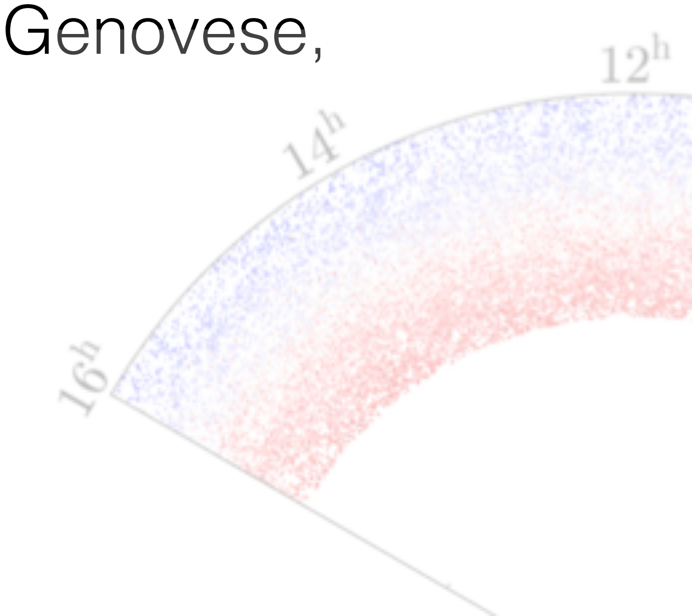
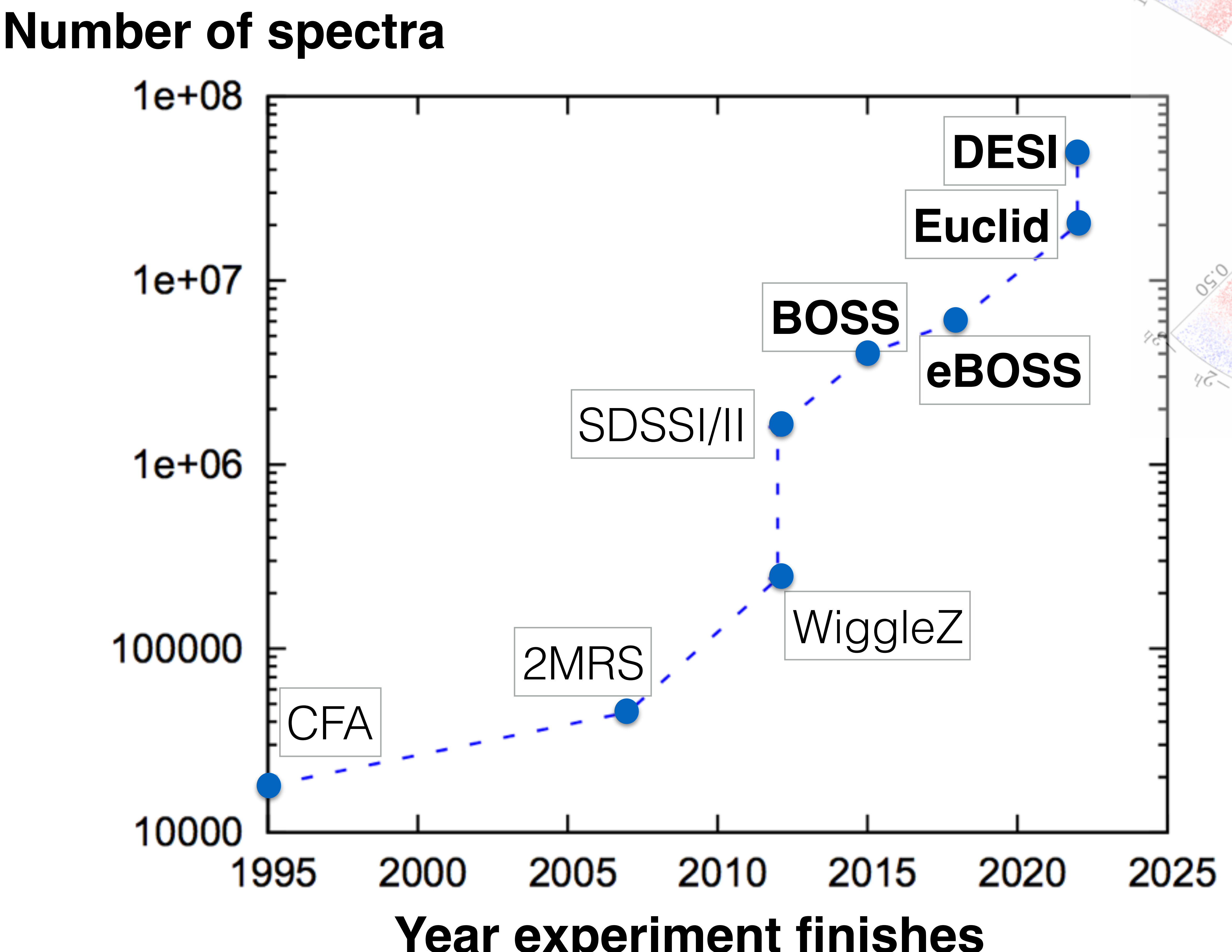
Digging in the Large Scale Structure of the Universe

Shirley Ho
Carnegie Mellon University

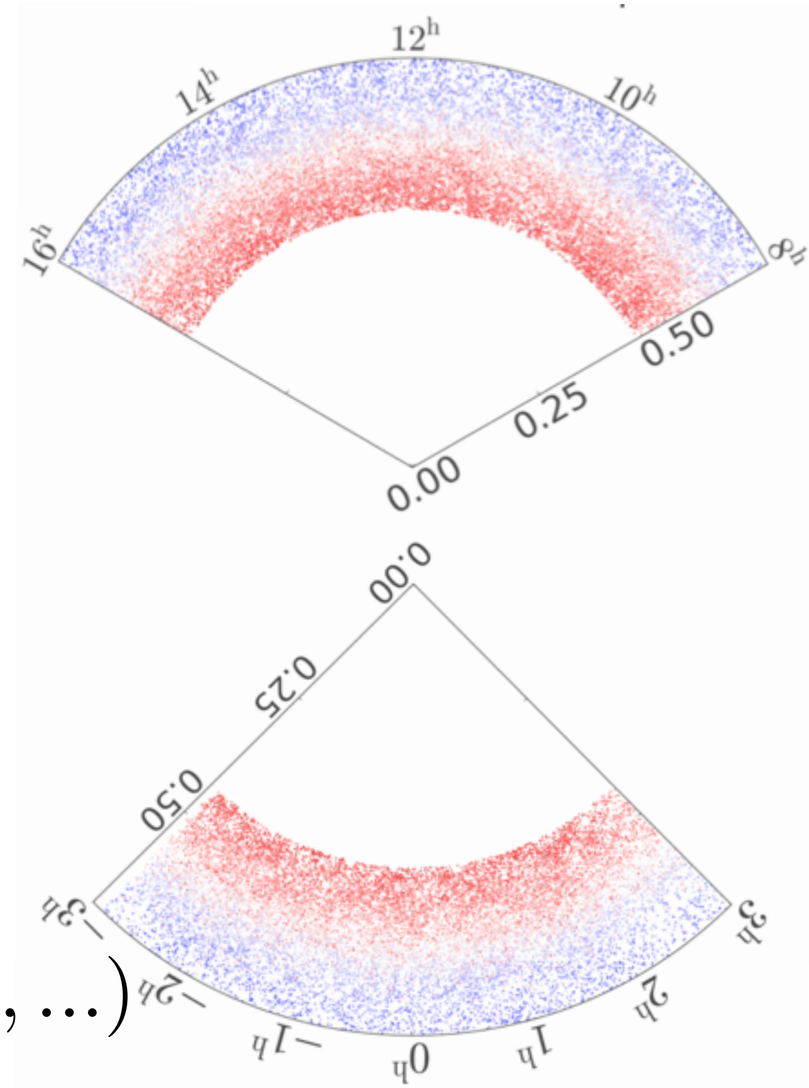
+Shadab Alam, Anthony Pullen, Siyu He, Mariana Vargas, Yen-Chi Chen, Simeon Bird, Roman Garnett, Siamak Ravanbakhsh, Layne Price, Francois Lanusse, Jeff Schneider, Chris Genovese, Rachel Mandelbaum, Barnabas Pocozs, Larry Wasserman
+ SDSS3-BOSS Collaboration



Madrid Cosmology Meeting, 2016


Number of spectra
1e+08
1e+07
1e+06
100000
10000
1995
2000
2005
2010
2015
2020
2025
Year experiment finishes
DESI
Euclid
BOSS
eBOSS
SDSSI/II
WiggleZ
2MRS
CFA

Large Scale Structure



- The contents and properties of our Universe affects the phase space distribution of the density field.

$$(x, y, z, v_x, v_y, v_z) = f(w_0, w_a, \Omega_m(z), H(z), \sum m_\nu, G, ...)$$

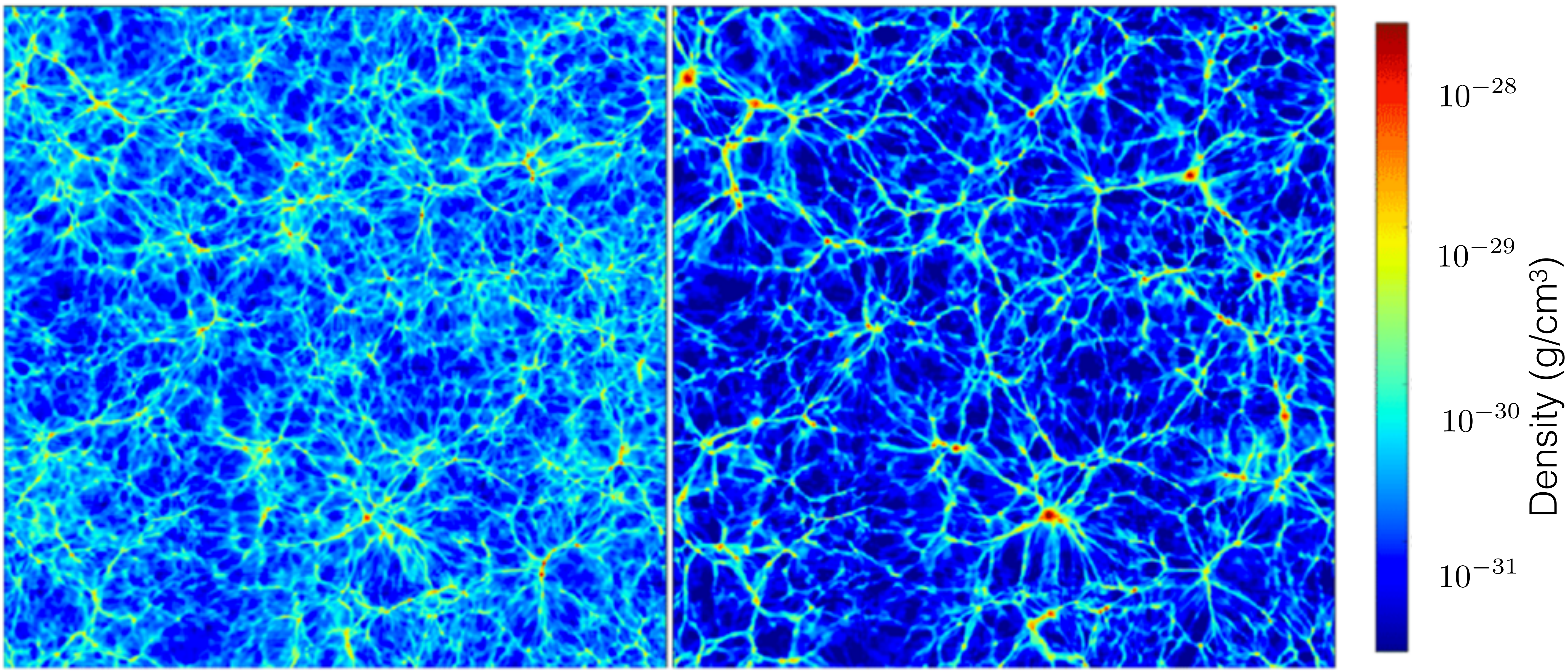
$$w = \frac{p}{\rho}$$ equation of state of dark energy

$$w(a) = w_0 + w_a(1 - a)$$ time-dependent equation of state

$\sum m_\nu$ is sum of neutrino masses

Eisenstein & Hu 1997; Eisenstein, Seo & White 2007; Kaiser 1987; Peacock 2001

How sum of neutrino masses affect the density field



$\sum m_\nu = 0$ eV

$\sum m_\nu = 1$ eV

Figure Credit: Agarwal & Feldman

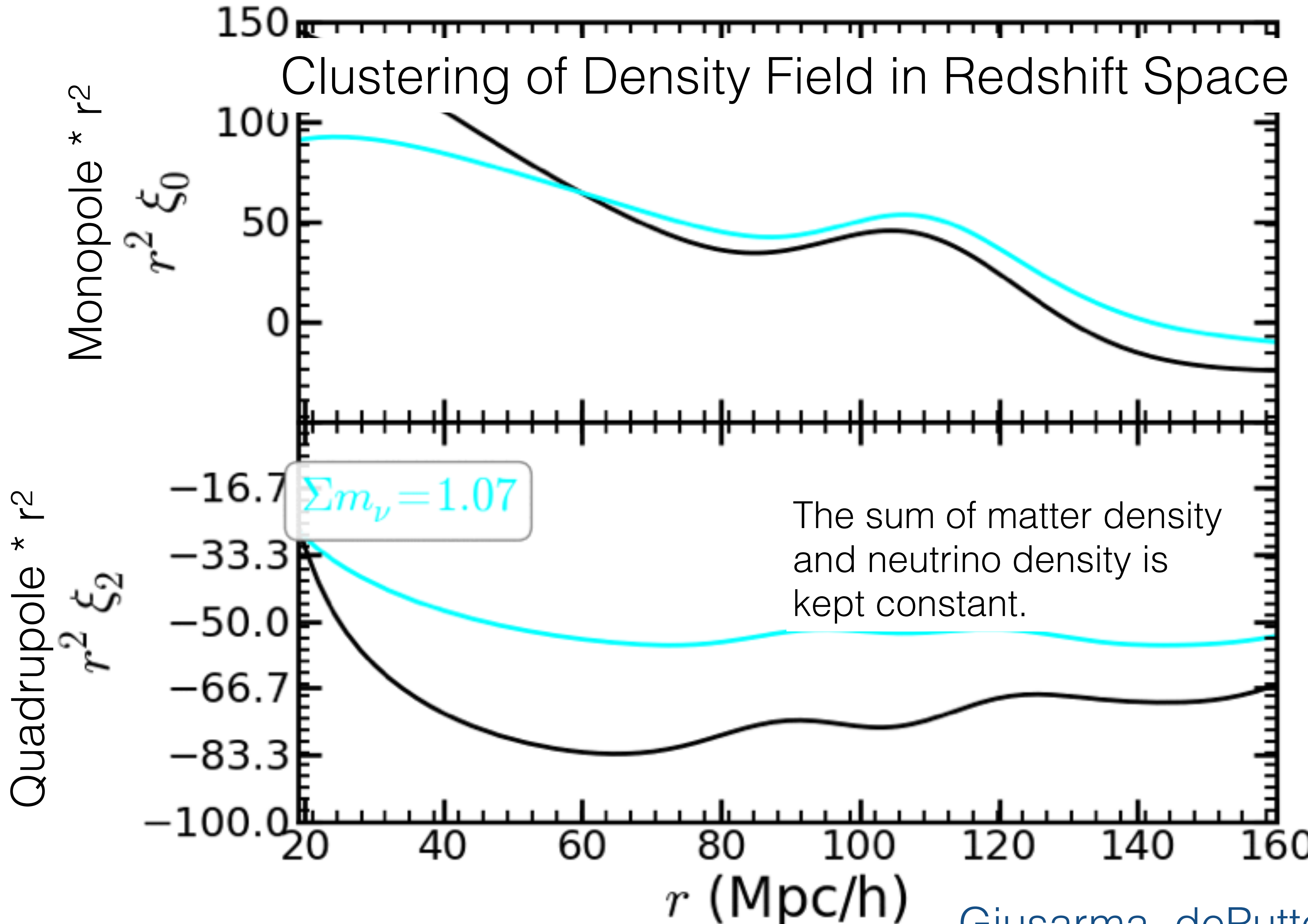
What do we do with all these interesting datasets?

- Standard analyses: 2 point correlation functions / clustering / stacking

- Going beyond standard analyses

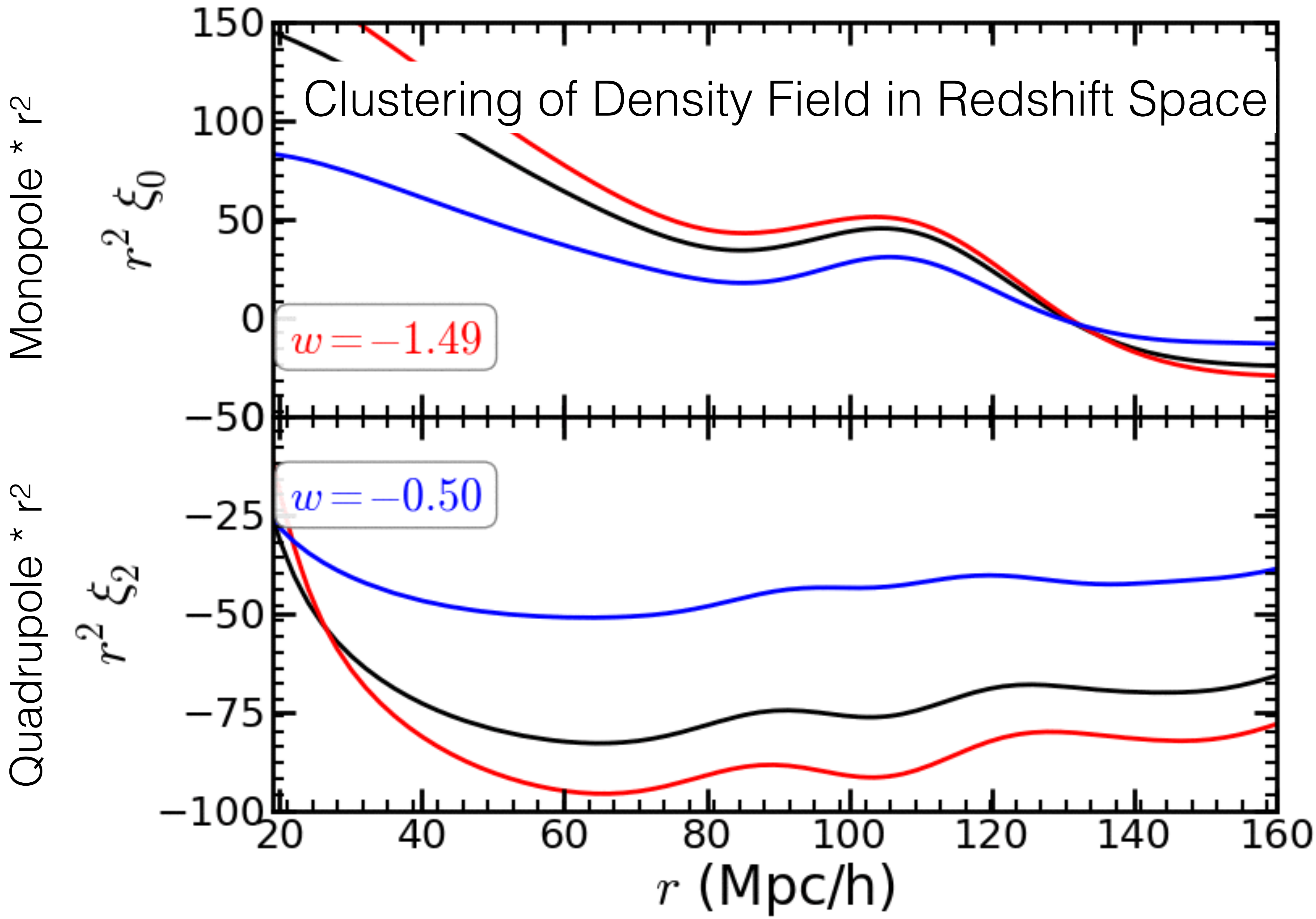
- Going beyond 3D information that these large scale structure surveys provide.

Standard Analyses:
Sum of neutrino masses affect the clustering of density field

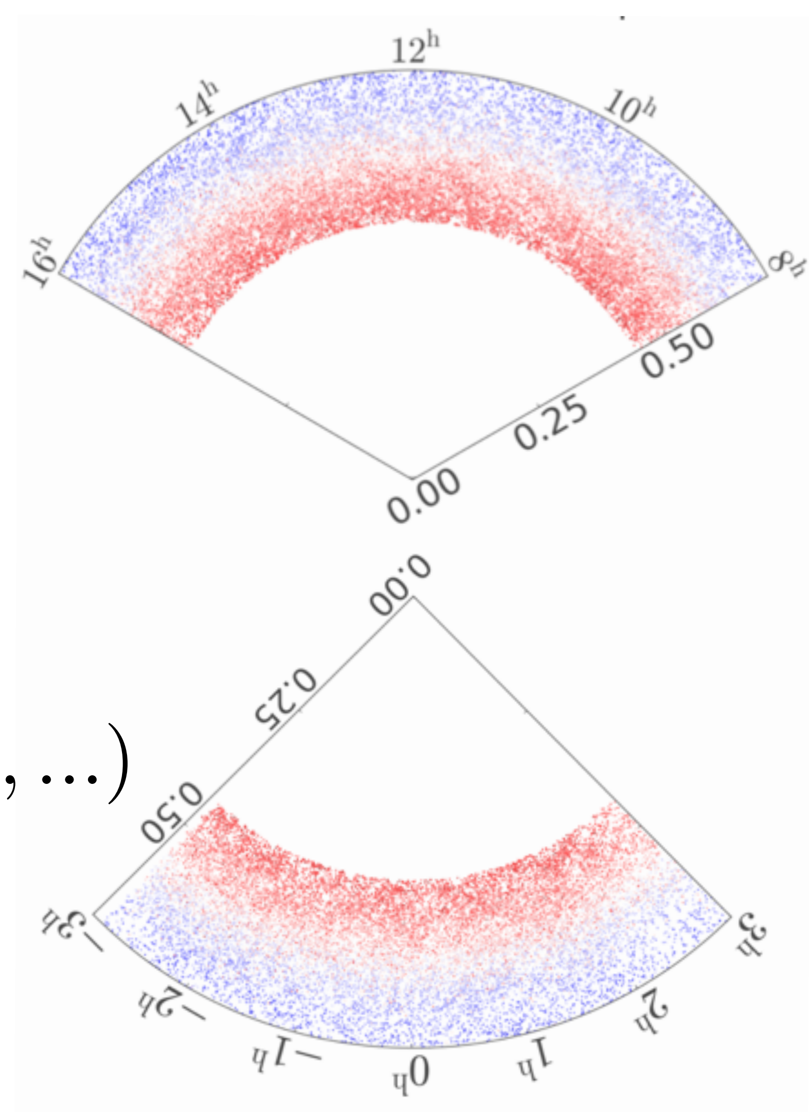


Giusarma, dePutter, **Ho** et al. 2013

Standard Analyses: Dark Energy equation of state affects the clustering of density field



Large Scale Structure



- The contents and properties of our Universe affects the phase space distribution of the density field.

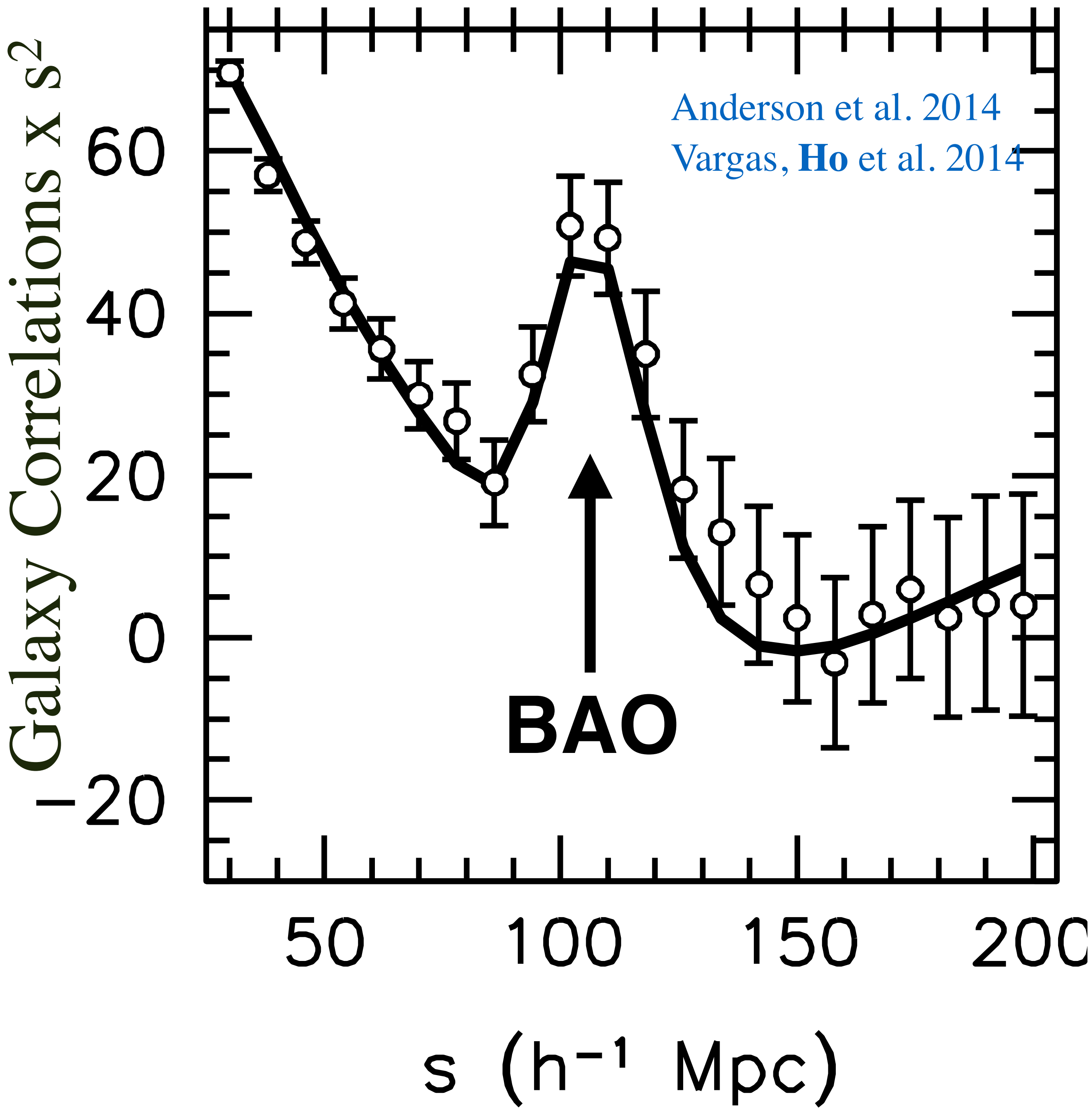
$$(x, y, z, v_x, v_y, v_z) = f(w_0, w_a, \Omega_m(z), H(z), \sum m_\nu, G, ...)$$

- The probe that focuses on **(x,y,z)** is **Baryon Acoustic Oscillations (BAO).**
- The probe that focuses on **(v_x, v_y, v_z)** is
- **Redshift Space Distortions (RSD).**

Eisenstein & Hu 1997; Eisenstein, Seo & White 2007; Kaiser 1987; Peacock 2001

Standard analyses

Correlation Function x s^2 in 1D



- Lots of good science are being done with standard 2 point correlation function/ power-spectrum analyses!
- **Baryon Acoustic Oscillations**

Baryon Acoustic Oscillations in Radiation and Matter

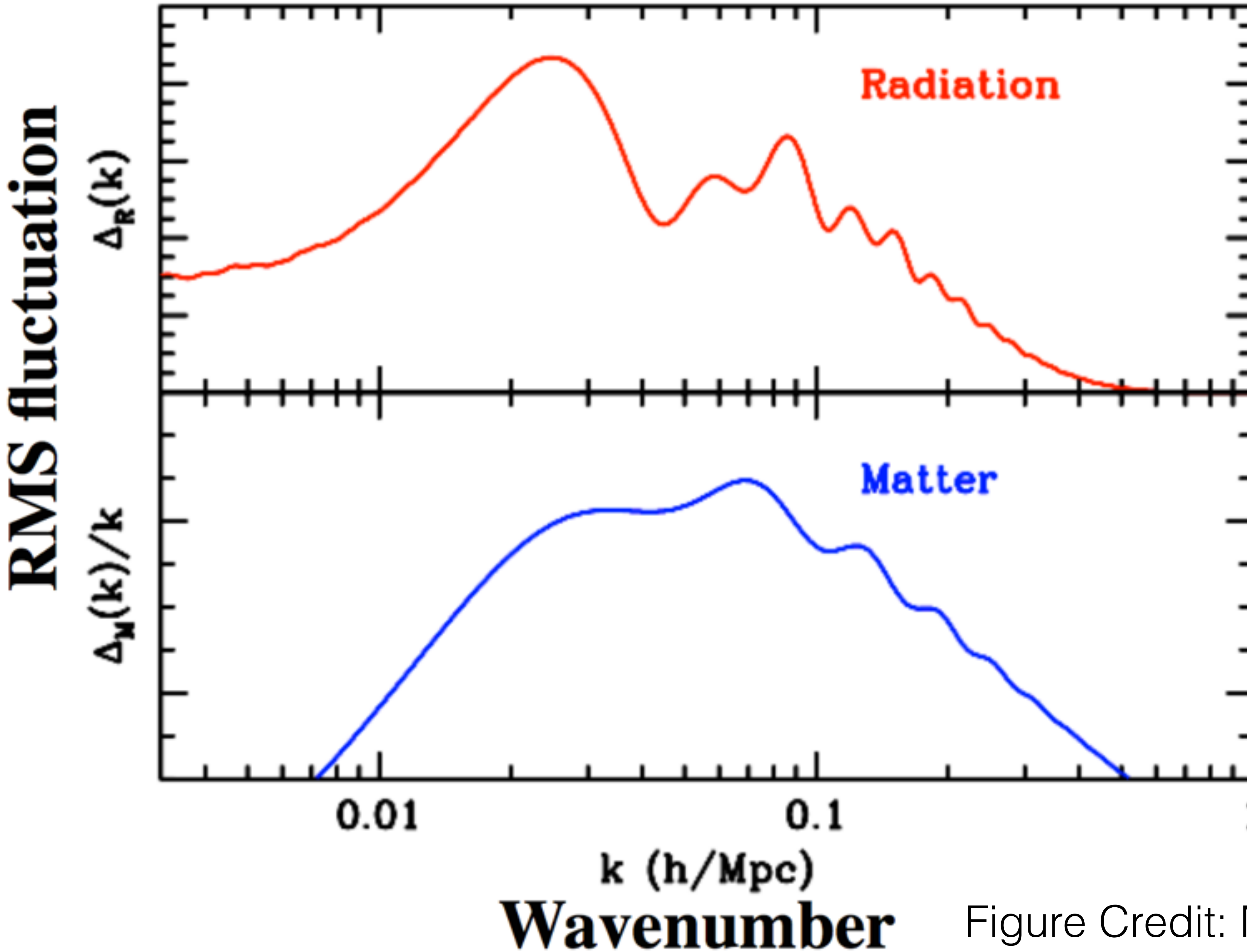
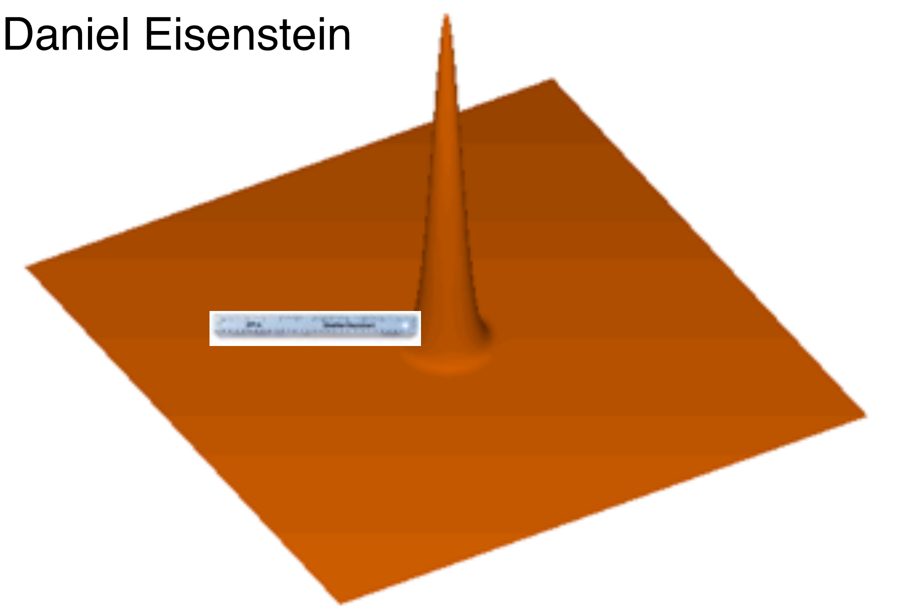


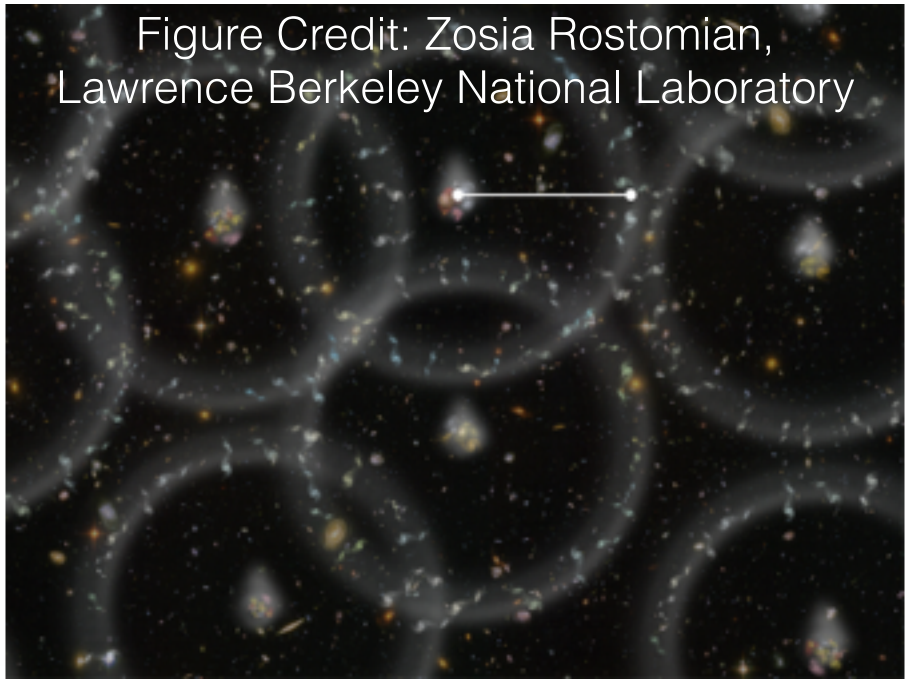
Figure Credit: Martin White

Multiple Clean Probes
Baryon Acoustic Oscillations in Large Scale Structure

- Each initial overdensity has excess pressure, leading to an outward-going sound wave.
- At recombination, these sound waves halt, depositing their gas in a spherical shell 500 million light-years (110 Mpc/h) from the original location.
- An overdensity at one location implies a small increase in the density 110 Mpc/h away.
- Small Statistical signal. We need sky surveys of large volume and high number density in order to detect it.

