Bayesian Forward Modelling of Galaxy Surveys

Creating the Universe's digital twin

First Nordic Cosmology Meeting
23–25 October 2023, Stockholm, Sweden

Jens Jasche
for
the Aquila Consortium

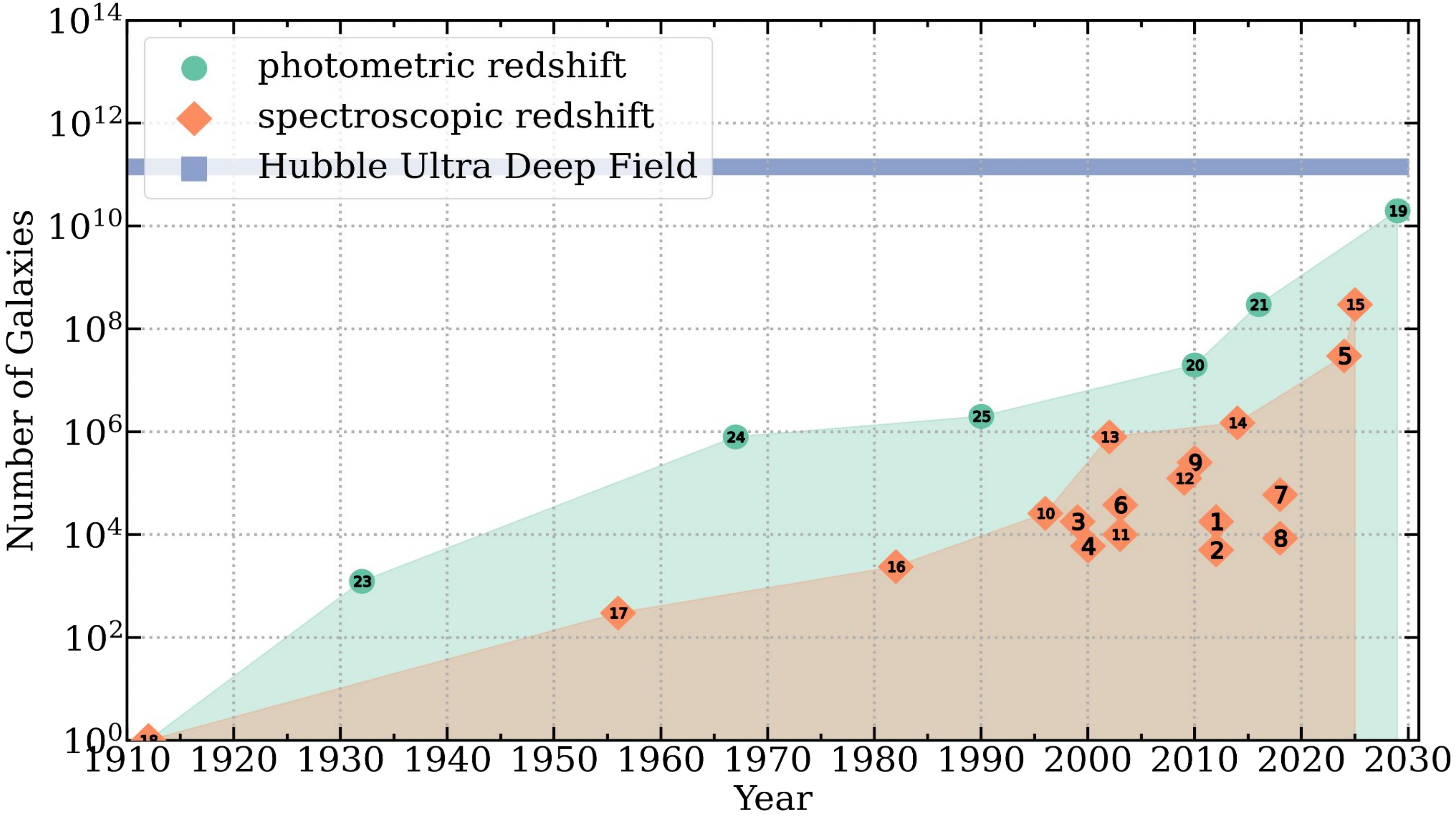
01

Motivation & Background

Bayesian forward modeling of
galaxy surveys

A complete characterization of the cosmic large scale structure

A CENTURY OF GALAXY SURVEYS

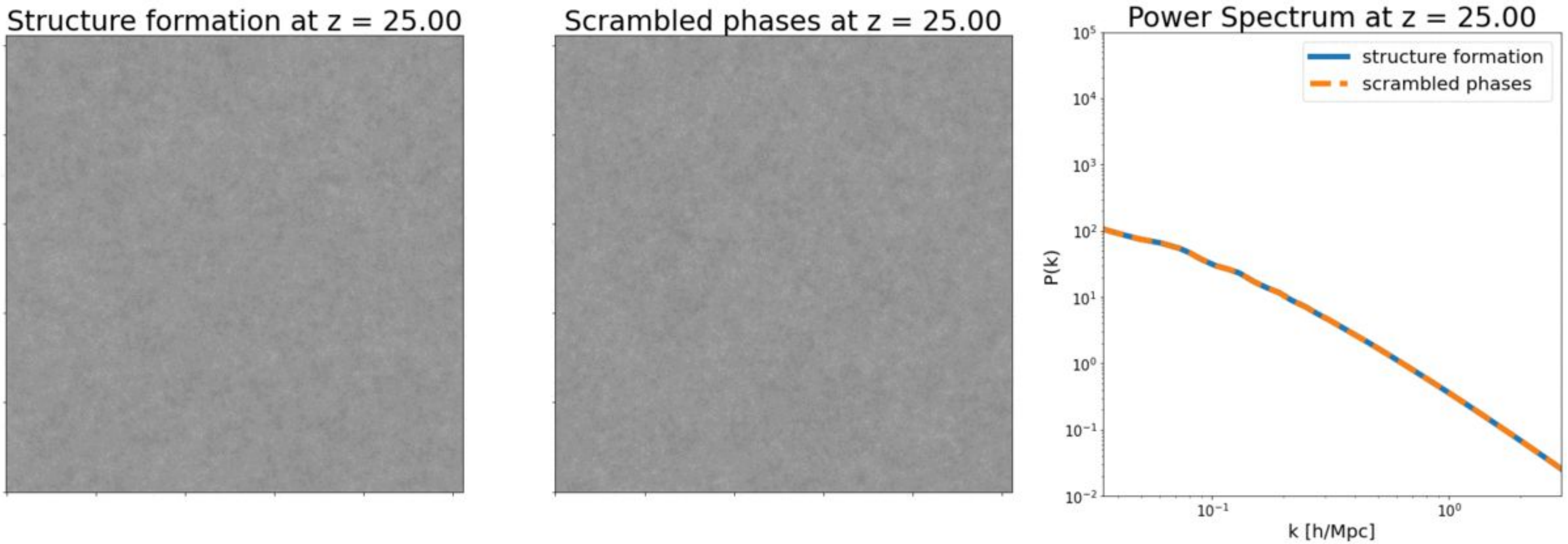


1 AGES
2 ACES
3 CfA2
4 CNOC2
5 DESI
6 DEEP2
7 DEVILS
8 DEIMOS
9 GAMA
10 LCRS
11 MGC
12 6dFGS
13 SDSS
14 BOSS
15 SPHERE-X
16 CfA
17 Humason
18 Slipher
19 LSST
20 WISE
21 DES
22 Las Campanas Redshift Survey
23 Shapley-Ames Catalog
24 Shane-Wirtanen Counts
25 APM

inspired by J. Peacock

The limits of summary statistics

Traditionally cosmic structures are analyzed via summary statistics (mostly 2PT).



There exists no closed form description of the 3D matter distribution:

- **The hierarchy of statistical moments is not guaranteed to close (e.g. [Carron & Neyrinck 2012](#))**

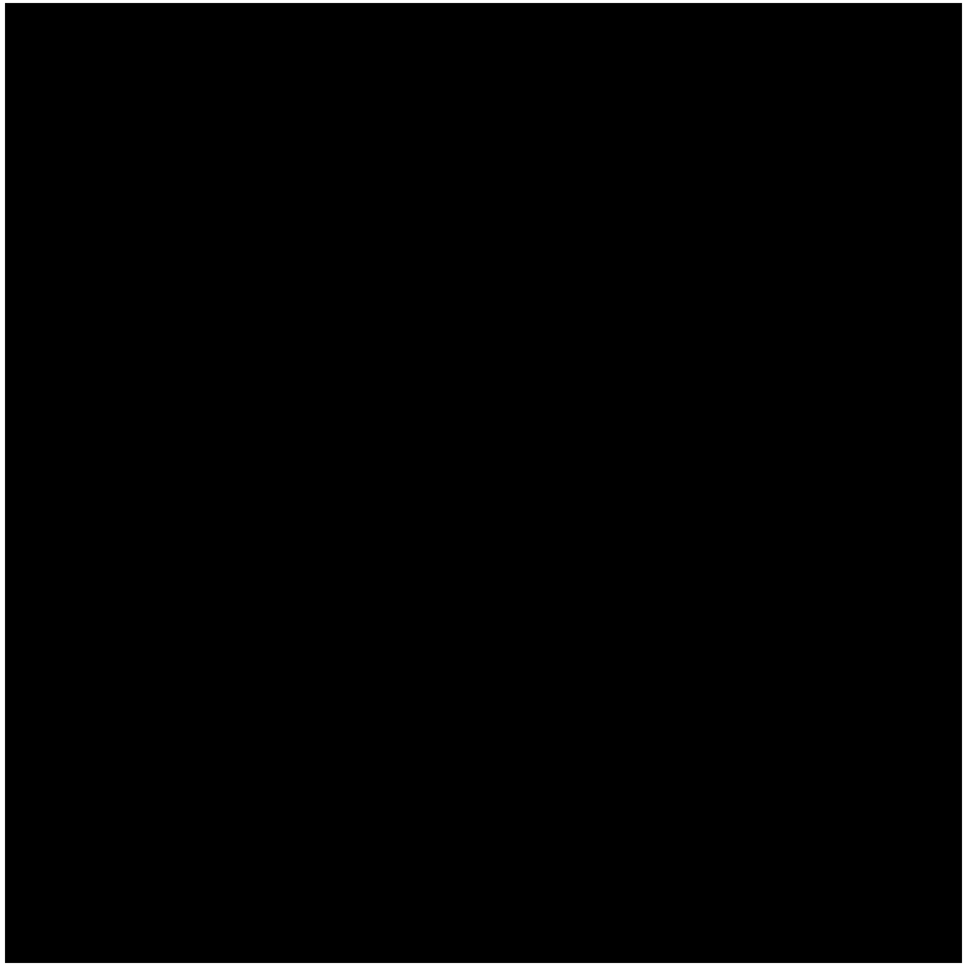
Should we rather use the entire field to do cosmology?

A complete characterization of the cosmic large scale structure

'What I cannot create, I do not understand.'
Richard P. Feynman, 1988

Bayesian Physical forward modeling

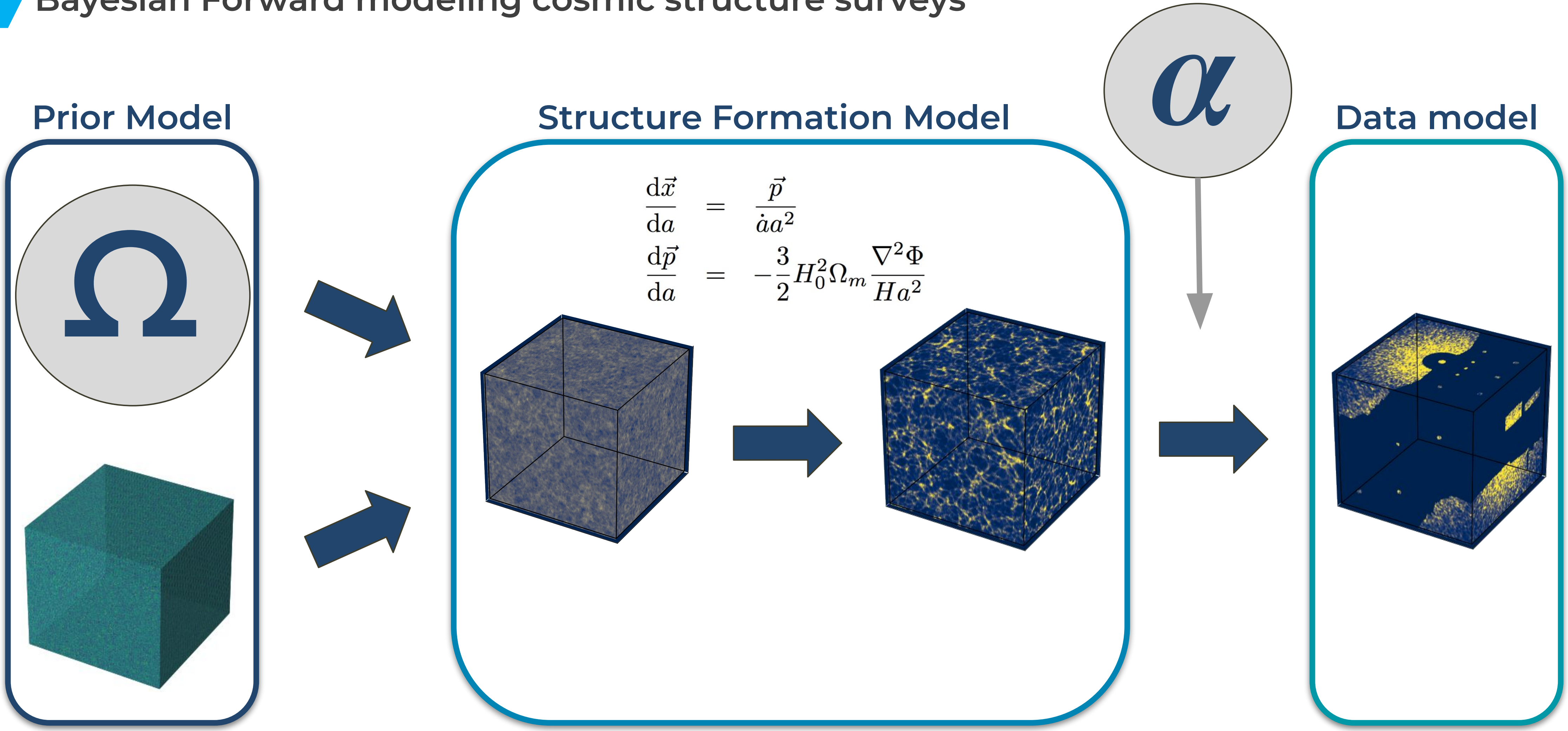
- **Field-level inference**
  - Beyond summary statistics
  - Beyond random realizations
- **Non-linear and dynamical inference**
  - Beyond linear structure growth
  - Redshift Distortions
  - Light-Cone effects
- **Causal inference**
  - Beyond associative analyses



Jasche & Lavaux 2019

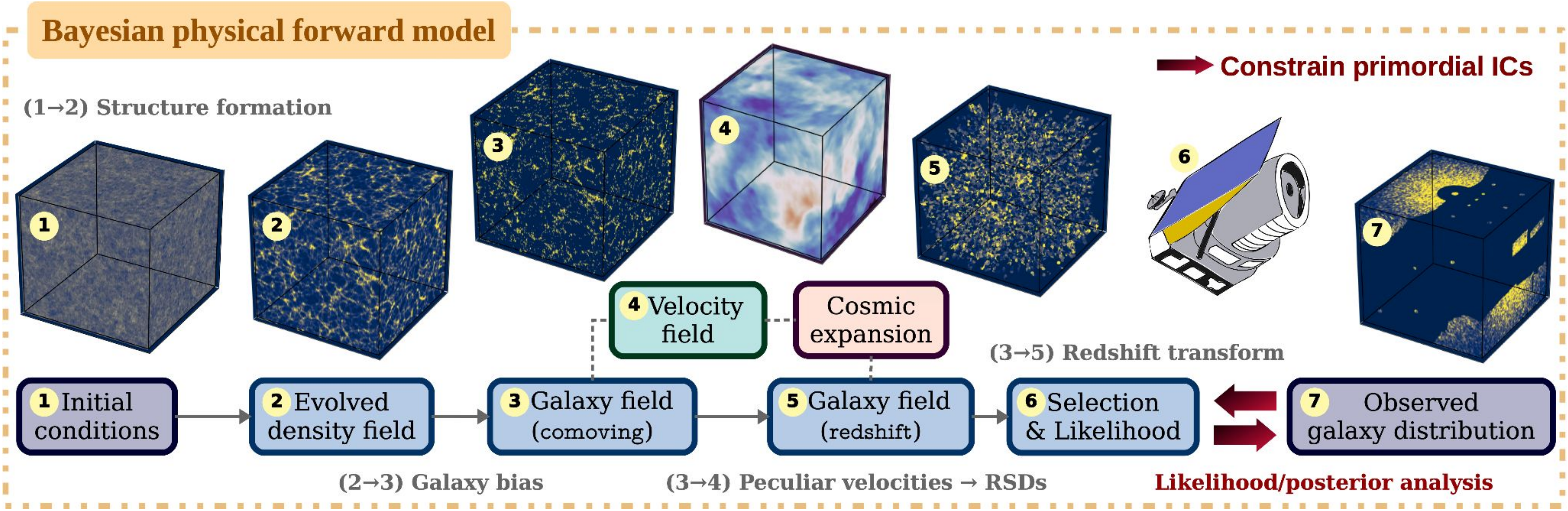
Aim: Create a Digital Twin of the Universe.

Bayesian Forward modeling cosmic structure surveys



$$\pi\left(\mathbf{x}, \boldsymbol{\Omega}\right) \qquad \pi\left(\rho_{\mathbf{m}} \middle| \mathbf{x}, \boldsymbol{\Omega}\right) \qquad \pi\left(\mathbf{N_g} \middle| \rho_{\mathbf{m}}, \alpha, \boldsymbol{\Omega}\right)$$

Bayesian Physical forward modeling

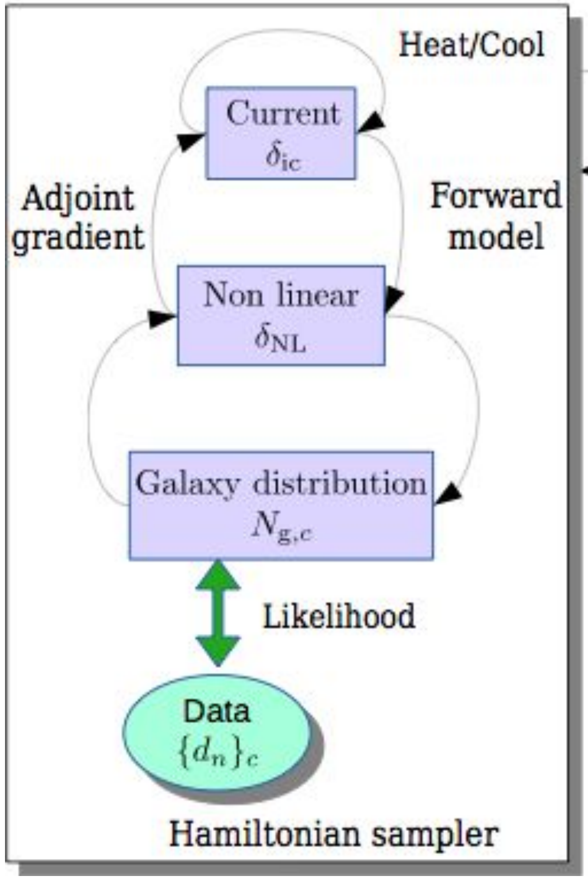


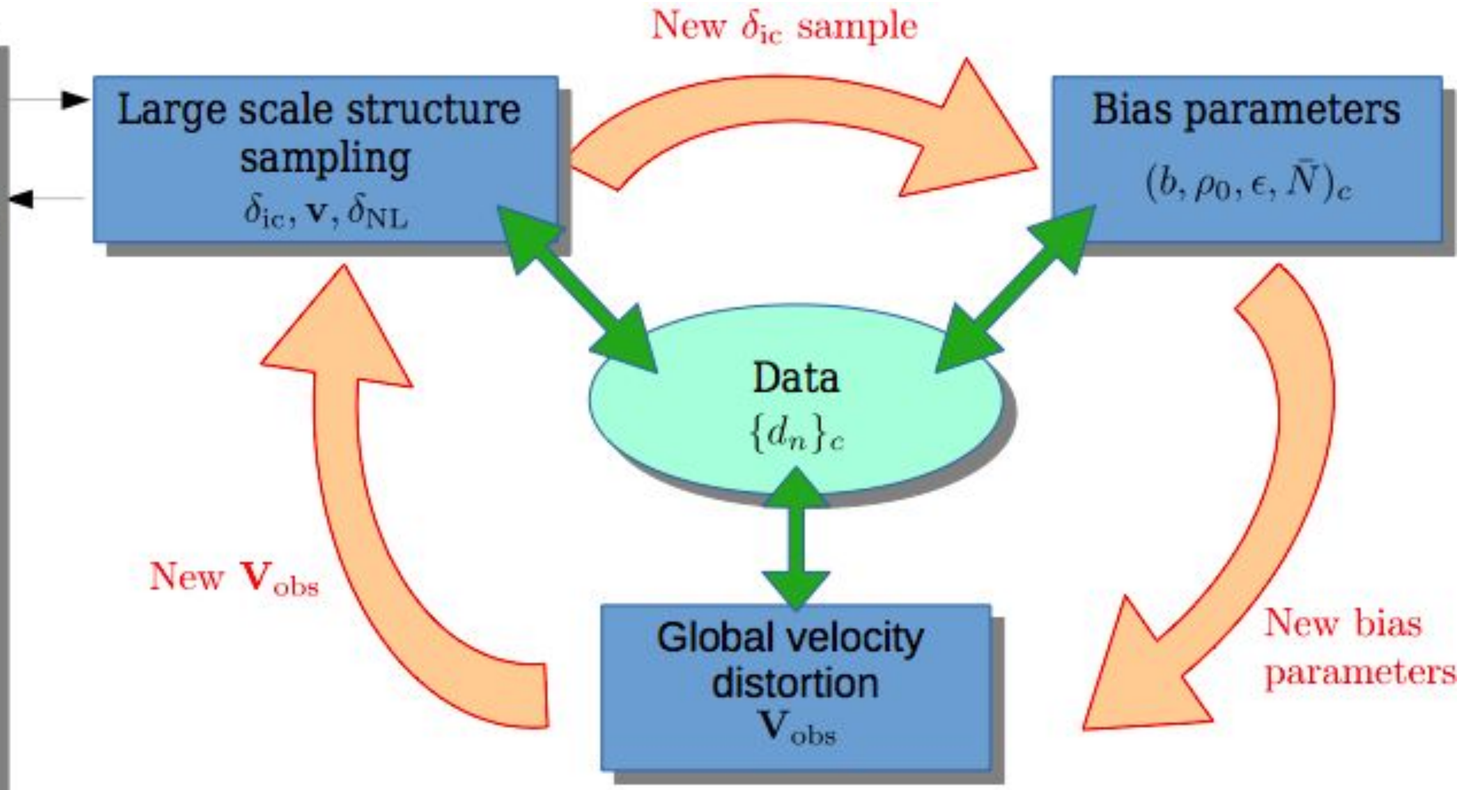
Now several groups are developing Forward Modeling approaches.

