixel noise
- D = data: observed image



arXiv:1411.2608

Our graphical model tells us how to factor the joint likelihood



- Use a probabilistic graphical model to encode the factorization of the joint probability distribution of variables in the model.
- We don't care about e^{sh} for cosmology, so integrate it out.

$$\Pr\left(g, \sigma_e \middle| \{\mathrm{PSF}\}_j, \{\sigma_{n,j}, \{D_{ij}\}\}\right)$$

$$\propto \int d^{n_{\mathrm{gal}}} \{e_i^{\mathrm{sh}}\} \left[\prod_{ij} \Pr\left(D_{ij} | \mathrm{PSF}_j, \sigma_{n,j}, e_i^{\mathrm{sh}}\right)\right] \left[\prod_i \Pr\left(e_i^{\mathrm{sh}} | g, \sigma_e\right) \Pr(g) \Pr(\sigma_e)\right]$$

Huge complicated integral to compute for every posterior evaluation.

Importance Sampling allows tractable divide & compute : The pseudo-marginal likelihood

Ongoing research question:
How many interim samples are needed?

Want:

Galaxy dist. Likelihood

$$\Pr(\mathbf{d}|\alpha) \propto \boxed{\prod_{n=1}^{n_{\rm gal}}} \int d\omega_n \boxed{\Pr(\omega_n|\alpha)} \boxed{\Pr(\mathbf{d}_{n,i}|\omega_n)}$$

Have samples from:

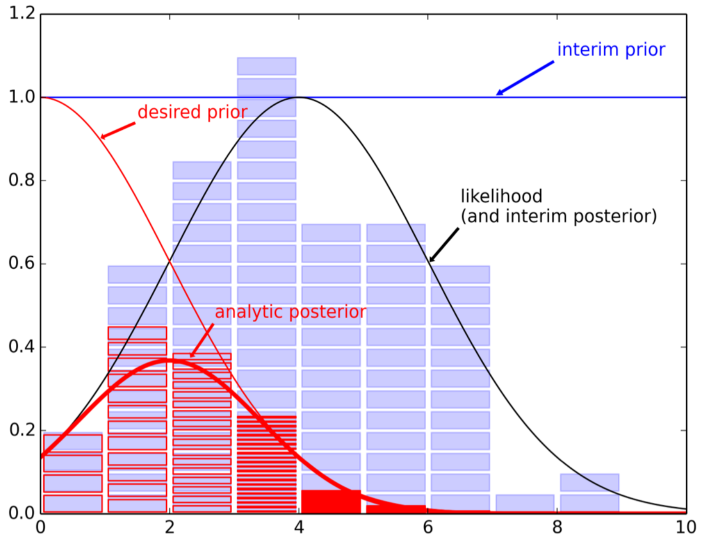
'Interim prior' specification

$$\boxed{\Pr\left(\omega_n|\mathbf{d}_n, I_0\right)}$$

Importance sampling:

$$\Pr(\mathbf{d}_n|\alpha) \approx \frac{Z_n}{K}\sum_k \frac{\Pr(\omega_{nk}|\alpha)}{\Pr(\omega_{nk}|I_0)},$$

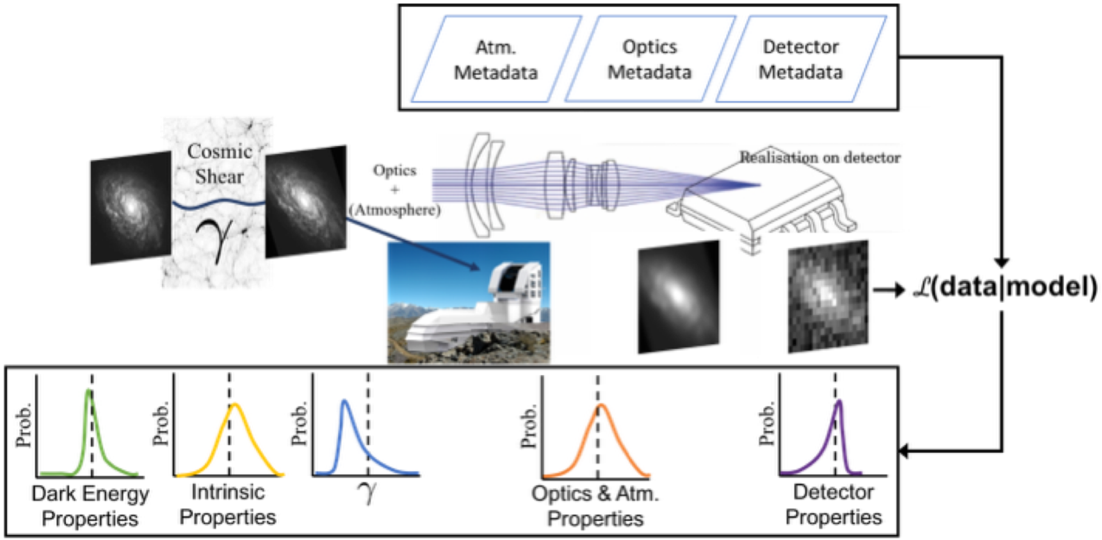
$$\Pr(\mathbf{d}|\alpha) = \prod_{n=1}^{n_{\rm gal}} \Pr(\mathbf{d}_n|\alpha).$$