Animation credit:
Daniel Eisenstein



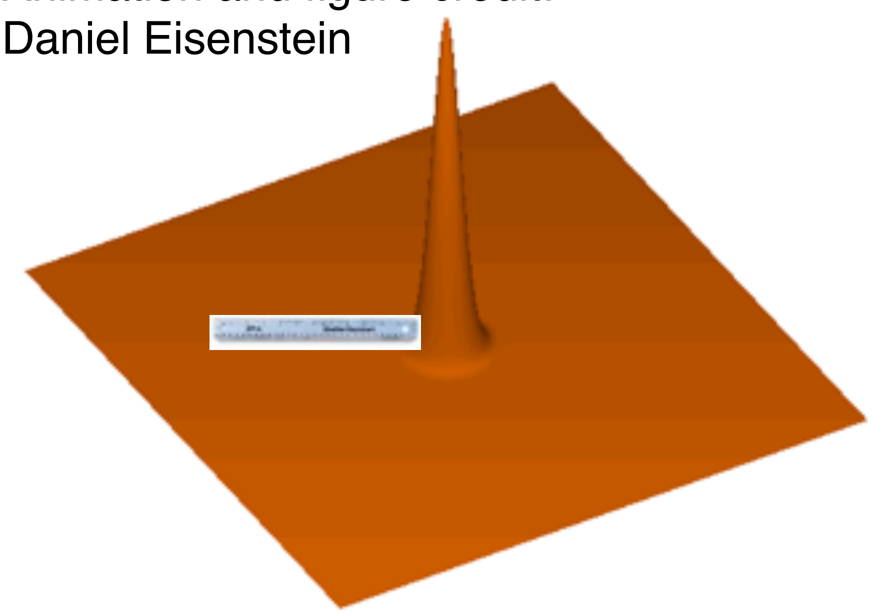


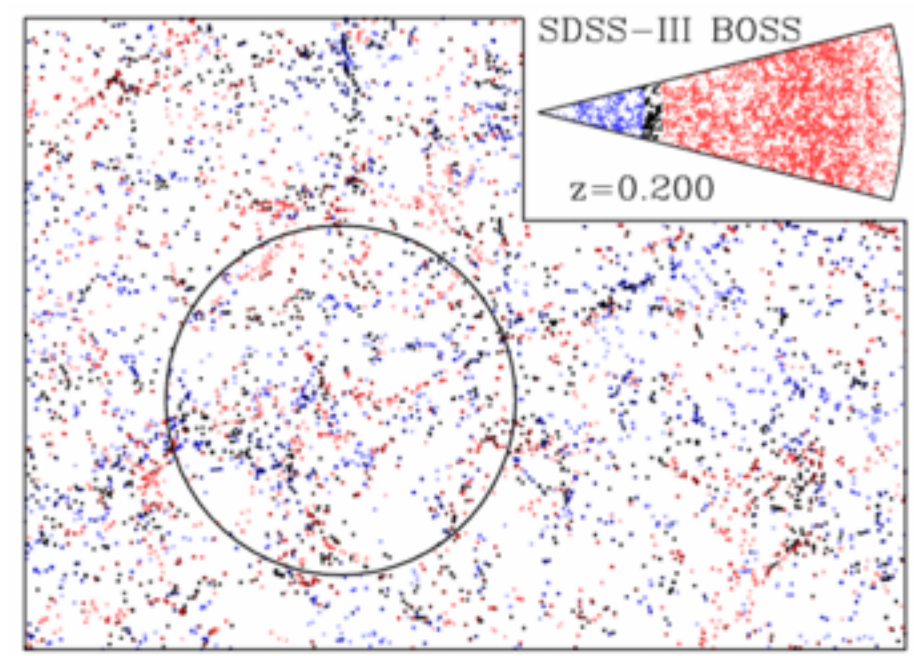
Eisenstein & Hu 1997; Eisenstein, Seo & White 2007

Multiple Clean Probes
Baryon Acoustic Oscillations in Large Scale Structure

- Each initial overdensity has excess pressure, leading to an outward-going sound wave.
- At recombination, these sound waves halt, depositing their gas in a spherical shell 500 million light-years (110 Mpc/h) from the original location.
- An overdensity at one location implies a small increase in the density 110 Mpc/h away.
- Small Statistical signal. We need sky surveys of large volume and high number density in order to detect it.

Animation and figure credit:
Daniel Eisenstein





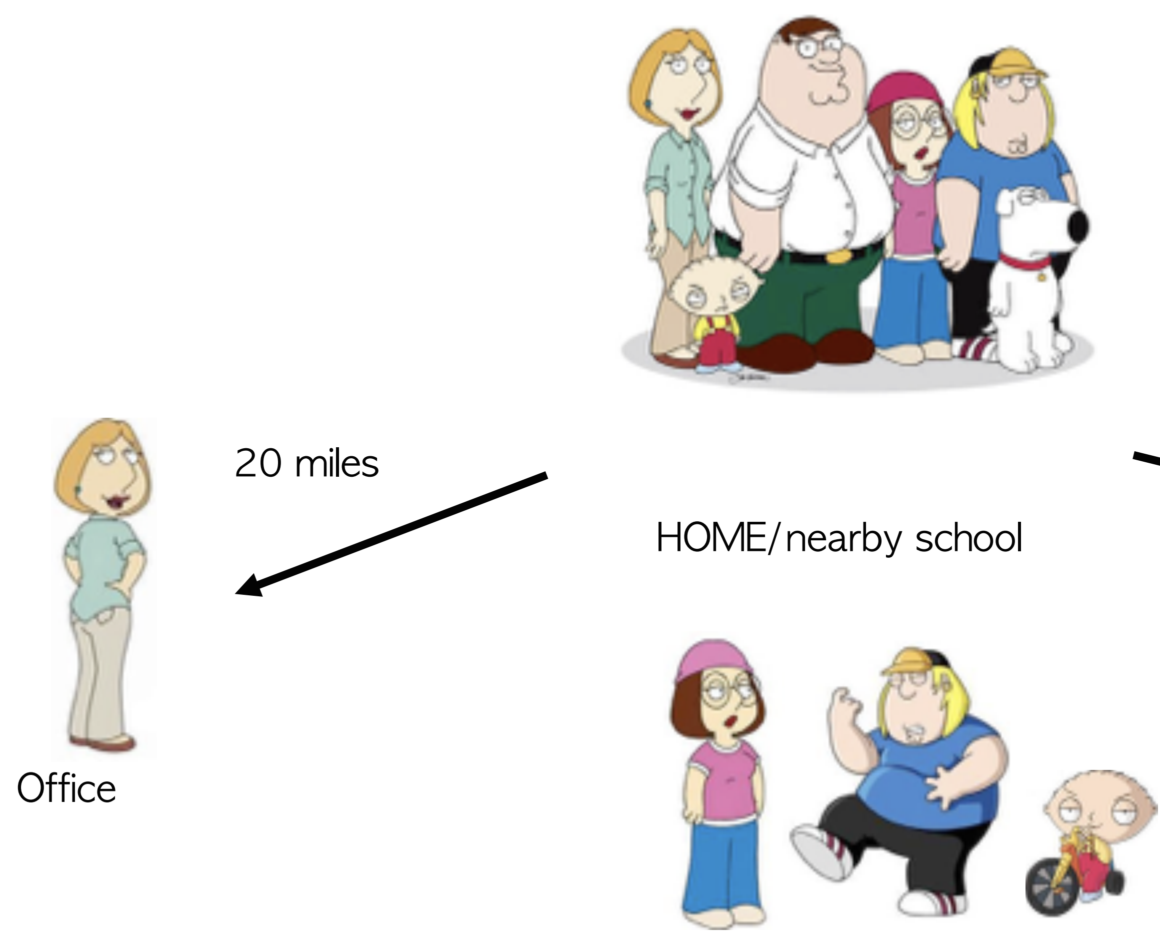
Eisenstein & Hu 1997; Eisenstein, Seo & White 2007

What are Baryon Acoustic Oscillations?

To measure BAO, we usually calculate the correlation function

What are Baryon Acoustic Oscillations?

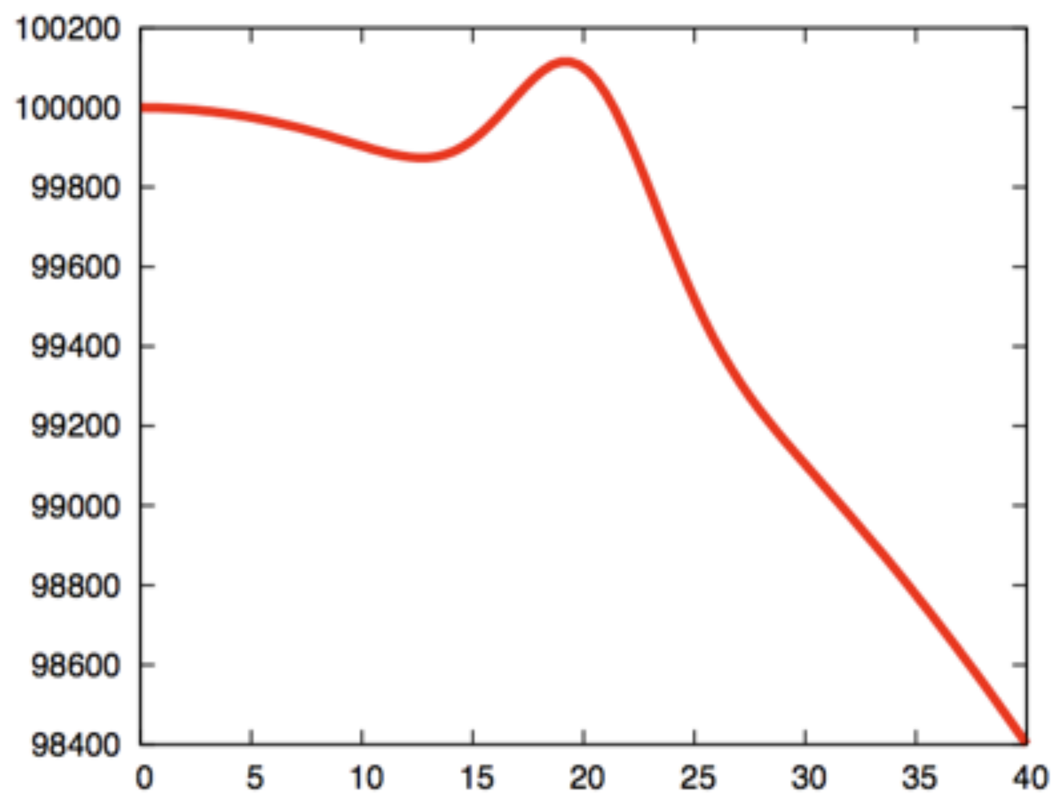
What is the correlation function of population during the day?



What are Baryon Acoustic Oscillations?

What is the correlation function of population during the day?

Office

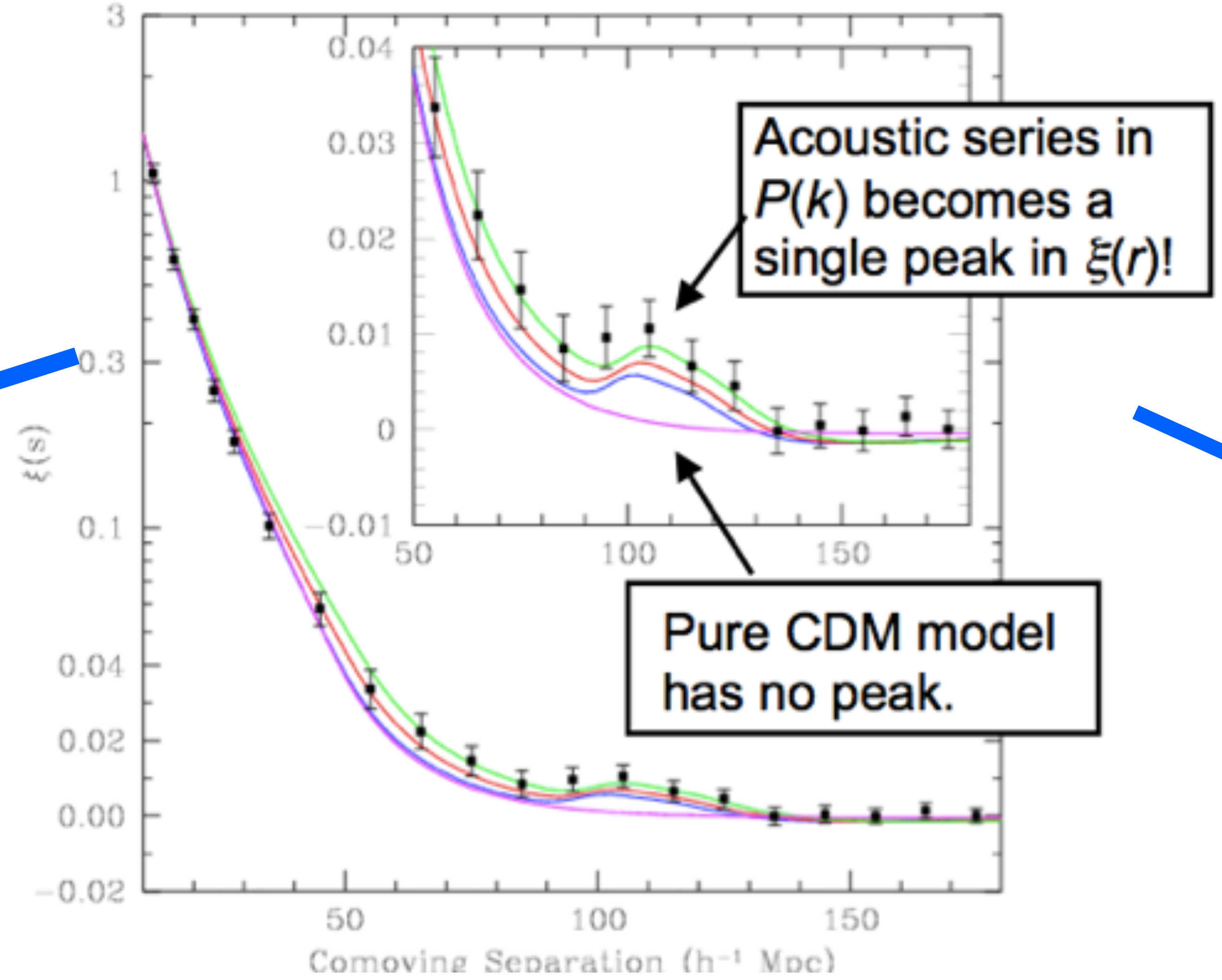


Office

A bump in 20 miles!

What are Baryon Acoustic Oscillations?

To measure BAO, we first calculate the correlation function



A bump at ~150Mpc

Baryon Acoustic Oscillations

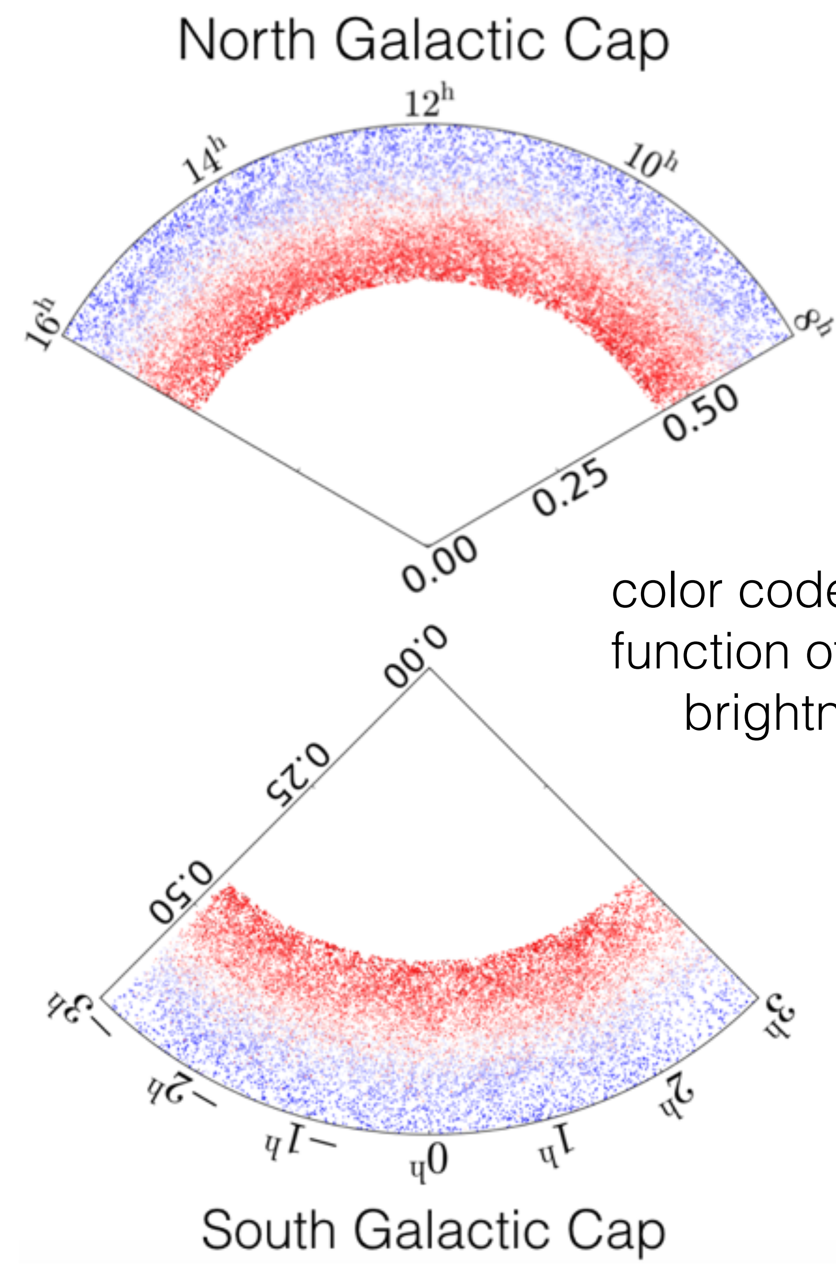
Steps towards a cosmological distance measurement

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples

- Potential issues include:
  - Variations of target sample caused by changes of targeting algorithm or properties of the targeting data
  - The spatial pattern of missing spectroscopic data points
  - Changes of galaxy densities due to observational systematics
- The sample creation methodology and pipeline is now adopted in SDSS IV/eBOSS to make other large scale structure catalogs



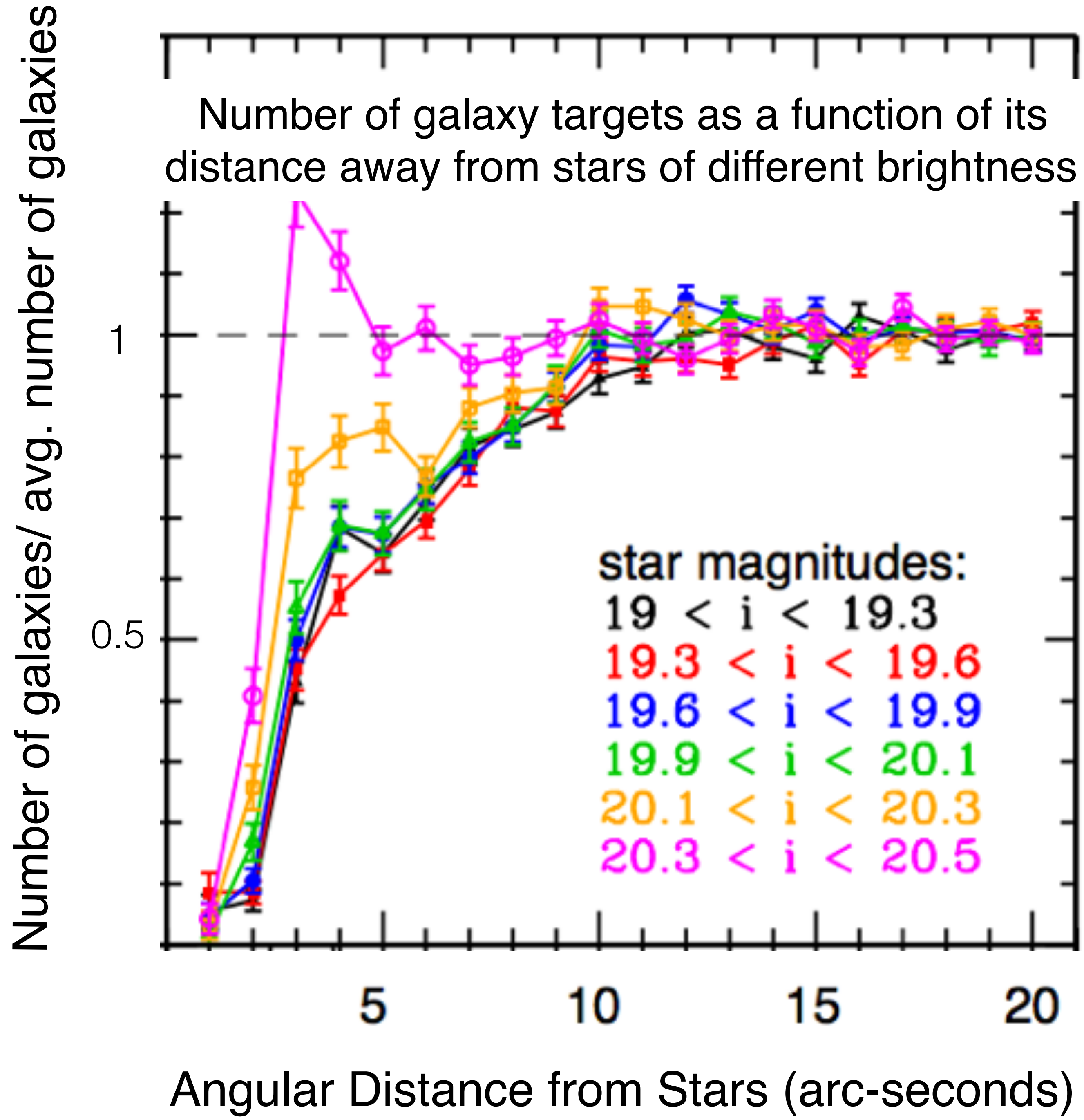
color coded as a function of tracer brightness

BOSS CMASS sample
White et al. 2010; Reid, **Ho** et al. 2015

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples
- **Correct for observational systematics**

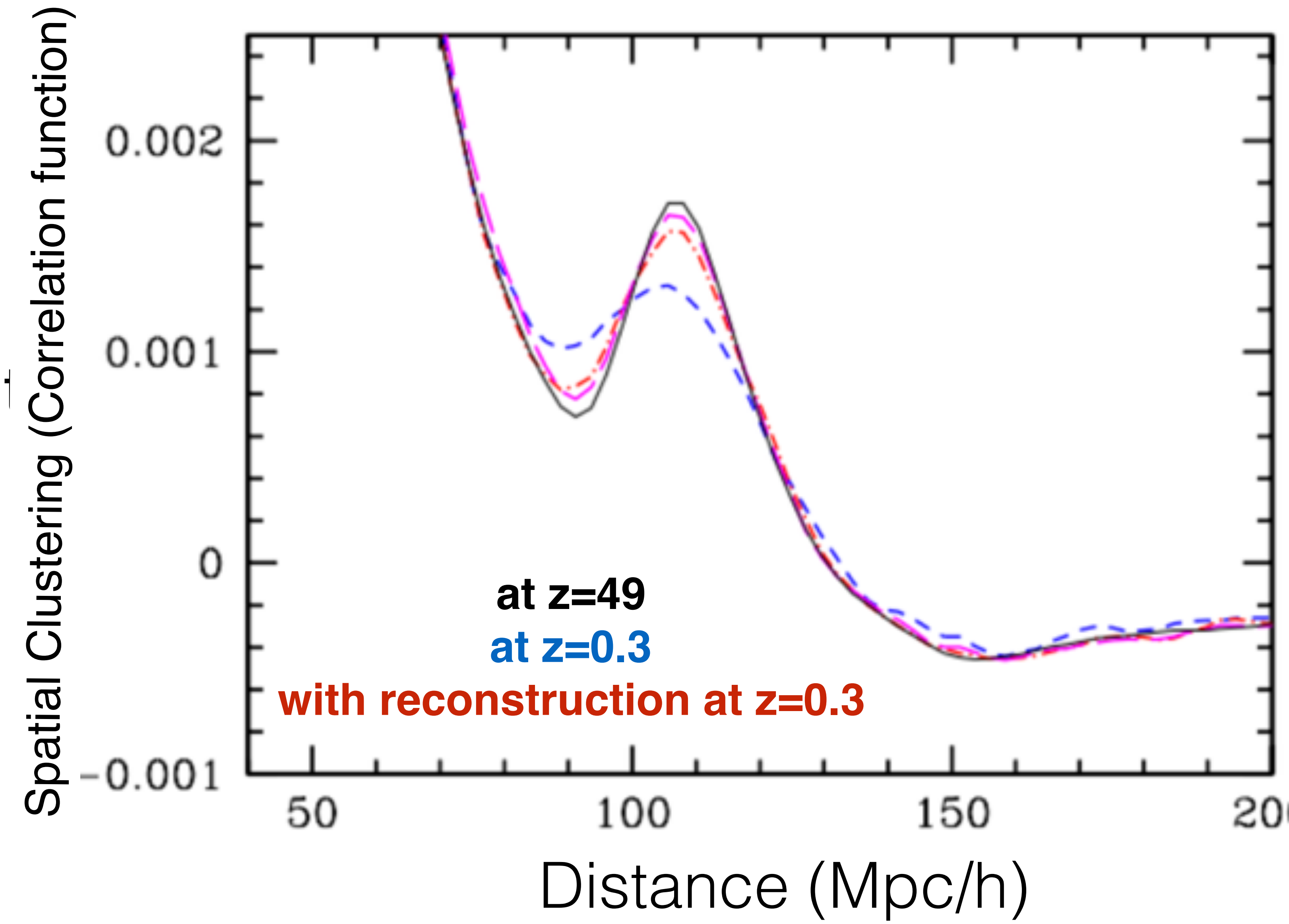


Ross, **Ho** et al. 2011
Ross, Percival, Samushia, **Ho** et al. 2012

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples
- Correct for observational systematics
- **Reconstruction: Move the galaxies backwards in time t increase signal-to-noise of measurement.**

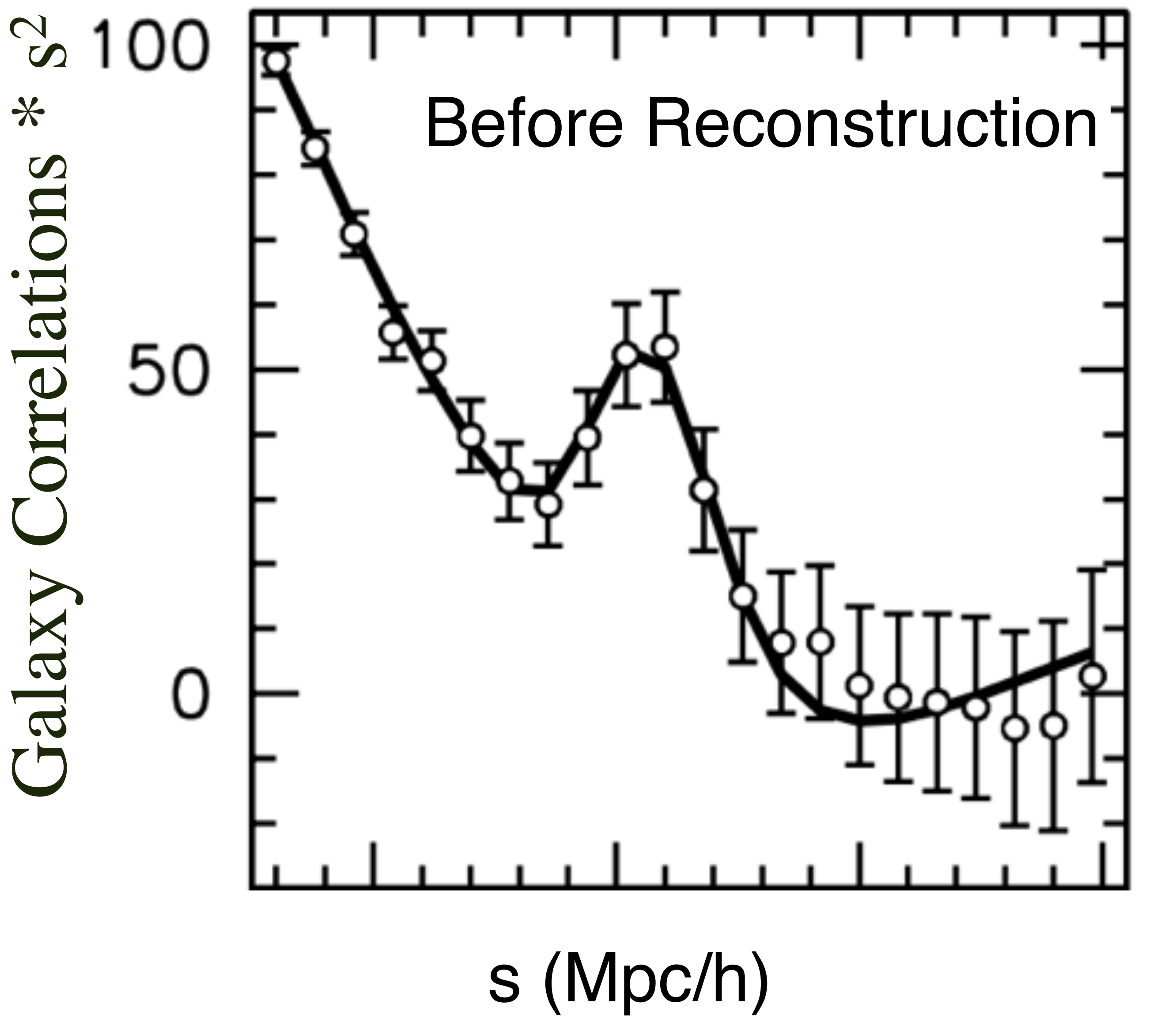


Eisenstein, Seo, Sirko & Spergel 2006;
Padmanabhan, White & Cohn 2008

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples
- Correct for observational systematics
- **Reconstruction: Move the galaxies backwards in time to increase signal-to-noise of measurement.**

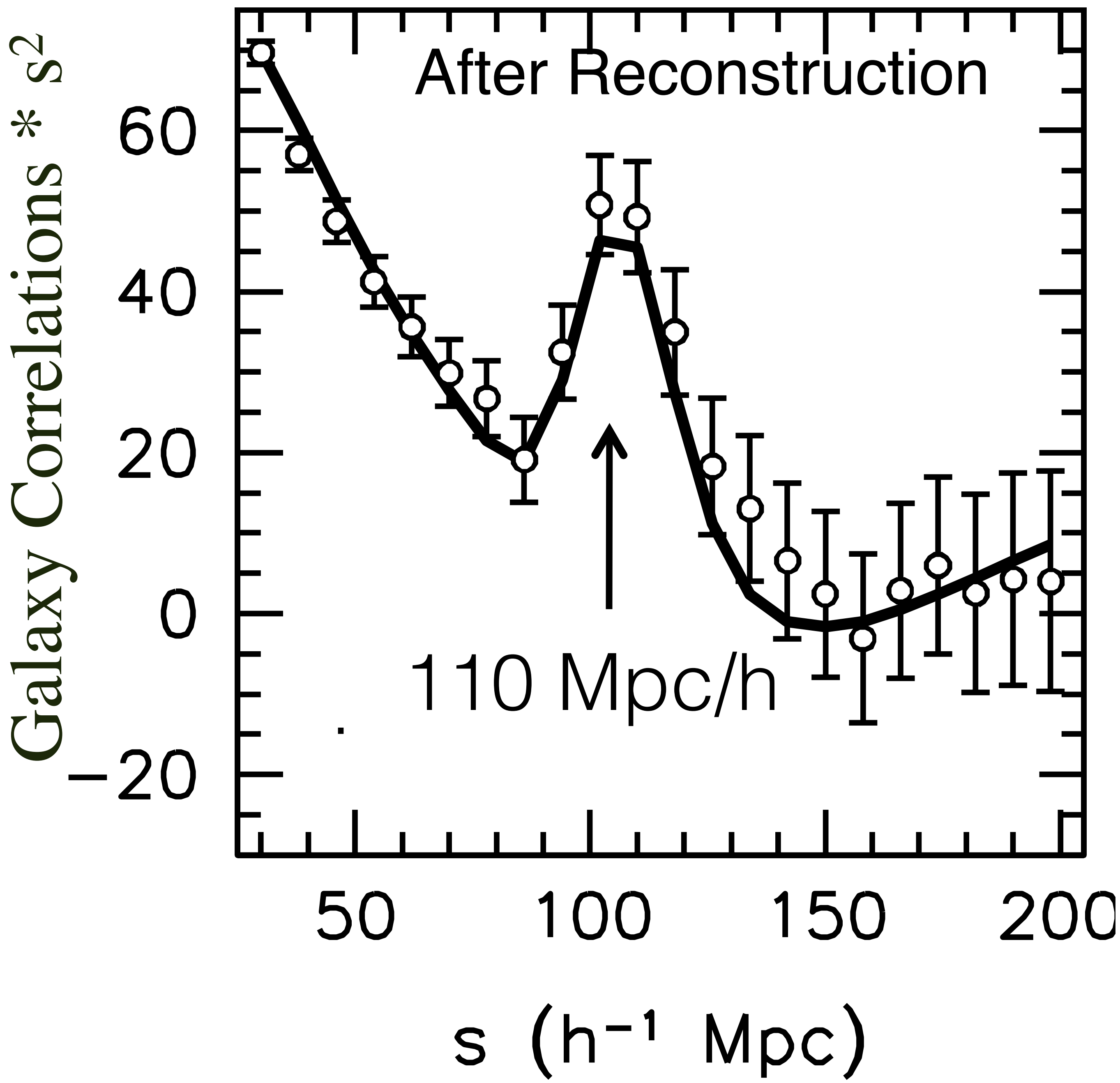


Anderson et al. 2014
Vargas, **Ho** et al. 2015

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples
- Correct for observational systematics
- Reconstruction: Move the galaxies backwards in time to increase signal-to-noise of measurement.