Jasche & Wandelt 2014
Wang et al 2014
Jasche, Leclercq, Wandelt 2015
Lavaux & Jasche 2016
Jasche & Lavaux 2019
Kitaura et al 2021
Ata et al 2022
Kostic et al. 2022

BORG's MCMC framework allows building flexible data models

- Hierarchical Bayes and block sampling
- Efficient Hamiltonian Monte Carlo technique
- **Fully differentiable physics forward model**

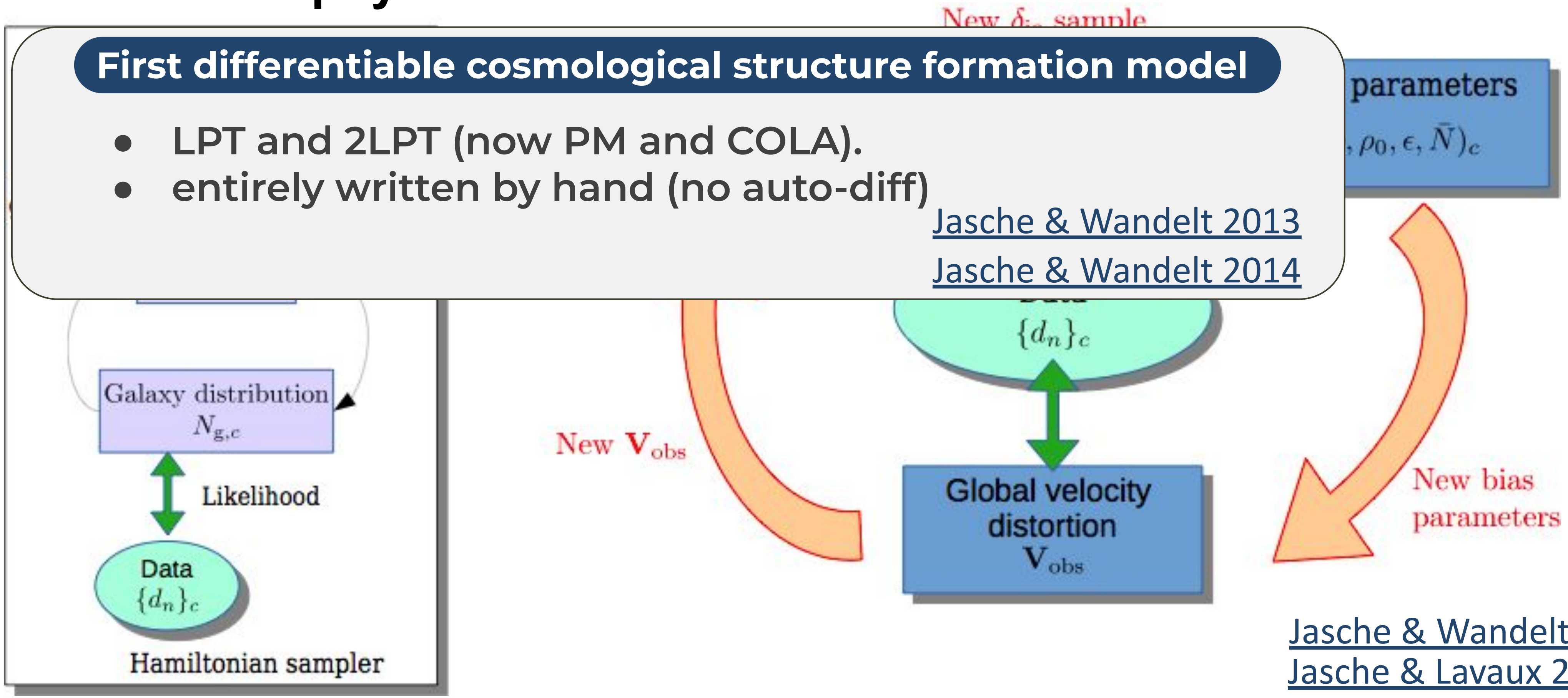




Jasche & Wandelt 2014
Jasche & Lavaux 2019

BORG's MCMC framework allows building flexible data models

- Hierarchical Bayes and block sampling
- Efficient Hamiltonian Monte Carlo technique
- **Fully differentiable physics forward model**



First differentiable cosmological structure formation model

- LPT and 2LPT (now PM and COLA).
- entirely written by hand (no auto-diff)

Jasche & Wandelt 2013
Jasche & Wandelt 2014

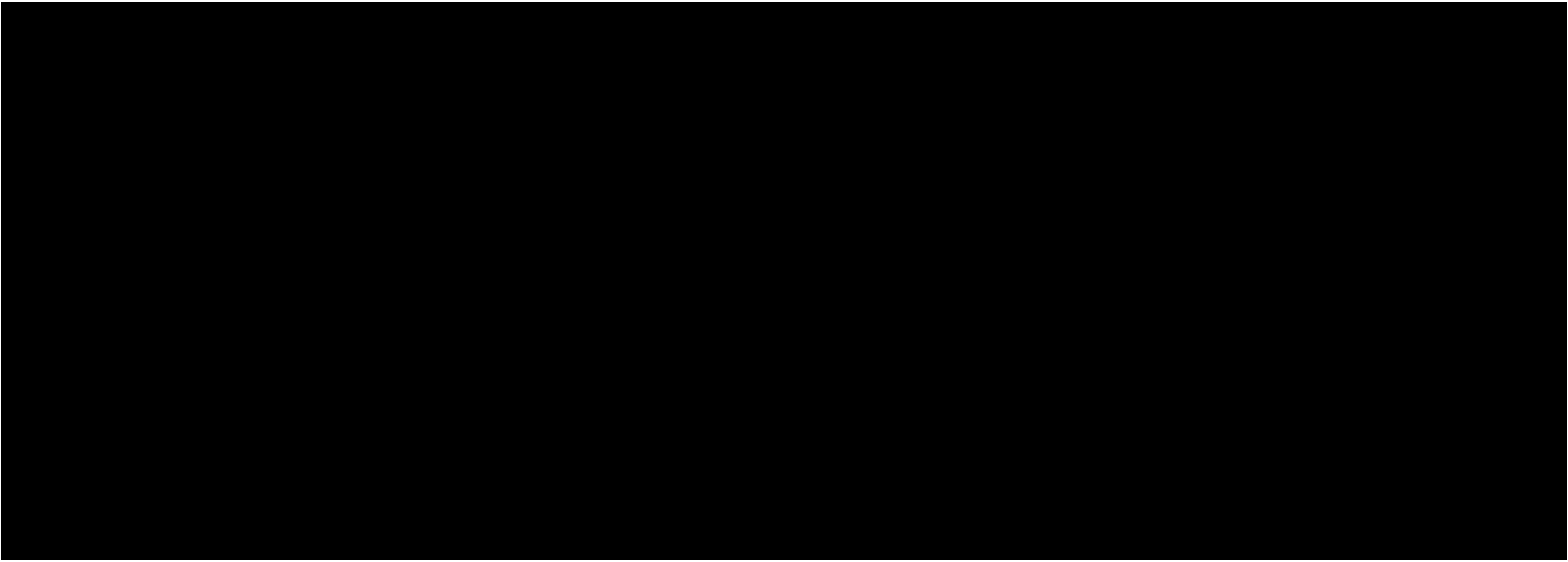
Jasche & Wandelt 2014
Jasche & Lavaux 2019

02

Non-linear inference of the Nearby Universe

Past and Present cosmic structures in the 2M++ galaxy compilation

A numerical cosmic structure posterior distribution



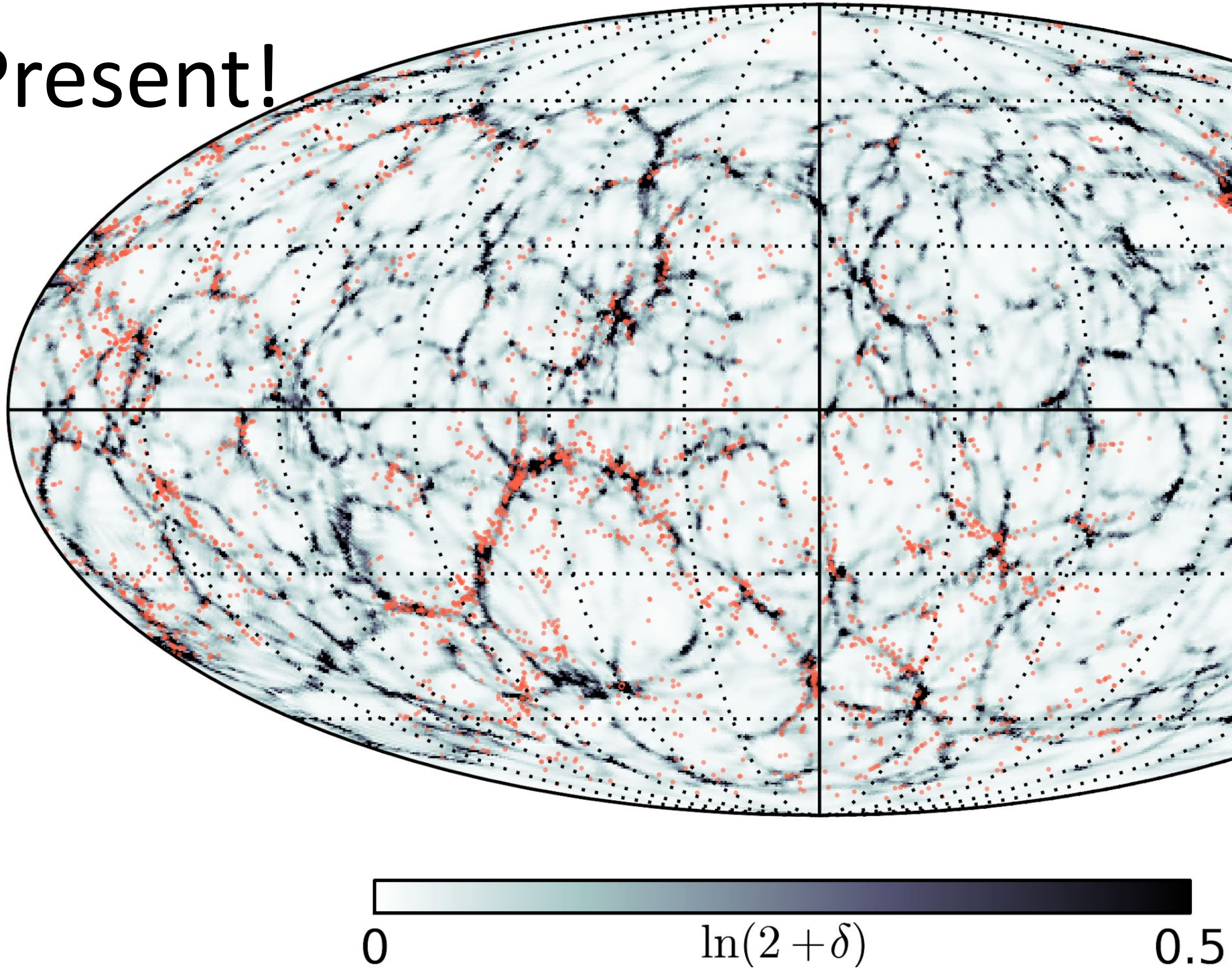
Movie credit: Florent Leclercq

Leclercq et al 2017

Jasche & Lavaux 2019

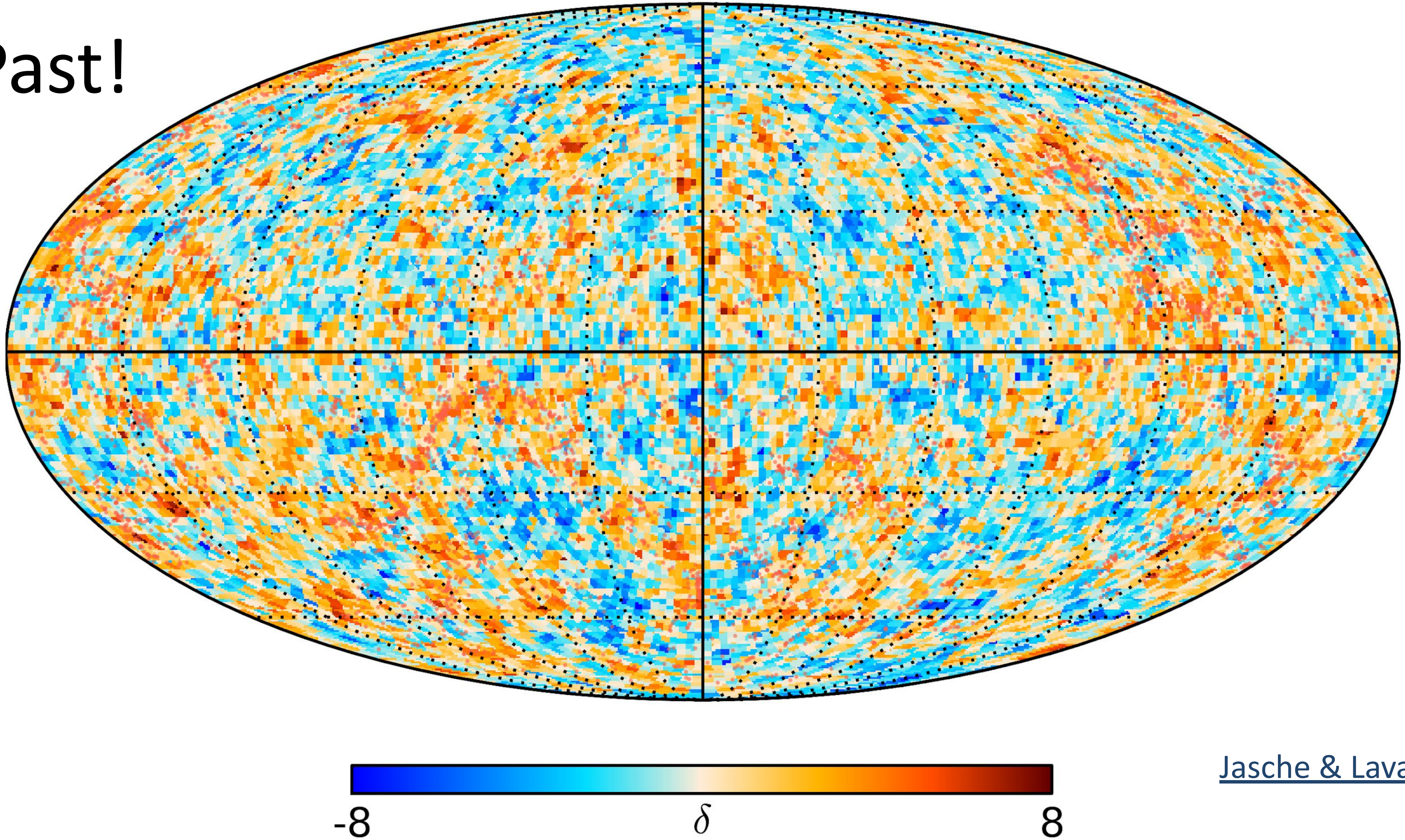
Bayesian forward modeling of the 2M++ galaxy compilation

The Present!



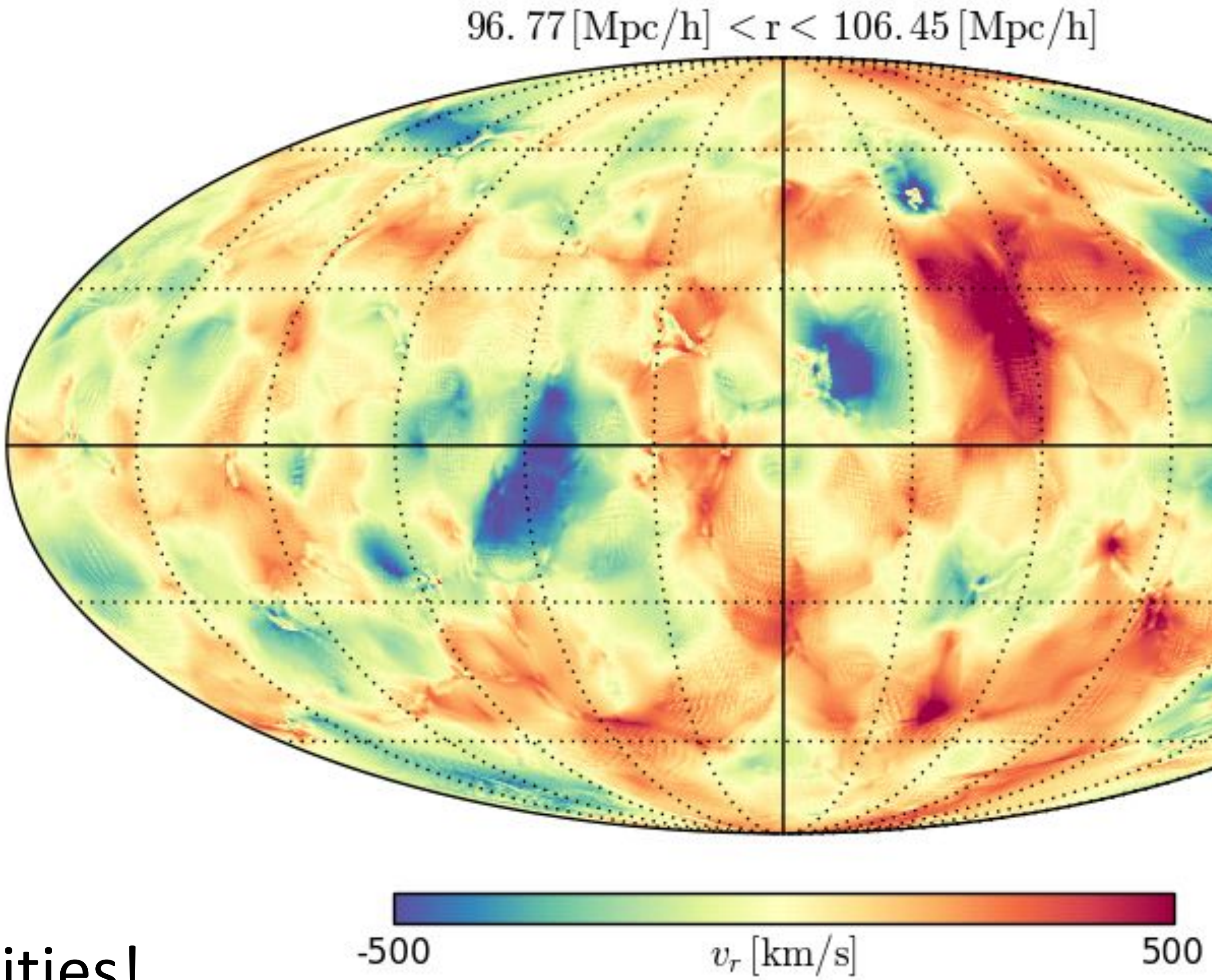
Jasche & Lavaux 2019

The Past!



