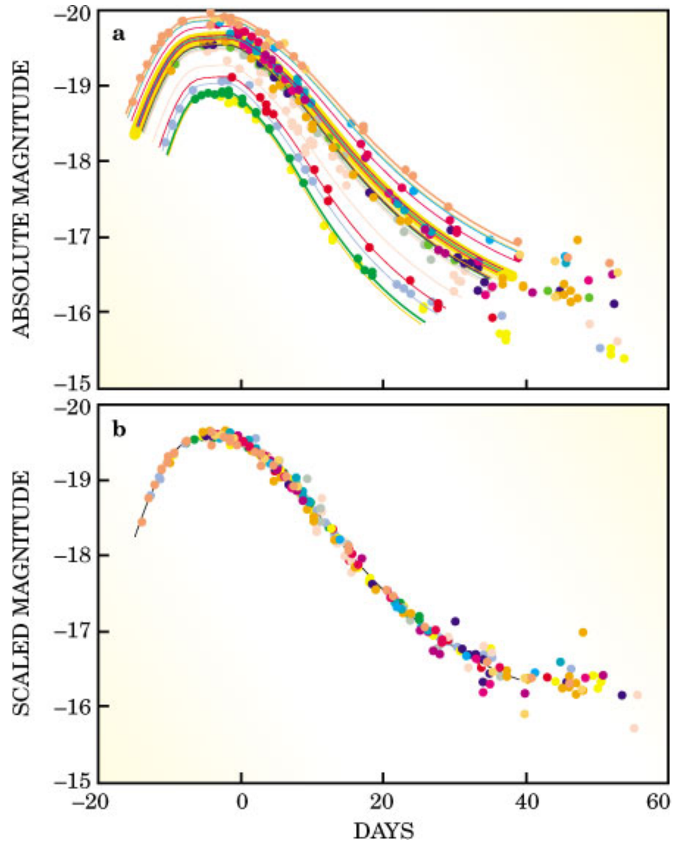
Bayesian Hierarchical Models

Alan Heavens

ICIC
Imperial Centre
for Inference & Cosmology

Bayesian Hierarchical Models

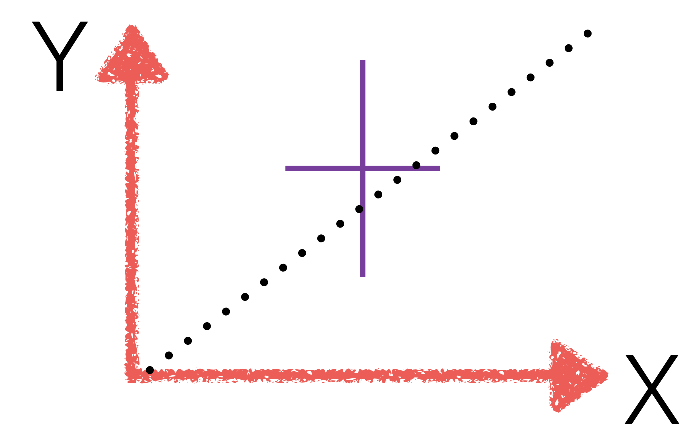


- Many inference problems are complex, with many layers
- How can we sample from the posterior?
- e.g. Supernovae. Not standard candles, but there are corrections due to colour, and 'stretch'
- Colours have errors; stretch has errors
- Redshifts have errors (small, if spectroscopic)
- Magnitudes have errors
- How can we make sure all the errors are propagated correctly to the posterior?
- Bayesian Hierarchical Models split the problem into stages, and can do this in many cases. They also expose what you need to know or assume.

BHM example

Model: y = mx. We measure X,Y, but they *both* have gaussian errors. What is the posterior for m?

