

# Non-linear bayesian inference of cosmic fields in SDSS3 and 2M++ and their application to Nearby cosmology

Guilhem Lavaux (IAP/CNRS)  
and Aquila Consortium

eBOSS meeting 2018



From pictures... to physics of Universe at large



From pictures... to physics of Universe at large

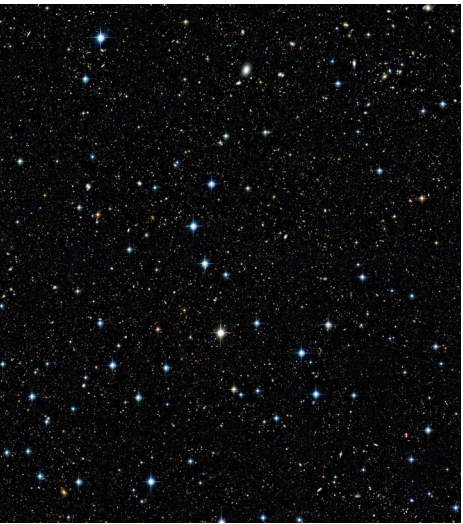
In a handful of model parameters



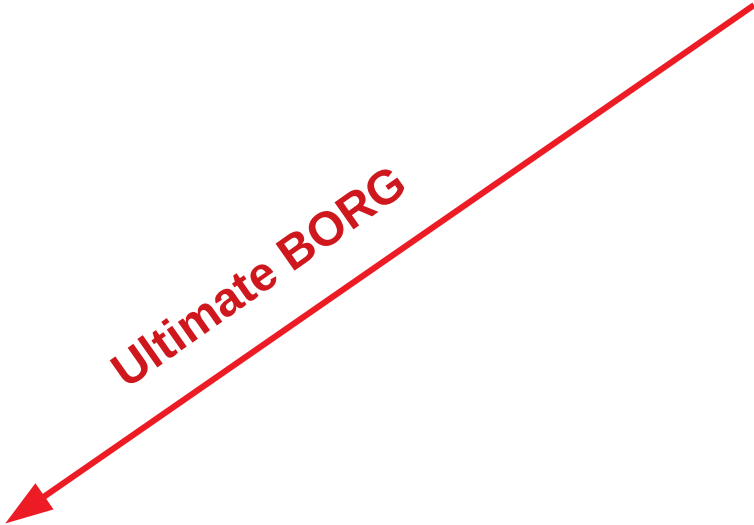
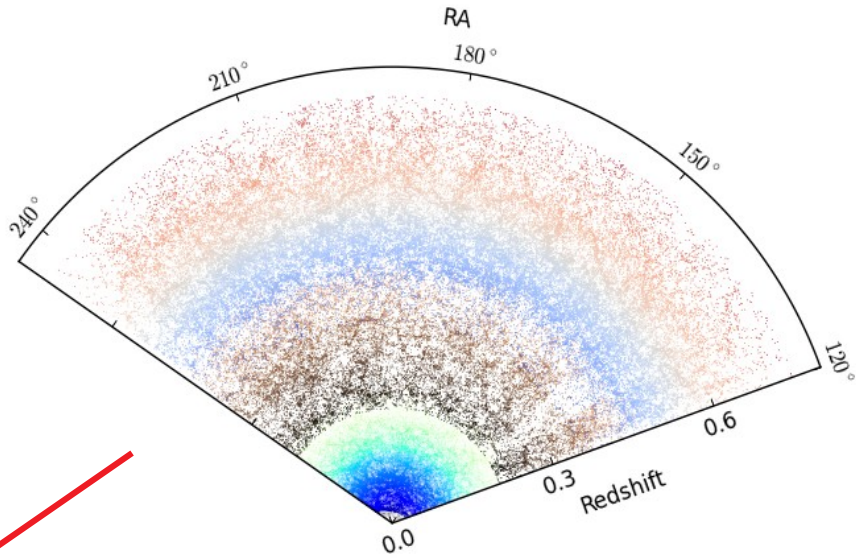
Ultimately: we want to fit a model to this kind of picture, and the pixel by pixel spectrum



Ultimately: we want to fit a model to this kind of picture, and the pixel by pixel spectrum  
That's very challenging, probably impossible → we reduce those datasets



Reduce



Massage, adjust,  
compute correlations



?

Do something



Fit model to  
even further  
reduced data

$\Omega$

Cosmological parameters  
Global cosmological properties

# Outline



**The statistical framework**



**The 2M++ compilation**

**(presentation, clusters, velocity fields, applications)**



**SDSS3 BOSS**

**(more modeling challenges, density field)**



**Conclusion**

# From theory to observations...

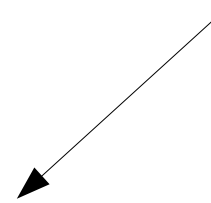
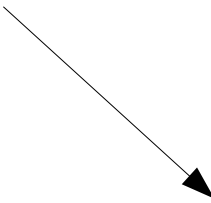
## Model

- Perfect
- Complete description
- Full knowledge of physics
- Did I say perfect ?



## Observations

- Great but messy
- We do not understand the physics
- Systematics not fully known
- Good attempt by observers to seemingly make our life easier end up bad



Various hacking to make sense of data





# From theory to observations...

## Model

- Perfect
- Complete description
- Full knowledge of physics
- Did I say perfect ?



## Observations

- Great but messy
- We do not understand the physics
- Systematics not fully known
- Good attempt by observers to seemingly make our life easier end up bad

## BORG3

Still far too perfect though... (see later)



... or IMNN

Charnock et al. (2018, PRD)

# The BORG3 inference framework

$$\pi(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k |\hat{\delta}_k|^2 / P_k\right)$$

**Initial conditions**

# The BORG3 inference framework

$$\pi(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k |\hat{\delta}_k|^2 / P_k\right)$$

**Initial conditions**

**Total evolved matter density**  $\rho_m = \mathcal{F}[\delta]$

# The BORG3 inference framework

$$\pi(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k |\hat{\delta}_k|^2 / P_k\right)$$

**Initial conditions**

**Total evolved matter density**

$$\rho_m = \mathcal{F}[\delta]$$

**Biased galaxy distribution**

$$\rho_g \propto \rho_m^\alpha \exp\left(-(\rho_m/\rho_0)^{-\epsilon}\right)$$

**Selected/contaminated sample**  $\rho_g^s(\vec{x}) = S(\vec{x})\rho_g(\vec{x})$

**Random extraction**

**(i.e. observational metric)**

$$N_g^s \leftarrow \mathcal{P}(\rho_g^s) \quad (\text{Poisson, Negative binomial, ...})$$

# The BORG3 inference framework

$$\pi(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k |\hat{\delta}_k|^2 / P_k\right)$$

**Initial conditions**

**Total evolved matter density**

$$\rho_m = \mathcal{F}[\delta]$$

**Biased galaxy distribution**

$$\rho_g \propto \rho_m^\alpha \exp\left(-(\rho_m/\rho_0)^{-\epsilon}\right)$$

**Selected/contaminated sample**

$$\rho_g^s(\vec{x}) = S(\vec{x})\rho_g(\vec{x})$$

**Random extraction**

**(i.e. observational metric)**

$$N_g^s \leftarrow \mathcal{P}(\rho_g^s) \quad (\text{Poisson, Negative binomial, ...})$$

Forward and adjoint model

# The BORG3 inference framework

$$\pi(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k |\hat{\delta}_k|^2 / P_k\right)$$

**Initial conditions**

**Total evolved matter density**

$$\rho_m = \mathcal{F}[\delta]$$

**Biased galaxy distribution**

$$\rho_g \propto \rho_m^\alpha \exp\left(-(\rho_m/\rho_0)^{-\epsilon}\right)$$

**Selected/contaminated sample**  $\rho_g^s(\vec{x}) = S(\vec{x})\rho_g(\vec{x})$

**Random extraction**

**(i.e. observational metric)**

$$N_g^s \leftarrow \mathcal{P}(\rho_g^s) \quad (\text{Poisson, Negative binomial, ...})$$

Easily exchangeable to try  
your favorite differentiable model

# The BORG3 inference framework

$$\pi(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k |\hat{\delta}_k|^2 / P_k\right)$$

**Initial conditions**

**Total evolved matter density**  $\rho_m = \mathcal{F}[\delta]$

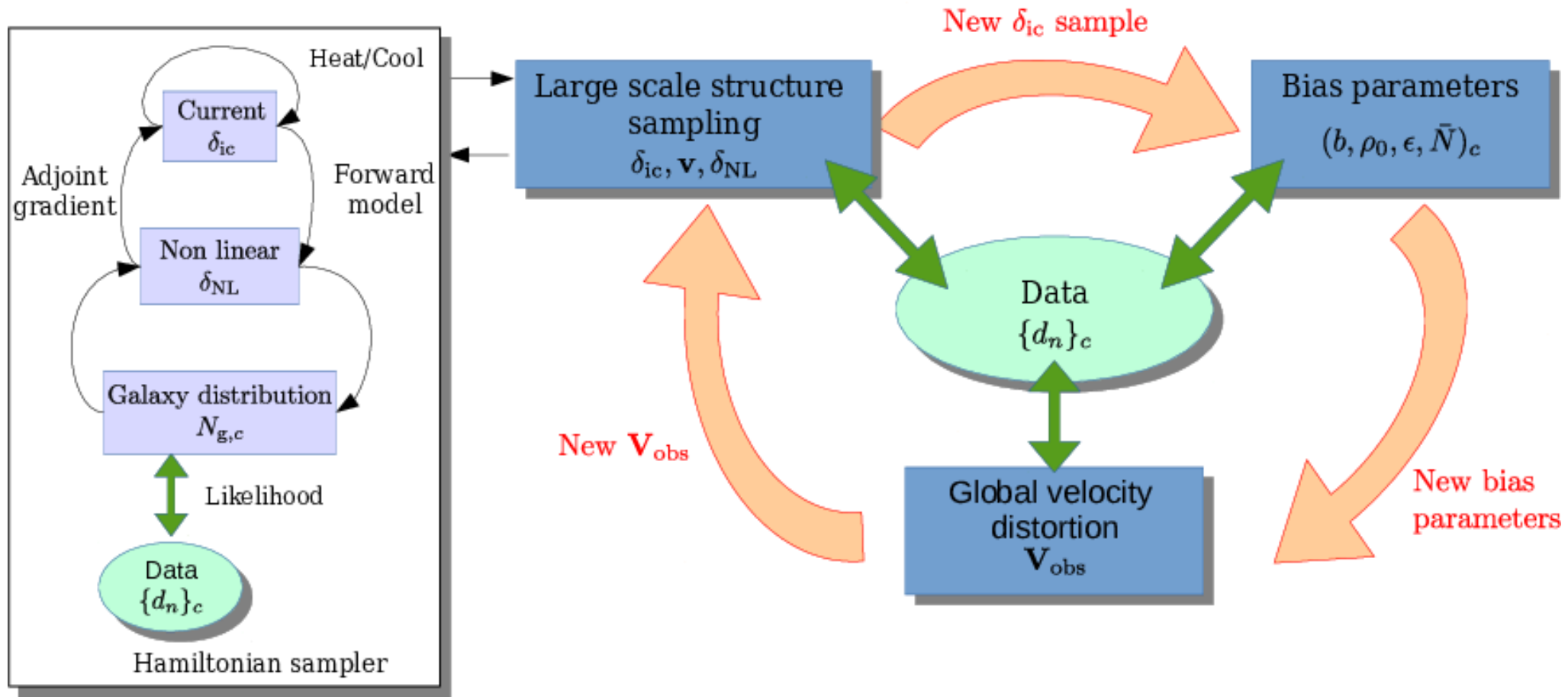
**Biased galaxy distribution**  $\rho_g \propto \rho_m^\alpha \exp\left(-(\rho_m/\rho_0)^{-\epsilon}\right)$

**Selected/contaminated sample**  $\rho_g^s(\vec{x}) = \mathbf{S}(\vec{x})\rho_g(\vec{x})$

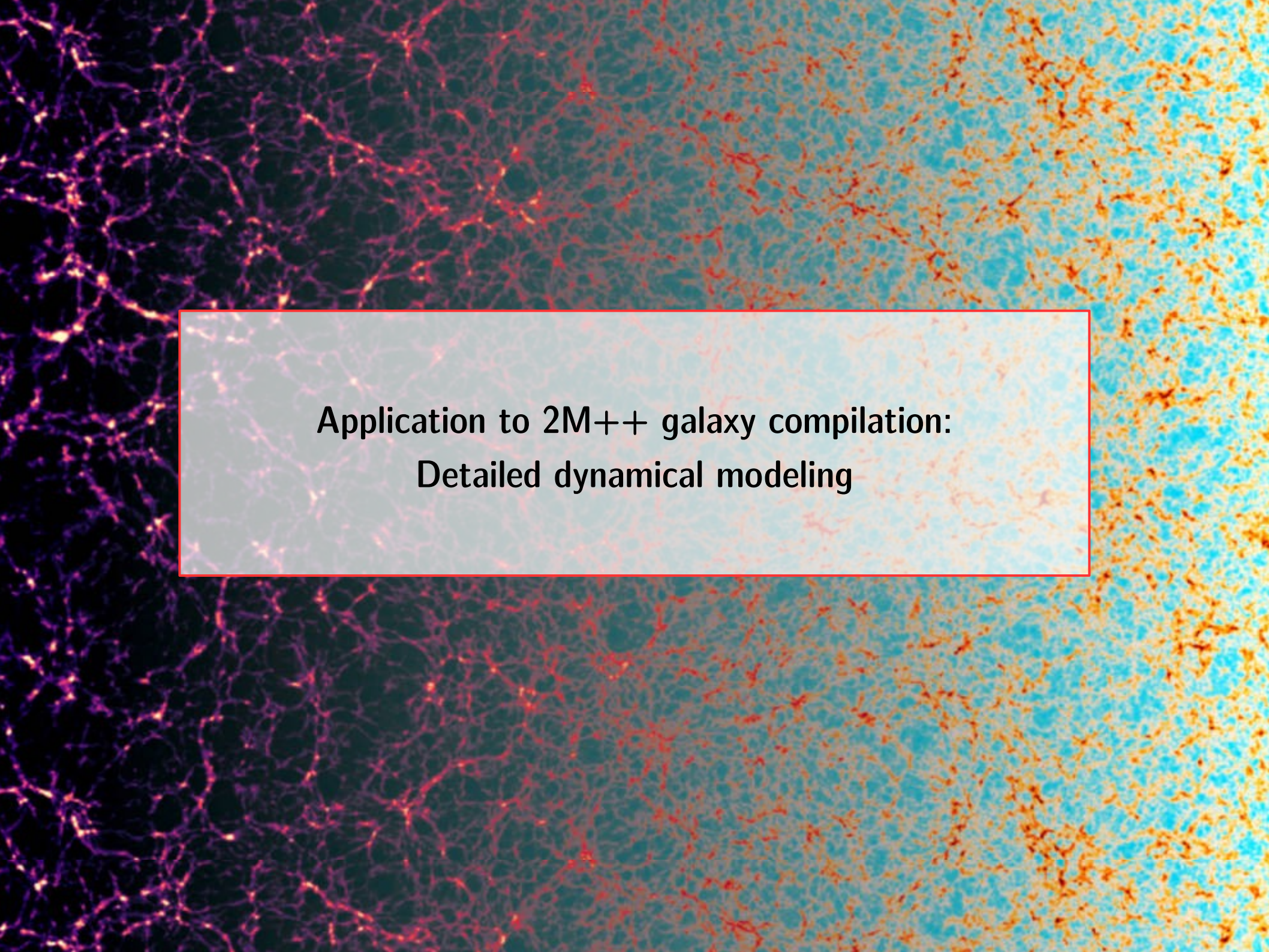
**Random extraction  
(i.e. observational metric)**  $N_g^s \leftarrow \mathcal{P}(\rho_g^s)$  (Poisson, Negative binomial, ...)

**Encode survey systematic effects with expansions:**  $S(\hat{x}) = S_0(\hat{x}) \prod_{f=1}^N (1 + \alpha_f F_f(\hat{x}))$

# The BORG3 machine

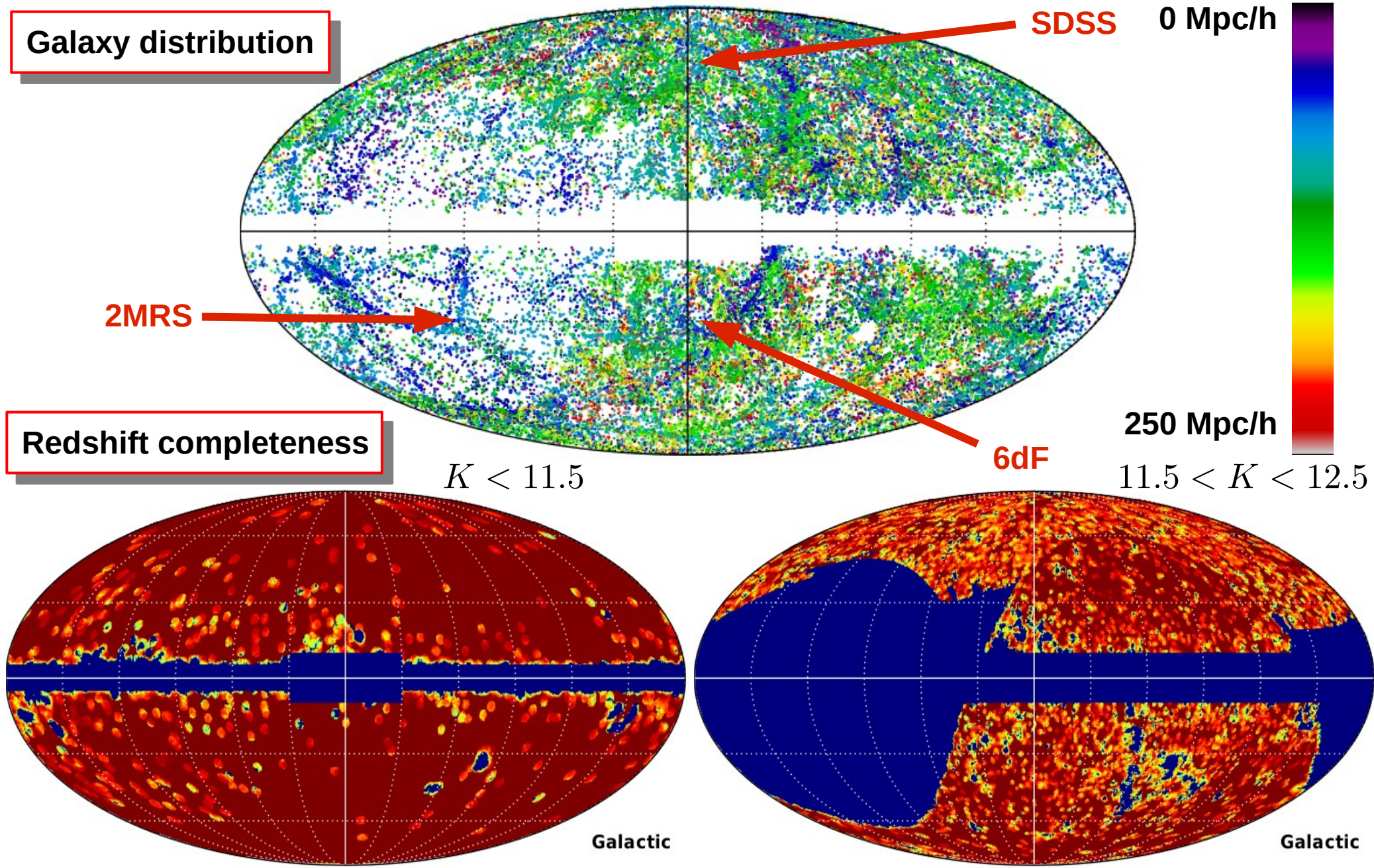




A visualization of the cosmic web, showing a complex network of filaments and nodes. The left side is dominated by dark purple and blue filaments, while the right side transitions into a lighter blue and yellowish-orange background with more diffuse, filamentary structures.

