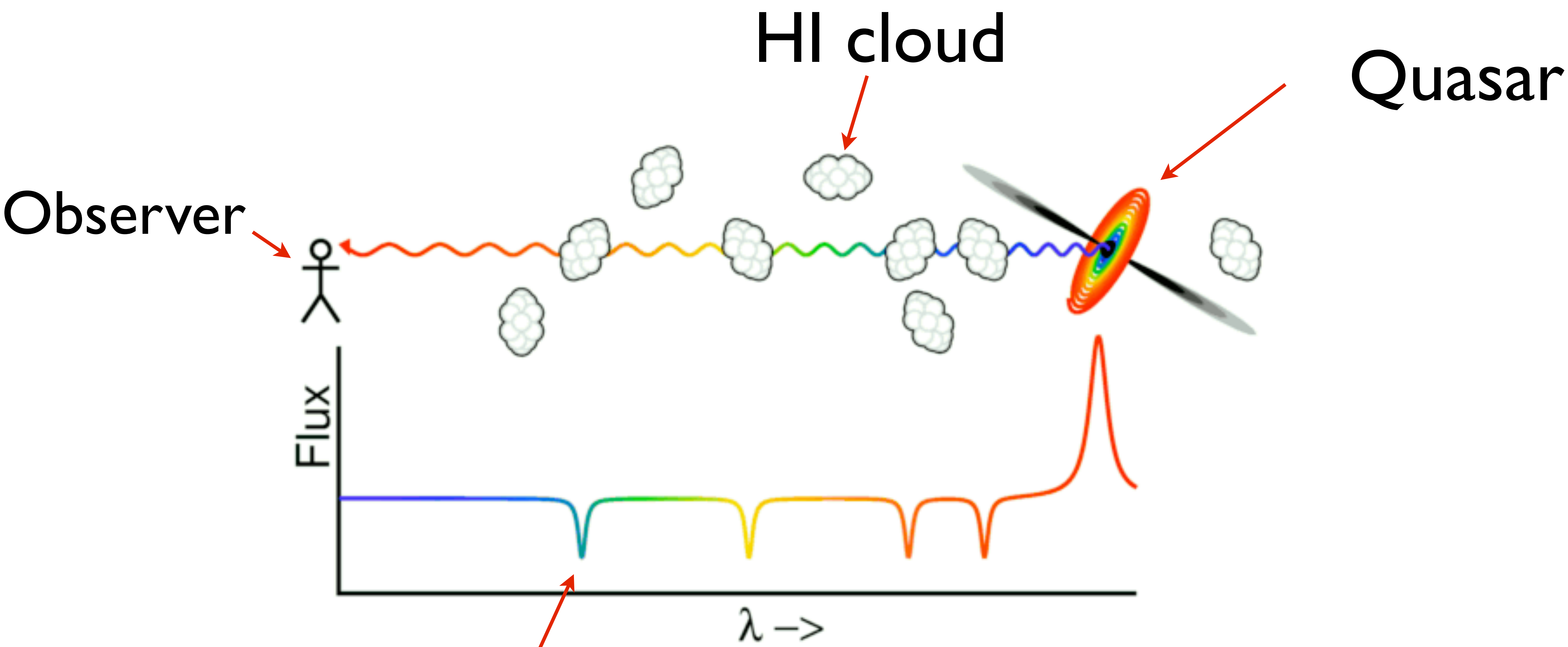
Minimally Parametric Power Spectrum Constraints from Lyman-α

Simeon Bird
Institute of Astronomy, Cambridge
Collaborators: Hiranya Peiris, Matteo Viel and Licia Verde

Bird et al (2010), arxiv:1010.1519

Lyman-α forest

Neutral hydrogen clouds scatter quasar light

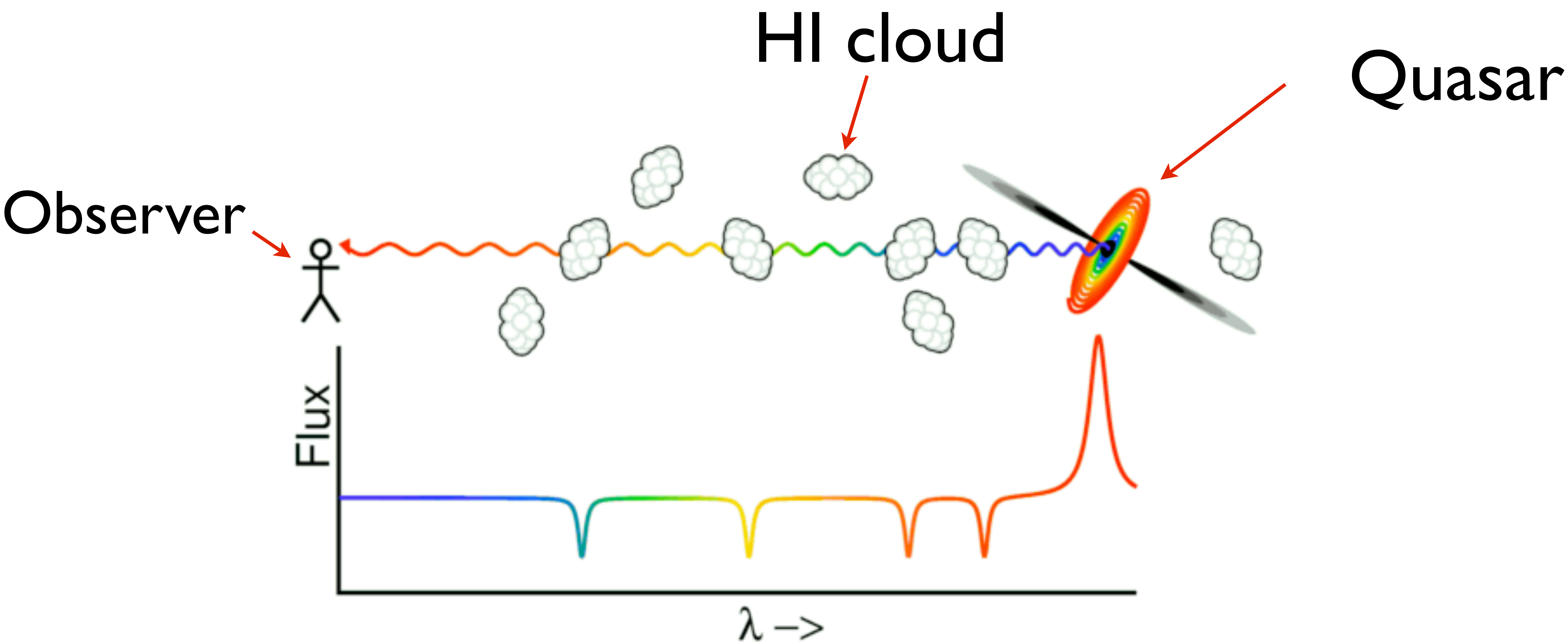


Trough where light is scattered
out of line of sight

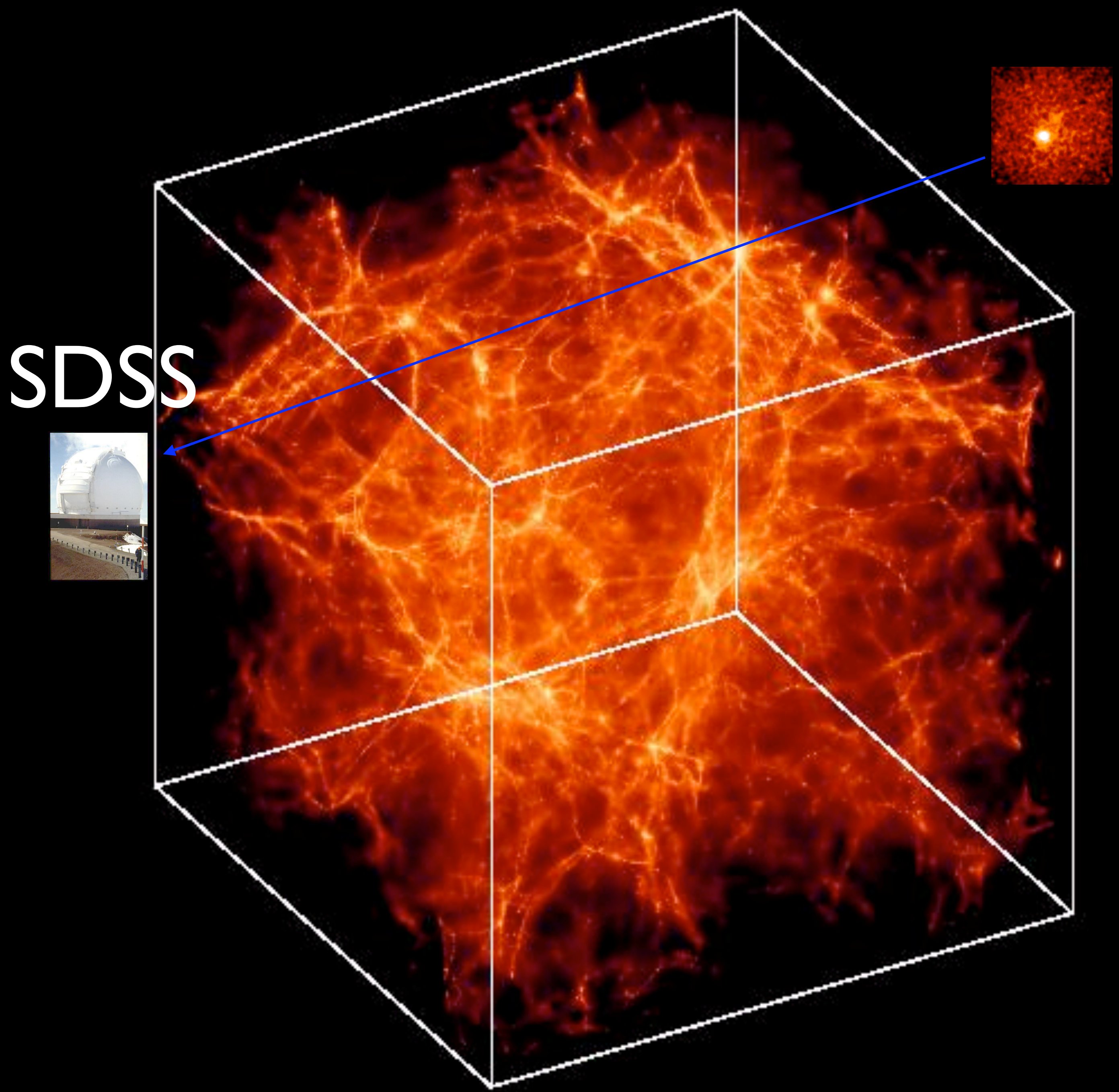
Image: E. Wright

Lyman-α forest

3D map of neutral hydrogen: traces baryons



Lyman-α forest



Baryons trace dark matter

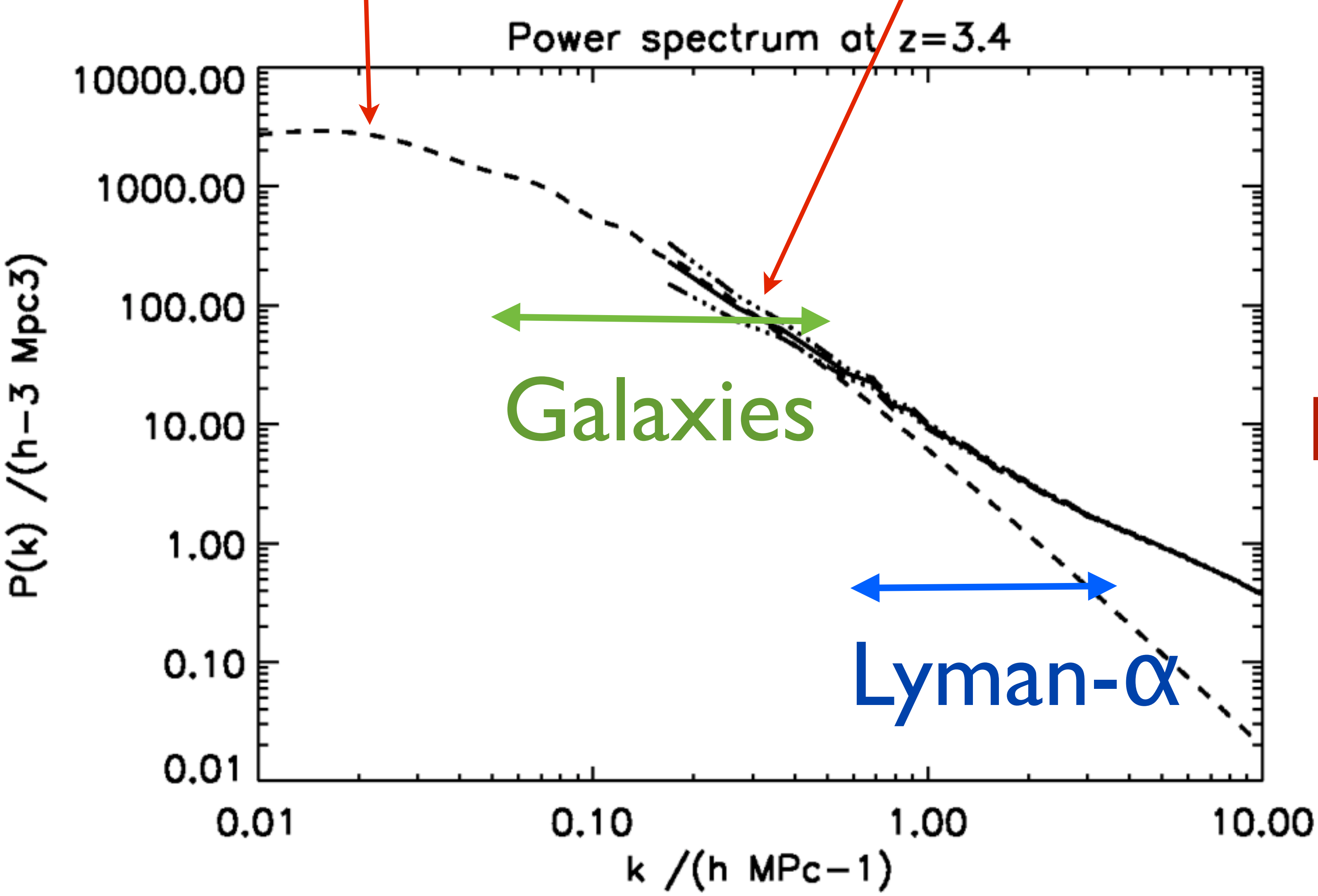
Lyman-α gives 3D map of dark matter clustering over time

