Inferring the Properties of Galaxies and the Universe with Hierarchical Dirichlet Process Prior

Will Dawson, Michael Schneider (LLNL), David Hogg (NYU), Phil Marshall, Debbie Bard, Josh Myers (SLAC), Dustin Lang (CMU)

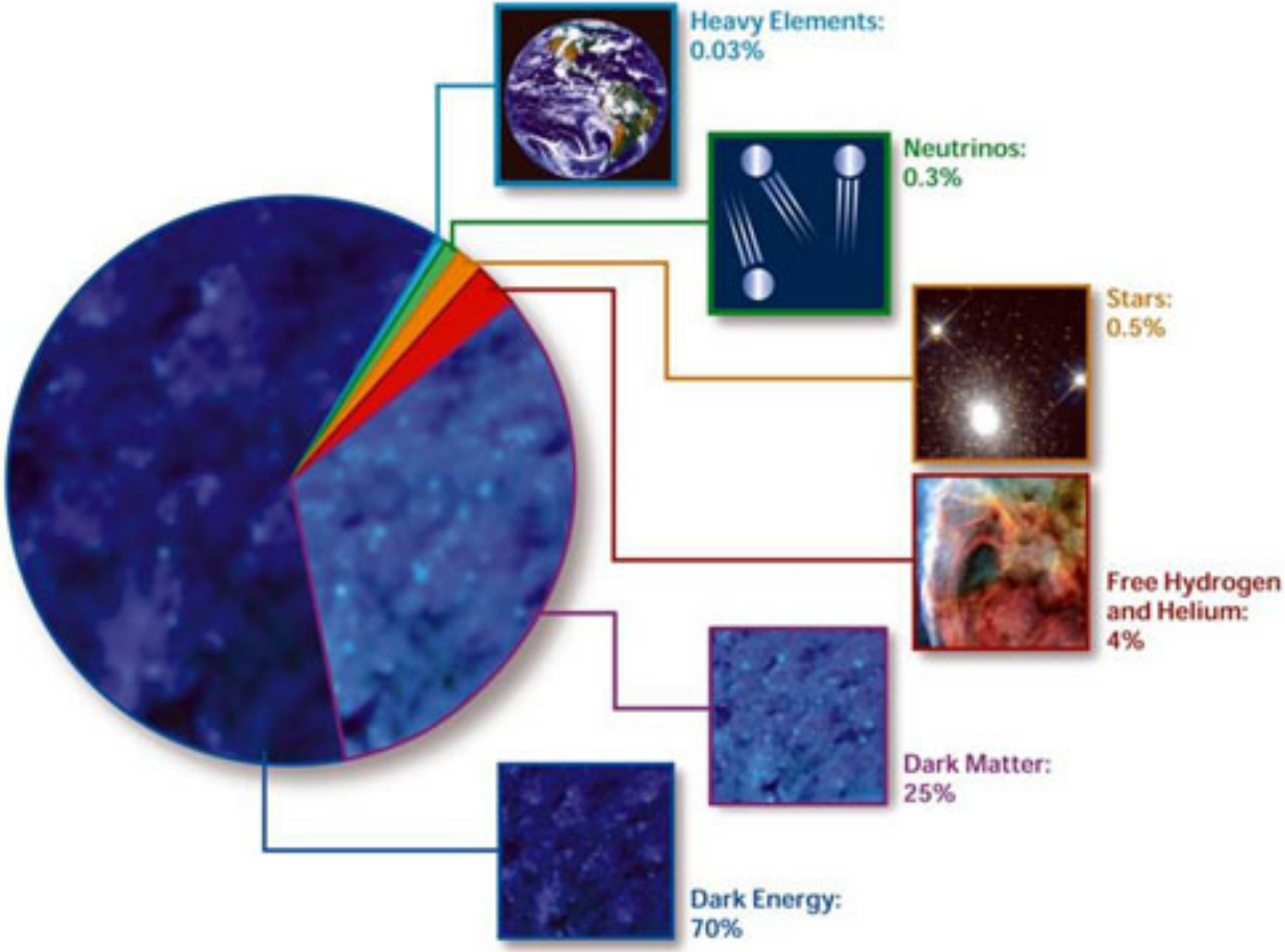
May 12 2015

Lawrence Livermore National Laboratory

