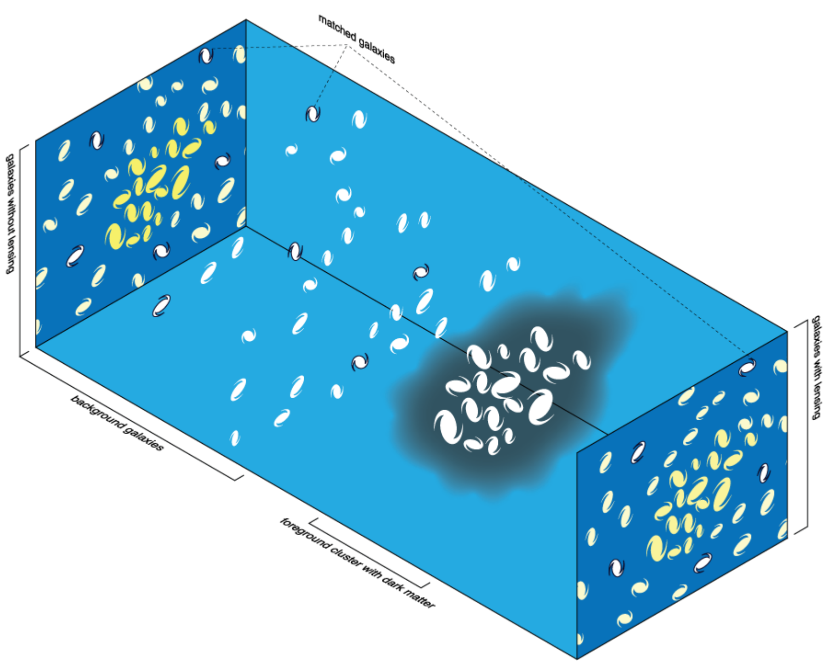
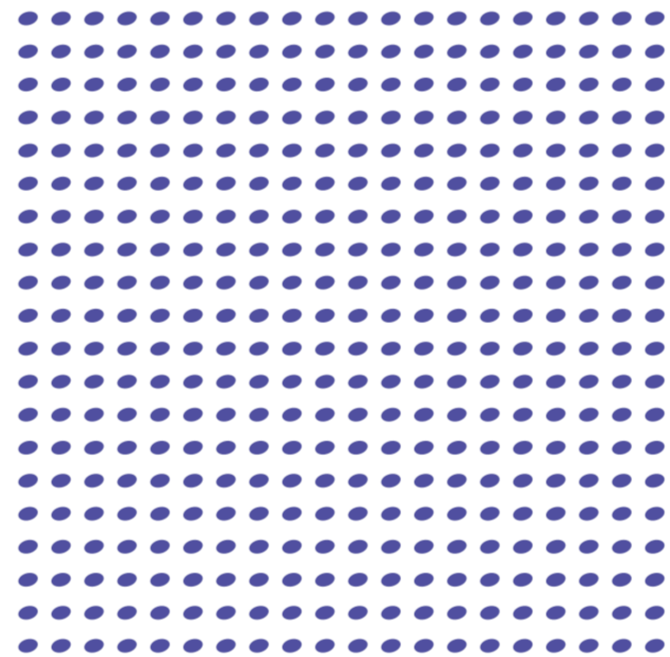
Hierarchical Probabilistic Inference of Multivariate Galaxy Properties

Karan Shah
Michael D. Schneider
(Lawrence Livermore National Lab)

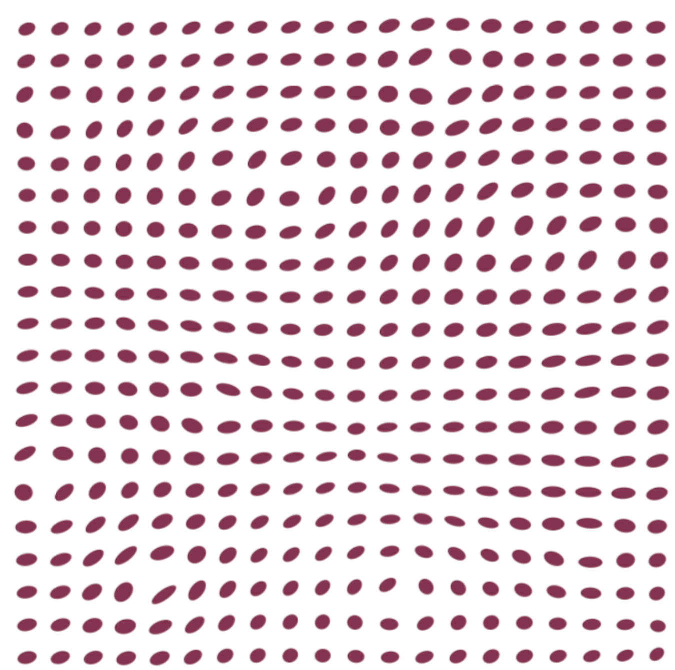
Bay Area LSST & Machine Learning Workshop
12/20/2018