Rule 1: we want ***p(m | X,Y)***

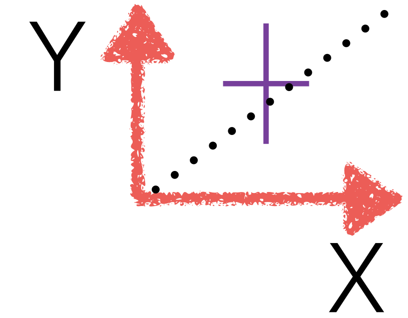


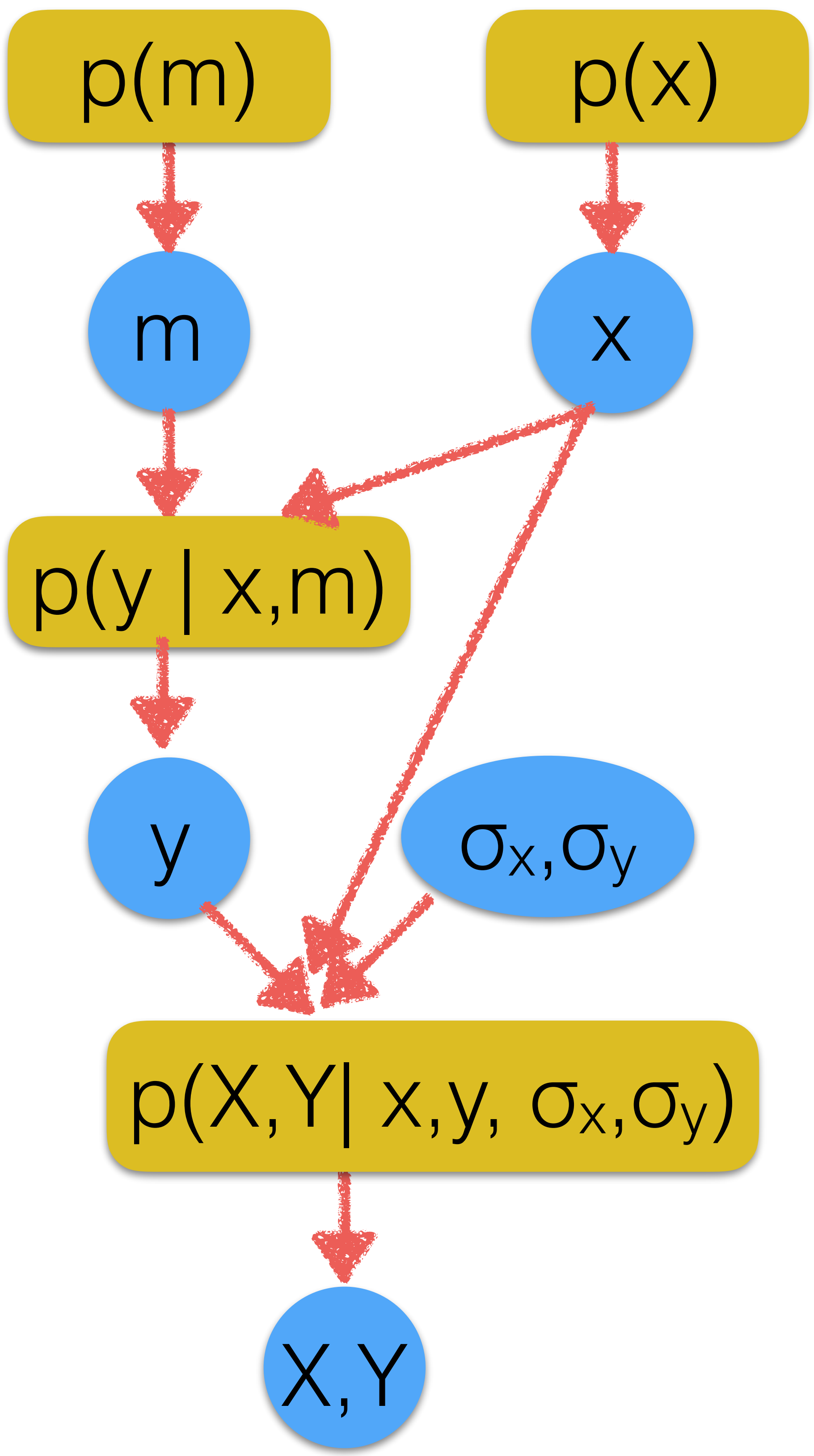
Bayes: ***p(m | X,Y)*** α ***p(X, Y | m) p(m)***

There are two unknown ('latent') variables in the problem:

the true values x,y

Forward (generative) model





Priors

δ(y-mx)

Does not depend on σ_x,σ_y

Error

$$X \sim \mathcal{N}(x, \sigma_x^2)$$
$$Y \sim \mathcal{N}(y, \sigma_y^2);$$

Data

Does not depend on m

Build the statistical model

Bayes: ***p(m | X,Y)*** α ***p(X, Y | m) p(m)***

- Marginalise over latent x and y:

$$p(m|X,Y) \propto \int p(X,Y,x,y|m)\, p(m)\, dx\, dy$$

- Use product rule to expand the first probability:

$$\propto \int p(X,Y|x,y,m)\, p(x,y|m)\, p(m)\, dx\, dy$$

$$p(M|X,Y) \propto \int p(X,Y|x,y,m)\, p(x,y|m)\, p(m)\, dx\, dy$$

X and Y depend only on x,y (not on m):

$$p(X,Y|x,y,m) = p(X,Y|x,y)$$

x and y distribution is:

$$p(x,y|m) = p(y|x,m)p(x|m)$$

The physical relation applies to x,y, *not* X,Y:

$$p(y|x,m) = \delta(y - mx)$$

Prior on x does not depend on m:

$$p(x|m) = p(x)$$

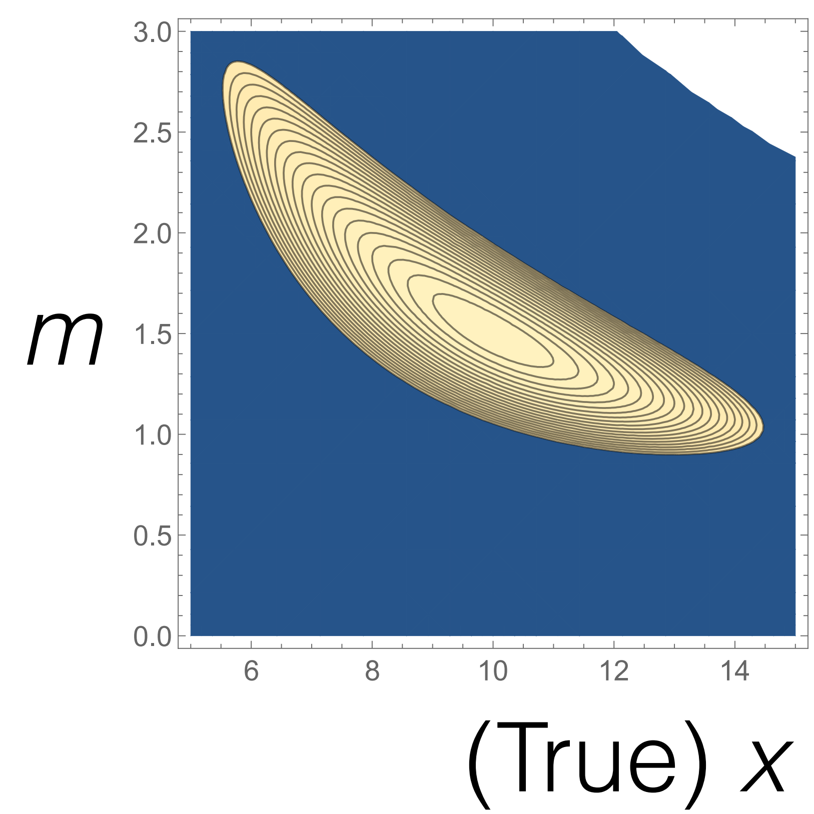
Integrate over y:

$$p(m|X,Y) \propto \int p(X,Y|x,mx)\, p(x)\, p(m)\, dx.$$

$$p(m|X,Y) \propto \int p(X,Y|x,mx)\,p(x)\,p(m)\,dx.$$

If the error distribution is gaussian, with zero mean and unit variance, and if we take a uniform prior in x and m:

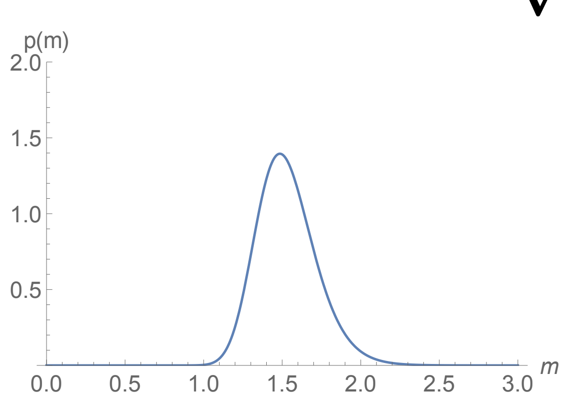
$$p(m|X,Y) \propto \int e^{-\frac{1}{2}(X-x)^2} e^{-\frac{1}{2}(Y-mx)^2}\,dx$$



X=10, Y=15

Integrating gives the posterior for m:

$$p(m|X,Y) \propto \frac{1}{\sqrt{1+m^2}}\, e^{-\frac{(-mX+Y)^2}{2(1+m^2)}}.$$



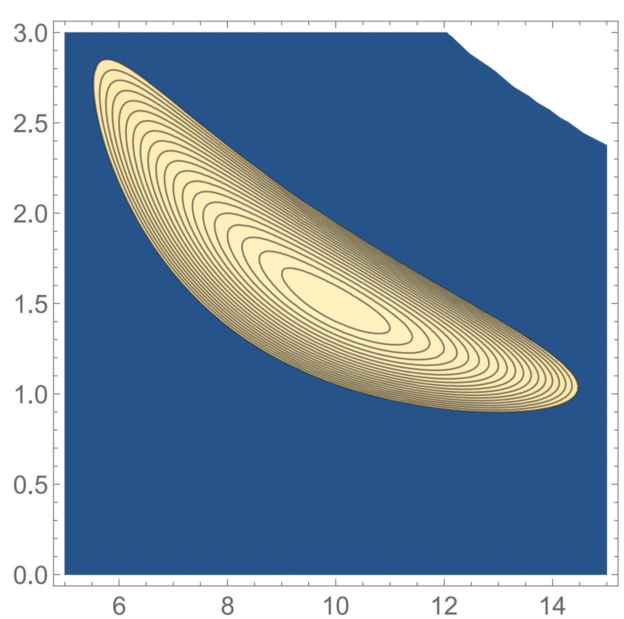
Gibbs sampling

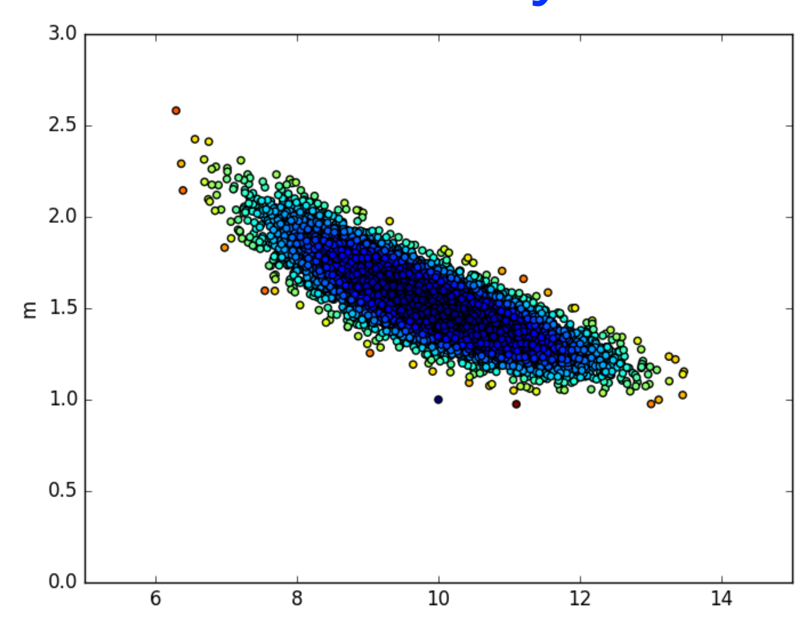
At fixed x: $$p(m|X,Y,x) \propto \exp\left[-\frac{(X-x)^2}{2}\right]\exp\left[-\frac{(Y-mx)^2}{2}\right]$$
$$\propto \exp\left[-\frac{x^2\left(m-\frac{Y}{x}\right)^2}{2}\right]$$

i.e. $$p(m|X,Y,x) \sim \mathcal{N}\left(\frac{Y}{x},\frac{1}{x^2}\right)$$

At fixed m: $$p(x|X,Y,m) \sim \mathcal{N}\left(\frac{X+Ym}{1+m^2},\frac{1}{1+m^2}\right)$$