Image: J. Bolton

Lyman-α forest

Linear power spectrum

Simulation power spectrum

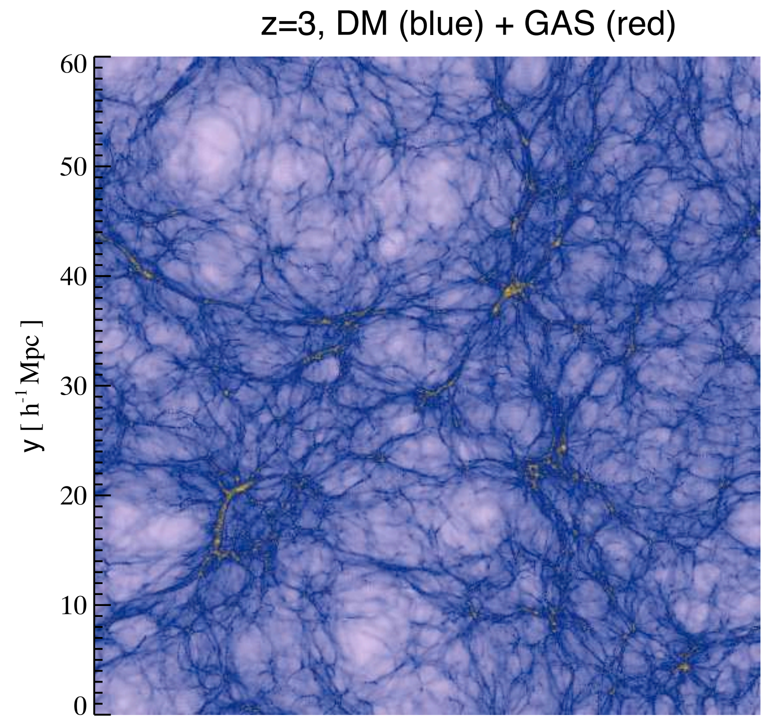


Lyman-α probes smallest scales

Mildly nonlinear physics

Lyman-α forest

- Structure growth tells us the initial conditions
- Primordial power spectrum



Lyman-α forest

Absorption with optical depth, τ

$$\mathrm{Flux} = \exp(-\tau)$$

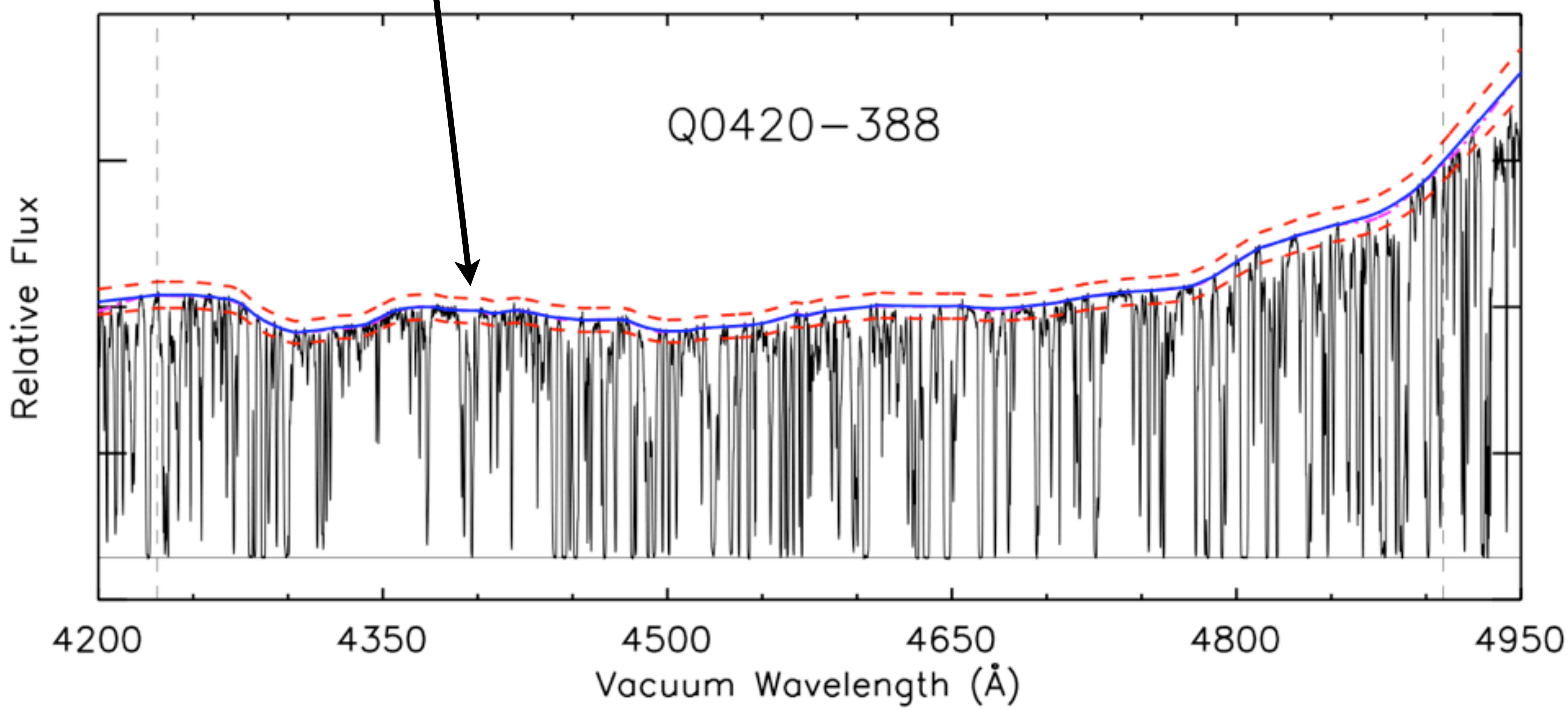


Image: Kim et al

Lyman-α forest

Observable: Flux Power spectrum

Fairly insensitive to small-scale structure

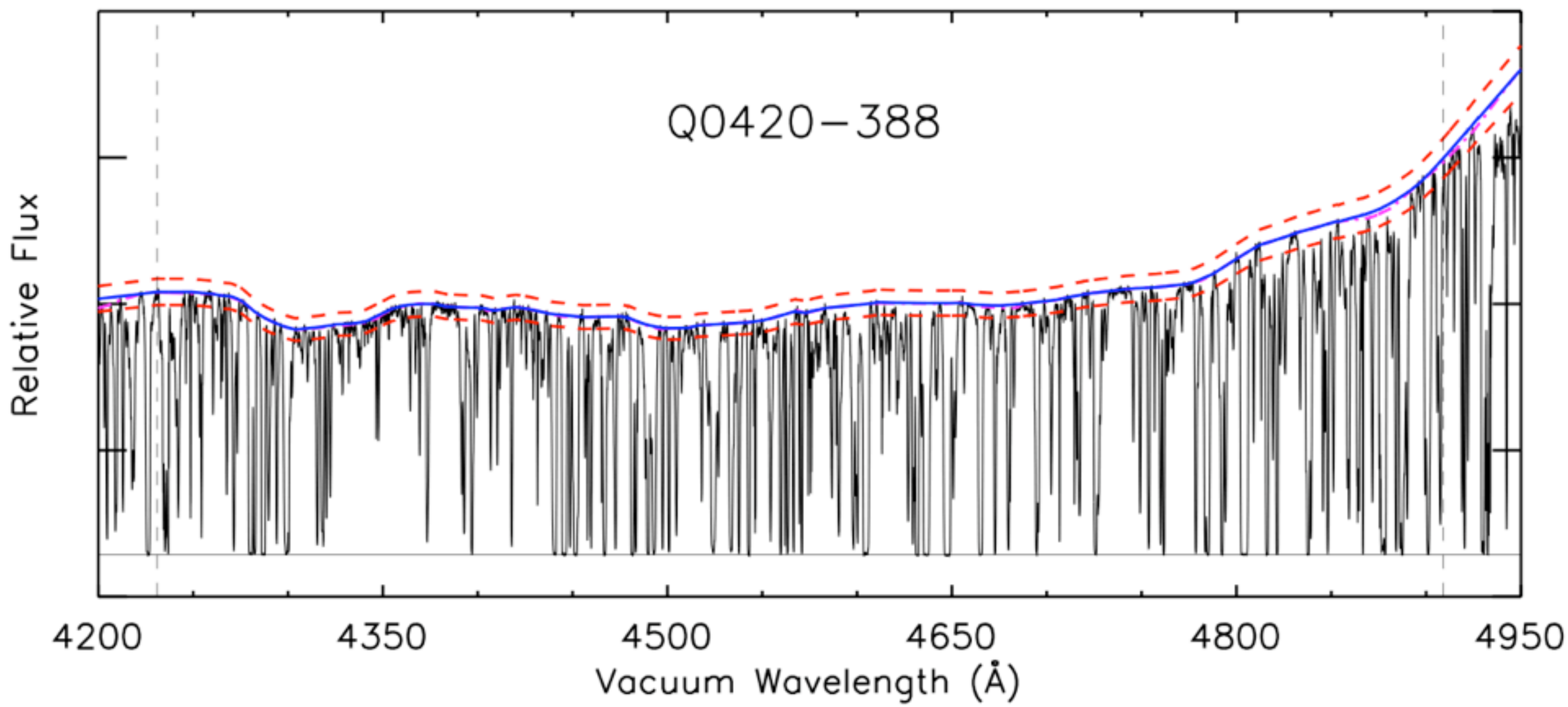


Image: Kim et al

SDSS

- 2.5m telescope in Apache Pt, New Mexico
- Takes enormous number of spectra
- Quantity over quality

SDSS

- Spectra need not have high S/N
- Instead need sky coverage and high density

Motivation

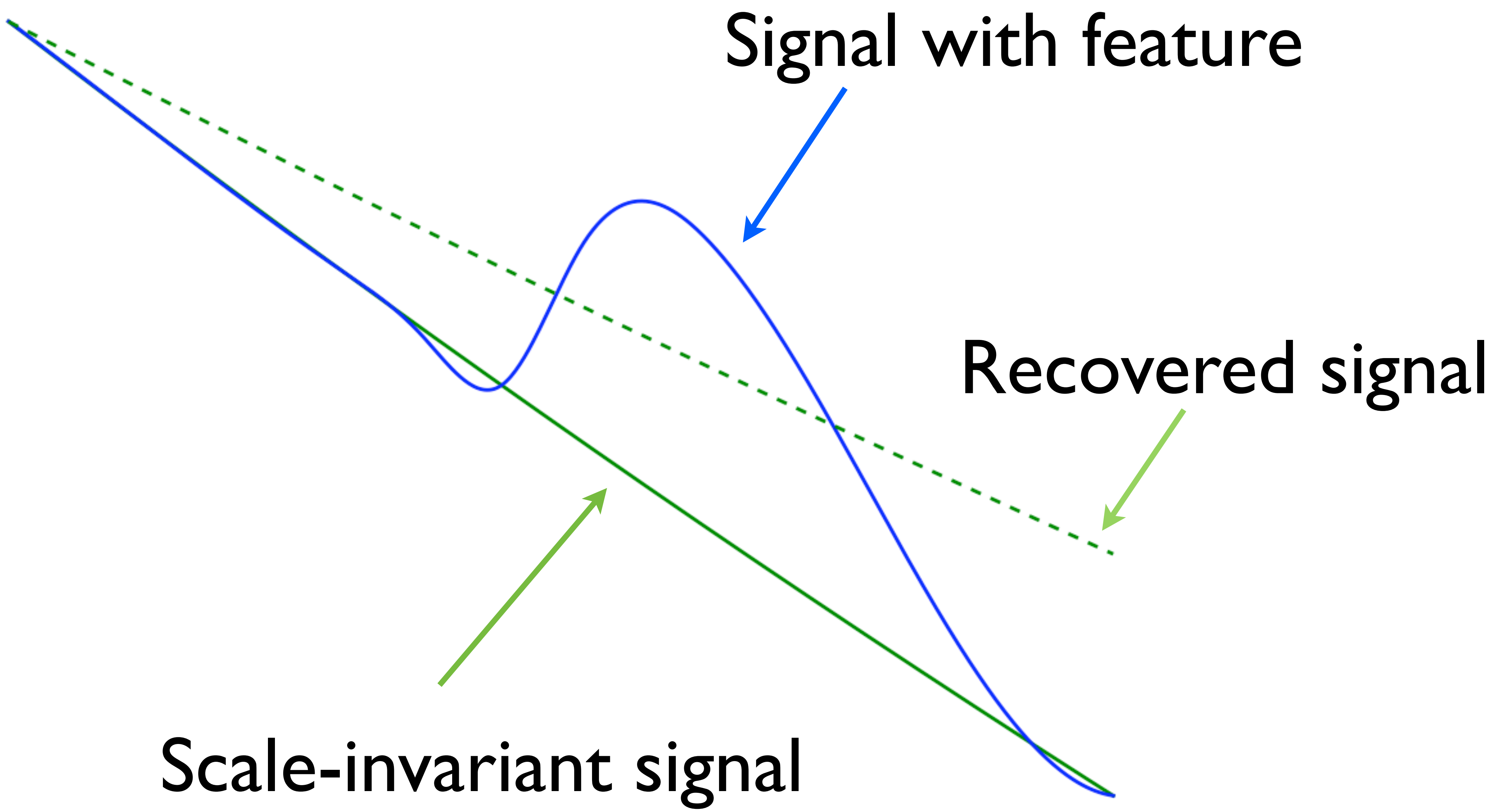
- Inflation predicts a nearly scale invariant smooth power law power spectrum.
- How strongly does the data support this?

Motivation

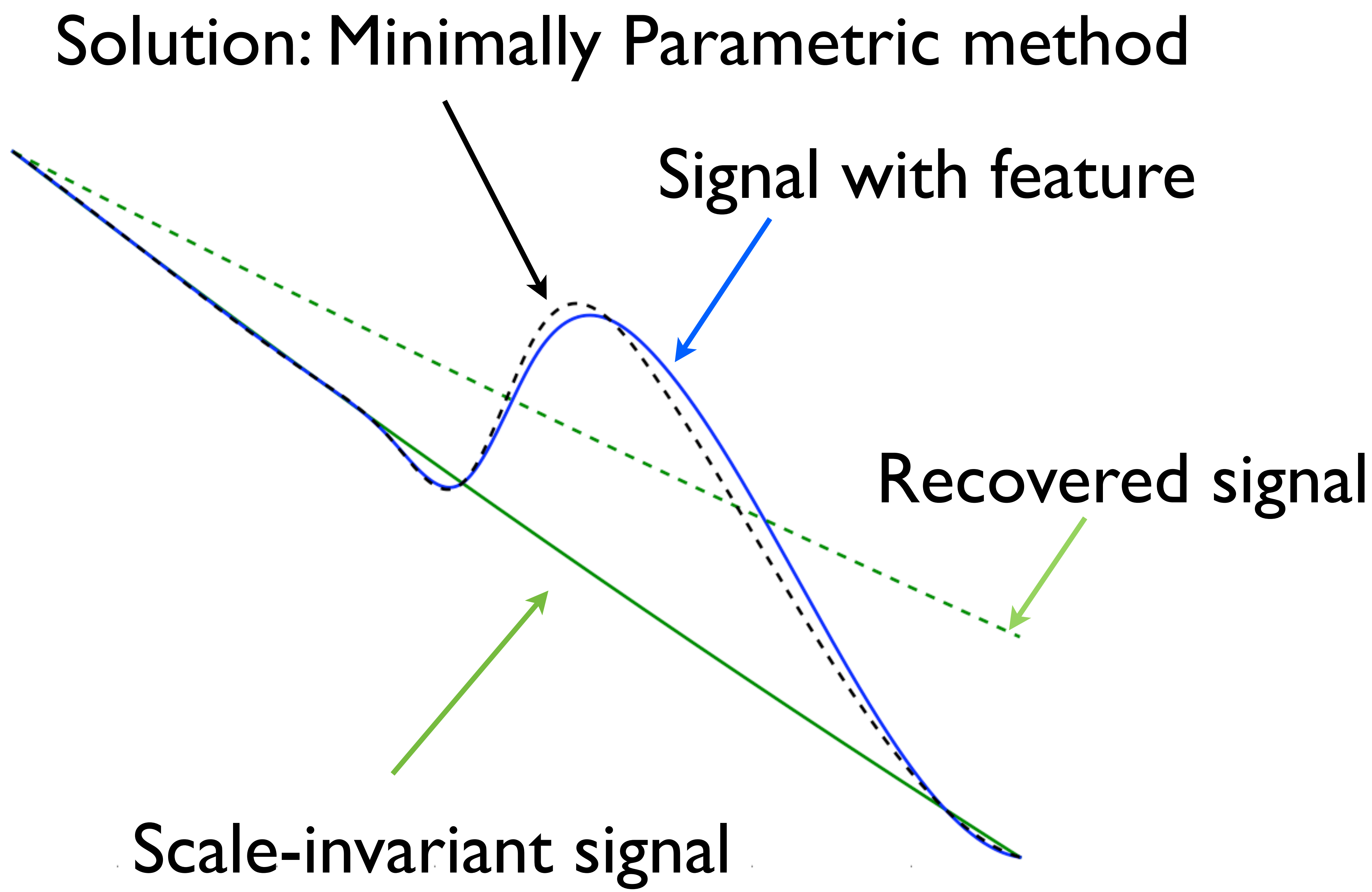
- Inflation predicts a nearly scale invariant smooth power law power spectrum.
- How strongly does the data support this?
- Lyman-α currently only direct probe of small-scale power spectrum.