Jasche & Lavaux 2019

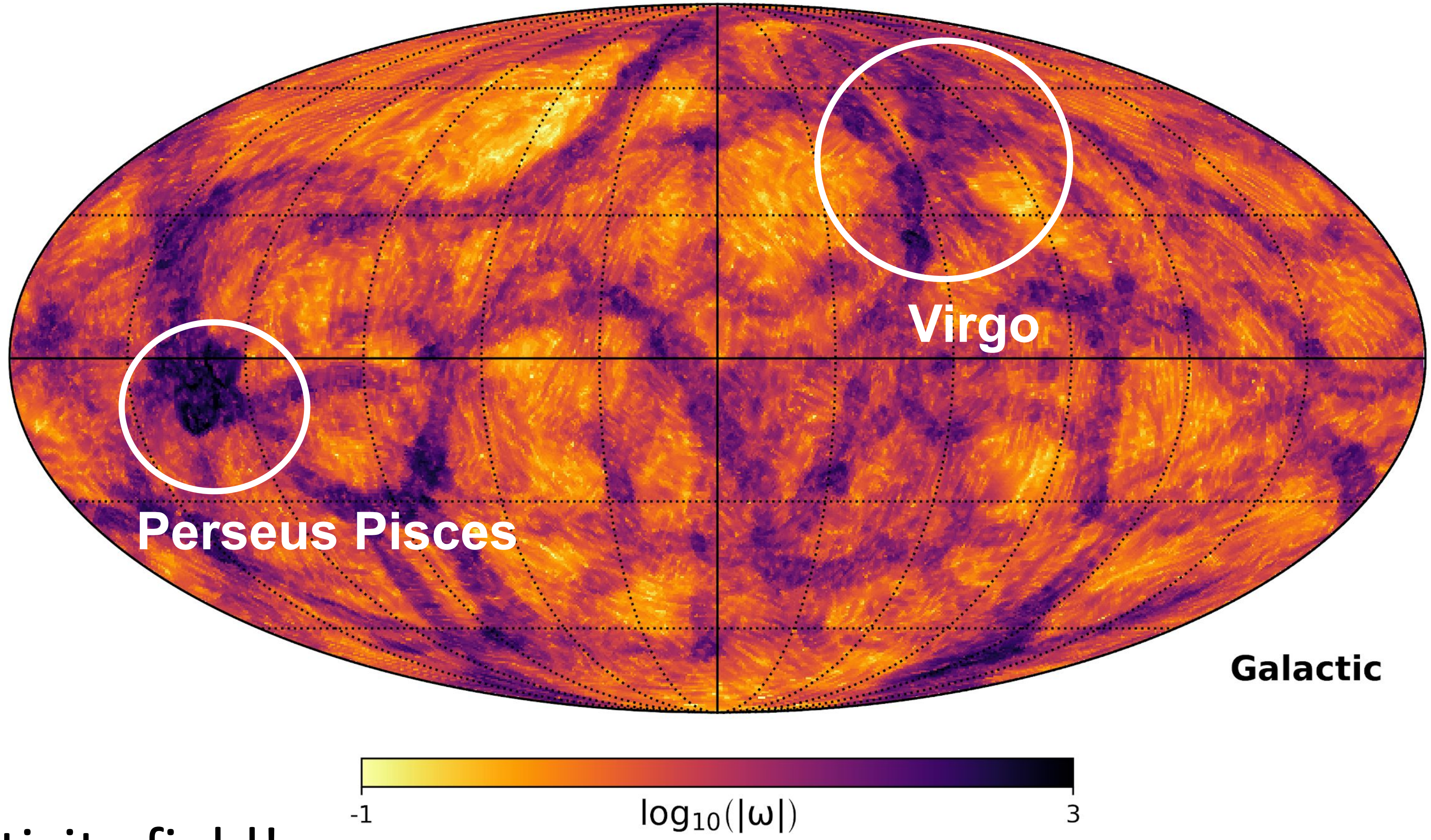
Bayesian forward modeling of the 2M++ galaxy compilation



Peculiar velocities!

Jasche & Lavaux 2019

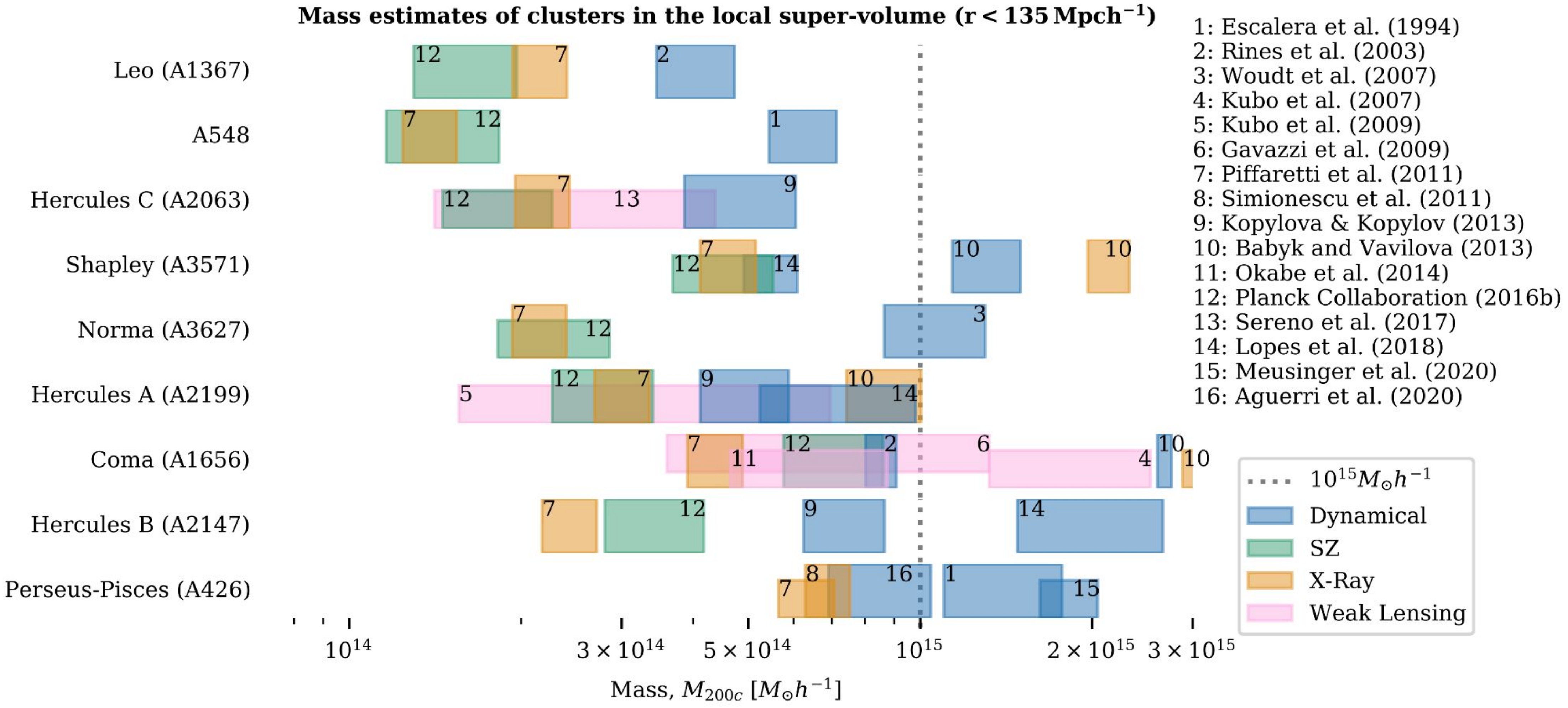
Bayesian forward modeling of the 2M++ galaxy compilation



Velocity vorticity field!

Jasche & Lavaux 2019

How rare is the local super volume?



An intriguing situation: Is the local super-volume compatible with ΛCDM or not?

Stopyra et al 2021

How rare is the local super volume?

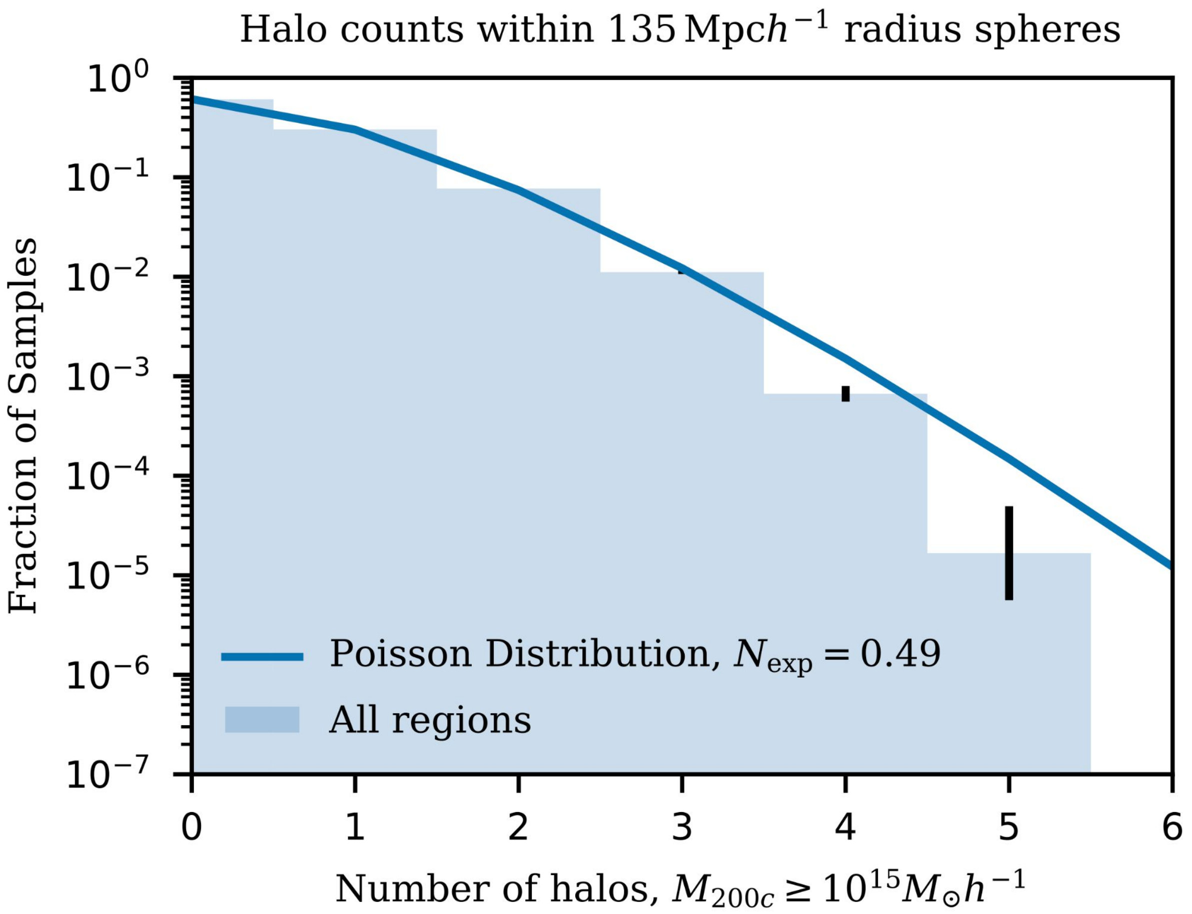
Prevalence of most massive nearby halos:

- Likelihood of finding N clusters with $M_{200c} \geq 10^{15} M_\odot h^{-1}$:

$$\mathcal{L}(N|N_{\text{exp}}) = \frac{N_{\text{exp}}^N \, e^{-N_{\text{exp}}}}{N!}$$

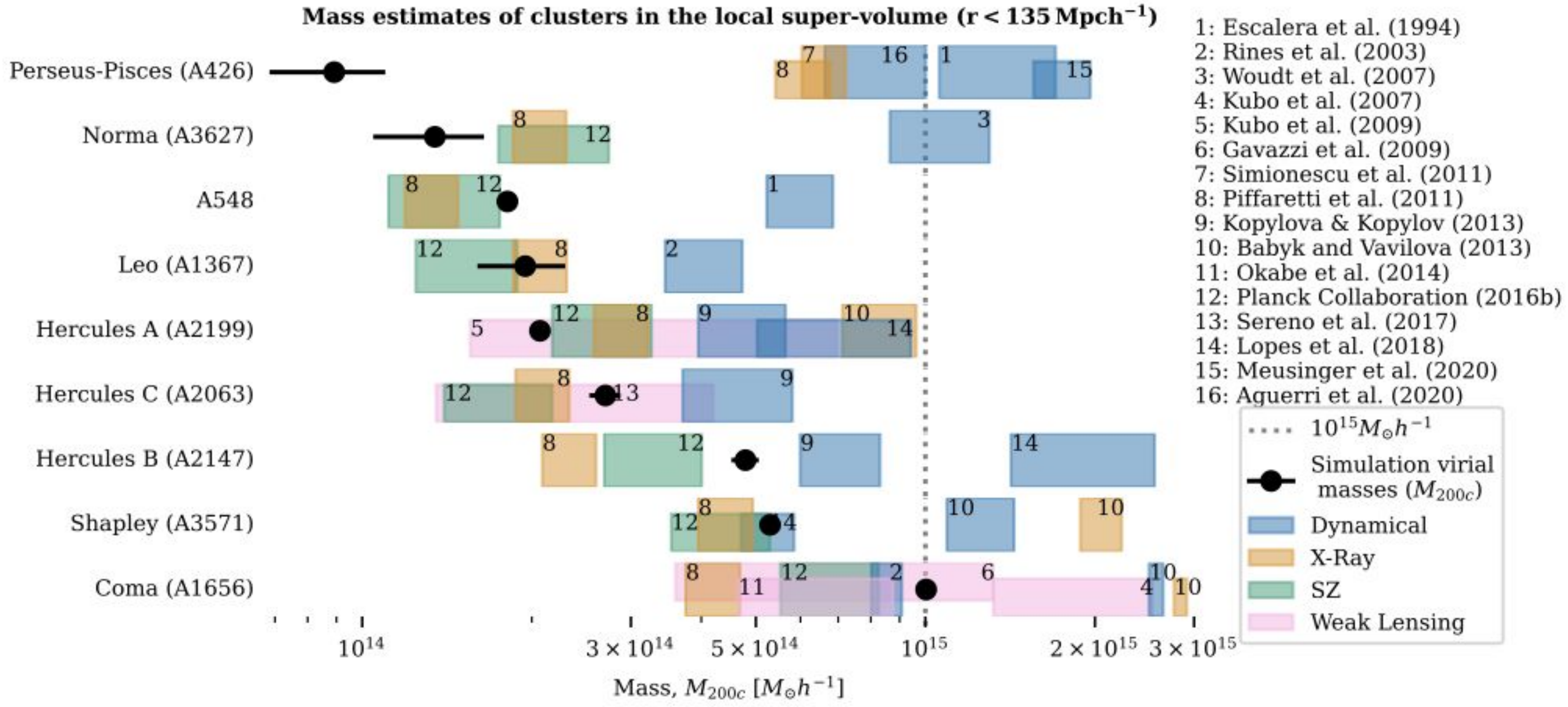
with N_{exp}= 1.

- A few high-mass halos can pose a significant challenge to ΛCDM. [Frenk et al 1990]



[Stopyra et al 2021]

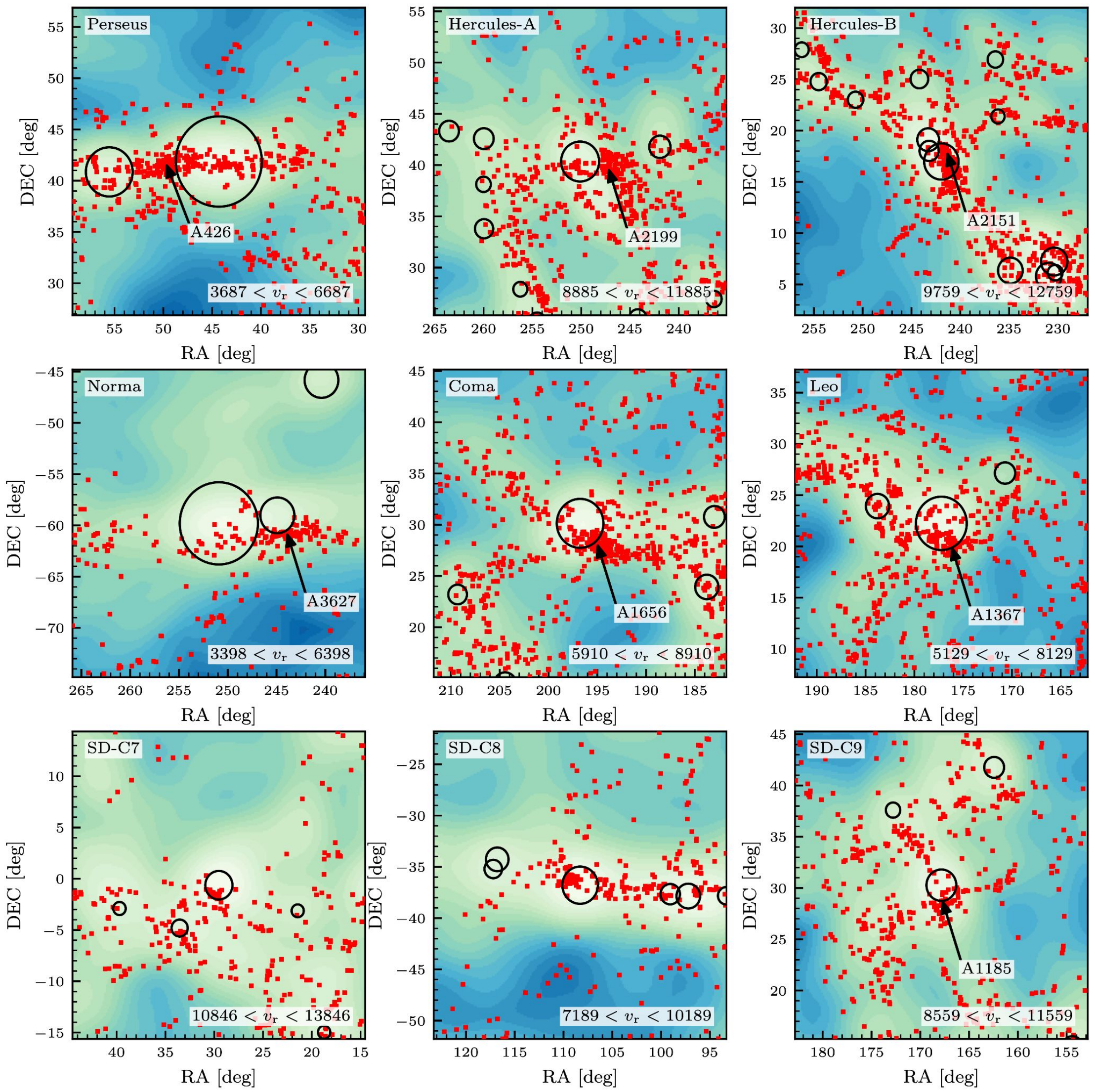
How rare is the local super volume?



Stopyra et al 2023

How rare is the local super volume?

- **Posterior mean density field**
  - Intricate detail
  - Recover non-linear structure
- **Improvements**
  - Sensitivity to cosmic variance
    - environment
    - rotation
  - No simplifying assumptions
  - Plausible formation history
- **Potential applications**
  - Test galaxy formation
  - ML training data
    - e.g. learn cluster members



McAlpine et al 2022

03

Testing cosmology

inferring cosmological parameters

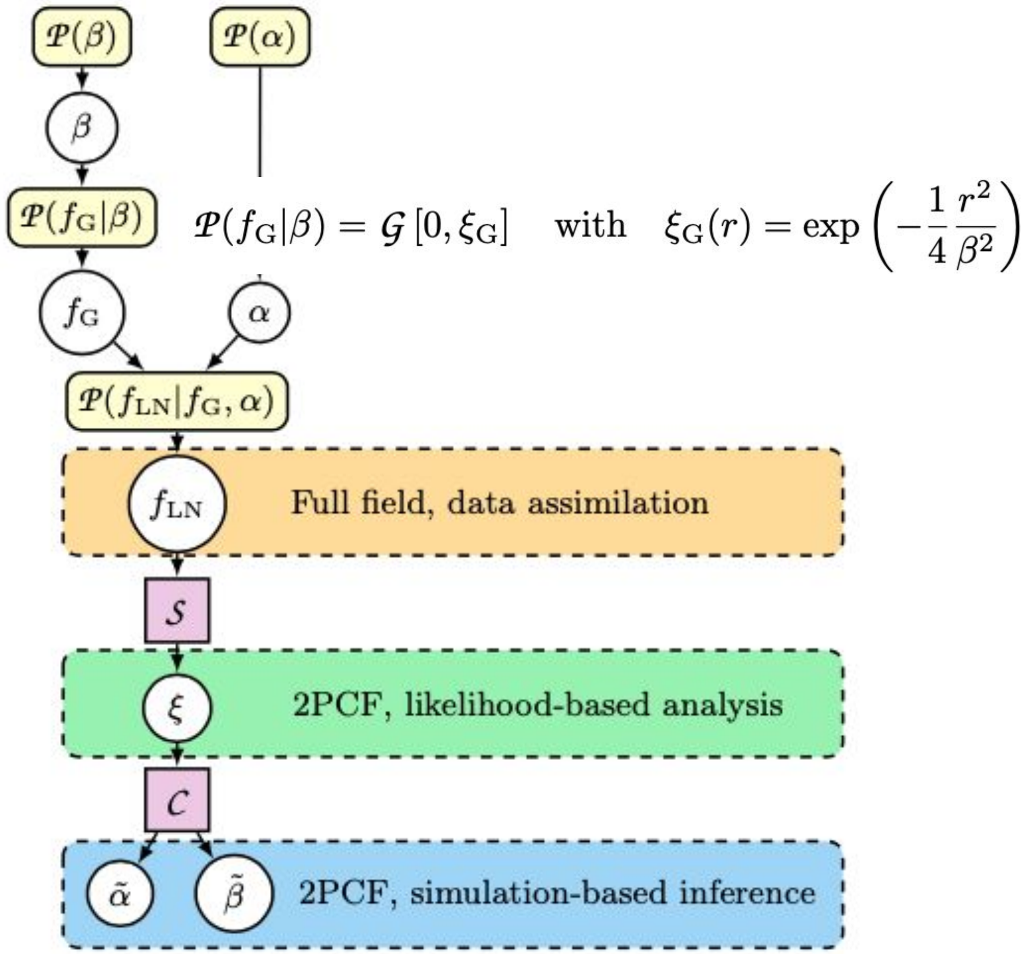
What analysis method to choose?