**Application to 2M++ galaxy compilation:  
Detailed dynamical modeling**

# The 2M++ galaxy compilation



Galaxy distribution

SDSS 0 Mpc/h

2MRS

Redshift completeness

$K < 11.5$

6dF

250 Mpc/h

$11.5 < K < 12.5$

Galactic

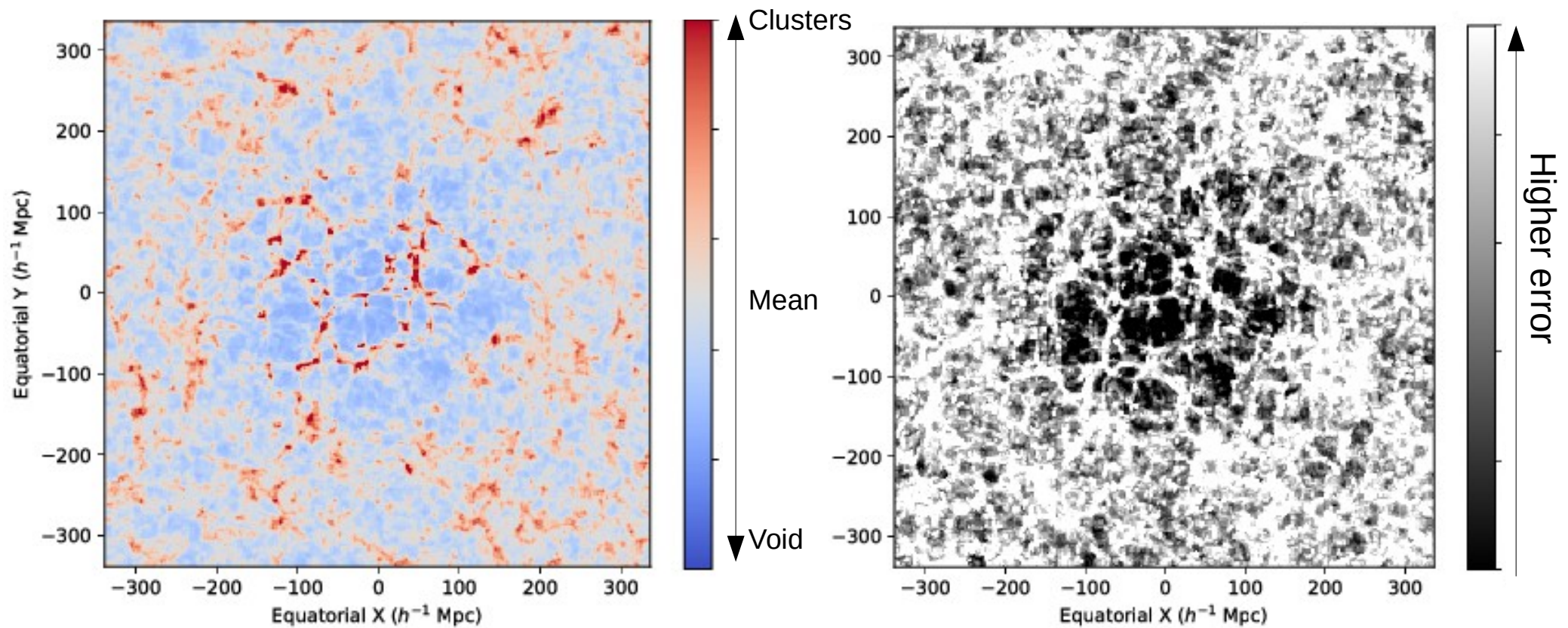
Galactic

~70 000 galaxies

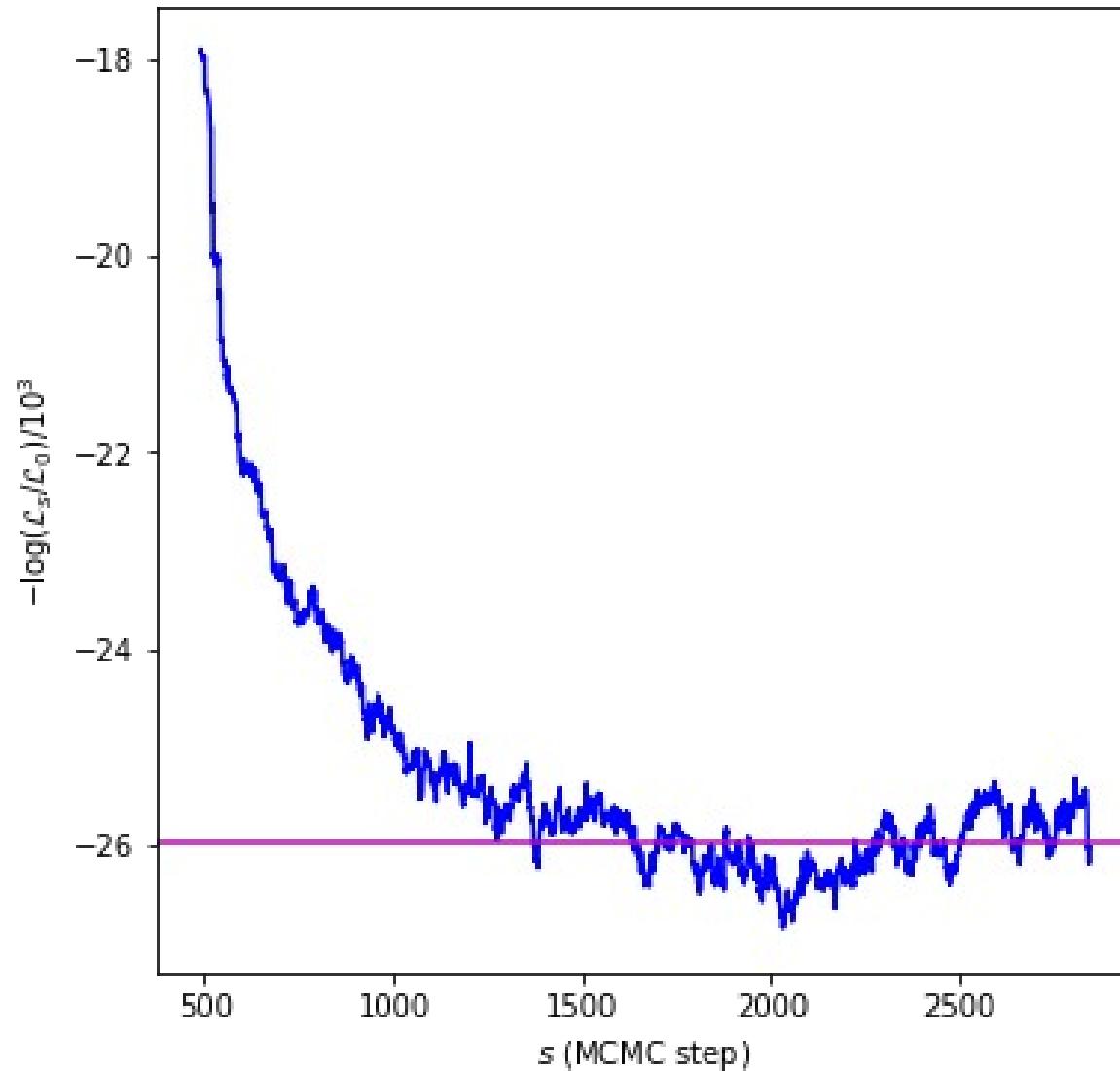
Lavaux & Hudson (MNRAS, 2011)

# Inferred density fields

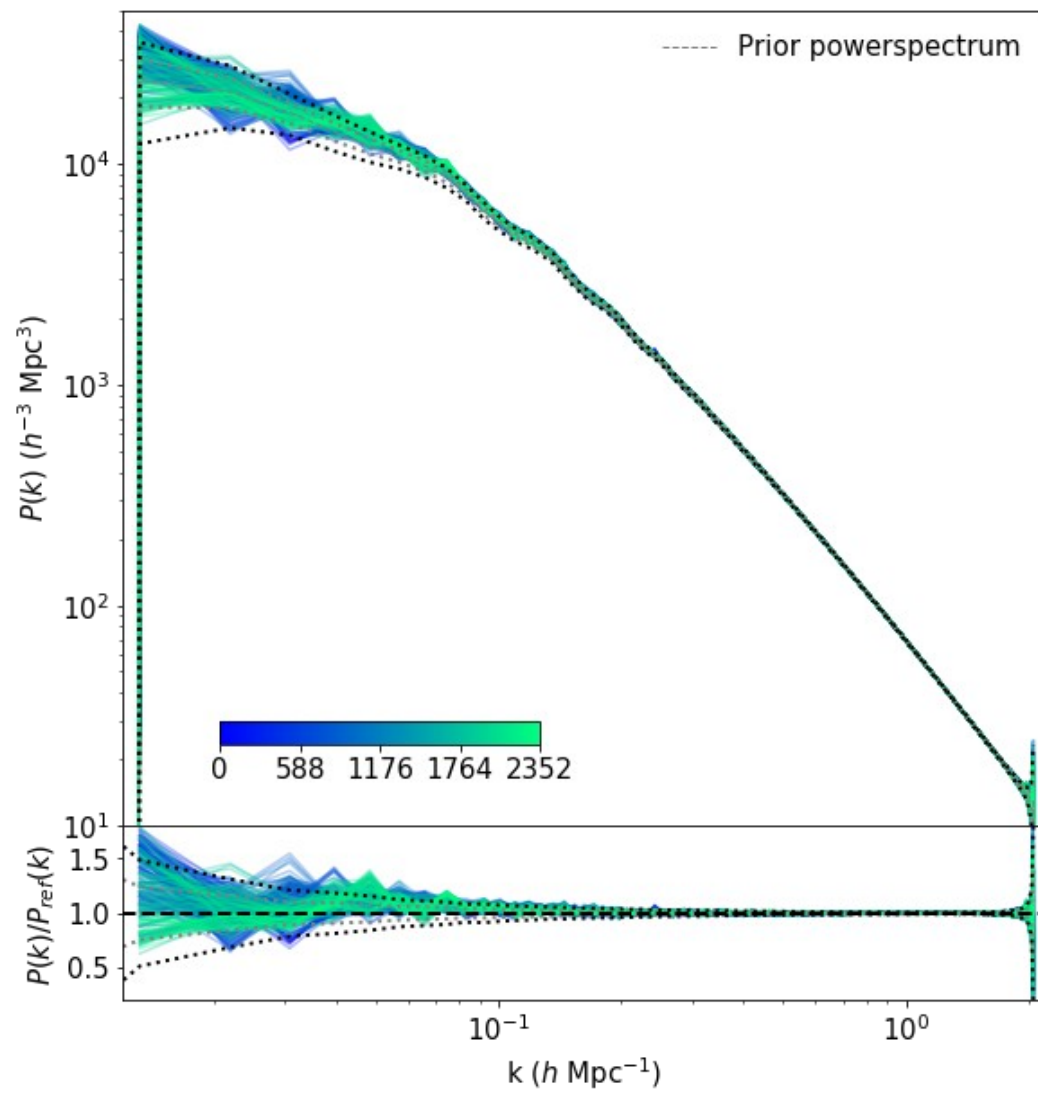
Ensemble average density fields at  $z=0$