Motivation

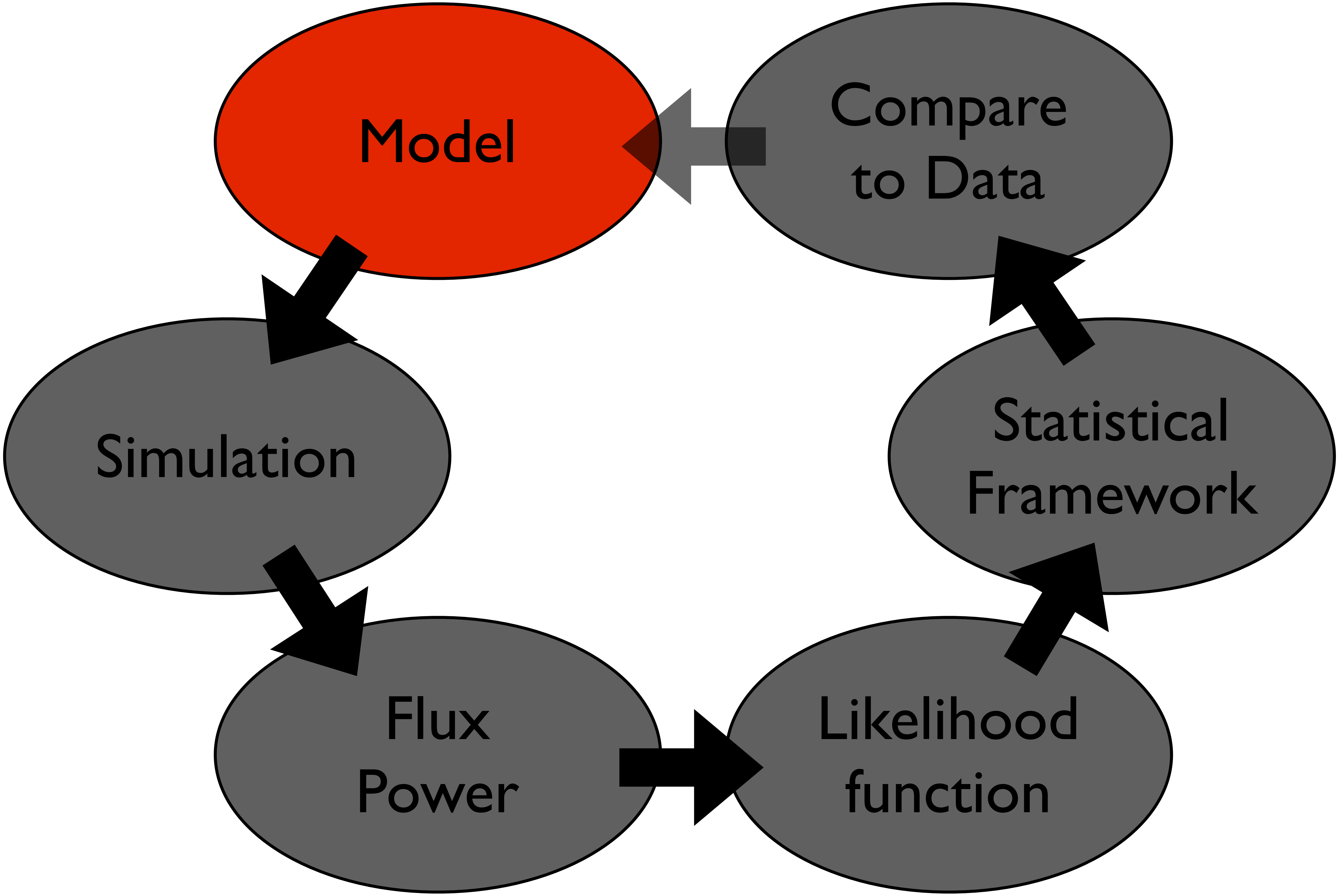
Local feature may bias recovered parameters



Motivation



Need to ensure robustness


Model
Compare to Data
Simulation
Statistical Framework
Flux Power
Likelihood function

Power Spectrum Reconstruction

Power law primordial power spectrum:

$$P(k) = A_s \left(\frac{k}{k_0} \right)^{n_s - 1}$$

Do parameter estimation.

Power Spectrum Reconstruction

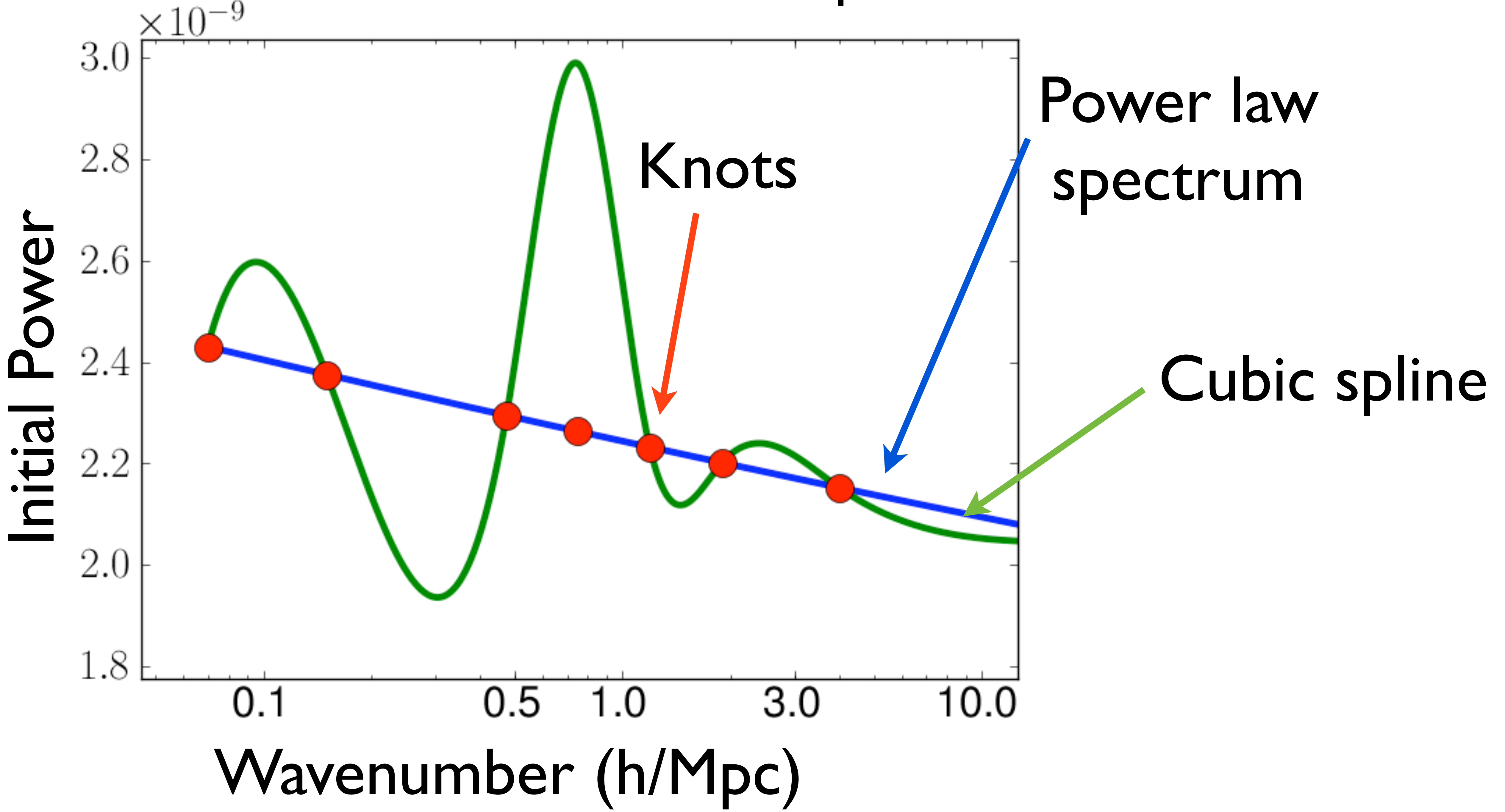
Power law primordial power spectrum:

$$P(k) = A_s \left(\frac{k}{k_0} \right)^{n_s - 1}$$

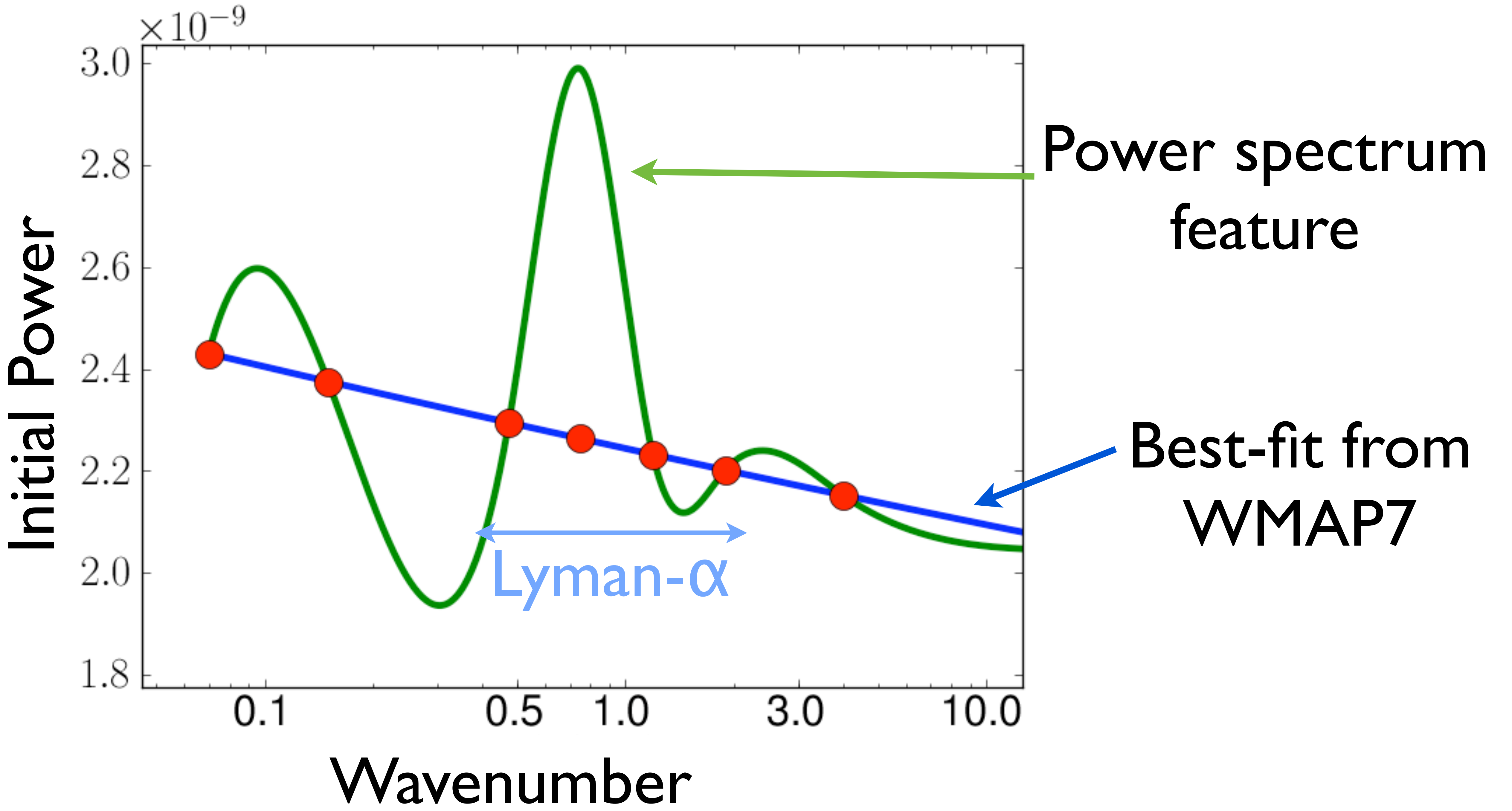
Do parameter estimation.