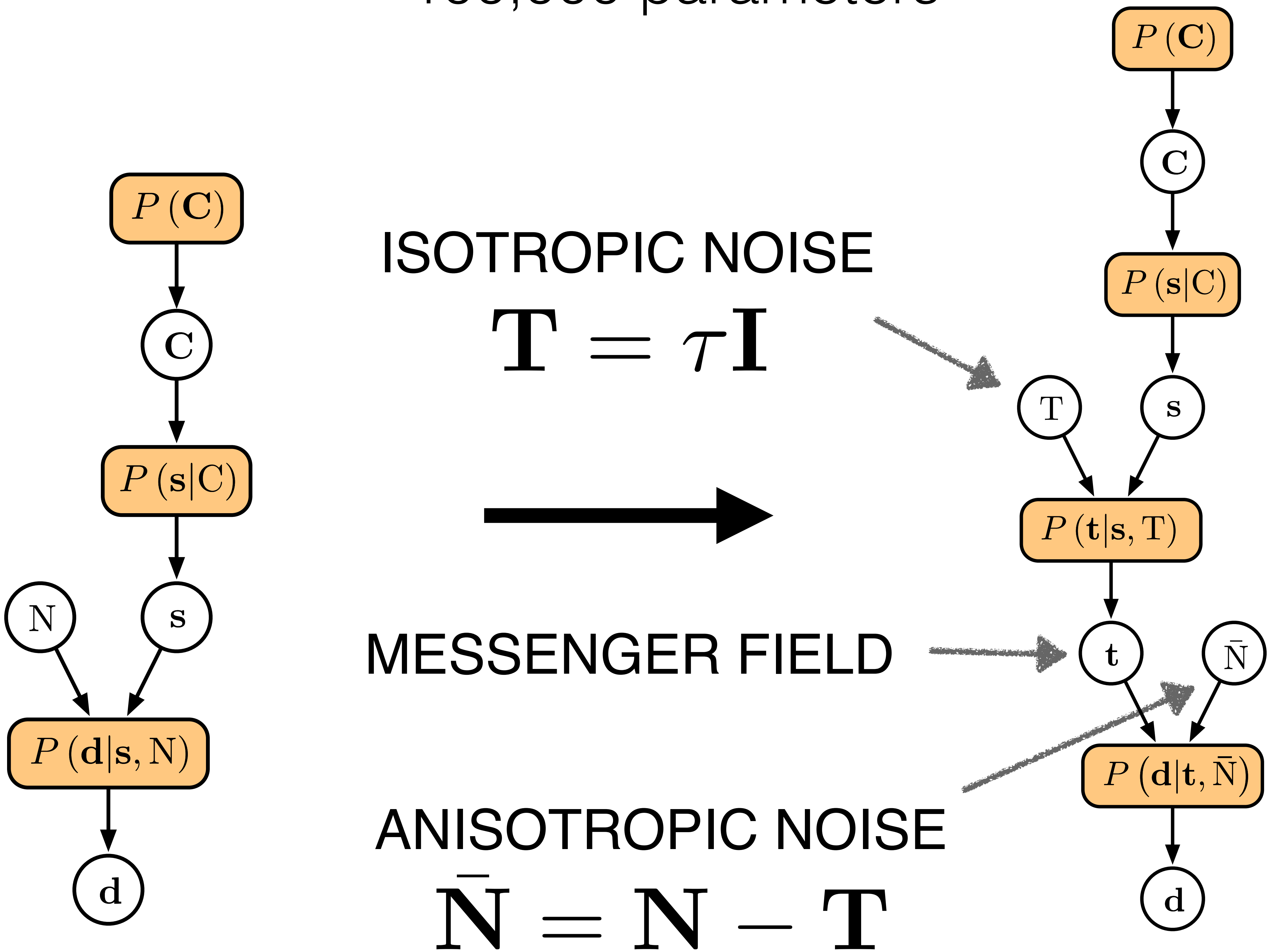
Draw alternately from *m* and *x*

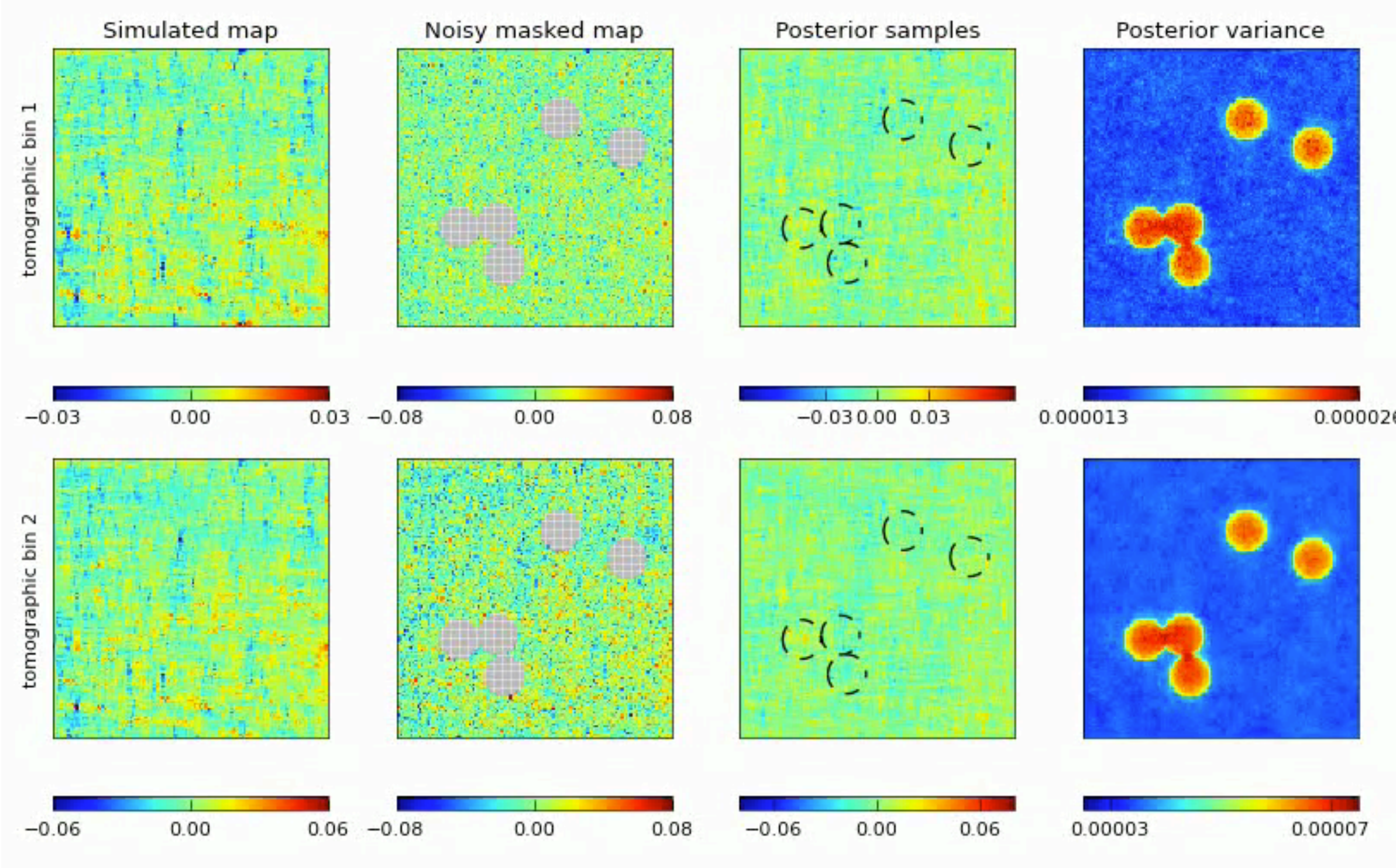




BHM for weak lensing (Alsing et al 2015).