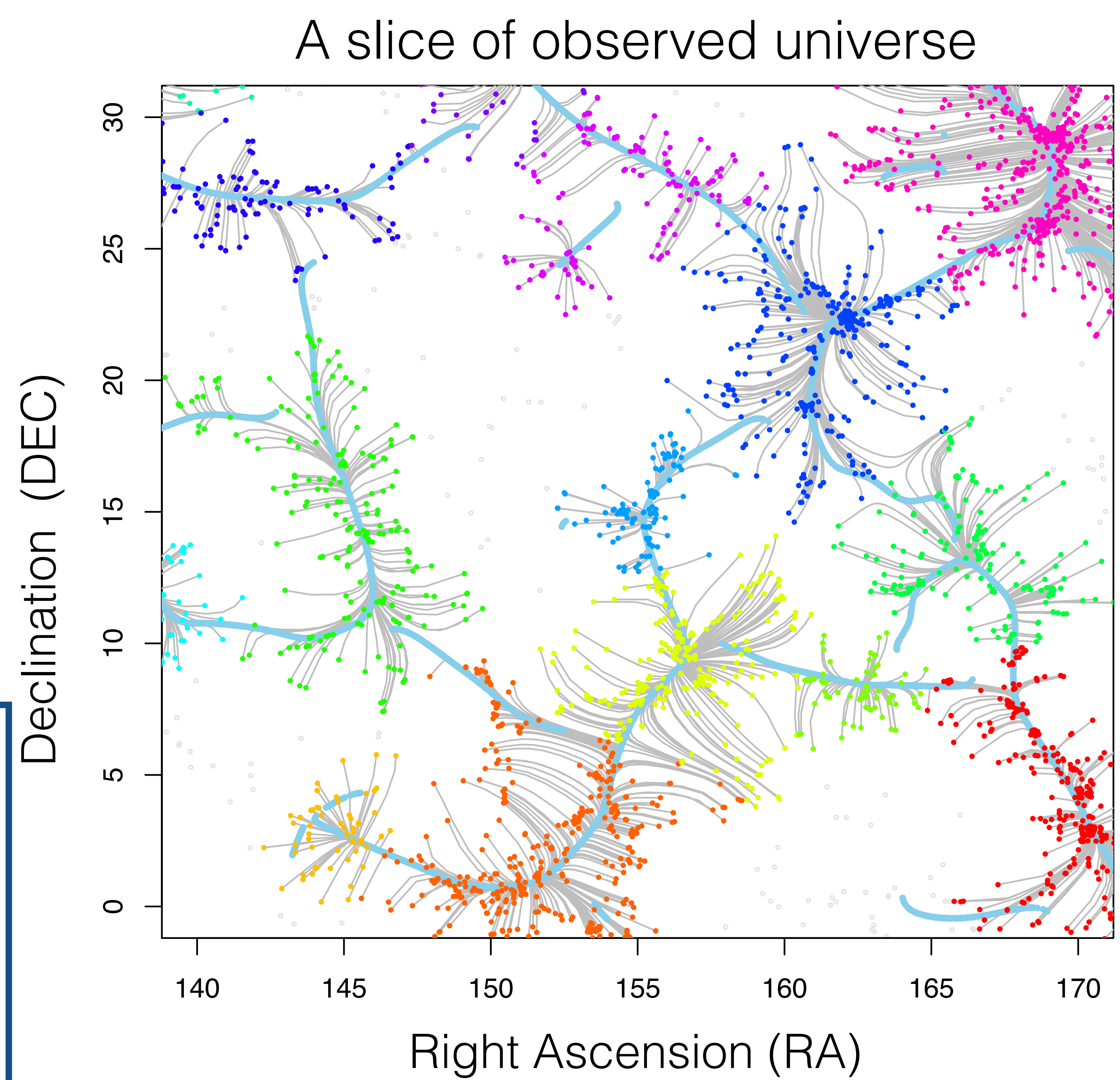


Anderson et al. 2014
Vargas, **Ho** et al. 2015

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples
- Correct for observational systematics
- Reconstruction: Move the galaxies backwards in time to increase signal-to-noise of measurement.



- New ways to improve reconstruction
  - Calculate the Hessian Matrix at each tracer's position
  - Steepest ascent up the potential
  - Good at high density region
  - Combine with traditional reconstruction method to find best velocities to move galaxies back in time.

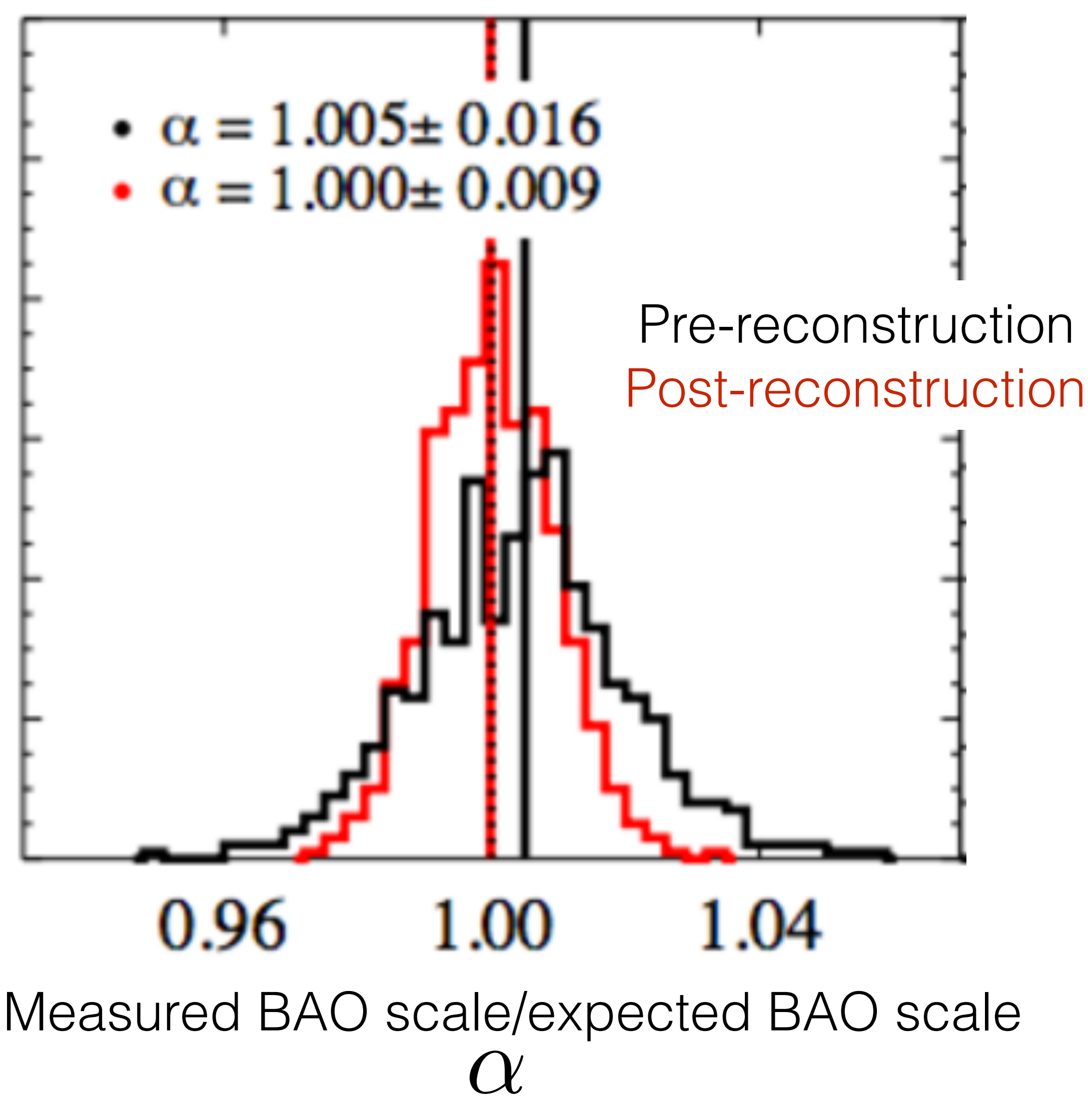
Preliminary results from : **Ho, Chen, Vargas** et al.
Vargas, Ho et al. 2015

Baryon Acoustic Oscillations

Steps towards a cosmological distance measurement

- Create Clean Samples
- Correct for observational systematics
- Reconstruction: Move the galaxies backwards in time to increase signal-to-noise of measurement.
- **Thorough test on fitting methods to achieve accurate BAO positions.**

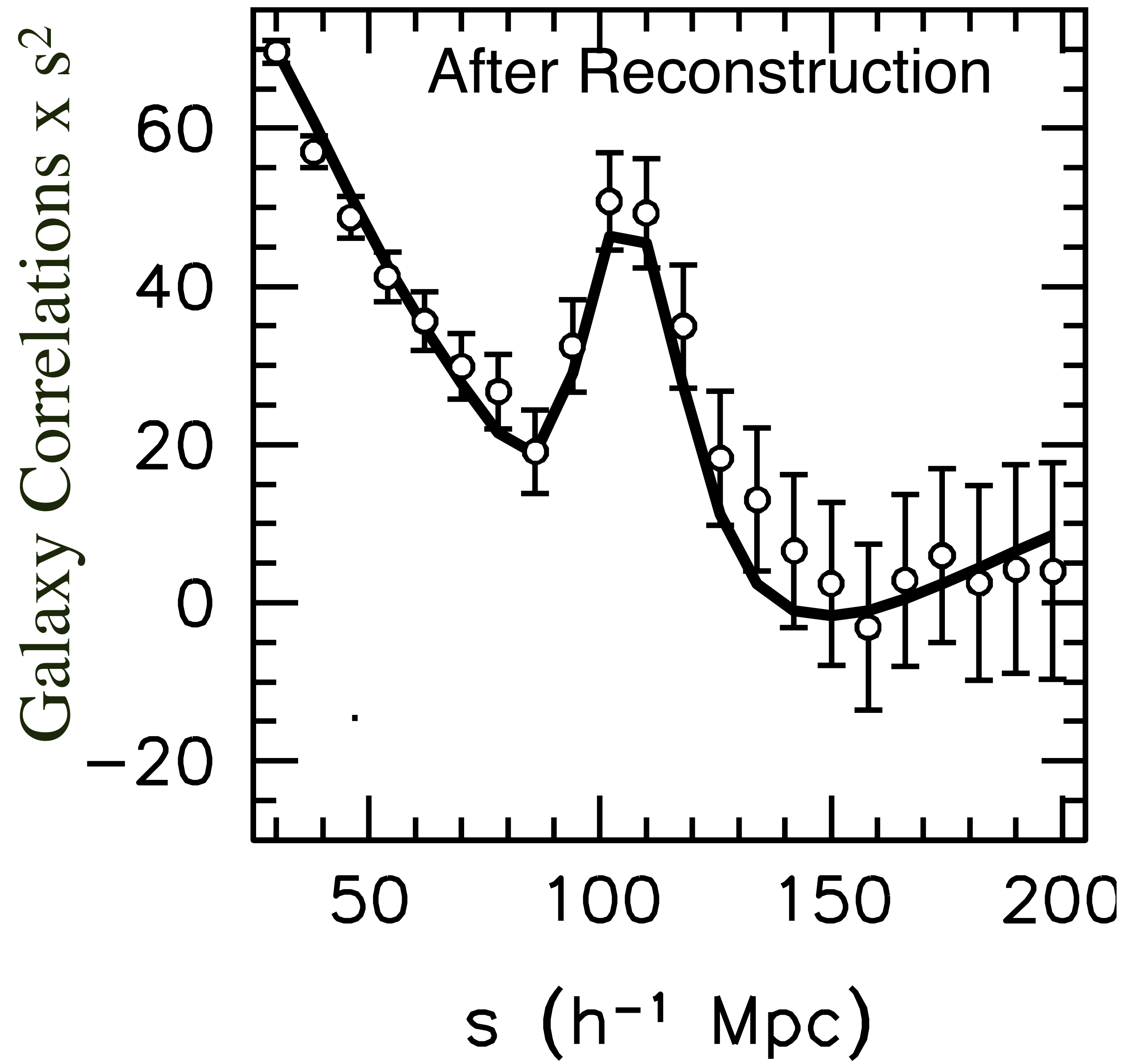
The distribution of normalized BAO scale after fitting 600 mock surveys



Anderson et al. 2014
Vargas, **Ho** et al. 2014

Baryon Acoustic Oscillations
Measurement of Distances at multiple redshifts

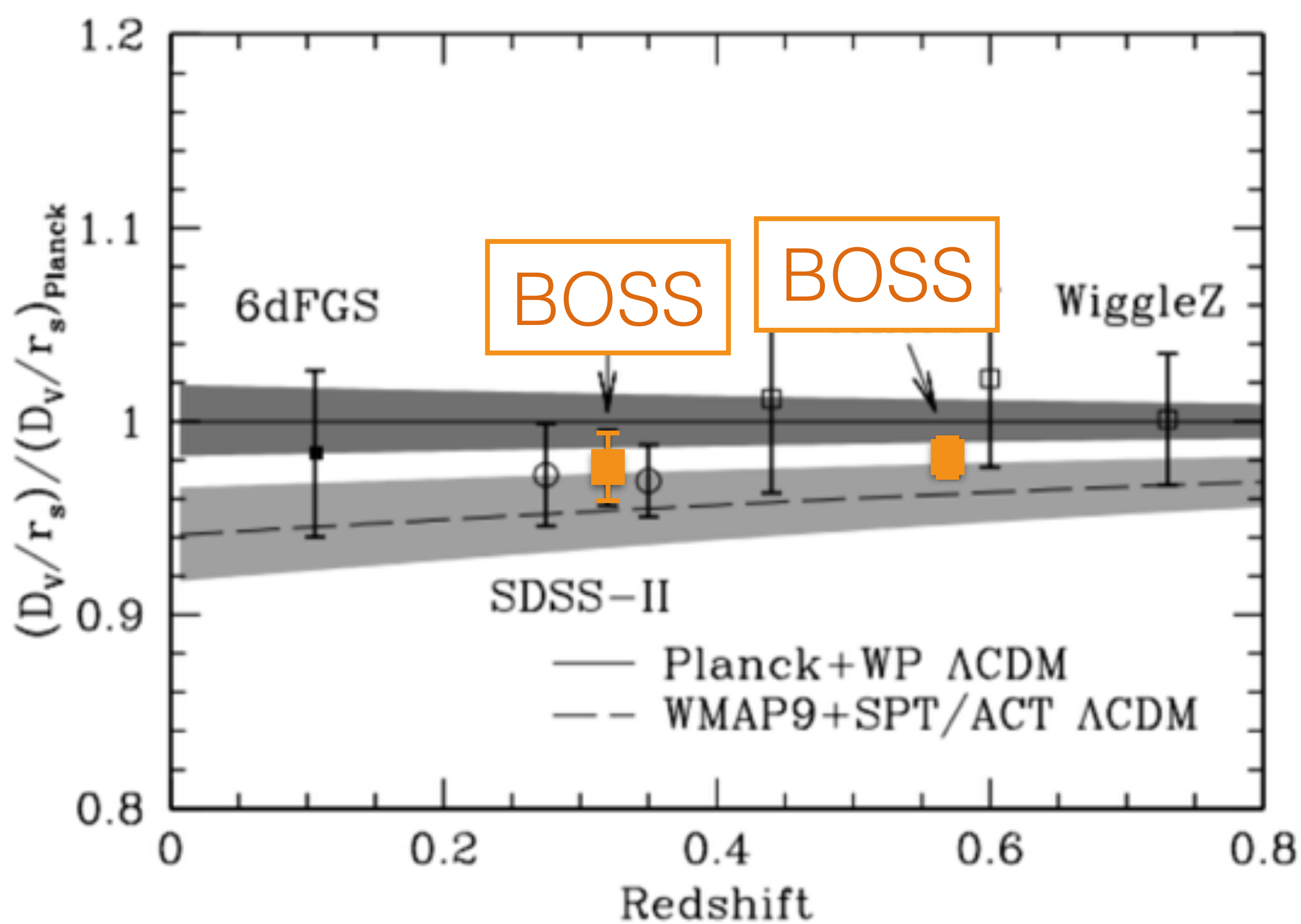
- Clustering Analysis of the BOSS galaxy sample has produced the world's best detection of the late-time acoustic peak.



Anderson et al. 2014
Vargas, **Ho** et al. 2014

Baryon Acoustic Oscillations
Measurement of Distances at multiple redshifts

- Clustering Analysis of the BOSS galaxy sample has produced the world's best detection of the late-time acoustic peak.
- The peak location is measured to
  - 1.0% in z = 0.57 sample and
  - 2.1% in z = 0.32 sample

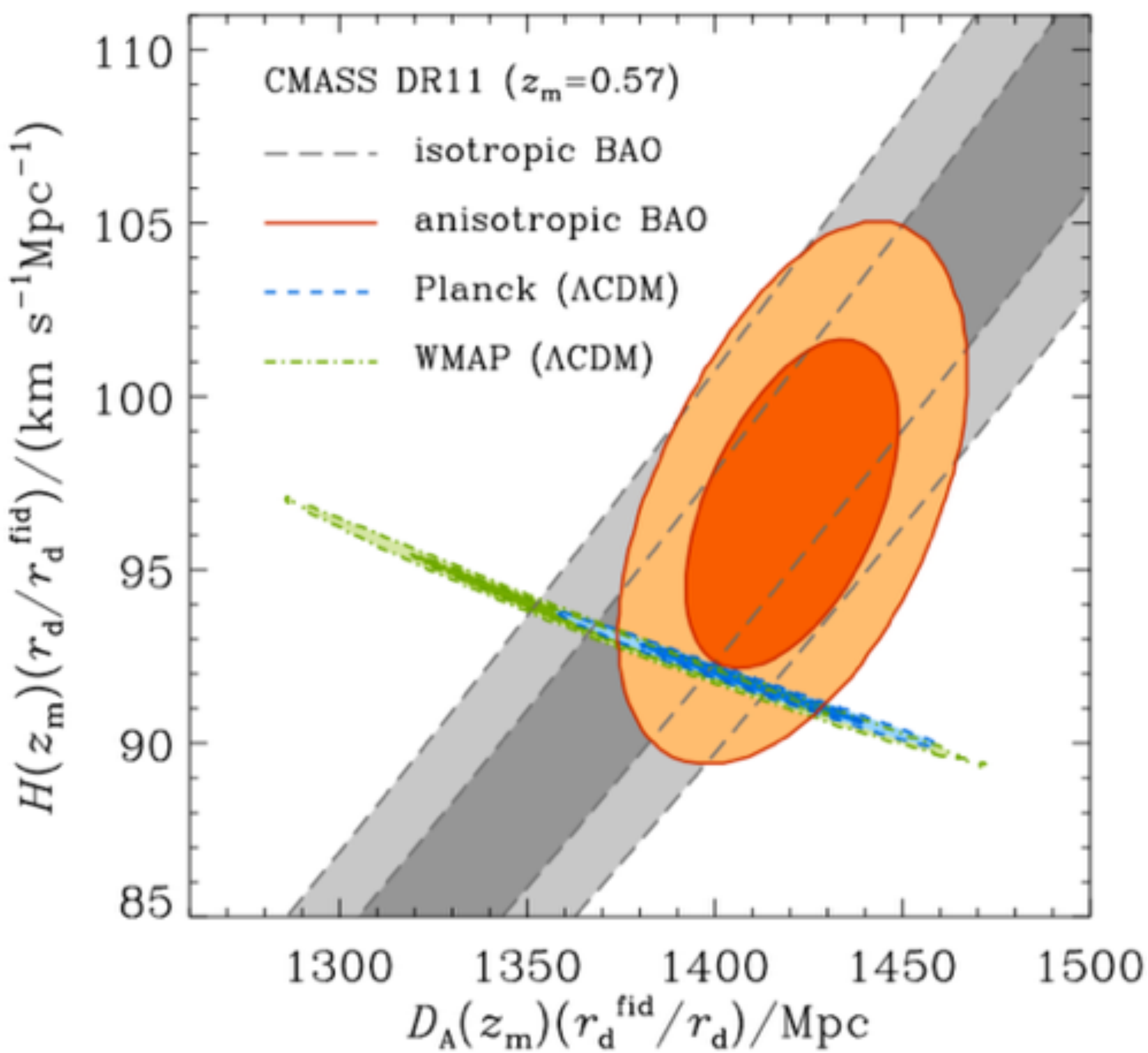


Anderson et al. 2014
Vargas, **Ho** et al. 2014

Baryon Acoustic Oscillations

Taking into account of anisotropy of our observed quantities

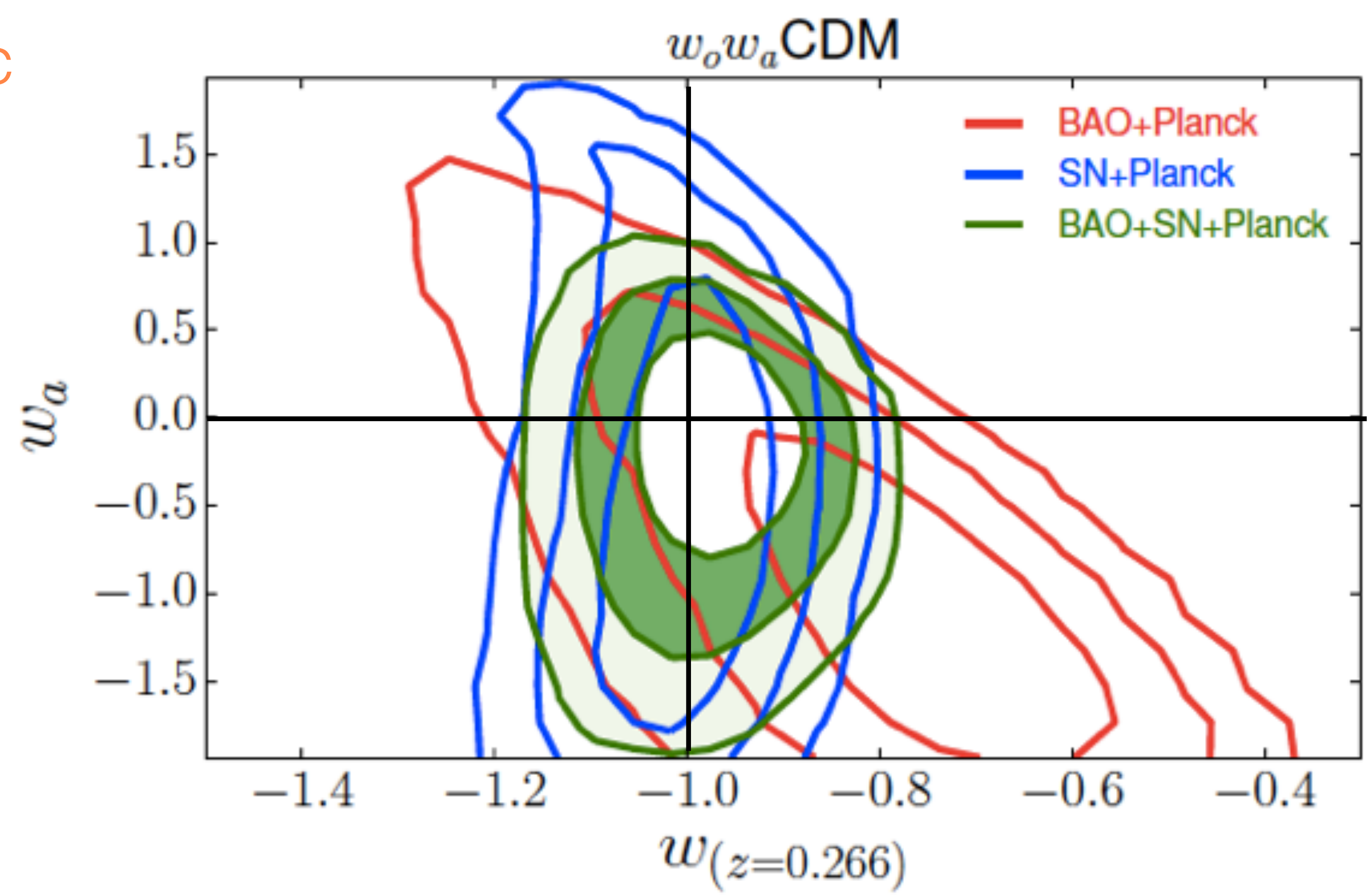
- Clustering Analysis of the BOSS galaxy sample has produced the world's best detection of the late-time acoustic peak.
- The peak location is measured to
  - 1.0% in $z = 0.57$ sample and
  - 2.1% in $z = 0.32$ sample
- Taking into account of the anisotropy of the observation, we improve our constraints on the distance scale in both transverse and radial direction.



Anderson et al. 2014
Vargas, **Ho** et al. 2014

Baryon Acoustic Oscillations
Towards cosmological constraints

- Clustering Analysis of the BOSS galaxy sample has produced the world's best detection of the late-time acoustic peak.
- The peak location is measured to
  - 1.0% in $z = 0.57$ sample and
  - 2.1% in $z = 0.32$ sample
- Taking into account of the anisotropy of the observation, we improve our constraints on the distance scale in both transverse and radial direction.
- Places strong constraint on cosmological parameters.



BOSS collaboration 2014
Anderson et al. 2014
Vargas, **Ho** et al. 2014

Two point functions only capture all information when the field is Gaussian.
But our universe is not that gaussian…

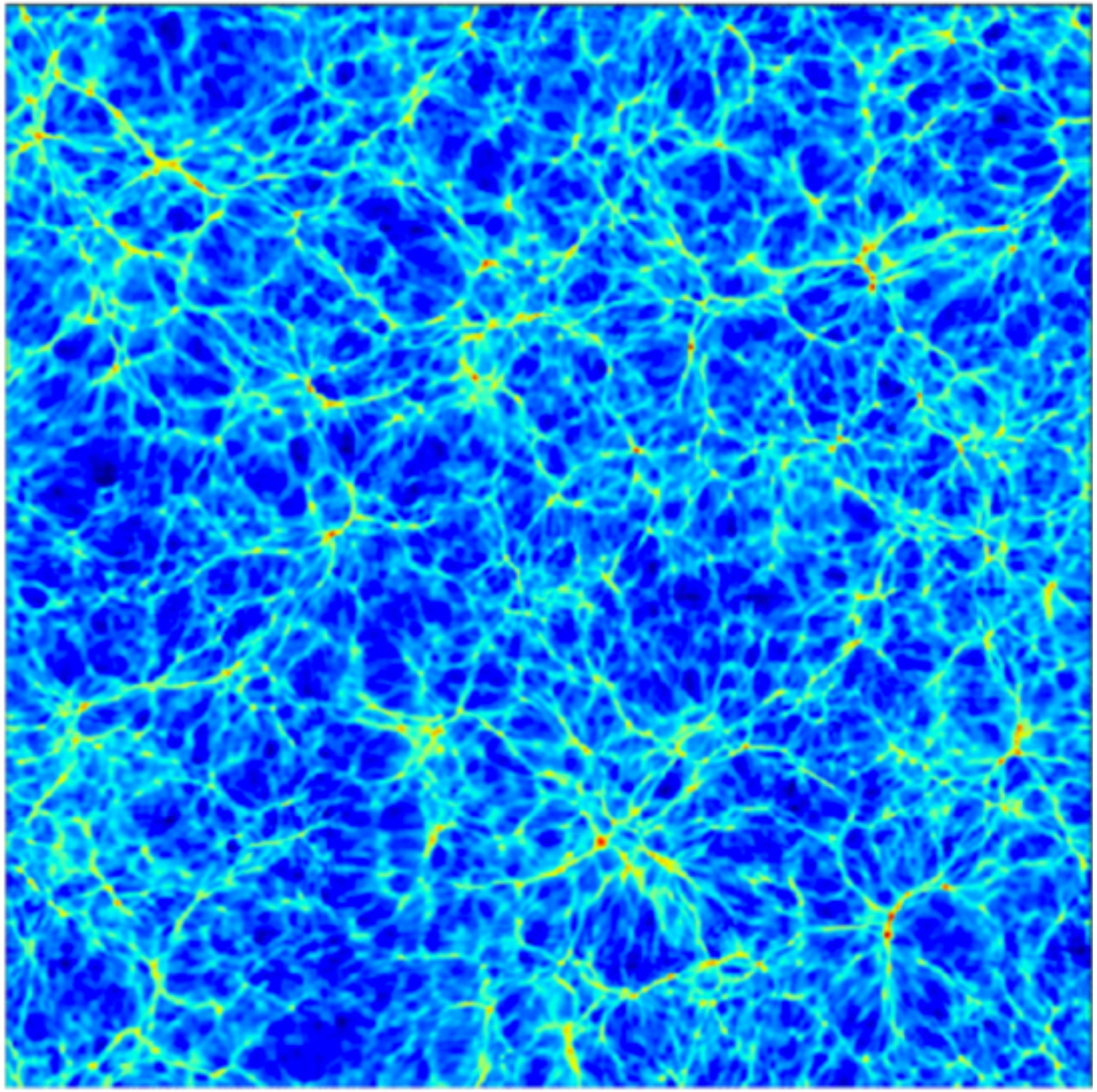


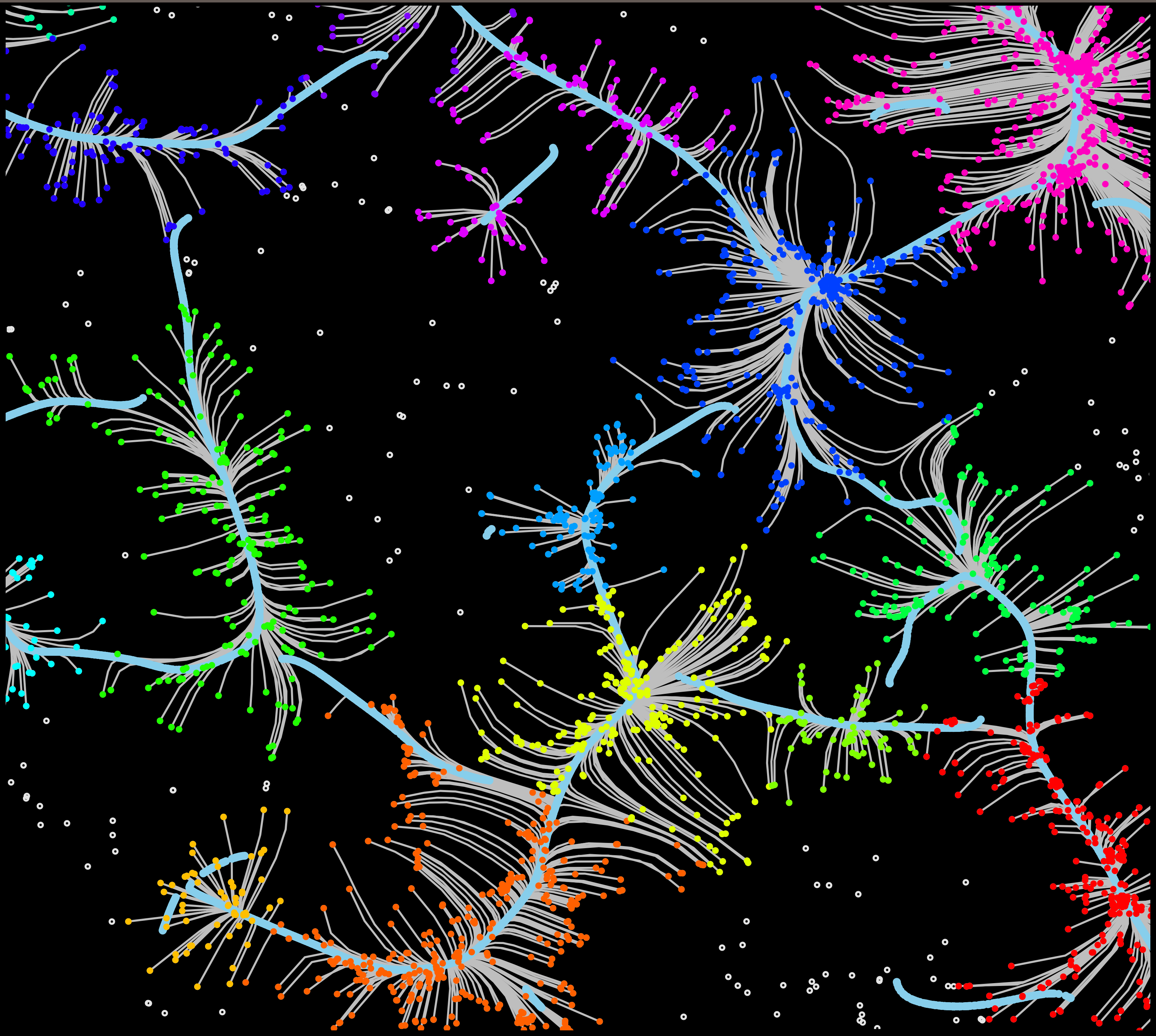
Figure Credit: Agarwal & Feldman

What do we do with all these interesting datasets?

- Standard analyses: 2 point correlation functions / clustering / stacking
- **Going beyond standard analyses**
- Going beyond 3D information that these large scale structure surveys provide.

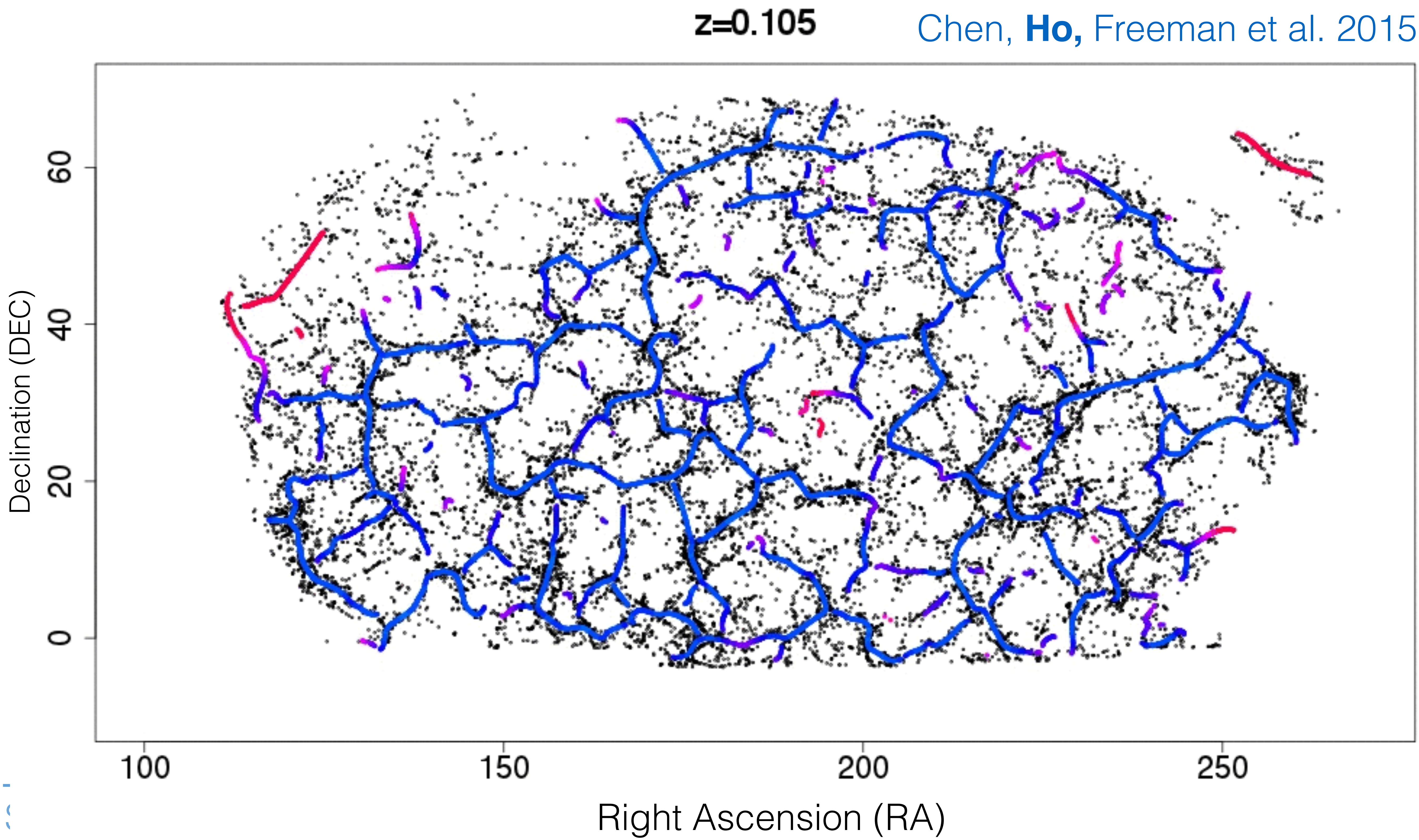
Going beyond Gaussian field with Subspace Constrained Mean Shift (SCMS)

Basic idea is to find ridges in the density field by looking at the Hessian matrix at each point.



Ozertem & Erdogmus, 2011
Chen, Genovese & Wasserman, 2014a
Chen, Genovese & Wasserman, 2014b
Chen, Ho, Freeman, Genovese, Wasserman, 2015

Cosmic Web reconstructed from SDSS



What do we learn that we didn't know already?