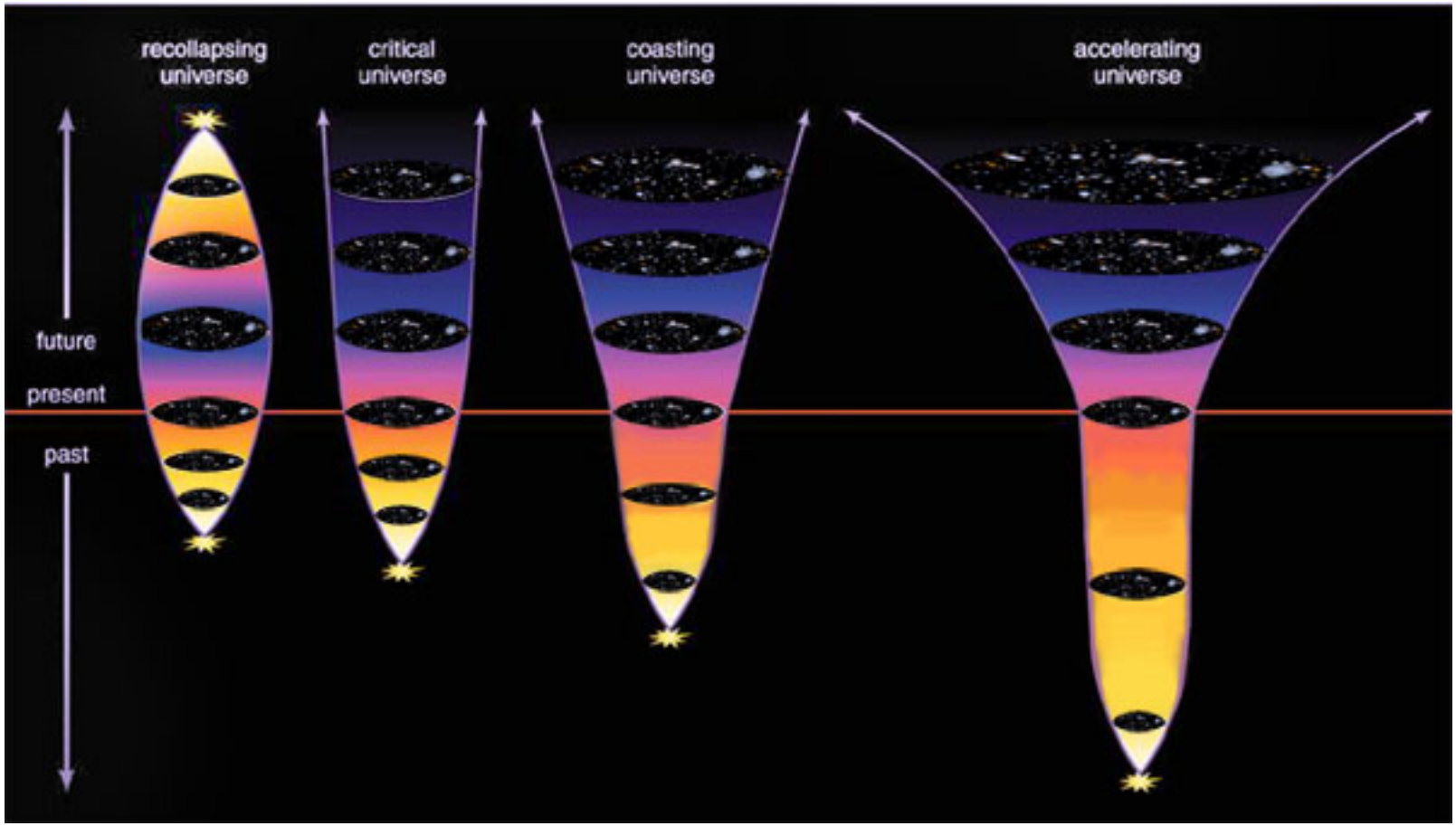
LLNL-PRES-669888

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

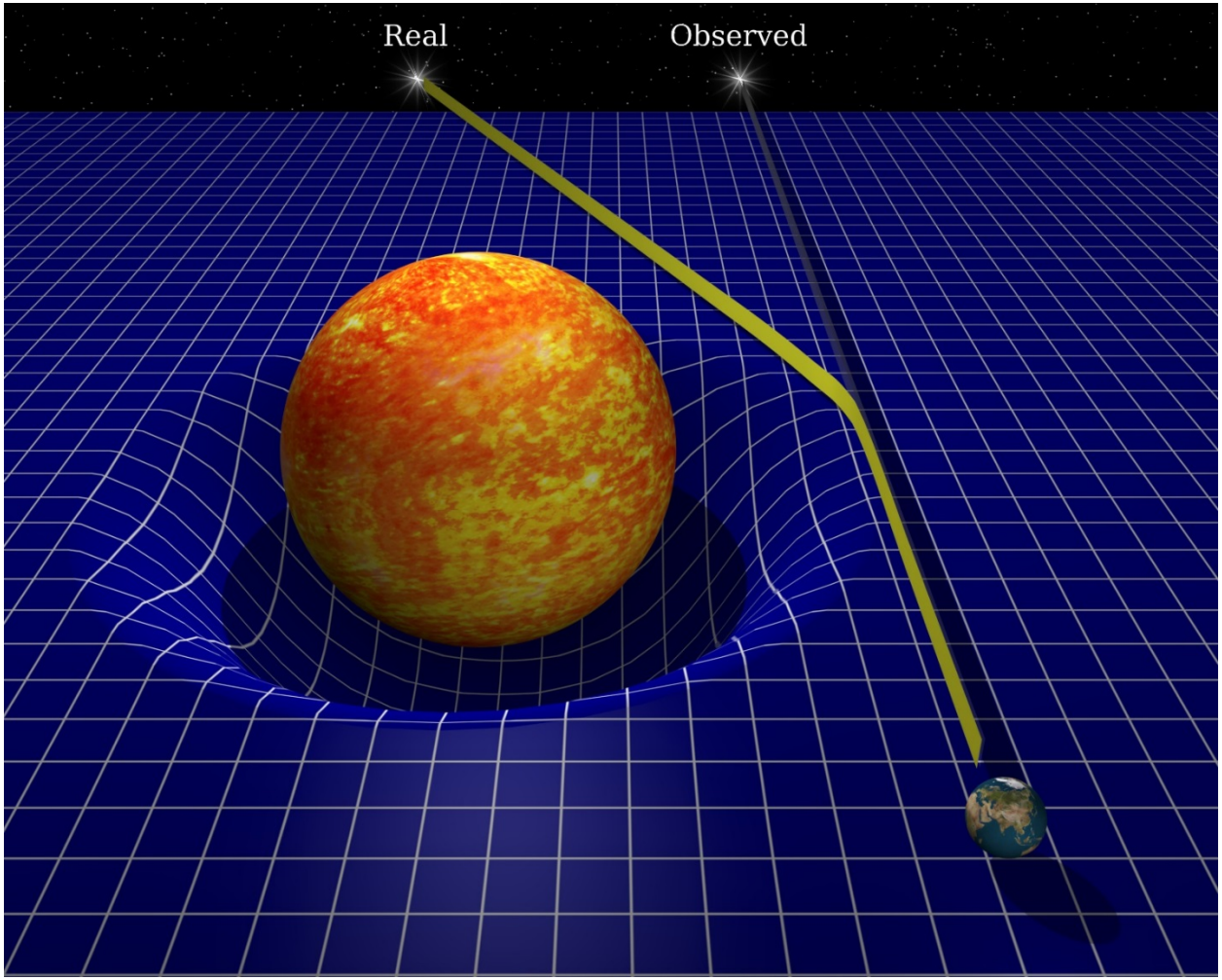
COMPOSITION OF THE COSMOS



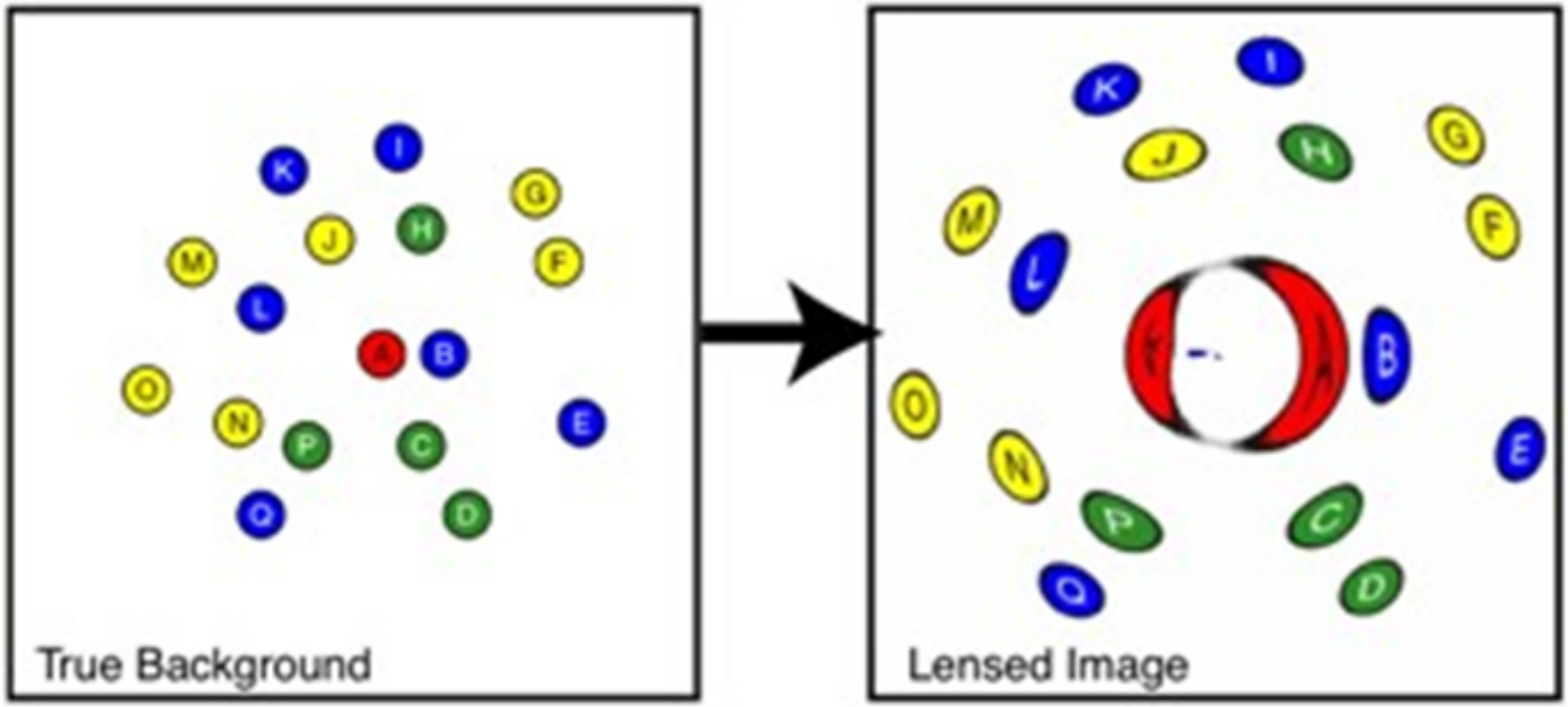
The fate of the Universe



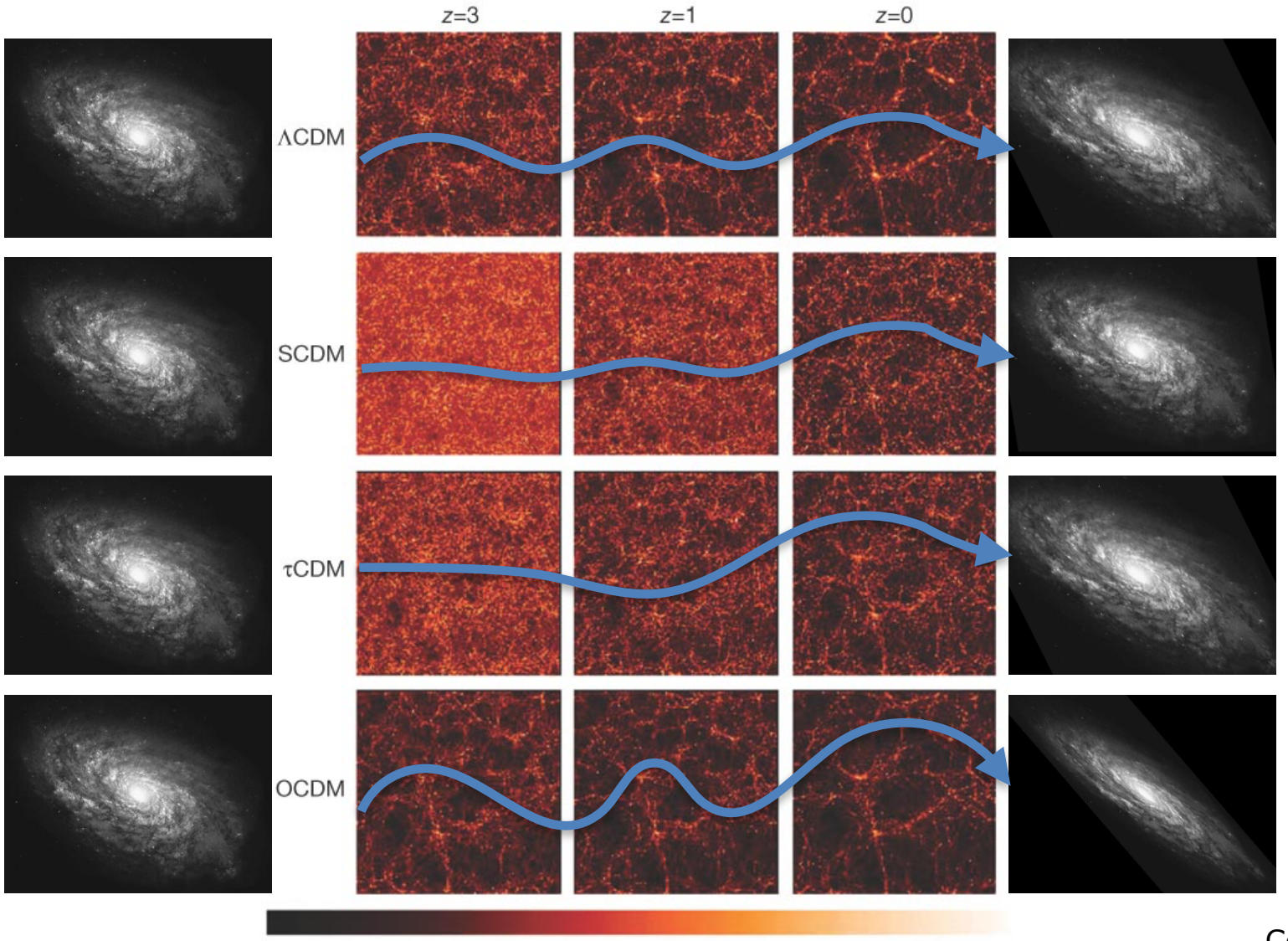
Mass warps space-time and alters the path of light