Compare 3 standard approaches:

- **Standard likelihood-based analysis (LBA)** of the two-point correlation function (2PCF), assuming a Gaussian distribution and fixed covariance matrix

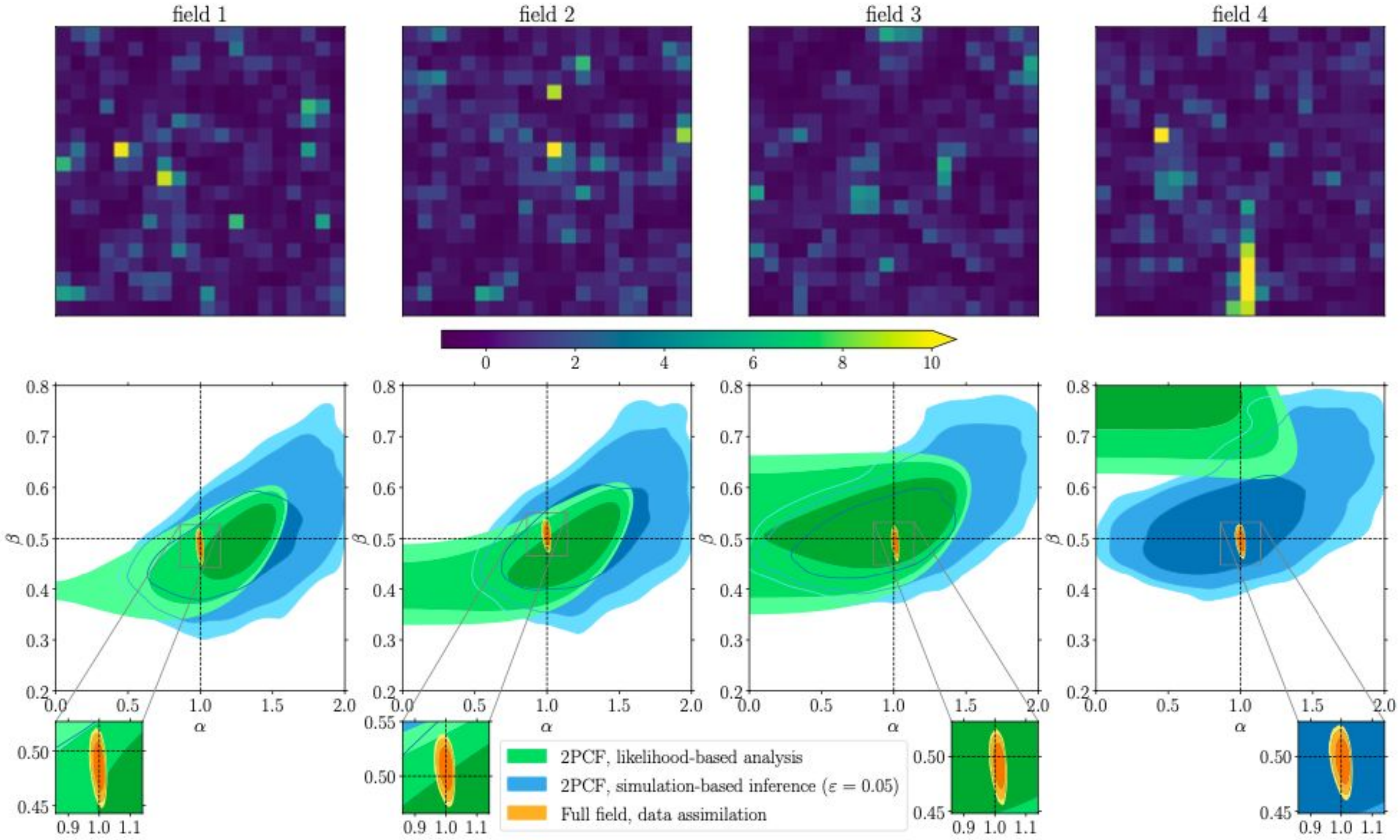
- **Simulation-based inference (SBI)**, aka likelihood-free inference, ABC, based on the 2PCF

- **Field level data assimilation (DA)** technique, e.g. Bayesian forward modeling, BORG



Leclercq & Heavens 2021

Beyond summary statistics: Field level inference

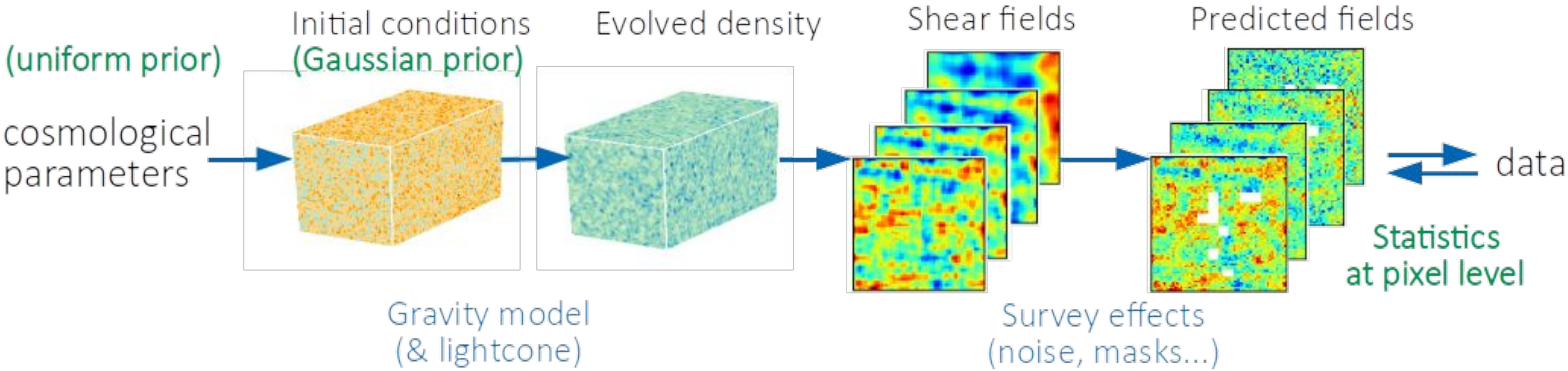


Field level inference uses many more constraints of the data! (higher-order statistics)

Leclercq & Heavens 2021

(Also see McQuinn (2021))

Sampling cosmology with BORG-WL (Natalia Porqueres)

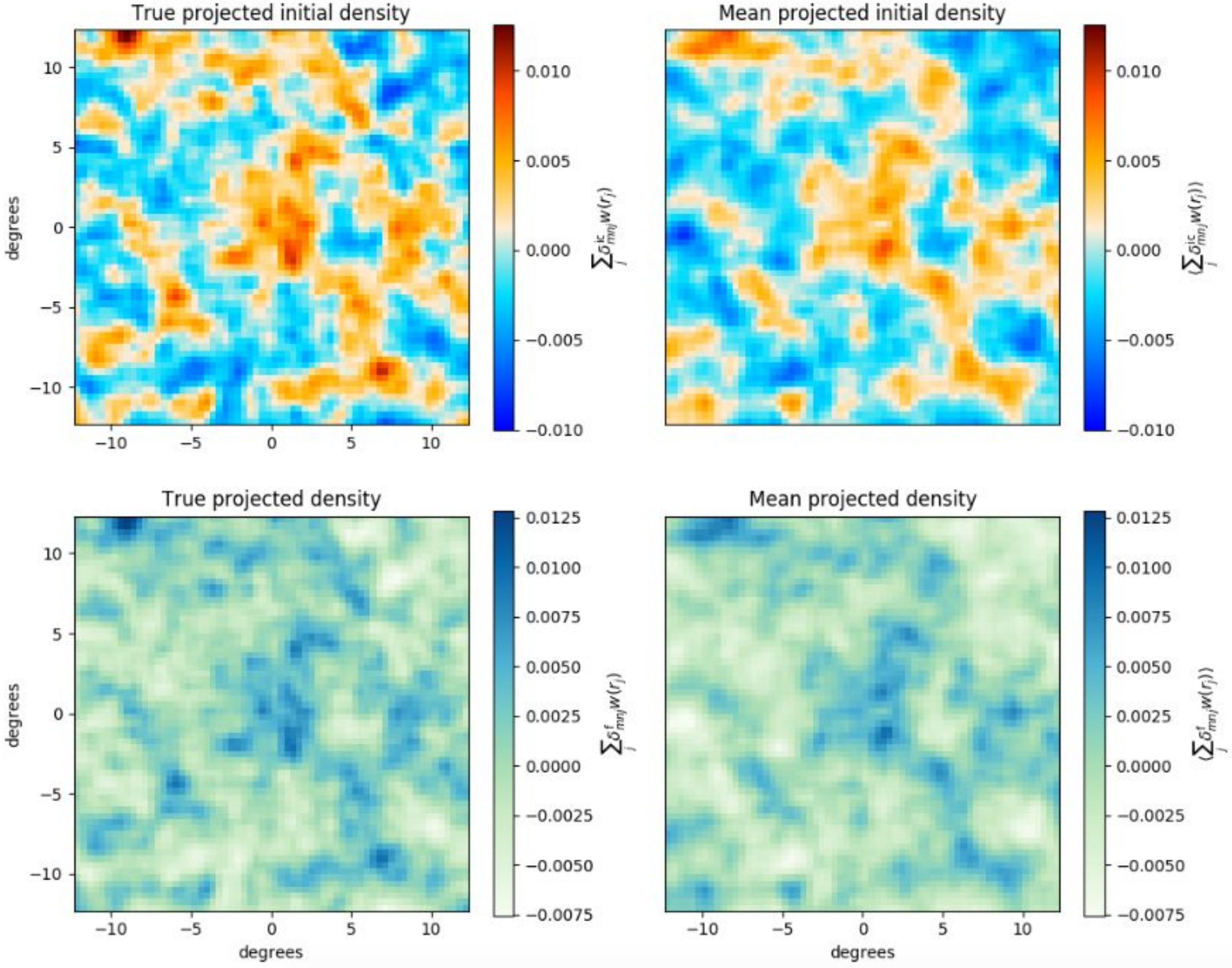


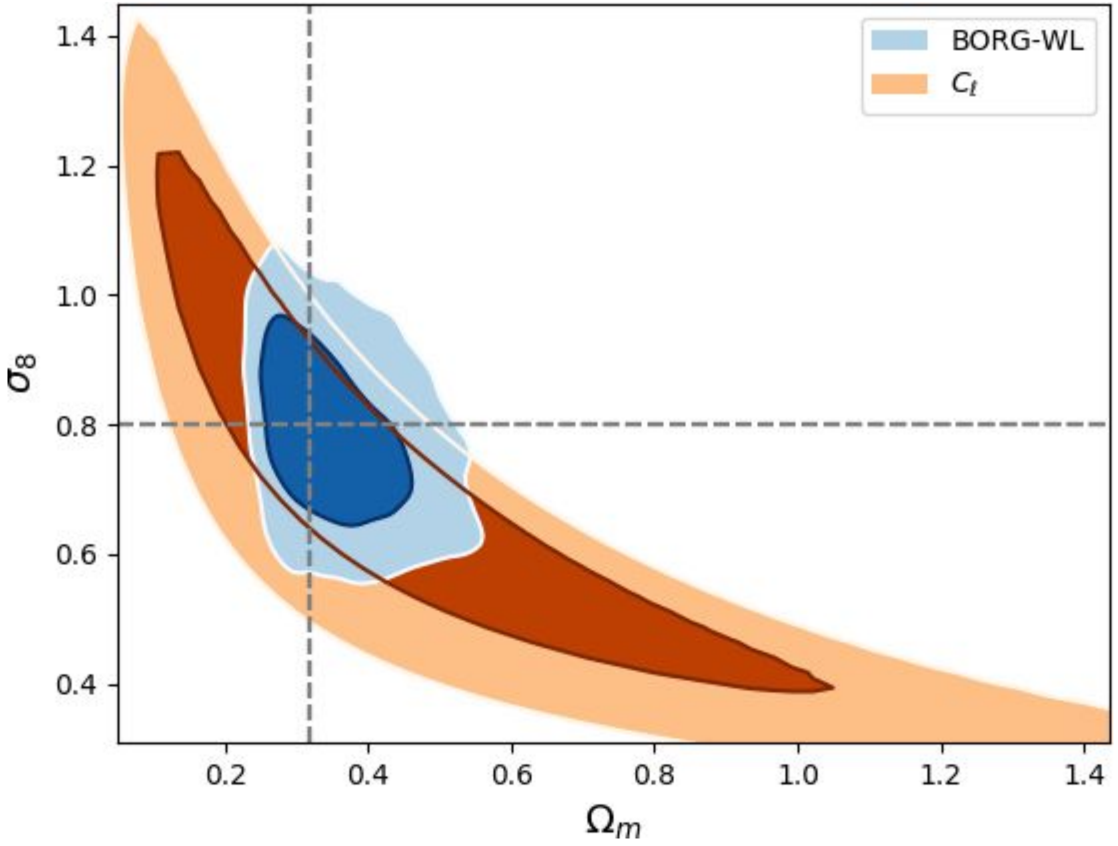
Cosmology updated consistently throughout the forward model, varying:

- initial matter power spectrum
- gravity model
- geometry (line-of-sight integration)

Porqueres et al. 2022

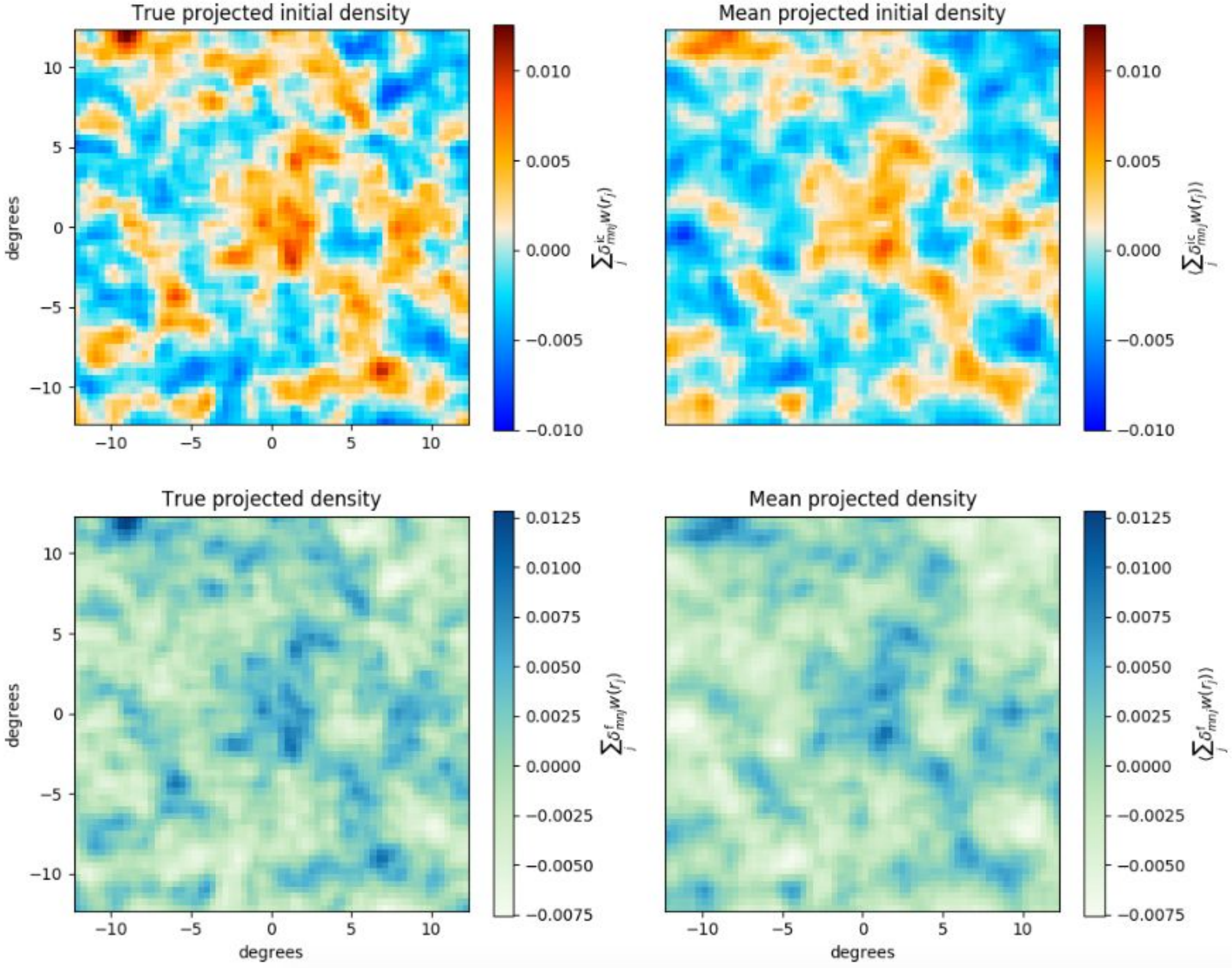
Sampling cosmology with BORG-WL (Natalia Porqueres)

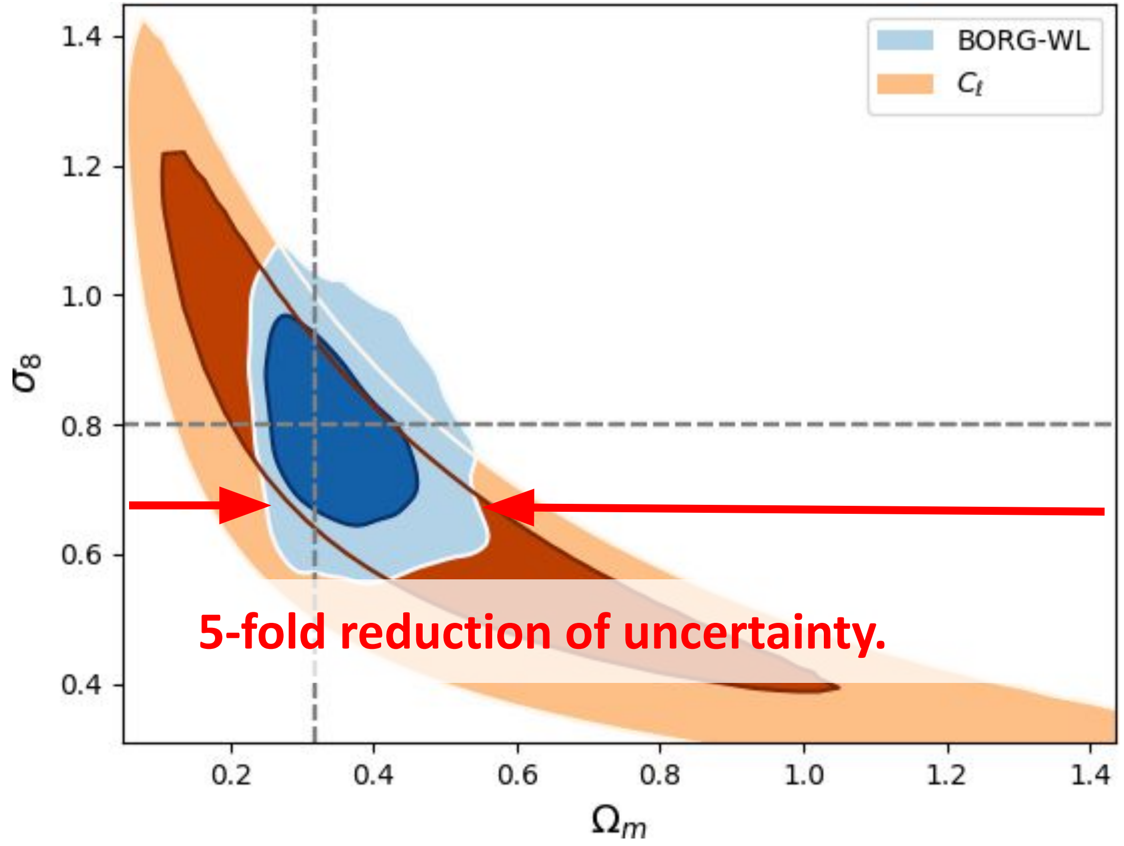




Porqueres et al. 2022

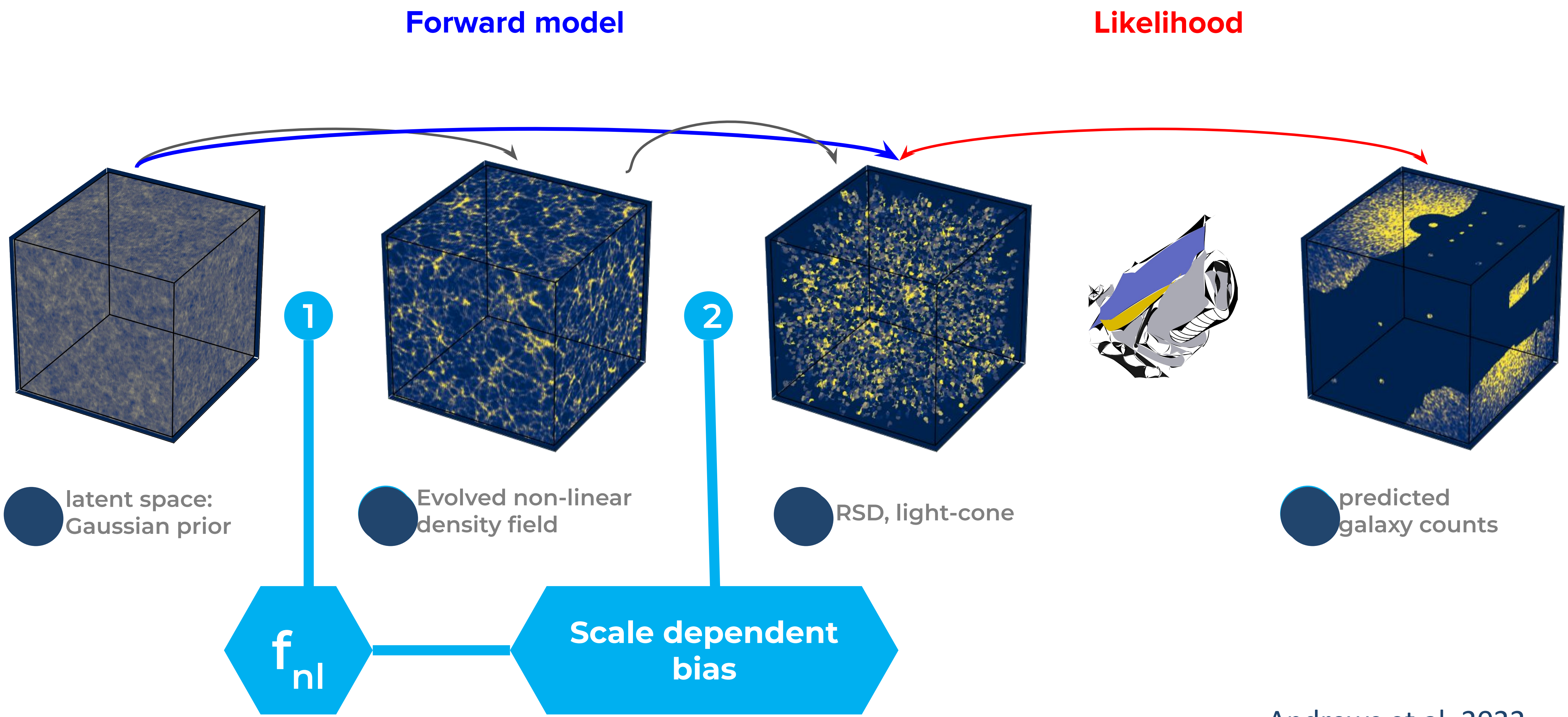
Sampling cosmology with BORG-WL (Natalia Porqueres)