Gravitational Lensing





Unlensed sources



Weak lensing

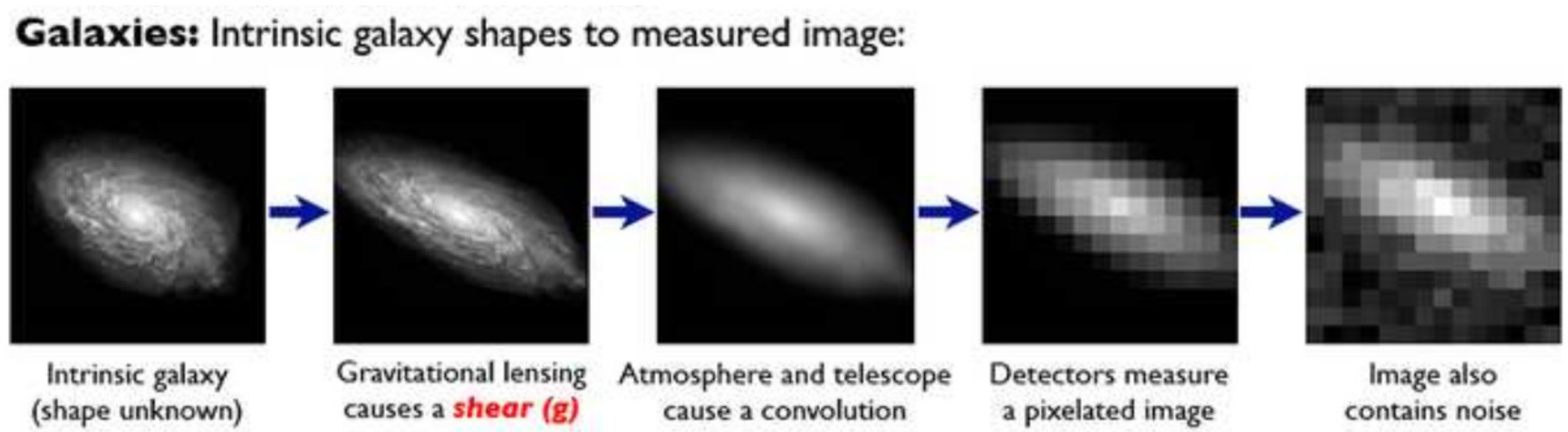
Cosmic shear signal is comparable to ellipticity of the Earth, ~0.3%

- D. Wittman

Weak Shear Regime

$$\epsilon_{sh} \approx \epsilon_{int} + g$$

Hierarchical Bayesian Model



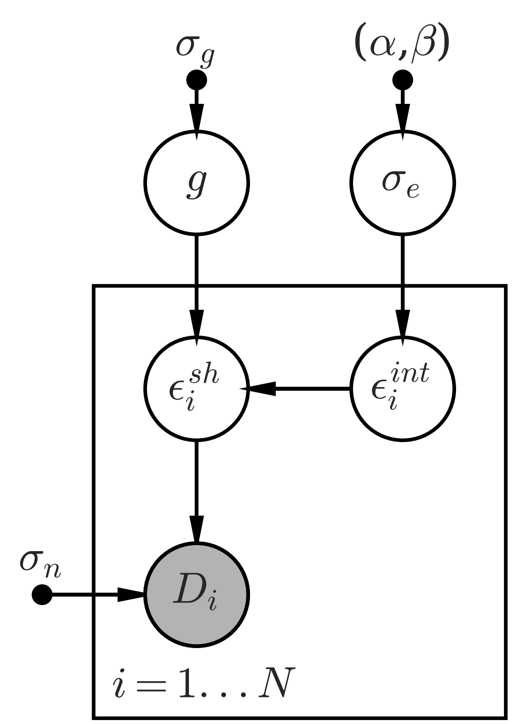
$$P(\gamma \,|\, \{D_i\}_1^N) \propto P(\{D_i\}_1^N \,|\, \gamma) P(\gamma)$$

Toy Models

$$P(D\,|\,\epsilon_{sh}, \sigma_n) = N_D(\epsilon_{sh}, \sigma_n^2)$$

$$P(\epsilon_{int}) = N_{\epsilon_{int}}(0, \sigma_e^2)$$

$$\epsilon_{sh} \approx \epsilon_{int} + g$$



Toy Models

For a single galaxy,

$$P(D|\epsilon_{int}, g) = P(D|\epsilon_{sh}(\epsilon_{int}, g))$$

We can marginalize out ϵ_{int}

$$P(D_n|g) \propto \int d\epsilon_{(int,n)} P(D_n|\epsilon_{(int,n)}, g) P(\epsilon_{(int,n)}|\sigma_e^2)$$

We specify an interim prior for ϵ_{sh} **as** $P(\epsilon_{sh}) \sim I$

$$P(D_i|\sigma_e^2) \propto \int d\epsilon_{sh} P(D_i|\epsilon_{sh}) P(\epsilon_{sh}|I) \frac{P(\epsilon_{int}(\epsilon_{sh}, g)|\sigma_e^2)}{P(\epsilon_{sh}|I)}$$

Toy Model

$$P(D_i\,|\,\sigma_e^2) \propto \int d\epsilon_{sh} P(D_i\,|\,\epsilon_{sh}) P(\epsilon_{sh}\,|\,I) \frac{P(\epsilon_{int}(\epsilon_{sh}, g)\,|\,\sigma_e^2)}{P(\epsilon_{sh}\,|\,I)}$$