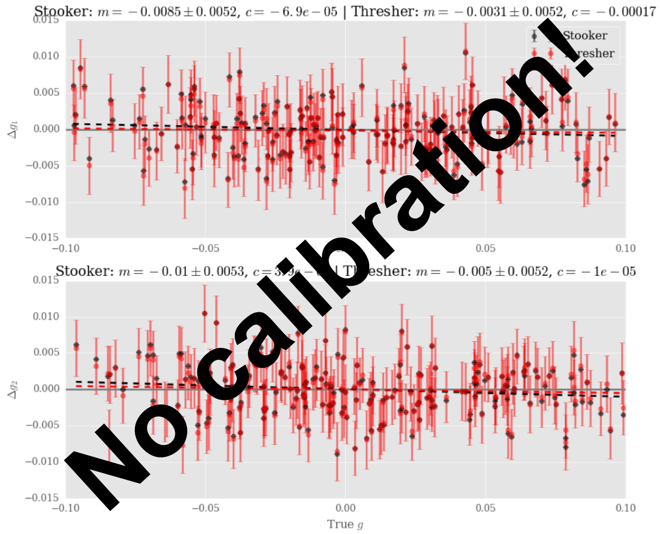
Credit: J. Meyers

Our hierarchical Bayesian forward models can meet LSST systematics tolerances for galaxy shear when the model is accurate enough

The forward model of our galaxy image data



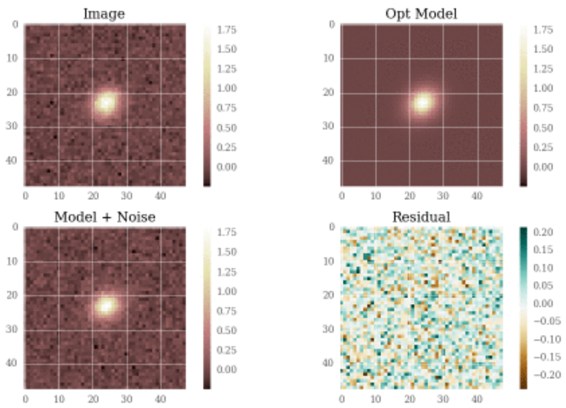
Our approach works:

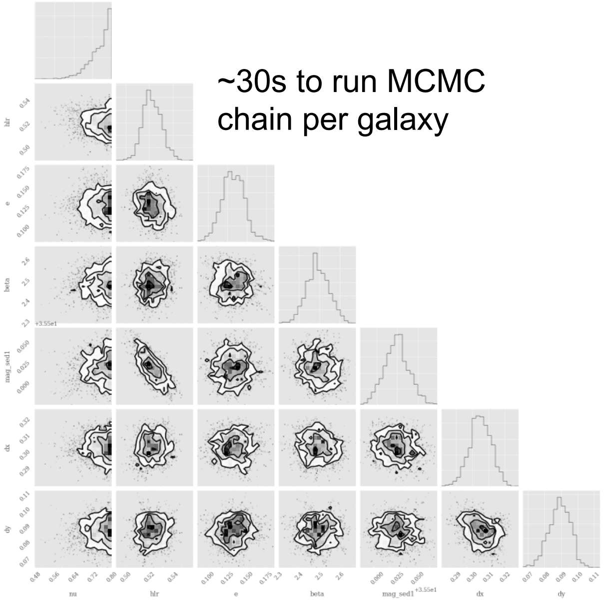


Sensitivity analyses with (simplified) galaxy simulation suites show biases well below LSST tolerances