Porqueres et al. 2022

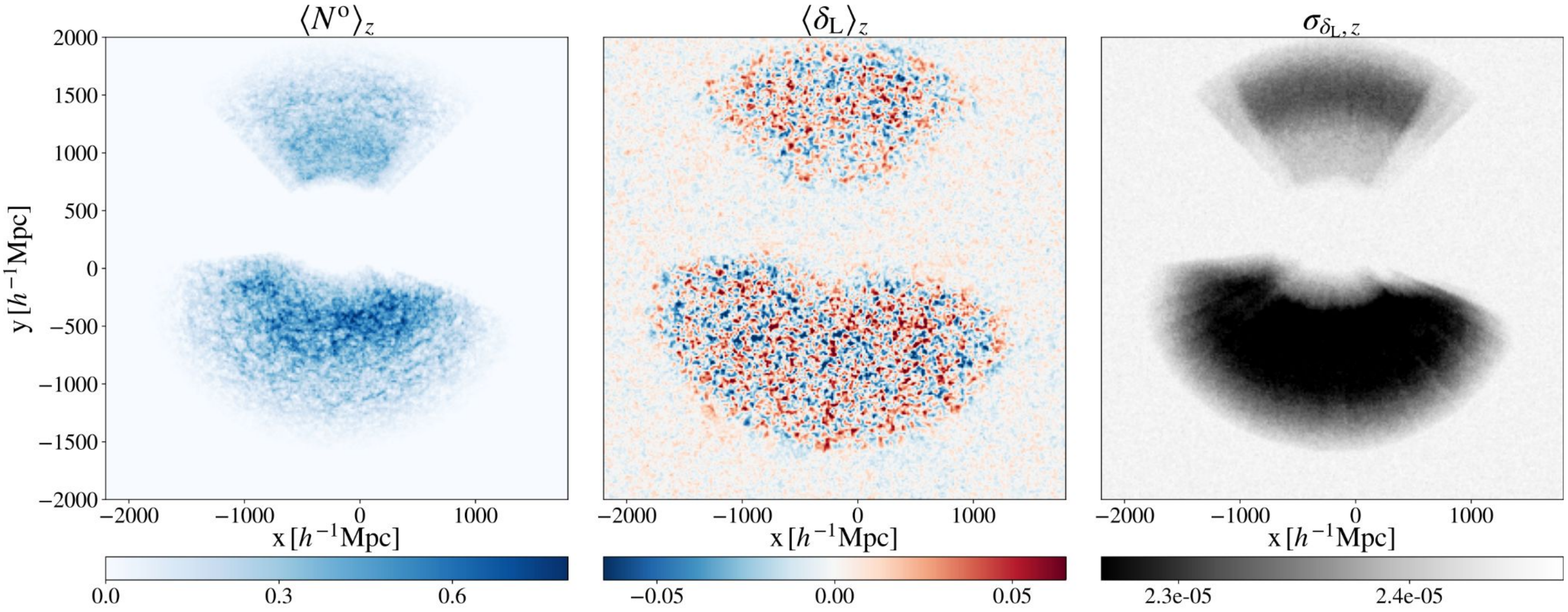
Bayesian forward modeling of Primordial Non-Gaussianity Non-Gaussianity (Adam Andrews)



Andrews et al. 2022

Physical forward modeling of PNG at the level of 3d fields (Adam Andrews)

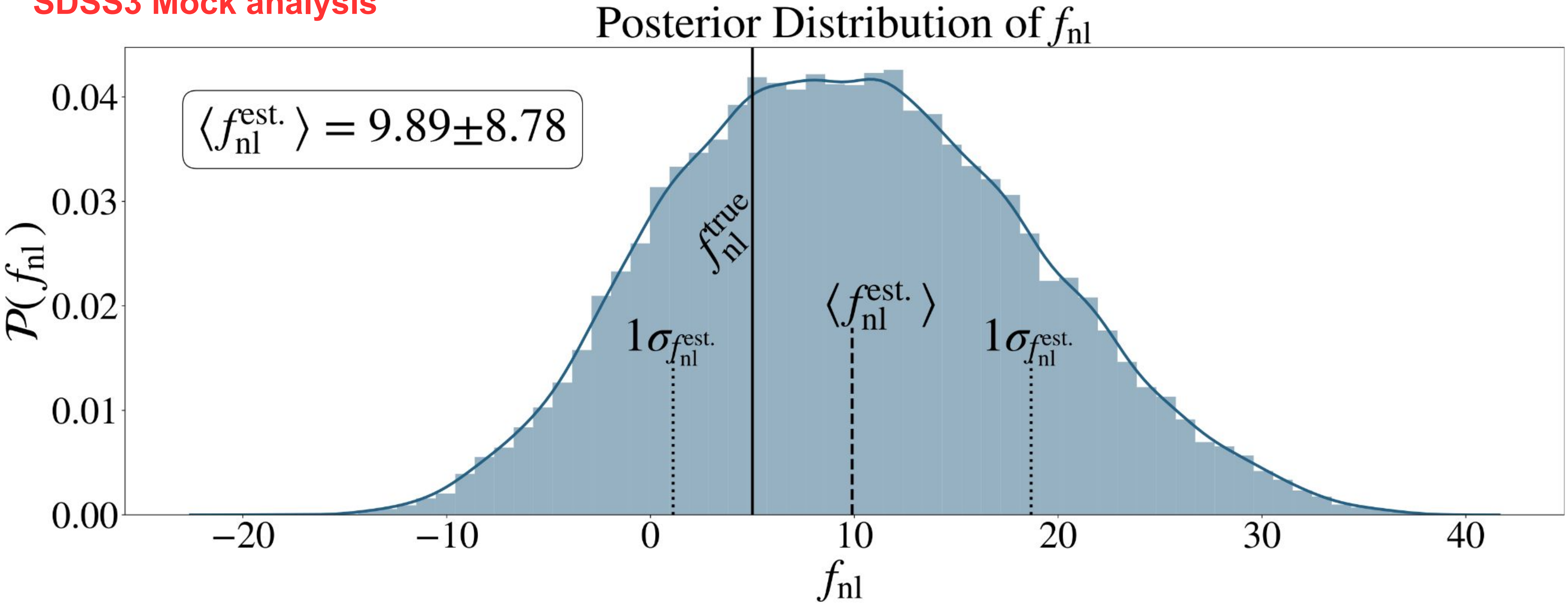
SDSS3 Mock analysis: Primordial fluctuations field inference from galaxy mock data



Andrews et al. 2022

Physical forward modeling of PNG at the level of 3d fields (Adam Andrews)

SDSS3 Mock analysis

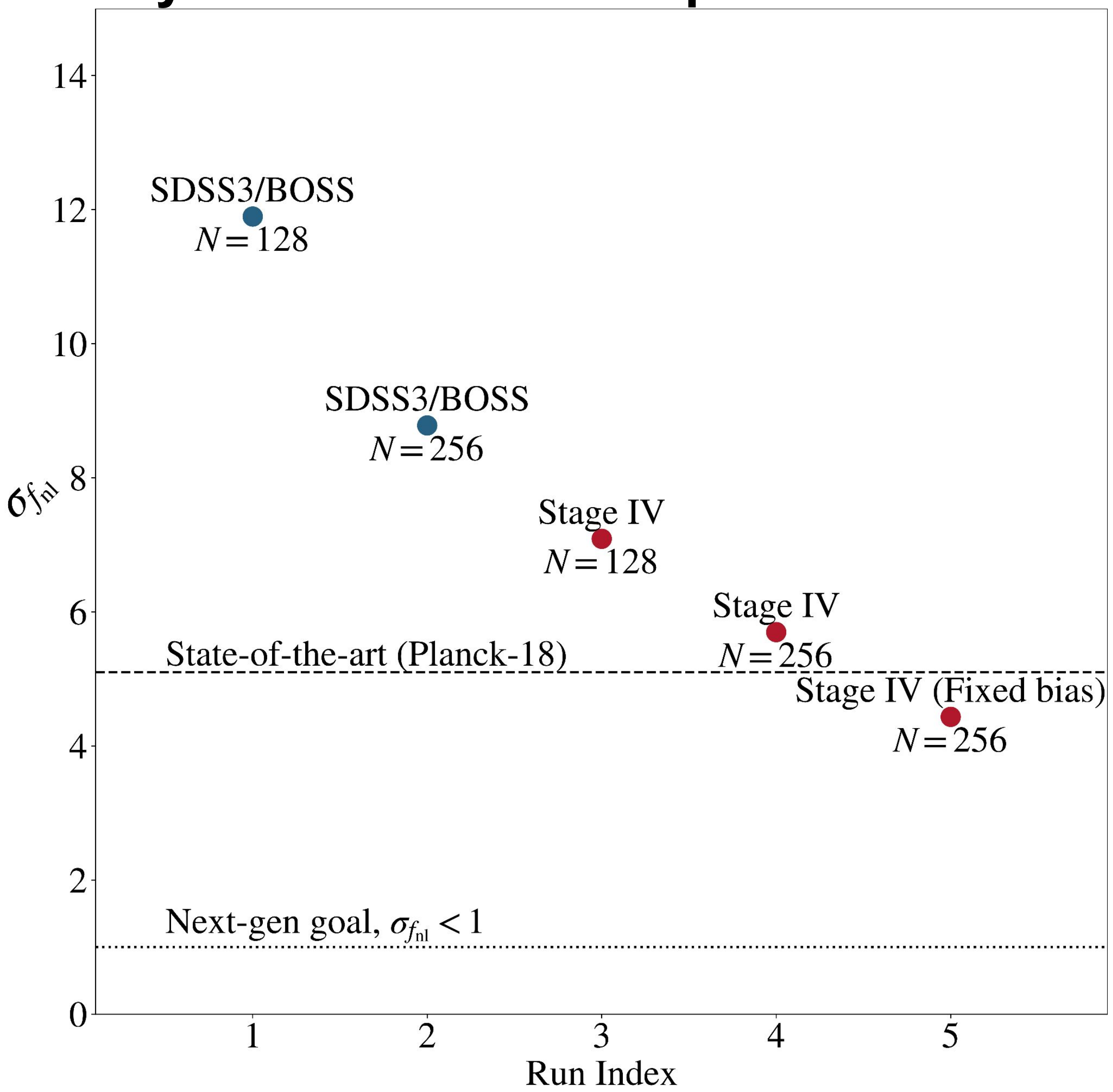


Current state-of-the-art (SDSS3): $f_{nl} = -30 \pm 29$ (D'Amico et al 2022)

Andrews et al. 2022

Physical forward modeling of PNG at the level of 3d fields (Adam Andrews)

Higher resolution of density reconstructions improves PNG constraints.



Andrews et al. 2022

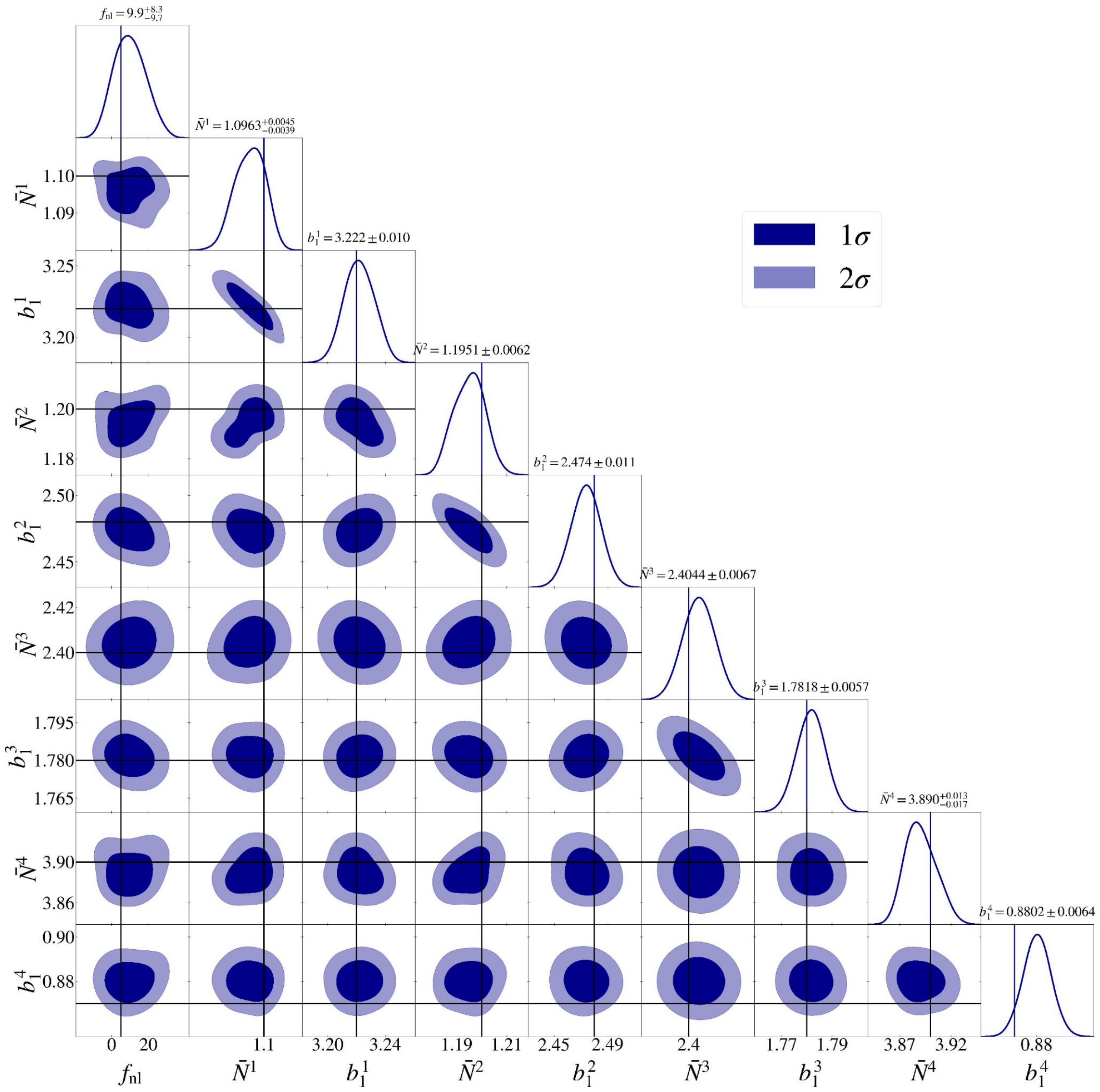
Physical forward modeling of PNG at the level of 3d fields (Adam Andrews)

BORG provides full statistical control

- Covariance matrices of all parameters of the data model
- Marginalizing out nuisance parameters

BORG handles various effects

- Survey Geometries
- Noise
- RSDs
- Light cone effects



Andrews et al. 2022

04

SIBELIUS-Dark

Posterior simulations of the Universe

SIBELIUS-DARK: a constrained simulation of the local volume (Stuart McAlpine)

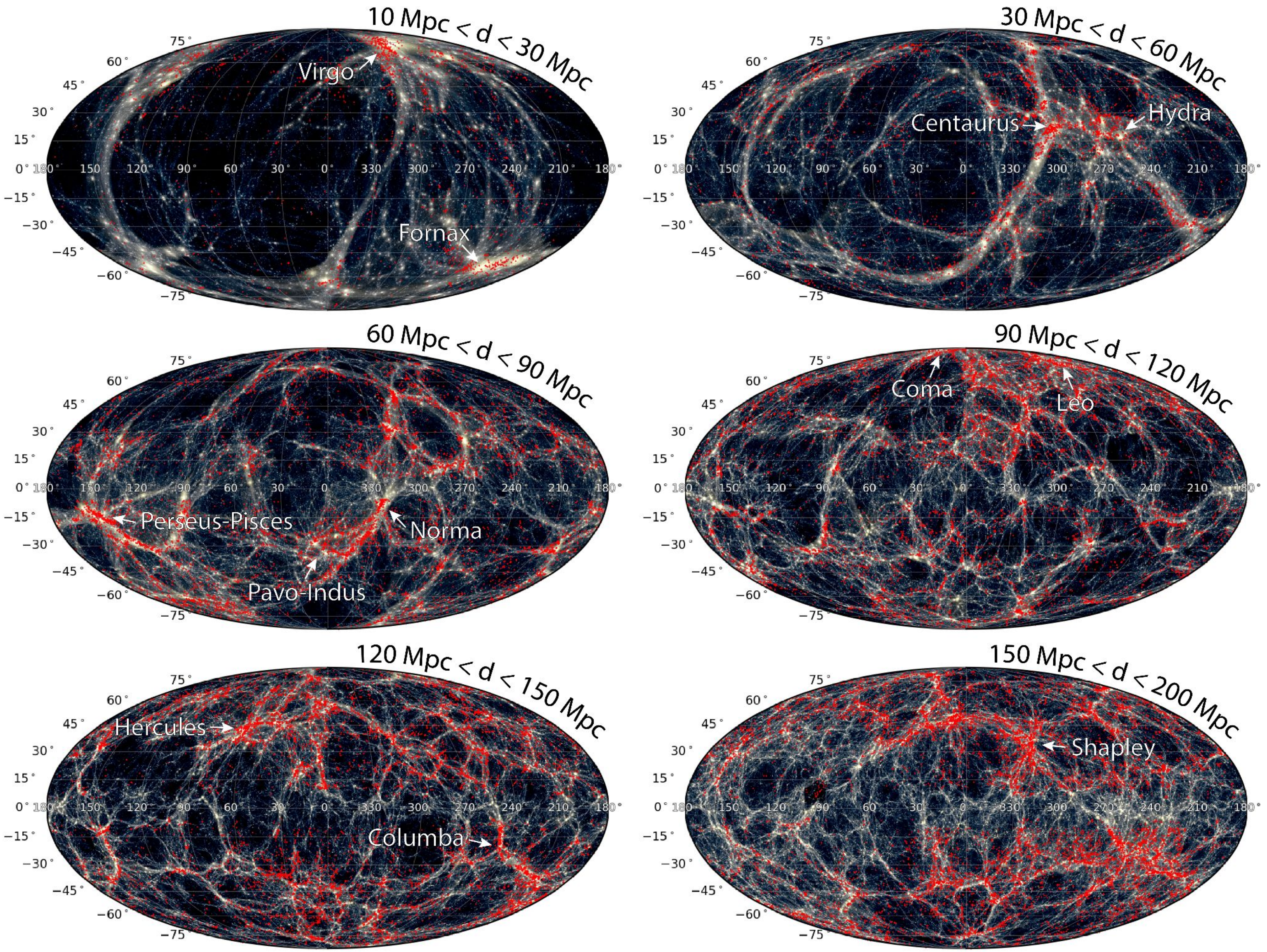
Re-simulate our Nearby Universe (d<200 Mpc)

Constrained Initial conditions

- BORG constraints at scales >3.9 Mpc
- Random fluctuations at scales < 3.9 Mpc
- Add local group candidate