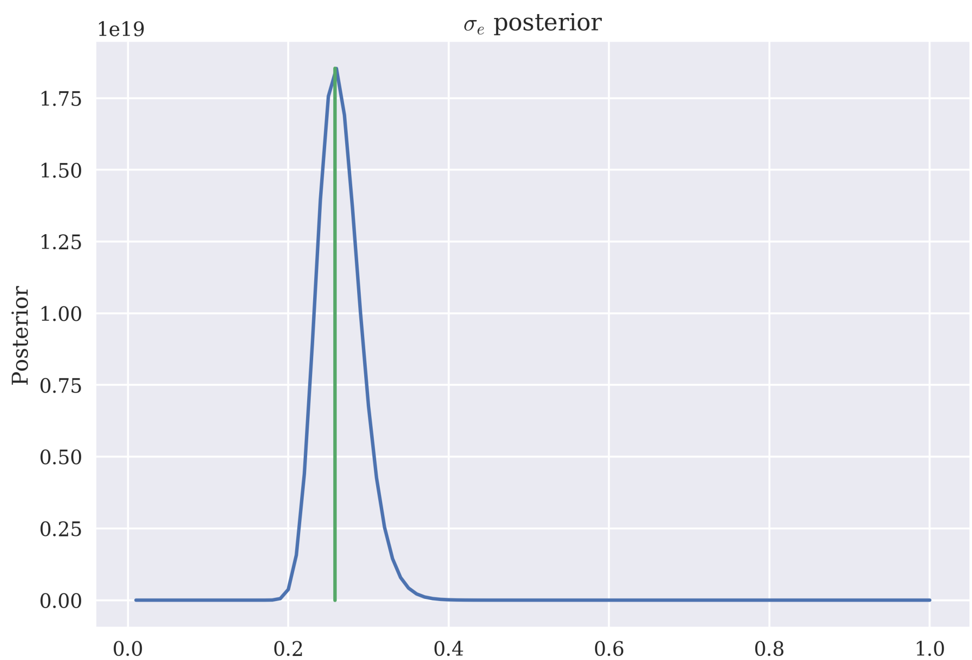
Importance Sampling:

$$P(D_i\,|\,\sigma_e^2) \propto \frac{Z_i}{K} \sum_{j=1}^{K} \frac{P(\epsilon_{int}(\epsilon_{sh,j}, g)\,|\,\sigma_e^2)}{P(\epsilon_{sh,j}\,|\,I)}$$

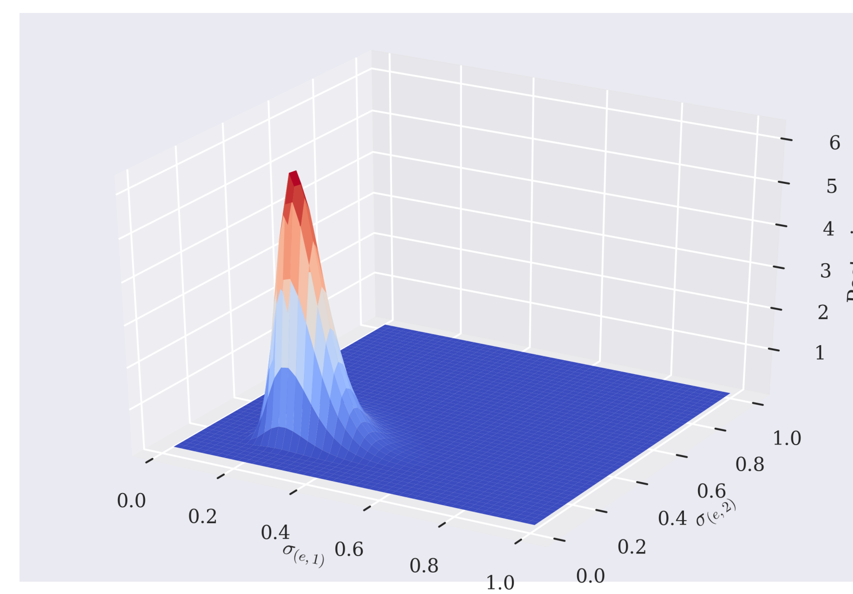
For N galaxies:

$$P(\{D_i\}_{i=1}^{N}\,|\,\sigma_e^2) \propto \prod_{i=1}^{N} \frac{Z_i}{K} \sum_{j=1}^{K} \frac{P(\epsilon_{int}(\epsilon_{sh,j}, g)\,|\,\sigma_e^2)}{P(\epsilon_{sh,j}\,|\,I)}$$