- Effects of filaments on galaxy masses
- Constraining 'intrinsic alignment' model of galaxies
- Filaments as a tracer of large scale structure
- Finding missing baryons with filaments
- Filaments help find dimmer galaxies
- Filaments can probe models of gravity

Effects of filaments on the stellar mass of the galaxies

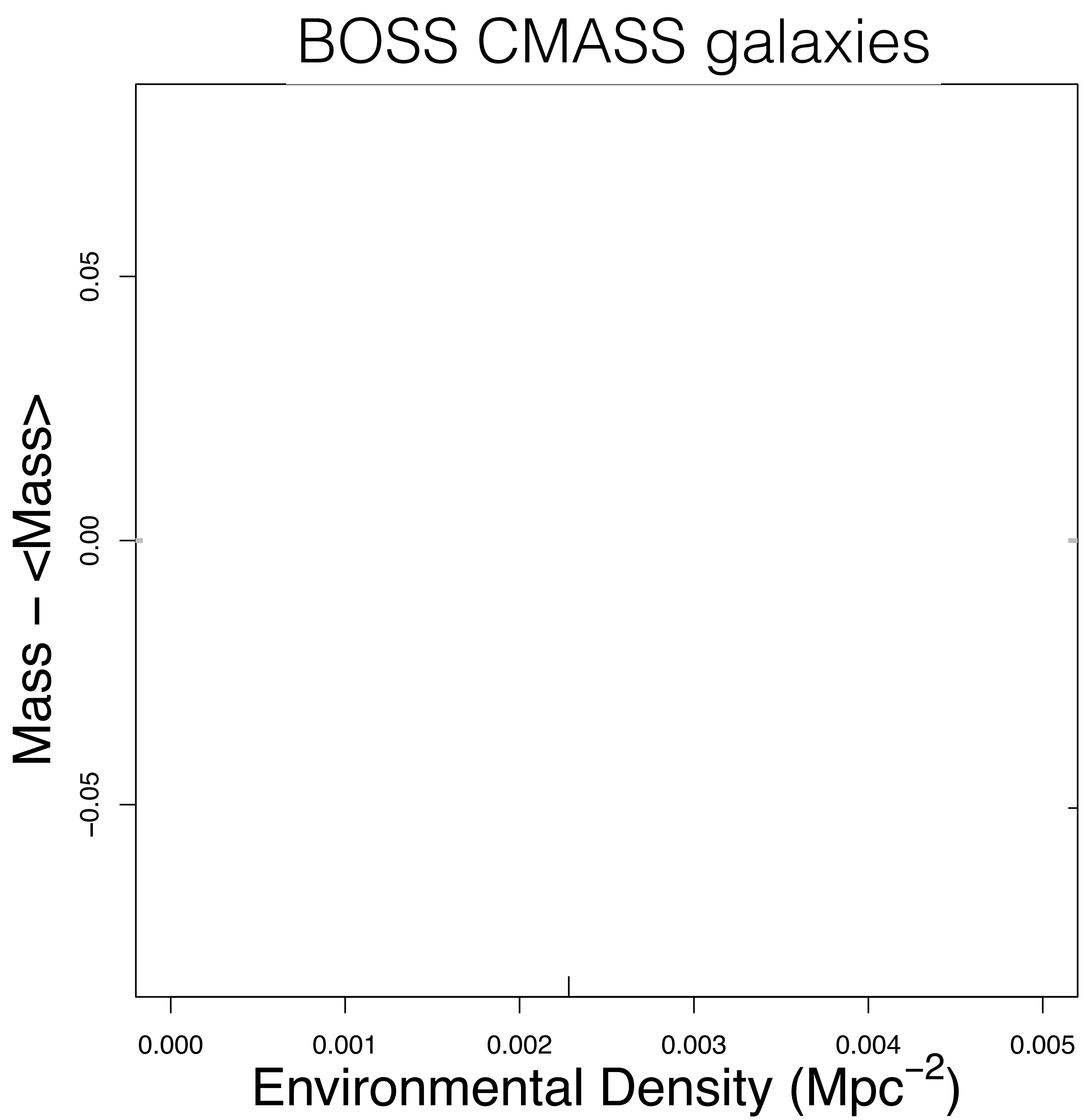
With the current standard scenario, we only need to know

1) the halo mass
2) the environmental density and
3) possibly the evolution of the parent halos

to understand the basic properties of galaxies.

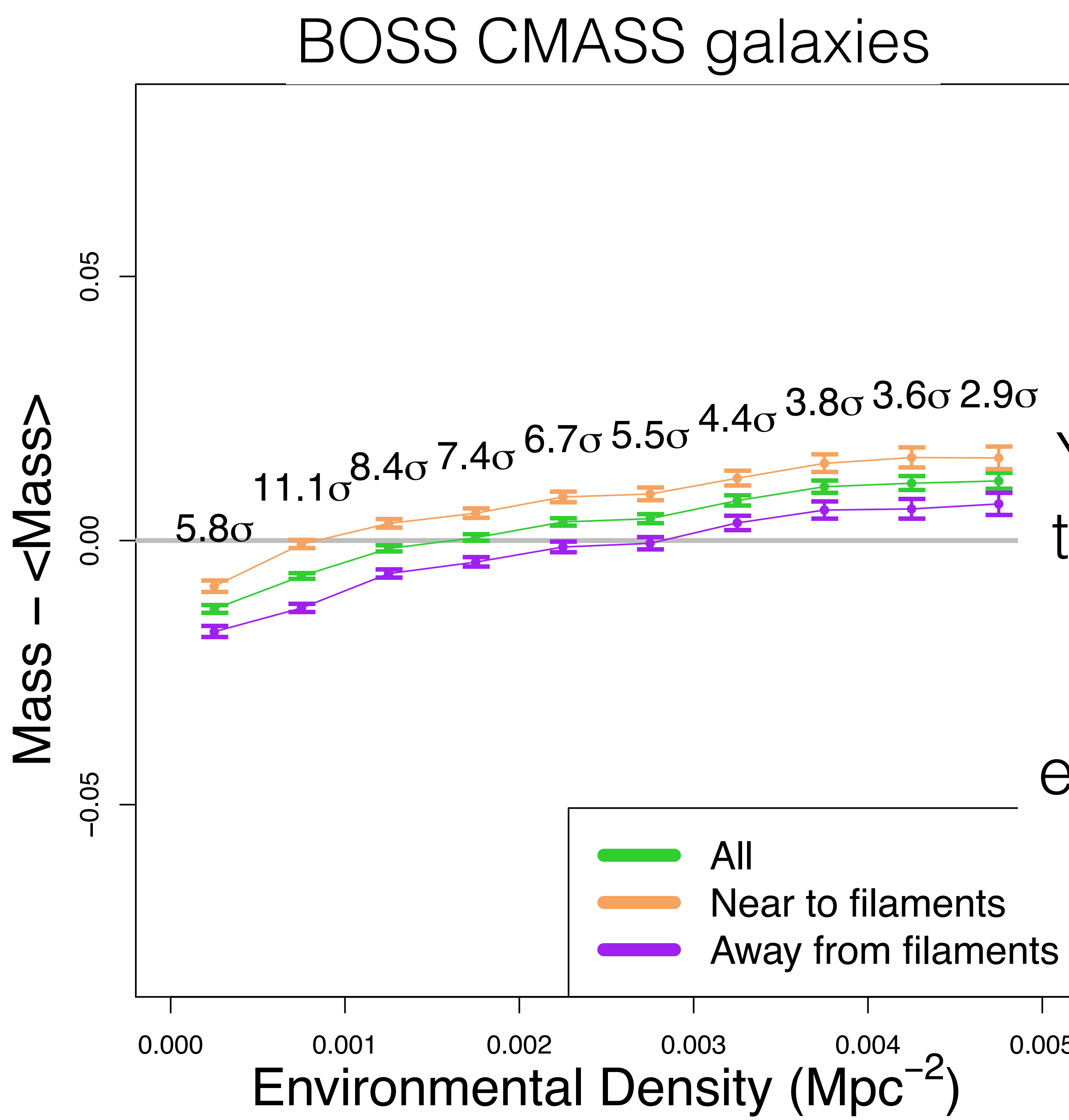
A good question to ask would be: Does filaments around the affect the galaxies? Or is it just the environment that matters?

Effects of filaments on the stellar mass of the galaxies



Chen, **Ho,** Mandelbaum et al. 2015

Effects of filaments on the stellar mass of the galaxies

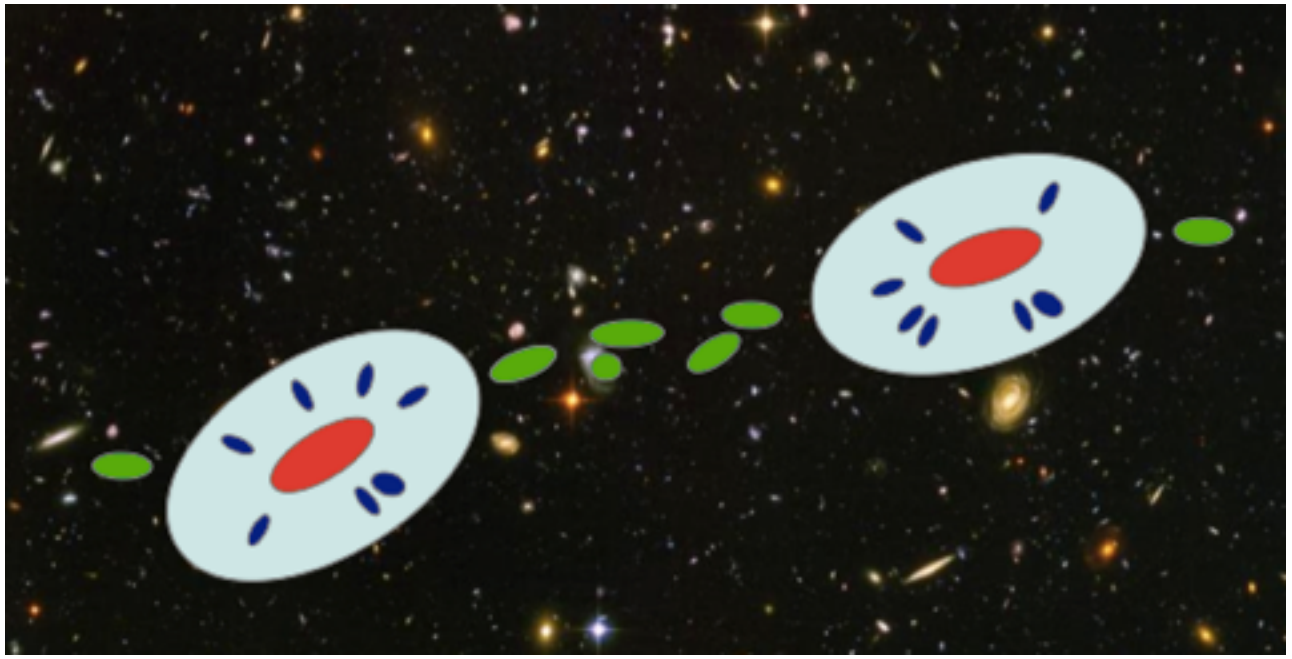


Yes! Filaments affect the galaxies in a way that is independent from the environmental density

Chen, **Ho,** Mandelbaum et al. 2015

Constraining “intrinsic alignment model” of galaxies
What is intrinsic alignment?

In weak gravitational lensing, one of the most challenging astrophysical systematic is called “intrinsic alignment”, which basically means galaxies are “intrinsically” aligned with each other due to interactions with the larger scale tidal fields.

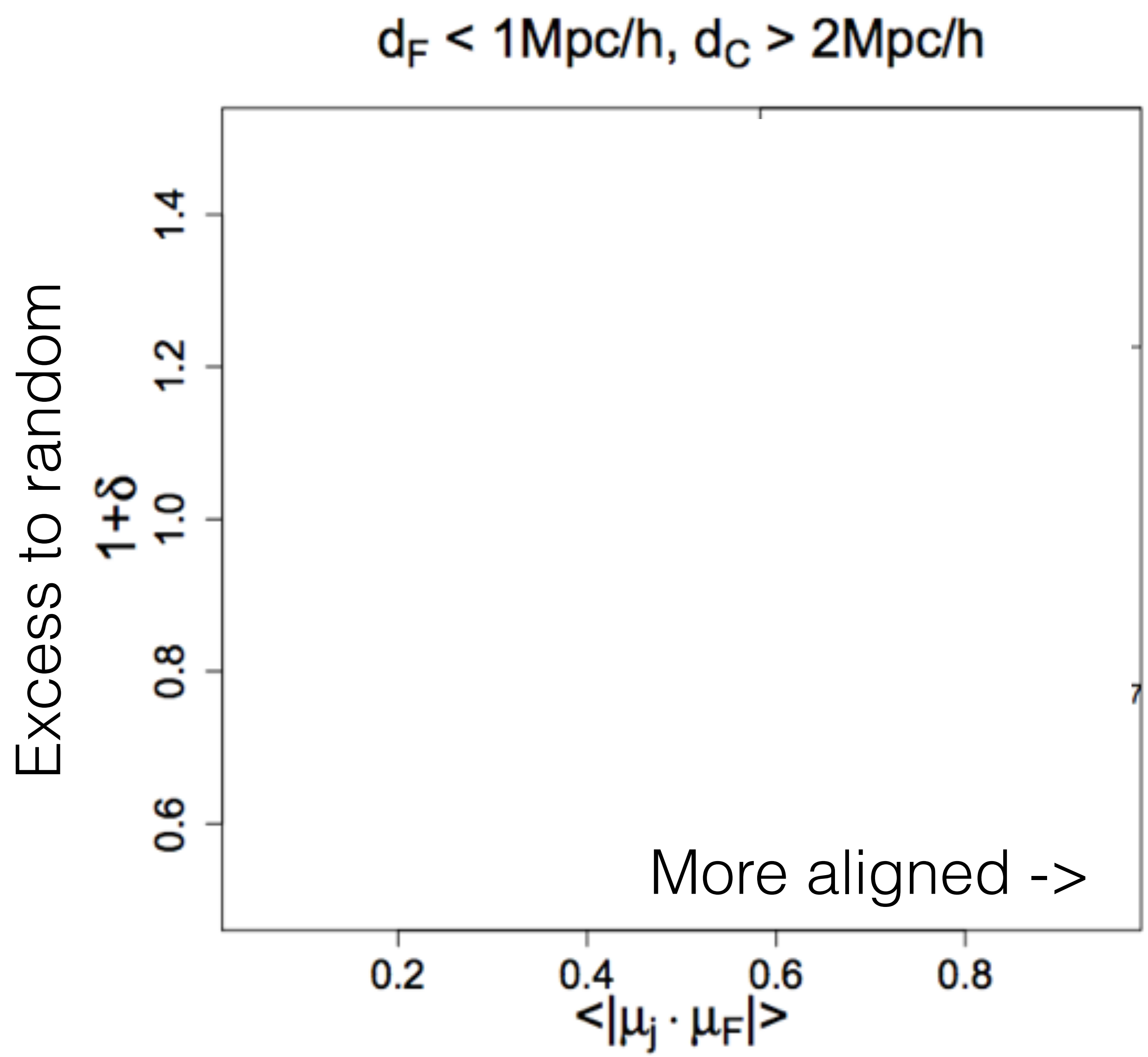


Constraining "intrinsic alignment model" of galaxies
What is intrinsic alignment?

In weak gravitational lensing, one of the most challenging astrophysical systematic is called "intrinsic alignment", which basically means galaxies are "intrinsically" aligned with each other due to interactions with the larger scale tidal fields.

If this systematic is left alone, this can significantly affect the final estimate of how much dark matter there is (and many other cosmological parameters).
Ex. Understanding the IA model increase our S/N on **dark energy equation of state** by factor of **2**;
-> Equivalent by **increasing the sky area by factor of 4** !
(Bridle & King 2007)

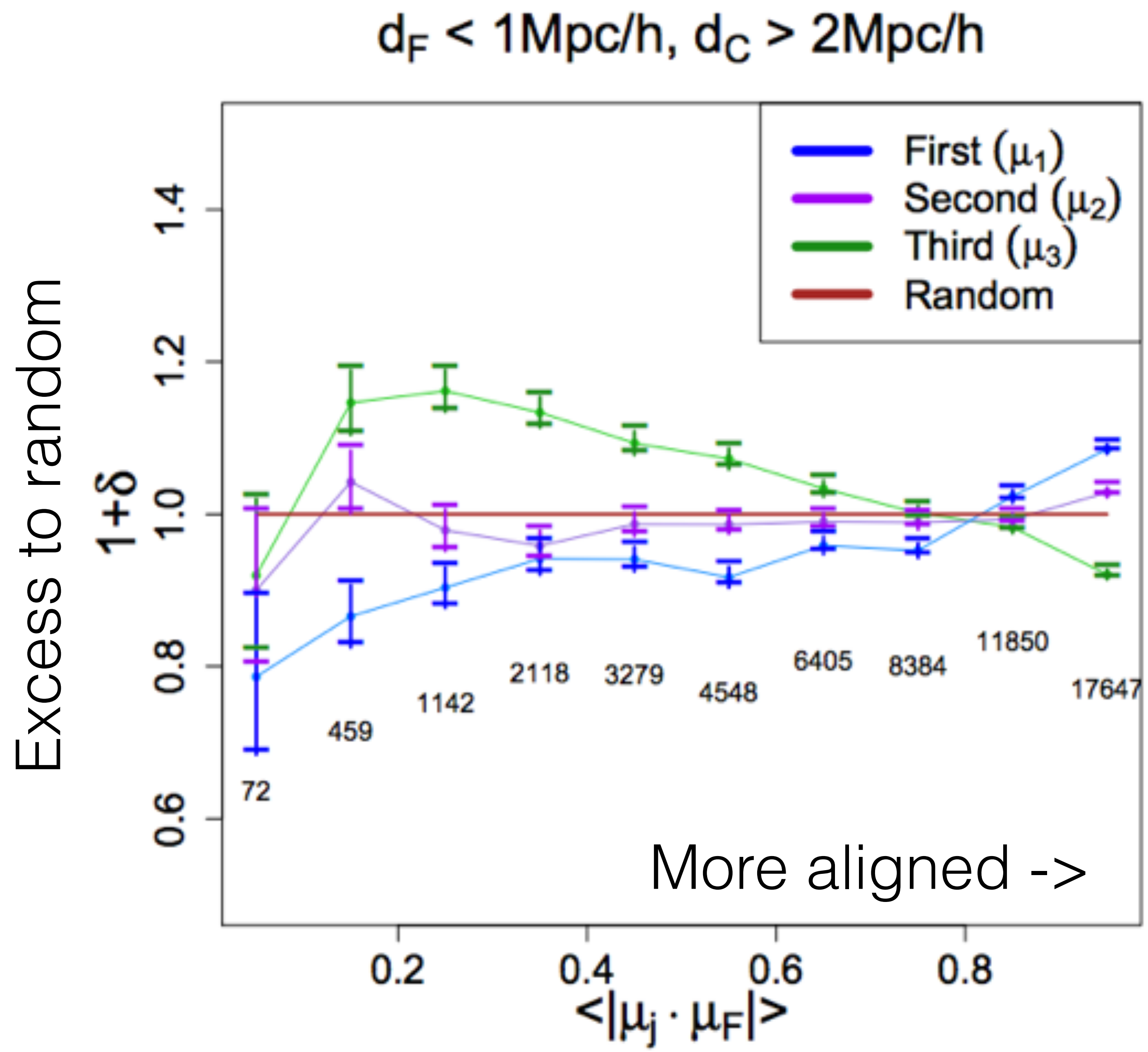
Constraining "intrinsic alignment model" of galaxies
Using filaments to trace tidal fields and thus shapes of galaxies



Chen, **Ho,** Tennetti et al. 2015

Using MassiveBalck II simulations, we study the Relationship between major axes of galaxies and the direction of filaments!

Constraining "intrinsic alignment model" of galaxies
Using filaments to trace tidal fields and thus shapes of galaxies



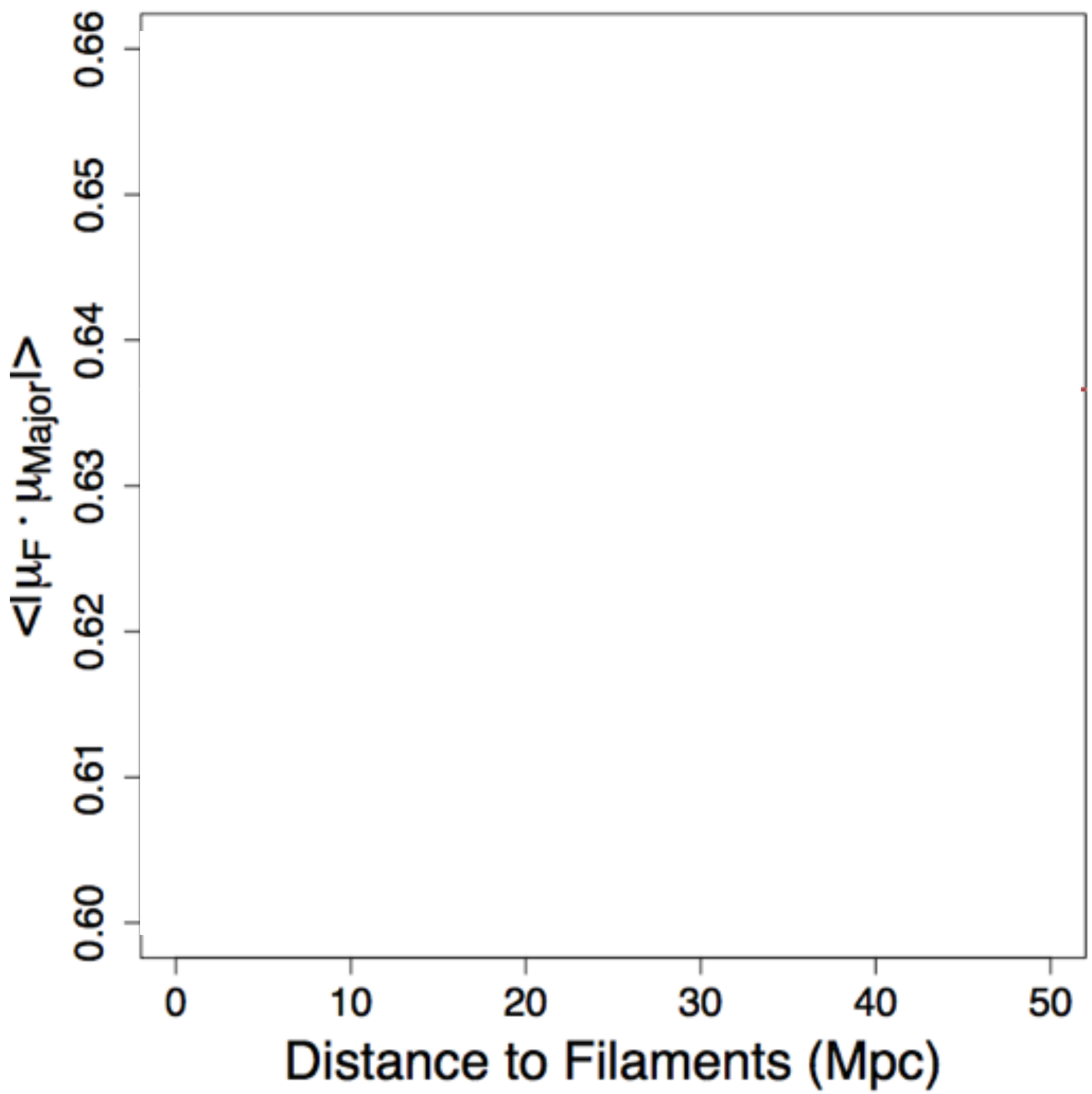
Chen, **Ho,** Tennetti et al. 2015

Using MassiveBalck II simulations, we study the Relationship between major axes of galaxies and the direction of filaments!

Constraining "intrinsic alignment model" of galaxies

How about real data? We next look at SDSS data

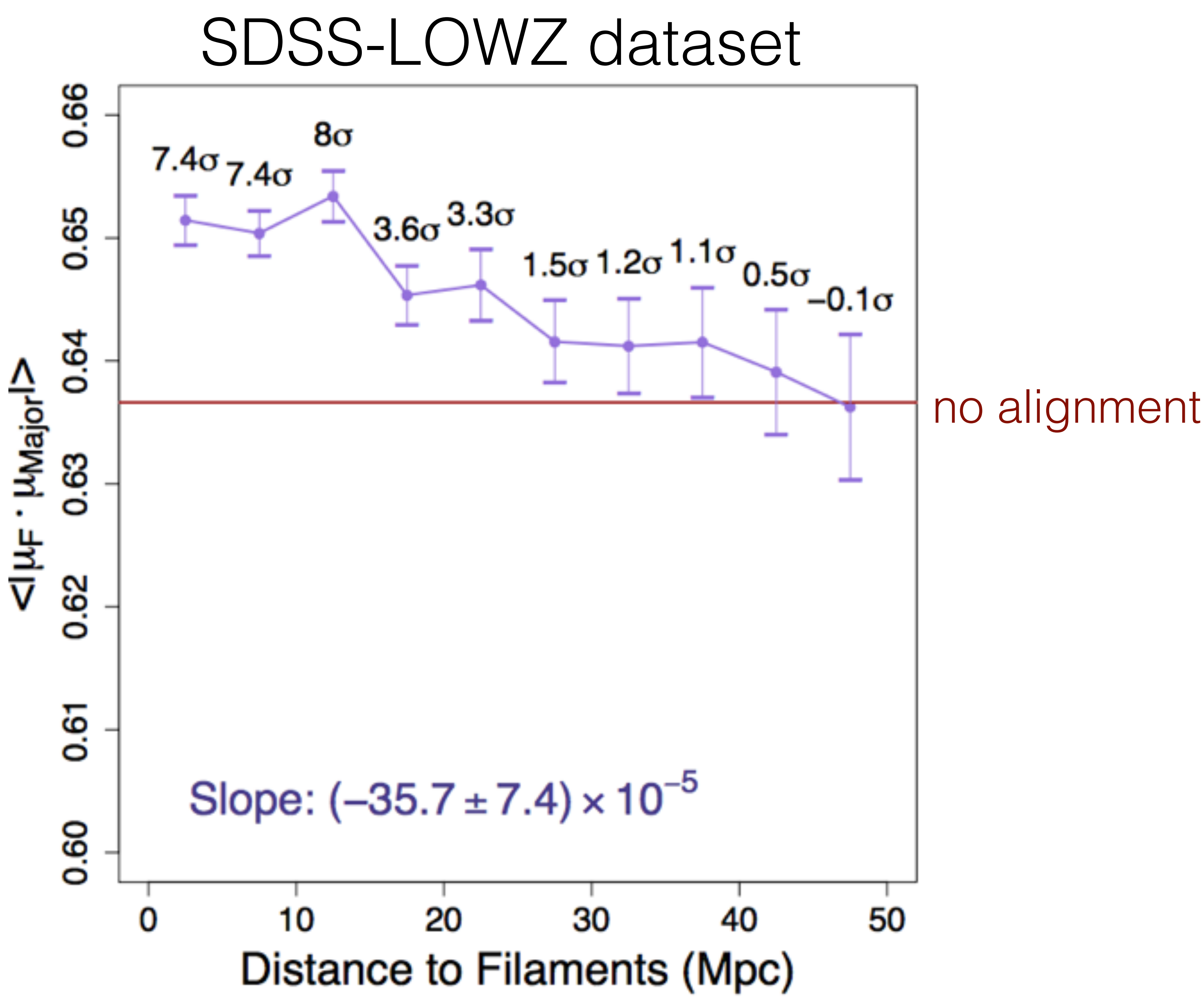
Dot product between filament direction and galaxy major axes



Preliminary results by Chen, Ho, Mandelbaum et al.

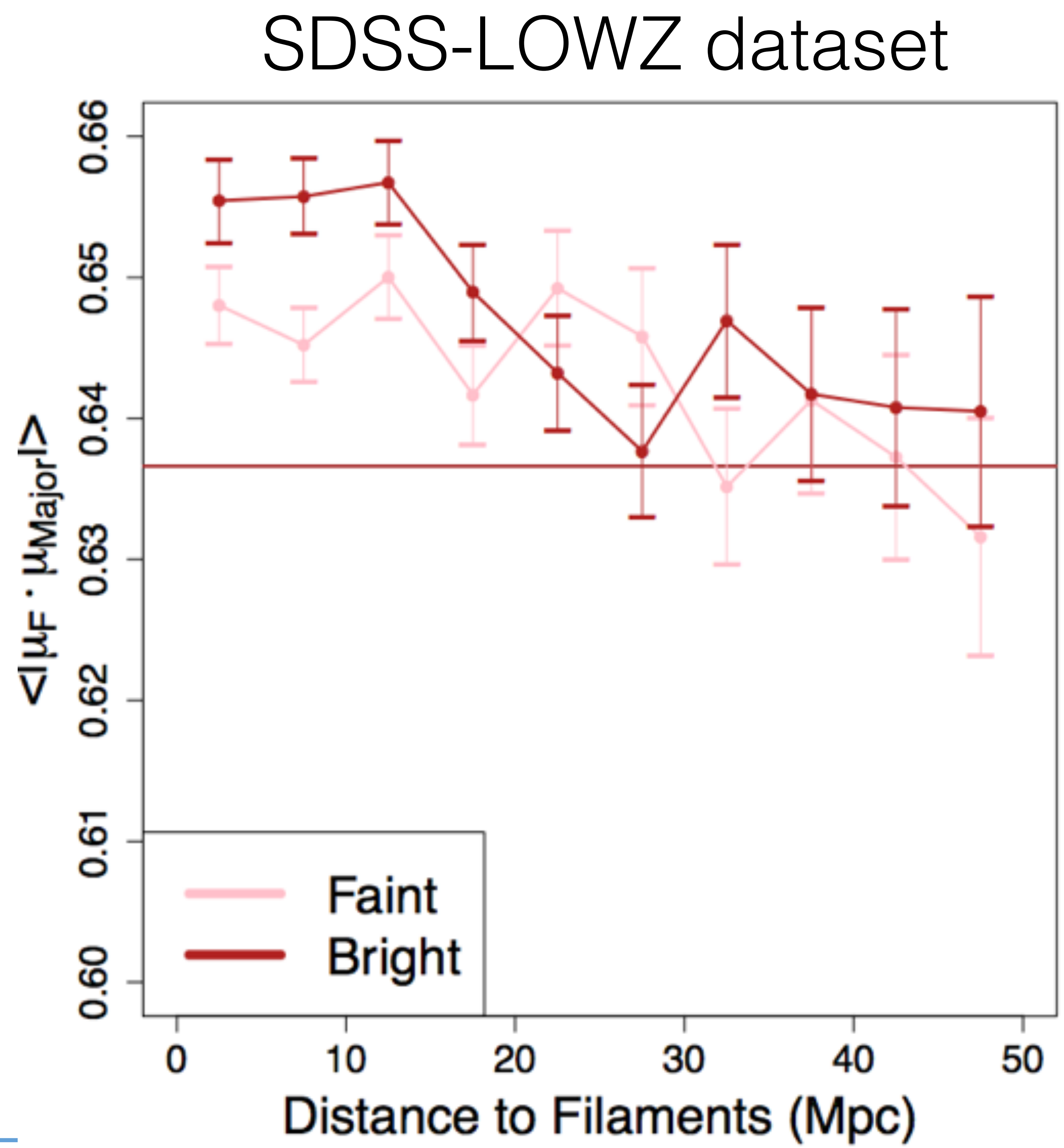
We find filaments in SDSS data and found significant alignments between the filament direction and the galaxy major axes

Dot product between filament direction and galaxy major axes



Preliminary results by Chen, Ho, Mandelbaum et al.

Alignment strength vs brightness of galaxies.



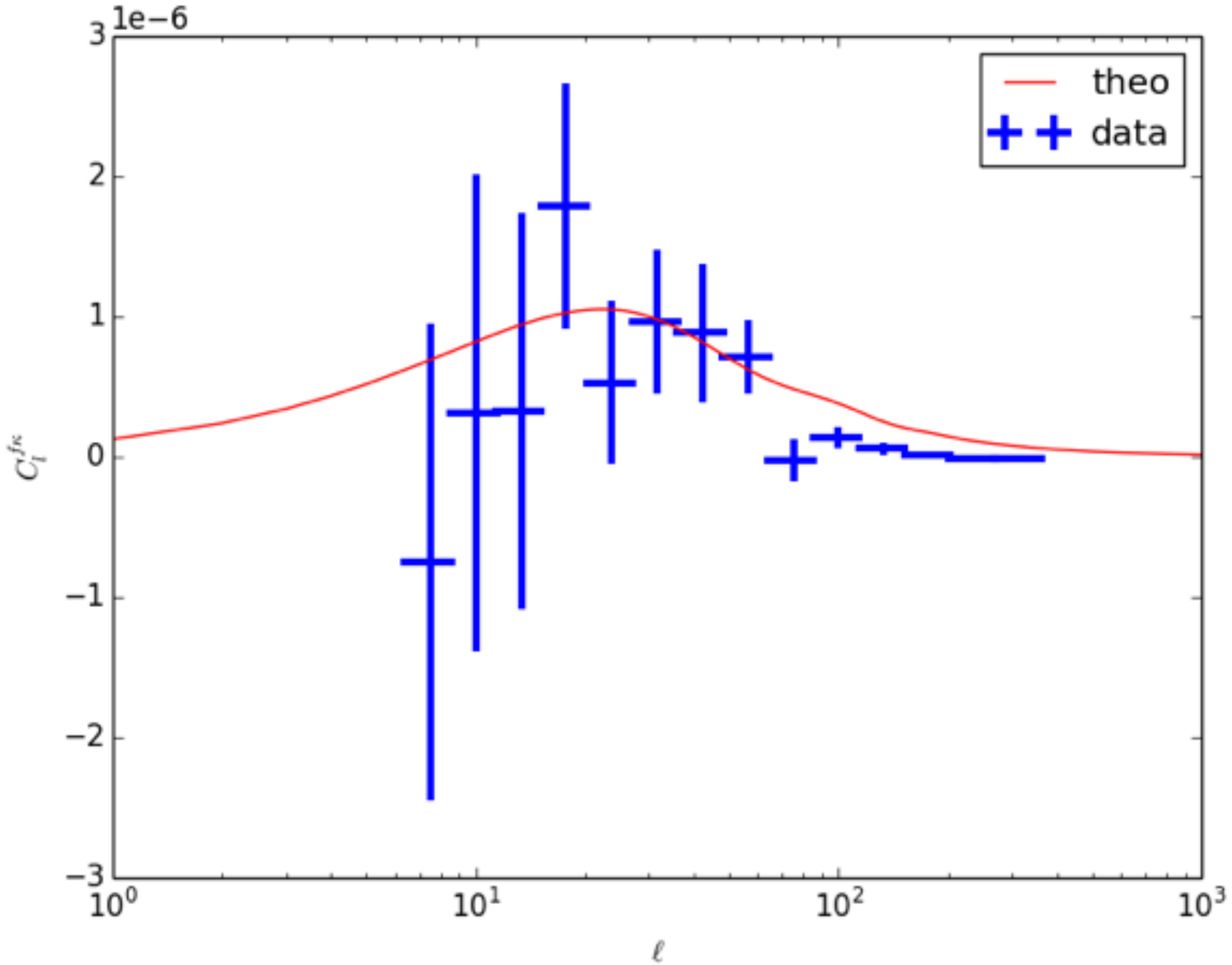
Preliminary results by Chen, Ho, Mandelbaum et al.