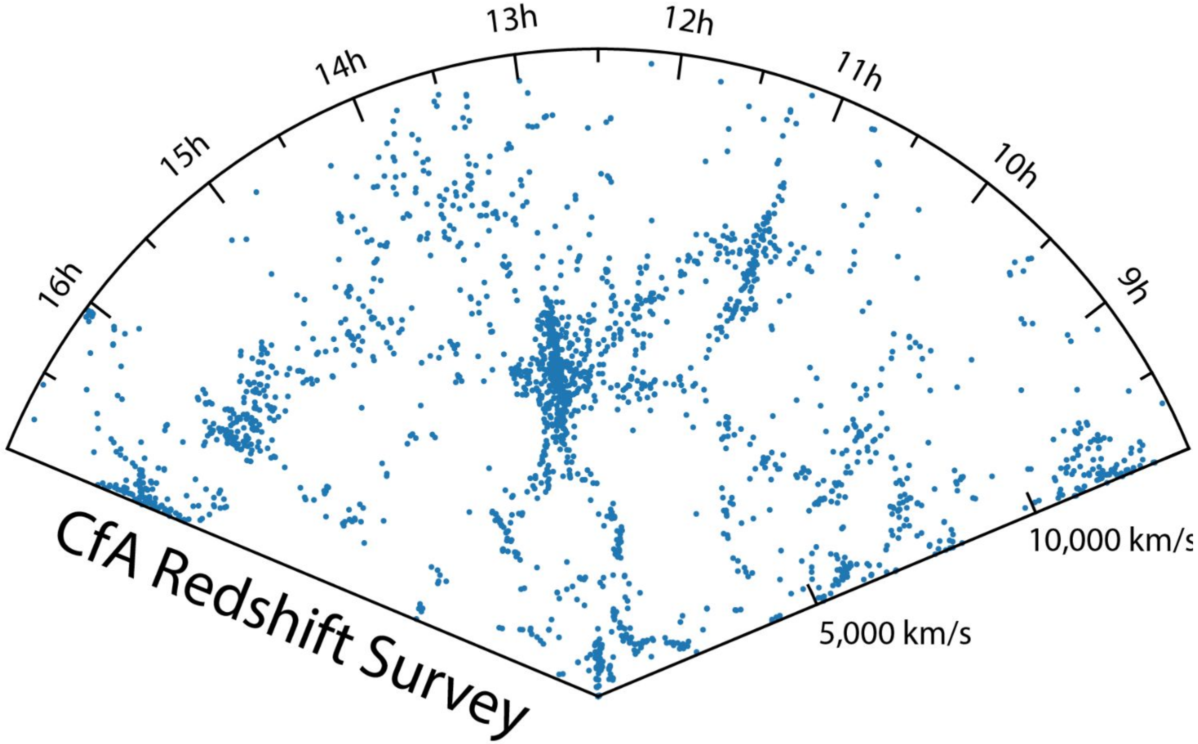
Simulation setup

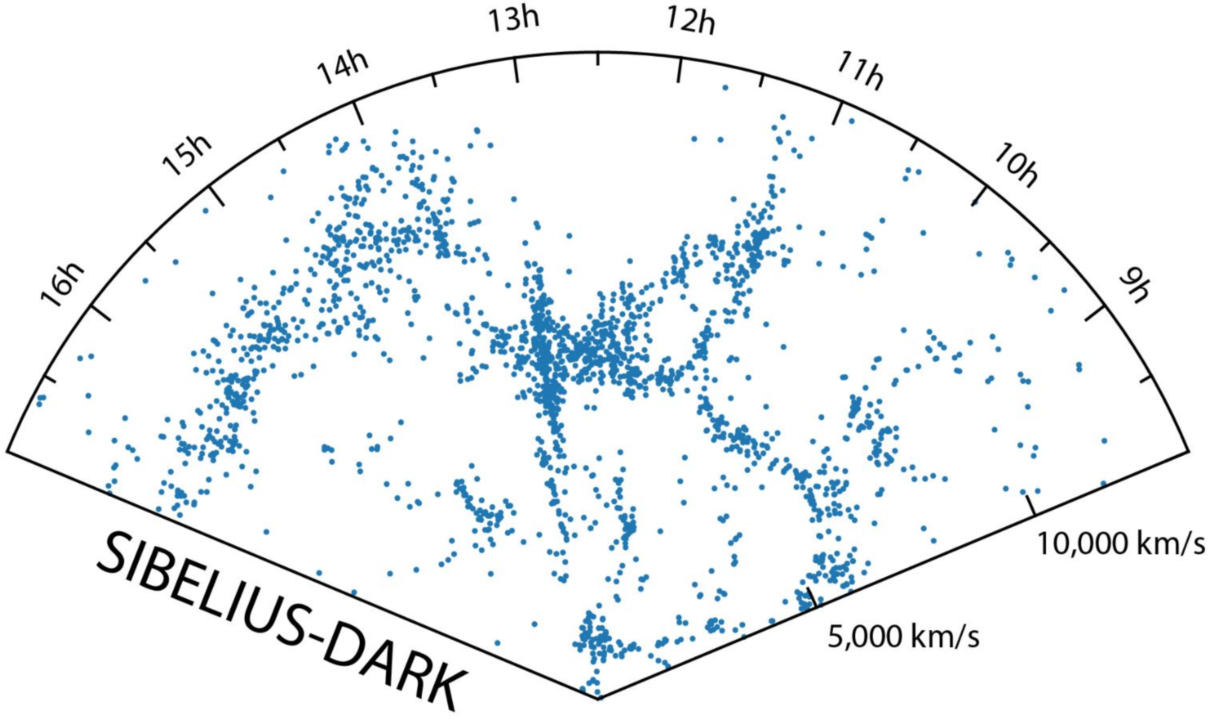
- SWIFT simulation code (Schaller et al 2018)
- Dark Matter Only (DMO)
- L = 1 Gpc
- "Zoom-in" simulation for the inner 200 Mpc
- $N_p^3 = 5078^3$
- 4489 compute cores and 3.5M CPU hours



McAlpine et al 2022

SIBELIUS-DARK: a constrained simulation of the local volume (Stuart McAlpine)





de Lapparent 1986

McAlpine et al 2022

SIBELIUS-DARK: a constrained simulation of the loc

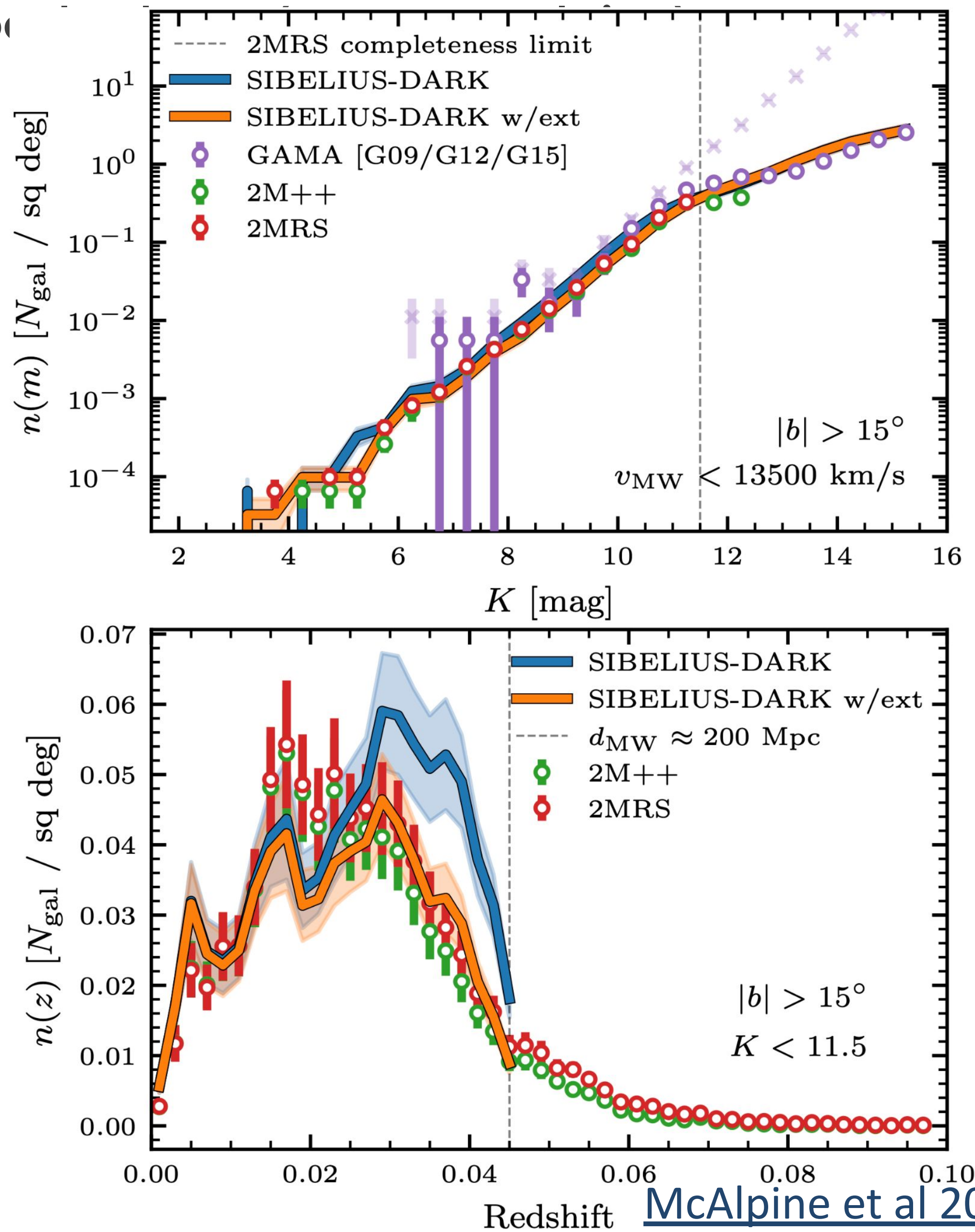
Posterior predictive tests with GALFORM(Lacey et al 2016)

Use **GALFORM** to predict semi-analytic galaxies:

- **SIBELIUS-DARK has a high cadence**

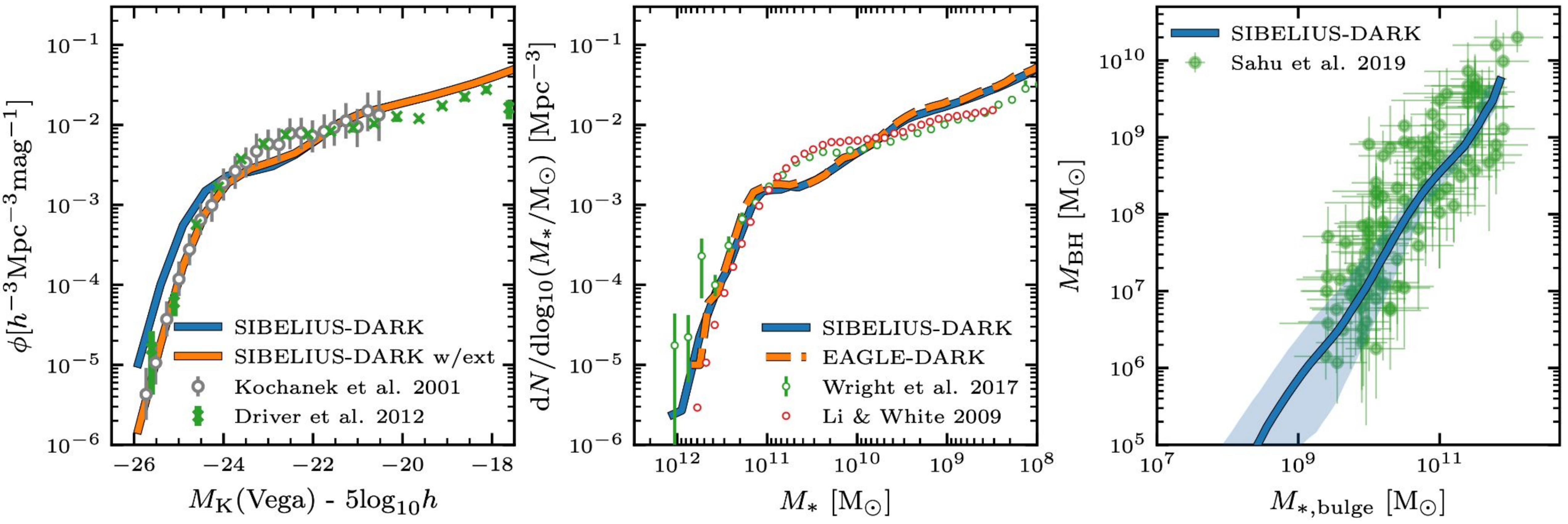
  (200 'snapshots' from 0 < z < 25)
- **Generate Halo catalogs and Merger trees**

  (22,904,767 halos with mass M >10 $M_\odot$)
- **Populate Halos with Galaxies**
- **Account for dust extinction and emission**



McAlpine et al 2022

SIBELIUS-DARK: a constrained simulation of the local volume (Stuart McAlpine)



McAlpine et al 2022

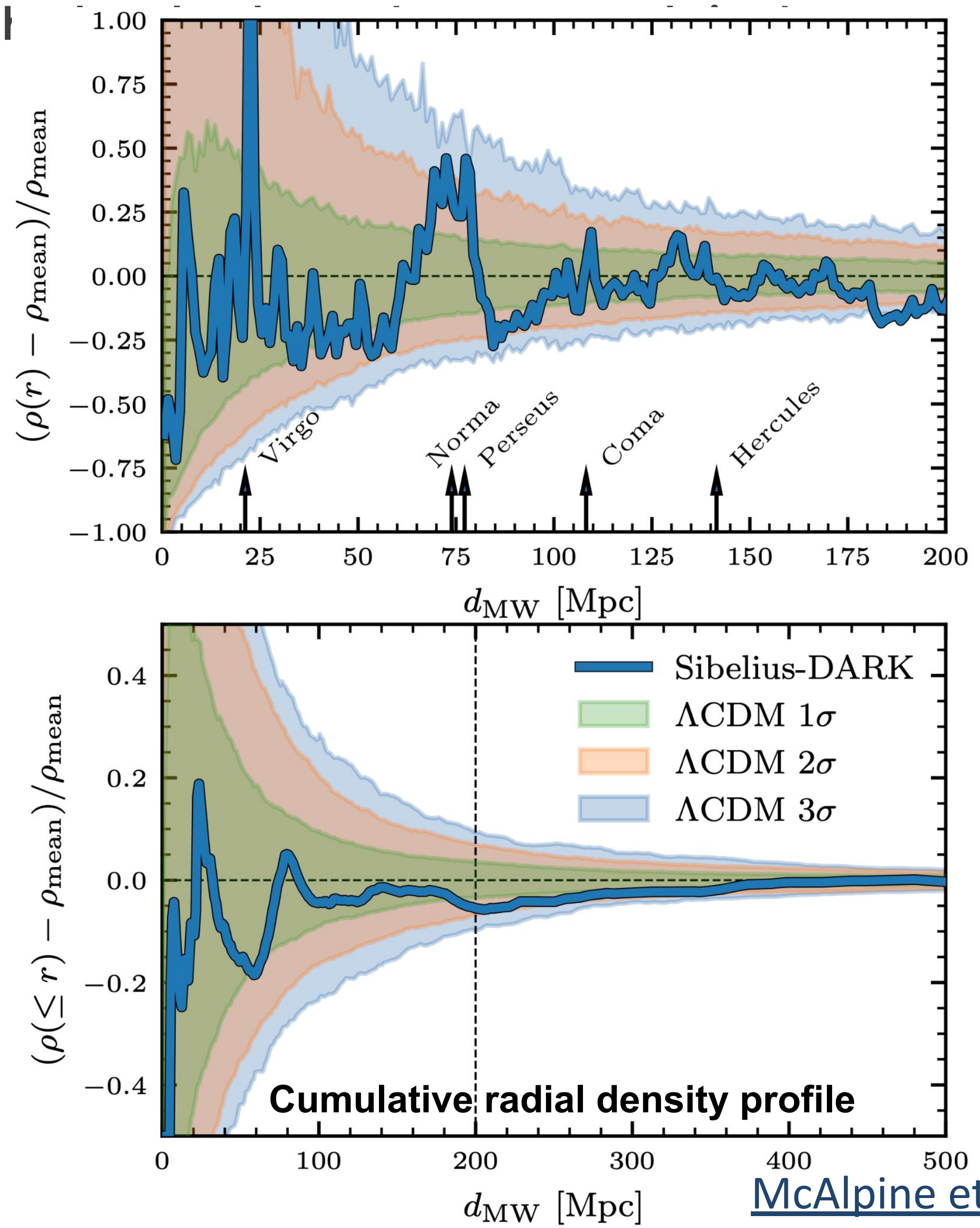
SIBELIUS-DARK: a constrained simulation of tl

A local underdensity of 20% of size 150-200 h^{-1} Mpc could alter reported values of H_0 by 5%.

- A 20% large scale underdensity in the galaxy distribution has been reported (see e.g. Whitbourn & Shanks 2014, Böhringer et al 2020, Wong et al 2021).
- Such an underdensity is unlikely in a ΛCDM Universe (e.g. Wu & Huterer 2017, Castello et al 2021).

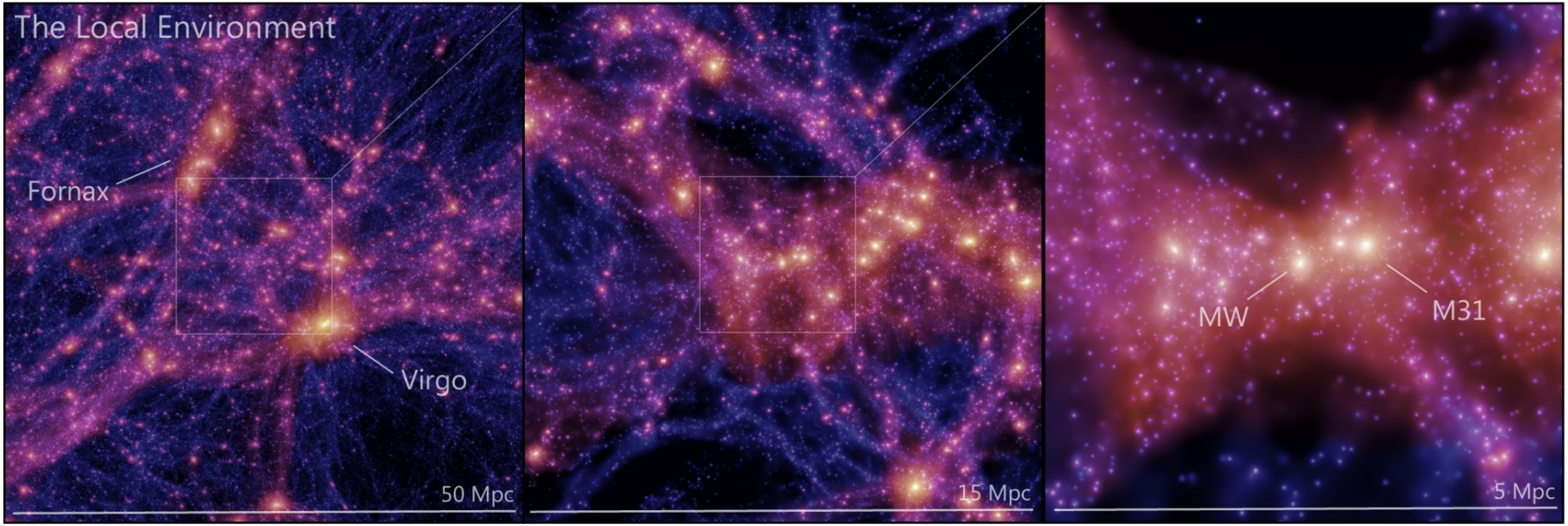
SIBELIUS-DARK finds:

- **No exceptionally large-scale under-density required to reproduce observed galaxy counts.**
- There is a slight 5% underdensity.
- Rare but no challenge for ΛCDM.



McAlpine et al 2022

SIBELIUS-DARK: a constrained simulation of the local volume (Stuart McAlpine)



SIBELIUS-DARK:

is the most comprehensive simulation of the local volume to date.
(For comparison see: Heß et al 2013, Wang et al 2016, Libeskind et al 2020, Sorce et al 2021)