Accurate modeling in an MCMC framework is more computationally demanding than traditional approaches – but still tractable

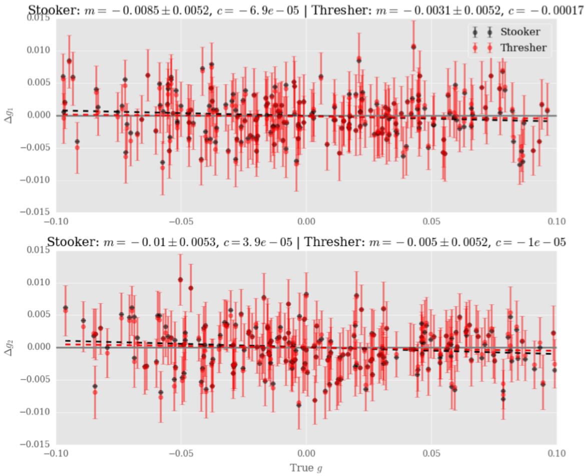
Every data point on the right is inferred from thousands of galaxy image model posterior inferences such as these:





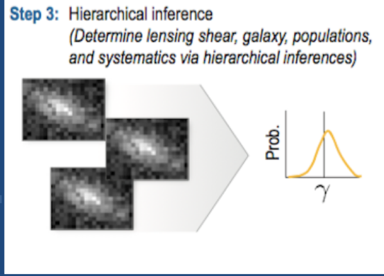
~30s to run MCMC chain per galaxy

Our approach works:



Sensitivity analyses with (simplified) galaxy simulation suites show biases well below LSST tolerances

LSST data volume: 4 billion galaxies, each seen 1000 times



Step 3: Hierarchical inference
Probabilistic cosmological one-point statistics

arXiv:1610.06673

Probabilistic cosmological mass mapping

with a maximum-entropy prior for computationally tractable probabilistic shear

Interpolate the unobserved lensing potential with GP

$$\psi_s \sim GP(0, \Sigma),$$

$\kappa, \gamma_1, \gamma_2$ are the second (spatial) derivatives of ψ_s

$$\mathrm{Cov}(\psi_{,ij}(\vec{x}), \psi_{,k\ell}(\vec{y})) = \Sigma_{,x_i x_j y_k y_\ell}(\vec{x}, \vec{y}).$$

GP kernels of $\kappa, \gamma_1, \gamma_2$ are linear combinations of the 4th (spatial) derivatives of the kernel of ψ_s

Zero E/B mode mixing by construction

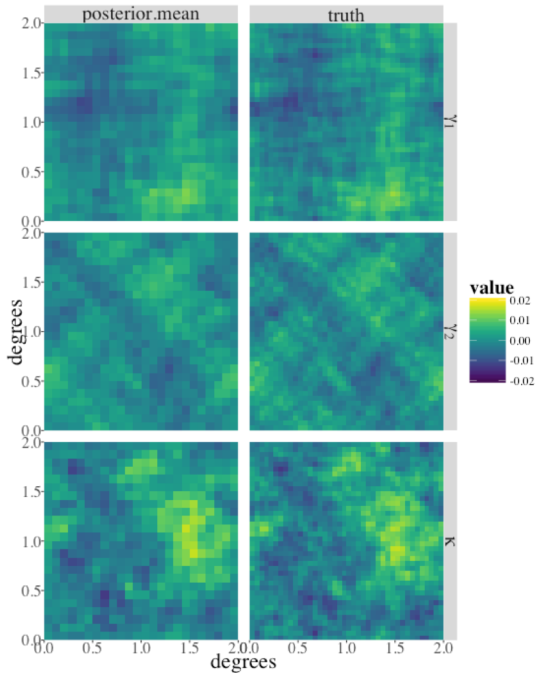
Objective: infer the late-time 3D gravitational potential and cosmological parameters

- Can make potential complex to model B-mode systematics
- IA models generate an additional (complex) potential
- Cosmology and systematics signals separable in a GP mixture