Gravitational lensing distorts the shapes of galaxies

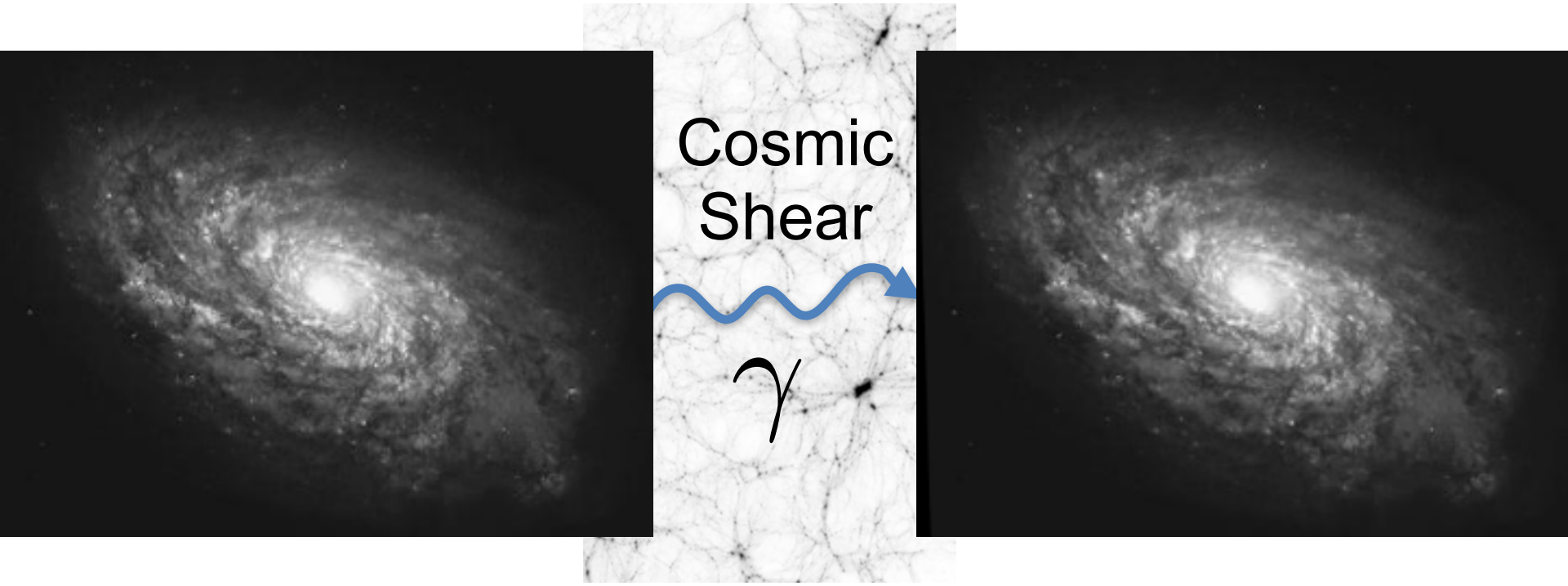


Different DE causes different shear



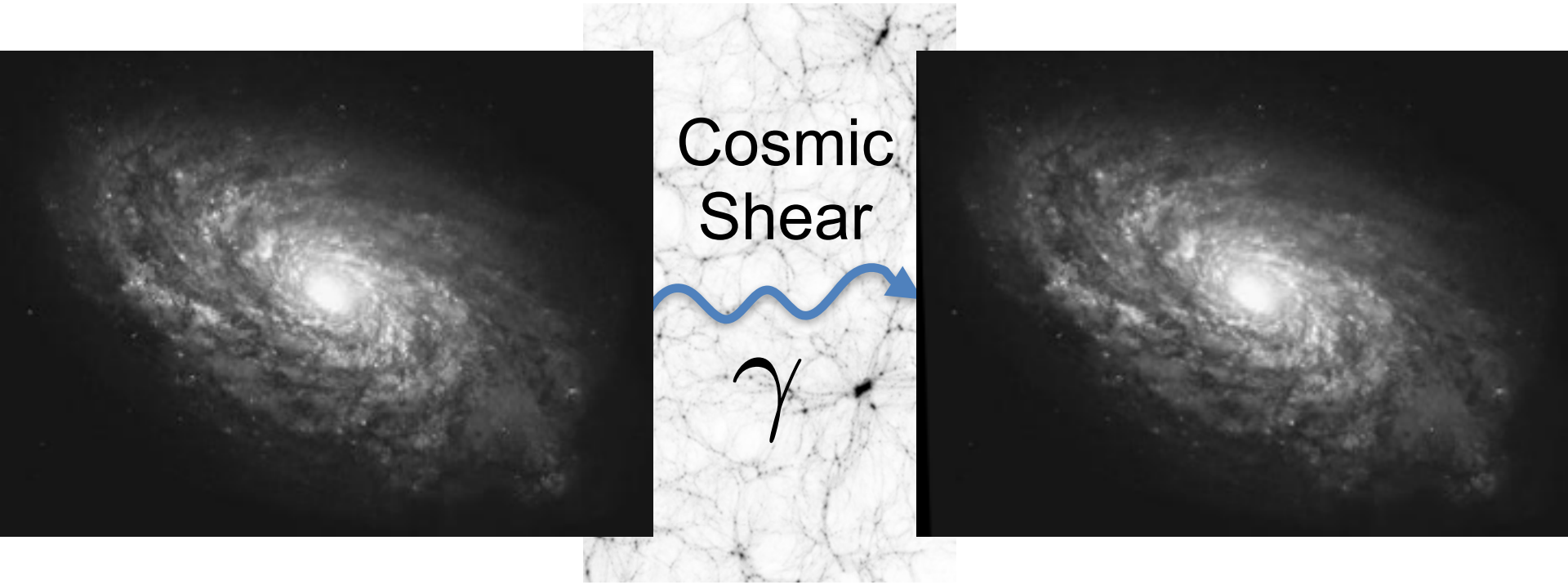
Coles 2005

Cosmic shear is a small signal



$$e_{\text{observed}} \approx e_{\text{intrinsic}} + \gamma$$

Small signal relative to shape noise
Thus need many galaxies

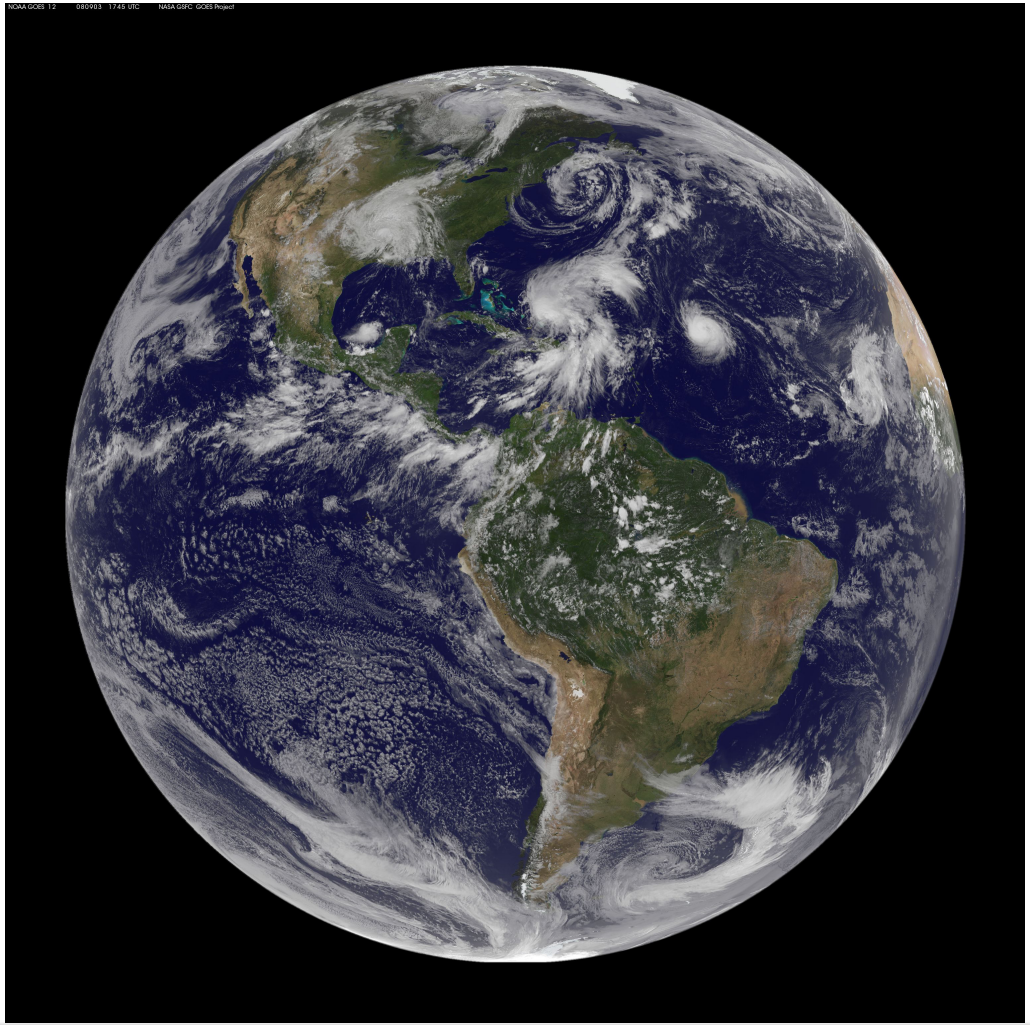


$$e_{\text{observed}} \approx e_{\text{intrinsic}} + \gamma$$

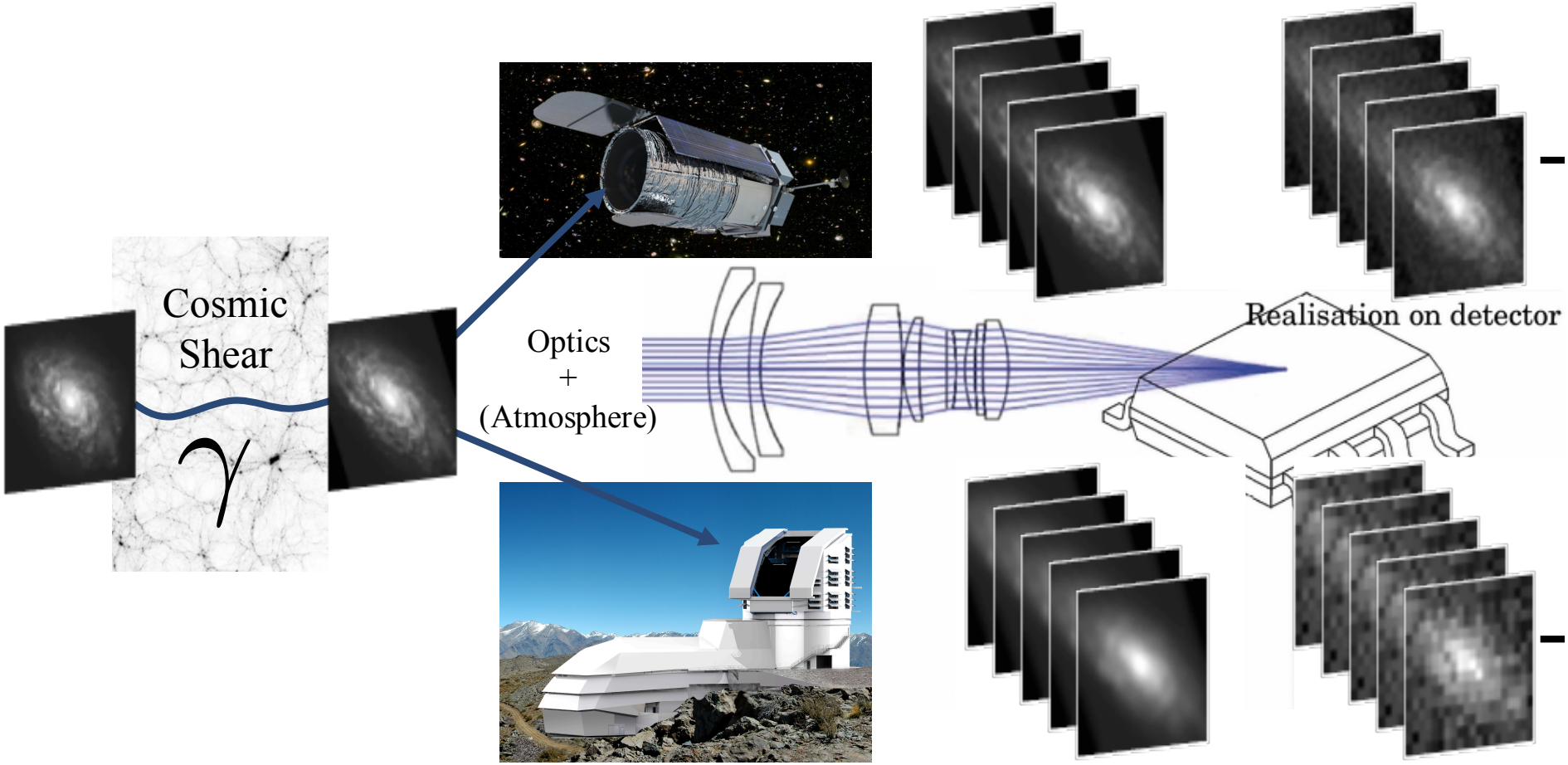
$$\sigma_{e_{\text{instinsic}}} \sim 0.3$$

$$\sim 0.03$$

Cosmic shear signal of order the ellipticity of Earth



Shape noise (intrinsic galaxy properties) is only part of the

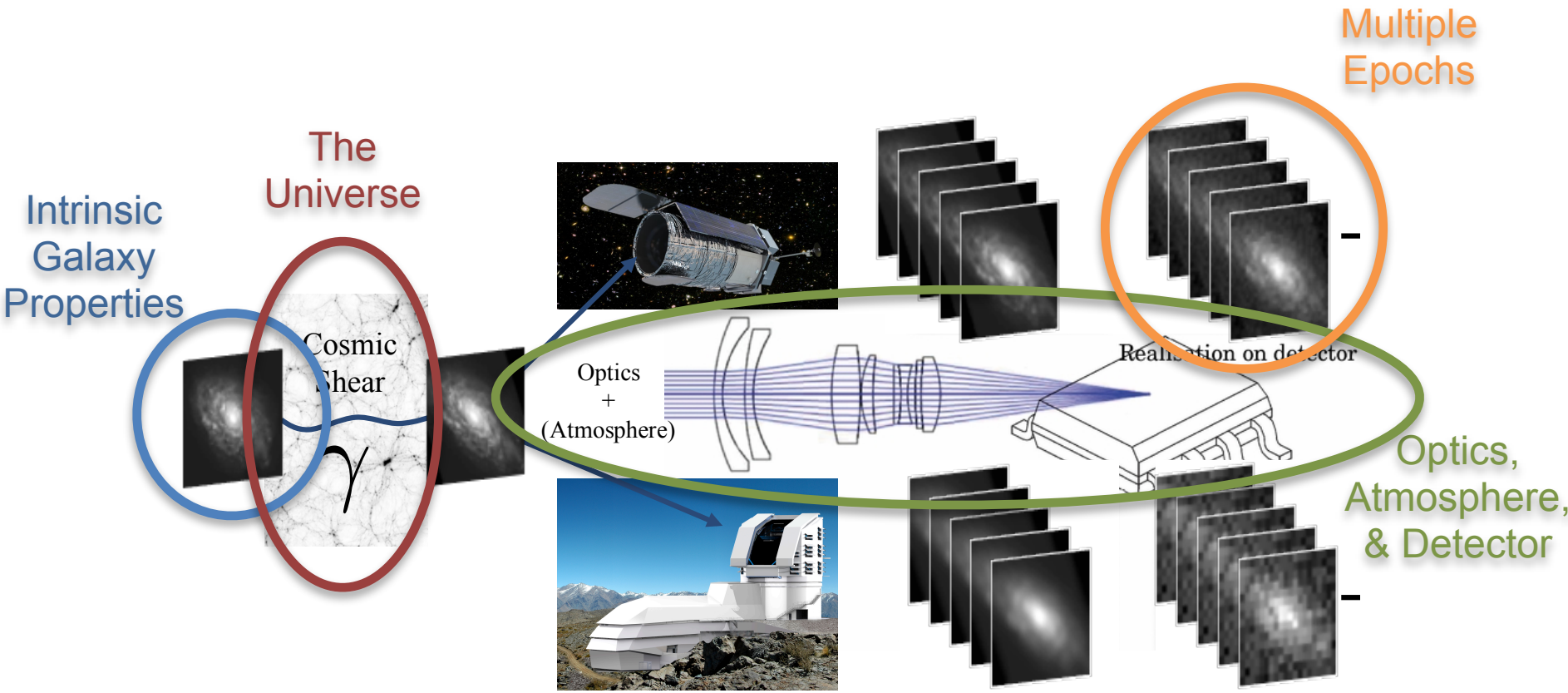


Technical challenges

- Statistical model