Filaments tracing LSS: therefore we expect filaments to lens !

Preliminary results: Siyu He, Yen-Chi Chen, Shadab Alam & S.H.

Filaments tracing LSS: therefore we expect filaments to lens !

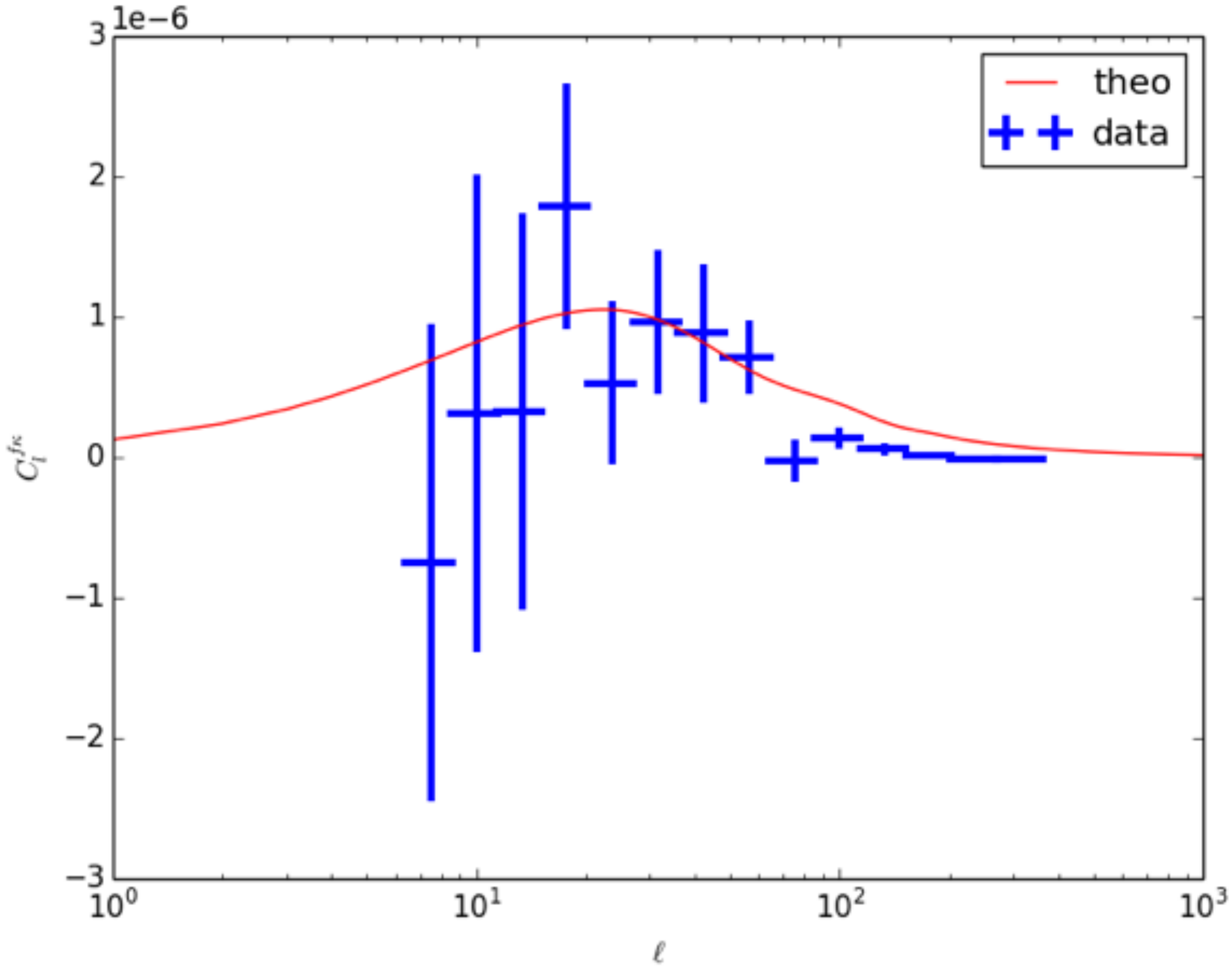
First filament X CMB lensing signal



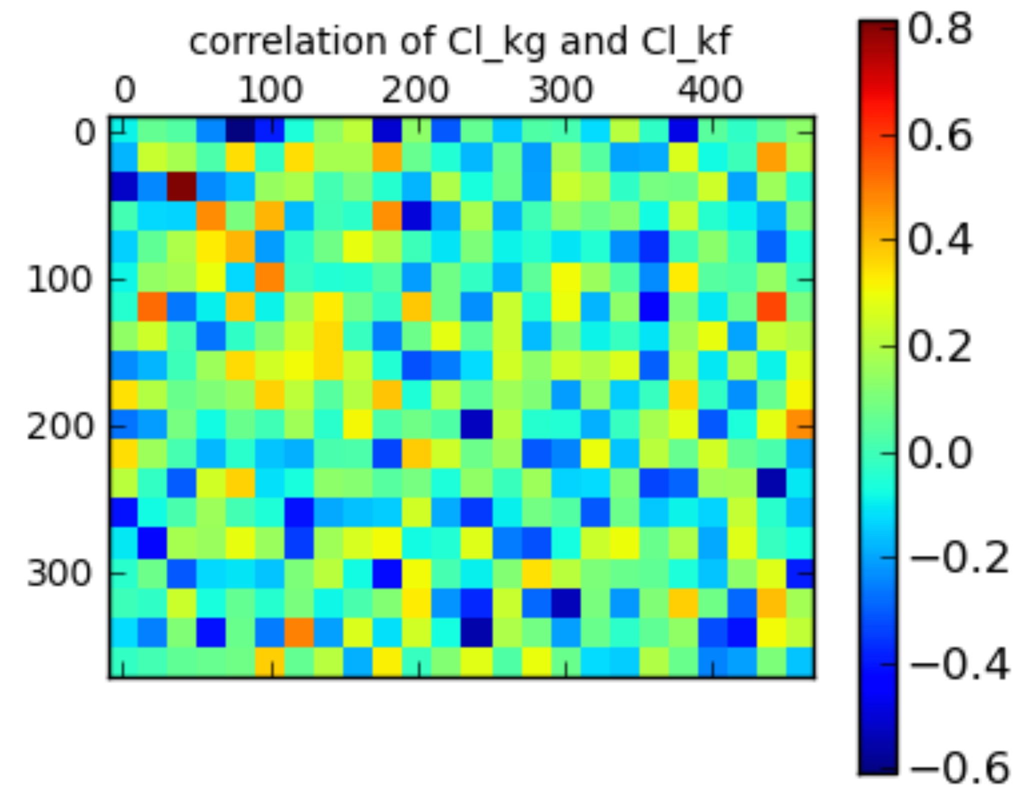
Preliminary results: Siyu He, Yen-Chi Chen, Shadab Alam & S.H.

Filaments tracing LSS: therefore we expect filaments to lens !

First filament X CMB lensing signal



Covariances between filaments X CMB Lensing and galaxies X CMB lensing



Preliminary results: Siyu He, Yen-Chi Chen, Shadab Alam & S.H.

Finding missing baryons with filaments

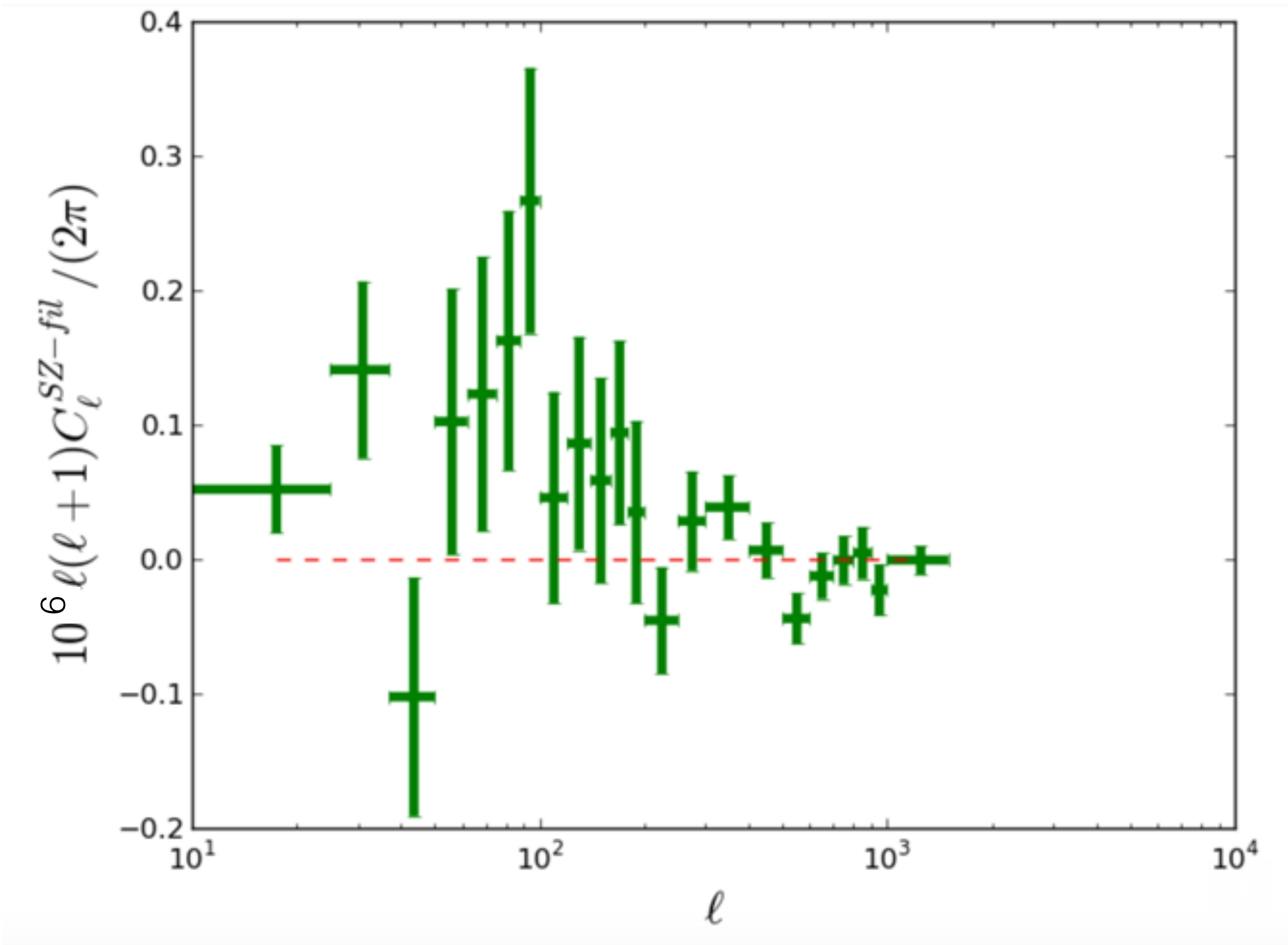
Missing Baryons are the expected amount of baryons that are not yet accounted for by looking at the “bright” stuff

Peebles & Fukugita (2004)

 Preliminary results: Anthony Pullen, Siyu He, Yen-Chi Chen & S.H.

Finding missing baryons with filaments

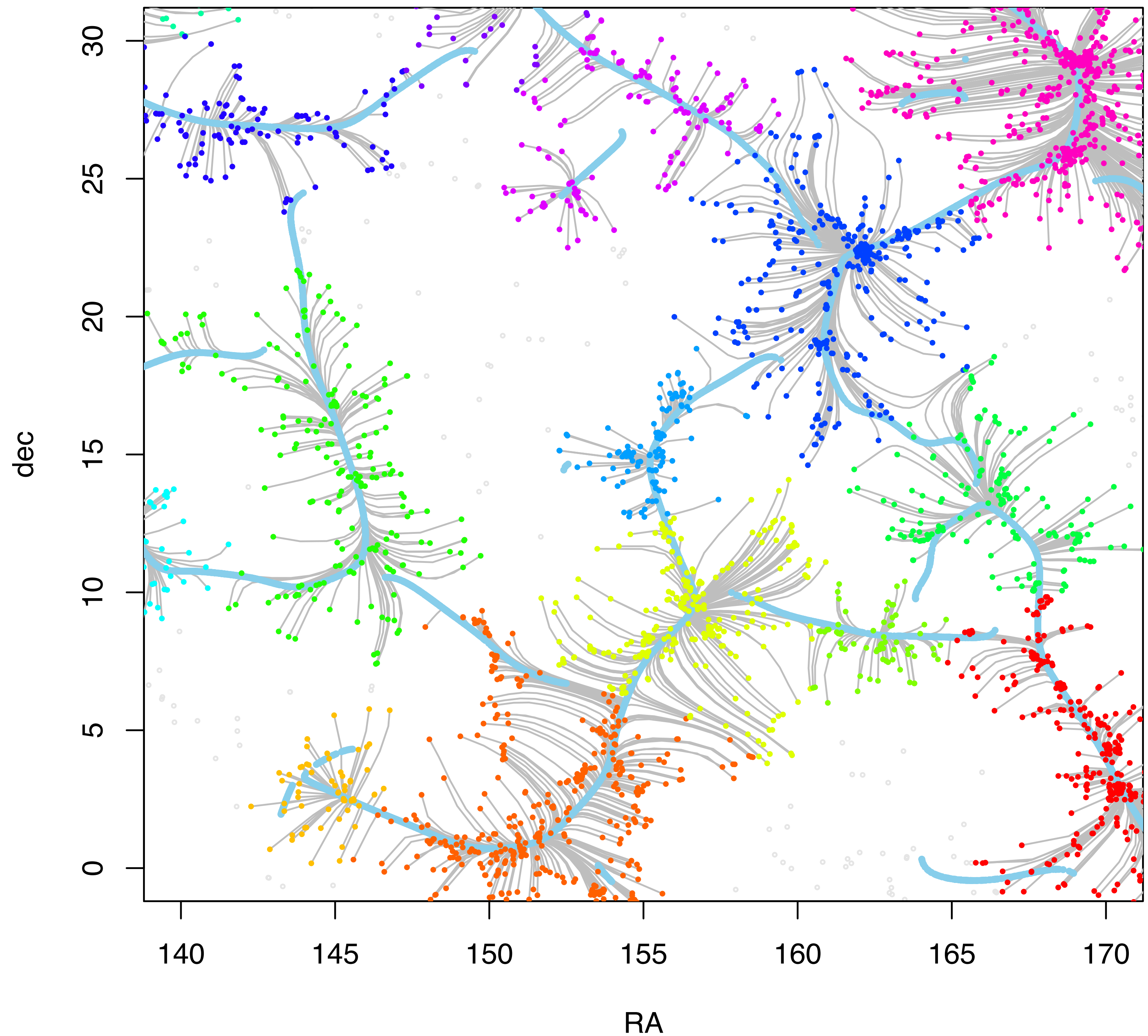
First filament X SZ signal with filaments at z=0.5-0.55



Preliminary results: Anthony Pullen, Siyu He, Yen-Chi Chen & S.H.

Finding dimmer galaxies

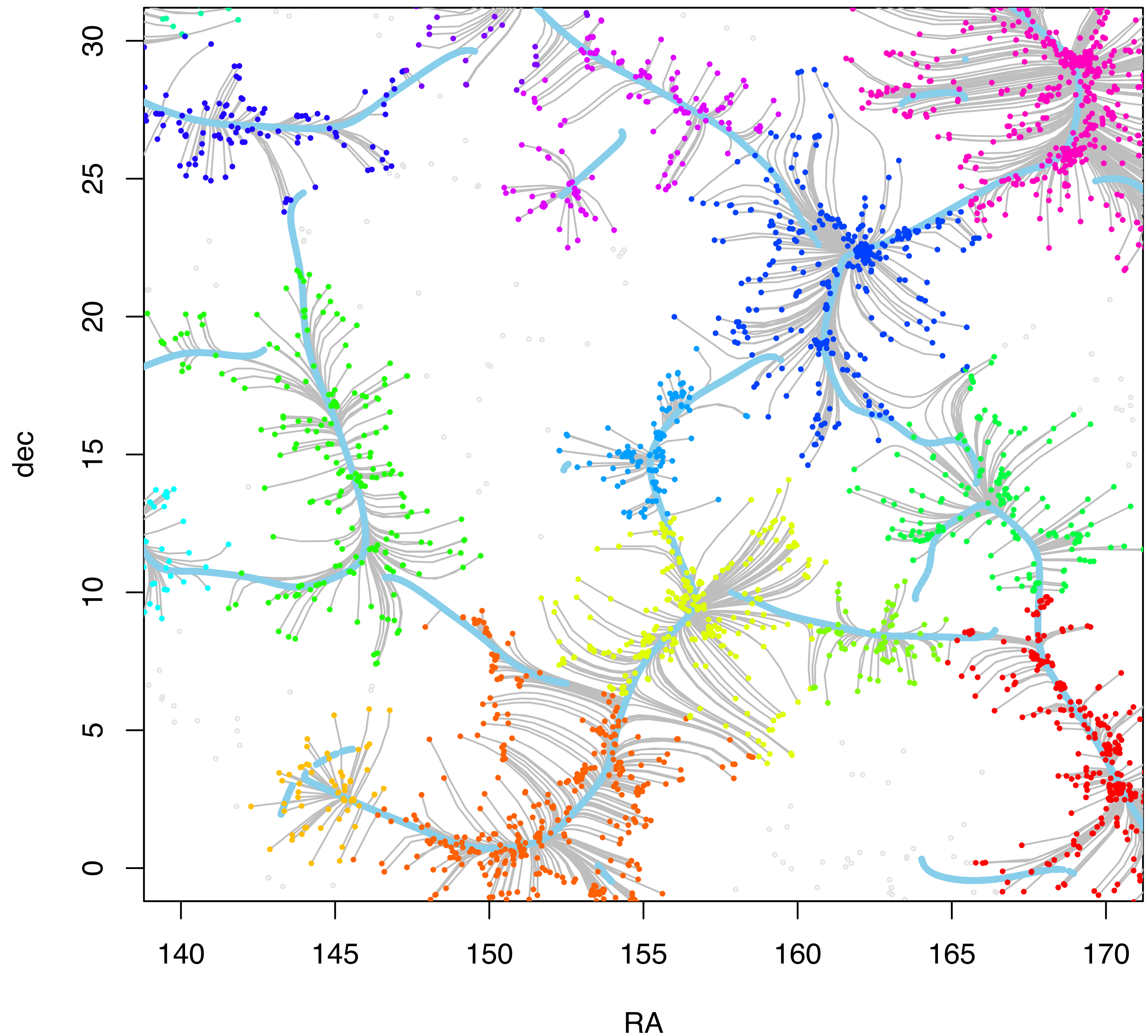
- We know that ridges of density field are regions of medium to high density region
- It is likely that they will harbor dimmer galaxies than the sample that we used to construct the filaments
- One can use these filaments as an independent catalog to constrain redshift distribution of photometric galaxies



Preliminary: Francois Lanusse, Yen-Chi Chen & S.H.

Finding dimmer galaxies

- We know that ridges of density field are regions of medium to high density region
- It is likely that they will harbor dimmer galaxies than the sample that we used to construct the filaments
- One can use these filaments as an independent catalog to constrain redshift distribution of photometric galaxies



Preliminary: Francois Lanusse, Yen-Chi Chen & S.H.

Going beyond standard analyses (mini-conclusion)

- Finding a new way to look at the density field opens up many possibilities of extracting information from the non-gaussian density field.
- Looking at filaments is just *one of the many ways of going beyond the traditional analyses.
- Very much open to discussion on other ways of looking at the non-gaussian fields: Level sets? modes? sheets? ??

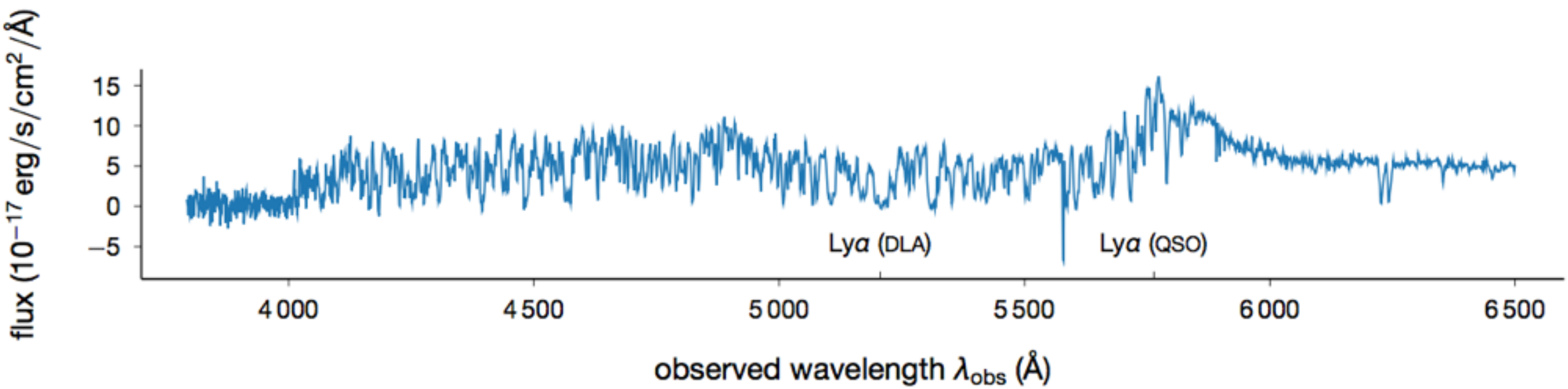
What do we do with all these interesting datasets?

- Standard analyses: 2 point correlation functions / clustering / stacking
- Going beyond standard analyses
- **Going beyond 3D information that these large scale structure surveys provide: Digging into the noise!**

Digging into the noise: What we do (and can do) with large spectroscopic datasets

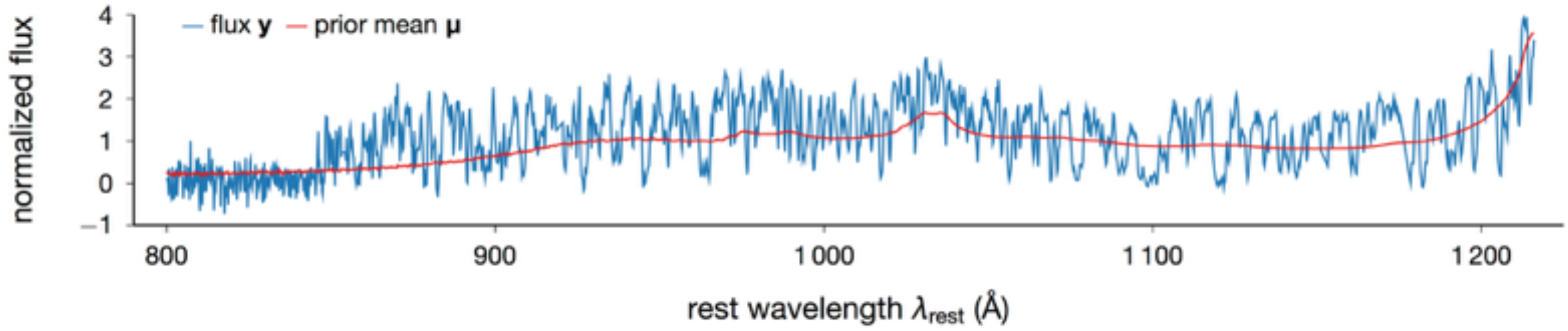
- ~2 million spectra: for each source, we have O(1000)s datapoint, each tells you how bright the source is at a very small range of wavelength
- We **visually inspect** > 1 million of all quasar (blackholes) targets. Takes approximately 6 person years.
- To find a) **special absorbers** such as Damped Lyman Alpha systems b) see if it is a **quasar** (blackhole) or not.
- We usually do not use much of the spectrum except to get its redshift, but we can and we should!

Here is the problem:
Finding dips in low signal to noise data



Processing example

We first compute the model evidence of the *null model;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\neg \mathrm{DLA}}) = -2\,589$.

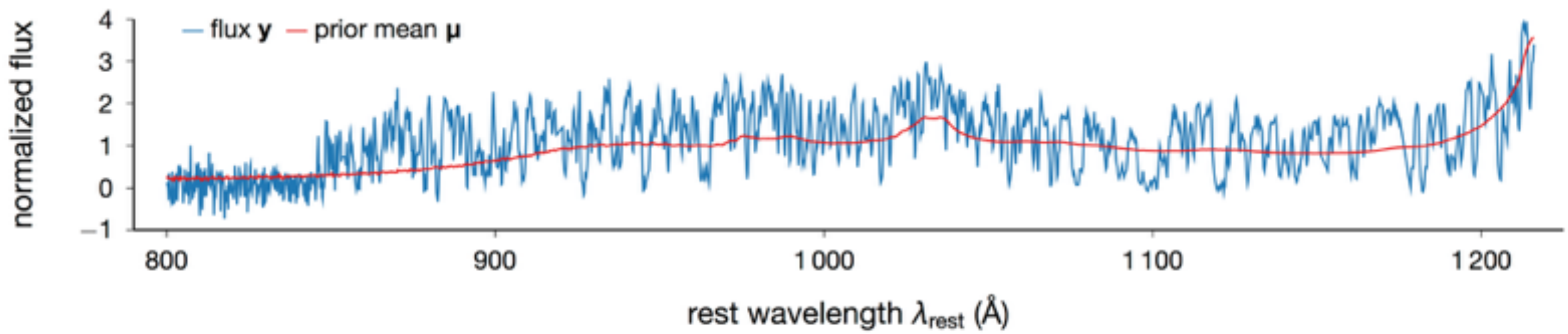


We approach this problem via *Bayesian model selection.* We build bespoke *Gaussian process* priors to model normalized QSO emission spectra in the range $\lambda \in [800, 1\,216]$ Å. We build a *null model* for spectra without intervening DLAs, then extend this model to the DLA case.

Processing example

Garnett, Ho & Schneider, International conference on Machine Learning 2015

We first compute the model evidence of the *null model;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\neg\mathrm{DLA}}) = -2\,589$.



We approach this problem via *Bayesian model selection.* We build bespoke *Gaussian process* priors to model normalized QSO emission spectra in the range $\lambda \in [800, 1\,216]$ Å. We build a *null model* for spectra without intervening DLAs, then extend this model to the DLA case.

Null model

We model QSO emission spectra without DLAs along the line of sight as having been generated independently from an unknown *joint Gaussian process* prior

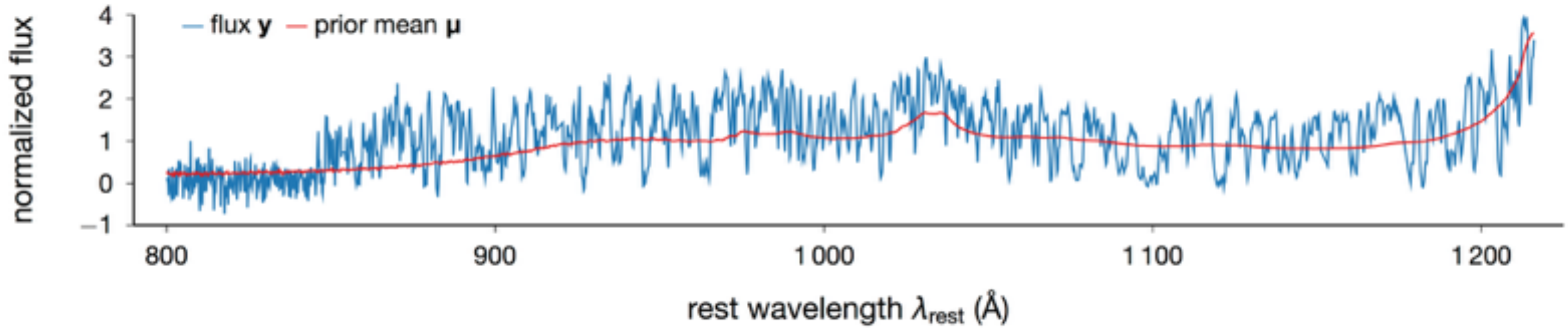
$$p(\mathbf{y}) = \mathcal{GP}(\mathbf{y}; \boldsymbol{\mu}, \mathbf{K} + \operatorname{diag} \mathbf{v}).$$

We learn the parameters of this Gaussian process from nearly *50 000* example SDSS spectra. With these fixed, we may compute the *null model evidence* in closed form.

Processing example

Garnett, Ho & Schneider, International conference on Machine Learning 2015

We first compute the model evidence of the *null model;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\neg\mathrm{DLA}}) = -2\,589$.



We approach this problem via *Bayesian model selection.* We build bespoke *Gaussian process* priors to model normalized QSO emission spectra in the range $\lambda \in [800, 1\,216]$ Å. We build a *null model* for spectra without intervening DLAs, then extend this model to the DLA case.

Null model

We model QSO emission spectra without DLAs along the line of sight as having been generated independently from an unknown *joint Gaussian process* prior

$$p(\mathbf{y}) = \mathcal{GP}(\mathbf{y}; \boldsymbol{\mu}, \mathbf{K} + \operatorname{diag} \mathbf{v}).$$

We learn the parameters of this Gaussian process from nearly *50 000* example SDSS spectra. With these fixed, we may compute the *null model evidence* in closed form.

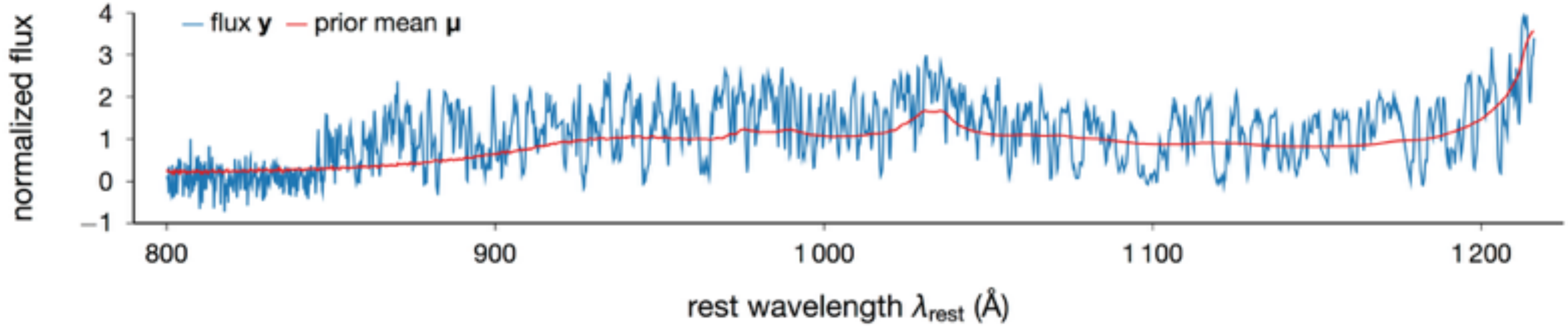
DLA model

Assume a DLA occurs along the line of sight to a quasar at redshift z_{DLA} with column density N_{HI}. The effect on observations is to *multiply by a known absorption function τ:*

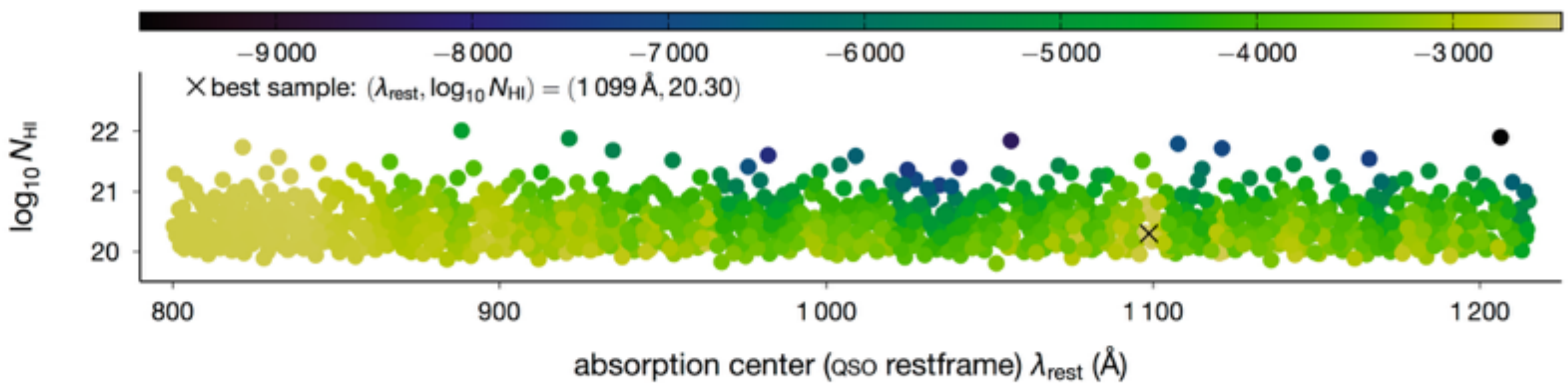
$$y(\lambda) = f(\lambda) \exp\bigl(-\tau(\lambda; z_{\mathrm{DLA}}, N_{\mathrm{HI}})\bigr) + \epsilon(\lambda).$$

Gaussian processes are *closed under linear transformations!* To compute the model evidence, we must *marginalize* the parameters $(z_{\mathrm{DLA}}, N_{\mathrm{HI}})$; here we use *quasi Monte Carlo.*

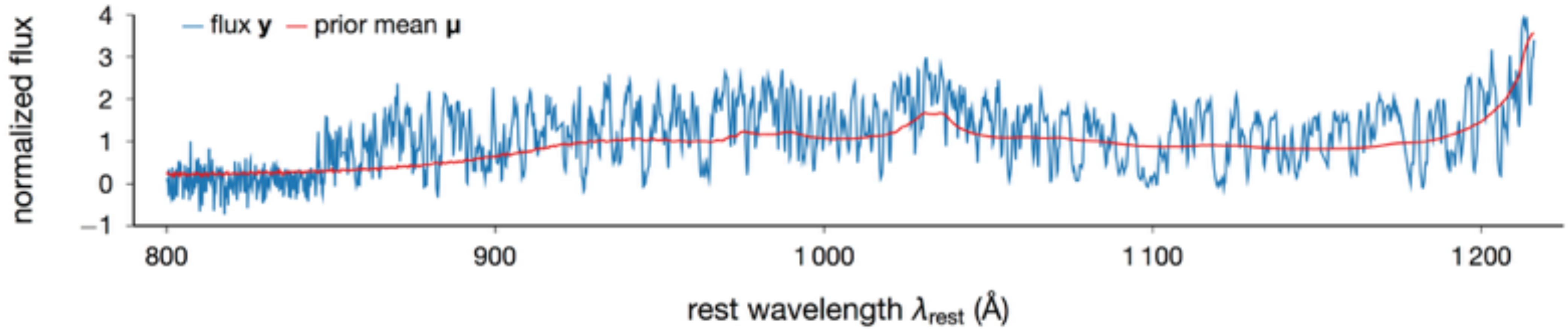
We first compute the model evidence of the *null model;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\neg\mathrm{DLA}}) = -2\,589$.