SIBELIUS-DARK data is available at:
https://virgodb.dur.ac.uk/

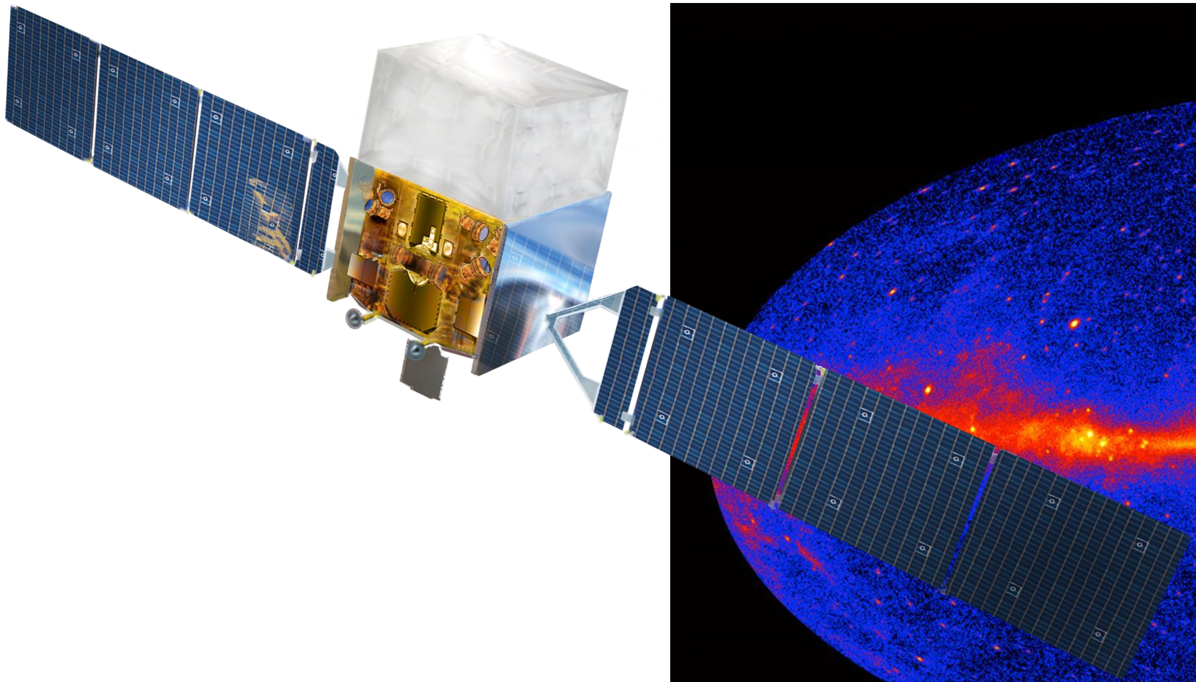
McAlpine et al 2022

05 Constraining Fundamental Physics

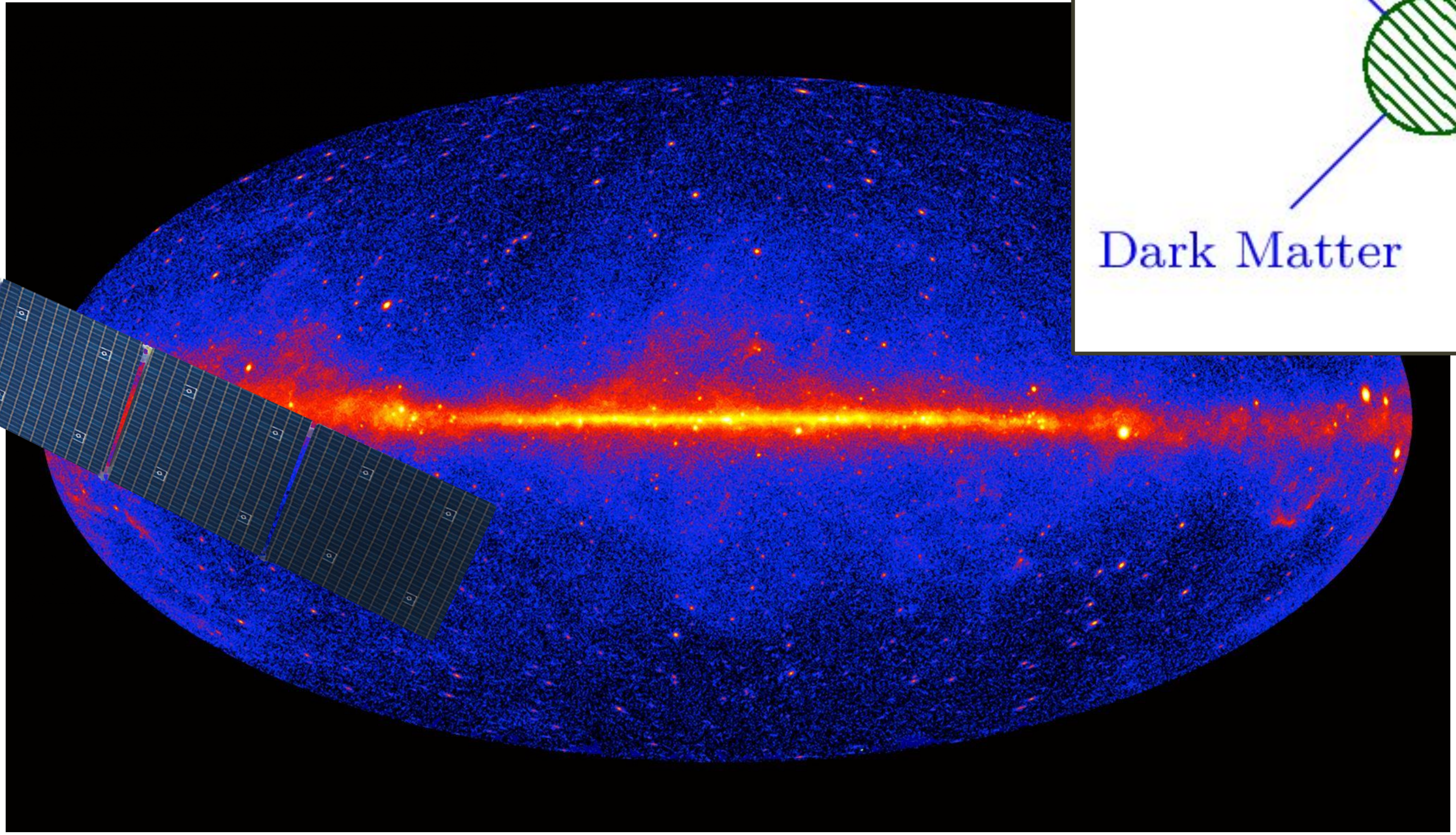
Turning the Universe into a lab

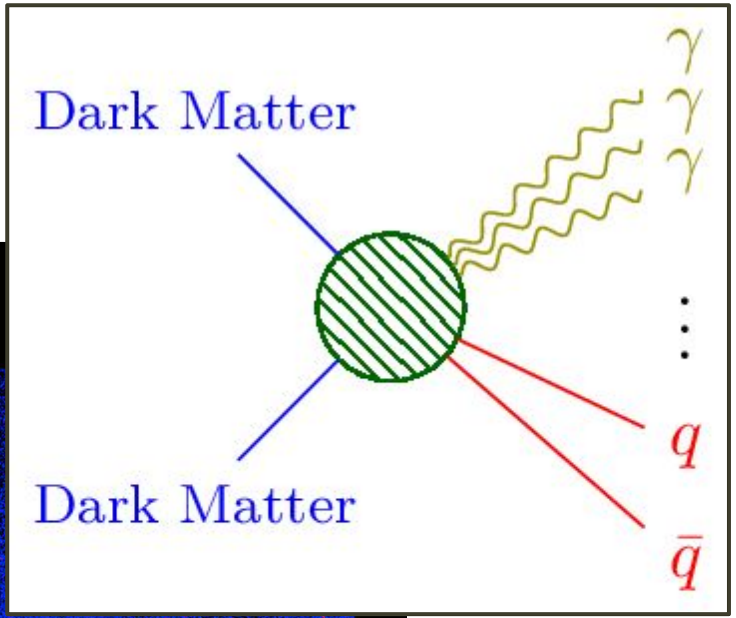
Constraining dark matter annihilation and decay

Fermi Gamma-ray Space Telescope



All-sky Gamma-ray map

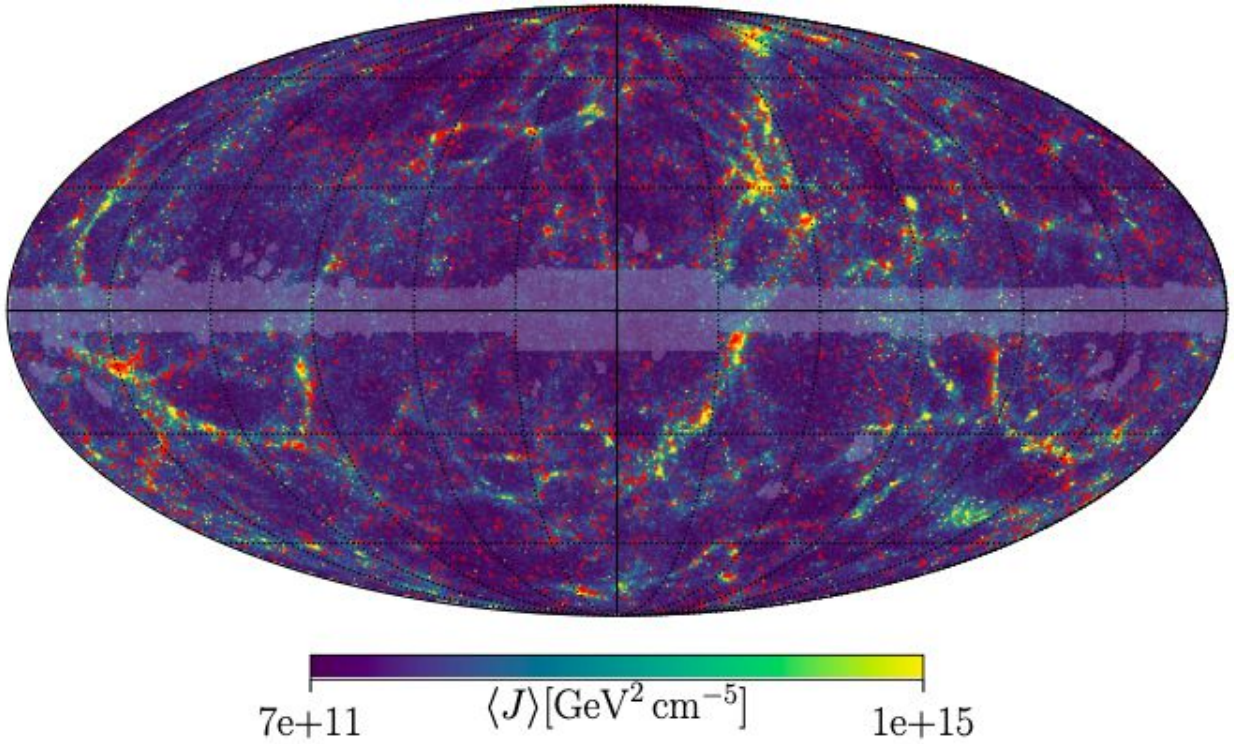




(Image credit: NASA/DOE/Fermi LAT Collaboration)

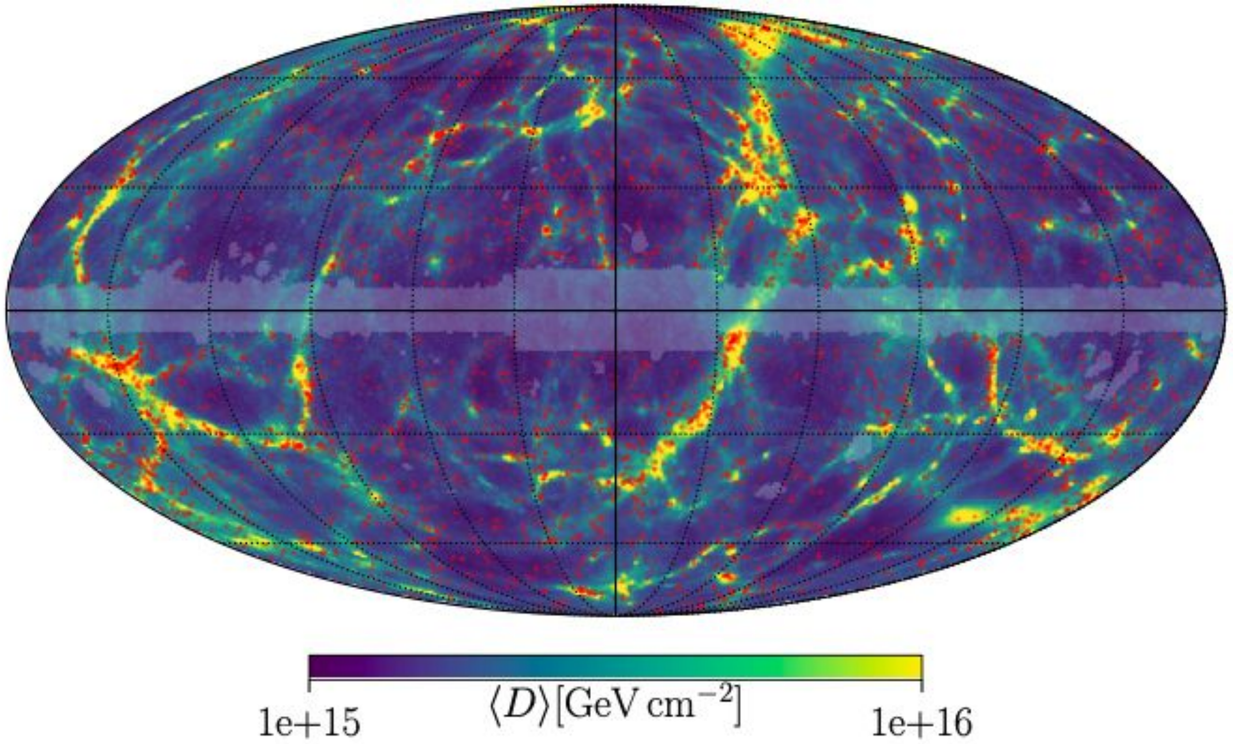
Constraining dark matter annihilation and decay

Dark Matter annihilation



$$\frac{\mathrm{d}J}{\mathrm{d}\Omega} = \int \rho_{\mathrm{DM}}^{2}\left(s,\Omega\right)\mathrm{d}s$$

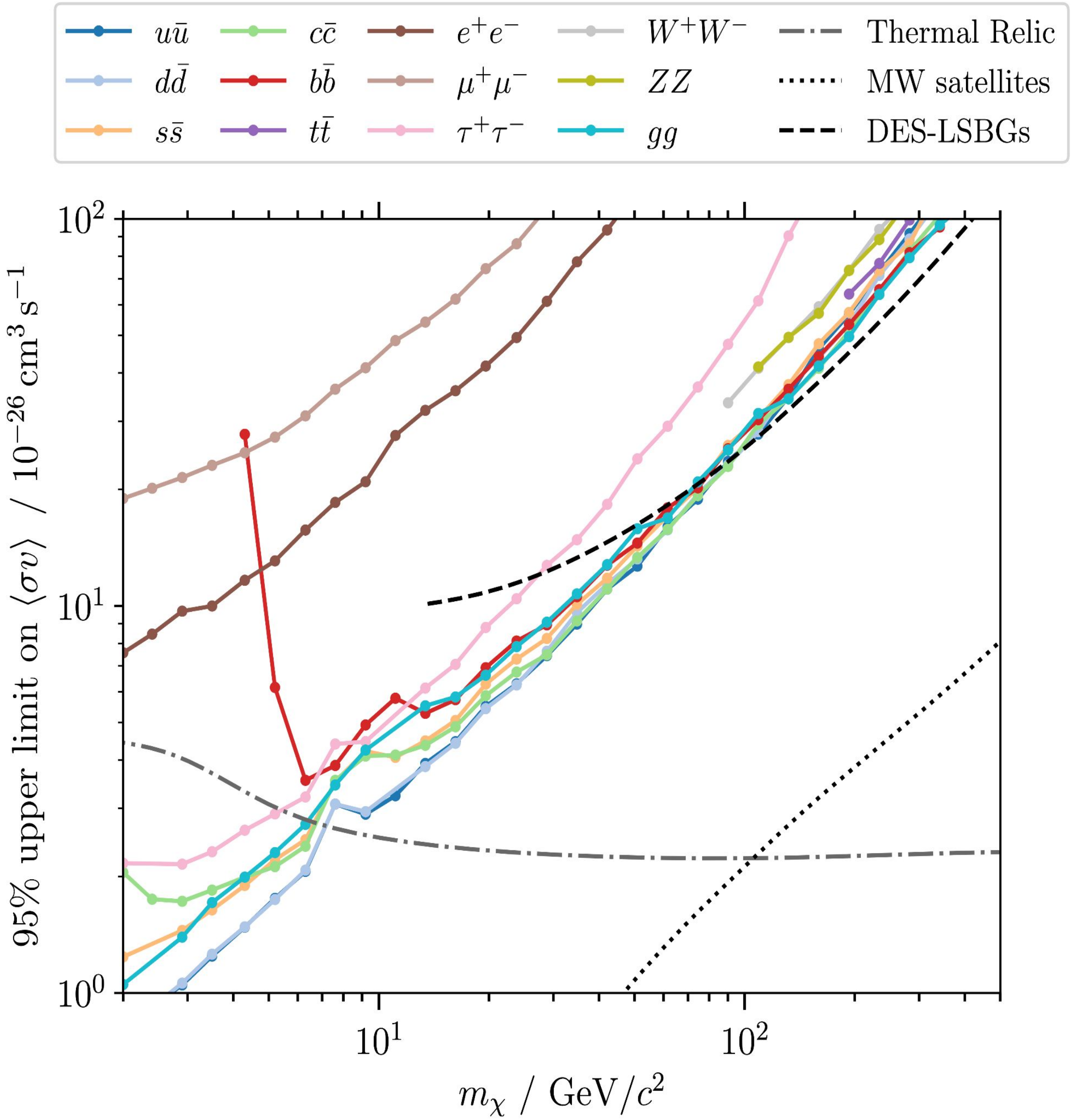
Dark Matter decay



$$\frac{\mathrm{d}D}{\mathrm{d}\Omega} = \int \rho_{\mathrm{DM}}\left(s,\Omega\right)\mathrm{d}s$$

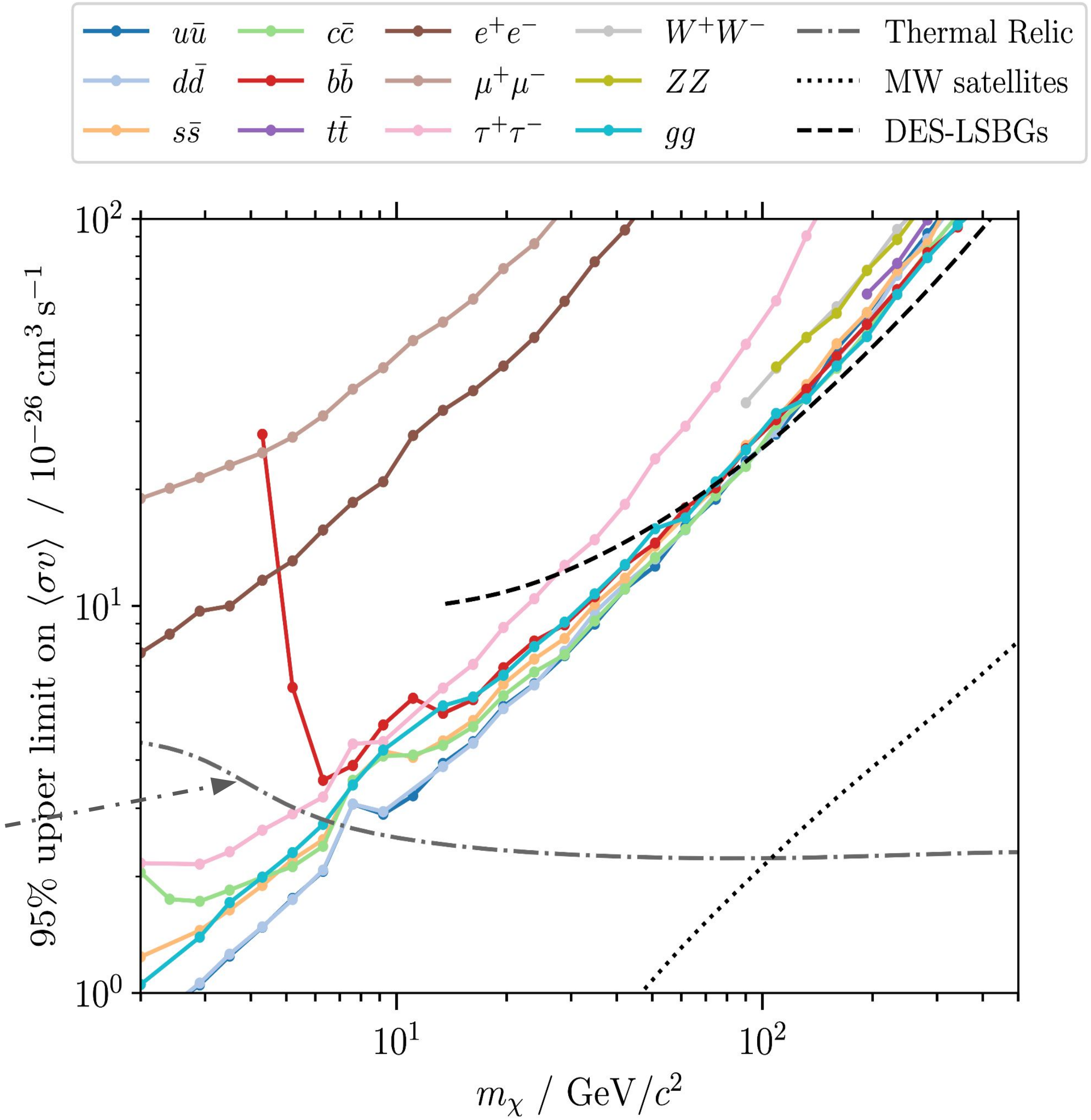
Bartlett et al 2022
Cuesta et al 2011

Constraining dark matter annihilation and decay



Bartlett et al 2022

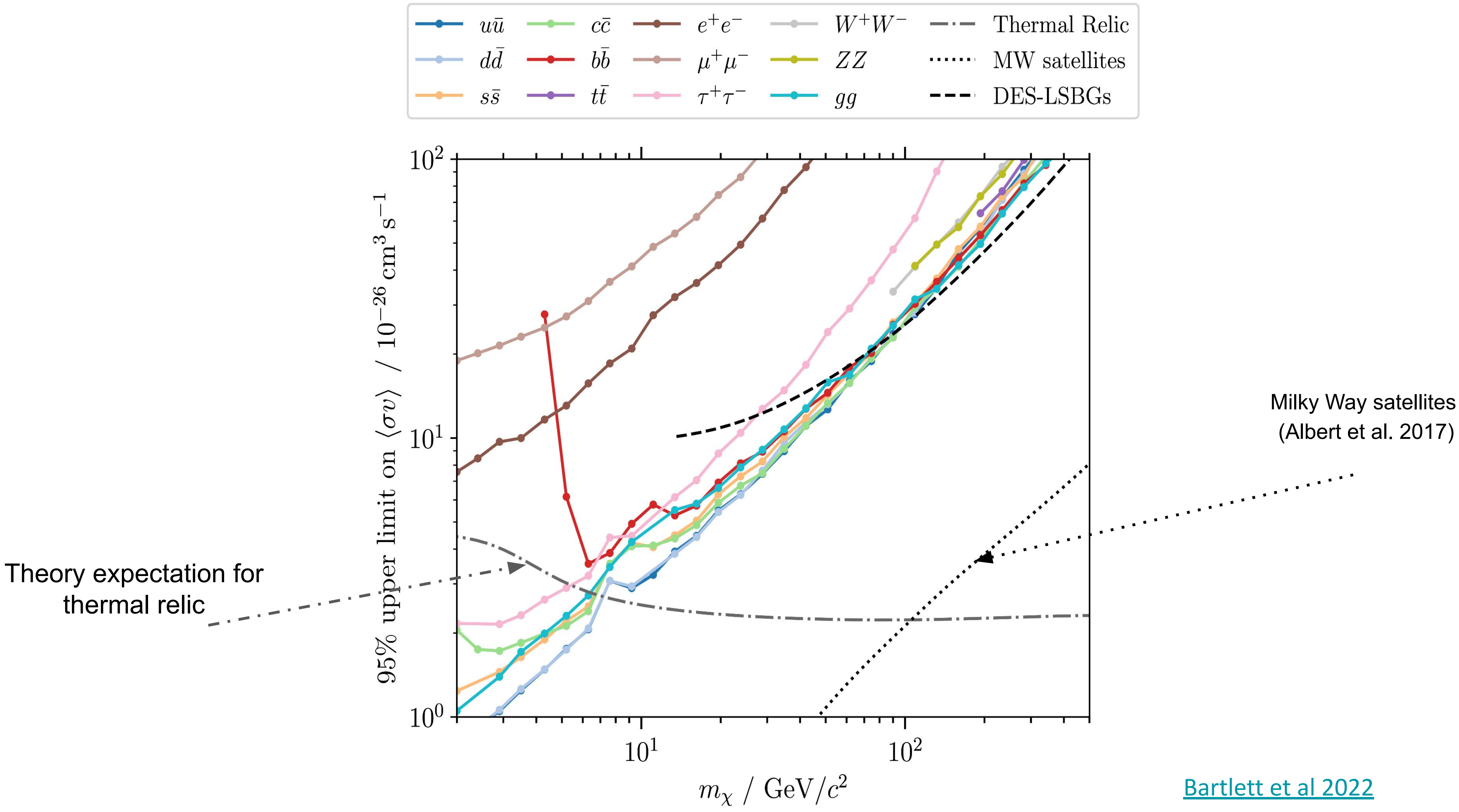
Constraining dark matter annihilation and decay