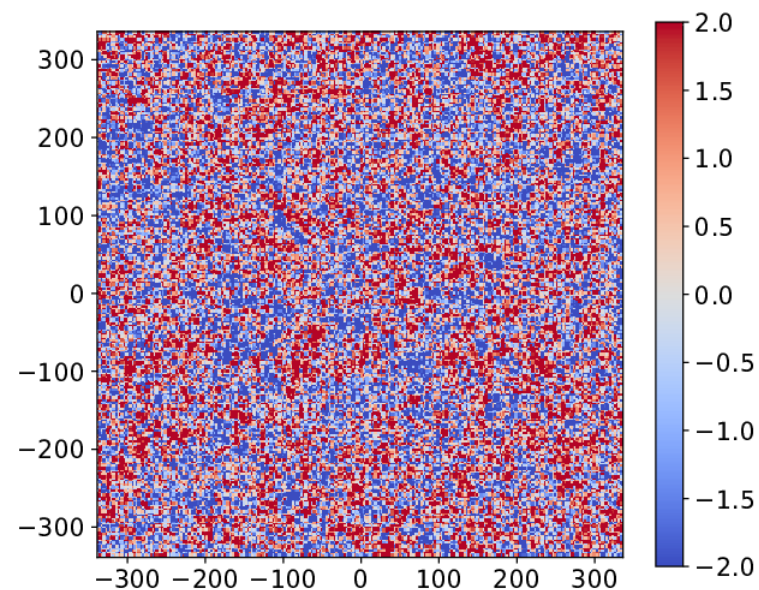
# Performance aspect: burnin



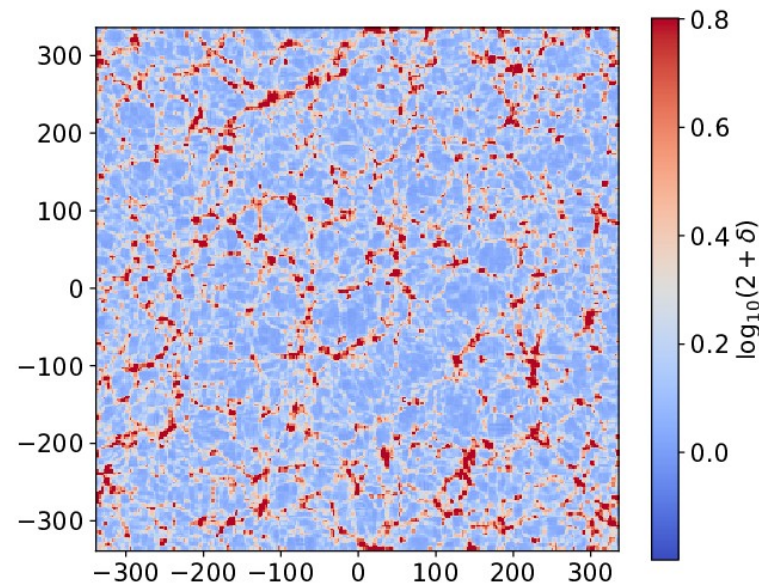
# Initial condition powerspectrum



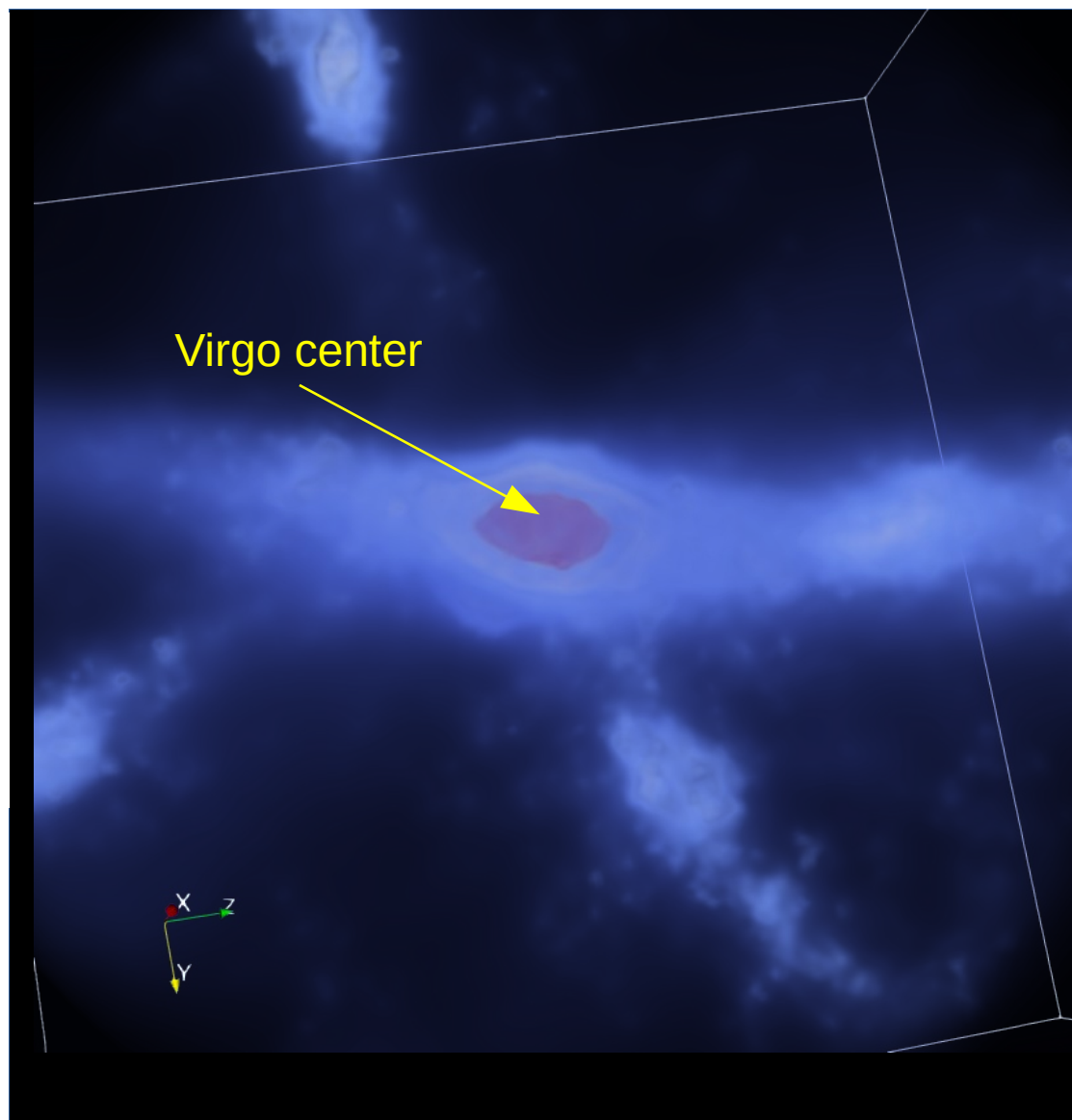
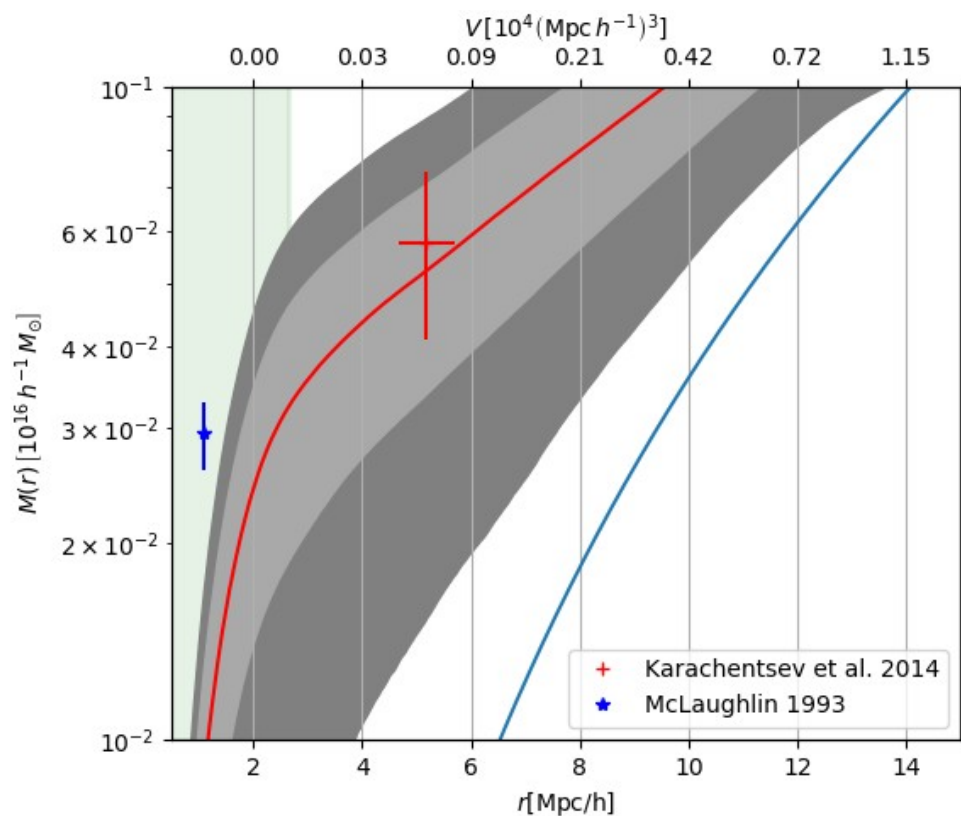
Initial conditions



Post PM simulation

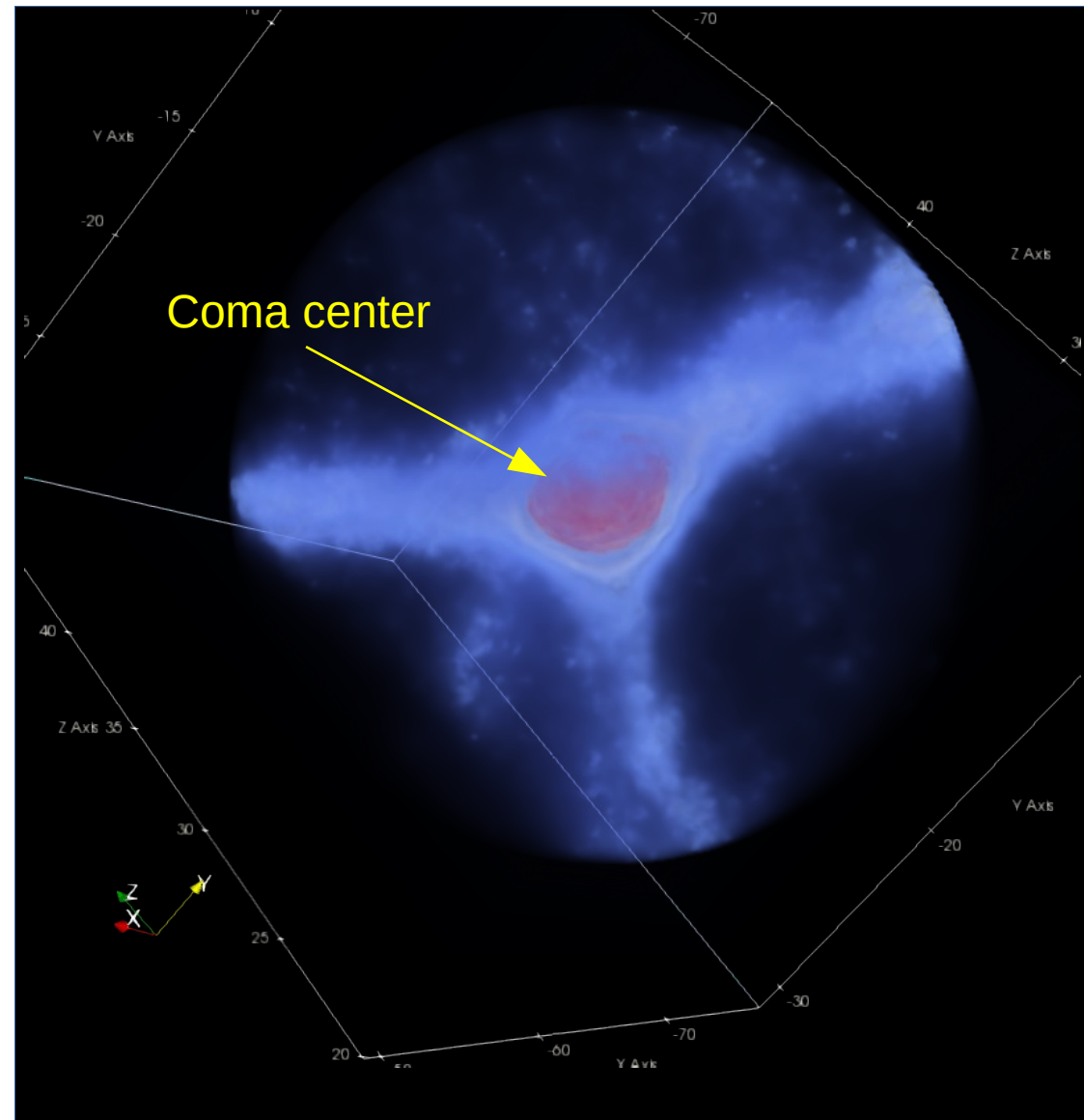
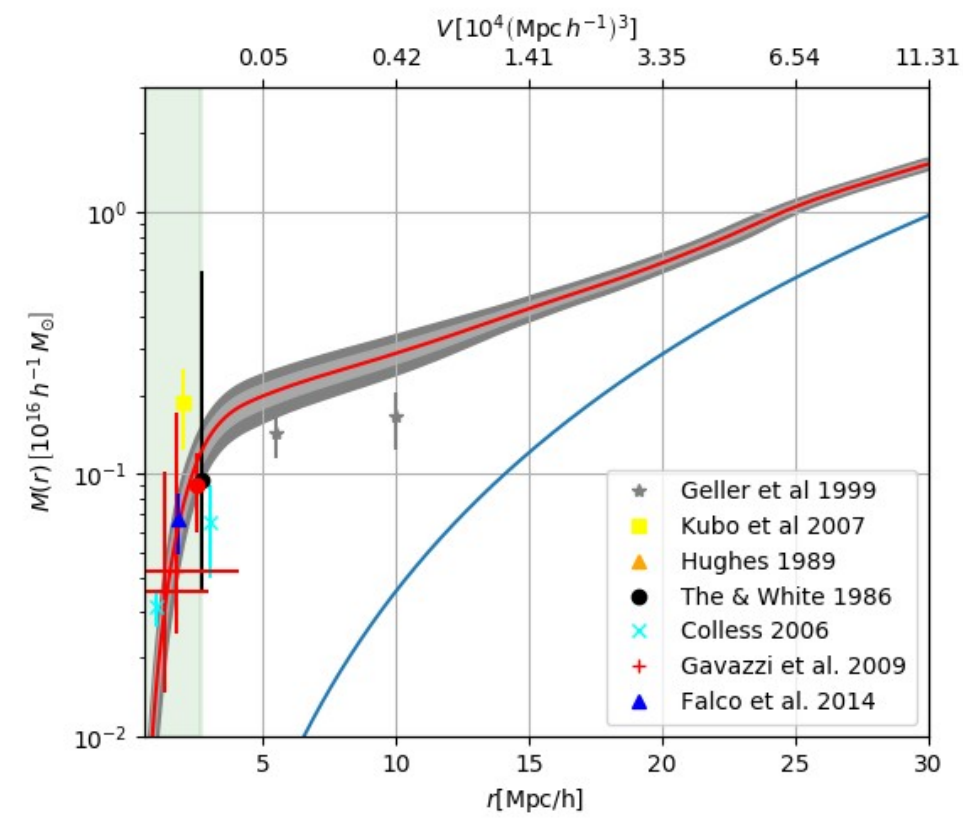


# Virgo cluster

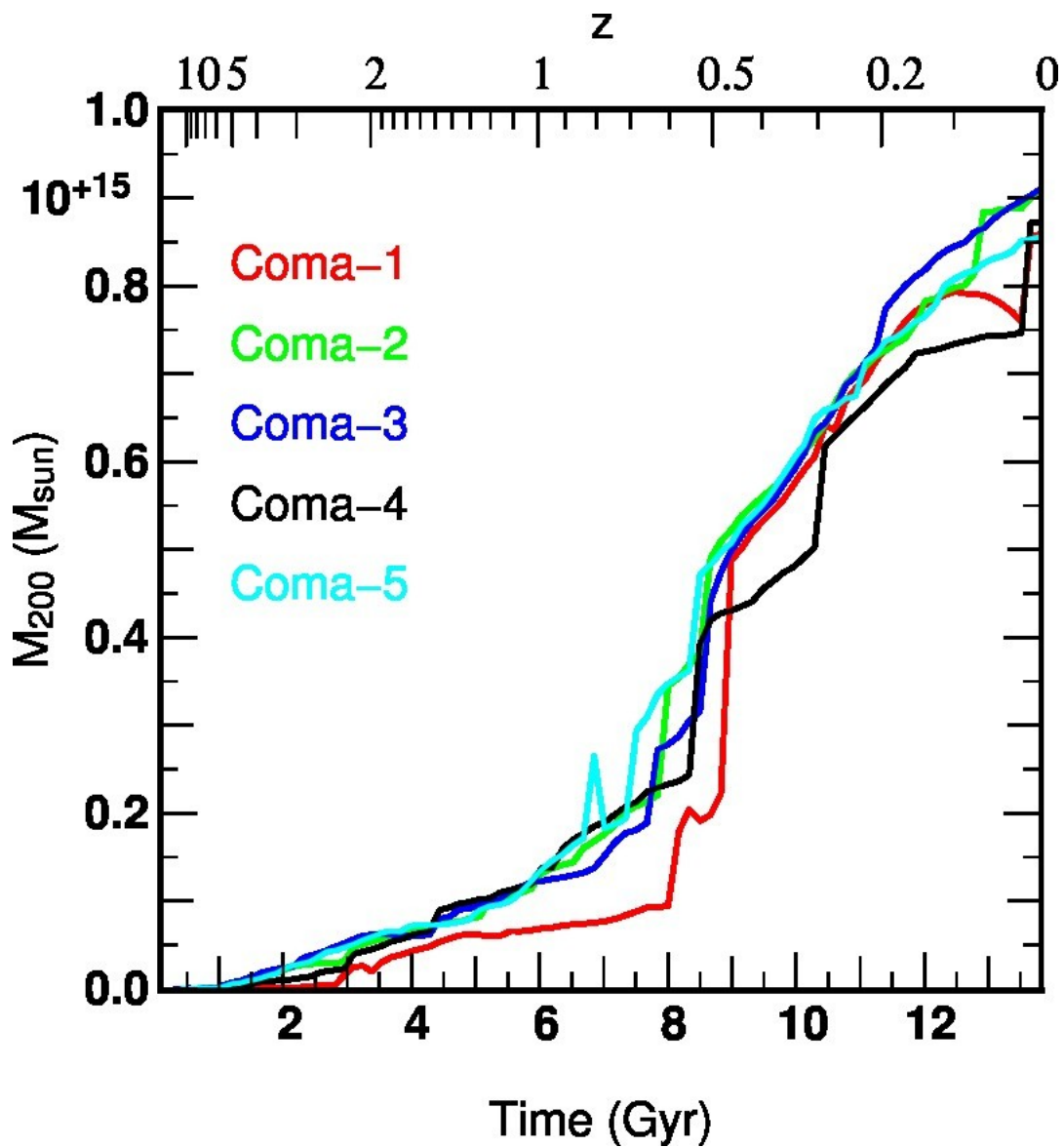
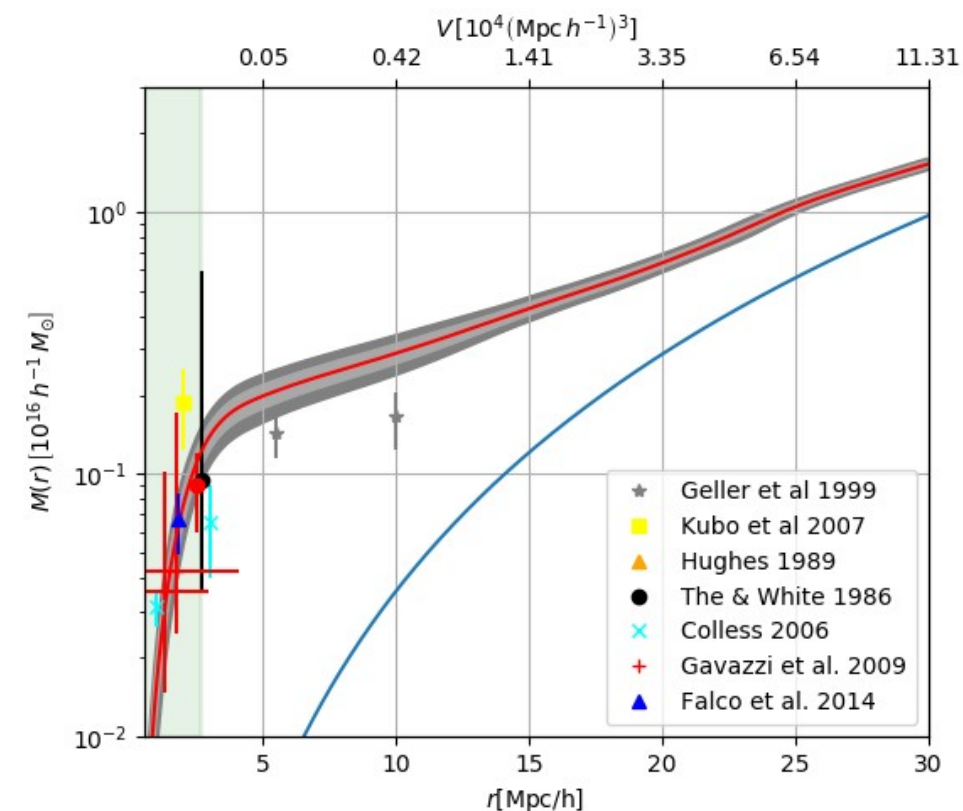