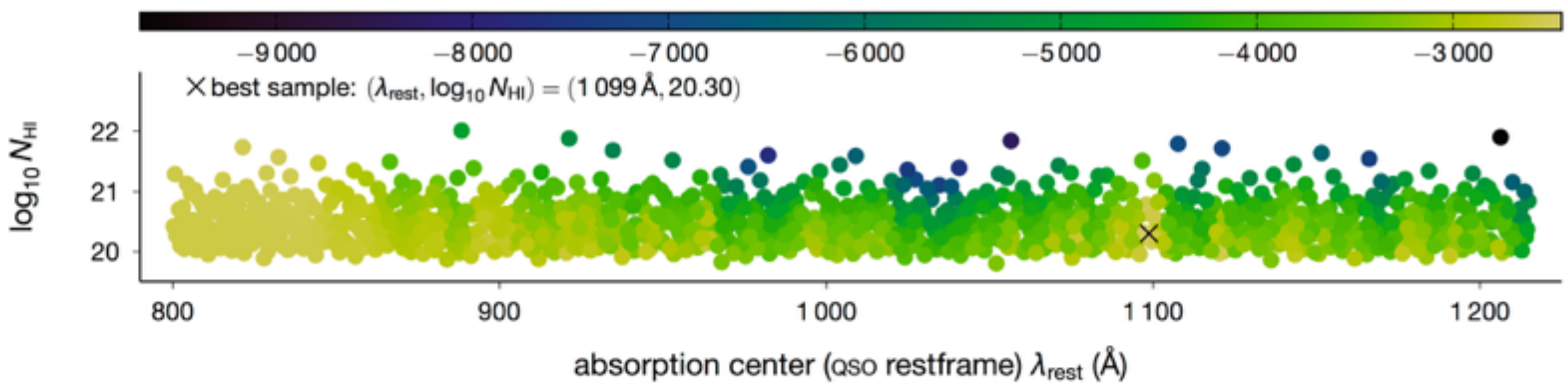
Next we estimate the model evidence for the *DLA model* via *quasi Monte Carlo;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\mathrm{DLA}}) = -2\,453$. The Bayes factor in favor of the DLA model is *overwhelming* (136 nats).



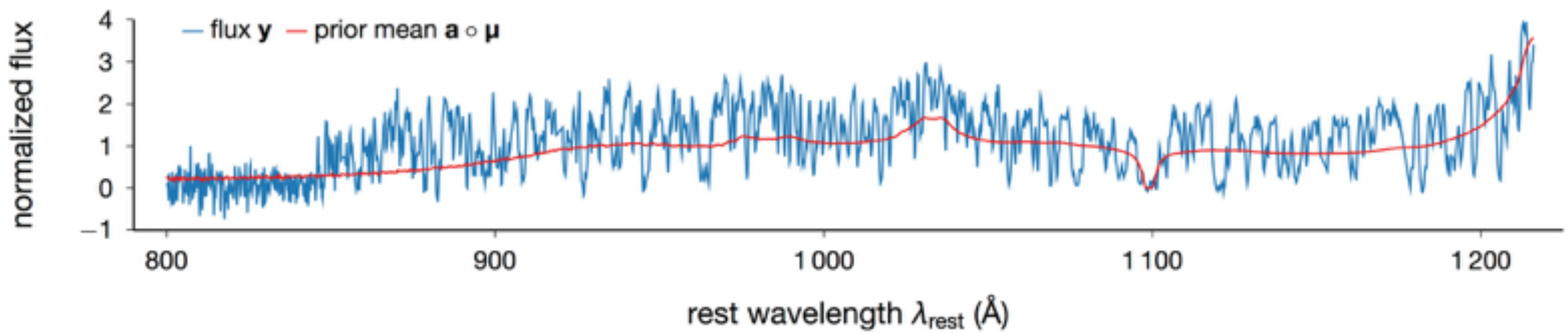
We first compute the model evidence of the *null model;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\neg\mathrm{DLA}}) = -2\,589$.



Next we estimate the model evidence for the *DLA model* via *quasi Monte Carlo;* for our example spectrum we have $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\mathrm{QSO}}, \mathcal{M}_{\mathrm{DLA}}) = -2\,453$. The Bayes factor in favor of the DLA model is *overwhelming* (136 nats).



The sample with the highest likelihood is an *extremely close match* to previously published values $(\lambda_{\mathrm{rest}}, \log_{10} N_{\mathrm{HI}}) = (1\,098\,\text{Å}, 20.29)$.



Garnett, Ho, Schneider, ICML 2015
Garnett, Ho, Bird & Schneider, 2016

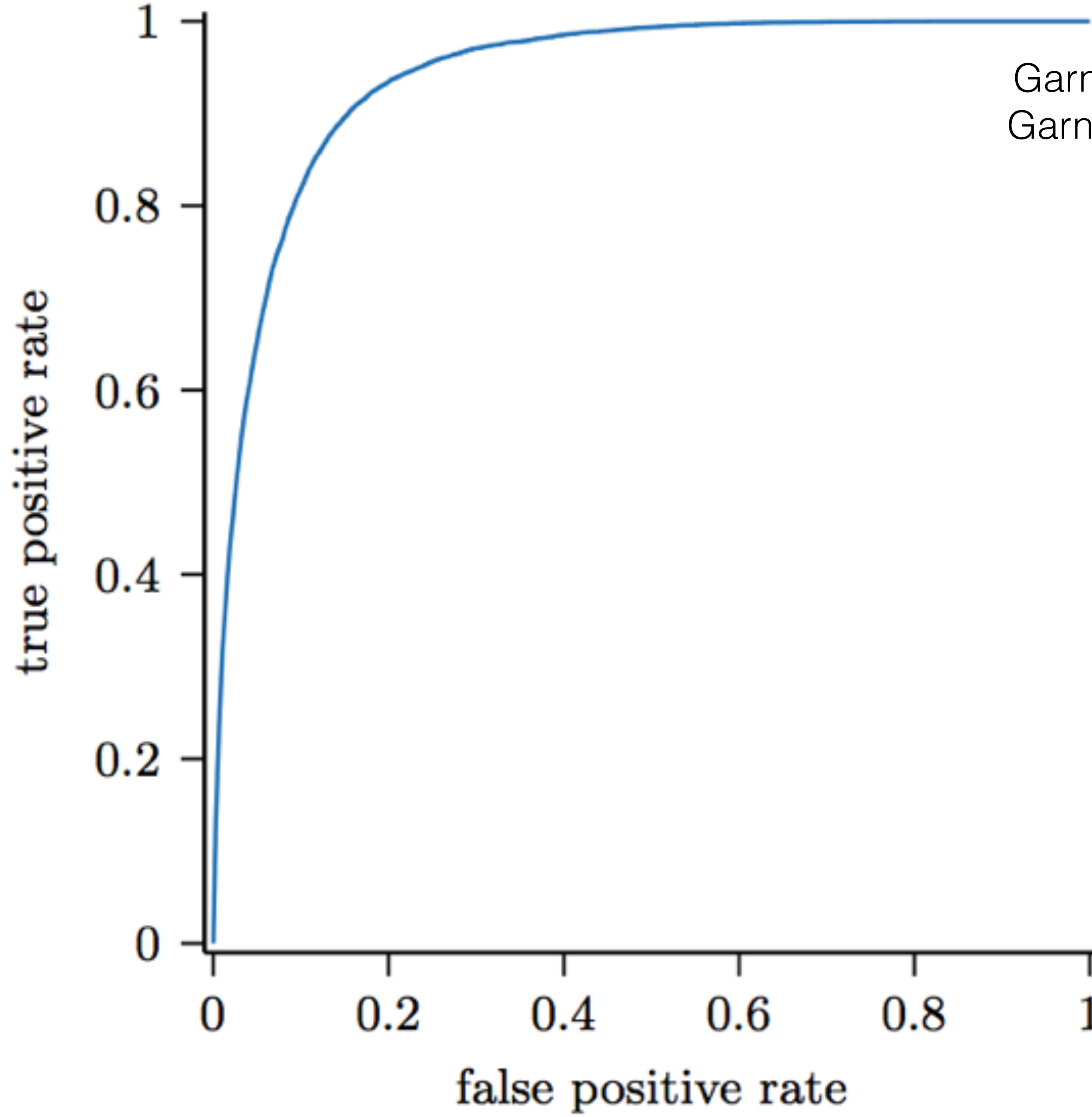
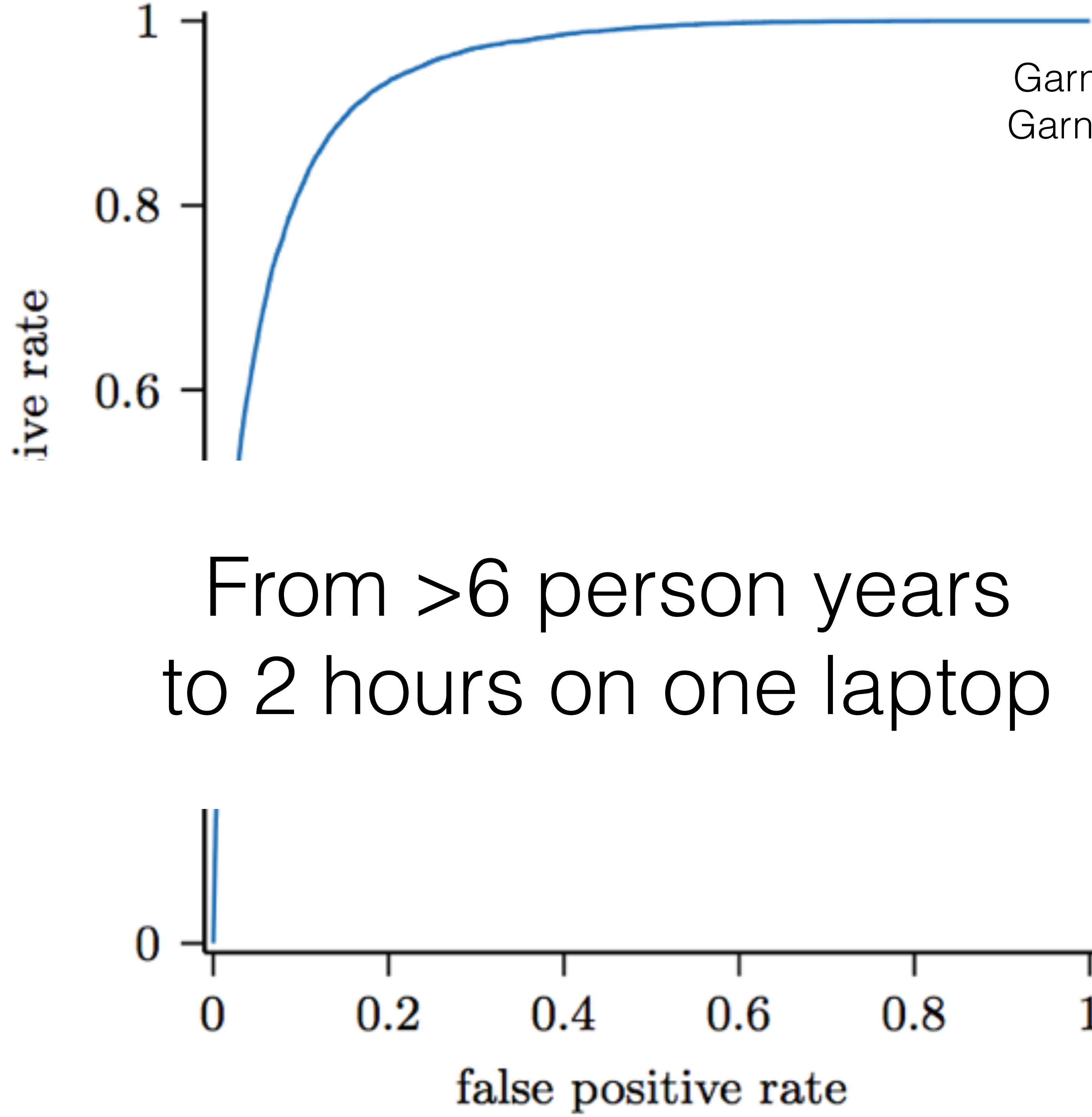


Figure 2. We processed over 100,000 Quasar sight-lines from SDSS and ranked them by posterior odds in favor of the DLA model. Ground truth is approximated by previous semi-automated approaches. The area under the ROC curve is 94.1%, and many “false positives” are likely to be newly discovered DLAs.

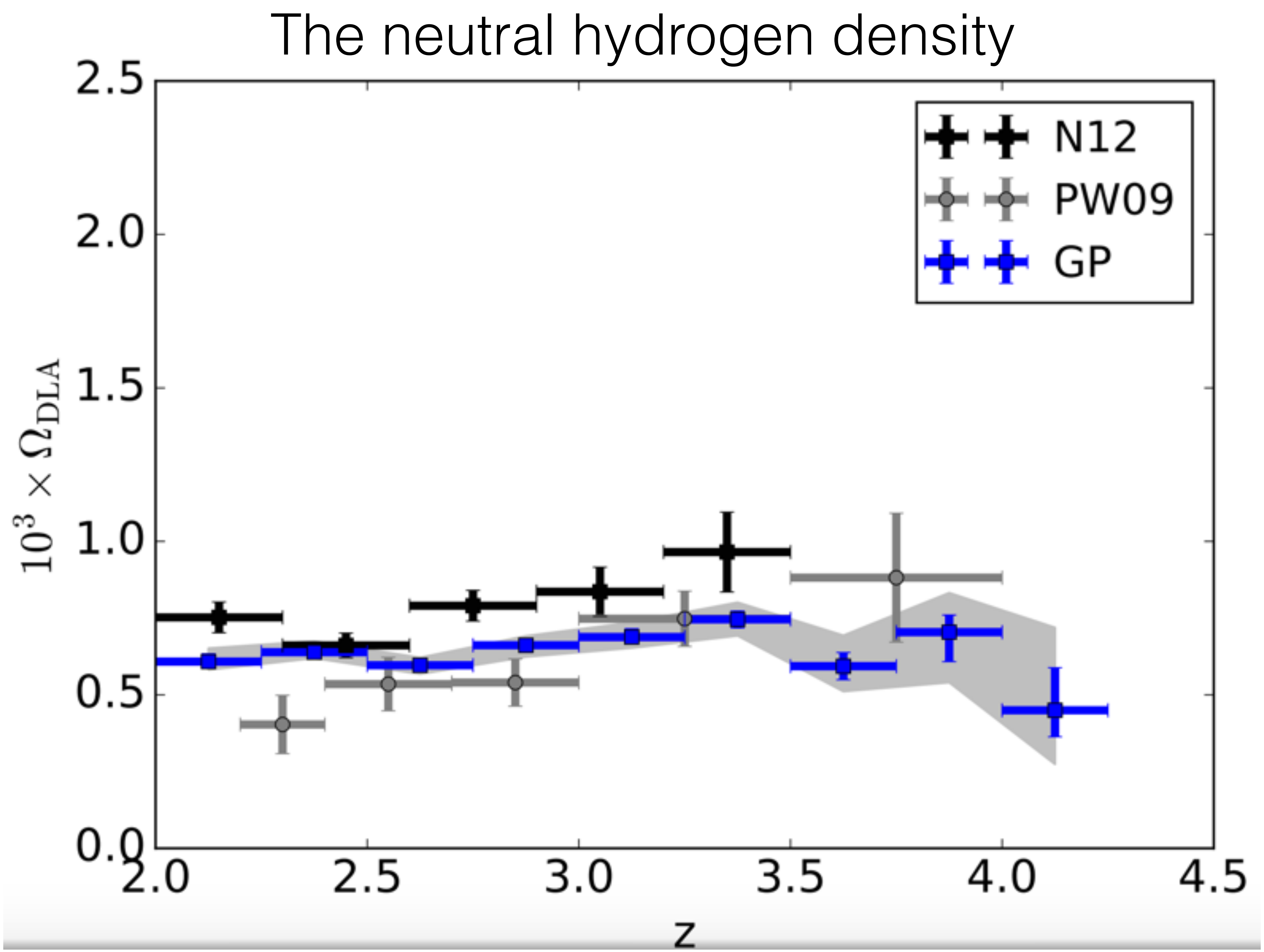
Garnett, Ho, Schneider, ICML 2015
Garnett, Ho, Bird & Schneider, 2016



From >6 person years
to 2 hours on one laptop

Figure 2. We processed over 100,000 Quasar sight-lines from SDSS and ranked them by posterior odds in favor of the DLA model. Ground truth is approximated by previous semi-automated approaches. The area under the ROC curve is 94.1%, and many “false positives” are likely to be newly discovered DLAs.

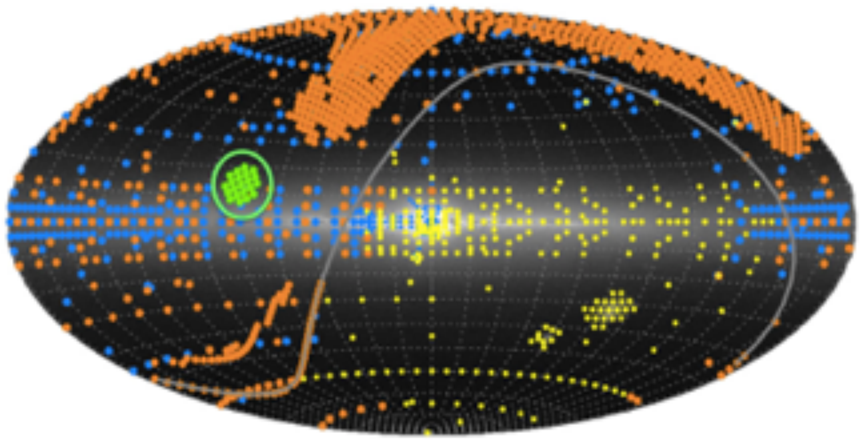
And we can do astrophysics with these objects!



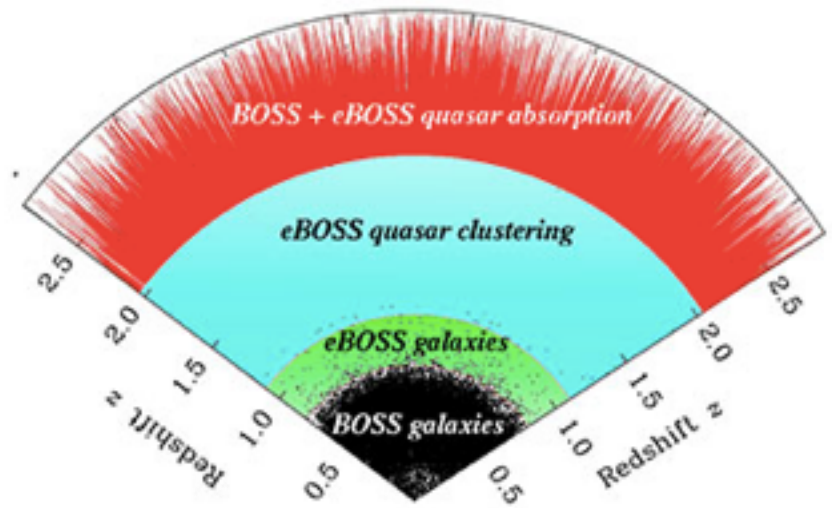
Preliminary results from **Simeon Bird, Roman Garnett & SH.

Looking forward:
Sloan Digital Sky Survey IV

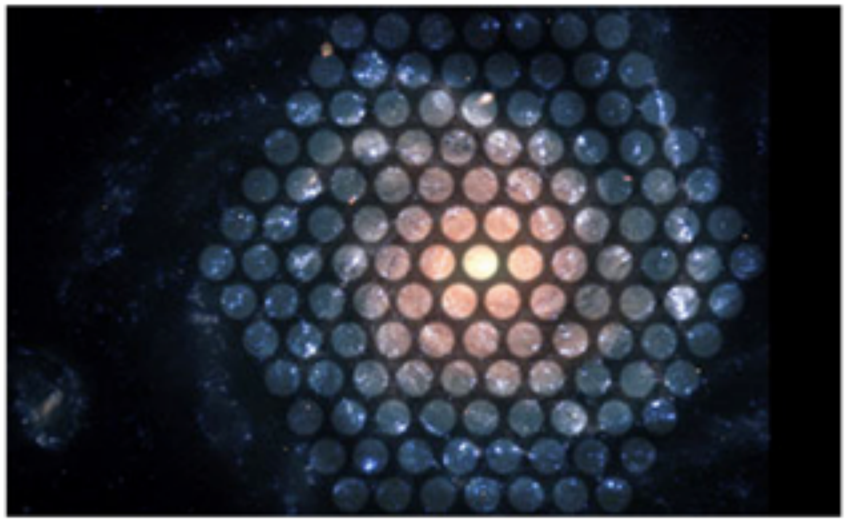
APOGEE-2



eBOSS

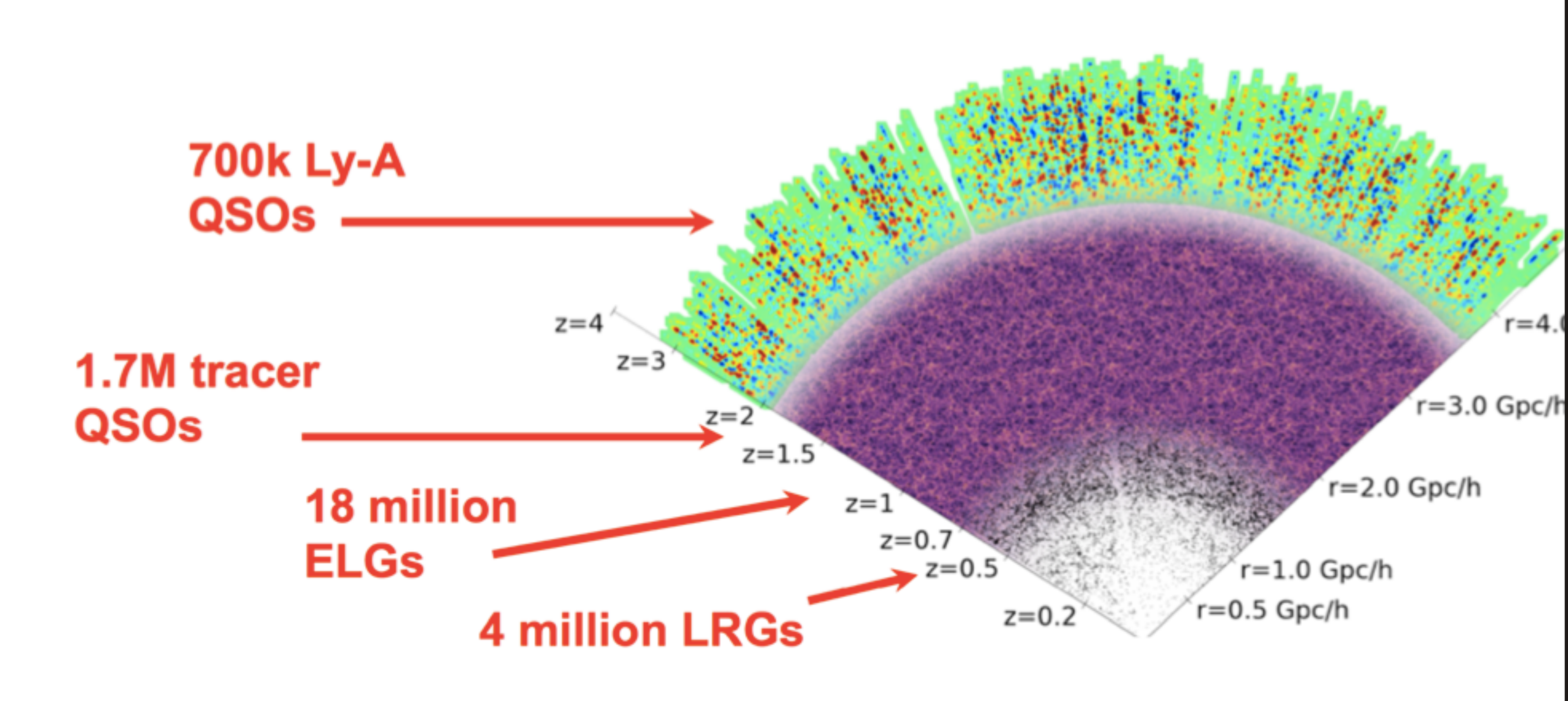


MaNGA



Looking forward

Dark Energy Spectroscopic Instrument (DESI) (2016-)
Each source has over 3000 data points !

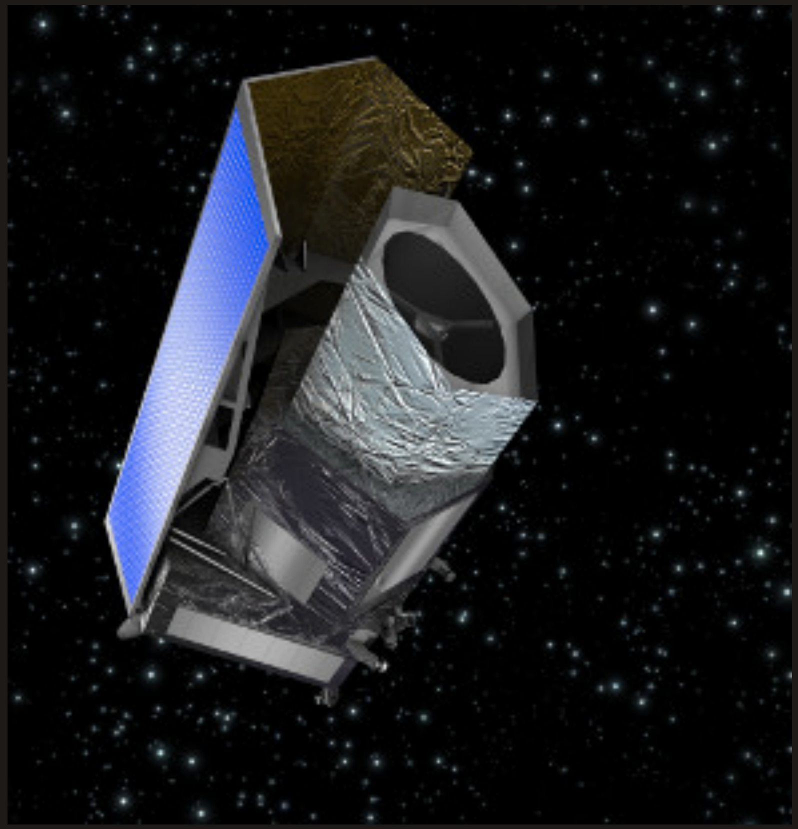


Looking forward

Euclid (ESA led space based mission) 2018-

Imaging: Visible : 30 gal/sq-arcmin
14000 square degrees -> 1.5 billion galaxies

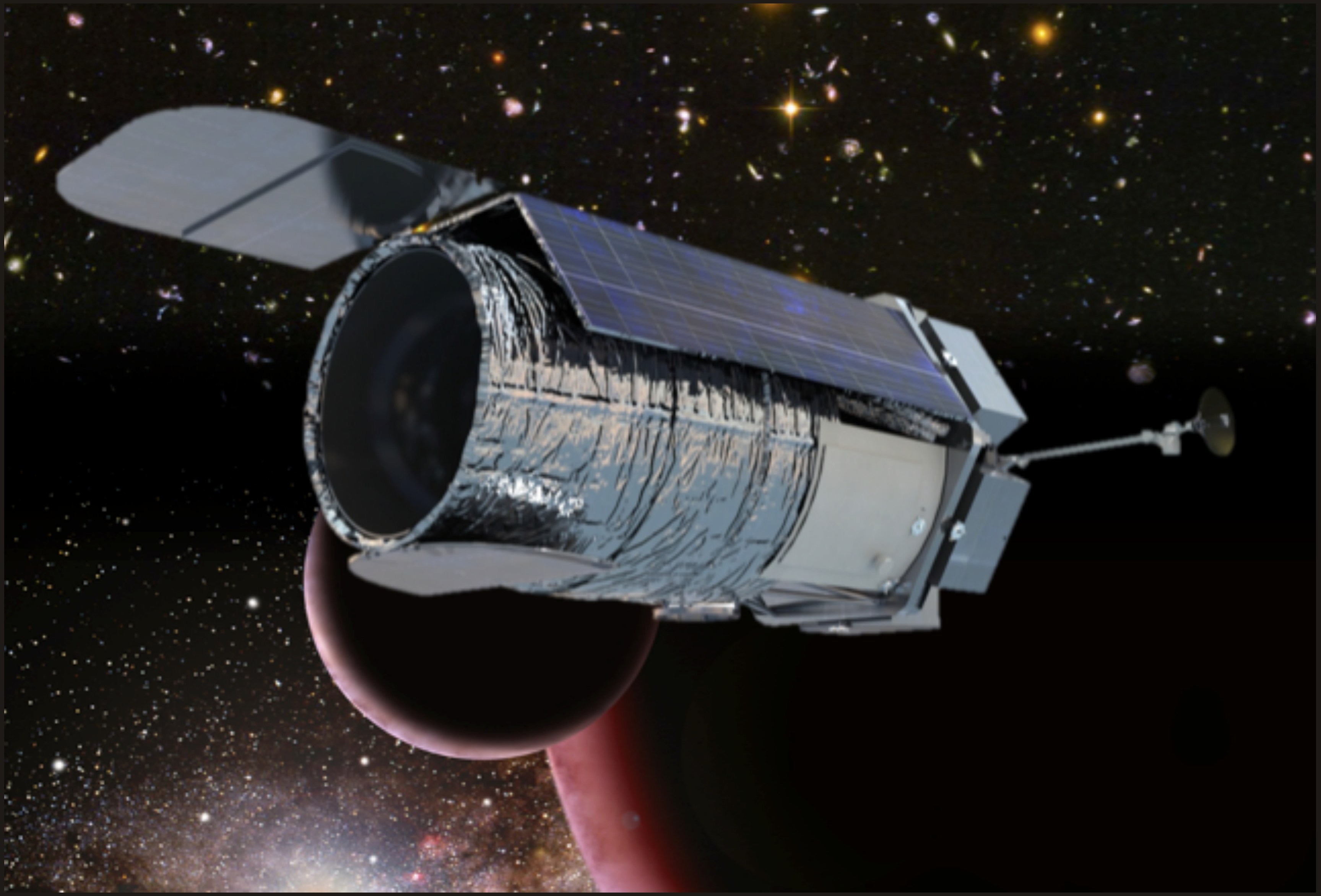
Spectroscopy: 10 million galaxies, with spectroscopy (think X1000)



Aims: Mapping the Geometry of the Dark Universe via Weak Lensing, Baryon Acoustic Oscillations, Redshift Space Distortions

Looking forward

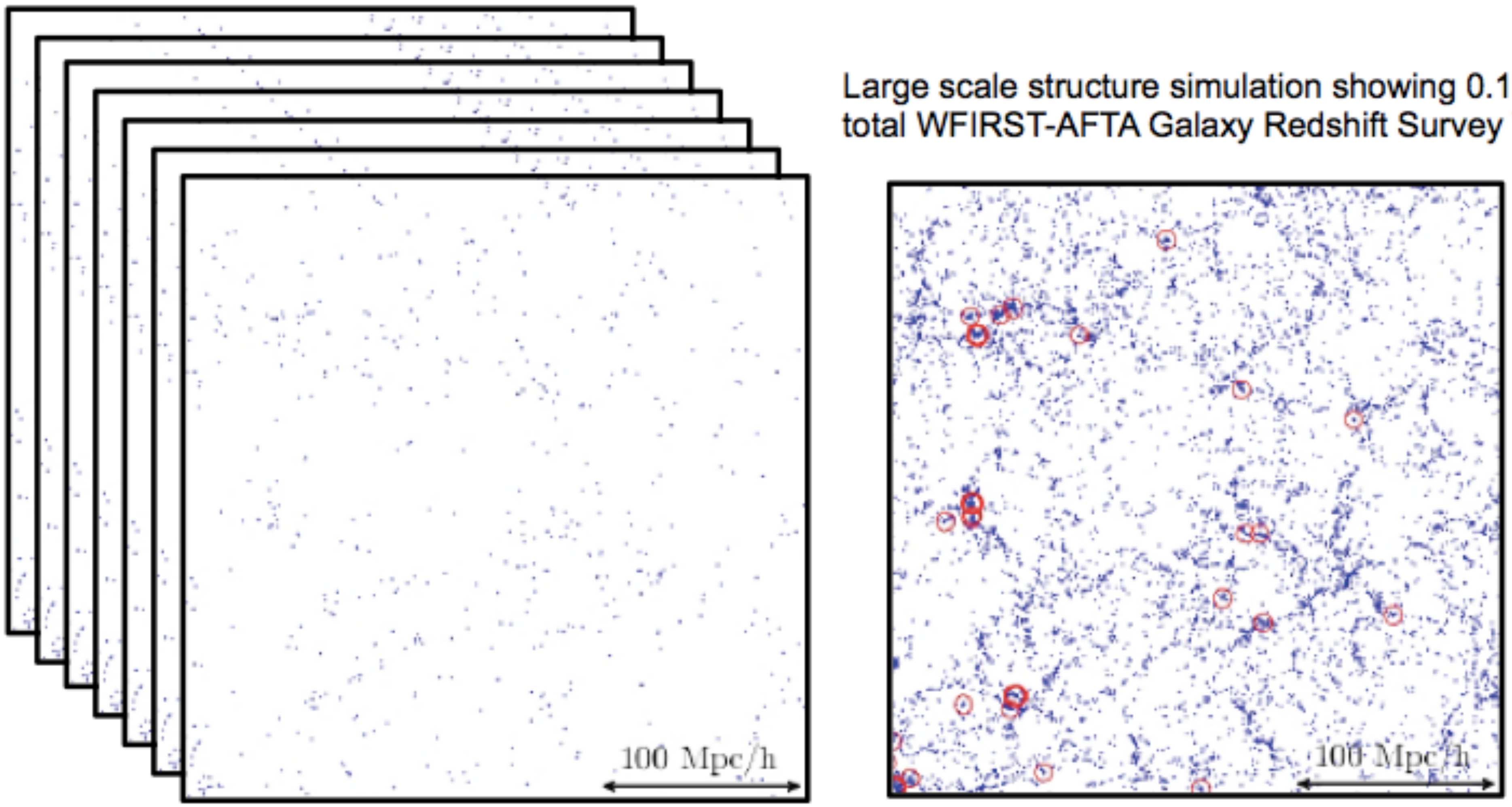
WFIRST-AFTA



Looking forward

Detailed 3D Map of Large Scale Structure at z = 1-2

Large scale structure simulation showing 0.1% of the total WFIRST-AFTA Galaxy Redshift Survey Volume



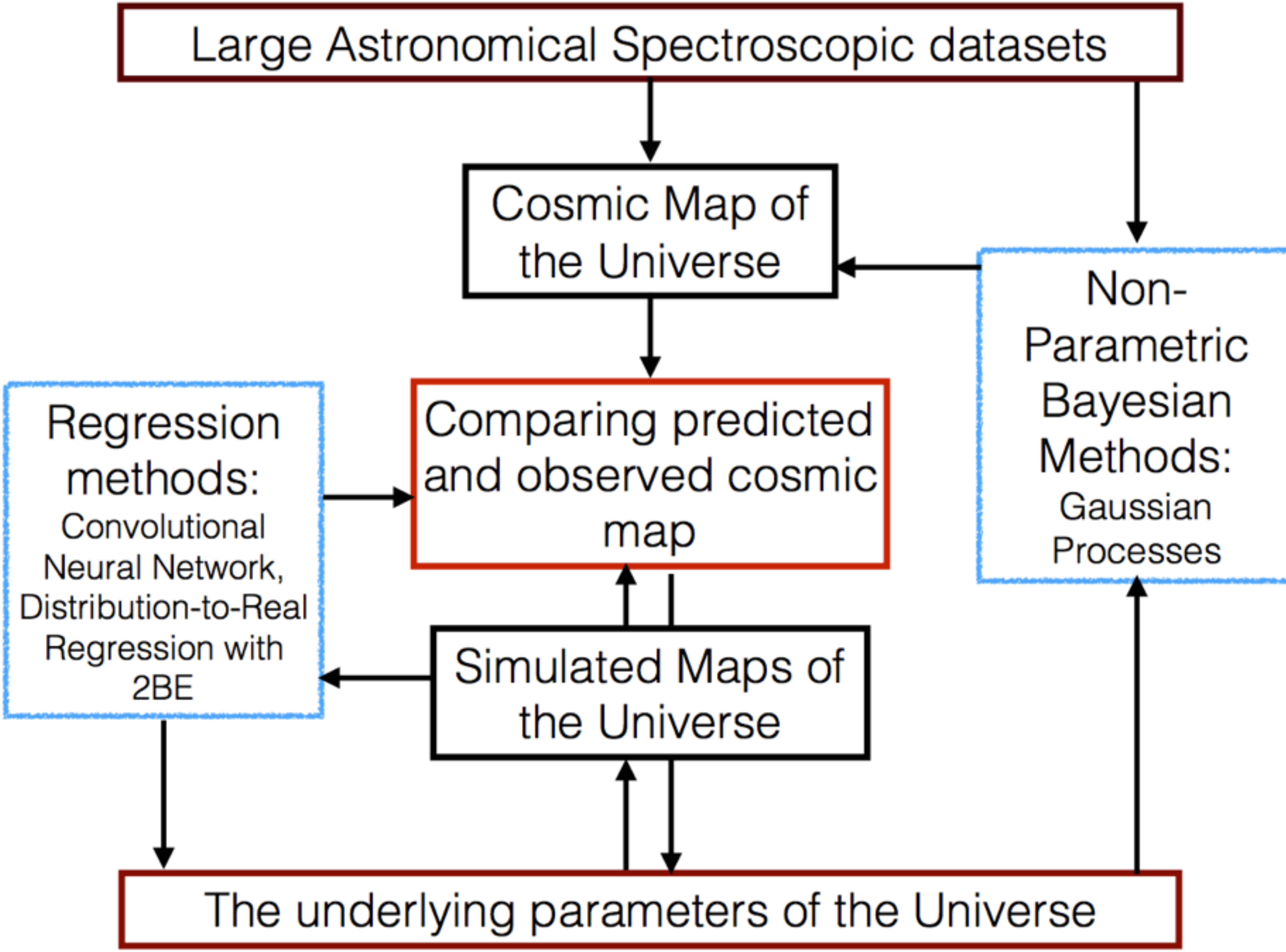
Euclid
15,000 deg^2 @ 1700 gal/deg^2

WFIRST
2,400 deg^2 @ 12,600 gal/deg^2

Large scale structure simulations from 2013 SDT Report – courtesy of Ying Zu
Thin and thick red circles mark clusters with masses exceeding 5 x 10^{13} M_{Sun} and 10^{14} M_{Sun}, respectively

Conclusion

- New tools and methodologies pushes forward alternative summary statistics that actually does something we have not done before.
- This allows us to make quantitative statement of our cosmology AND improve our current cosmological understanding !
- There are many interesting fronts that we can push forward in accelerating science with help from statisticians, computer scientists, applied mathematicians.