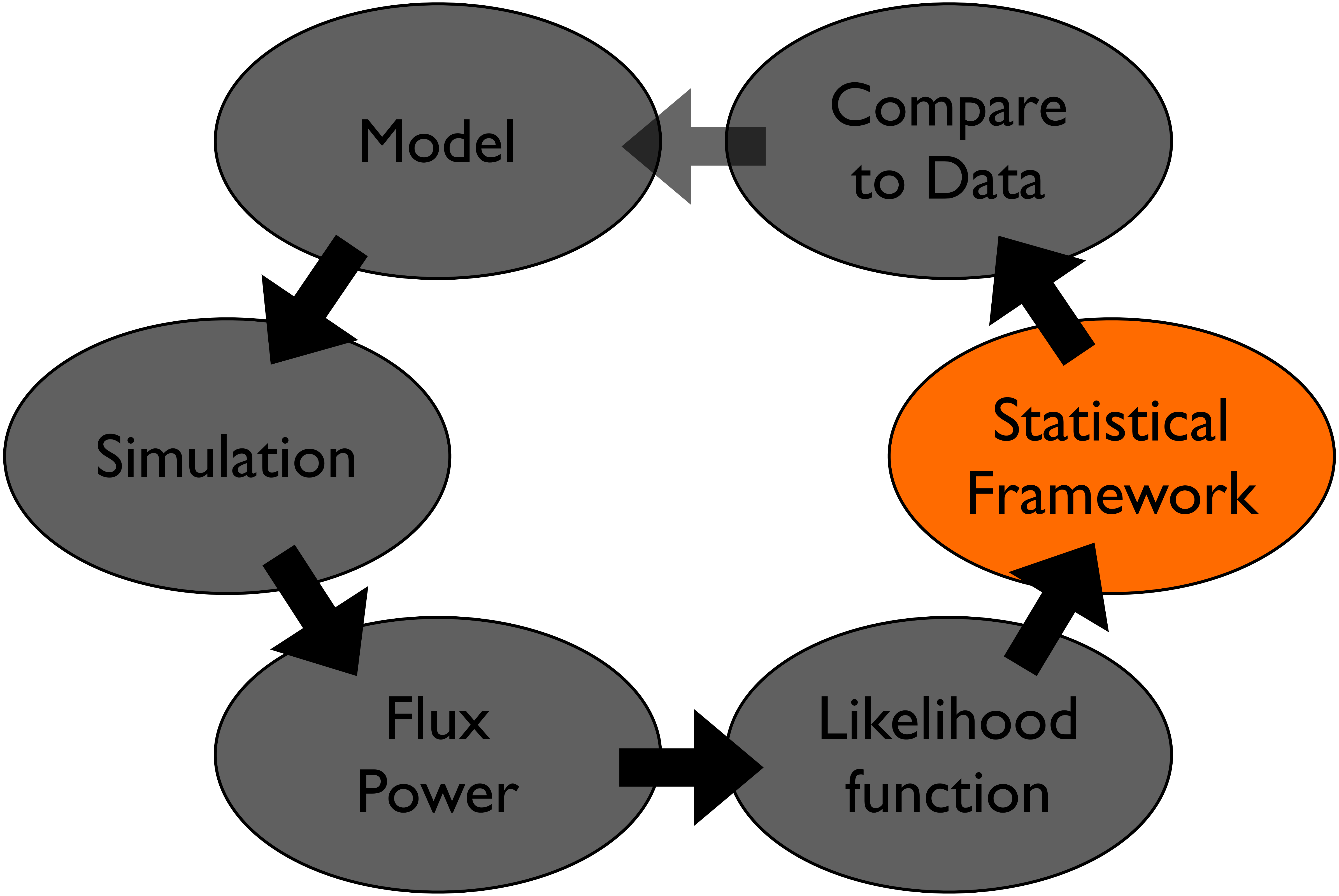
Reconstruction

Fit with cubic spline.



Reconstruction



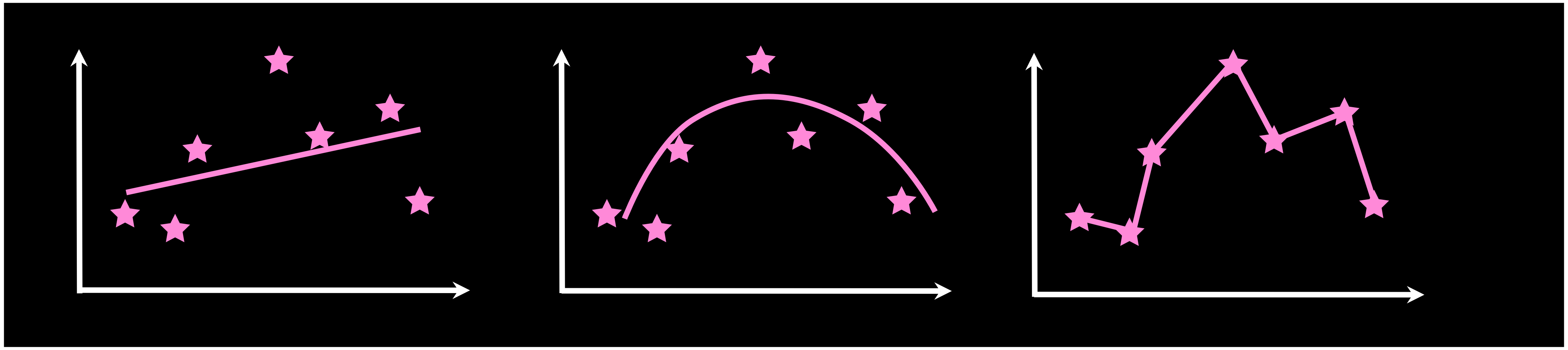

Model
Compare to Data
Simulation
Statistical Framework
Flux Power
Likelihood function

Reconstruction

We need to fit the signal, but NOT the noise

Use cross-validation: similar to jack-knifing.

Which is best?



Power Spectrum

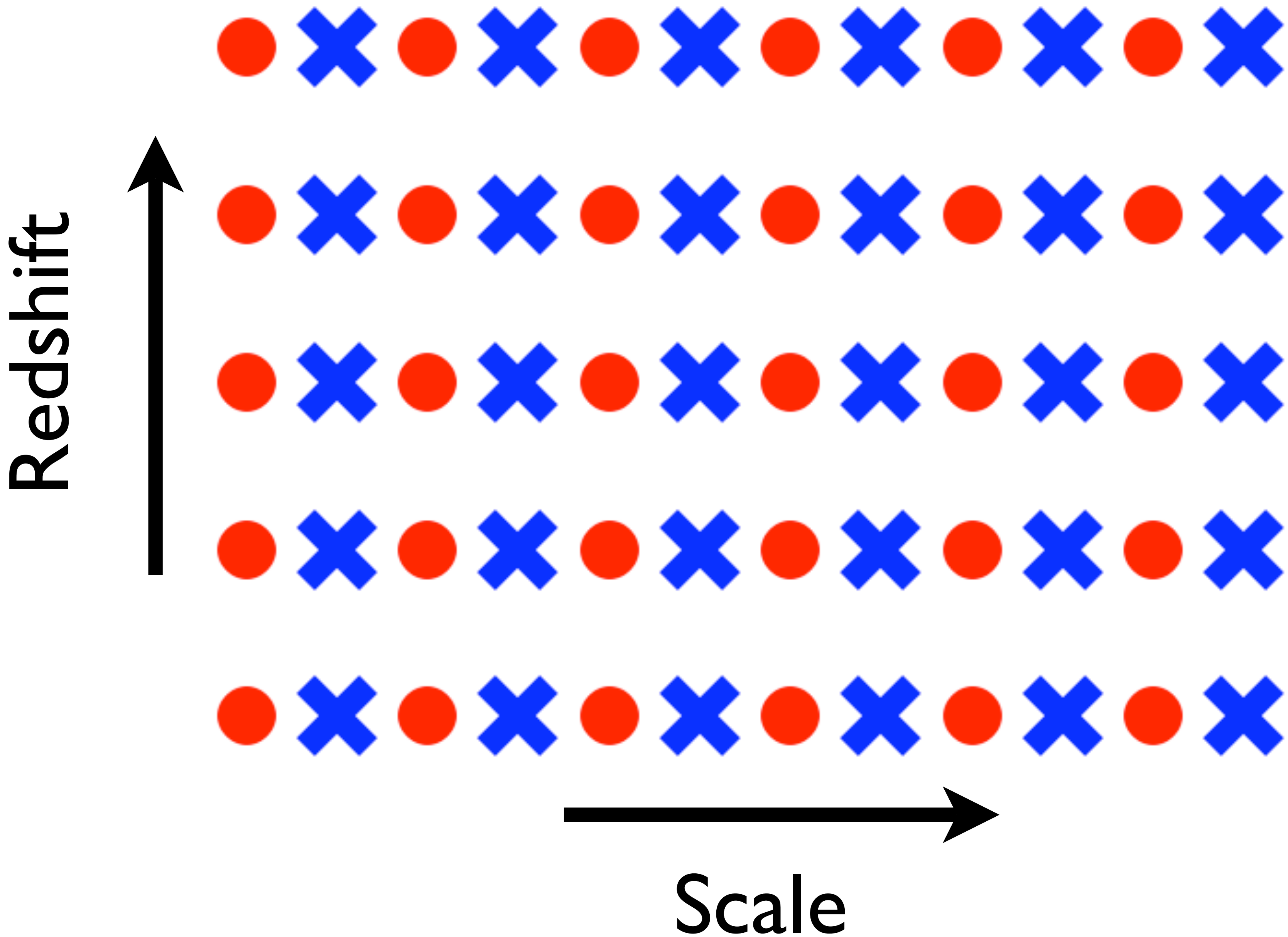
Noise is extra small-scale variation.

Likelihood to penalise "wiggly" shapes:

$$\log \mathcal{L} = \log \mathcal{L}(\mathrm{Data}|P(k)) + \lambda \int_k dk (P''(k))^2$$

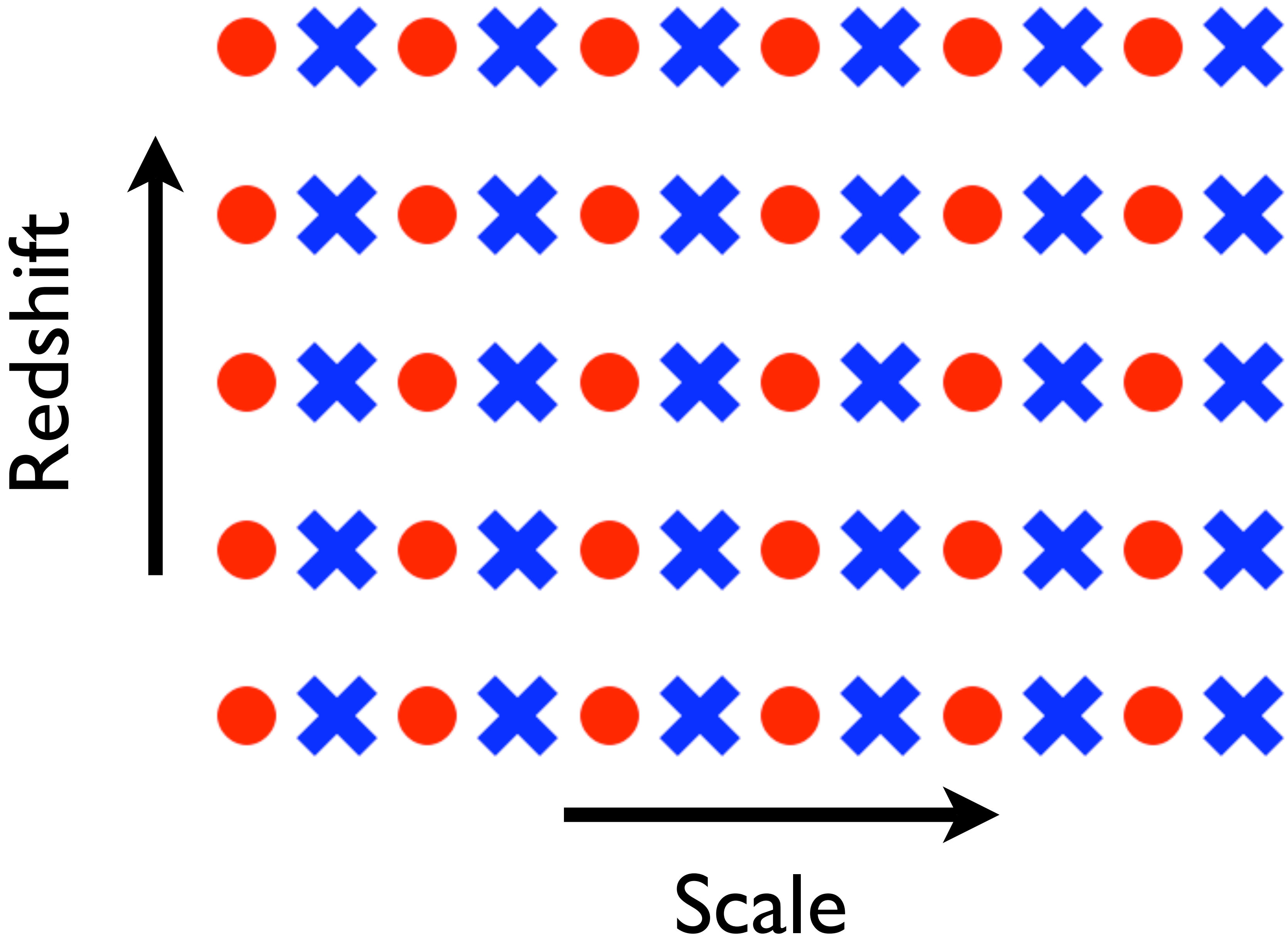
Cross-validation to choose penalty most accepted by data

Cross-Validation



Divide data into: training set (crosses) and validation set (circles)

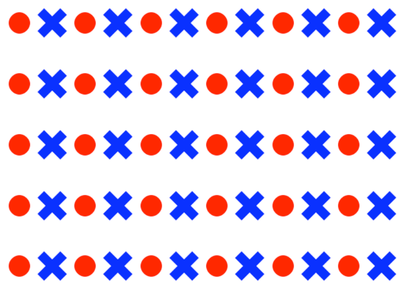
Cross-Validation



Training set should predict validation set

Cross-Validation

1. Pick penalty.
2. Find best fit to training set
3. Predict validation set from best-fit
4. Find penalty which best predicts validation set



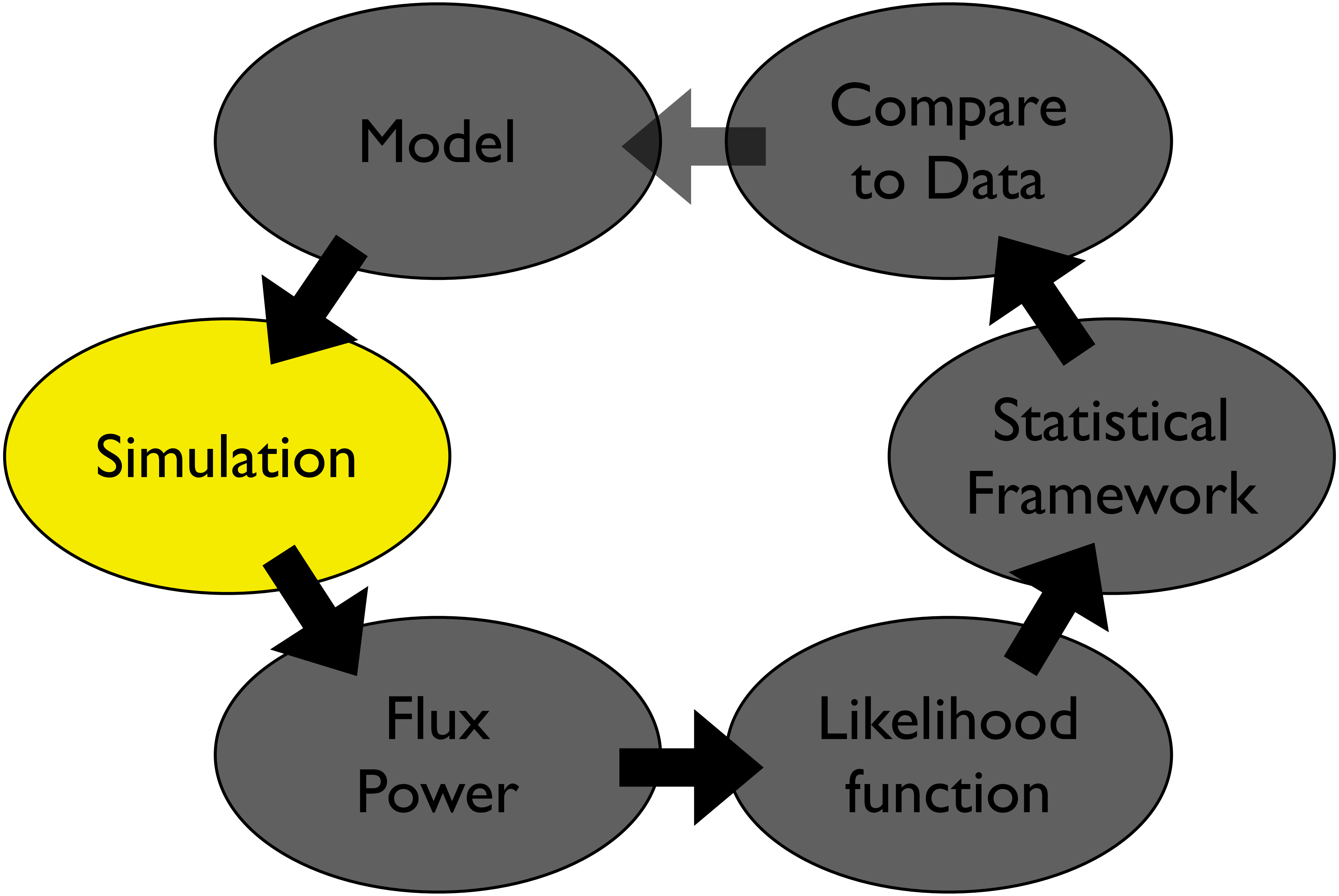
$$\log \mathcal{L} = \log \mathcal{L}(\mathrm{Data}|P(k)) + \lambda \int_k dk (P''(k))^2$$