PLAY

# Coma dynamical properties



# Coma dynamical properties



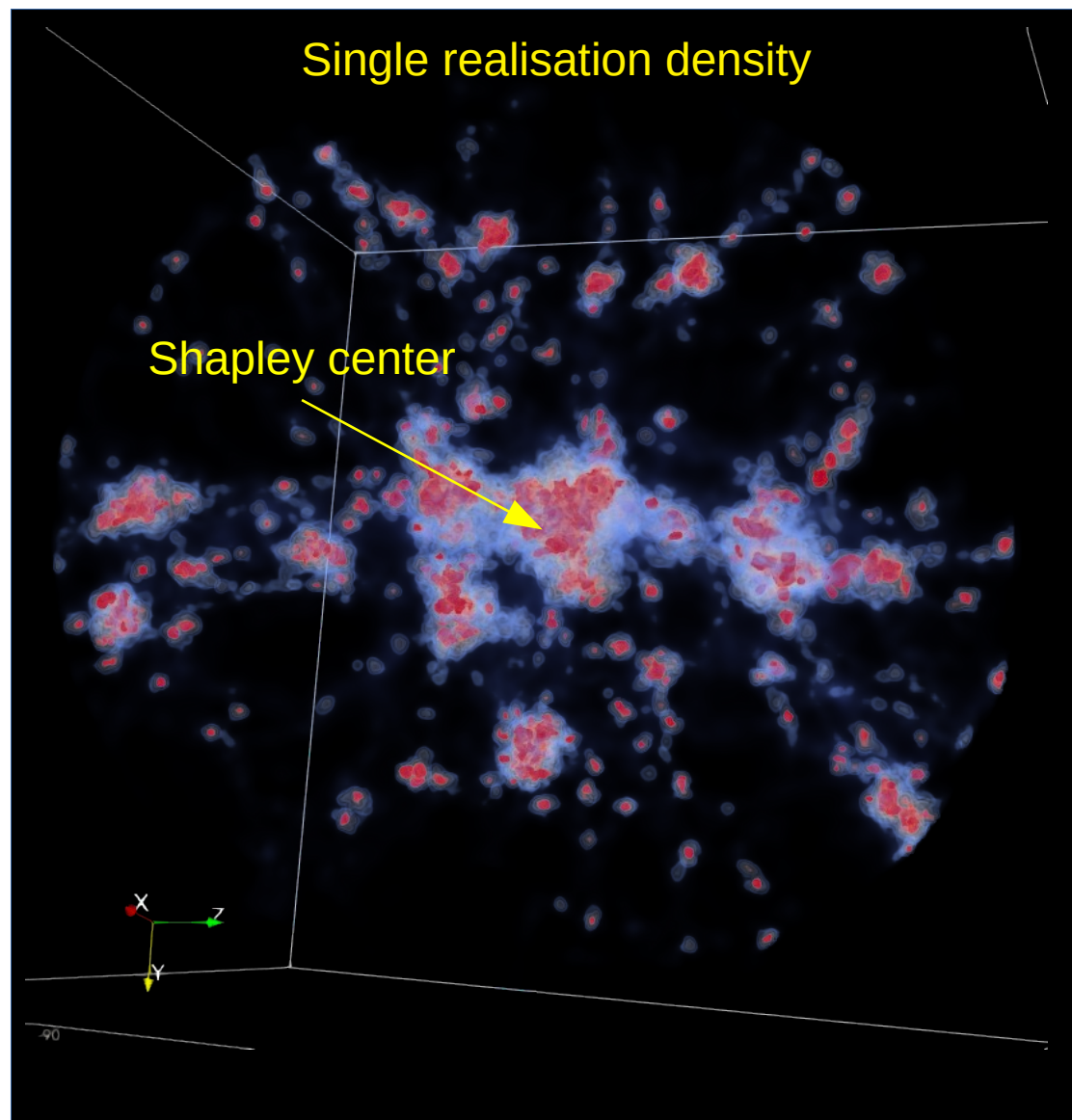
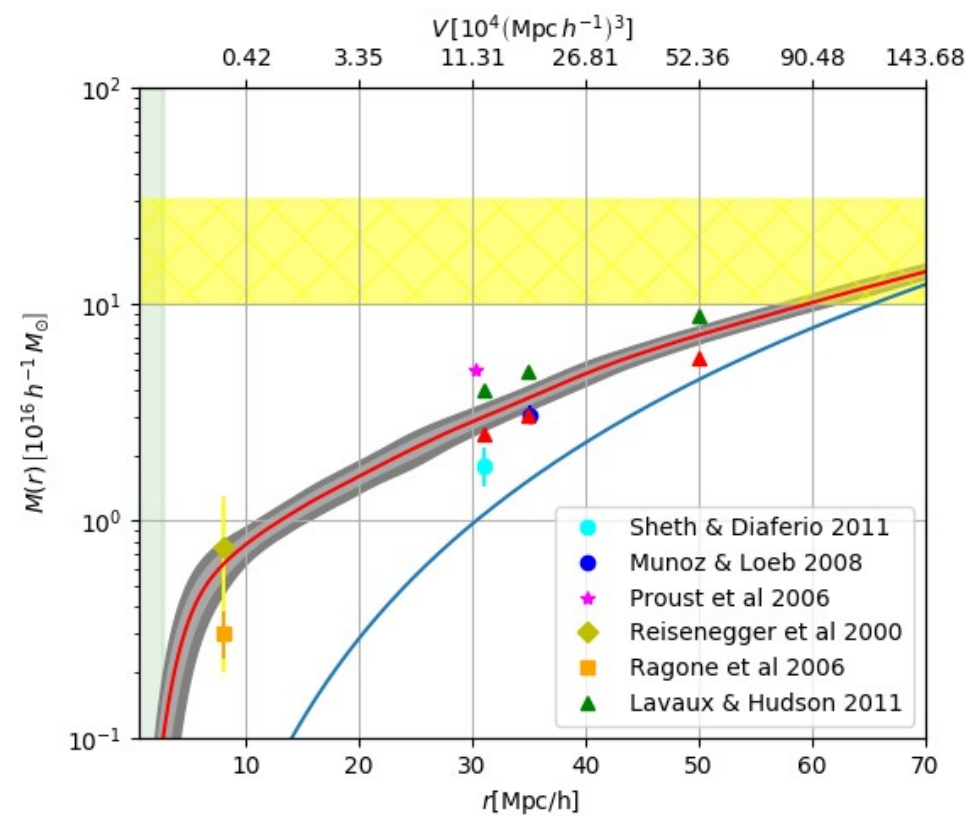
Zoom simulation on Coma  
 (~250 Mpart in zoom)

$4 \times 10^7 h^{-1} M_{\odot} / \text{part}$

PLAY

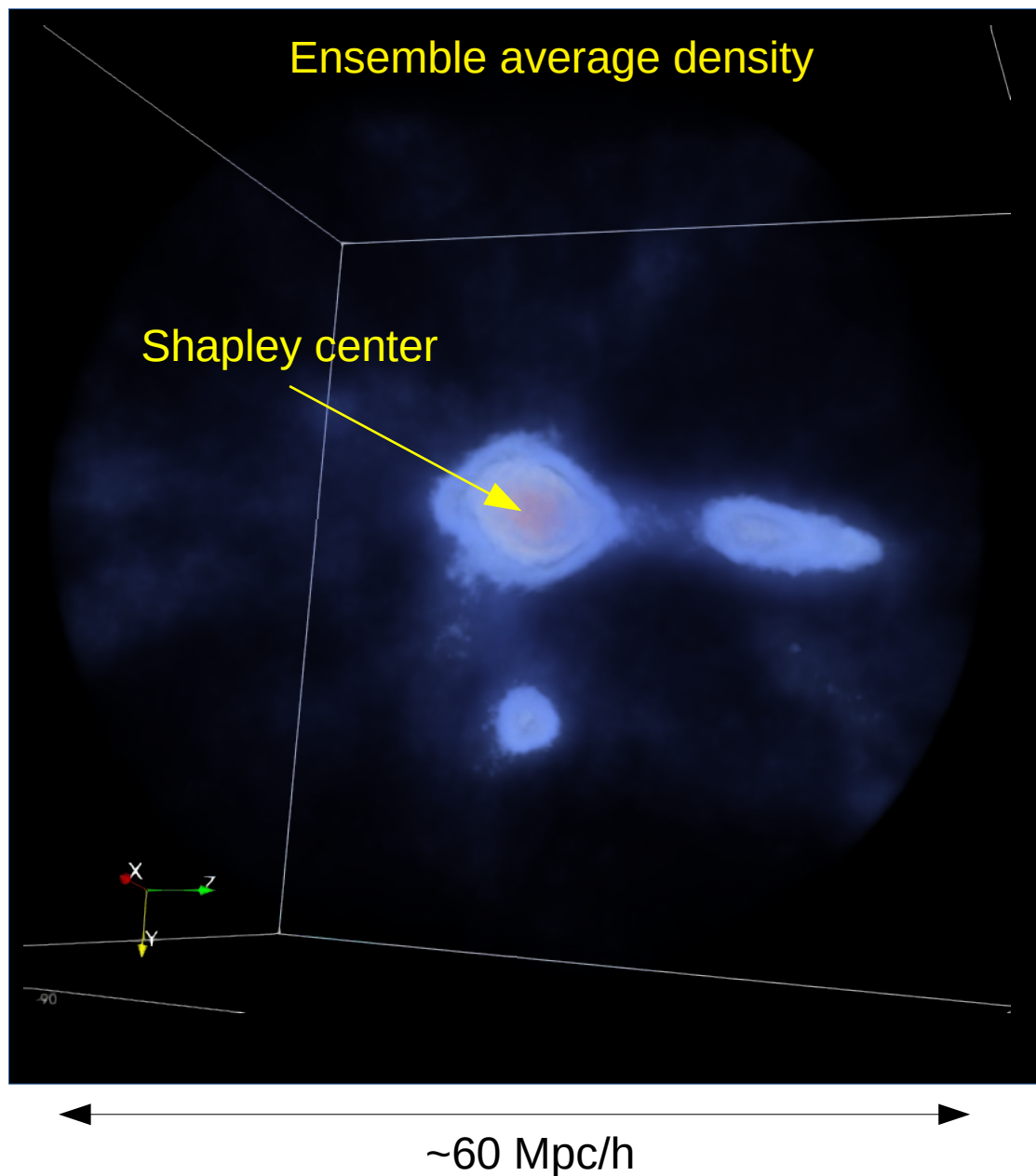
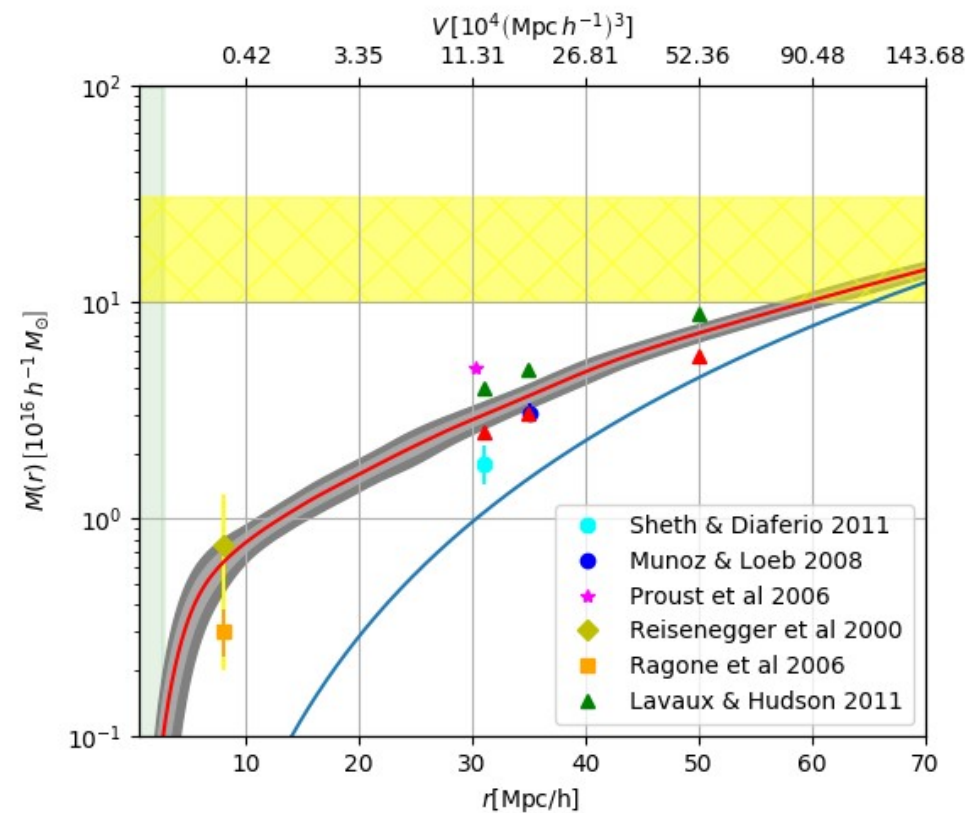