Large Astronomical Spectroscopic datasets
Cosmic Map of the Universe
Non-Parametric Bayesian Methods:
Gaussian Processes
Regression methods:
Convolutional Neural Network, Distribution-to-Real Regression with 2BE
Comparing predicted and observed cosmic map
Simulated Maps of the Universe
The underlying parameters of the Universe

Predicting the amount of dark matter with Machine Learning

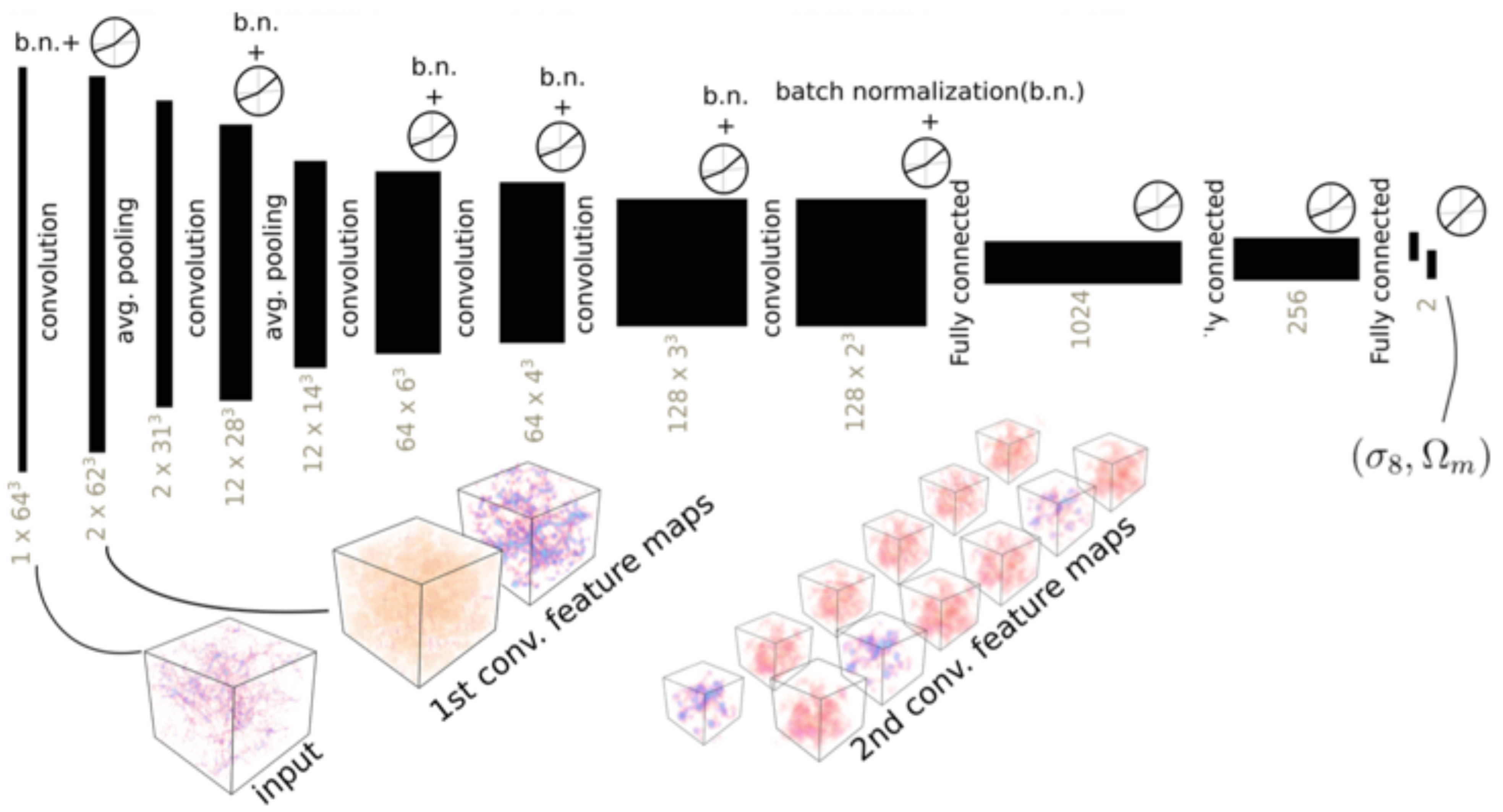
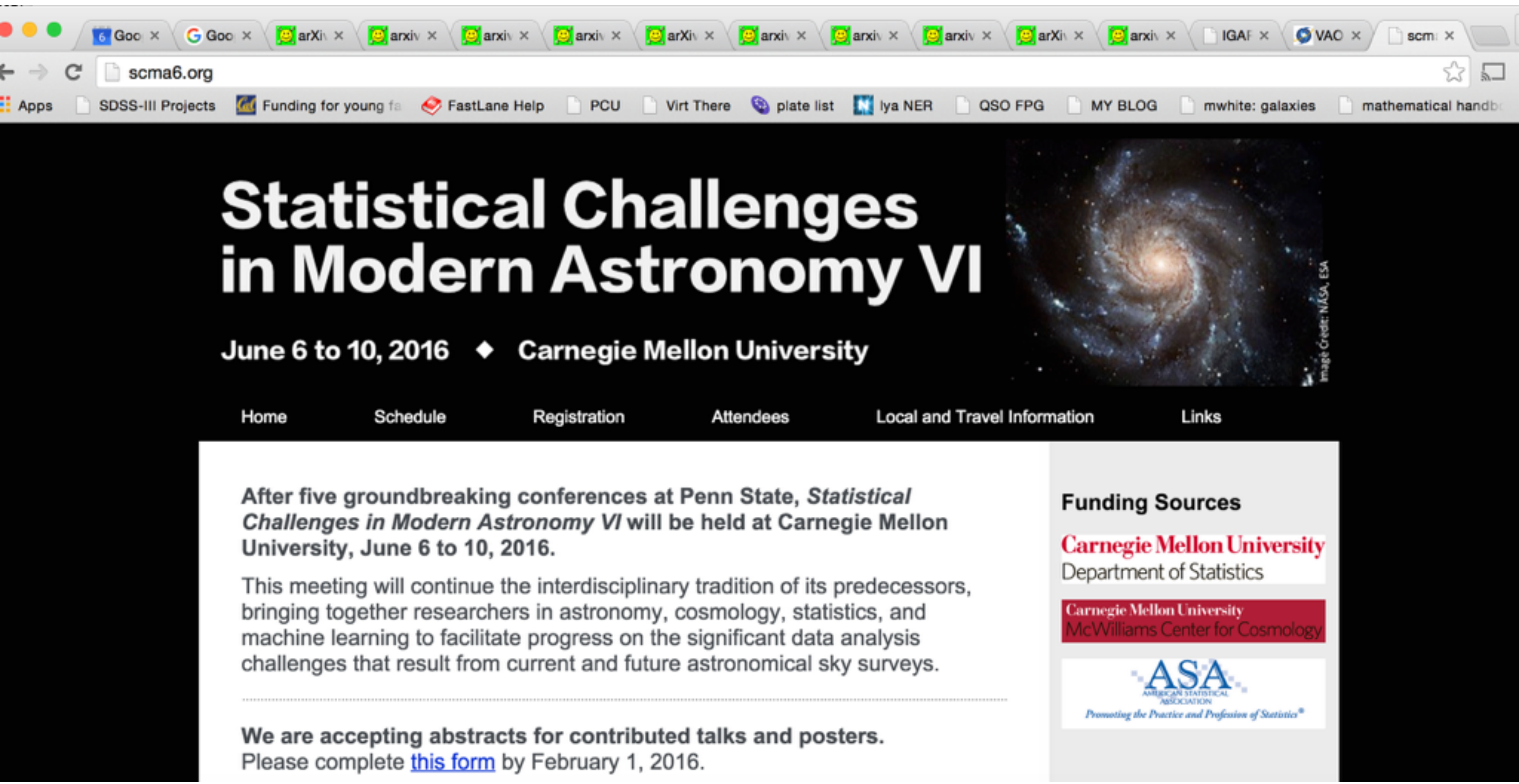


Figure 5: A possible proposed 3D conv-net architecture for predicting the parameters of cosmological simulations. The model has six convolutional and three fully connected layers. The first two convolutional layers are followed by average pooling. The layers can use leaky rectified linear units and batch-normalization (b.n.).

Preliminary work with Siamak Ravanbakhsh, Barnabas Pocozs, Layne Price

Advertisement



Advertisement

Speakers

The program is under development but tentatively includes

keynote talks by Zeljko Ivezic (Univ. of Washington) and Robert Tibshirani (Stanford);

invited talks by astronomers/cosmologists Coryn Bailer-Jones (Max Planck Institute for Astronomy), Rebekah Dawson (UC Berkeley), Laurent Eyer (Geneva), Daniel Foreman-Mackey (Univ. of Washington), Ashish Mahabal (Caltech), Rachel Mandelbaum (CMU), Phil Marshall (SLAC), Brice Menard (Johns Hopkins), Pavlos Protopapas (Harvard), Lucianne Walkowicz (Princeton), Risa Wechsler (Stanford);

and by experts in statistical and machine learning methods: Ethan Anderes (UC Davis), Jeremy Kubica (Google), Ann Lee (CMU), Thomas Lee (UC Davis), James Long (Texas A&M), Jon McAuliffe (UC Berkeley), Xiao Li Meng (Harvard), Bodhisattva Sen (Columbia), Robert Wolpert (Duke).

There will also be interdisciplinary research commentaries from members of the SOC.

Organizers

Co-Chairs: Shirley Ho (CMU) and Chad Schafer (CMU)

How would we compare data and theory in the age of data intensive astronomy research?

Observations | **Theory**

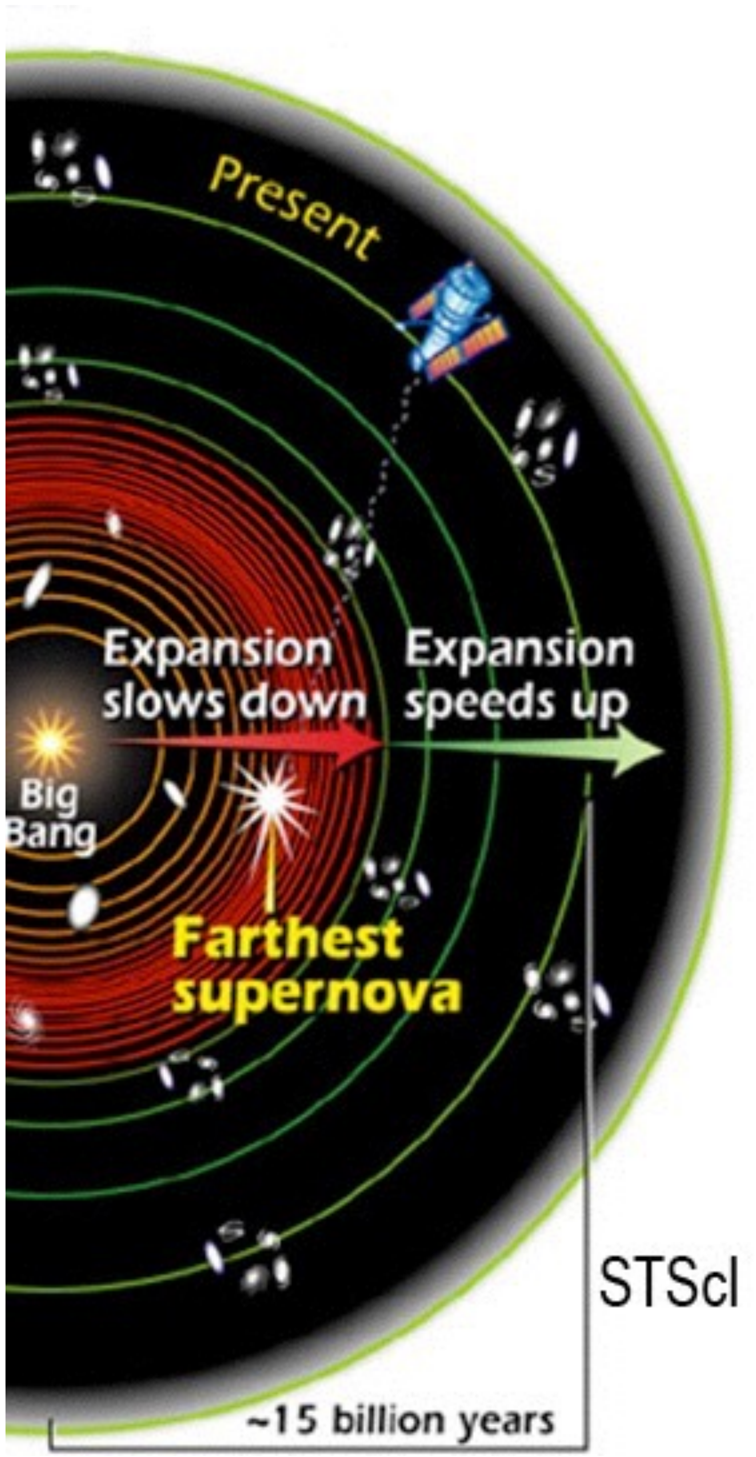
Improve methodologies in comparisons of theory and data.

Employ Tools from statistics, machine learning, computer science which can provide ability to uncover relations in data, isolate unusual classes of objects, find correlations in high dimensional space.

Establish new practices that will accelerate scientific discovery. Develop common code-base, new workflow that allow scientists to do what they do best: Transform scientific data and theory into discovery!

Mapping Our History

Like tree rings mapping out
the climate history of Earth,
cosmic distances expand with time,
slowing down and speeding up,
mapping out cosmic history,
which is affected by things like
Dark Energy, Dark Matter...

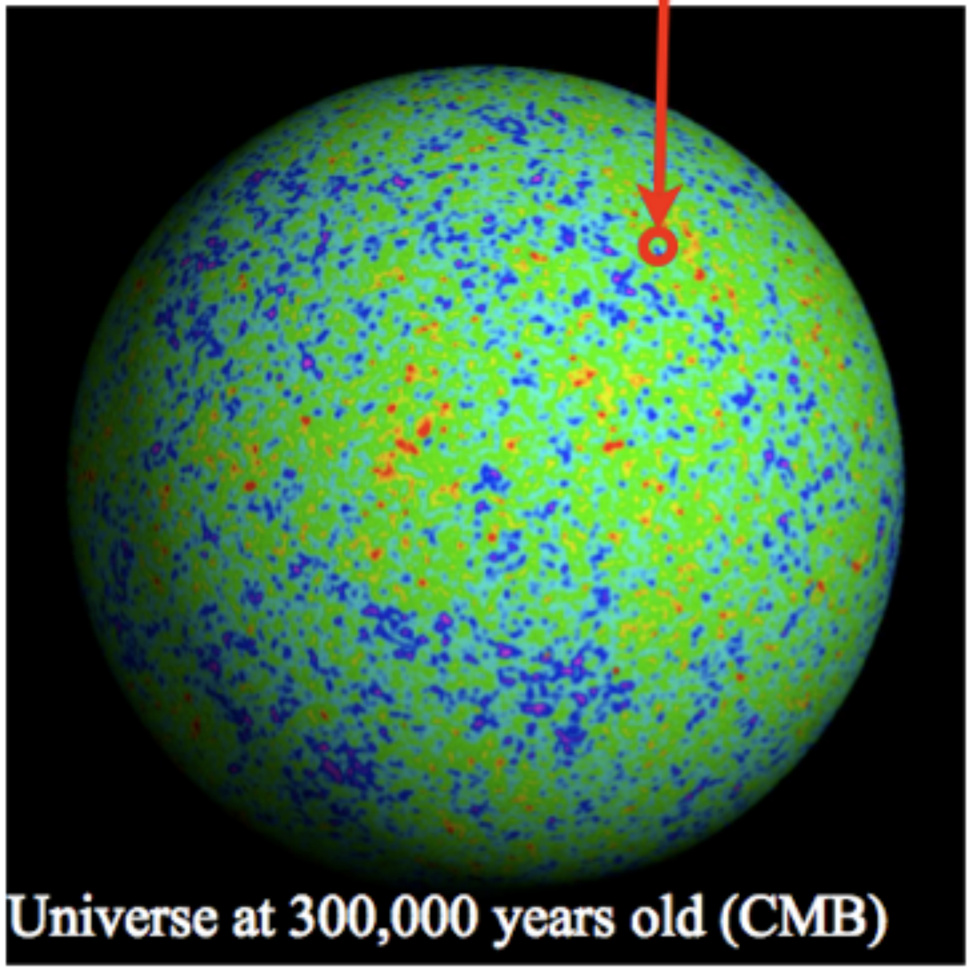


Our tree-rings:
"Baryon Acoustic Oscillations"

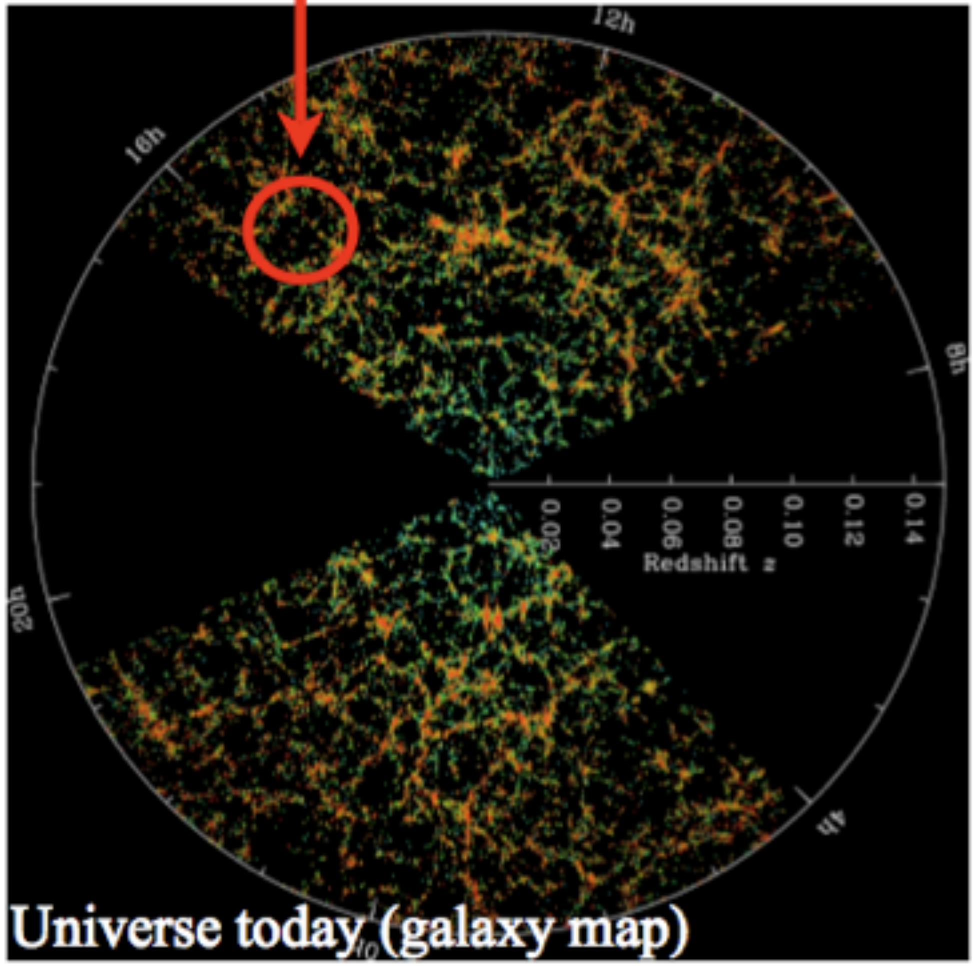
What are baryon acoustic oscillations (BAO)?

These fluctuations of 1 part in 10^5 gravitationally grow into...

...these ~unity fluctuations today



Universe at 300,000 years old (CMB)



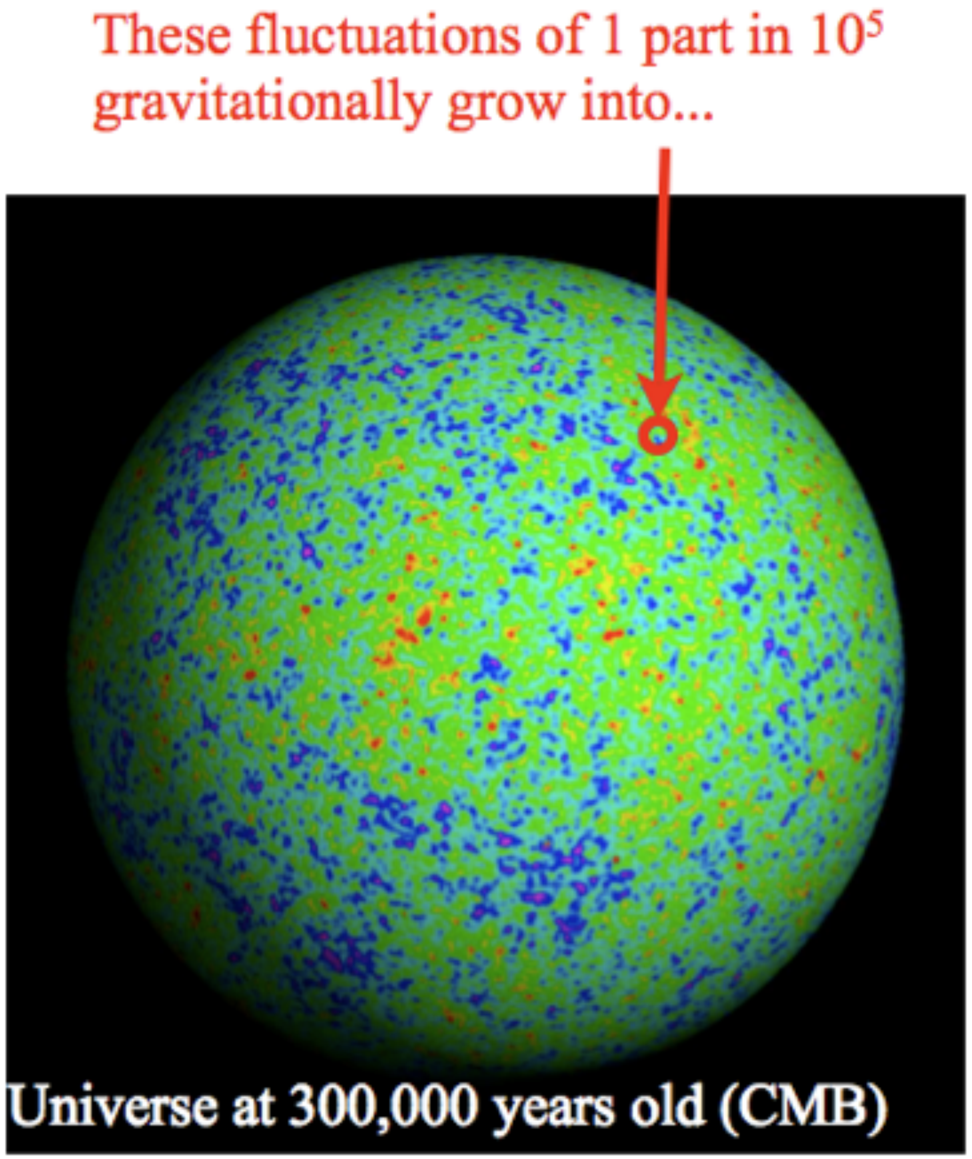
Universe today (galaxy map)

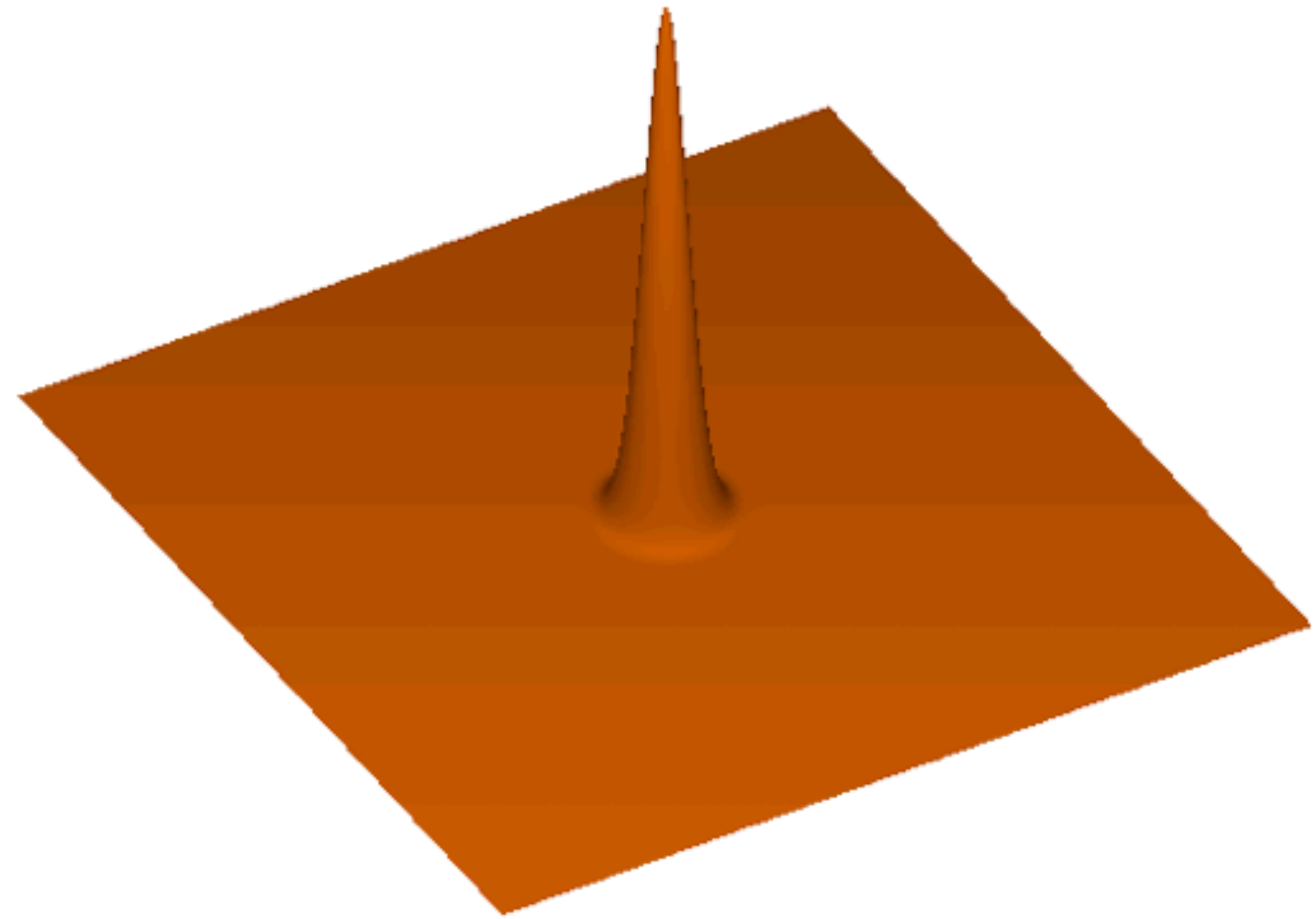
This sound wave can be used as a "standard ruler"
Dark energy changes this apparent ruler size

THE MCWILLIAMS
Center for Cosmology

Our tree-rings:
“Baryon Acoustic Oscillations”

What are baryon acoustic oscillations (BAO)?



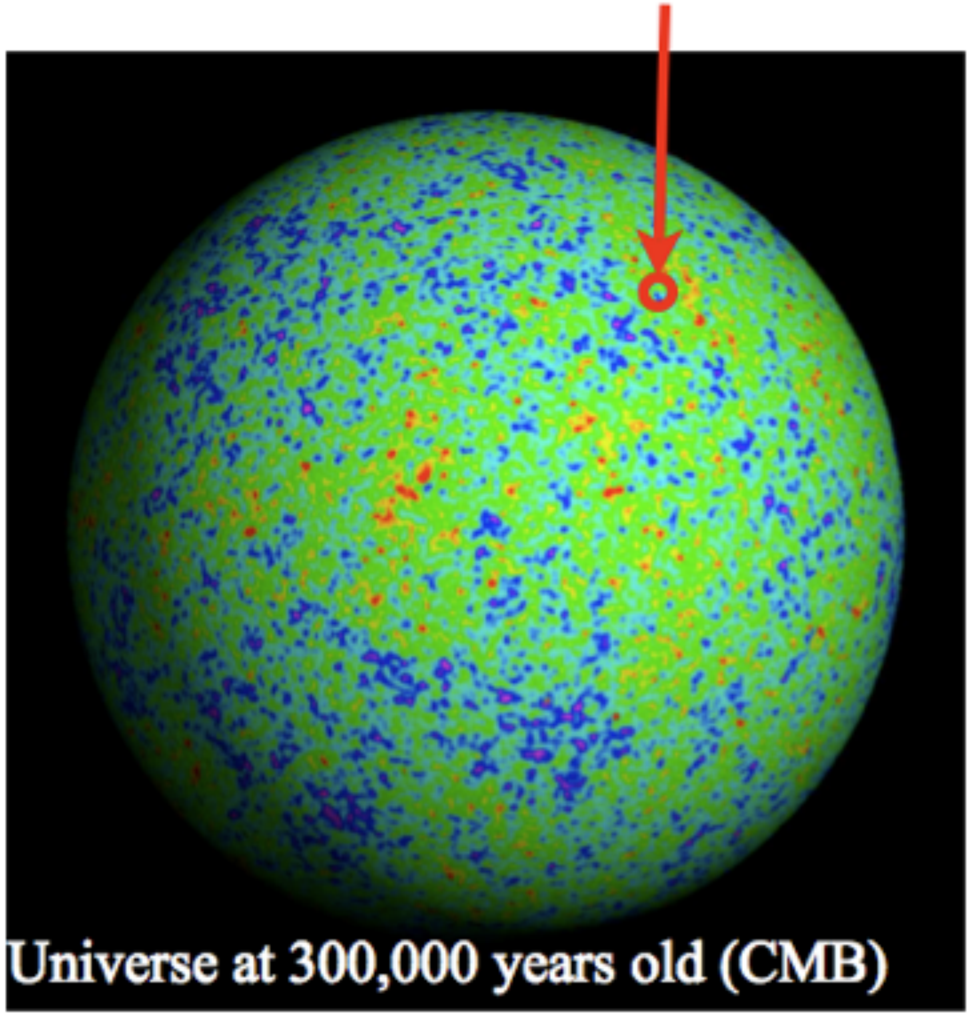


This sound wave can be used as a “standard ruler”
Dark energy changes this apparent ruler size

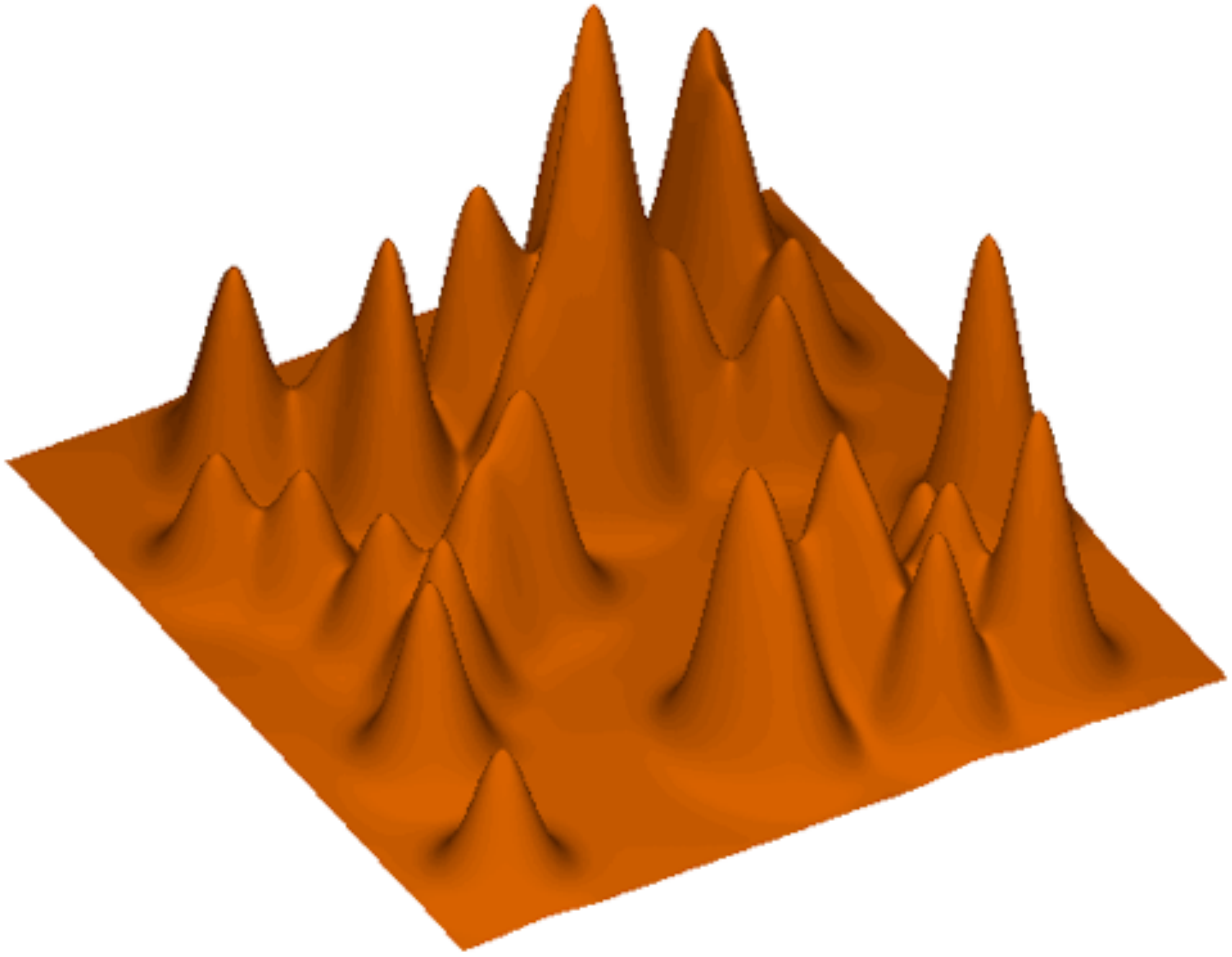
Our tree-rings: “Baryon Acoustic Oscillations”

What are baryon acoustic oscillations (BAO)?