Parameter Estimation

- Assume data Gaussian: N(μ,σ)
- Find μ, σ in best agreement with data

Minimally Parametric

- Choose some form F(μ,σ)
- Find μ, σ in agreement with training data
- Check how well F(μ,σ) predicts validation data



Model
Compare to Data
Simulation
Statistical Framework
Flux Power
Likelihood function

Motivation

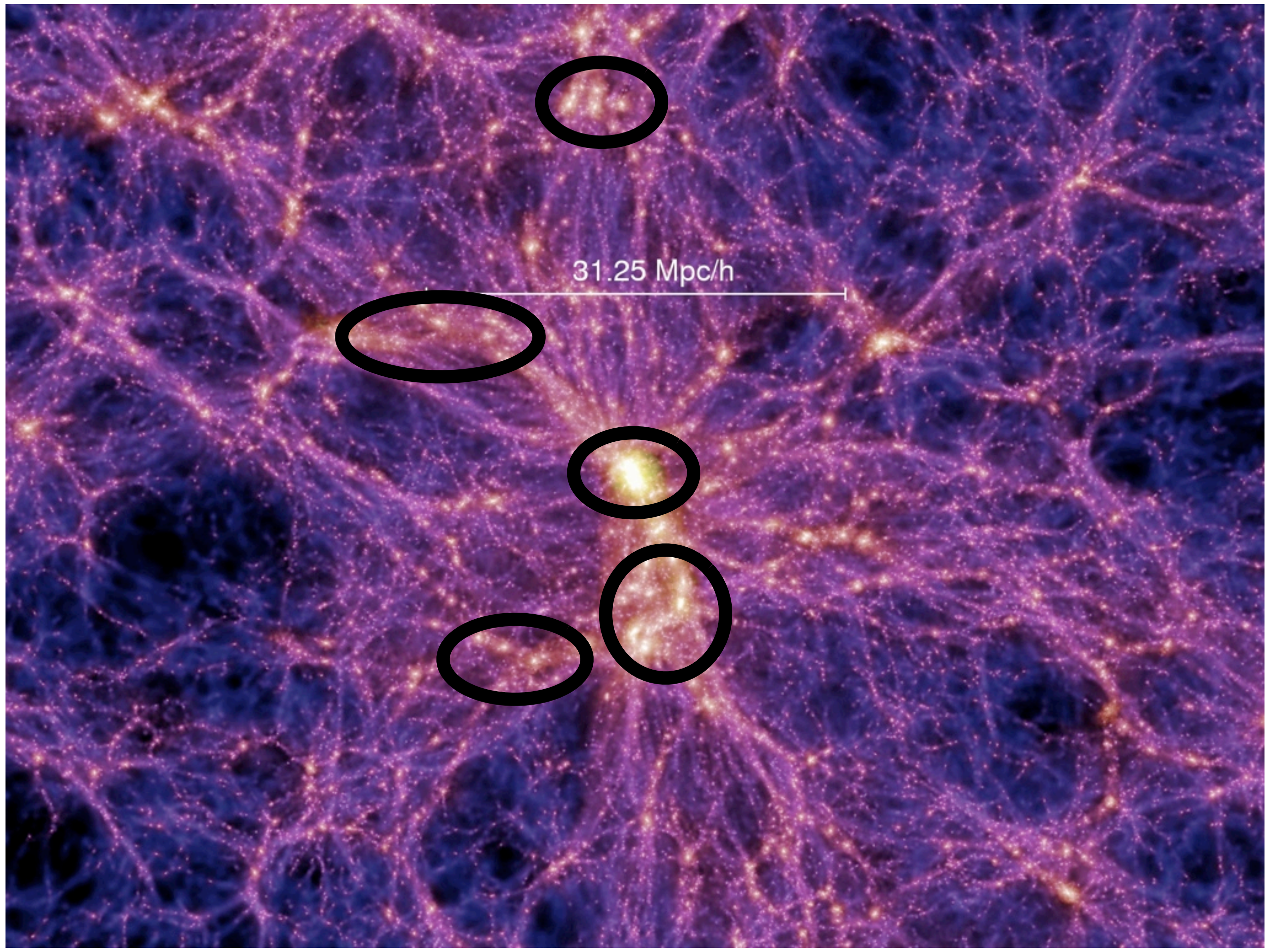
Why do we need new simulations?

- Structure nonlinear
- Need to construct a map between P(k) and flux statistics: depends on baryonic physics
- Previous map assumed scale-invariance

Simulation Setup

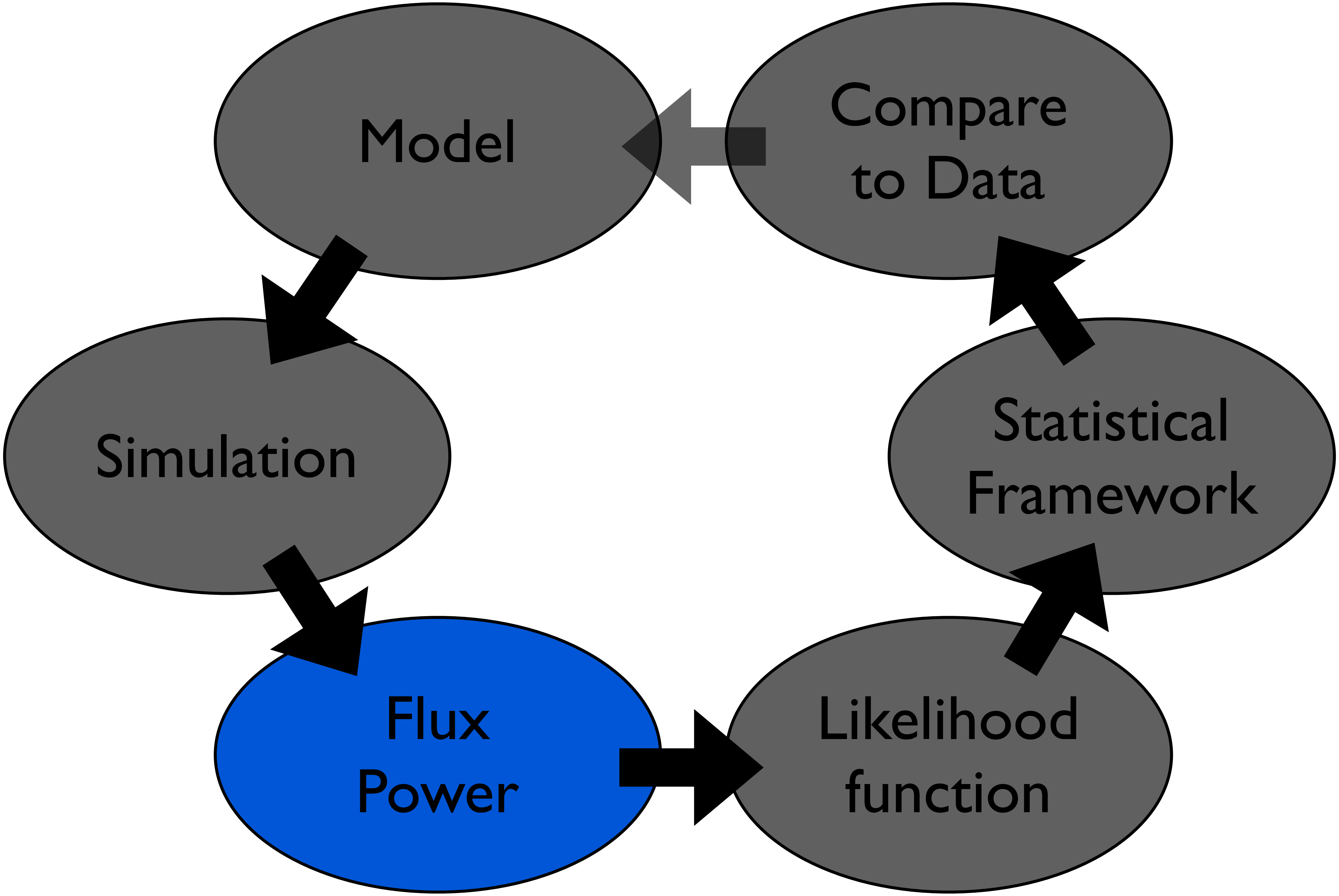
- 30+ hydrodynamic simulations using GADGET-II.
- 60 Mpc box, 2x400^3 particles
- 400^3 dark matter particles - collisionless
- 400^3 baryons - with cooling

Important Trick

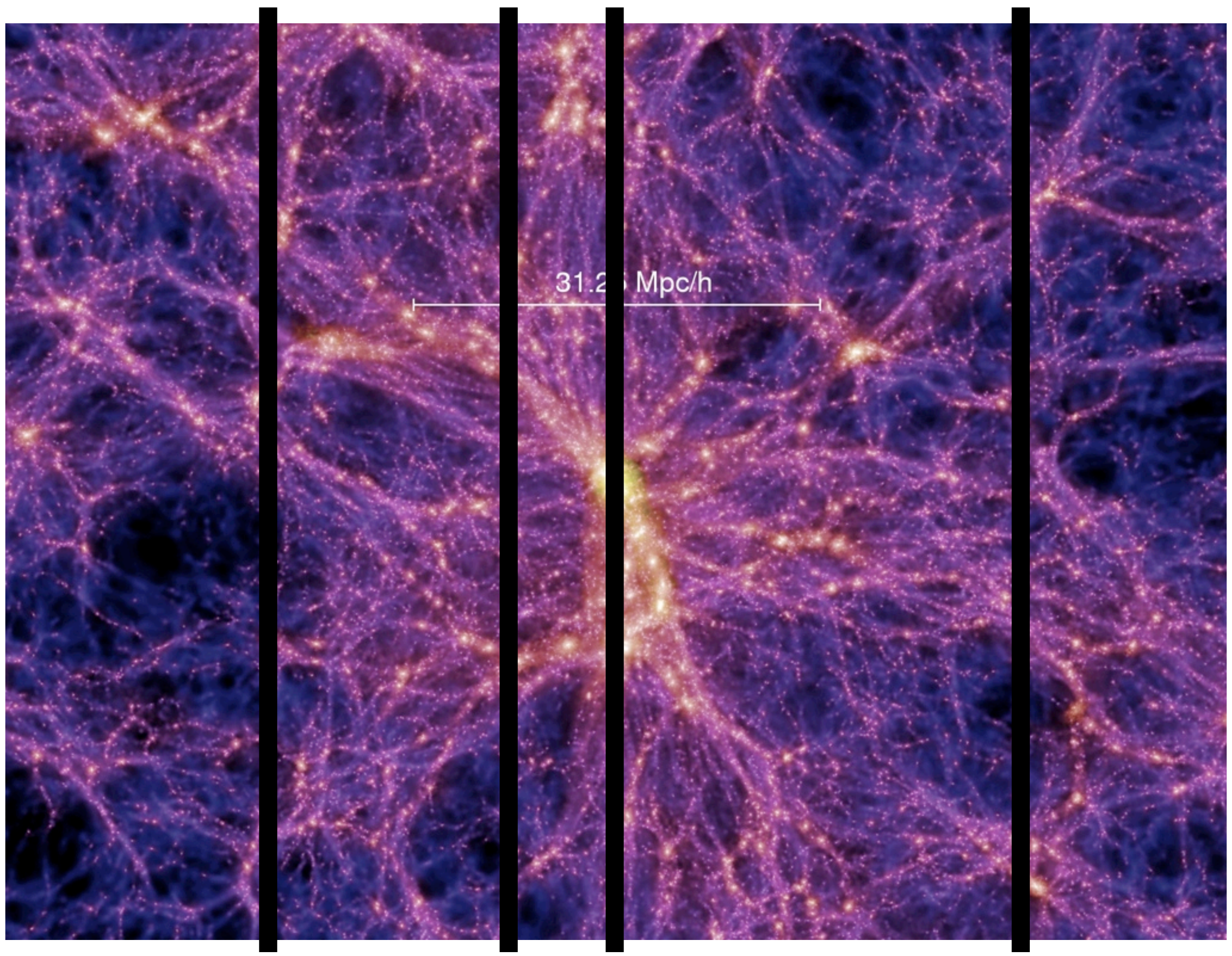


- Dense regions have many slow collisions
- Do not influence the Lyman-α forest
- Save time by making dense regions “stars”

Image: Millennium Simulation


Model
Compare to Data
Simulation
Statistical Framework
Flux Power
Likelihood function