# Shapley concentration

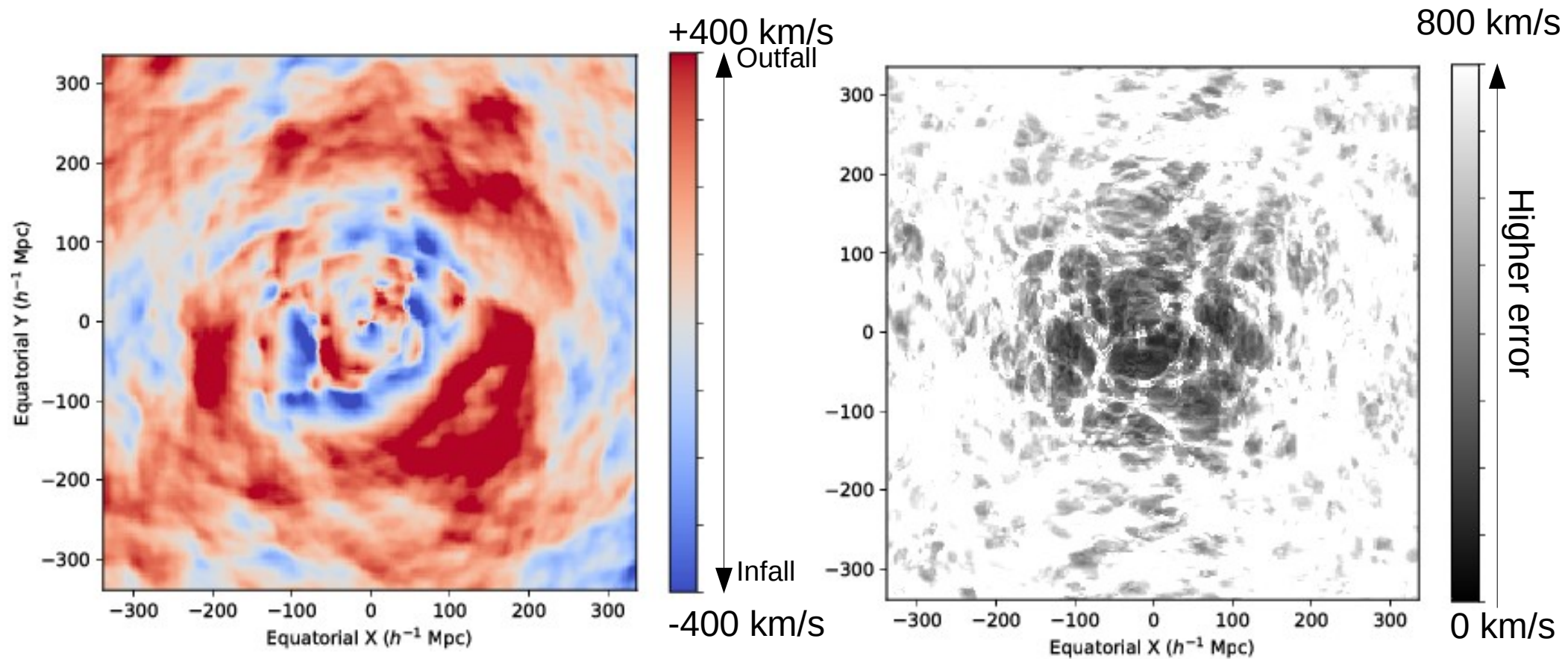


PLAY

# Shapley concentration

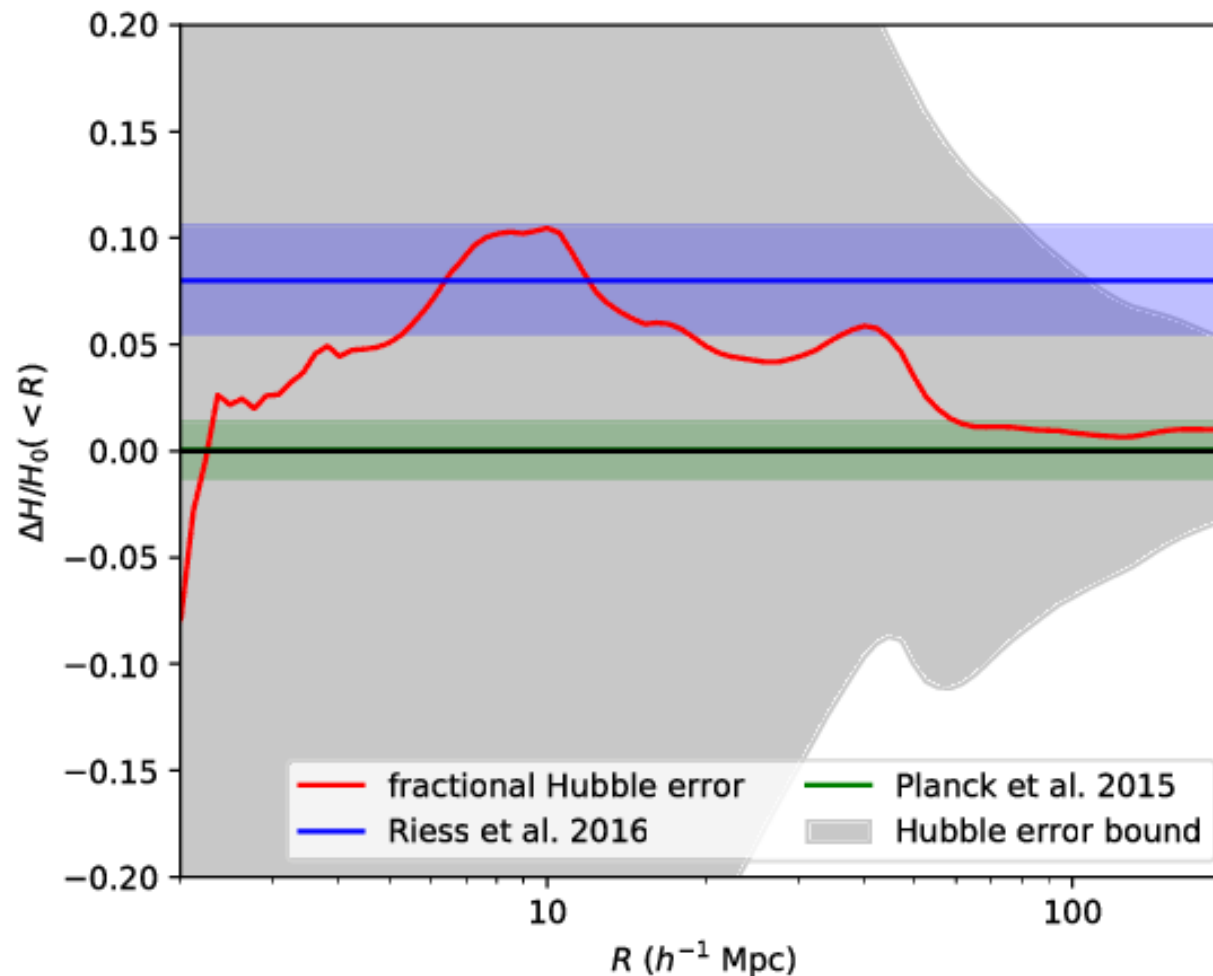


# Inferred velocity fields



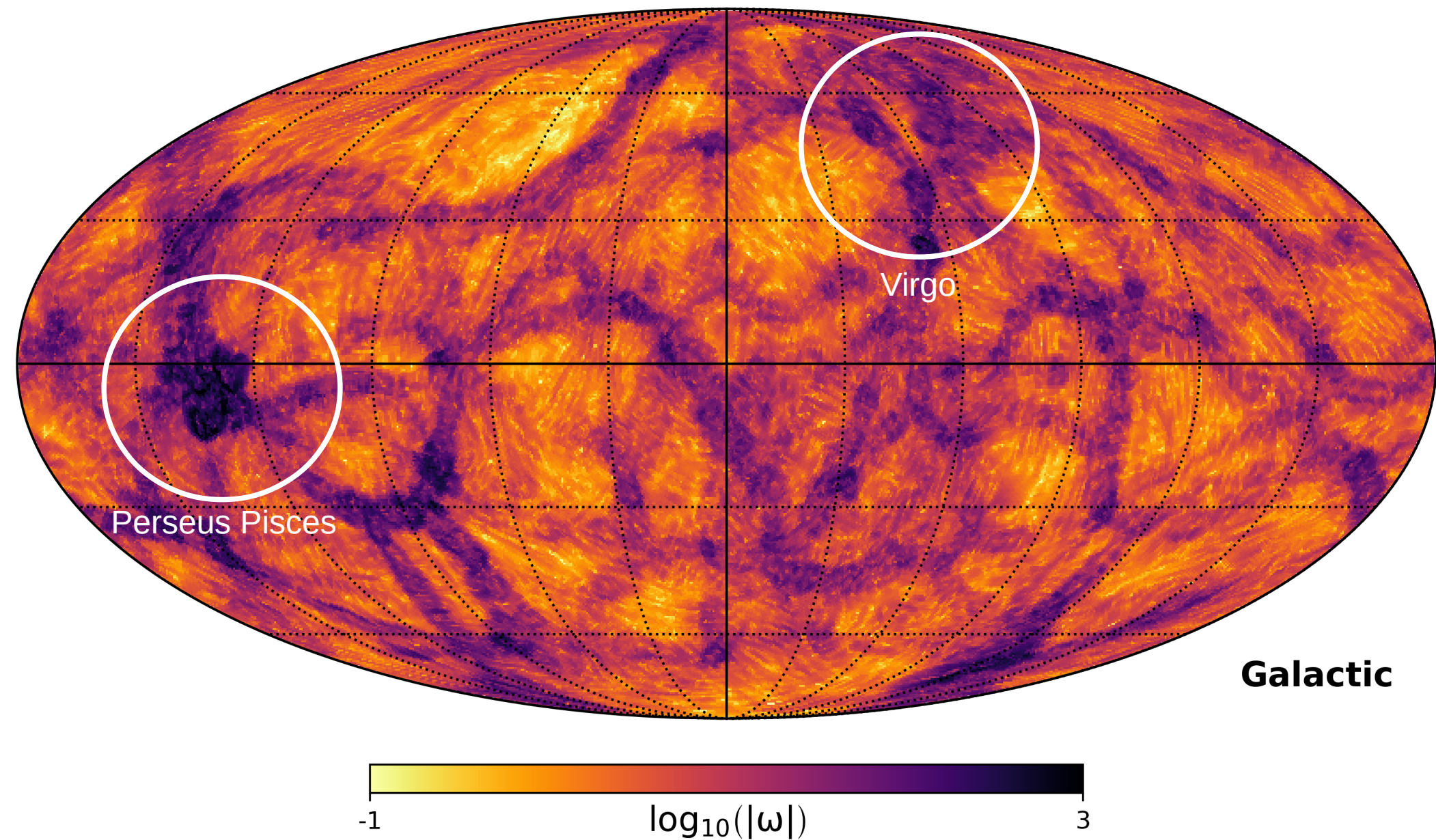
# Velocity field and Hubble constant

Mean error on Hubble measurement using tracers from observed large scale structures



**TODO: Compare all the flow models**

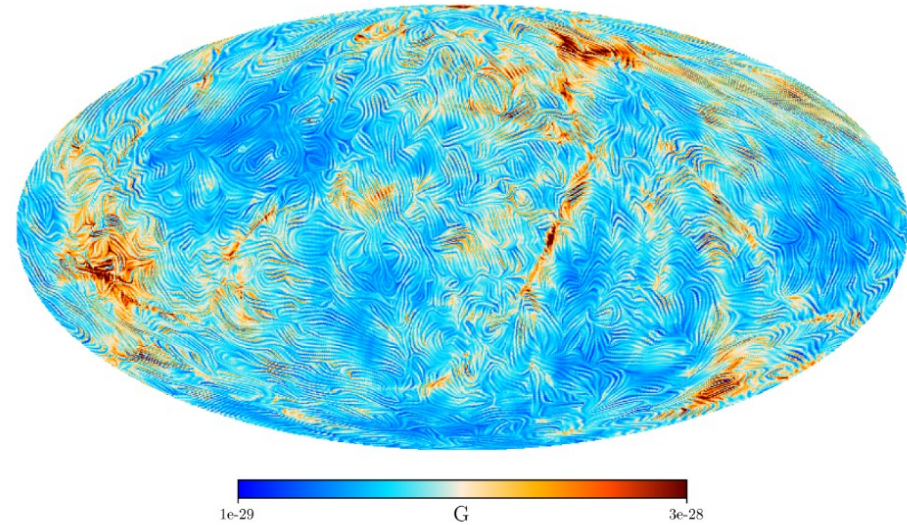
# Peculiar velocity field vorticity



# More applications

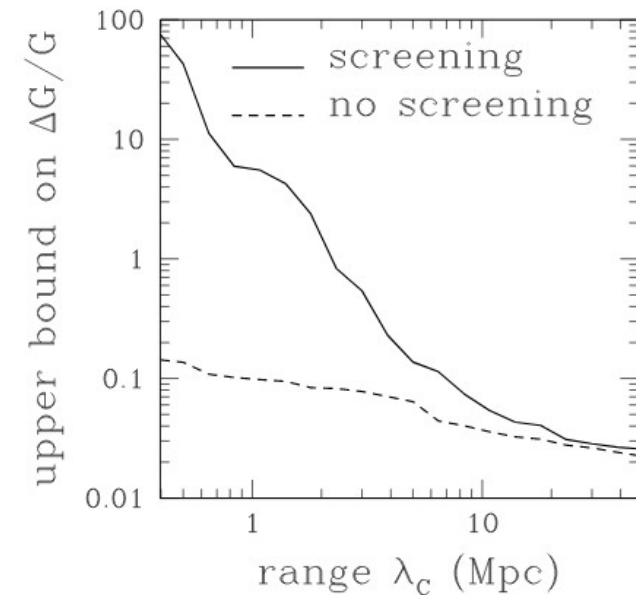
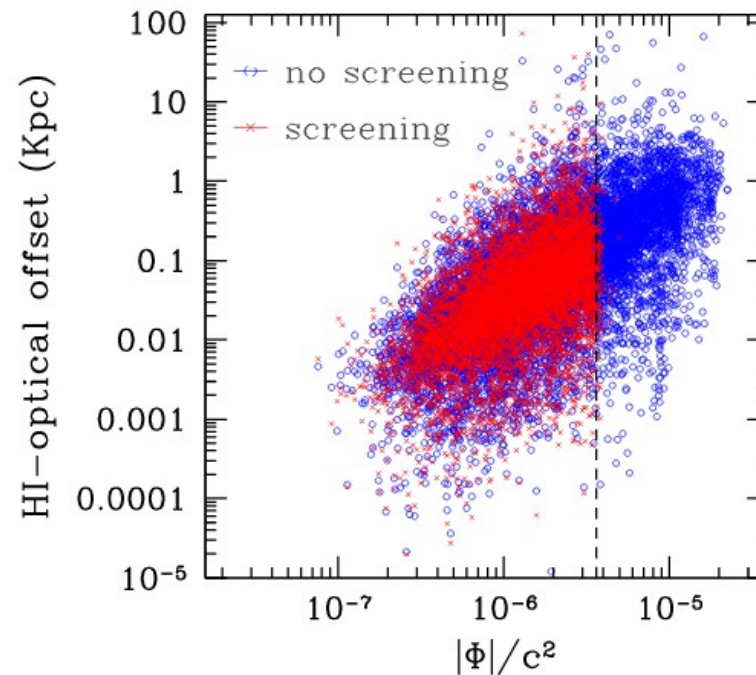
Magnetic field in our backyard generated by primordial mechanisms

Hutschenreuter et al. (2018, CQG)



“Fifth-force” constraints

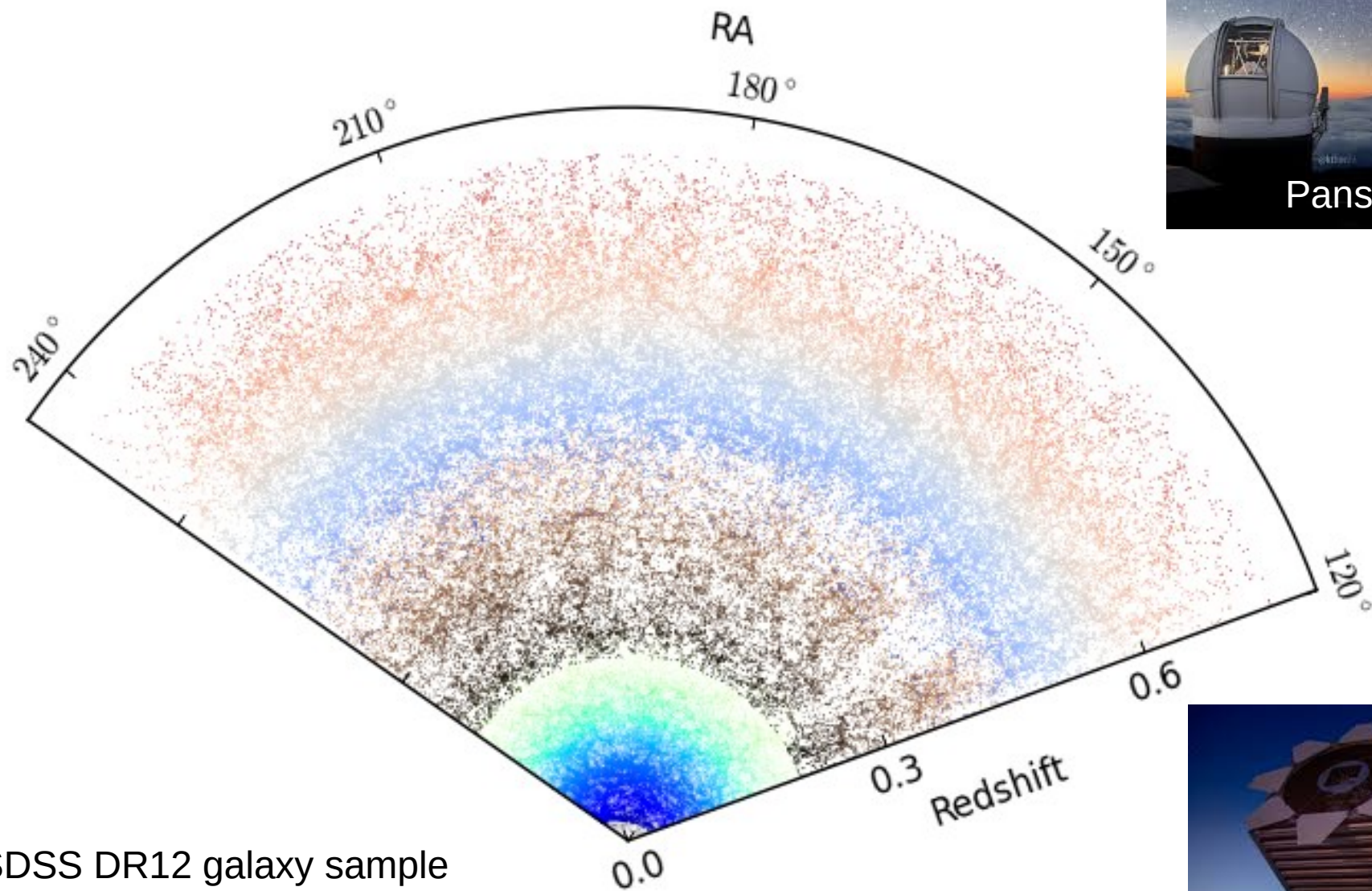
Desmond et al.  
(2018ab, PRD, PRL in review)



A visualization of the cosmic web, showing a complex network of filaments and nodes. The left side is dark purple and black, transitioning to red and orange filaments, and finally to a bright cyan and yellow background on the right. A semi-transparent white box with a red border is centered in the middle.

**Application to Sloan Digital Sky Survey III:  
Deep cosmological application**

# SDSS3 data



SDSS DR12 galaxy sample  
~1.6 millions of galaxies



Panstarrs



SDSS



# Forward model becomes more complex

## Cosmic expansion

Non-linear density remapping:  $\vec{x} \rightarrow \vec{z}$

$$\vec{z}(\vec{x}) = f(|\vec{x}|, \text{cosmology}) \times \vec{x}$$

→ Usual Hubble expansion at linear order

$$f(|x|, \dots) \simeq Hx + o(x)$$

## Cosmic growth of structures

Implemented so far for (2)LPT:

$$\vec{x}(\vec{q}, t) = \vec{q} + \Psi(\vec{q}, t) \underset{LPT}{\simeq} \vec{q} + D(t)\Psi(\vec{q})$$

→ Evolving density contrasts

# Forward model becomes more complex

## Cosmic expansion

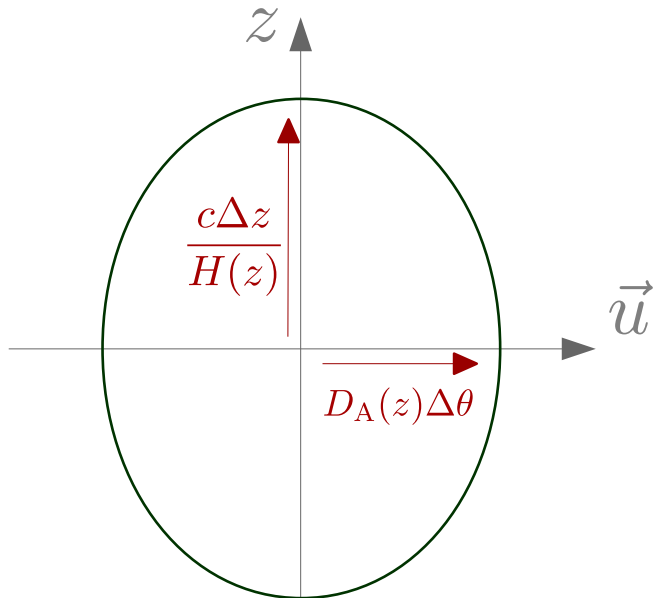
Non-linear density remapping:  $\vec{x} \rightarrow \vec{z}$

$$\vec{z}(\vec{x}) = f(|\vec{x}|, \text{cosmology}) \times \vec{x}$$

→ Usual Hubble expansion at linear order

$$f(|x|, \dots) \simeq Hx + o(x)$$

Do the ultimate Alcock-Paczynski test!



Kodi Ramanah, Lavaux, Jasche, Wandelt (2018, in prep.)