- Understand intrinsic galaxy properties to avoid bias
- How to infer billions of correlated parameters

Technical challenges

- Statistical model



Statistical framework summary

$$\Pr(\boldsymbol{\theta}, \{\psi_s\}, \{\Pi_i, \Omega_i\}, \{\omega_n, \alpha_n, \xi_n\}, \mathcal{A}, \tau, m, a_\eta | \{\mathbf{d}_n\}, \mathcal{X}) \propto$$

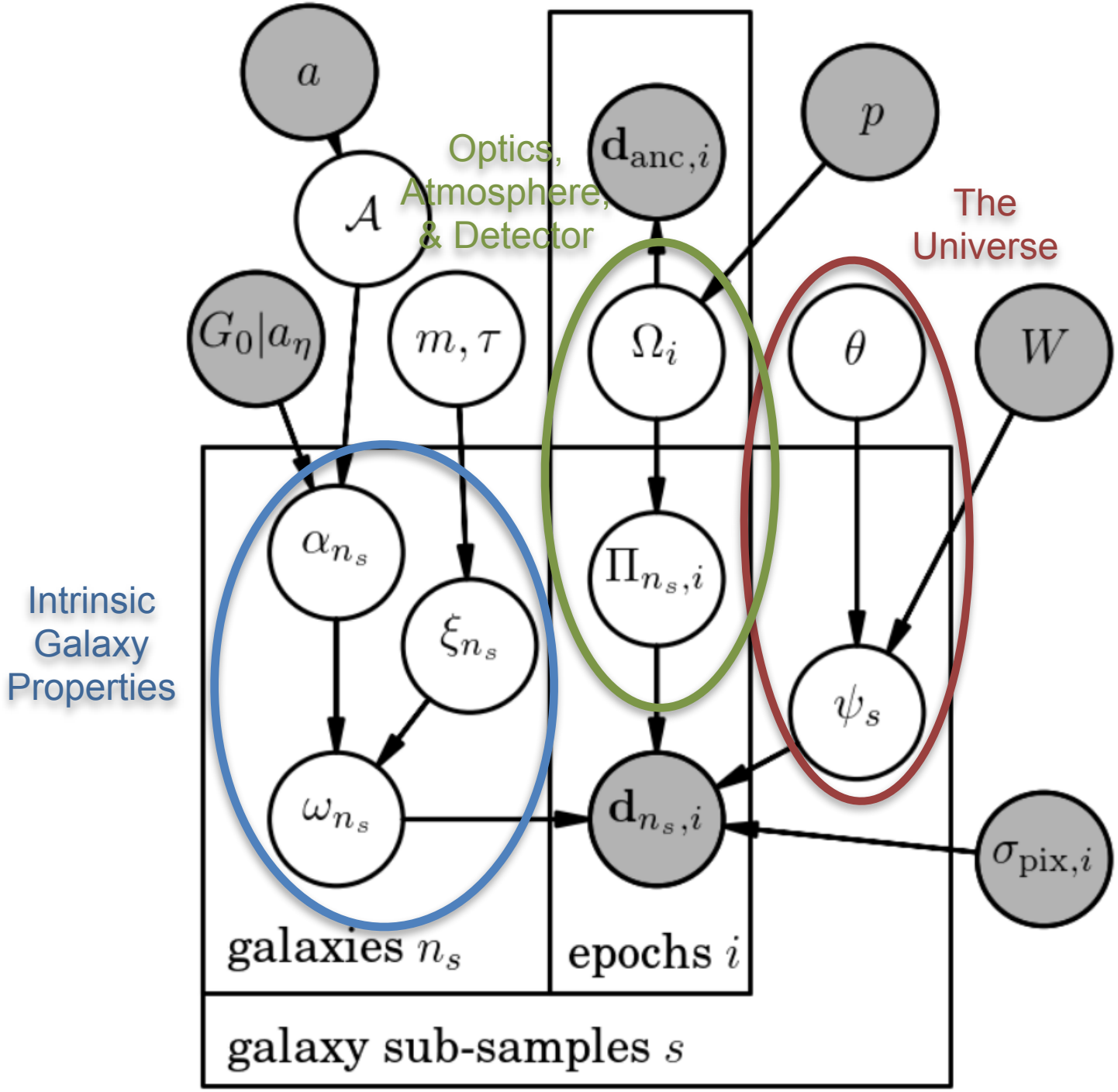
$$\Pr(\boldsymbol{\theta}) \cdot \Pr(\{\psi_s\} | W, \{s\}, \boldsymbol{\theta}) \cdot \Pr(\mathcal{A} | a)$$

$$\times \prod_{n=1}^{n_{\rm gal}} \Pr(\omega_n | \alpha_n, I) \cdot \Pr(\alpha_n | \mathcal{A}, a_\eta) \cdot \Pr(\xi_n | m, \tau)$$

$$\times \prod_{i=1}^{n_{\rm epoch}} \Pr(\Pi_{n,i} | \Omega_i, I) \cdot \Pr(\Omega_i | p, \mathbf{d}_{{\rm anc},i})$$