100,000 parameters

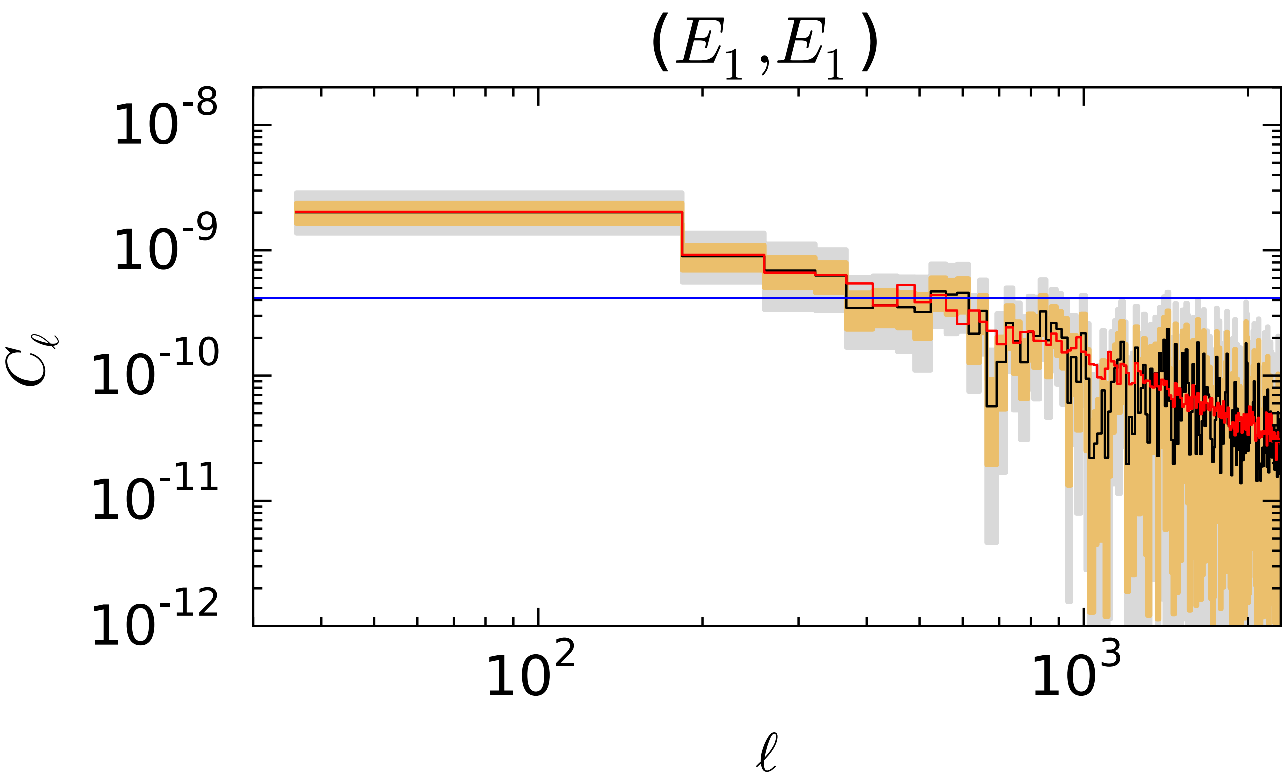


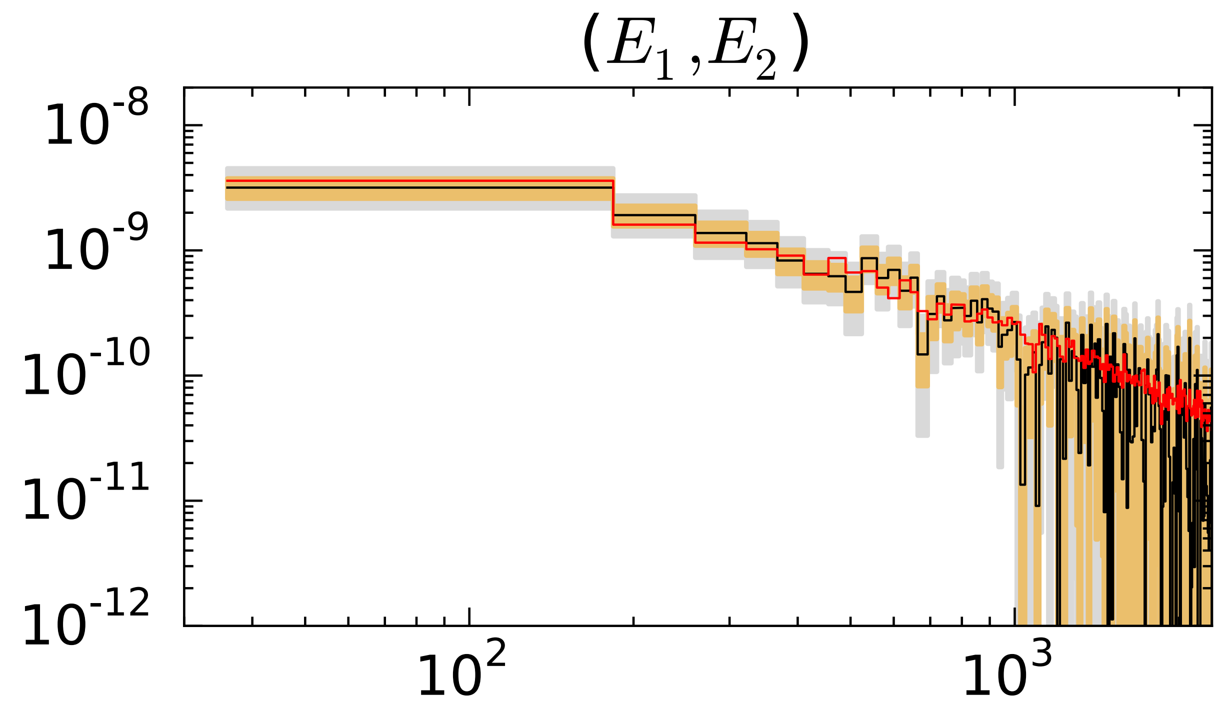


SUNGLASS simulations (Kiessling et al 2011)