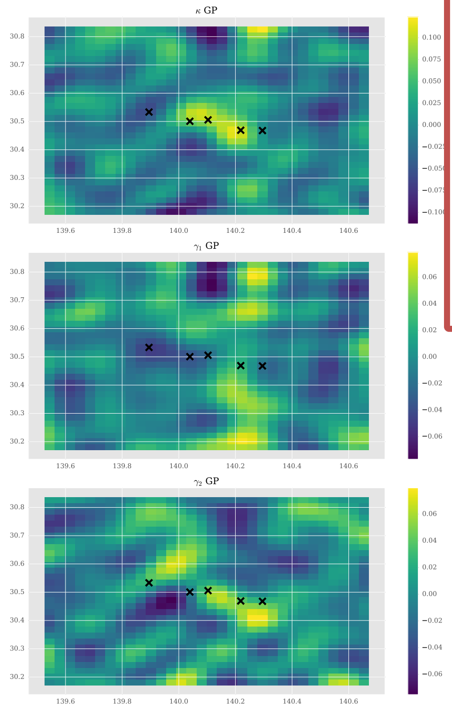
Hierarchical inference of cosmological lensing mass distributions

Akin to Wiener filter, but more appropriate for late-time mass density and systematics

<u>Validation with simulations</u>

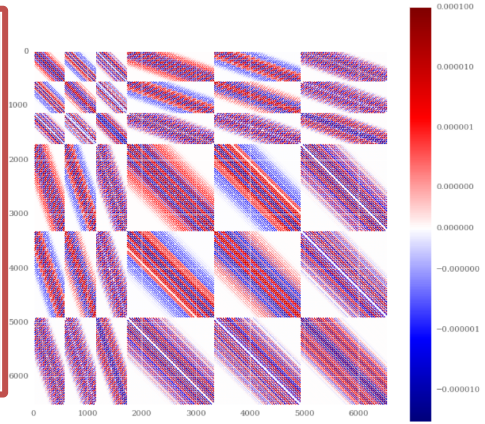


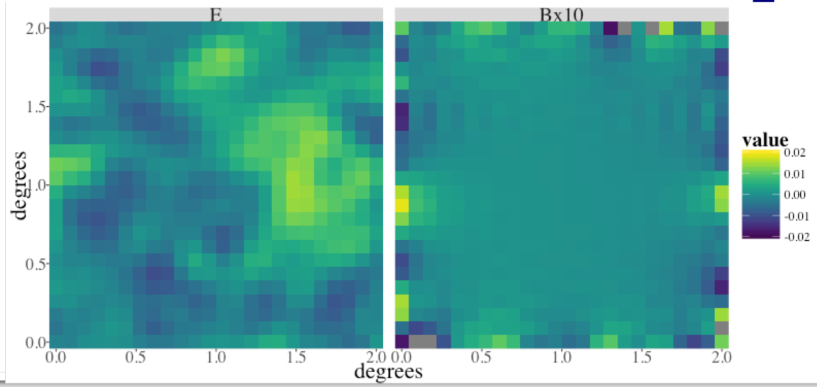
<u>A real merging galaxy cluster</u>



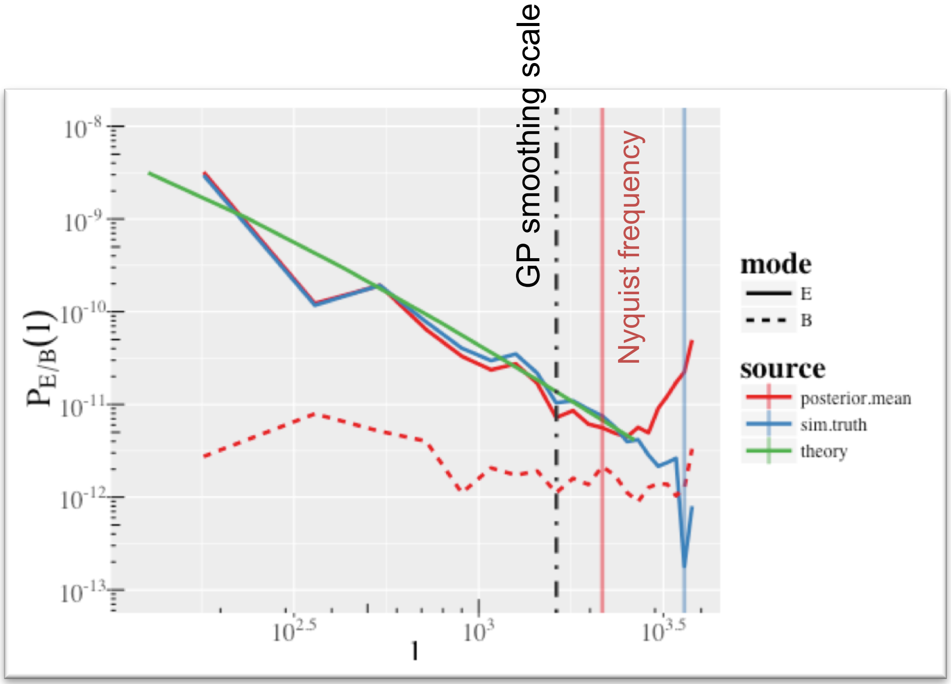
New:

- Linear and nonlinear scales reconstructed in one framework
- No E/B mode mixing by construction



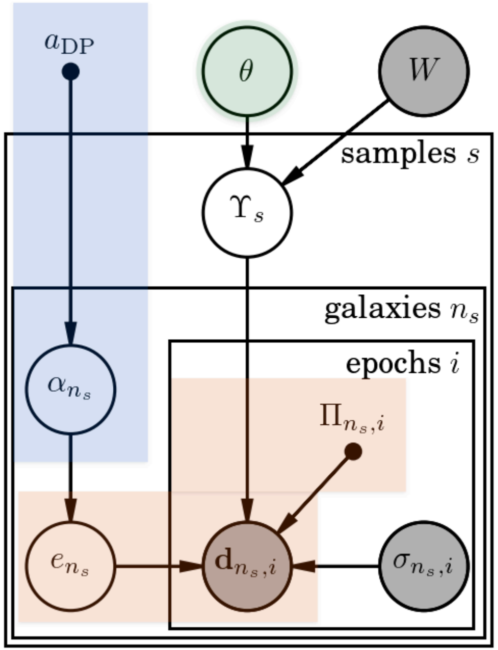


Lensing convergence maps with a GP prior yield accurate E/B mode power spectra

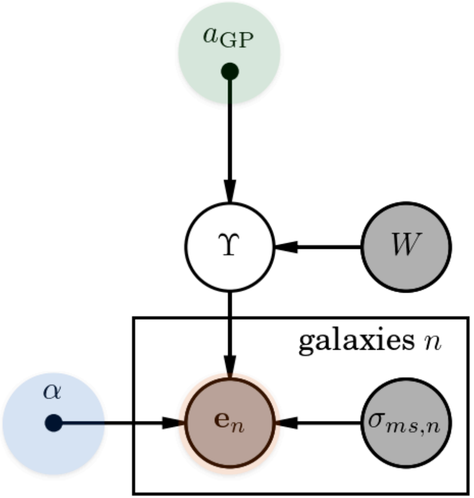


The GP prior for the lens potential can be used with existing galaxy ellipticity catalogs to explore new cosmological statistics

Simple model for pixel data



Approximate model for ellipticity catalog



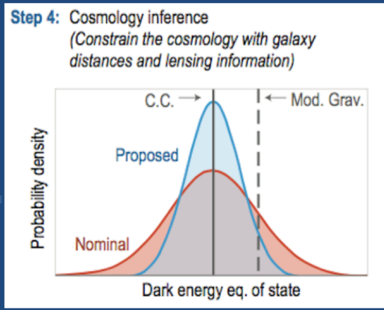
Sparse Gaussian Processes
(Adapted from S. Banerjee et al. arXiv:1406.7343)

- Let $\mathcal{S} = \{s_1, s_2, \ldots, s_k\}$ be a reference set of sky locations
  - Spatial random process: $(w(s_1), w(s_2), \ldots, w(s_k))^T \sim \mathcal{N}(0, K_\theta)$
  - Equivalent to auto-regressive process: $w(t) = \sum_{t=1}^{k} a_i(t) w(s_i) + \eta(t)$
- Independent 'noise': $\eta(t) \sim \mathcal{N}(0, \tau^2(t))$
- The $a_i(t) \neq 0$ **only if** t is a neighbor of s_i