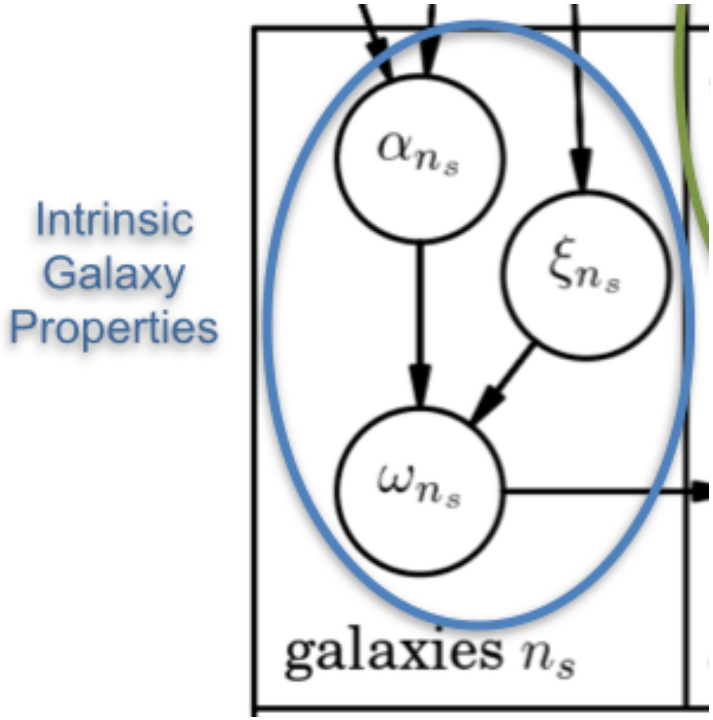
$$\times \Pr(\mathbf{d}_{n,i} | \omega_n, \xi_n, \psi_s, \Pi_{n,i}) \quad (18)$$

Schneider+ 2014

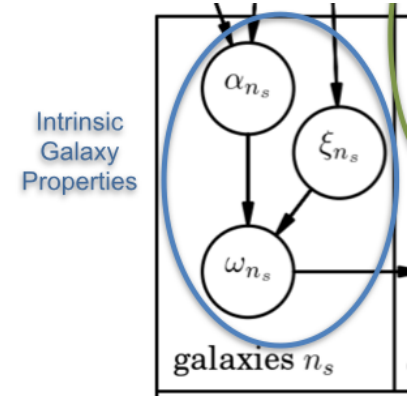

a
$\mathcal{A}$
$G_0|a_\eta$
m, τ
α_{n_s}
ξ_{n_s}
ω_{n_s}
$\mathbf{d}_{\mathrm{anc},i}$
Ω_i
$\Pi_{n_s,i}$
$\mathbf{d}_{n_s,i}$
p
θ
ψ_s
W
$\sigma_{\mathrm{pix},i}$
Optics, Atmosphere, & Detector
The Universe
Intrinsic Galaxy Properties
galaxies n_s
epochs i
galaxy sub-samples s

Technical challenges

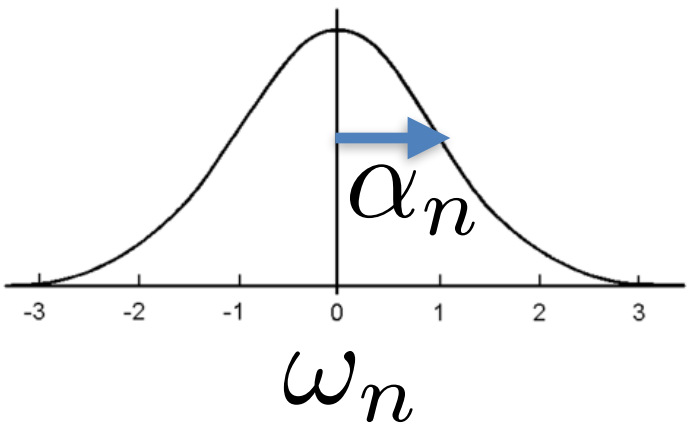
- Statistical model
- Understand intrinsic galaxy properties to avoid bias
- How to int d parameters

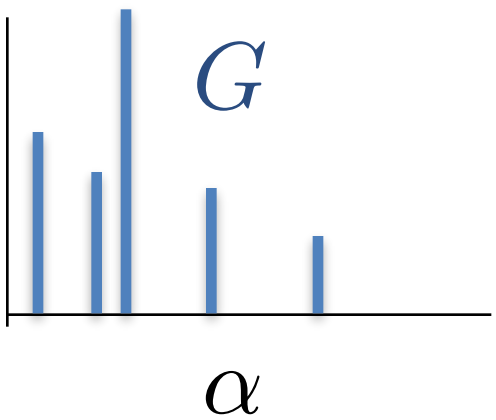


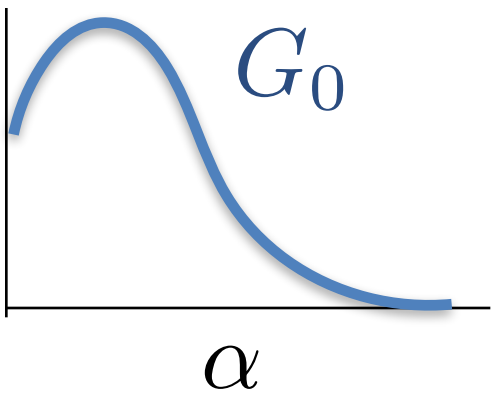
Dirichlet Process



$$\omega_n \sim \mathcal{N}(0, \alpha_n), \quad \alpha_n \sim G(\alpha_n|\mathcal{A}), \quad G \sim \mathrm{DP}(\mathcal{A}, G_0)$$

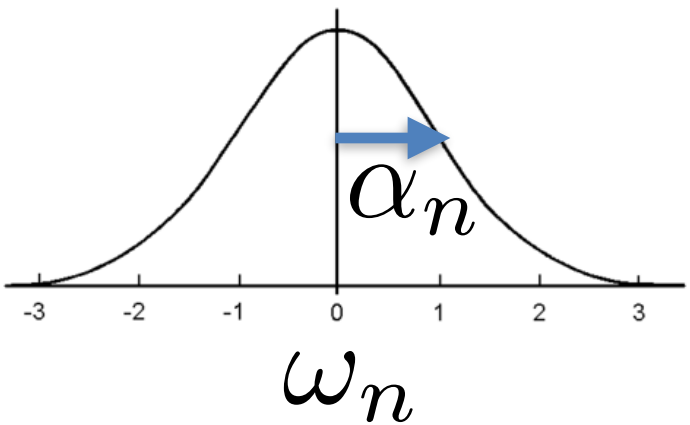


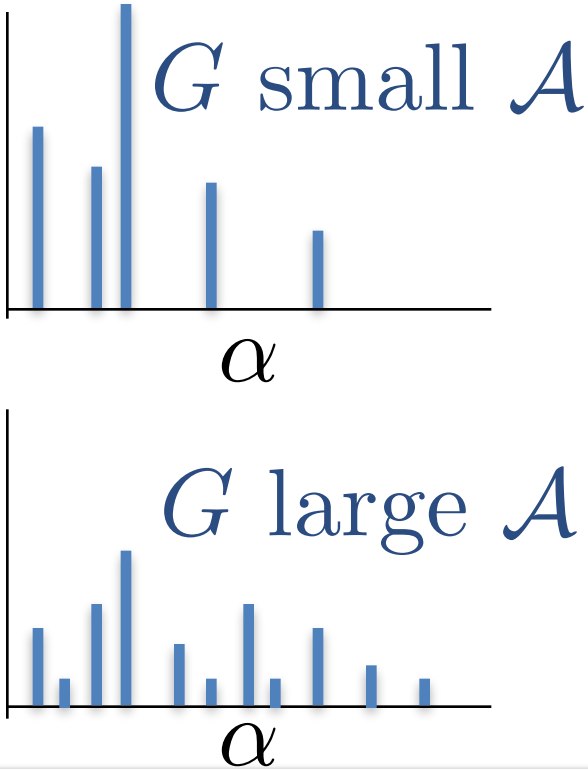


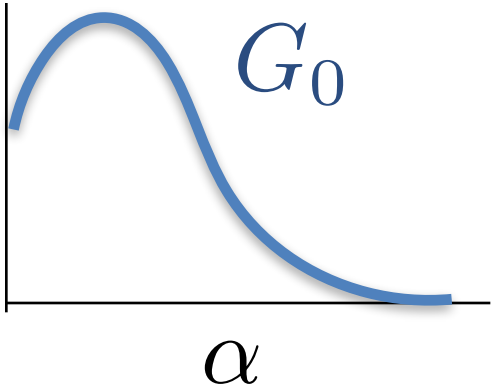


Dirichlet Process

$$\omega_n \sim \mathcal{N}(0, \alpha_n), \quad \alpha_n \sim G(\alpha_n | \mathcal{A}), \quad G \sim \mathrm{DP}(\mathcal{A}, G_0)$$





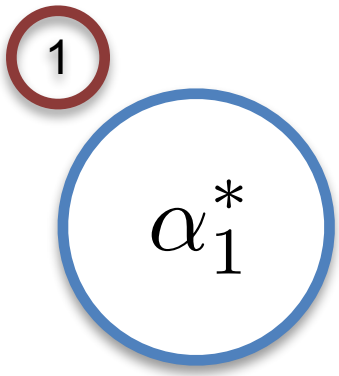


Dirichlet Process: Learning (Clustering) as you go

$$\alpha_n | \alpha_1, \ldots, \alpha_{n-1} \sim \frac{1}{n-1+\mathcal{A}} \left(\sum_{c=1}^{K} N_c \delta_{\mathrm{D}} \left(\alpha_c^* \right) + \mathcal{A} G_0 \left(\cdot \right) \right)$$