ICIC Imperial Centre for Inference & Cosmology

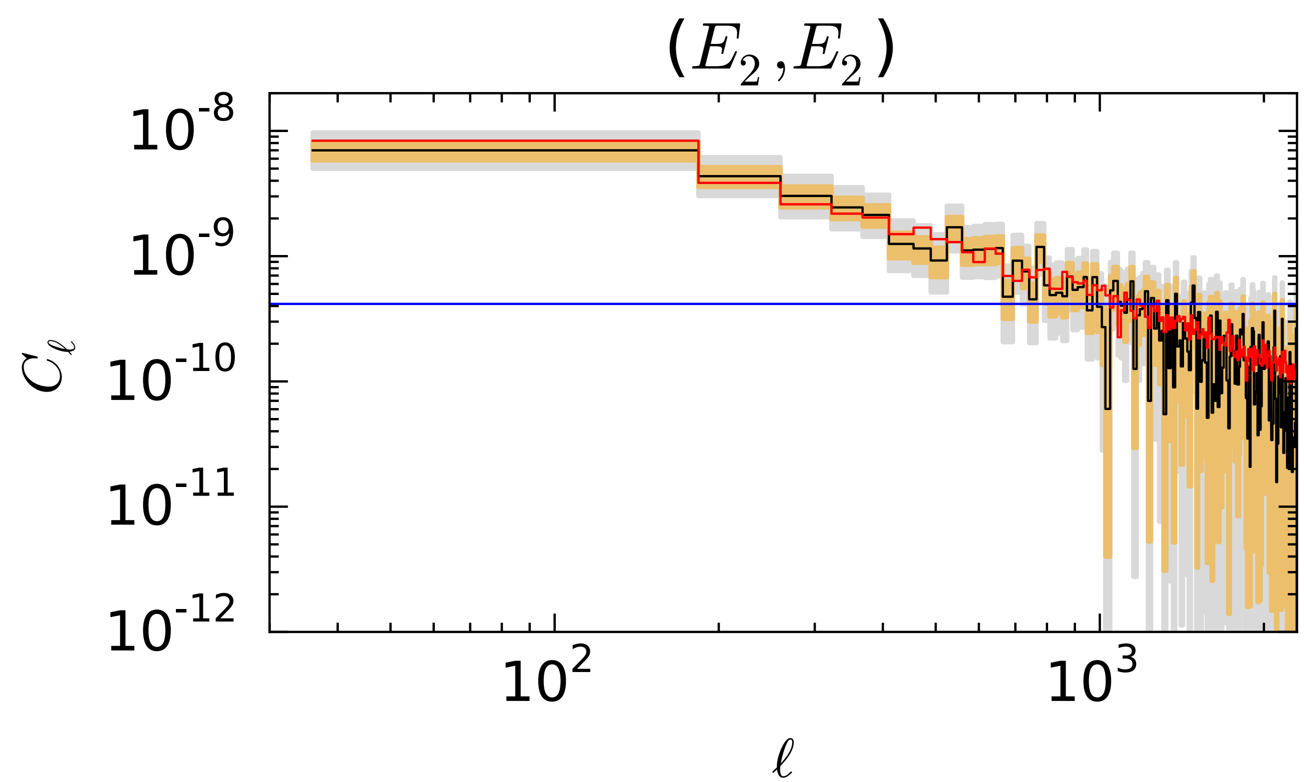
Imperial College London





Simulations (Kiessling et al 2011)