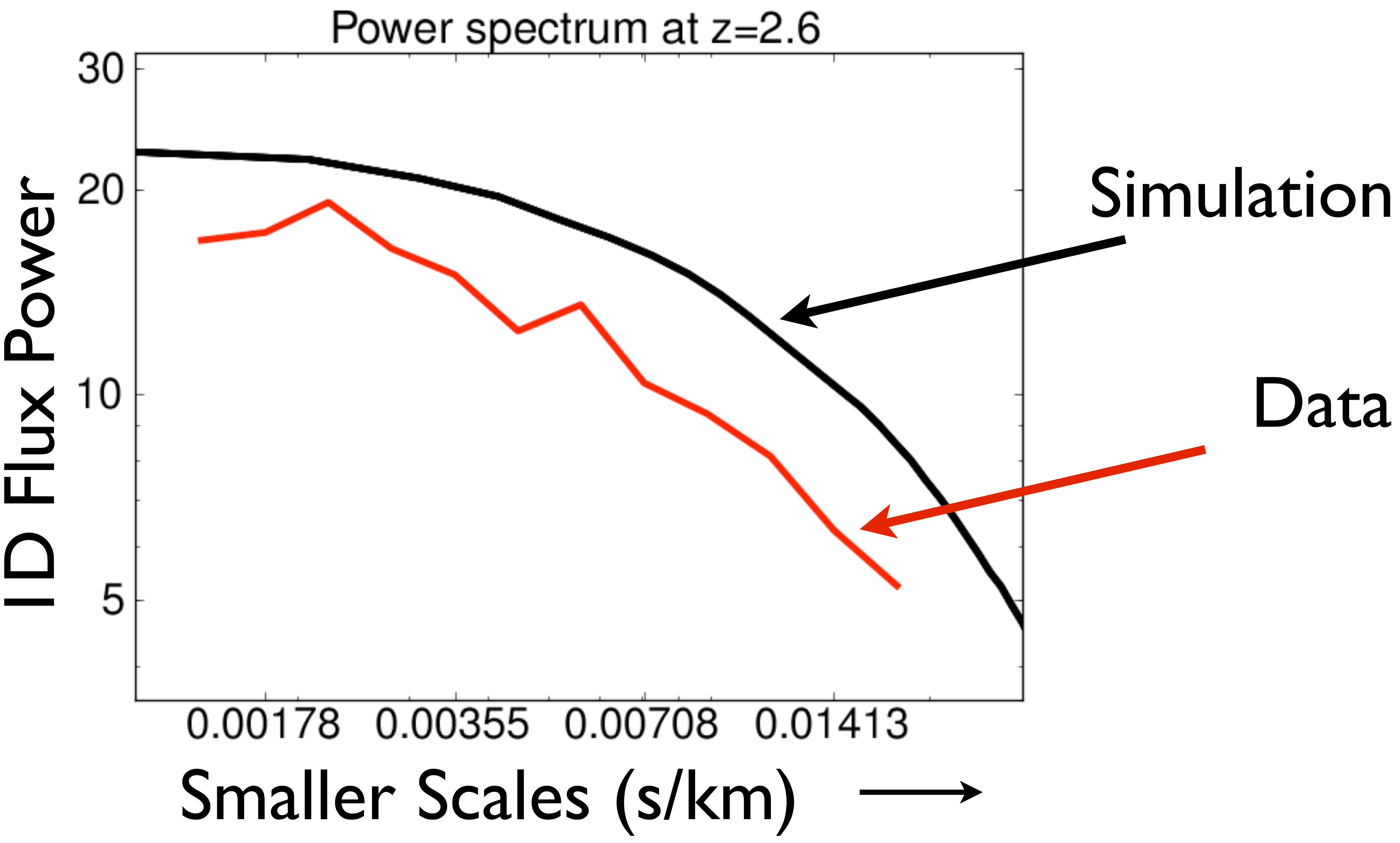
Simulated Spectra

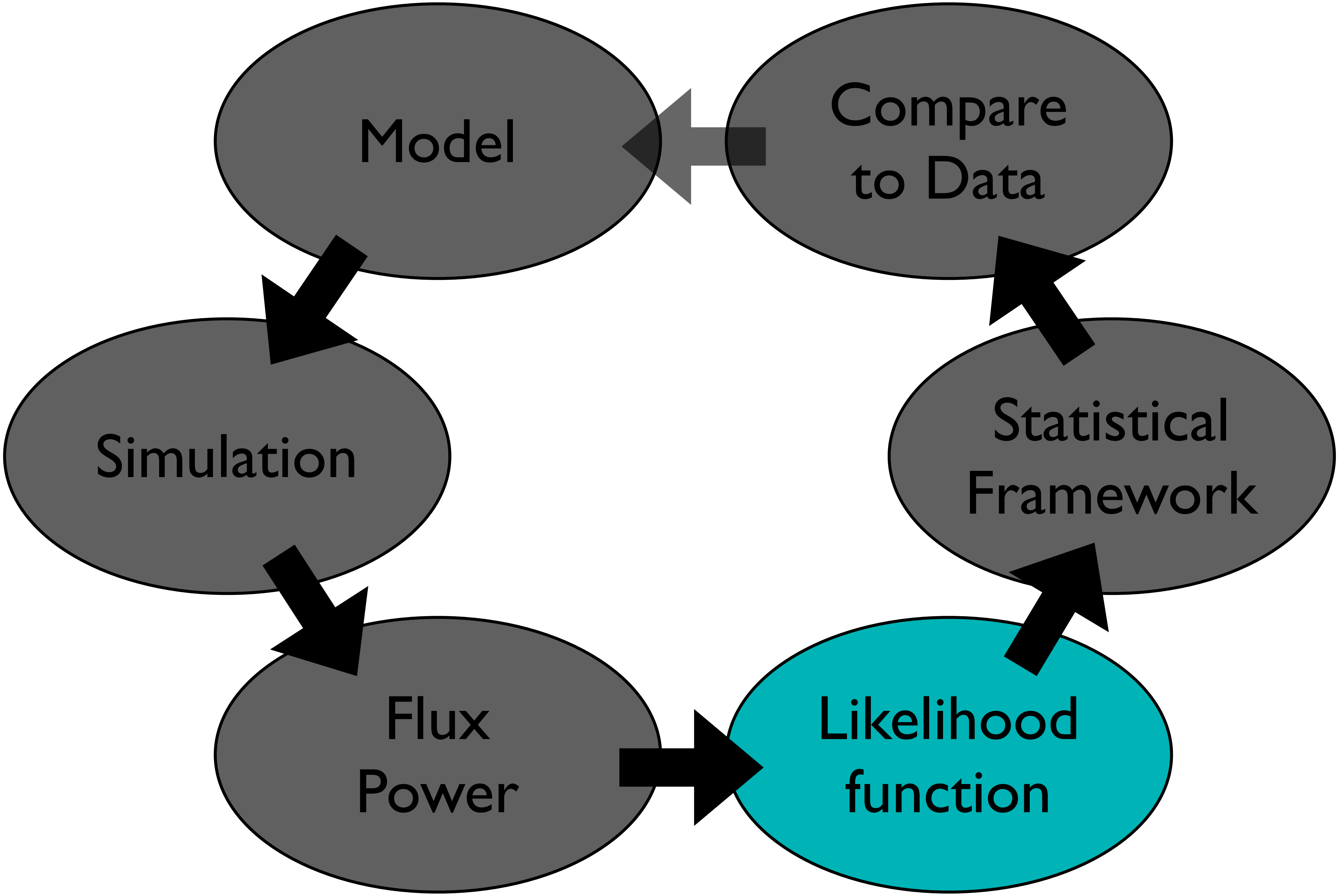


- Draw skewers through density field
- Calculate absorption along skewers
- Average of two-point statistics

Image: Millennium Simulation

Flux Power Spectrum





Model
Compare to Data
Simulation
Statistical Framework
Flux Power
Likelihood function

Likelihood Construction

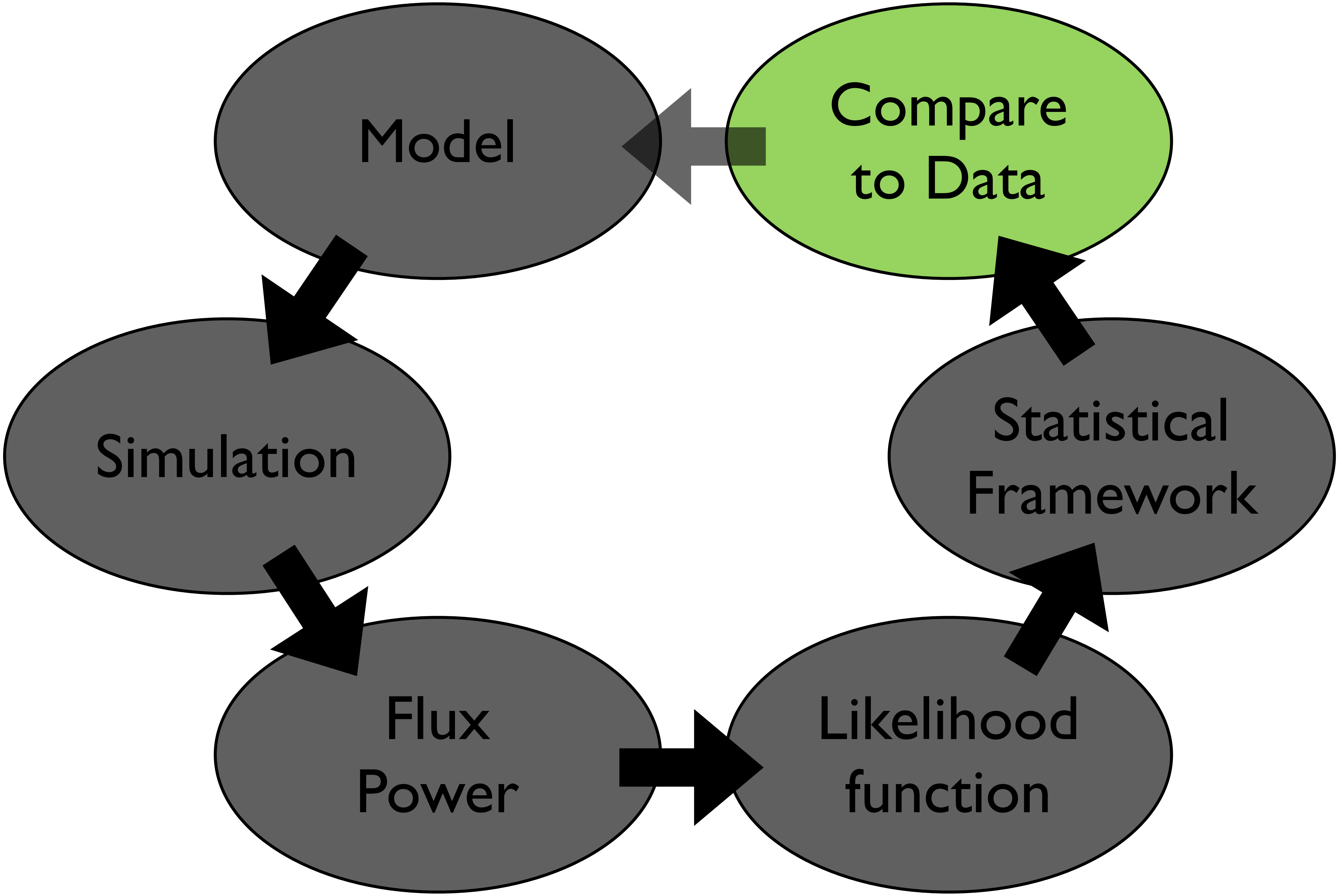
- Vary one parameter at a time.
- Fit change in flux power with a polynomial

$$\delta P_{\mathrm{F}}(p_i) = \Sigma_i \left(a\delta p_i^2 + b\delta p_i \right)$$

- Check accuracy with jack-knifing.

Likelihood Construction

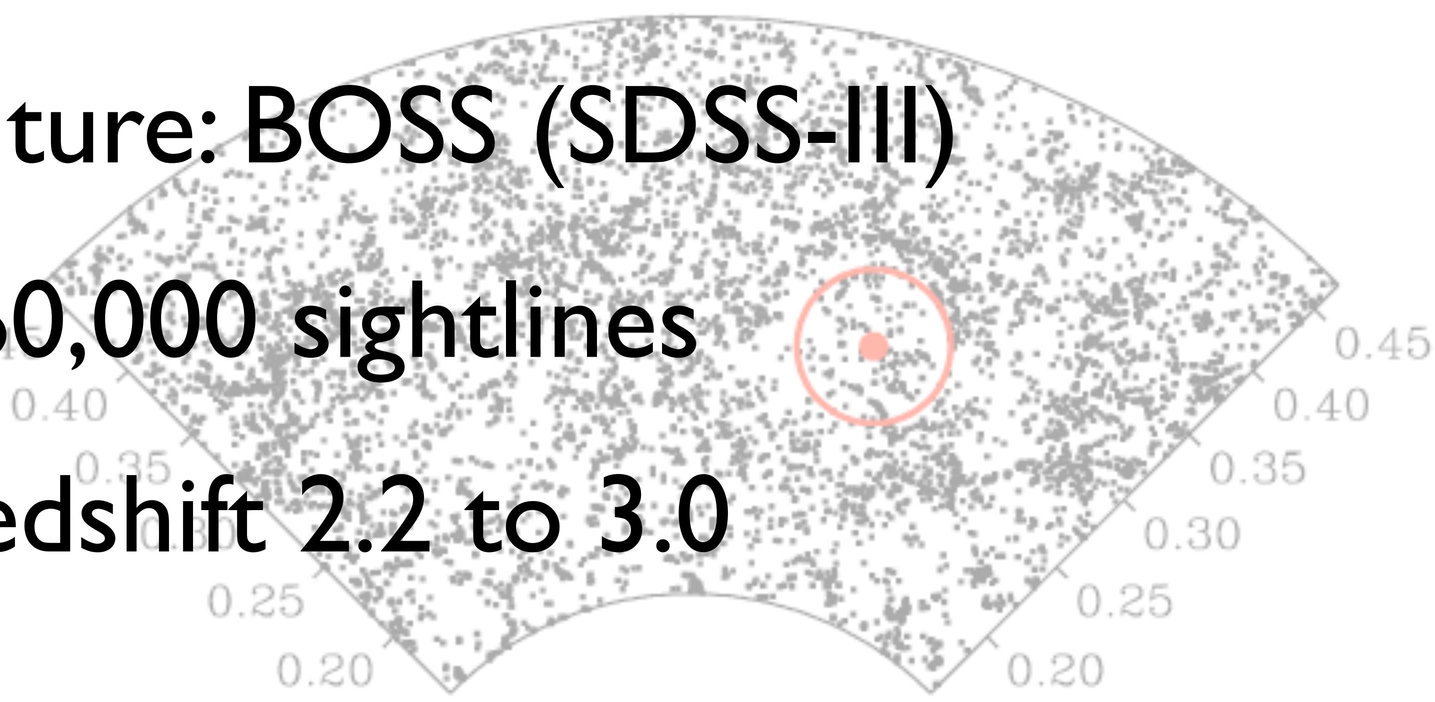
- Marginalise over thermal parameters:
  - Temperature
  - Temperature-density relation
  - Mean optical depth, aka ionising radiation density
- Correct for resolution and box effects, damping wings, SiIII, ...