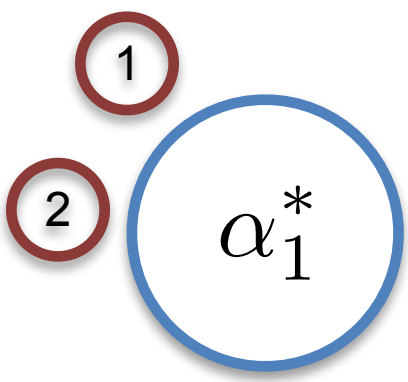
Dirichlet Process: Learning (Clustering) as you go

$$\alpha_n | \alpha_1, \ldots, \alpha_{n-1} \sim \frac{1}{n-1+\mathcal{A}} \left(\sum_{c=1}^{K} N_c \delta_{\mathrm{D}} \left(\alpha_c^* \right) + \mathcal{A} G_0 \left(\cdot \right) \right)$$



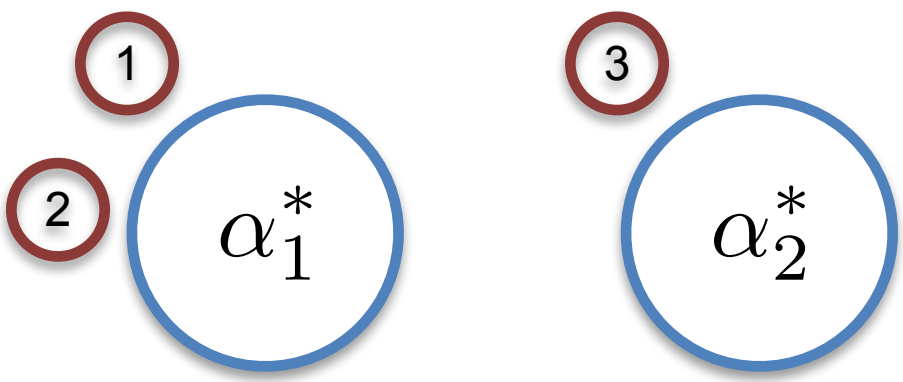
Dirichlet Process: Learning (Clustering) as you go

$$\alpha_n | \alpha_1, \ldots, \alpha_{n-1} \sim \frac{1}{n-1+\mathcal{A}} \left(\sum_{c=1}^{K} N_c \delta_{\mathrm{D}} \left(\alpha_c^* \right) + \mathcal{A} G_0 \left(\cdot \right) \right)$$



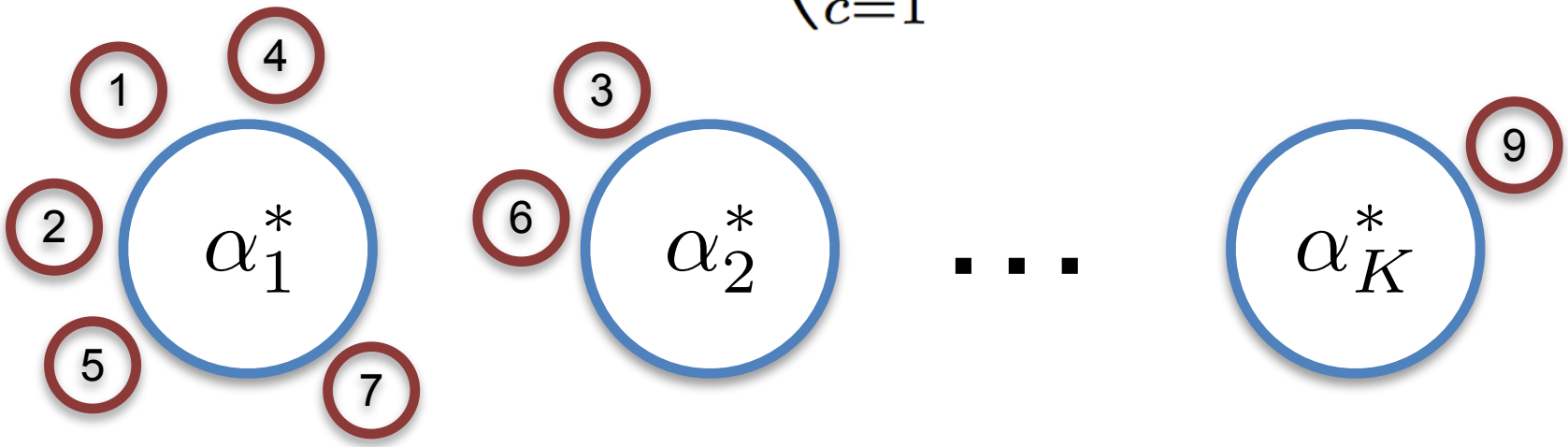
Dirichlet Process: Learning (Clustering) as you go

$$\alpha_n | \alpha_1, \ldots, \alpha_{n-1} \sim \frac{1}{n-1+\mathcal{A}} \left(\sum_{c=1}^{K} N_c \delta_{\mathrm{D}} \left(\alpha_c^* \right) + \mathcal{A} G_0 \left(\cdot \right) \right)$$



Dirichlet Process: Learning (Clustering) as you go

$$\alpha_n | \alpha_1, \ldots, \alpha_{n-1} \sim \frac{1}{n-1+\mathcal{A}} \left(\sum_{c=1}^{K} N_c \delta_{\mathrm{D}} \left(\alpha_c^* \right) + \mathcal{A} G_0 \left(\cdot \right) \right)$$



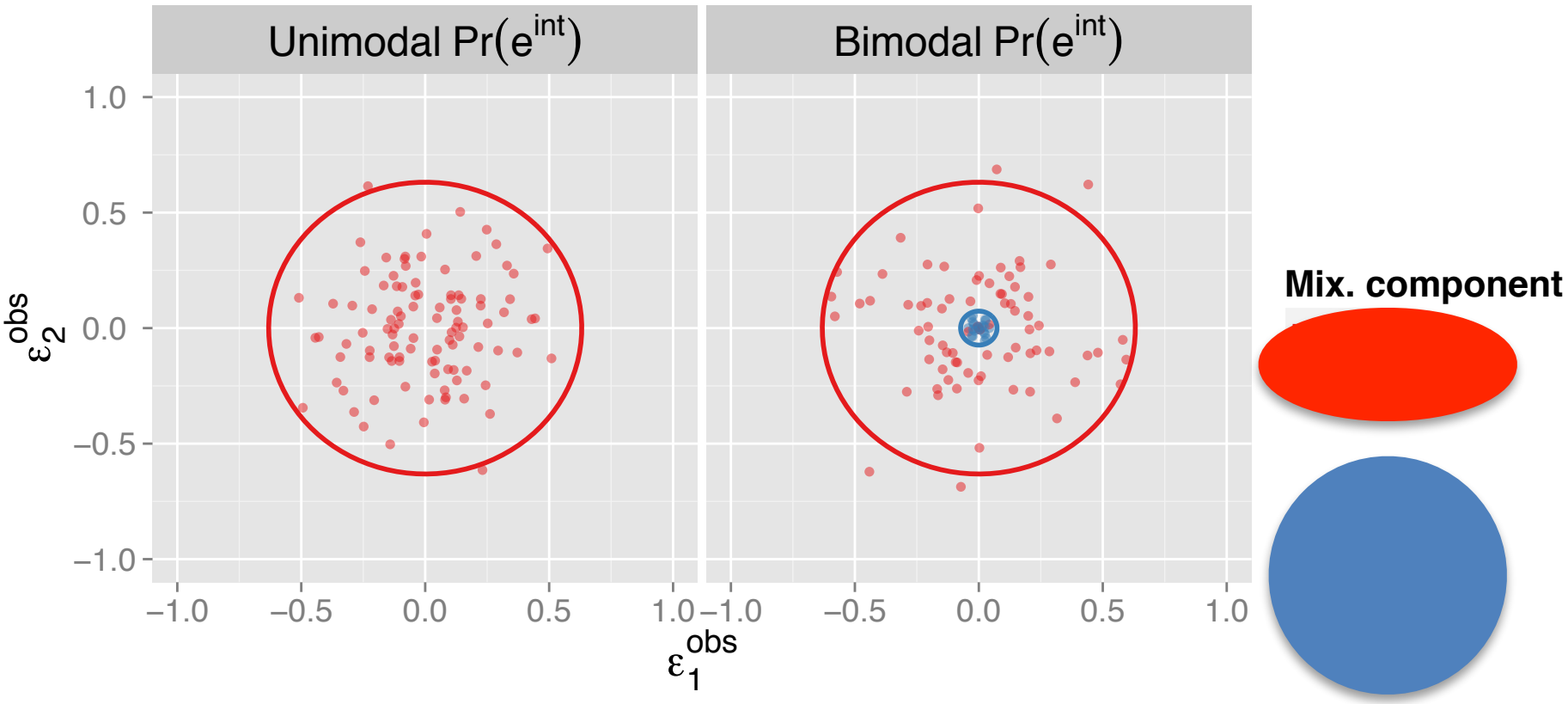
Listening to the data

$$\alpha_n|\alpha_1,\ldots,\alpha_{n-1} \sim \frac{1}{n-1+\mathcal{A}} \left(\sum_{c=1}^{K} N_c \delta_{\mathrm{D}}\left(\alpha_c^*\right) + \mathcal{A} G_0\left(\cdot\right)\right)$$