These fluctuations of 1 part in 10^5 gravitationally grow into...



...these ~unity fluctuations today



This sound wave can be used as a “standard ruler”
Dark energy changes this apparent ruler size

THE MCWILLIAMS
Center for Cosmology

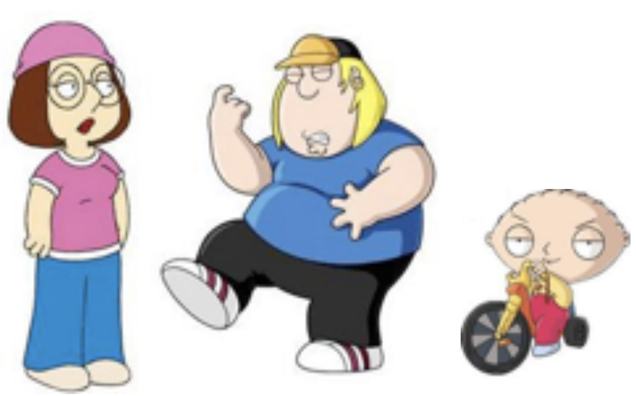
How do we detect Baryon Acoustic Oscillations?

THE MCWILLIAMS Center for Cosmology

To detect Baryon Acoustic Oscillations, calculate the Correlation Function

$$\xi_f(r) = < \delta_f(\hat{x})\delta_f(\hat{x}+\hat{r}) >$$

HOME/nearby school



20 miles

Office

20 miles

Office

Detecting
Baryon Acoustic Oscillations!

THE MCWILLIAMS
Center for Cosmology

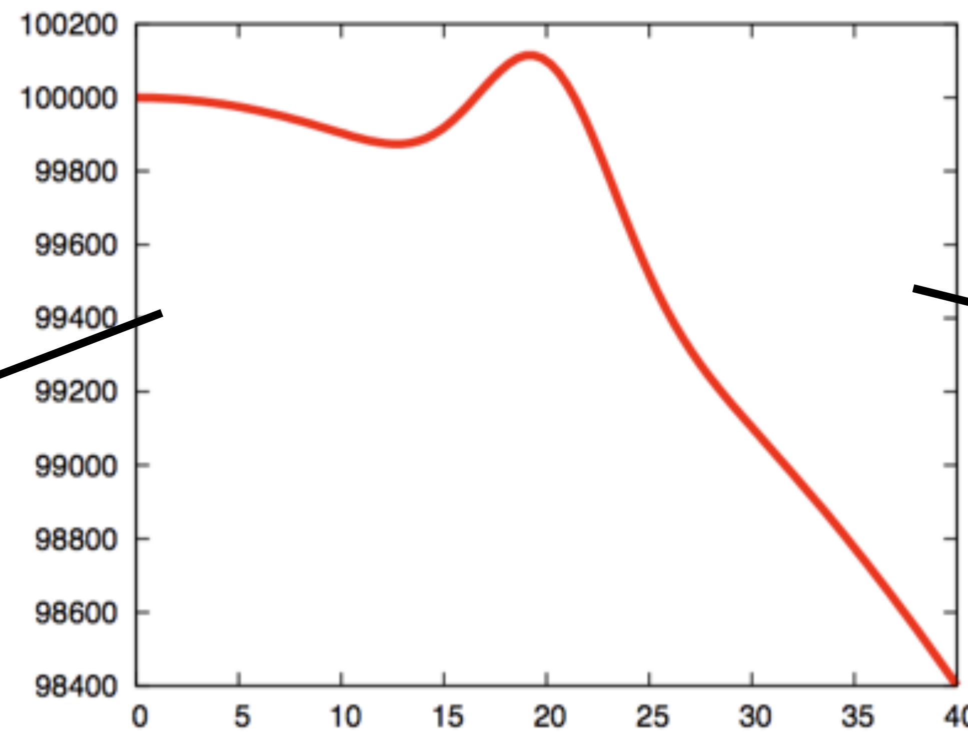
To detect Baryon Acoustic Oscillations, calculate the Correlation Function

$$\xi_f(r) =< \delta_f(\hat{x})\delta_f(\hat{x}+\hat{r}) >$$

The Correlation Function of the population (during the day) is:

Office

20 miles



20 miles

Office

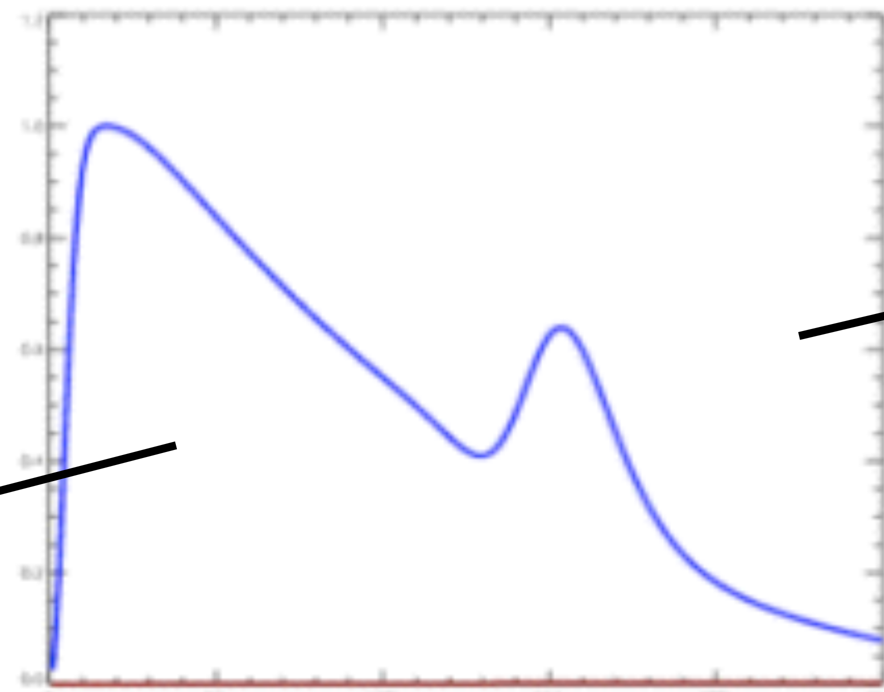
The bump will be at 20 miles!

Detecting
Baryon Acoustic Oscillations!

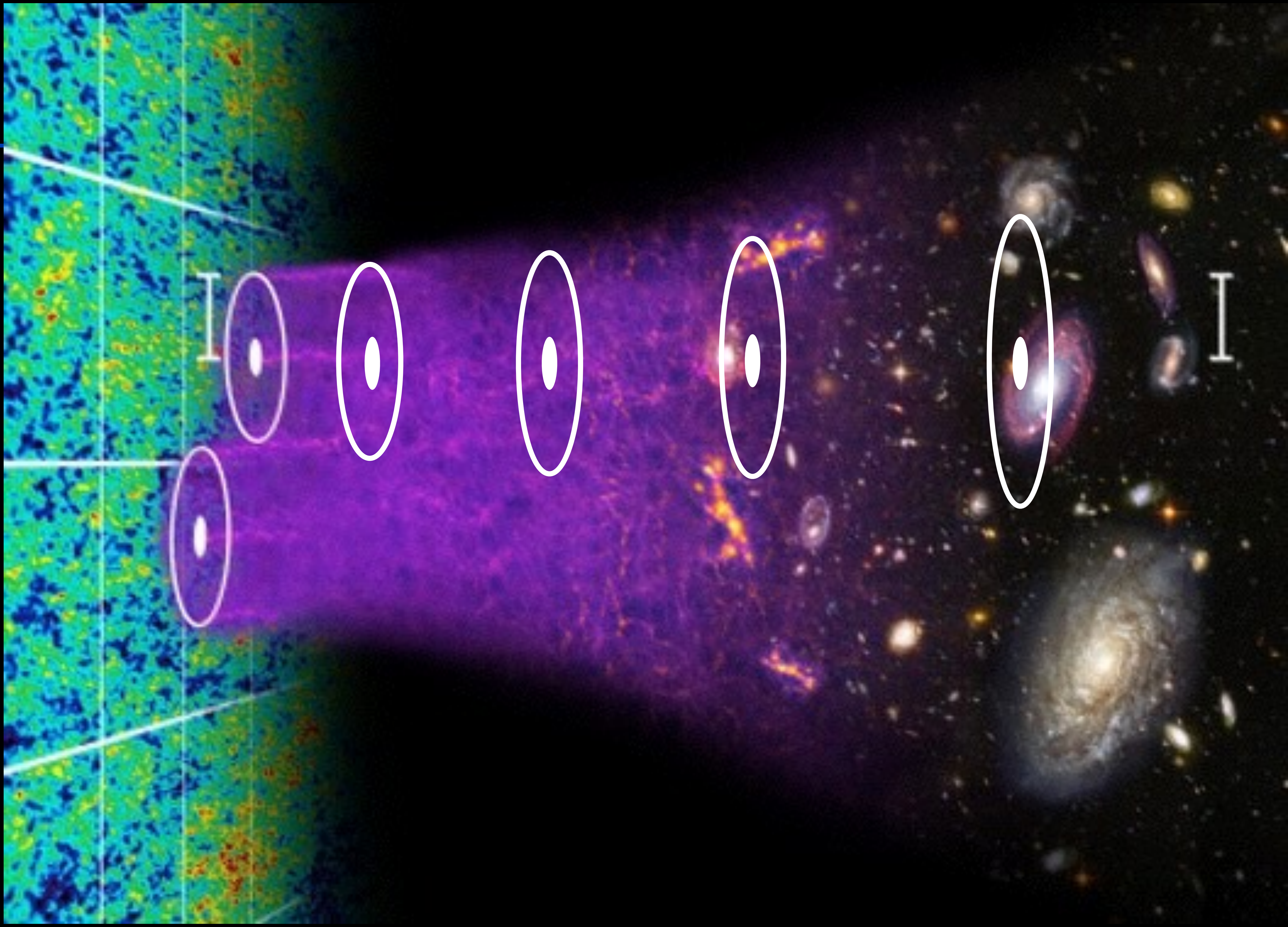
To detect Baryon Acoustic Oscillations, calculate the Correlation Function

$$\xi_f(r) =< \delta_f(\hat{x})\delta_f(\hat{x}+\hat{r}) >$$

The Correlation Function of galaxies is:

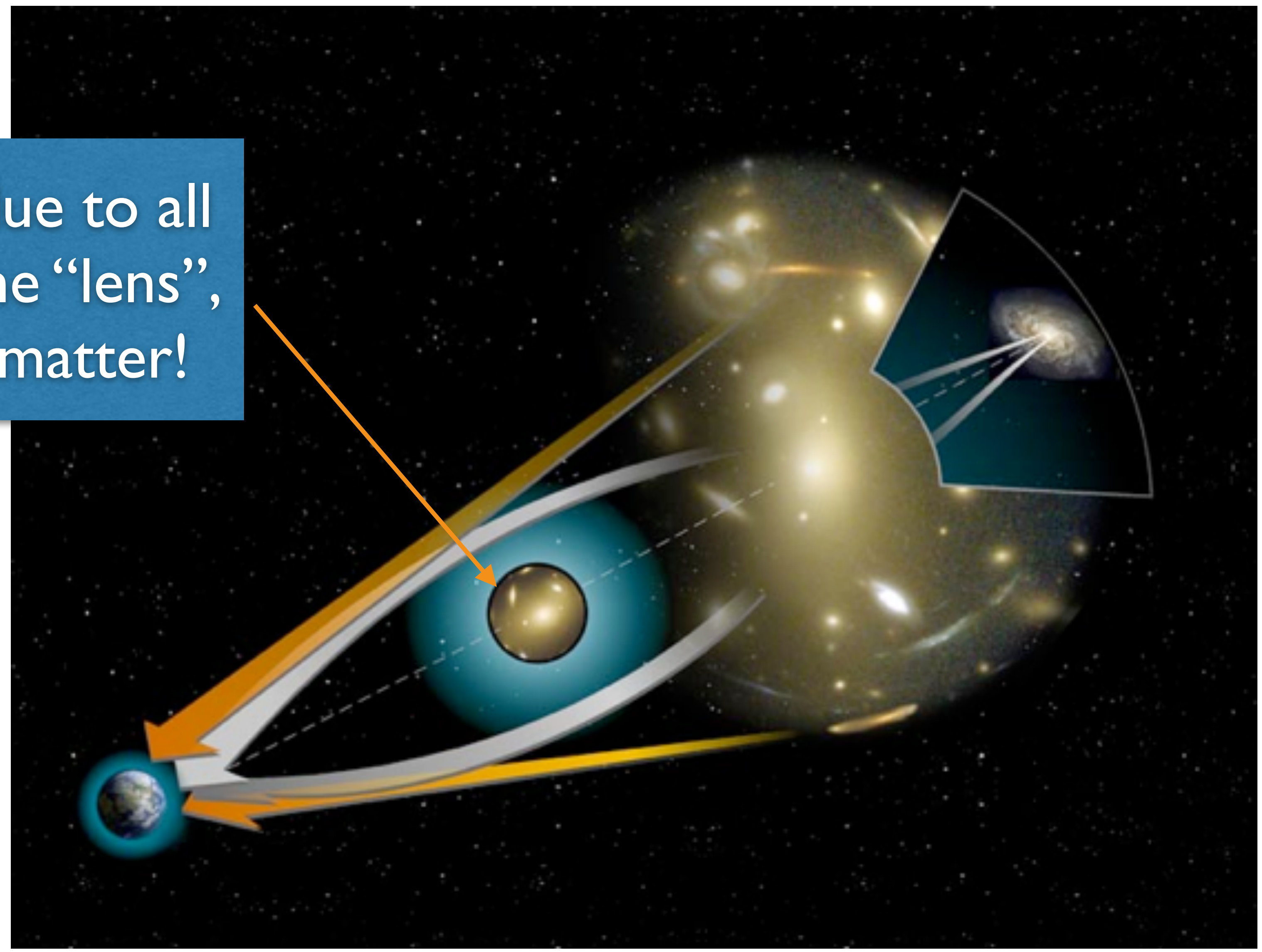


The bump will be at about 470,000,000 light years.
The effect is highly exaggerated for visual effects here,
It is a 1% effect if we survey at least 470,000,000
light years of the Universe.



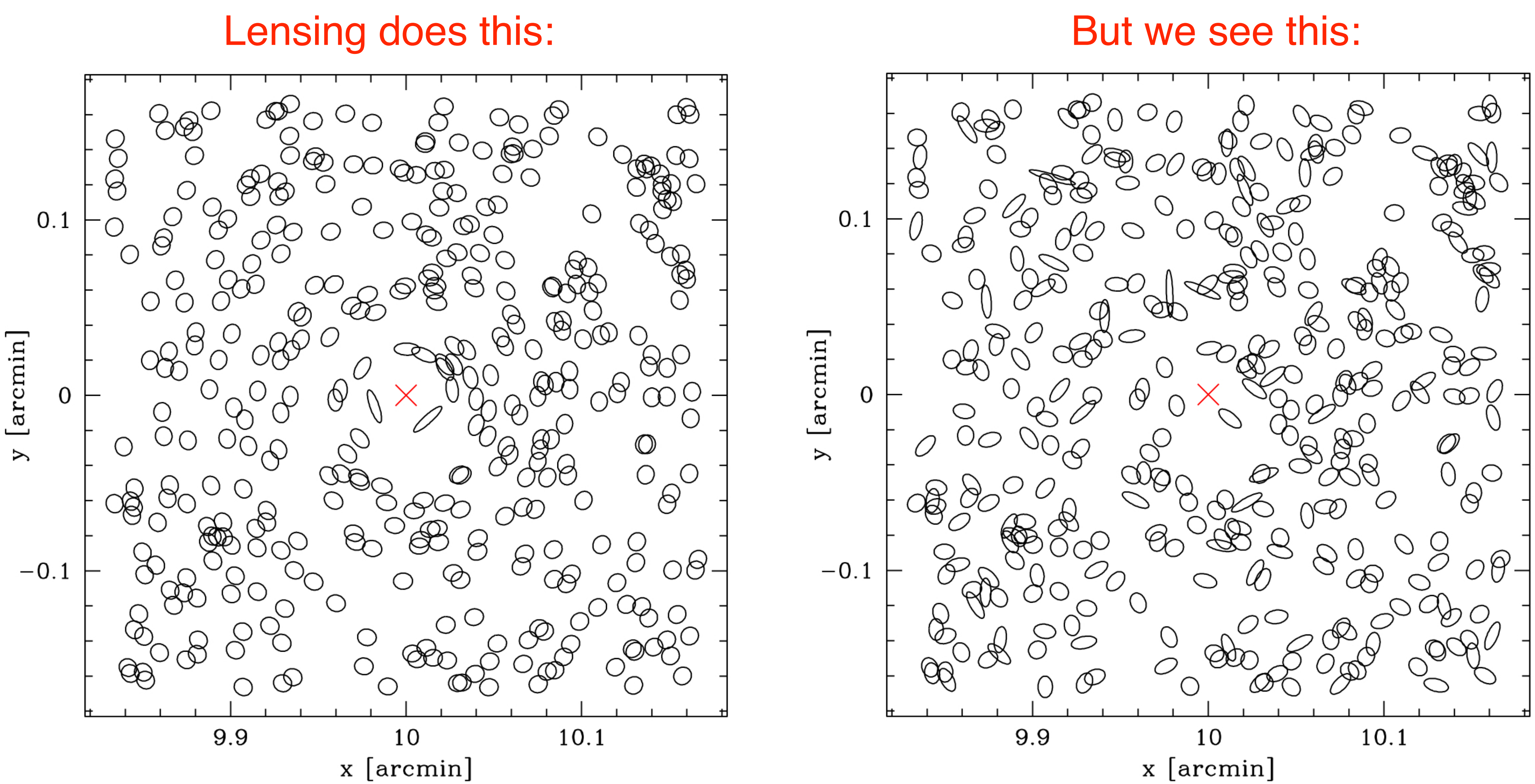
Gravitational lensing: mass deflects light rays!

Lensing is due to all matter in the “lens”, even dark matter!



In the extreme case: multiple images = very striking!
(“strong lensing”, very rare)

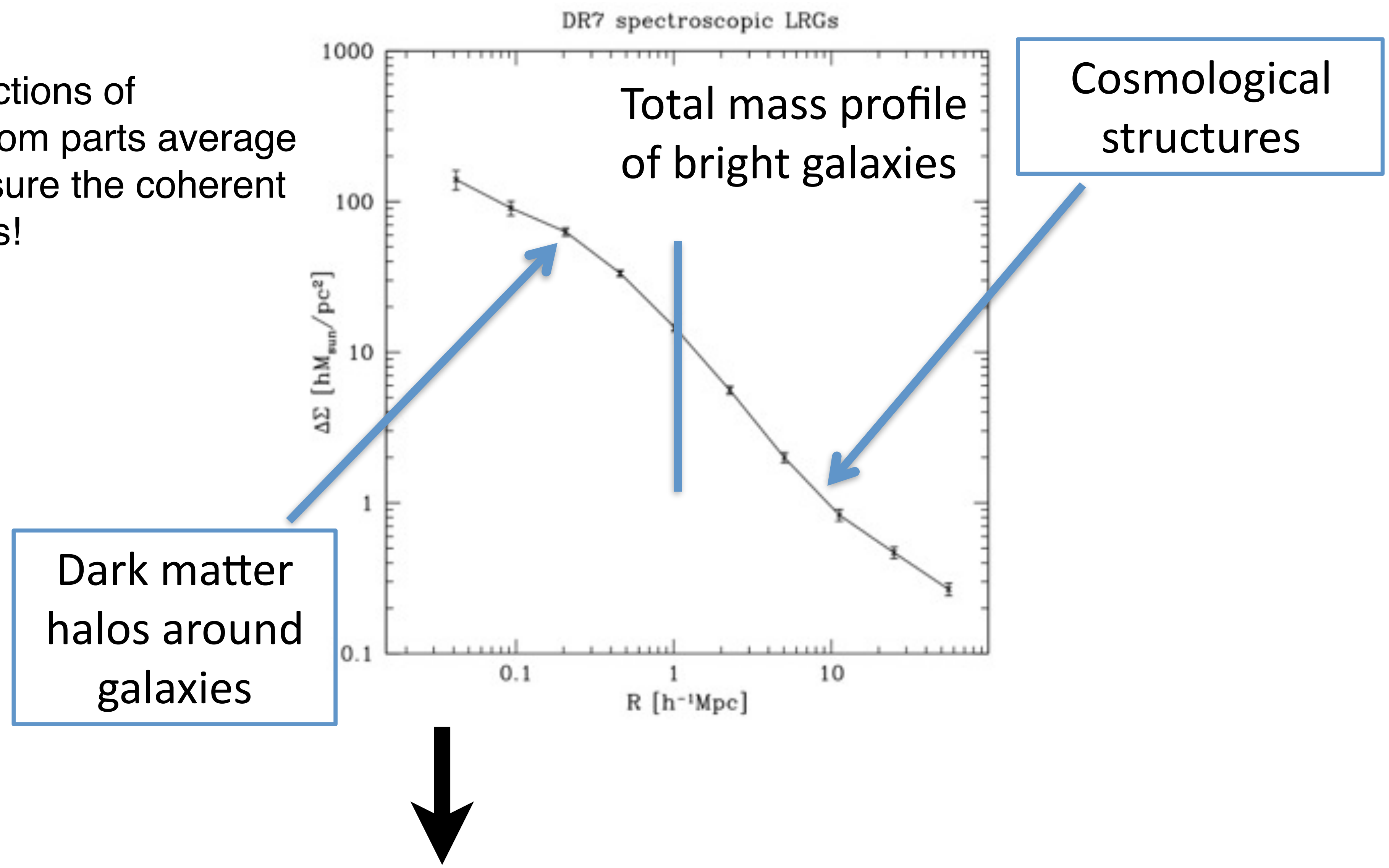
Gravitational lensing: mass deflects light rays!



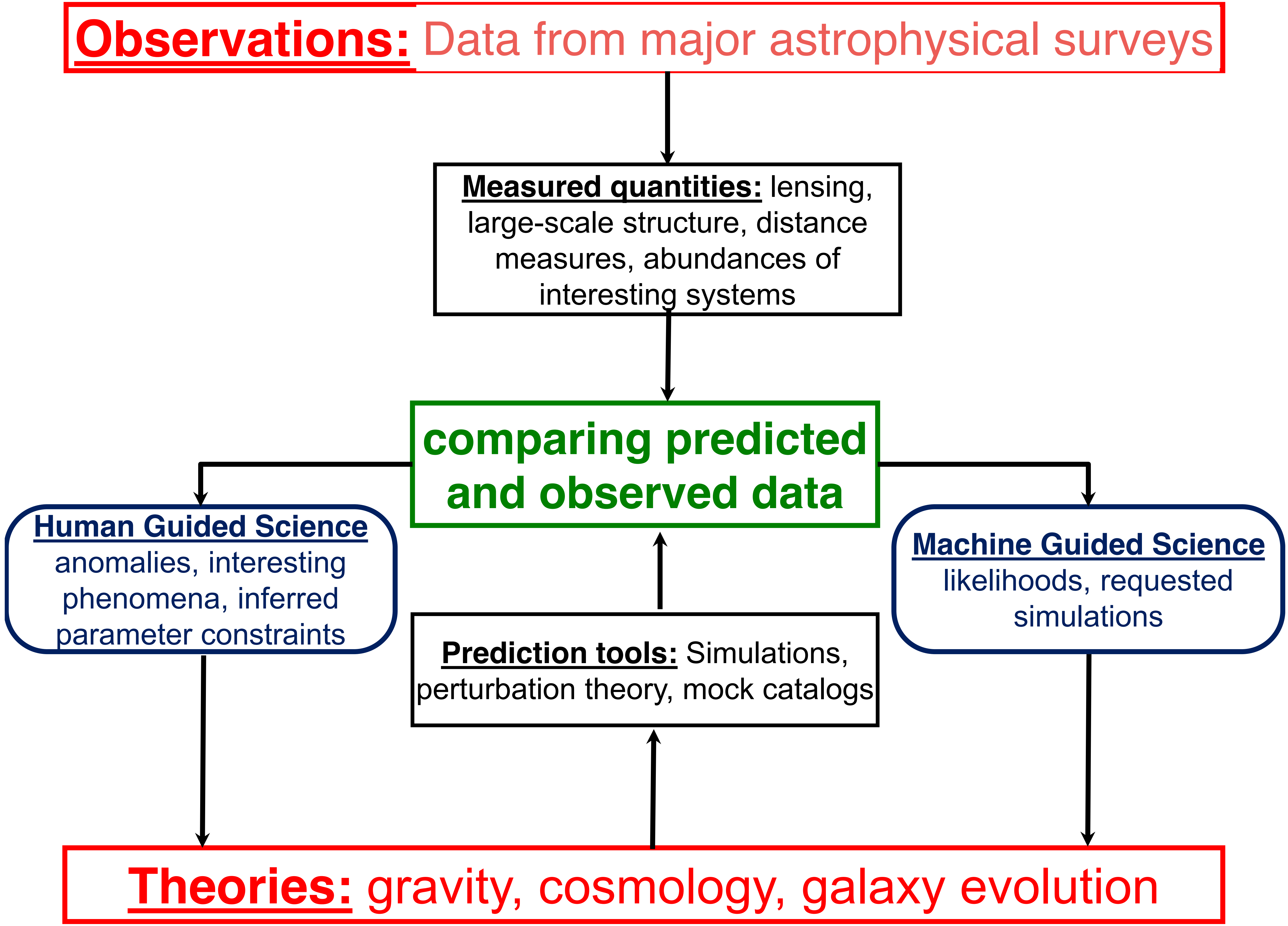
Weak lensing (galaxy shape distortions due to lensing) happens *everywhere*, but they are small = hard to measure

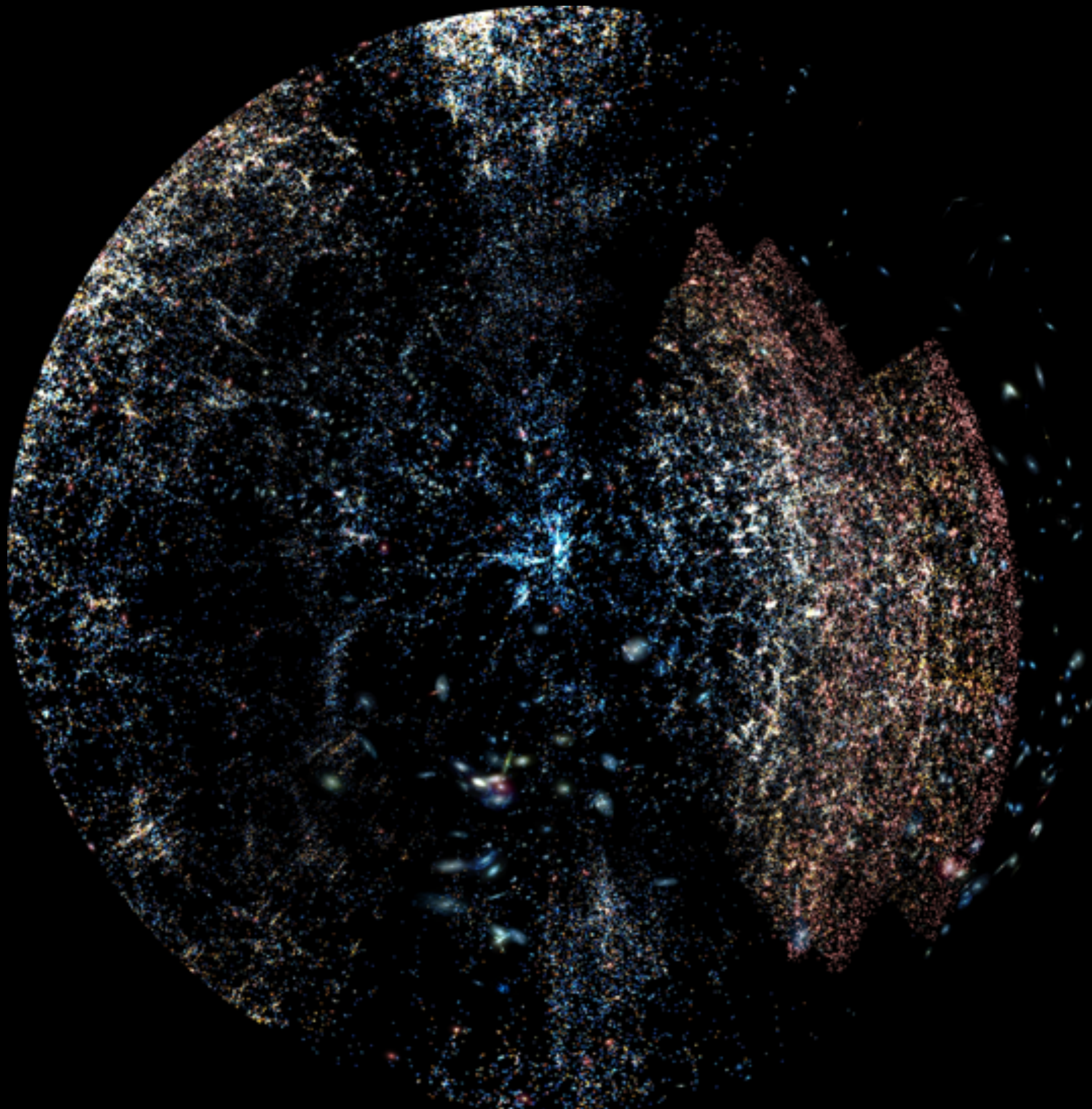
Gravitational lensing: mass deflects light rays!

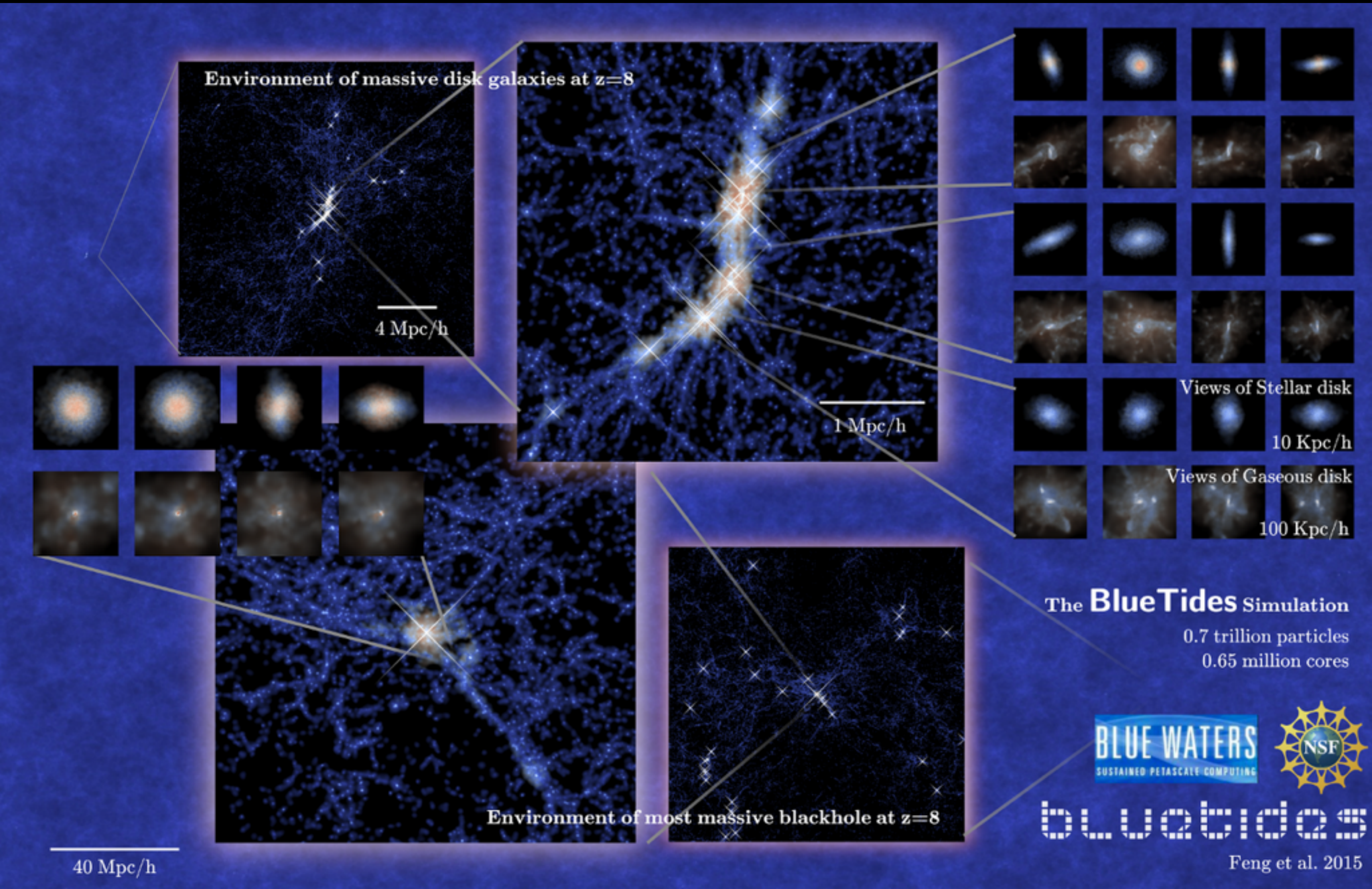
Measure correlation functions of galaxy shapes: the random parts average to zero, so we can measure the coherent lensing shape distortions!



Learn about galaxy-dark matter connection, dark energy and theory of gravity on cosmological scales!


Observations: Data from major astrophysical surveys
Measured quantities: lensing, large-scale structure, distance measures, abundances of interesting systems
comparing predicted and observed data
Human Guided Science
anomalies, interesting phenomena, inferred parameter constraints
Machine Guided Science
likelihoods, requested simulations
Prediction tools: Simulations, perturbation theory, mock catalogs
Theories: gravity, cosmology, galaxy evolution




Environment of massive disk galaxies at z=8
4 Mpc/h
1 Mpc/h
Views of Stellar disk
10 Kpc/h
Views of Gaseous disk
100 Kpc/h
The BlueTides Simulation
0.7 trillion particles
0.65 million cores
BLUE WATERS
SUSTAINED PETASCALE COMPUTING
NSF
bluetides
Environment of most massive blackhole at z=8
40 Mpc/h
Feng et al. 2015