Stores n small m x m matrices
Flop count is linear in n

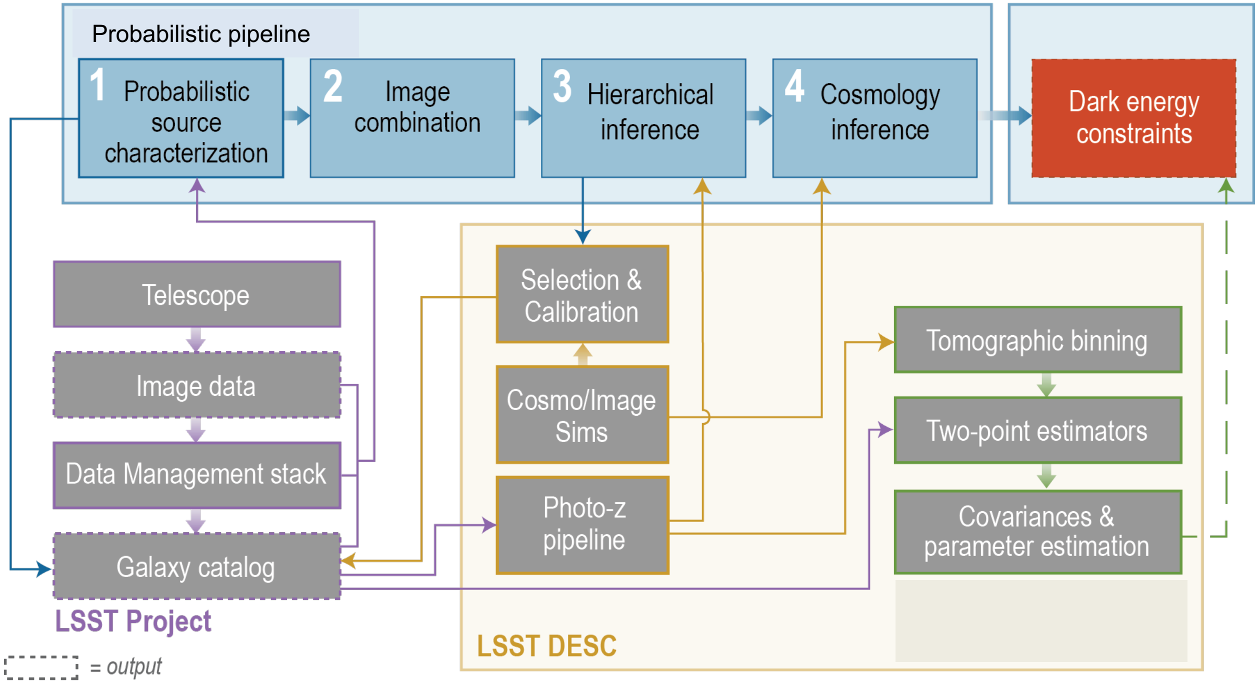
- Equivalent to the Generalized Cholesky Decomposition

See Daniels & Pourahmadi 2002
https://www.jstor.org/stable/4140601



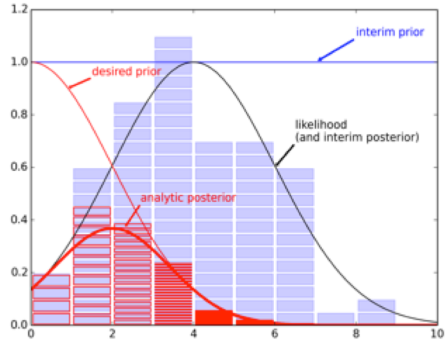
Step 4: Cosmology inference

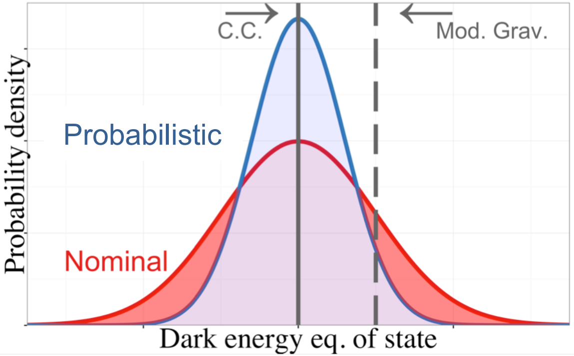
A probabilistic weak lensing workflow plan for LSST



Summary

- Cosmic shear is systematics limited & signal is dominated by PSF and astrophysics
  - A probabilistic approach is warranted to infer a small signal and mitigate biases
- A hierarchical probabilistic model for cosmic shear can trade bias for variance, but also can increase precision by learning latent structure in the galaxy distribution.
- Importance sampling methods allow tractable approaches to a probabilistic forward model of LSST imaging
  - With billions of galaxies and hundreds of epochs per galaxy modeling LSST imaging requires an approach to separating analyses of data subsets, even though statistically correlated
- We are able to sample from a probabilistic model with multiple hierarchies to marginalize both correlated image systematics and astrophysical properties of galaxies.





Lawrence Livermore National Laboratory