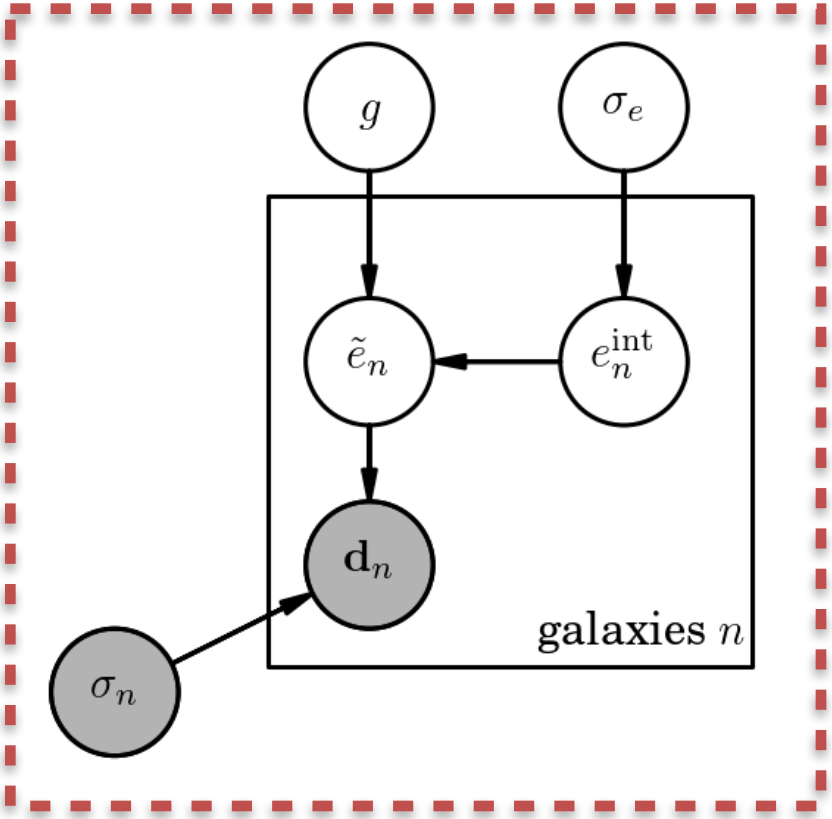
$$\Pr(c_n = c_\ell | c_{-n}, \omega_n, \alpha, \mathcal{X}) = b\, N_{-n,c} \Pr(\mathbf{d}_n | \alpha_{c_\ell}, \mathcal{X}), \quad \forall \ell \neq n \tag{26}$$

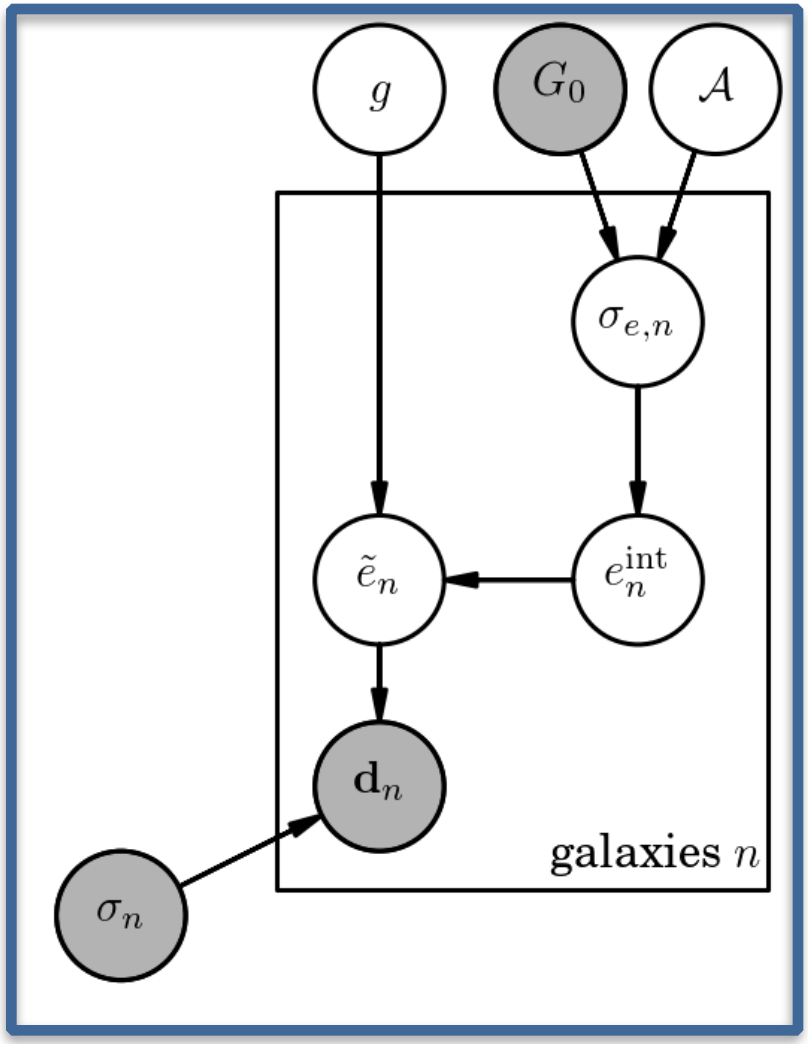
$$\Pr(c_n \neq c_\ell \forall \ell \neq n | c_{-n}, \omega_n, \mathcal{X}) = b\, \mathcal{A} \int \Pr(\mathbf{d}_n | \alpha, \mathcal{X})\, G_0(\alpha)\, d\alpha, \tag{27}$$

Toy Model

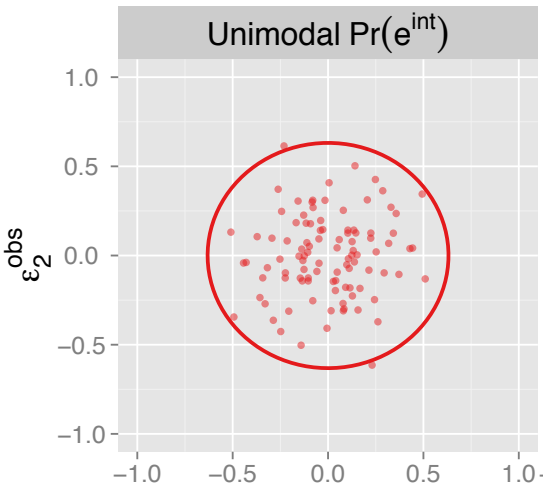

Unimodal Pr(e^int)
Bimodal Pr(e^int)
1.0
0.5
0.0
−0.5
−1.0
ε_2^obs
−1.0
−0.5
0.0
0.5
1.0
ε_1^obs
Mix. component