Toy Model

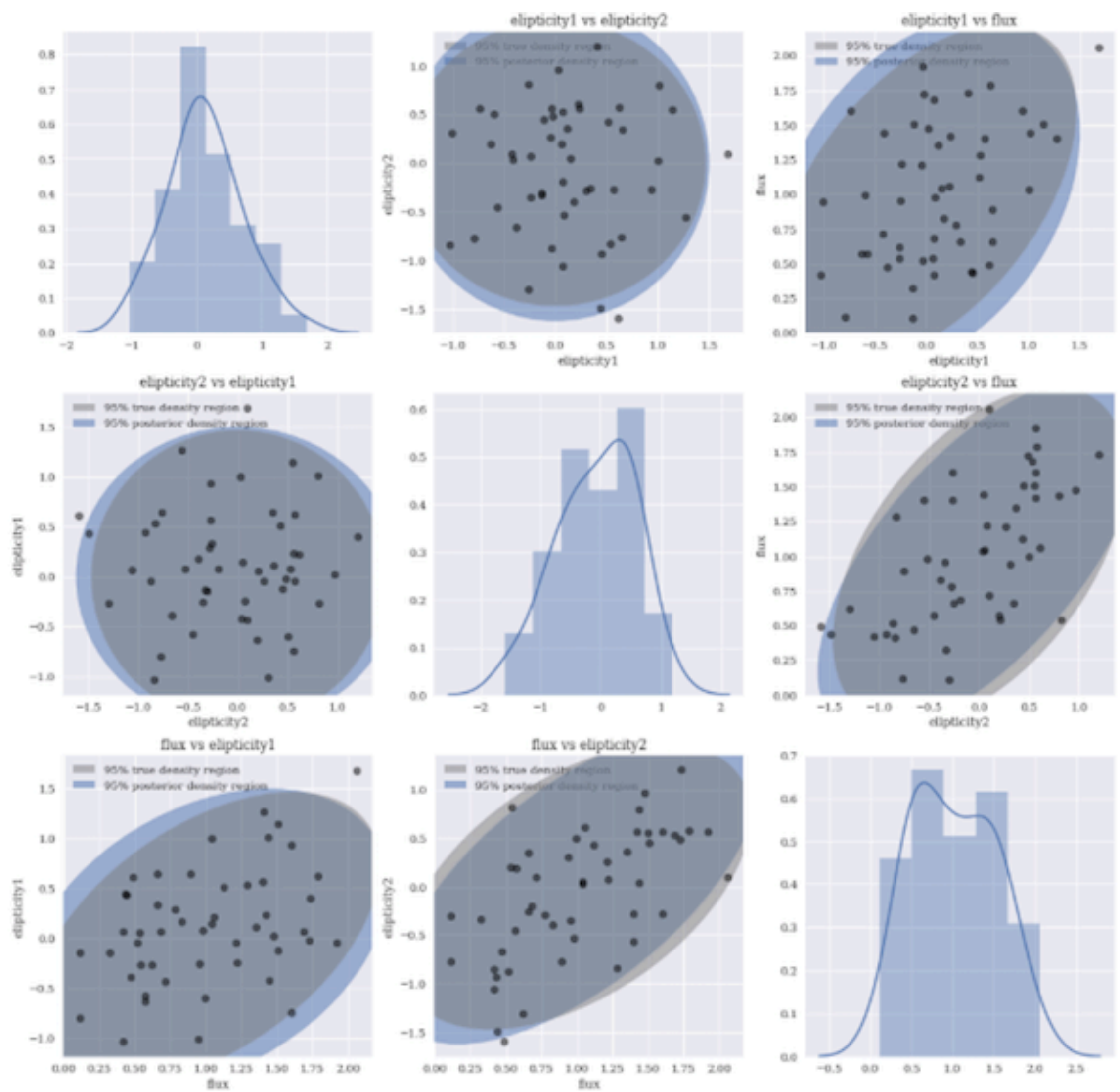


Estimation of $\sigma_{e,1}$ (Univariate)