## Cosmic growth of structures

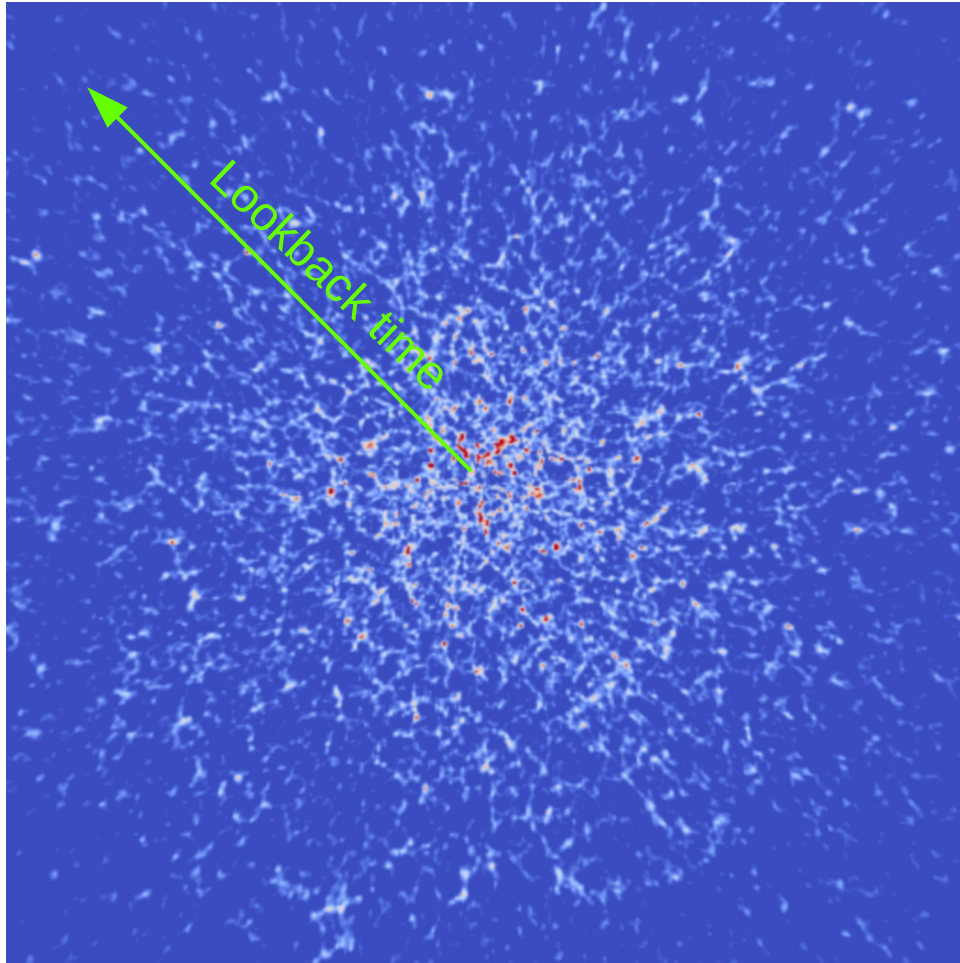
Implemented so far for (2)LPT:

$$\vec{x}(\vec{q}, t) = \vec{q} + \Psi(\vec{q}, t) \underset{LPT}{\simeq} \vec{q} + D(t)\Psi(\vec{q})$$

→ Evolving density contrasts

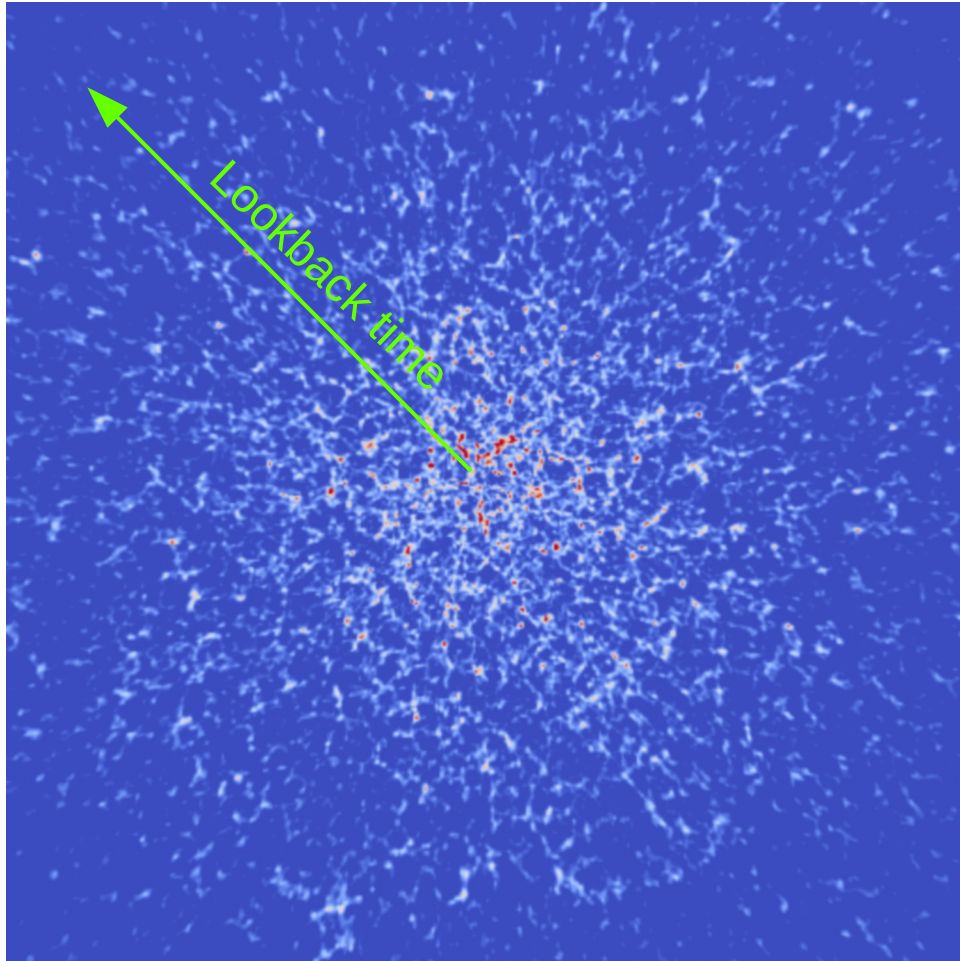
# Forward model becomes more complex

## Cosmic expansion

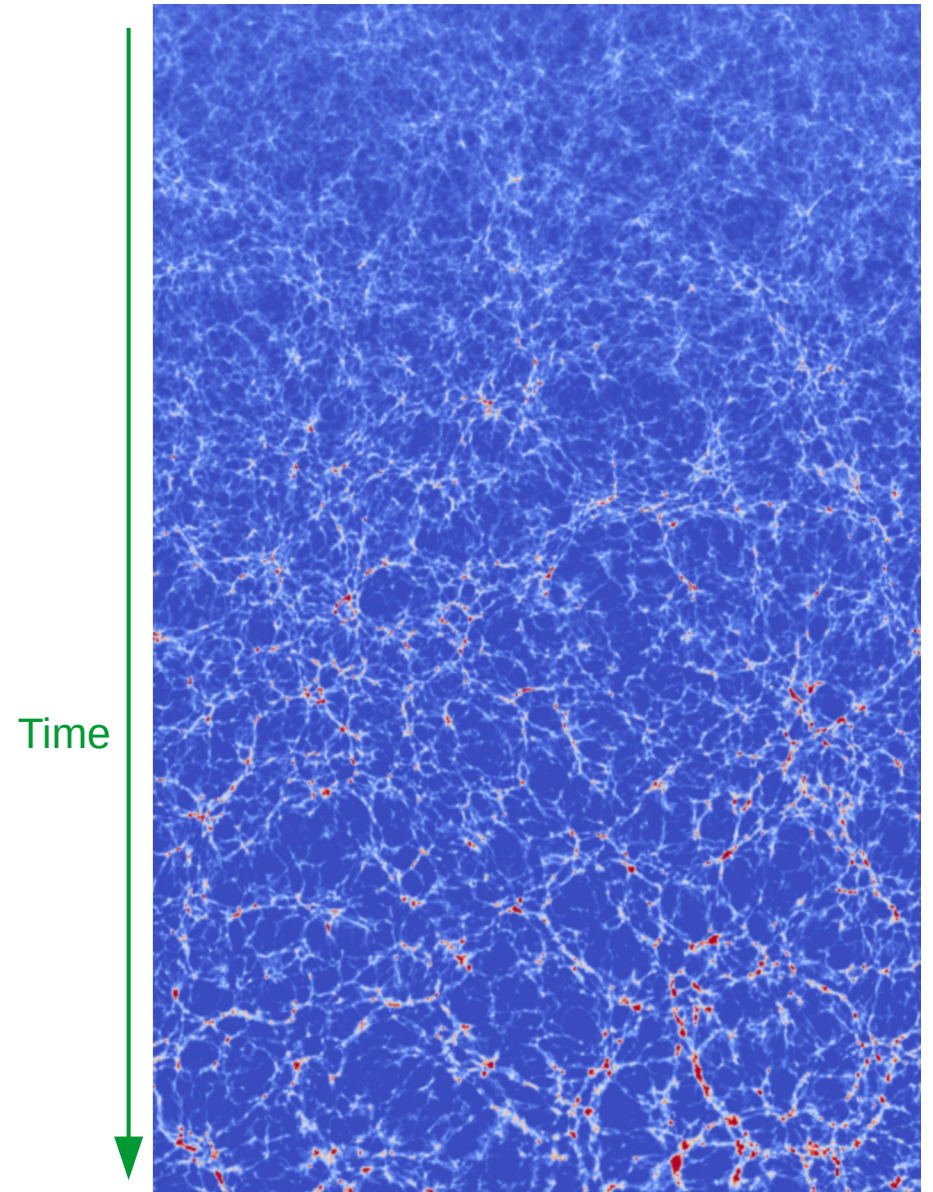


# Forward model becomes more complex

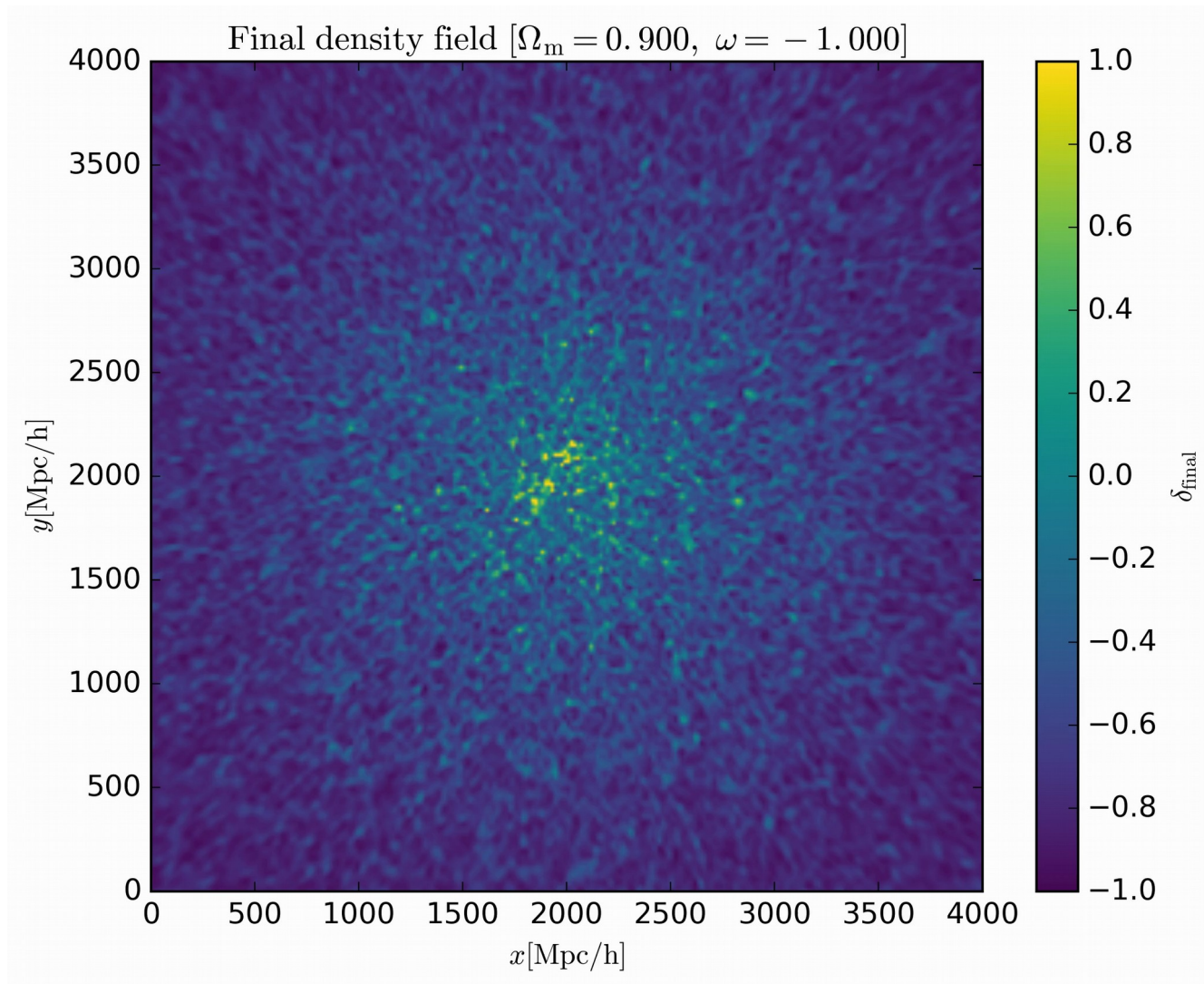
Cosmic expansion



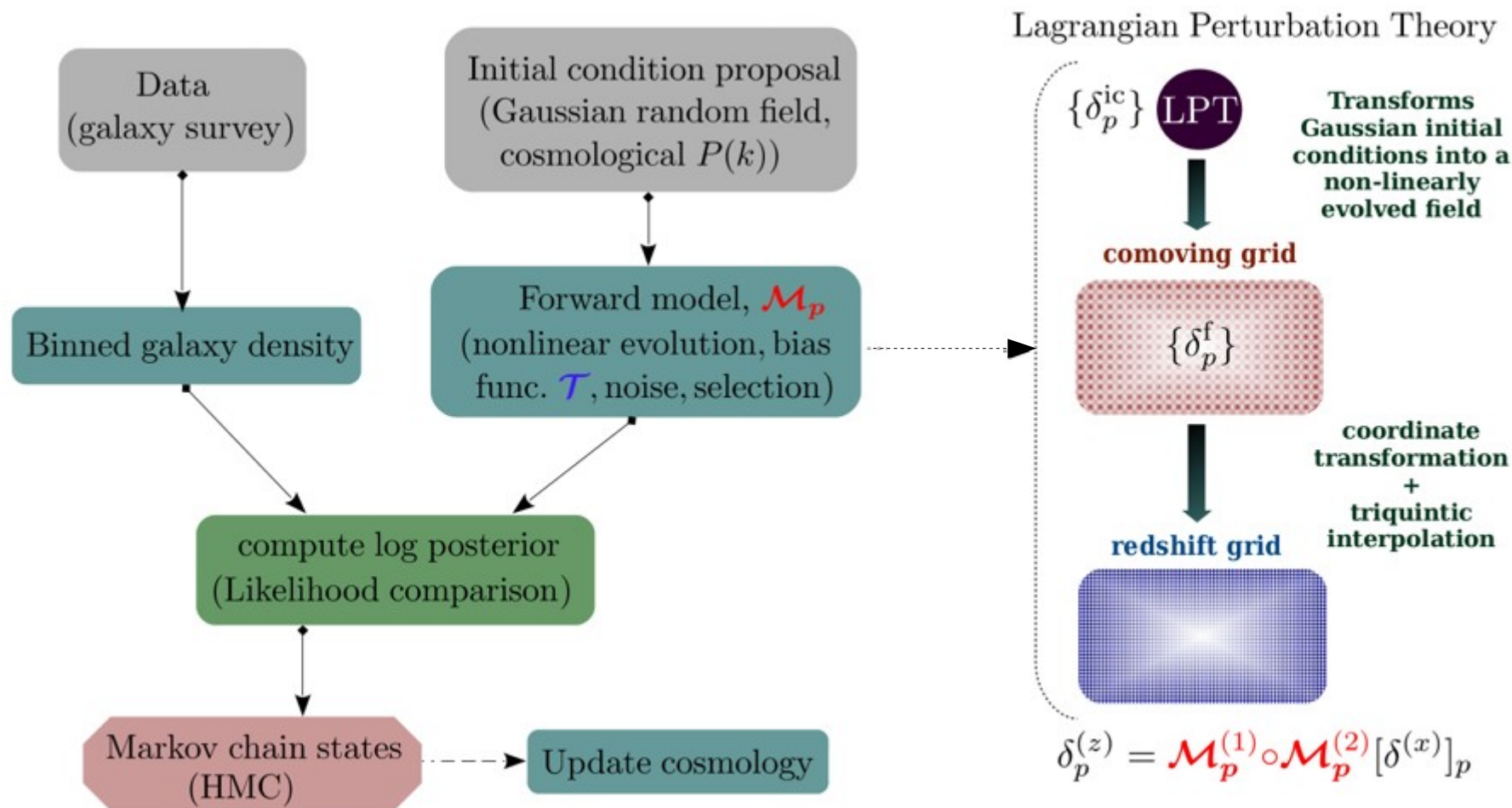
Cosmic growth of structures



# Impact of parameters on expansion



# An updated scheme: more coordinates transform



**Data model:** 
$$N_p^g = R(z_p) \bar{N}_p^g \mathcal{T} (1 + \mathcal{M}_p \{[\hat{\delta}_k]_{\tilde{p}}\}_p)$$

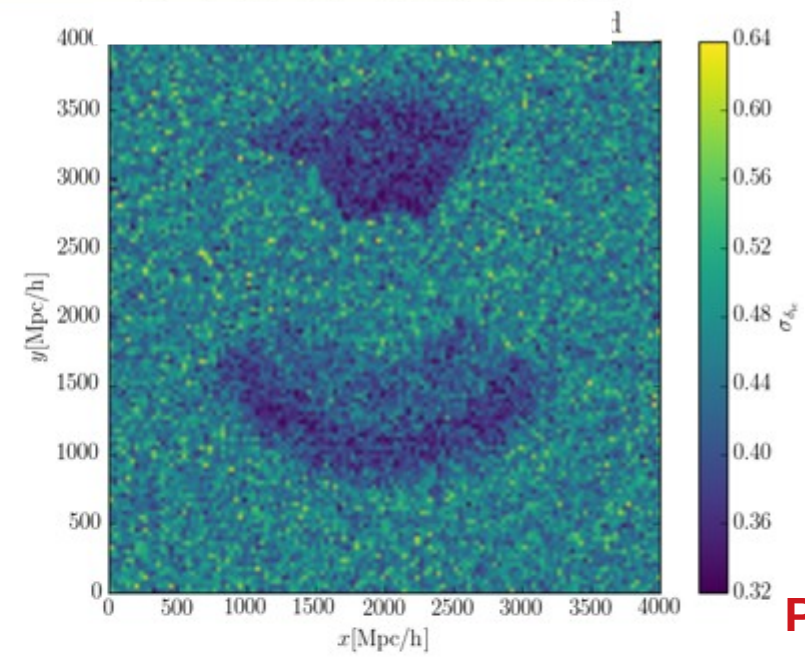
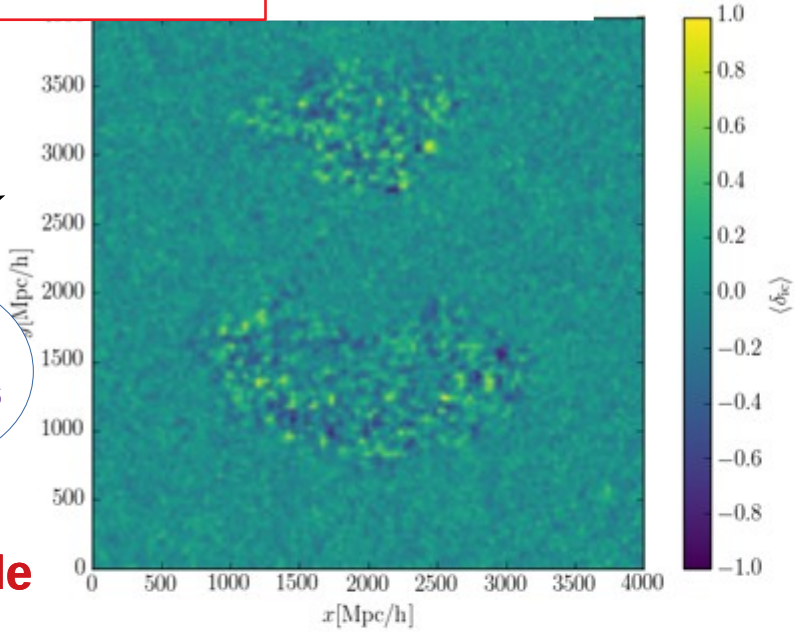
# Inference with a freed light cone