Model
Compare to Data
Simulation
Statistical Framework
Flux Power
Likelihood function

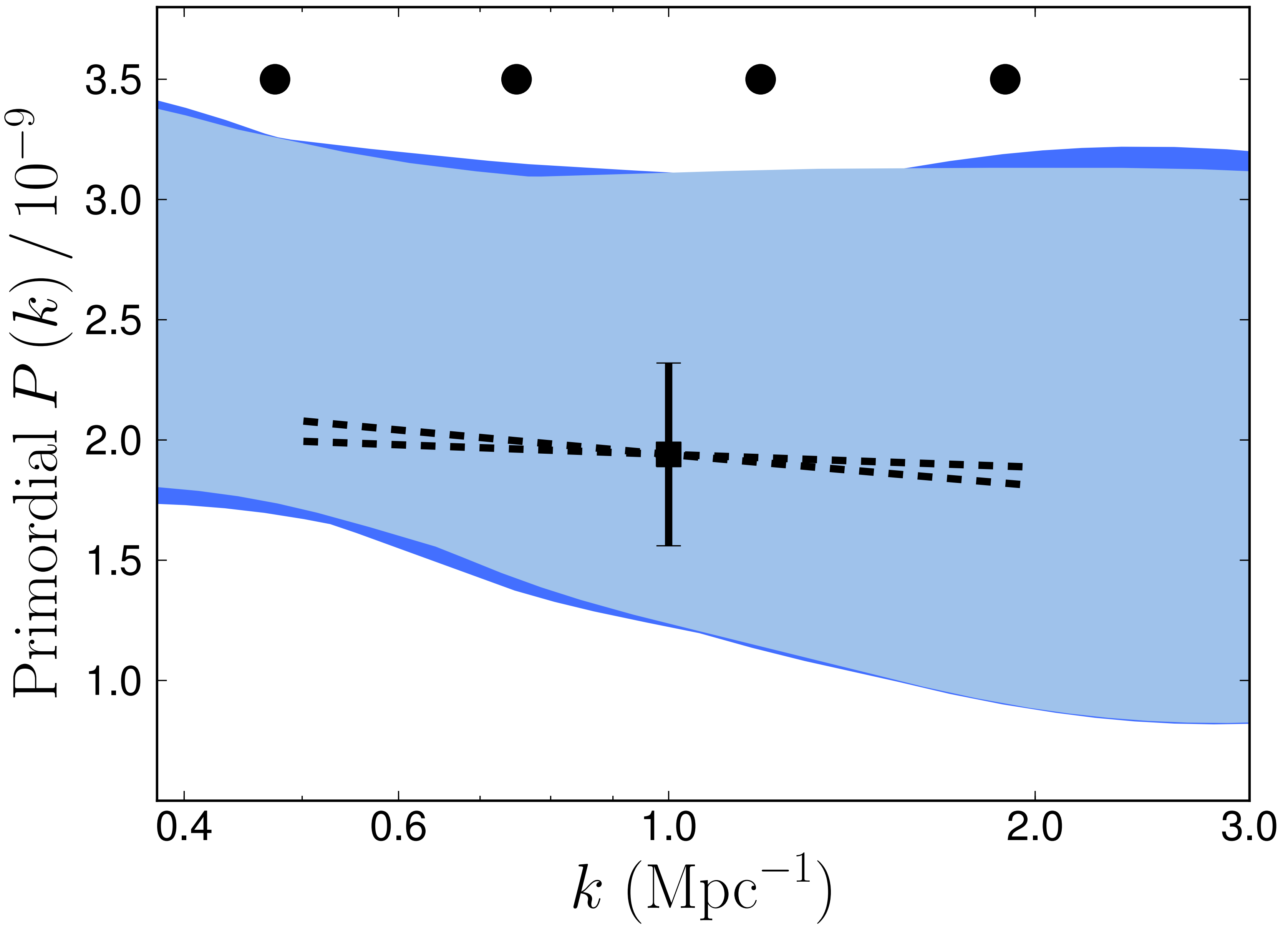
Data Comparison

- Current data: SDSS quasar flux power spectrum from McDonald et al 2005.
- ~3000 quasar sightlines
- Redshift 2.2 to 4.2

- Future: BOSS (SDSS-III)
- 160,000 sightlines
- Redshift 2.2 to 3.0



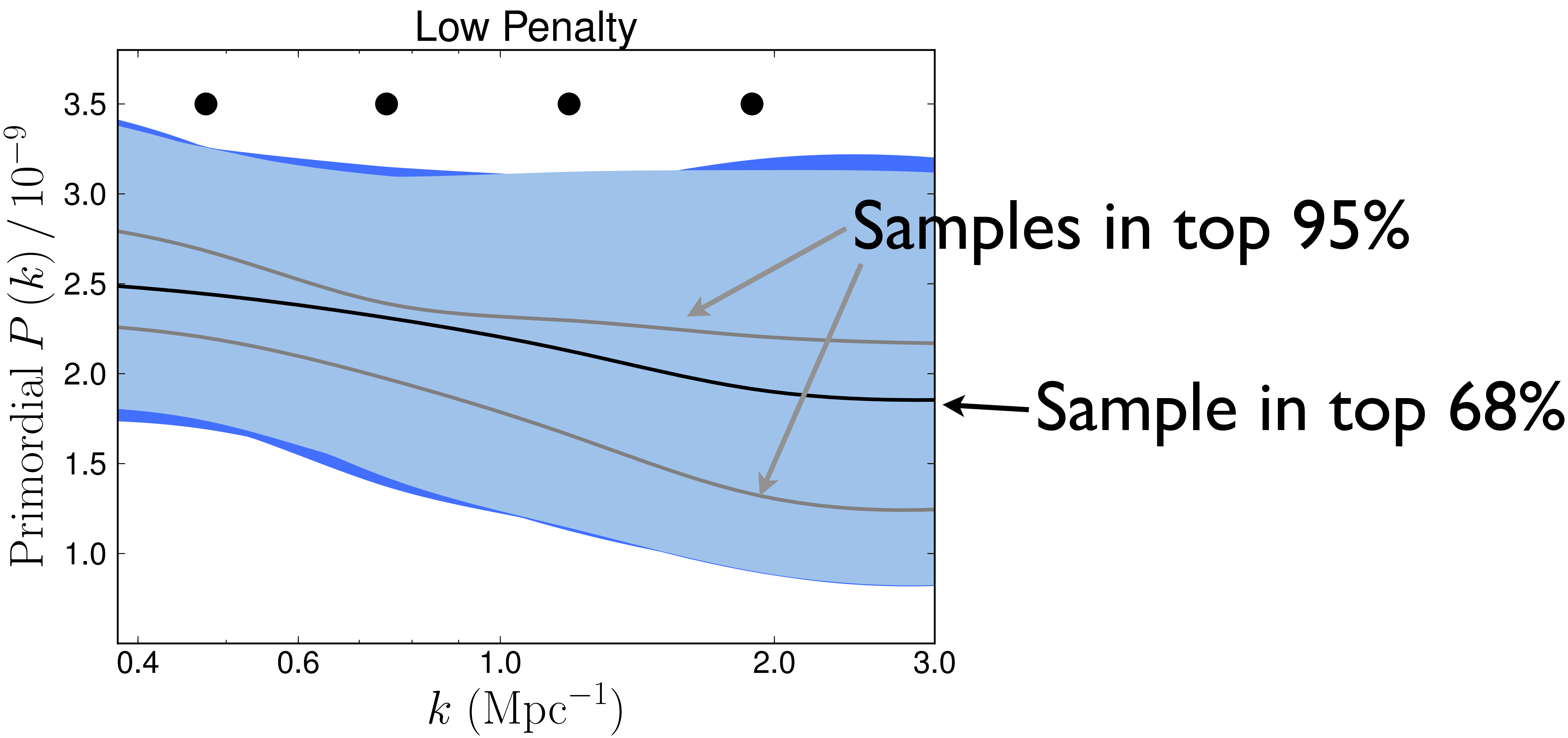
Results



- "Envelope" of splines with likelihood in top 95%.

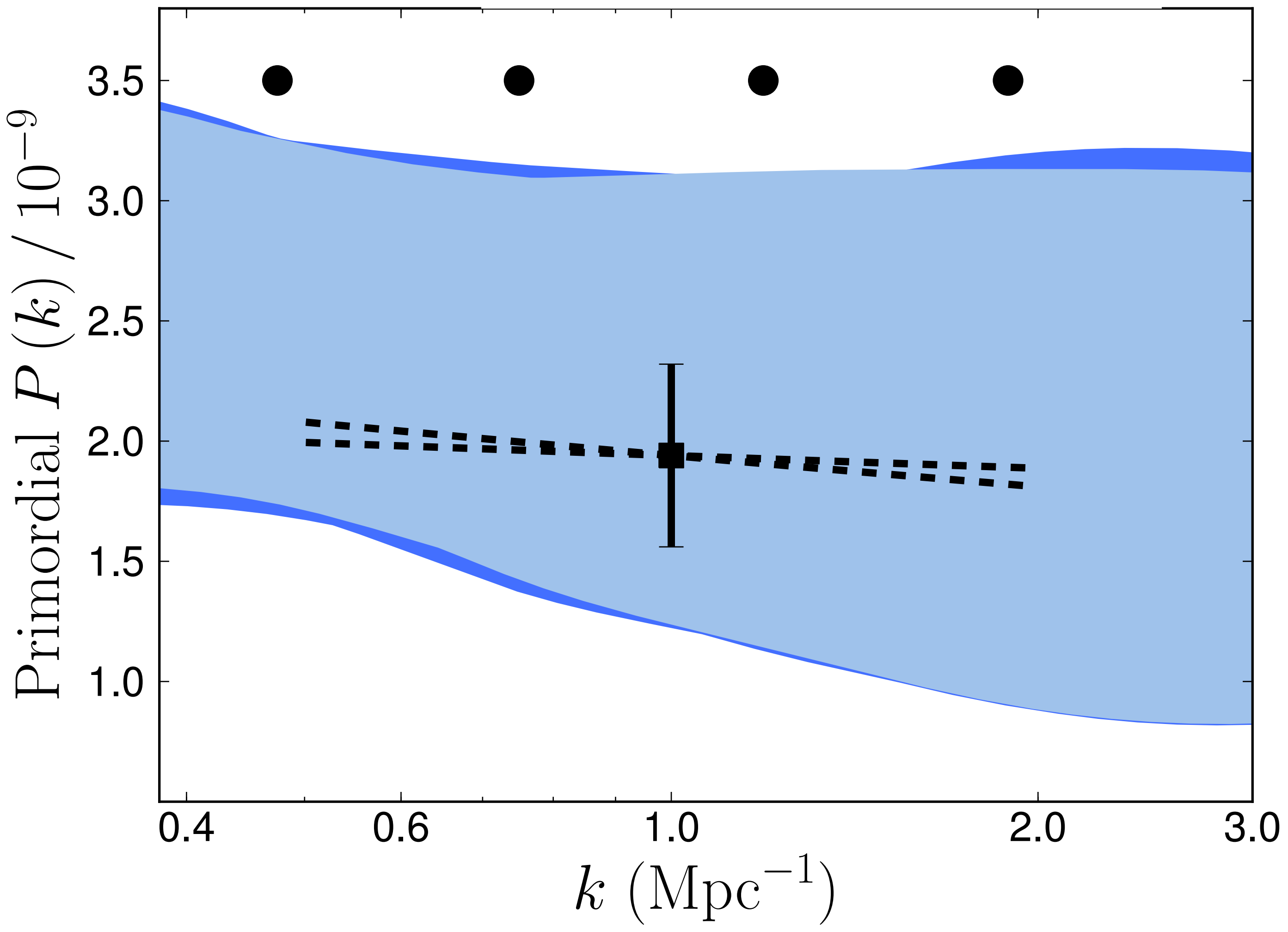
Bird et al (2010), arxiv:1010.1519

Results



- 68% and 95% have similar envelopes; lower likelihood splines have more features.