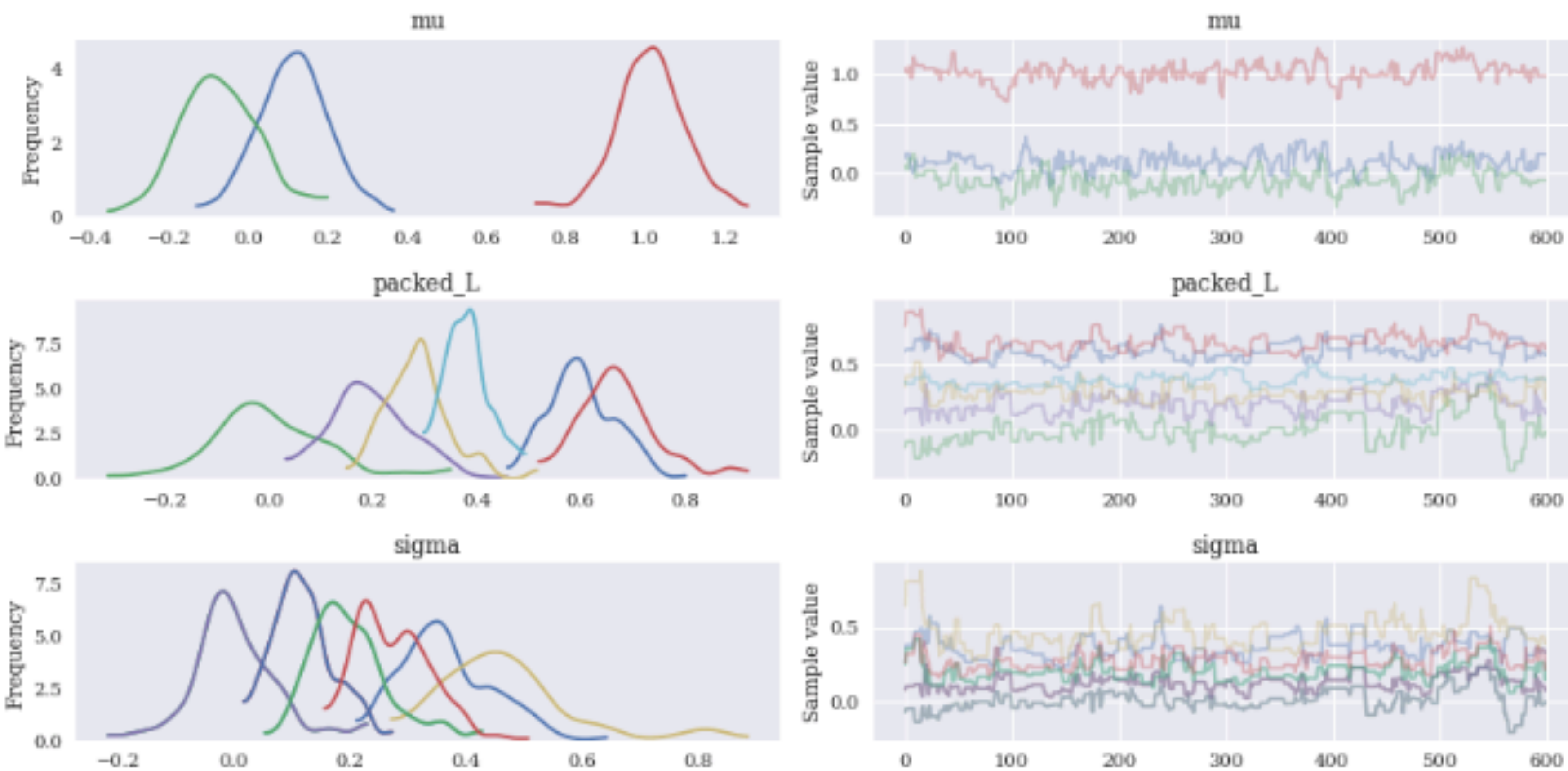
Estimation of $\sigma_{e,1}$ and $\sigma_{e,2}$(Bivariate)

Multivariate Toy Model with Covariance



Predicted Distribution vs Actual Distribution

Multivariate Toy Model with Covariance



MCMC distribution and chain

Including morphological properties

Intrinsic Properties ω

$$\omega = [\epsilon_1, \epsilon_2, \nu, r, \phi],$$

$$\epsilon_1 \sim N(0,\sigma_{e1})\,.$$

$$\epsilon_2 \sim N(0,\sigma_{e2})\,.$$

$$\nu \sim N(\mu_\nu, \sigma_\nu)\,.$$

$$r \sim lognorm(\mu_r, \sigma_r)\,.$$

$$\phi \sim powerlaw(\alpha)\,.$$

Including morphological properties

Non-Gaussian distributions transformed into approximately Gaussian distributions via,

$$g(r) = \log(r)$$
$$h(\phi) = \log(\frac{\phi^{\alpha}}{1-\phi^{\alpha}})$$

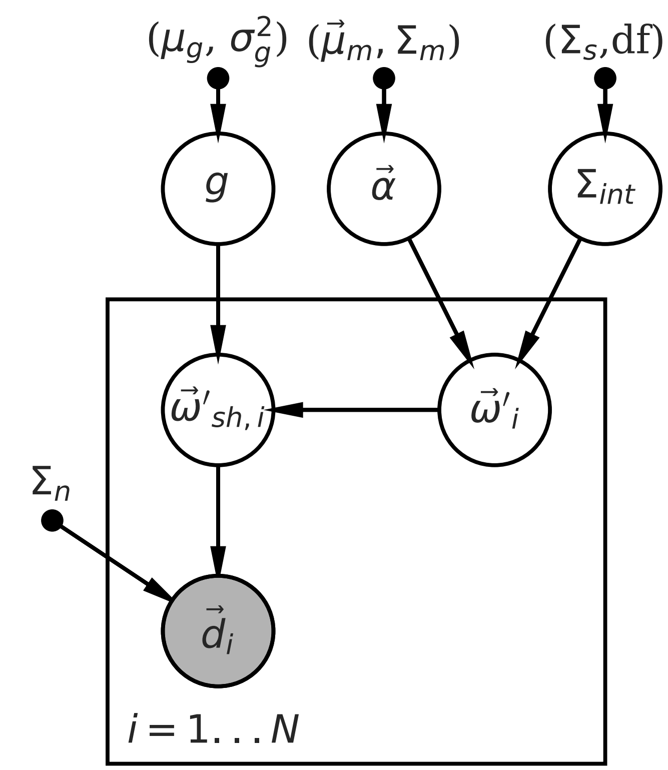
So the transformed data is,

$$\omega' = [\epsilon_1, \epsilon_2, \nu, g(r), h(\phi)]$$
$$\omega'_{sh} = [\epsilon_1 + g, \epsilon_2 + g, \nu, g(r), h(\phi)]$$