MOCK SDSS3 DATA

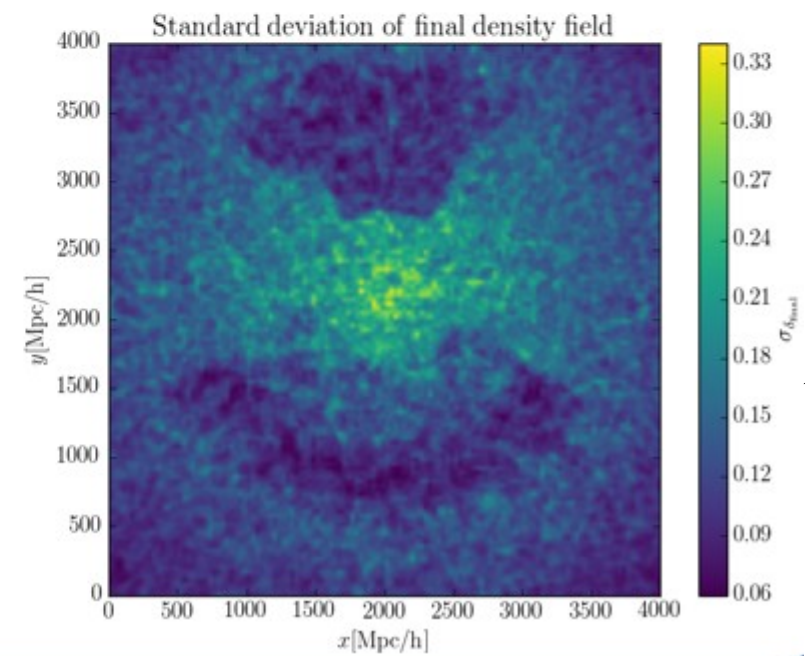
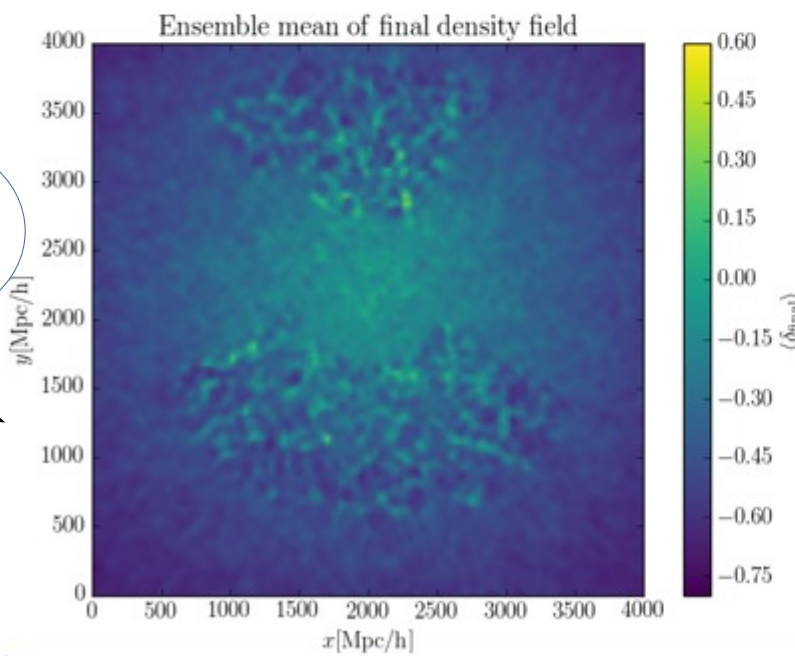
Initial conditions

Ensemble mean density

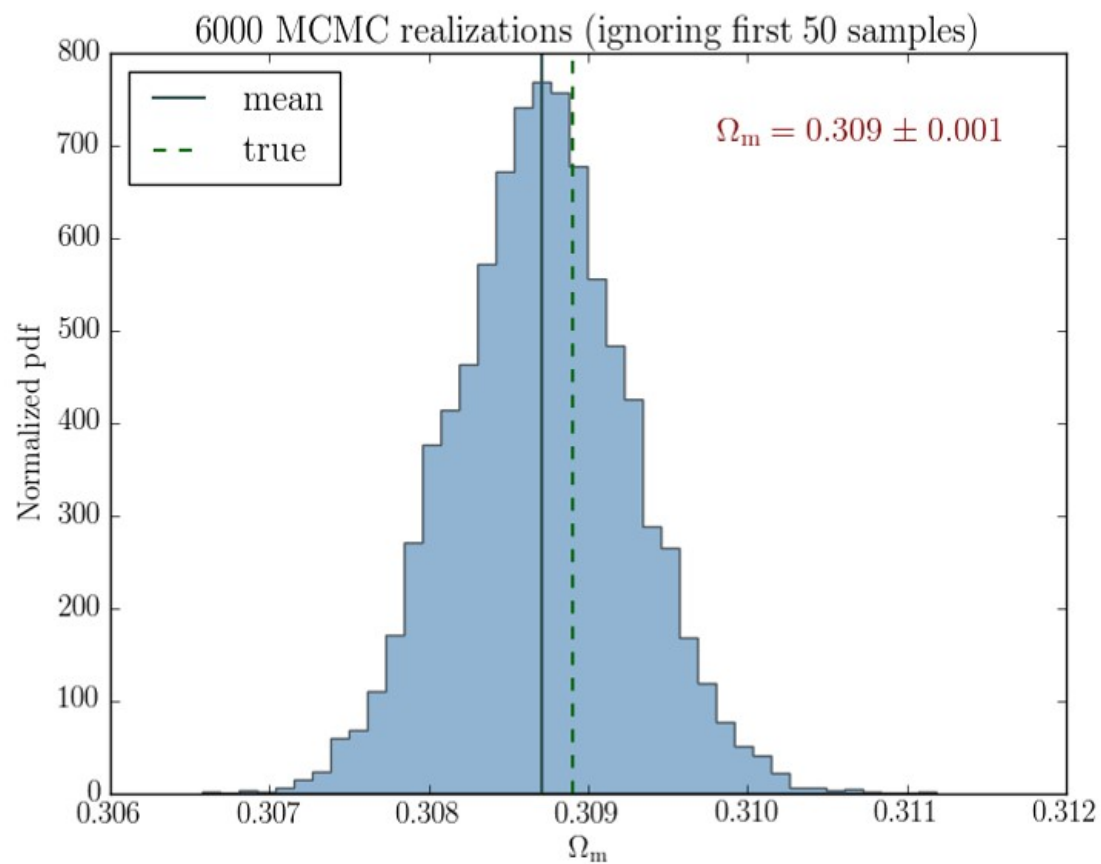
Final conditions



Per voxel error



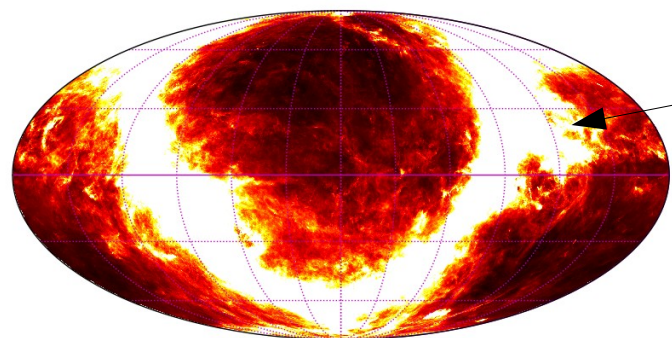
# Ultimate A/P: constraints on mean density



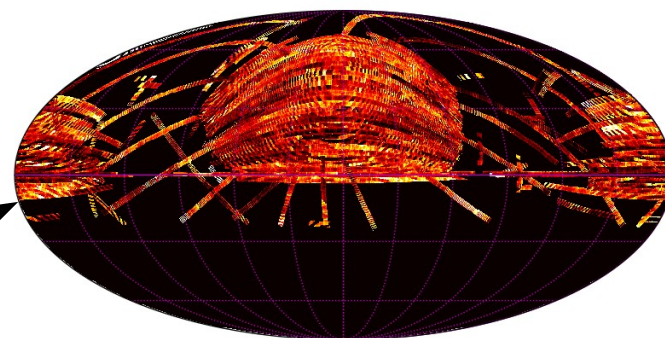


# ... and systematic cleaning ...

11 foregrounds (here only 8)... still much less than Leistedt & Peiris (2014) but improving

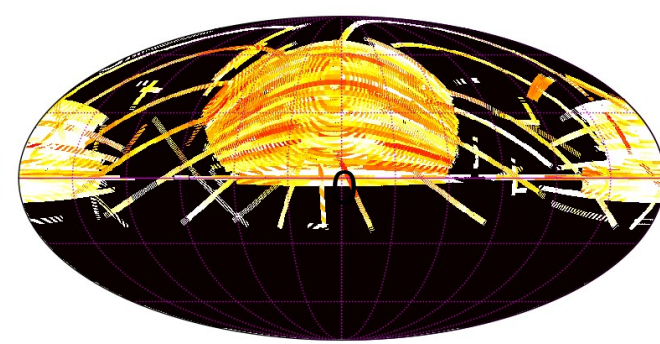
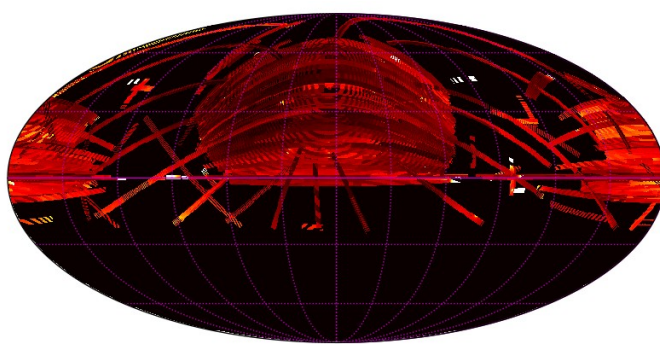
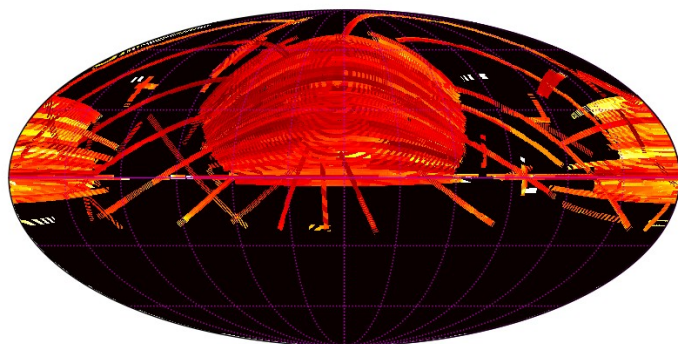


DUST



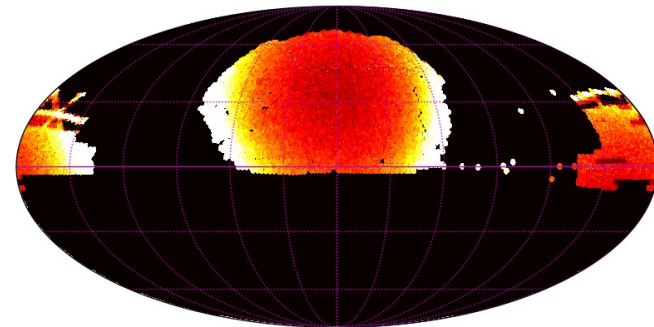
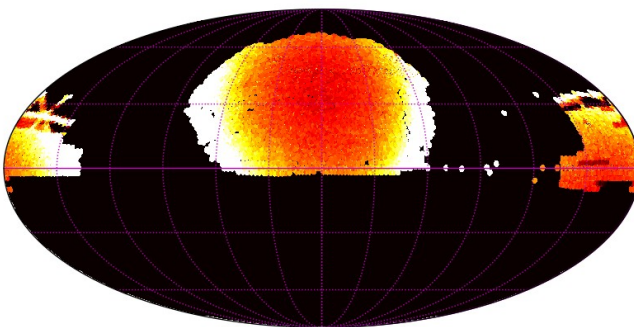
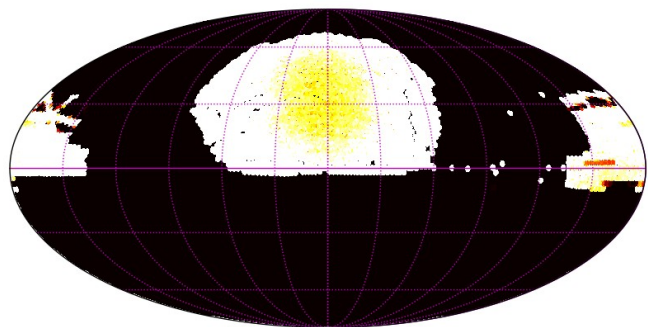
psfWidth

Sky fluxes



...

Star densities



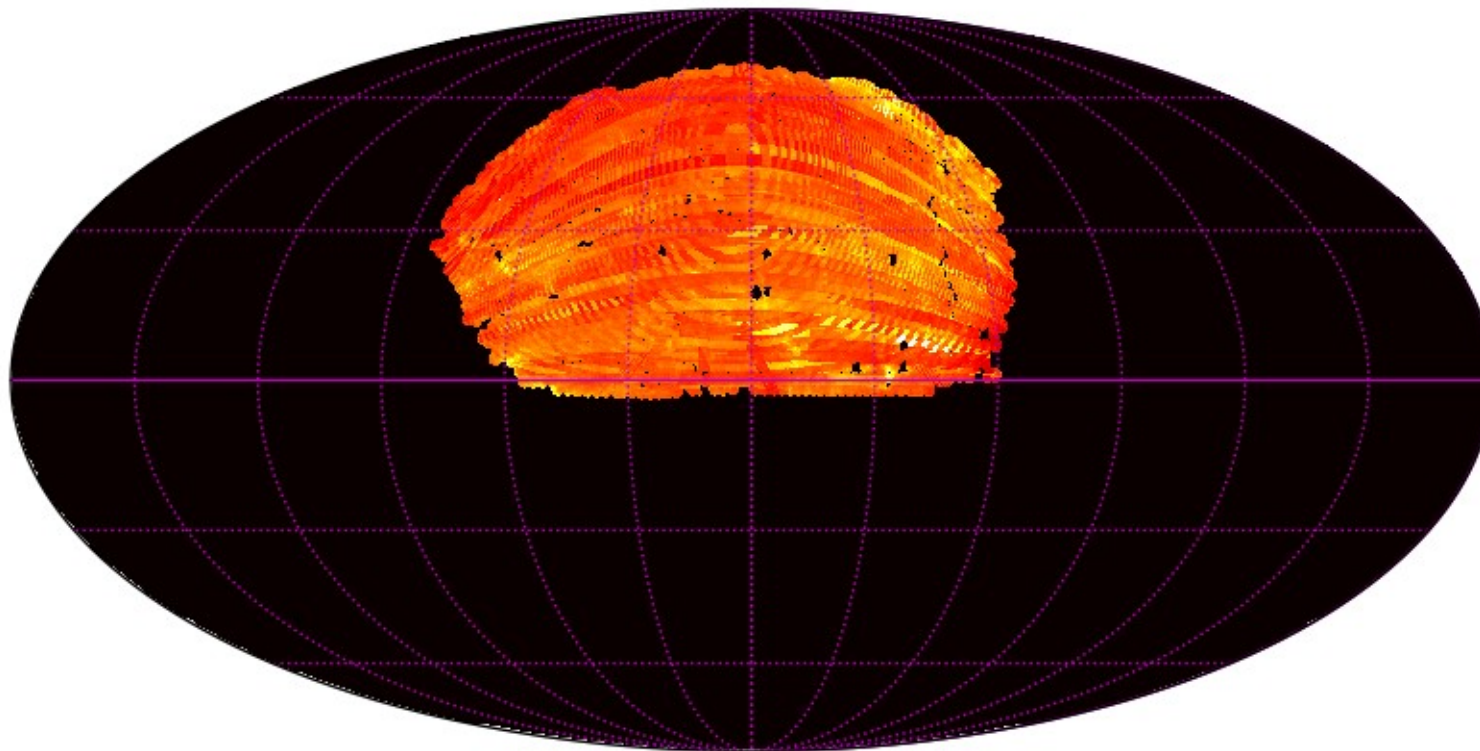
...