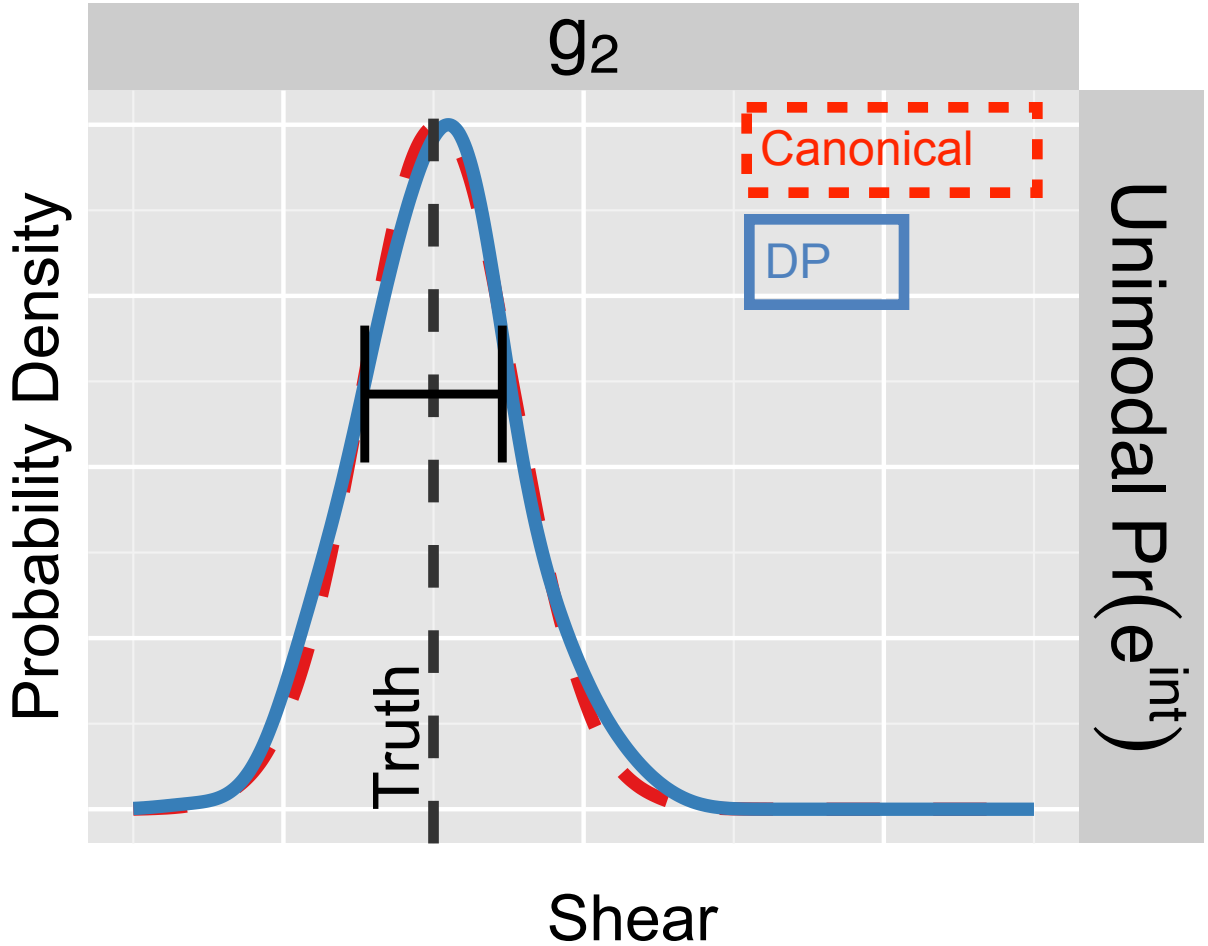
Two models: Canonical and DP



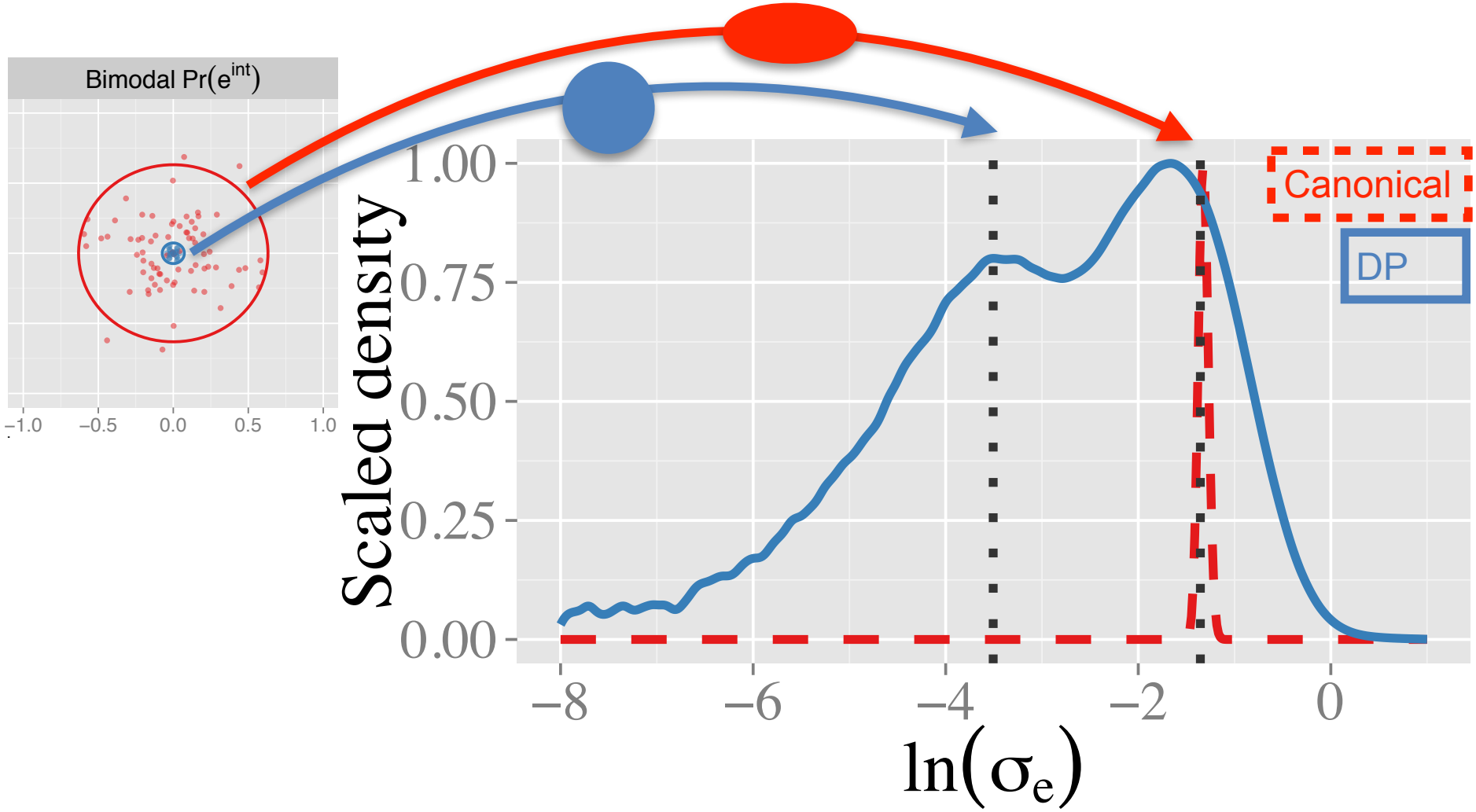


Both methods comparable

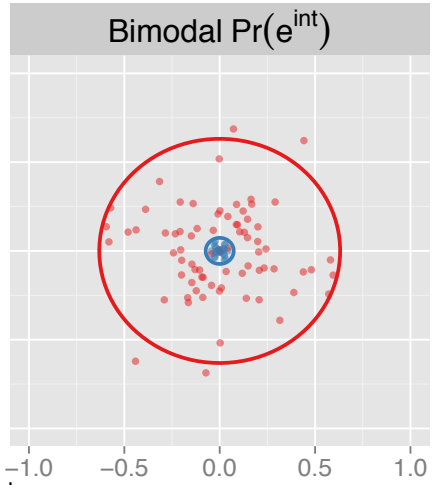


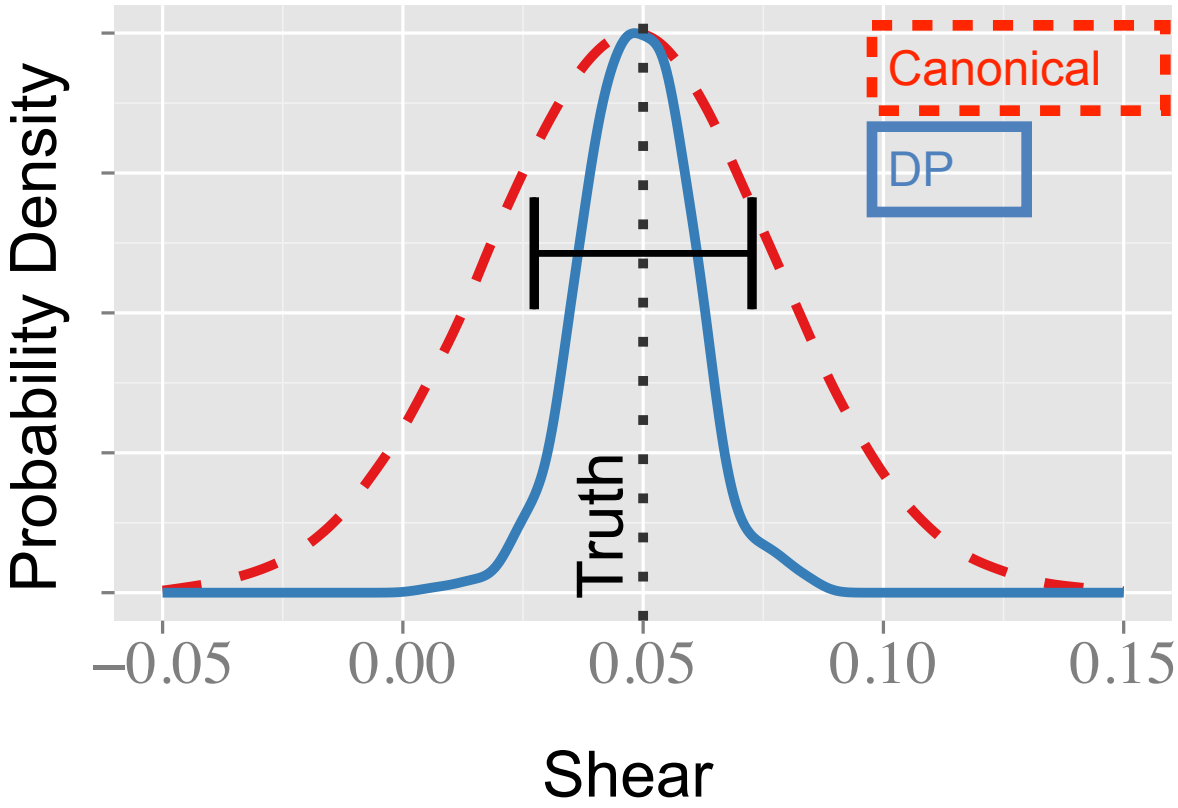


Dirichlet Process recognizes more than one type of galaxy



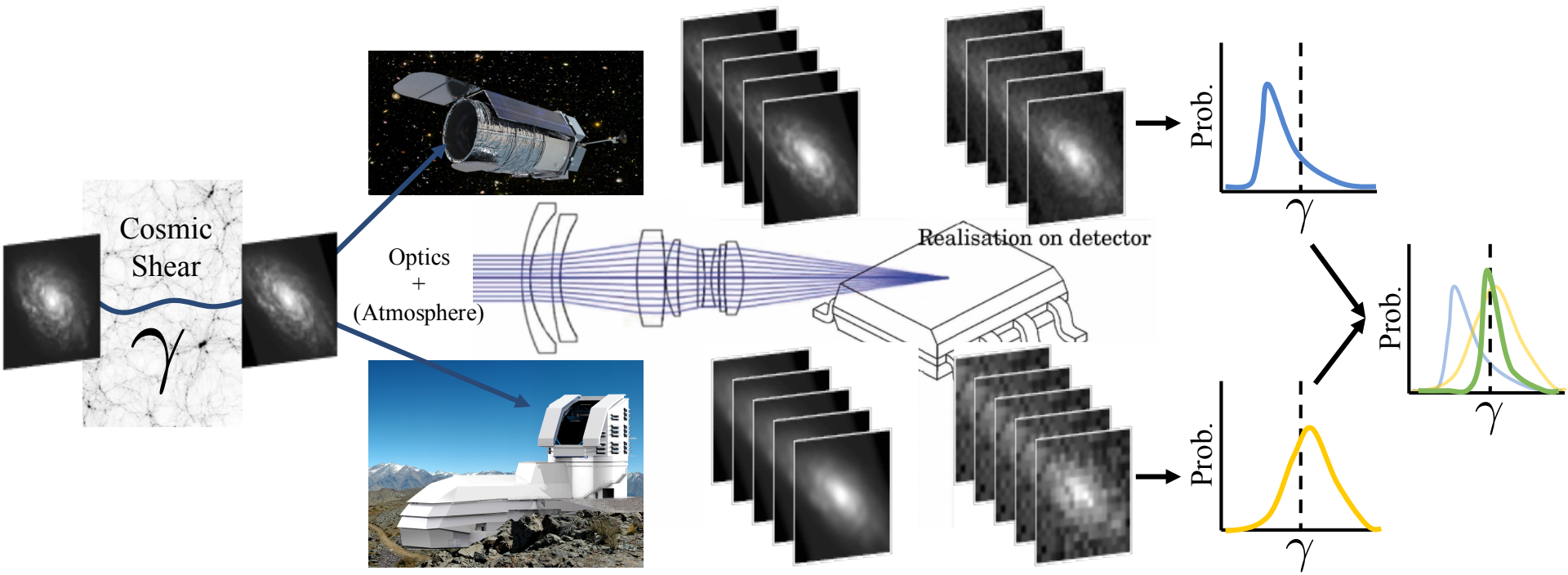
Because DP learned about galaxies it made more precise measurement





Looking to the future.

Combining observations from multiple probes



Lawrence Livermore National Laboratory