where, $\omega' \sim N(\alpha, \Sigma_{int})$

We define this transformation as

$$\omega' = f(\omega)$$



Including morphological properties

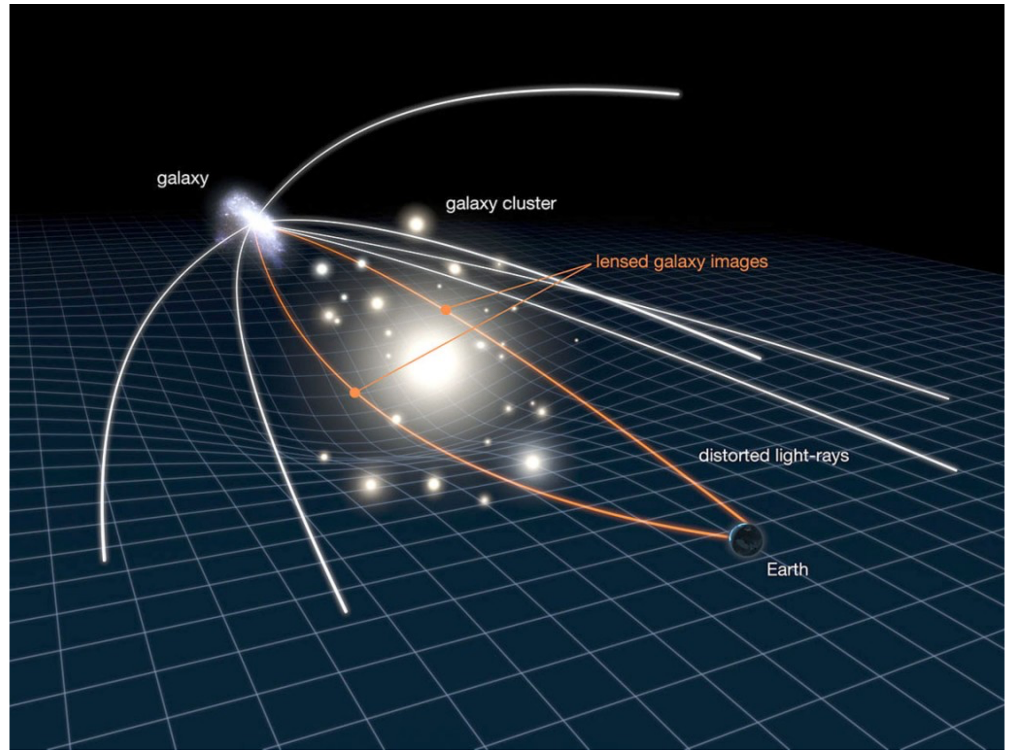
$$Pr(\{\mathbf{d}_i\}_{i=1}^{N}|\Sigma) \propto \prod_{i=1}^{N} \frac{Z_i}{K} \sum_{j=1}^{K} J^{-1}(\omega_{ij}) \frac{Pr(\omega'_{ij}(\omega'_{sh} = f(\omega_{sh}), g)\,|\,\alpha, \Sigma_{int}, g)}{Pr(\omega_{ij,sh}\,|\,I)}$$