# Example fitted composite...

11 foregrounds (here only 8)... still much less than Leistedt & Peiris (2014) but improving

Sky

Star

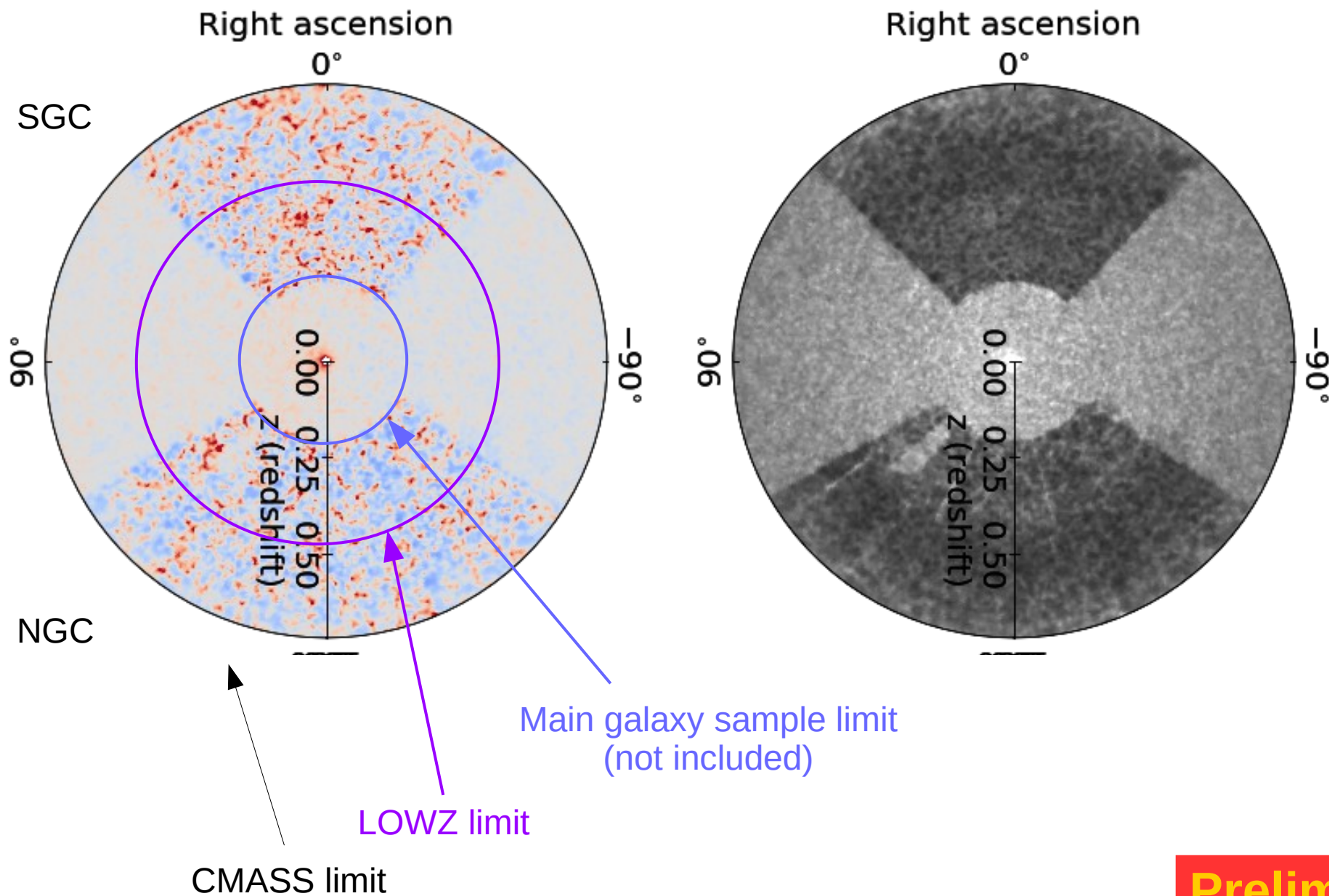


Preliminary

# Stage 1: Inferred density of SDSS3

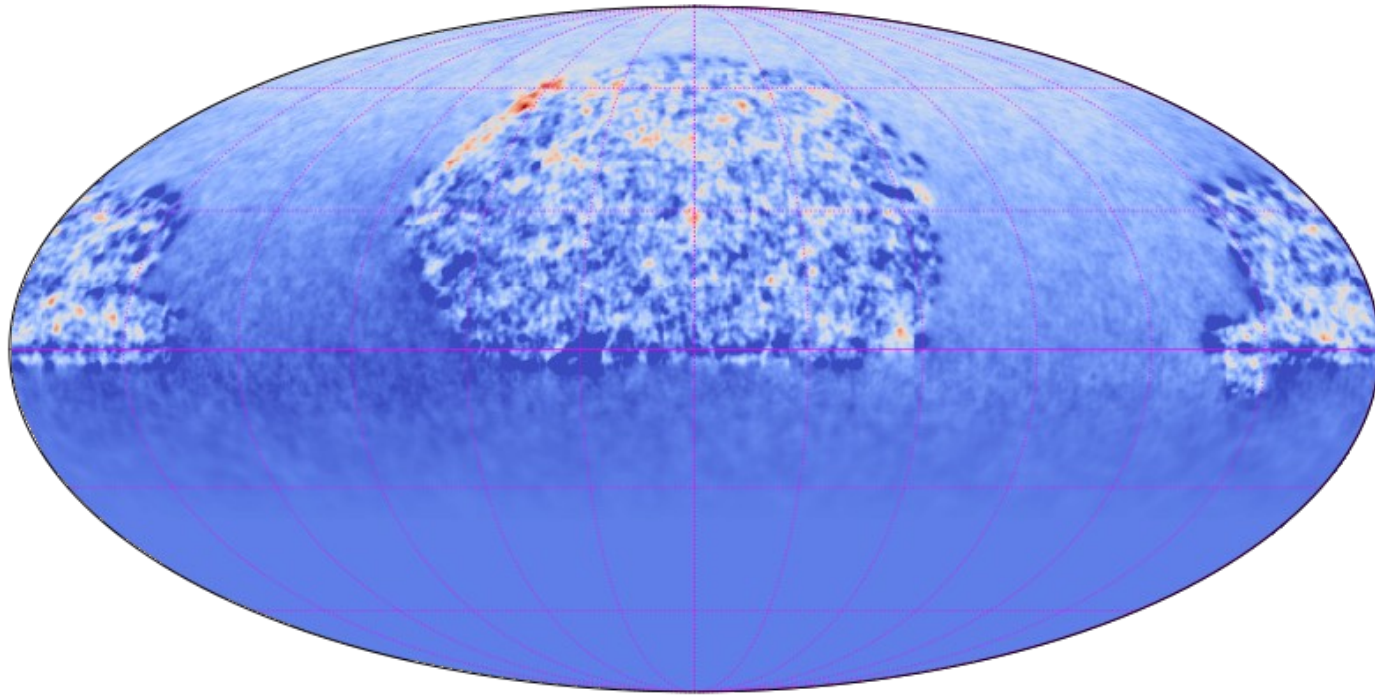
Ensemble density average

Error estimate from ensemble variance



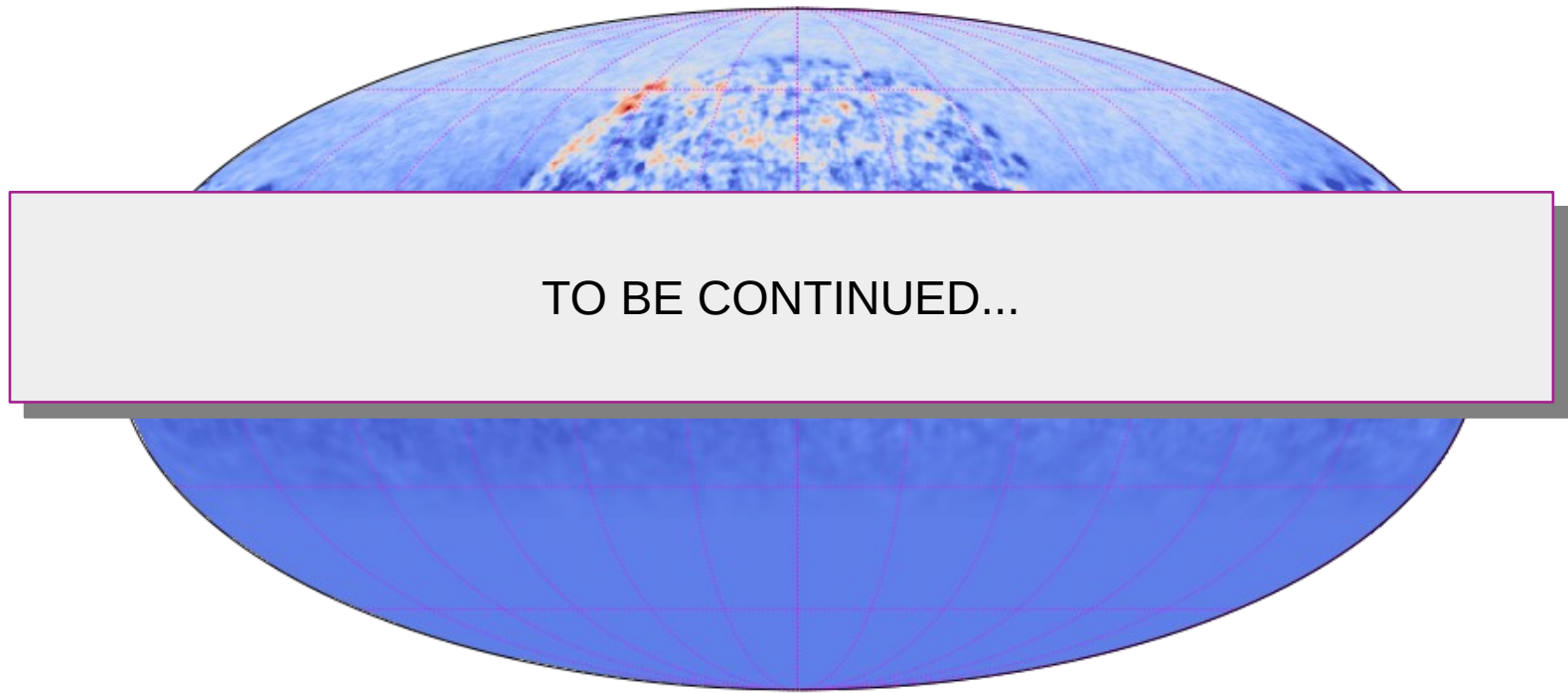
**Preliminary**

# Sky density



Preliminary

# Sky density



**Preliminary**

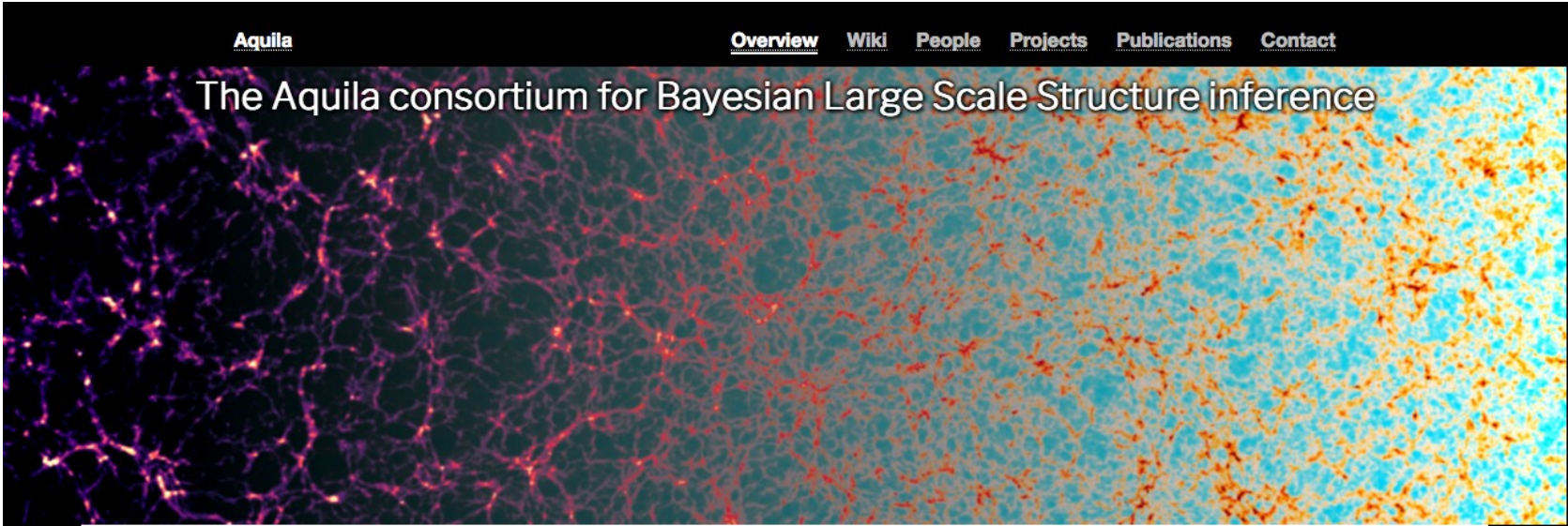
The background of the slide is an abstract, textured pattern. It features a dense network of thin, interconnected lines and shapes. The color palette is diverse, including deep purples, vibrant reds, and bright blues, all set against a dark, almost black background. The overall effect is reminiscent of a microscopic view of a complex material or a digital data visualization.

**Conclusion**

# The Aquila consortium

- Founded in 2016
- Gather people interested in working with each other on developing the Bayesian pipelines and run analysis on data.

<https://aquila-consortium.org/>



**Aquila**      Overview   Wiki   People   Projects   Publications   Contact

## The Aquila consortium for Bayesian Large Scale Structure inference

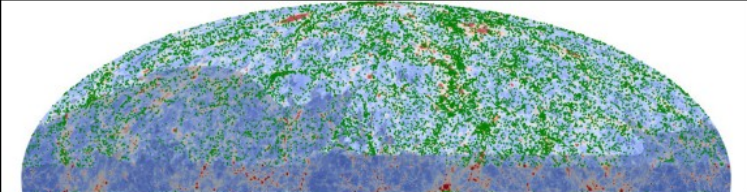
**Our mission: Data science meets the Universe**

The Aquila consortium is an international collaboration of researchers interested in developing and applying cutting-edge statistical inference techniques to study the spatial distribution of matter in our Universe. We embrace the latest innovations in information theory and artificial intelligence to optimally extract physical information from data and use derived results to facilitate new discoveries.

**Some results**

### Resimulating the Local Universe

To be updated. Copied from ILP. This picture shows the result of a high resolution N-body simulation which has been specifically designed to look like the Local Universe. More precisely it depicts what is the sky of an observer which would be located at the center of our galaxy and look at the entire sky. We use for that a Mollweide projection, which is another way of representing the surface of a full



# The Aquila consortium



- Founded in 2016
- Gather people interested in working with each other on developing the Bayesian pipelines and run analysis on data.

<https://aquila-consortium.org/>

**A biased list of Aquilians... check the website!**



Natalia Porqueres



Minh Nguyen



Doogesh Kodi Ramanah



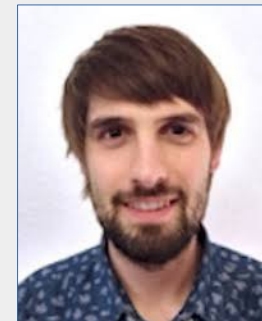
Tom Charnock



Harry Desmond



Franz Elsner



Florian Fuhrer

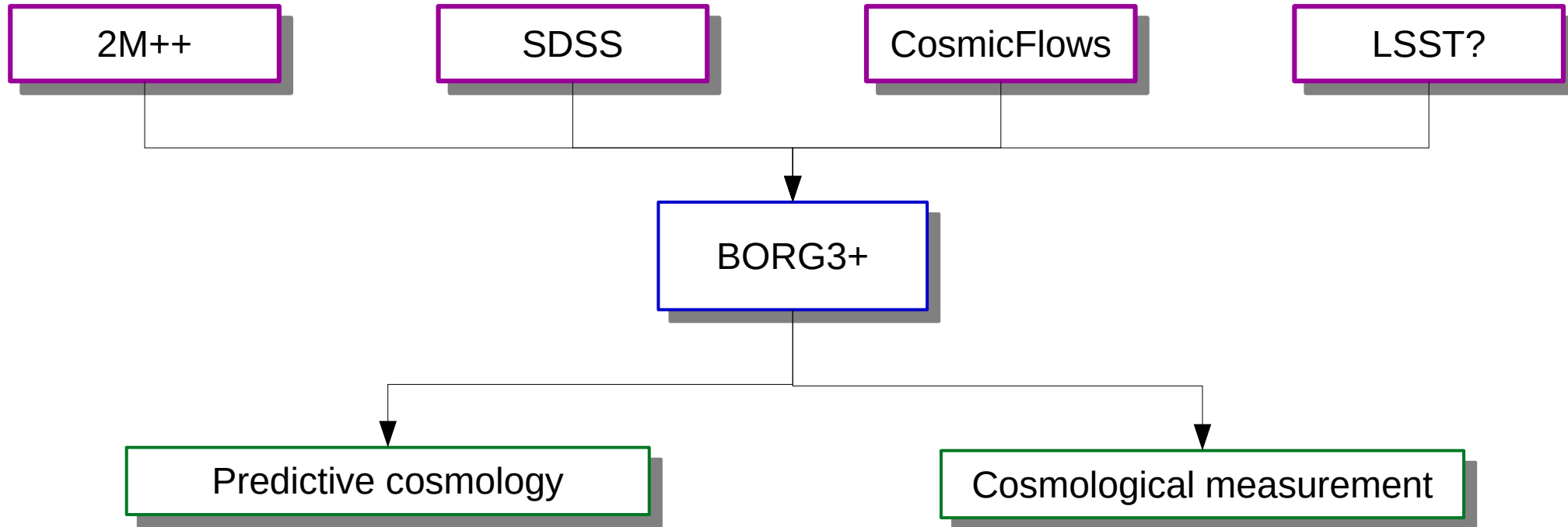


Florent Leclercq





# Conclusion: great future



- Velocity field (also VIRBIUS with F. Fuhrer)
  - X-ray cluster emission
  - Kinetic Sunyaev Zel'dovich
  - Rees-Sciama
  - Dark matter ?
- Cosmic expansion
  - Power spectrum (and governing parameters)
  - Gaussianity tests of initial conditions
  - Direct probe of dynamics

# Conclusion: great future and challenges