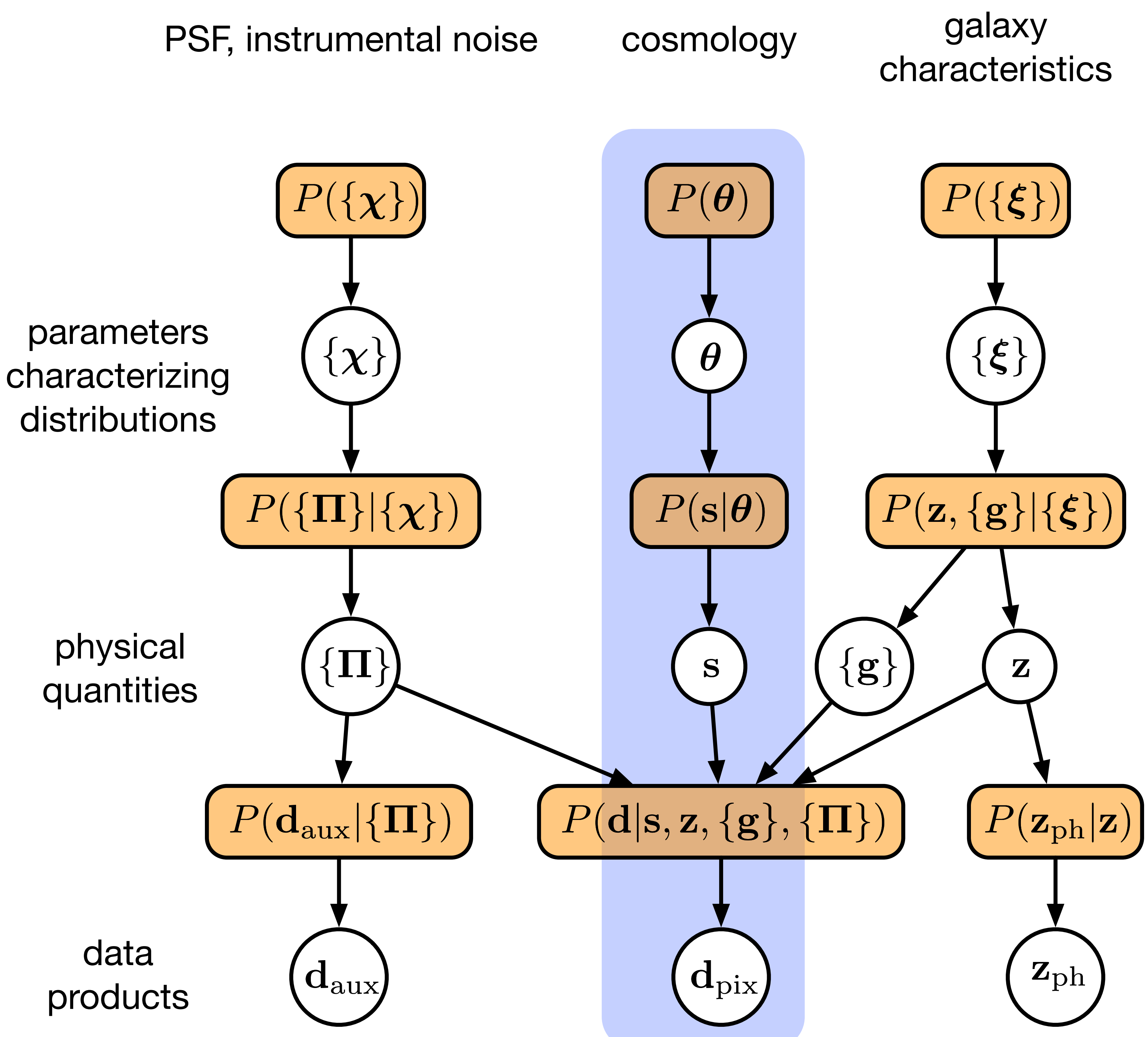
E-modes are recovered, well below the shot noise at high-*l*



ICIC Imperial Centre for Inference & Cosmology

Imperial College London

Global BHM:



Can include:

Mask
Intrinsic alignments
Baryon feedback
Shape measurement
Photometric redshifts

Alsing, J., AFH, et al., 2015
Schneider M., et al., 2014

ICIC Imperial Centre for Inference & Cosmology
Imperial College London

Conclusions

- With Bayesian hierarchical models: *we can sample the posterior probability density* - the object you really want for scientific inference
- Break problem into steps, with (known or unknown) conditional probabilities
- Sample with Gibbs, HMC, …
- Very large parameter spaces may be feasible
- BHM for SNe: Has been done (Mandel et al 2009; Shariff et al 2015)
- BHM for Large Scale Structure: (Wandelt, Leclercq, Elsner, Jasche, Lavaux, Ata, Kitaura…).
- By subdividing the problem, BH Modelling may be able to sample efficiently from the posterior - just what you want.