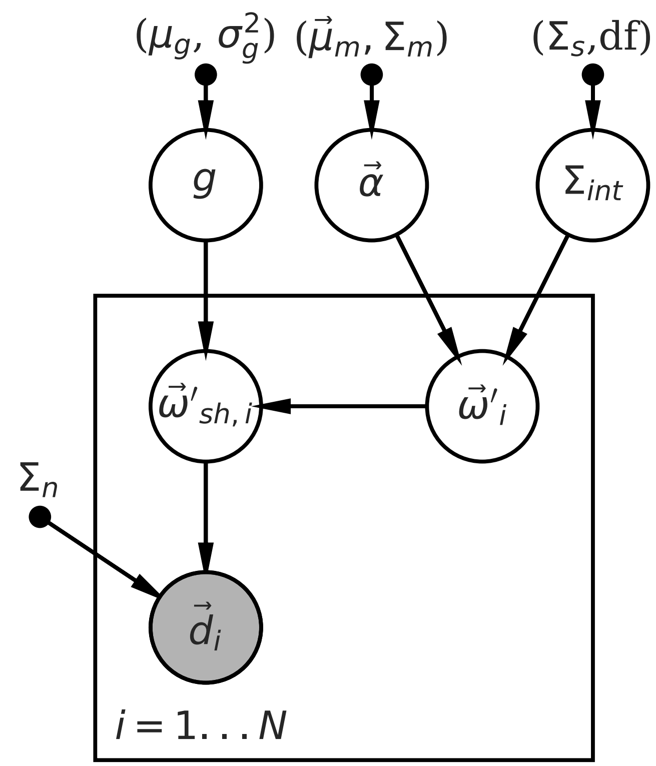
Current Challenges

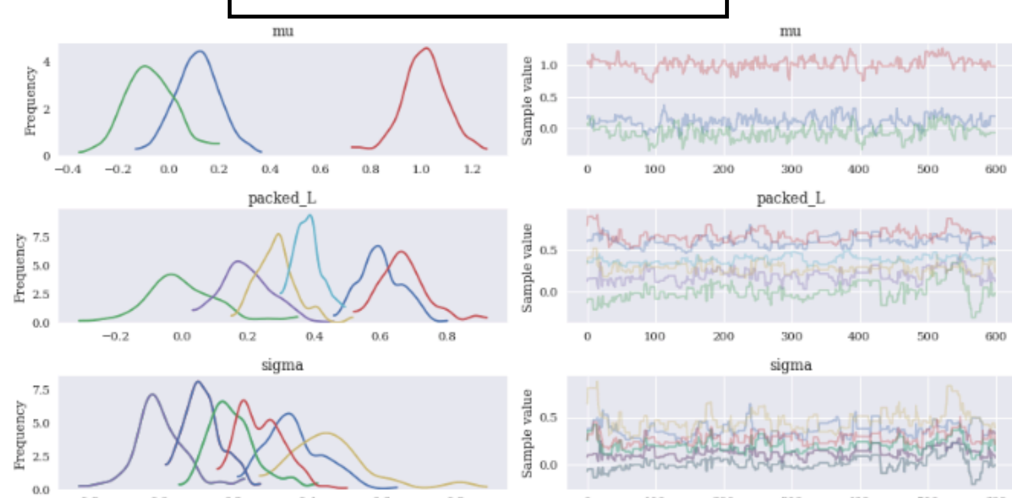
- Transformations from non-Gaussian to Gaussian
- Prior choice: LKJ vs Inv Wishart for Cov Matrix
- Other design factors: MCMC algorithm, interim sampling distributions etc.