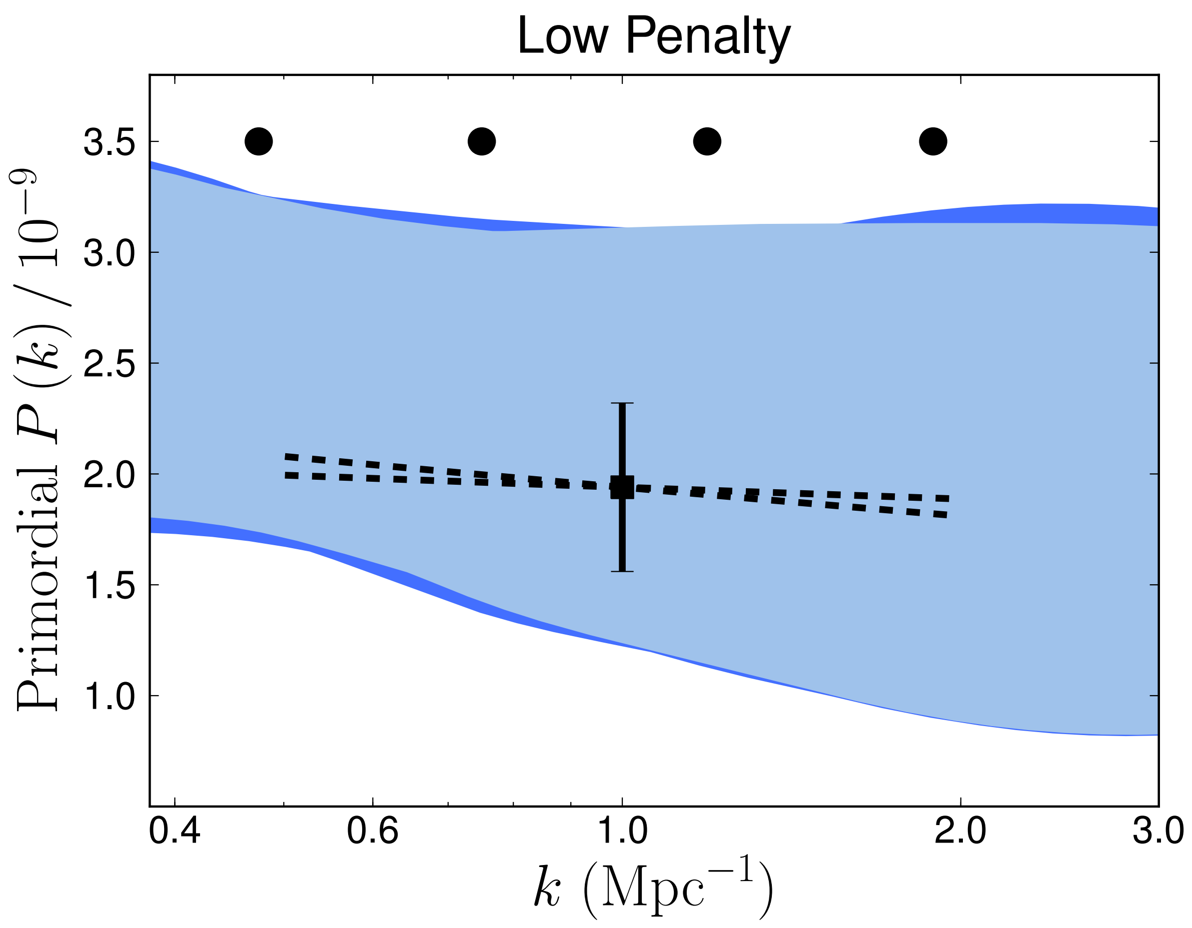
Results

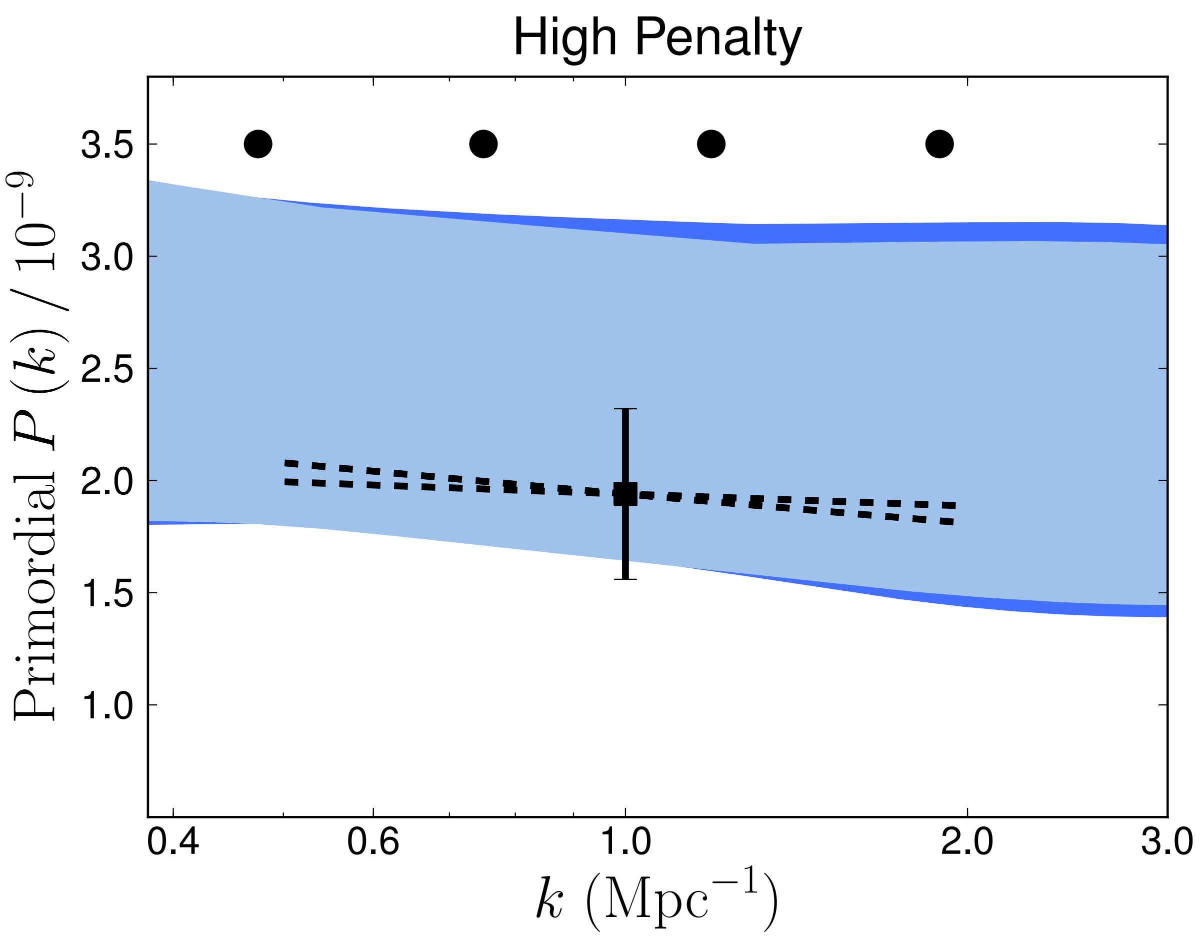


- Error bar shows constraints from parameter estimation
- Driven by prior assumption of power law form

A sufficiently high penalty reproduces the previous results.

Results



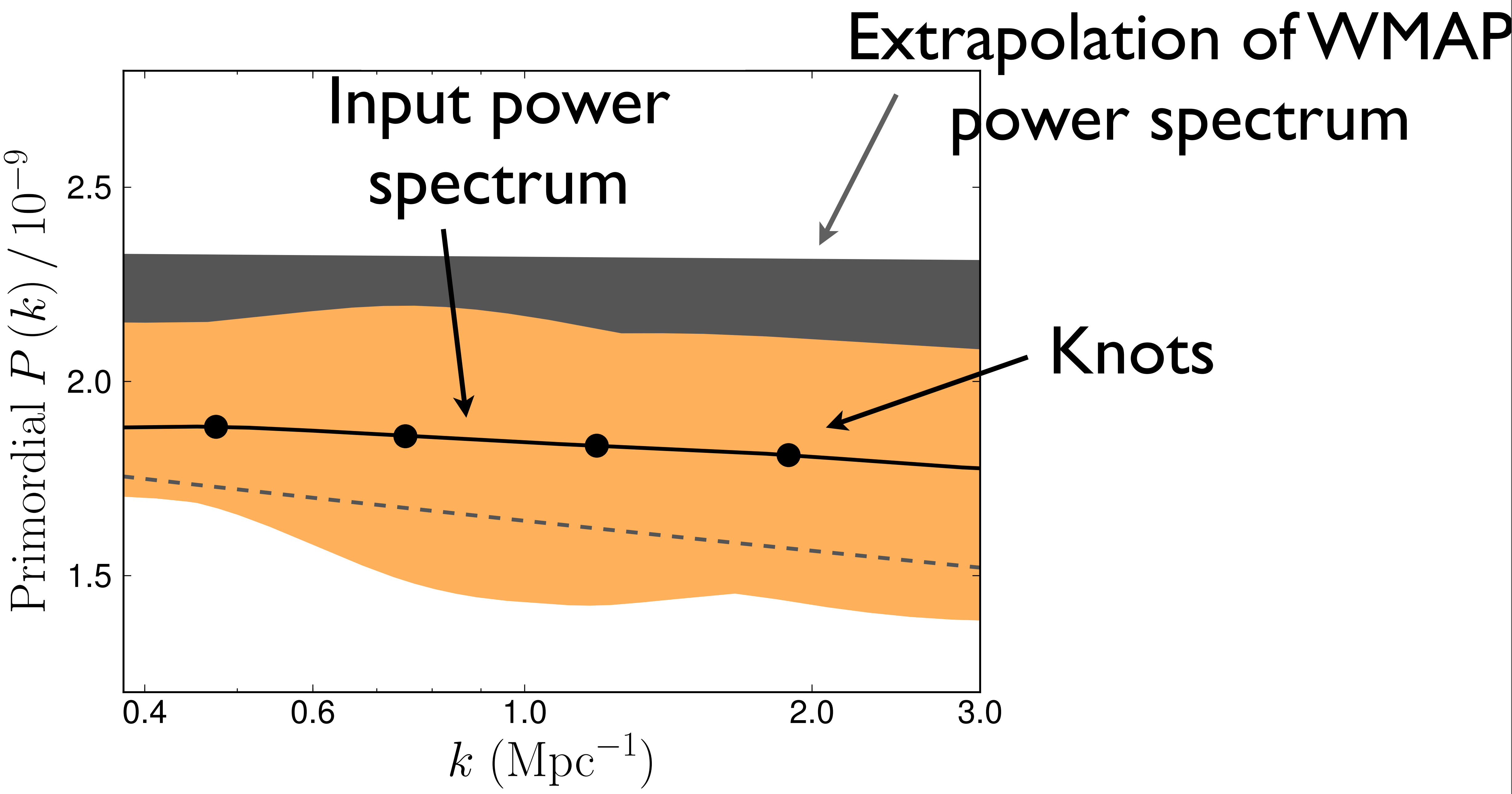


- CV score constant with penalty
- Cannot distinguish between above plots.

BOSS simulation

- Simulated flux power spectrum with theoretically motivated parameters
- Simulate BOSS covariance matrix by dividing SDSS-II covariance matrix by 80.
- Add Gaussian noise to simulated flux power spectrum
- Add SiIII, resolution...

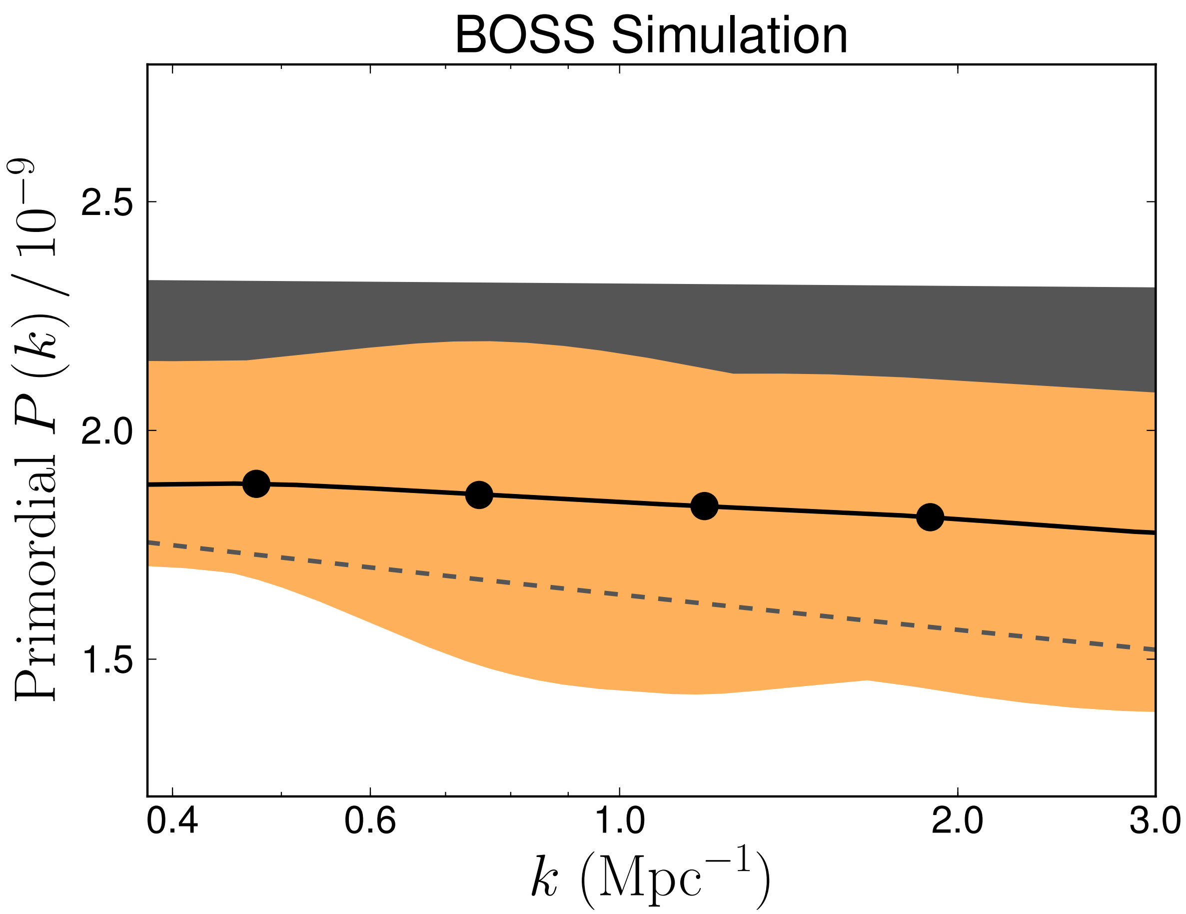
Results: BOSS Simulation



Comparable error bars to the CMB!

Bird et al (2010), arxiv:1010.1519

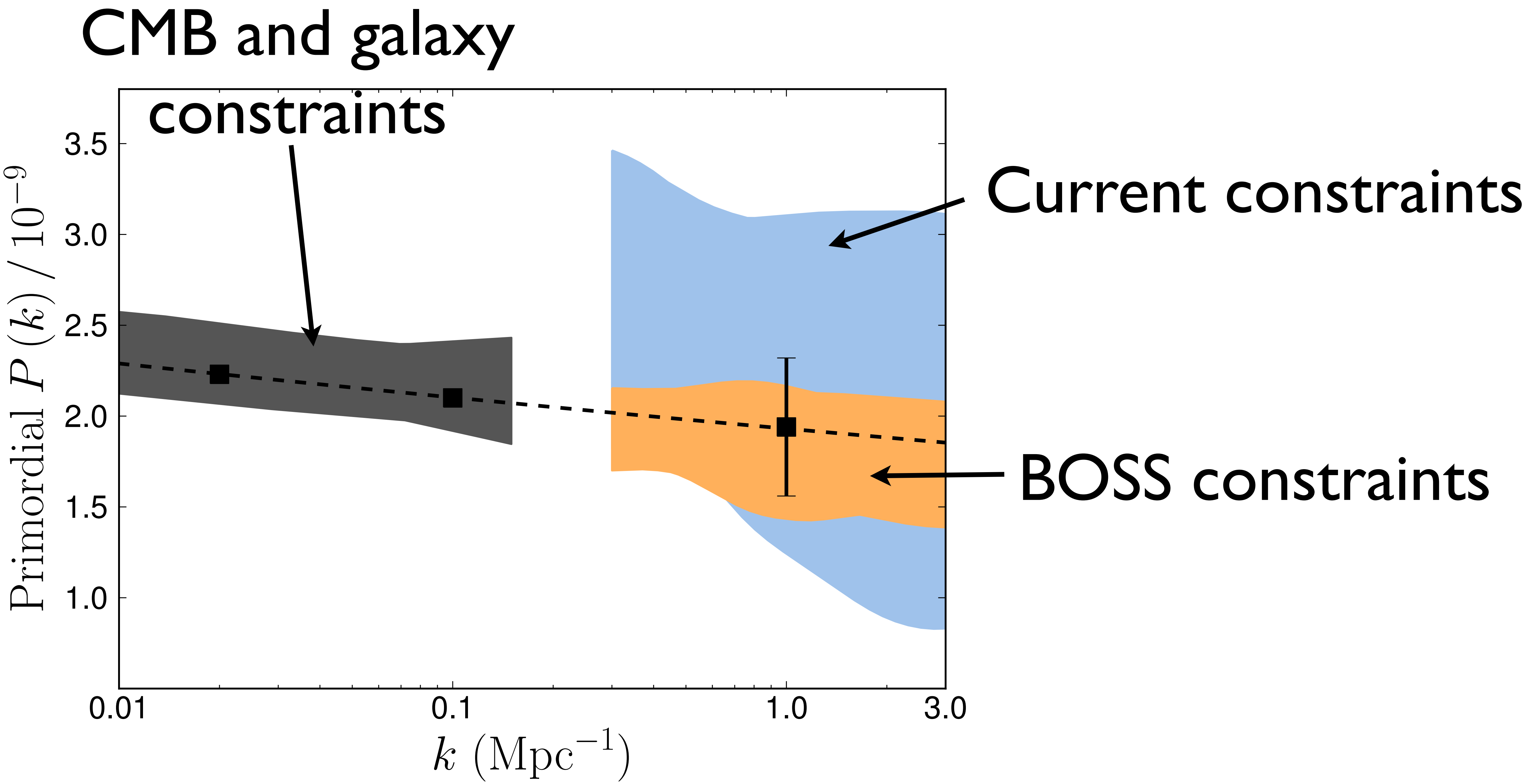
Results: BOSS



- Reproduce earlier results with SDSS covariance matrix
- CV score again constant with penalty
- Fixing thermal params finds preferred prior

No preferred prior for current data due to systematic and statistical error.

Conclusions



Ultimate goal: combine Lyman-α with large-scale datasets

Bird et al (2010), arxiv:1010.1519