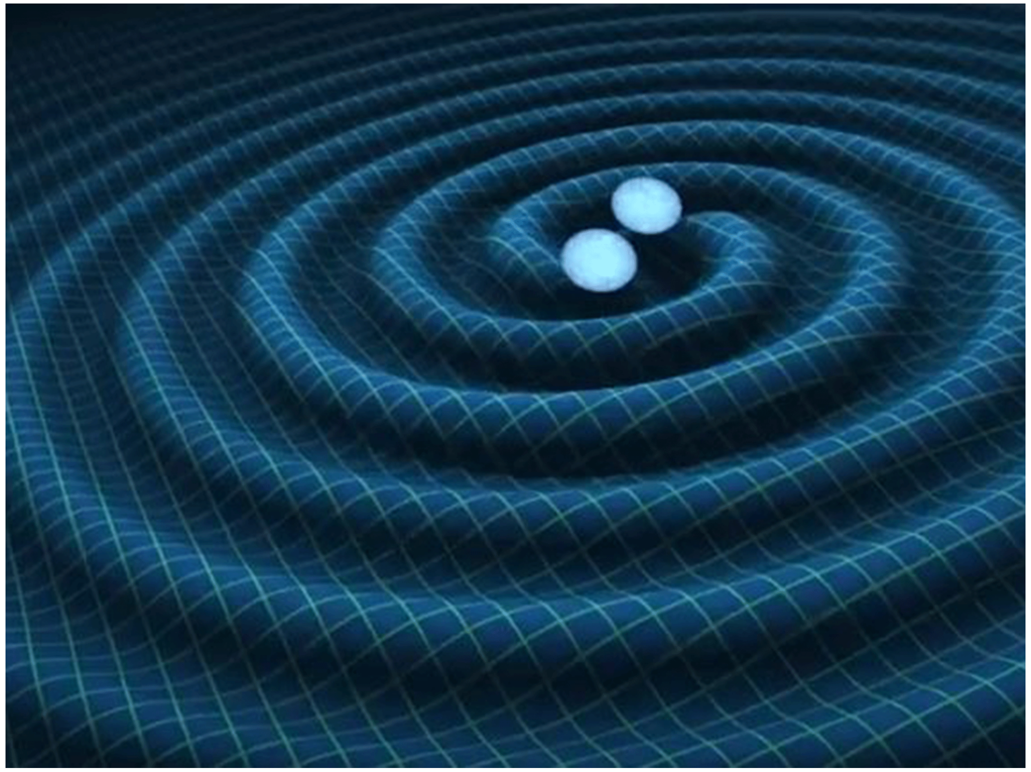
Future Work

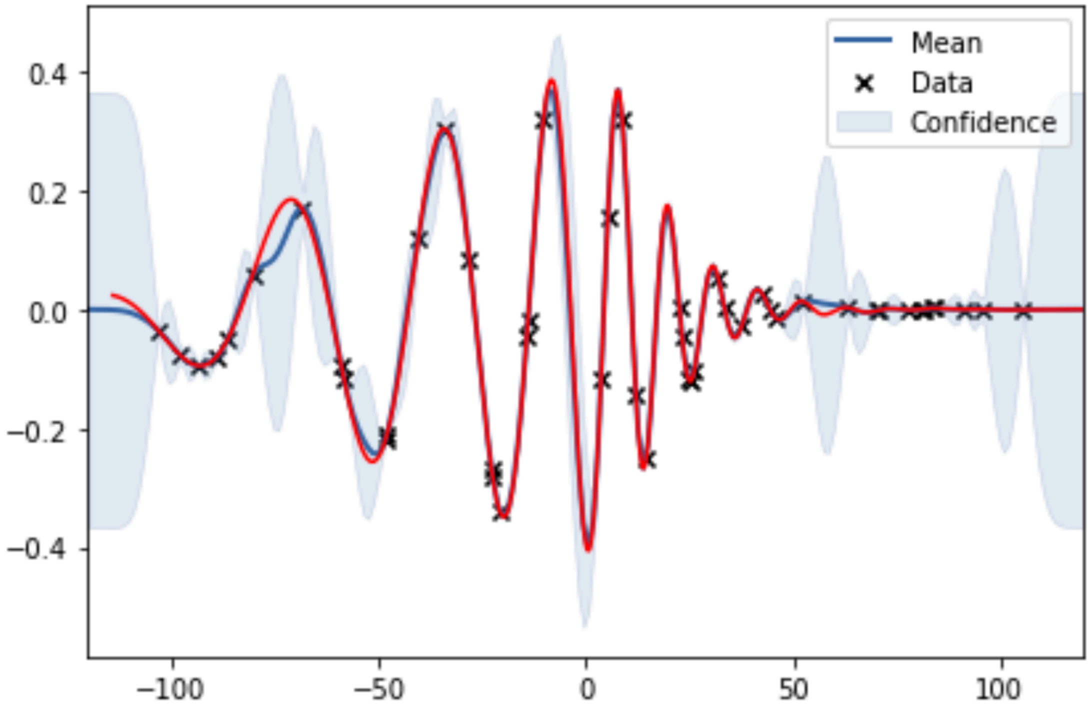
- Validate this approach on GalSim generated data
- Unbiased shear estimation through an extended model (Schneider et al. 2015)
- Determination of correlations in galaxy properties, morphological classification scheme


galaxy
galaxy cluster
lensed galaxy images
distorted light-rays
Earth

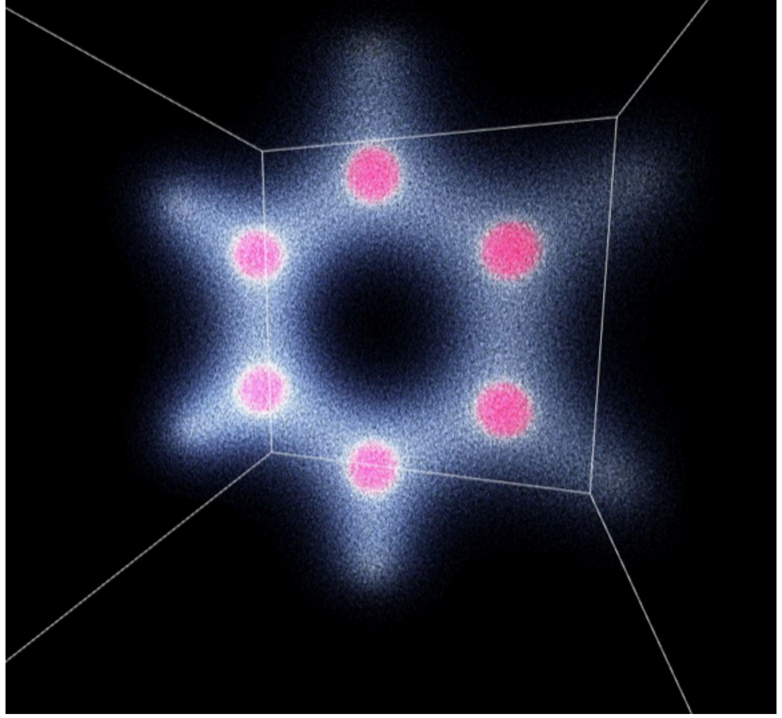

(μ_g, σ_g^2)
$(\vec{\mu}_m, \Sigma_m)$
(Σ_s, df)
g
$\vec{\alpha}$
Σ_{int}
$\vec{\omega}'_{sh,i}$
$\vec{\omega}'_i$
Σ_n
$\vec{d}_i$
$i = 1 \ldots N$

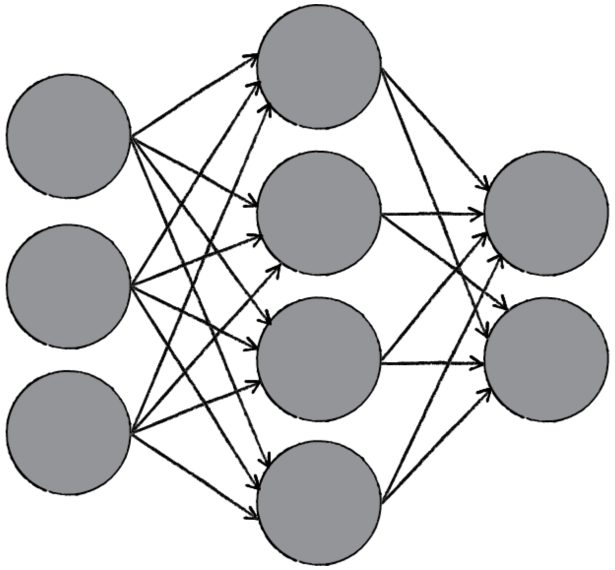

mu
Frequency
mu
Sample value
packed_L
Frequency
packed_L
Sample value
sigma
Frequency
sigma
Sample value





Mean
Data
Confidence
0.4
0.2
0.0
−0.2
−0.4
−100
−50
0
50
100






Uncertainty (std dev) with each residual point
residuals
4
3
2
1
0
−1
−2
−3
0
50
100
150
200
250
300

Georgia Tech

Lawrence Livermore National Laboratory