Theory expectation for thermal relic

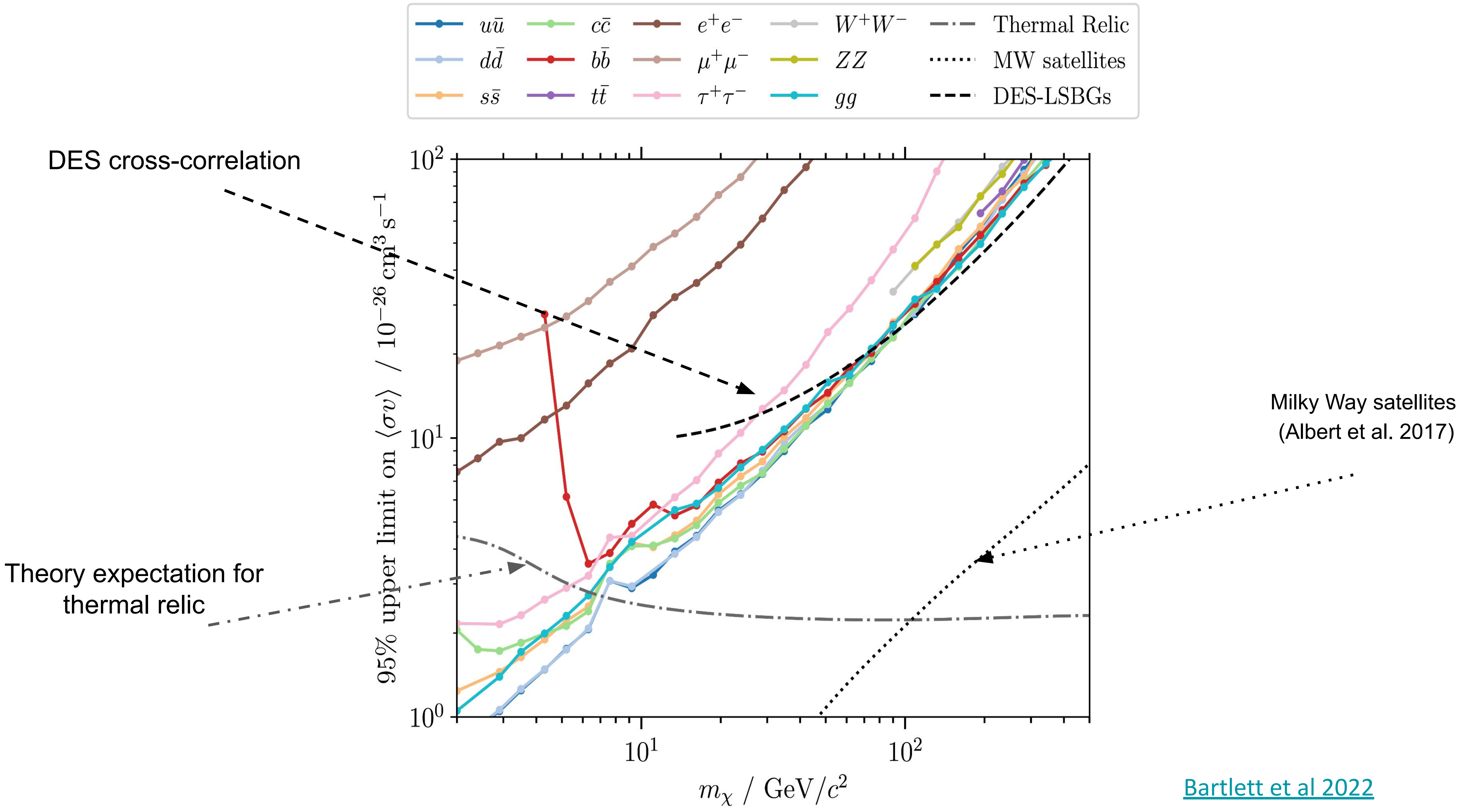
Bartlett et al 2022

Constraining dark matter annihilation and decay



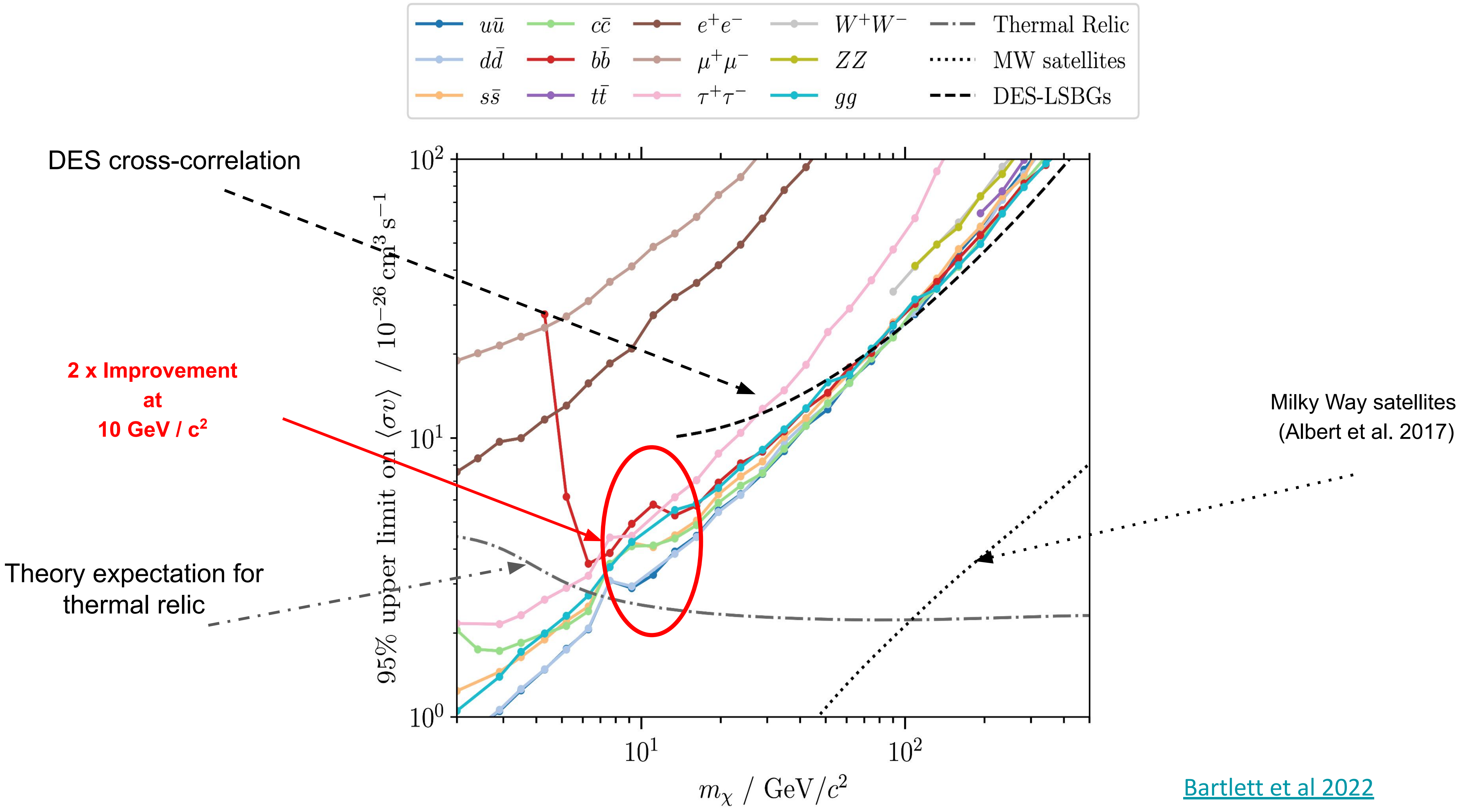
Bartlett et al 2022

Constraining dark matter annihilation and decay



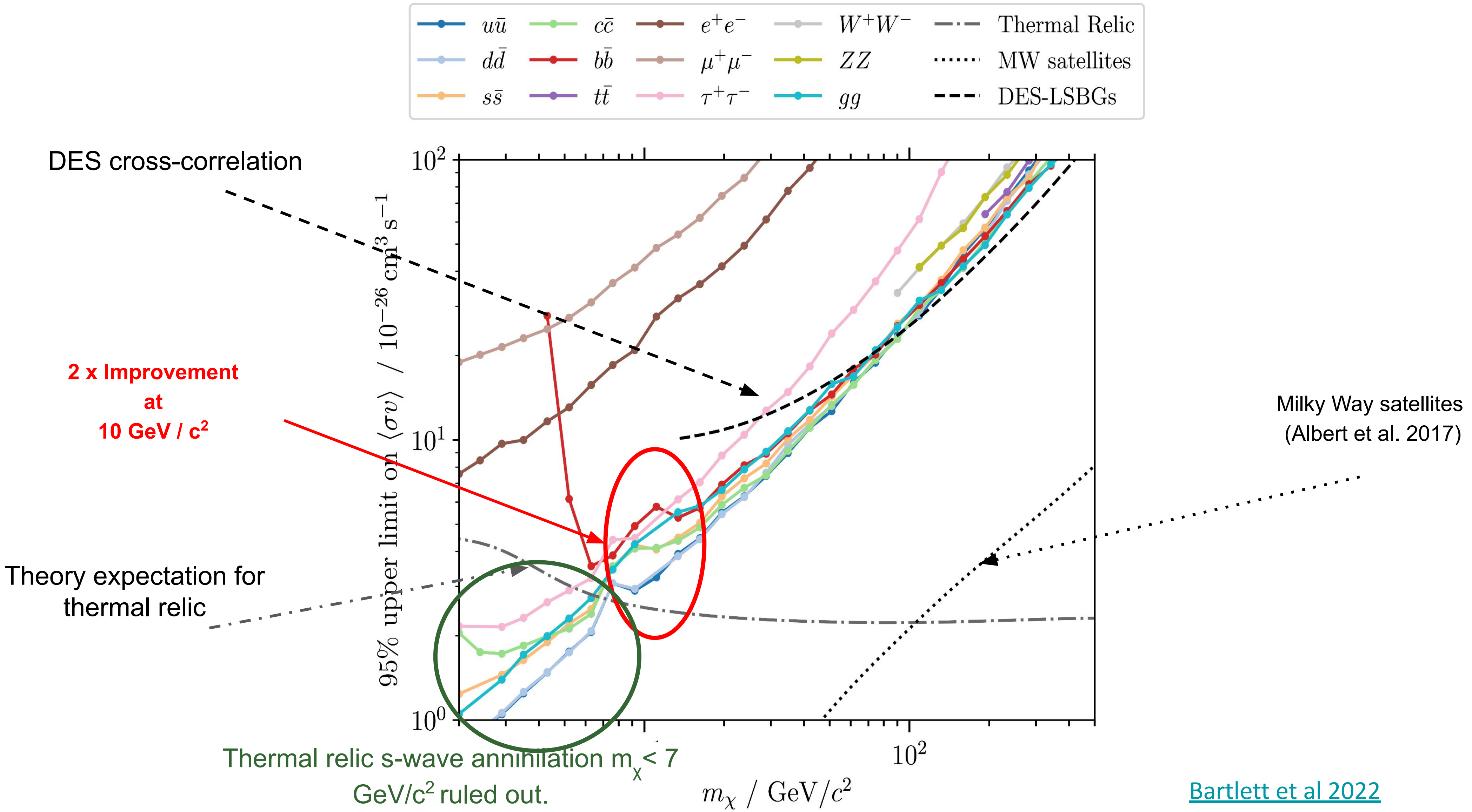
Bartlett et al 2022

Constraining dark matter annihilation and decay



Bartlett et al 2022

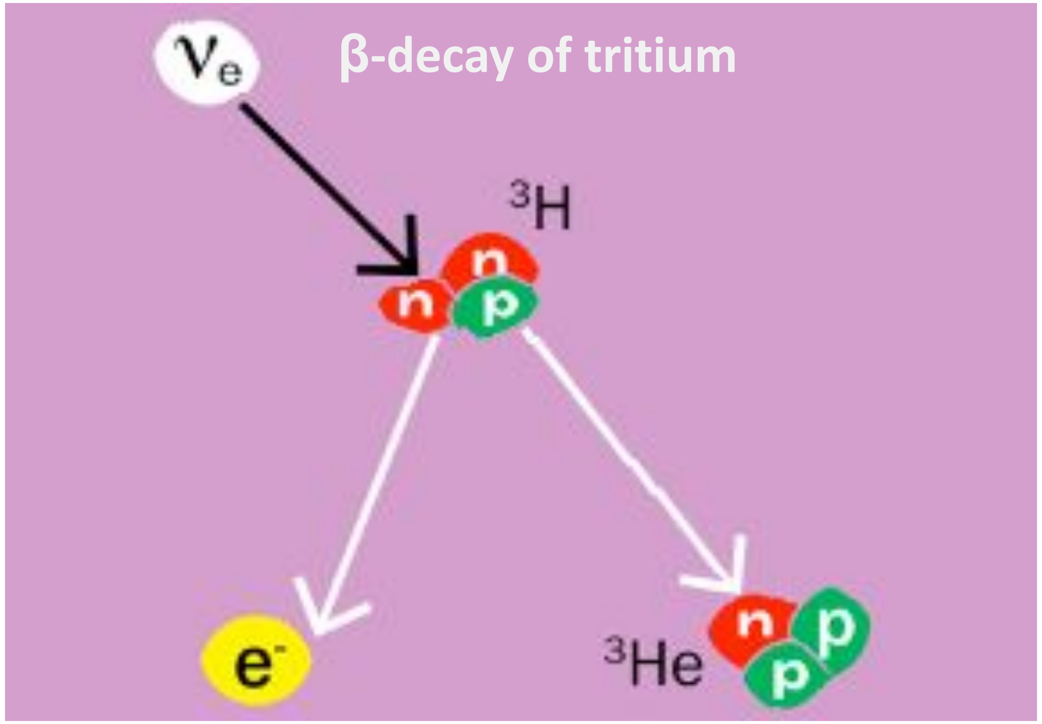
Constraining dark matter annihilation and decay



Bartlett et al 2022

PTOLEMY: detect CNB by capturing neutrinos through tritium inverse β-decay.

- Event rate depends on the local neutrino number density.
- Some experiments require the dipole moment and consider the velocity of neutrinos in the lab frame.

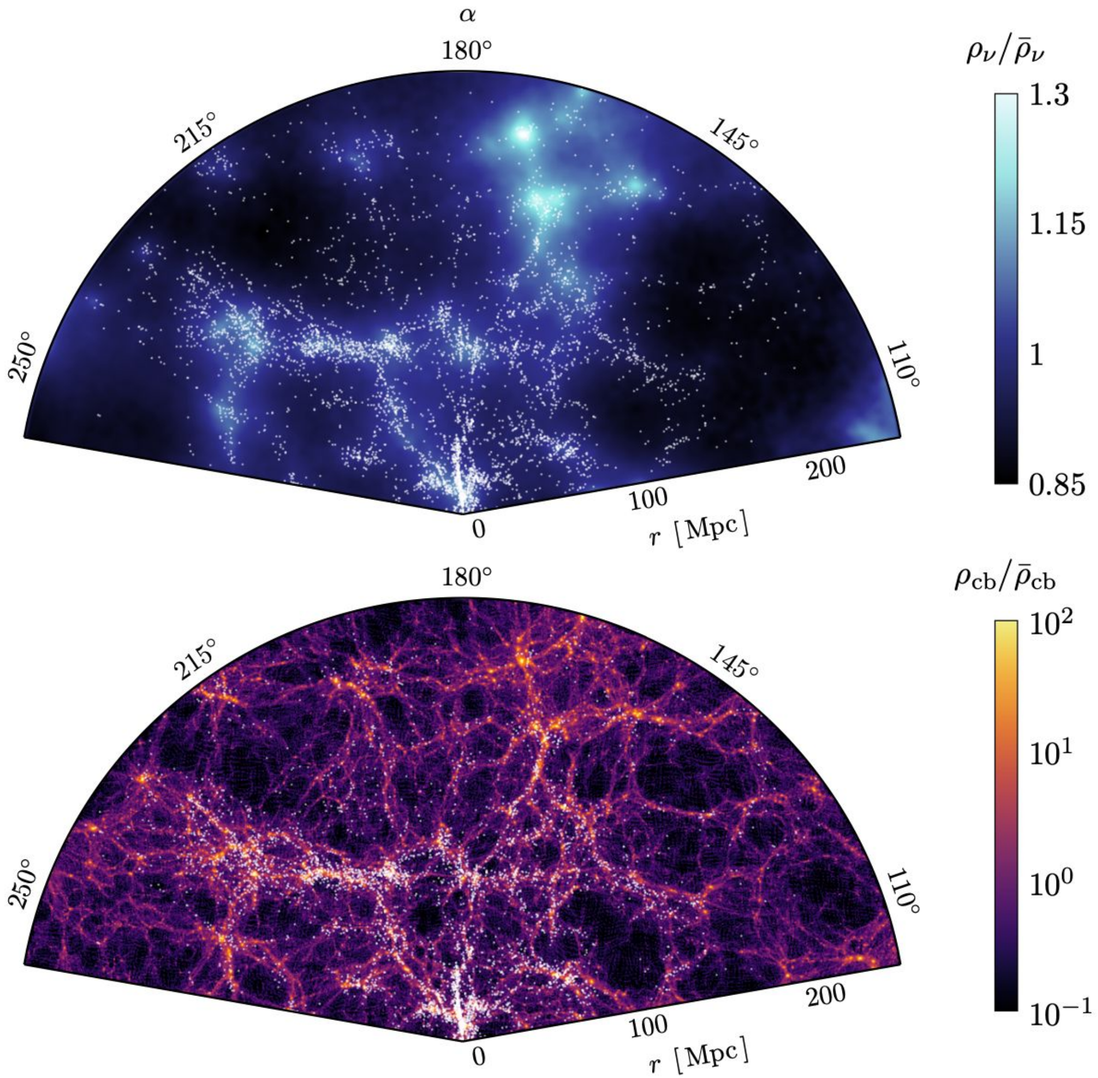


PTOLEMY collaboration 2019

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

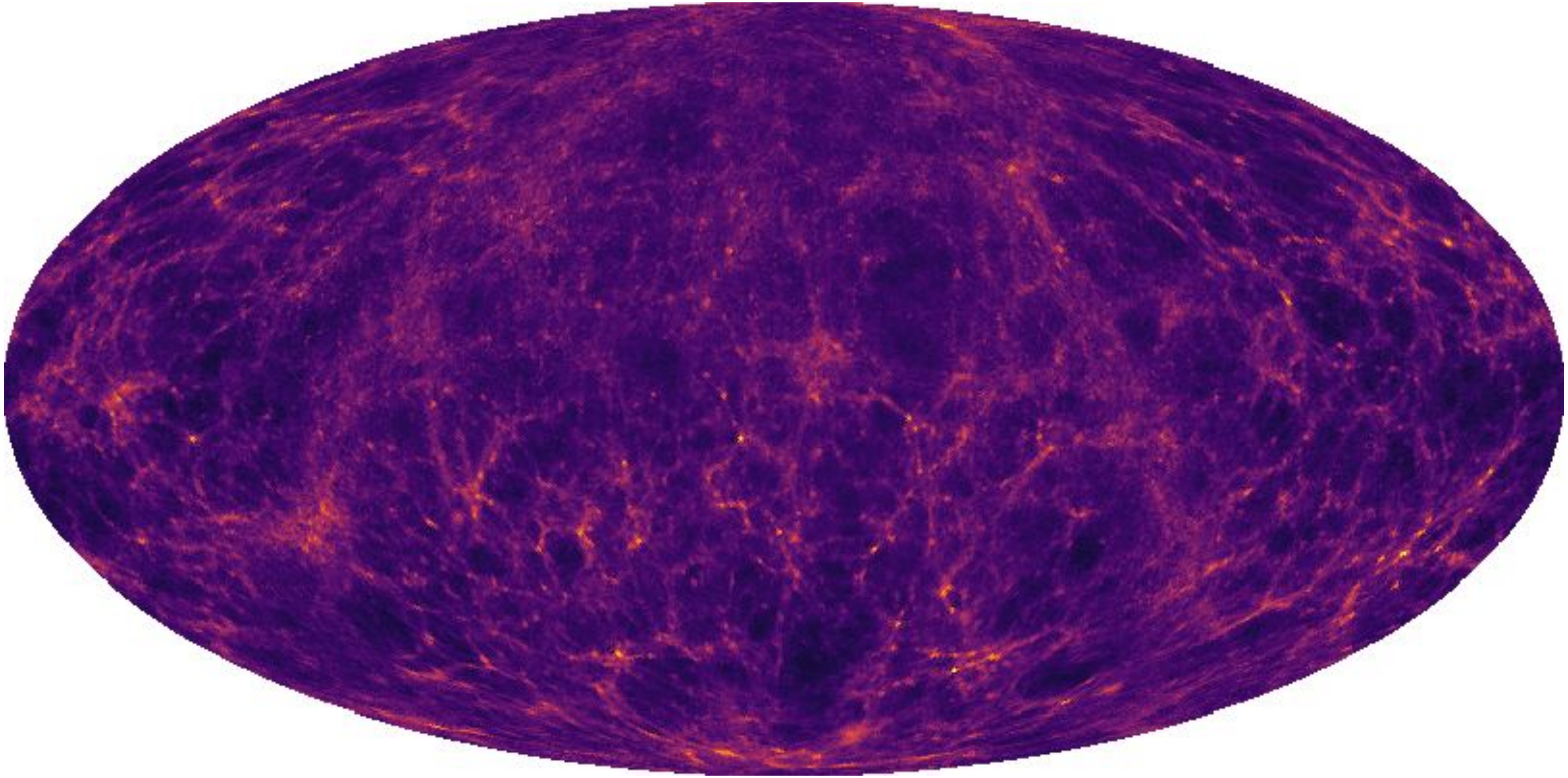
Goal: Determine the prospects of CNB detection

- Posterior simulations to jointly model the evolution of large-scale structures and the neutrino background.
- Non-linear treatment of massive neutrinos, including the gravitational effects of the neutrinos themselves.
- Full scale large-scale distribution of matter within $200h^{-1}$ Mpc
- Compute the expected density, velocity, and direction of relic neutrinos, as well as expected event rates for PTOLEMY.



Elbers et al 2023

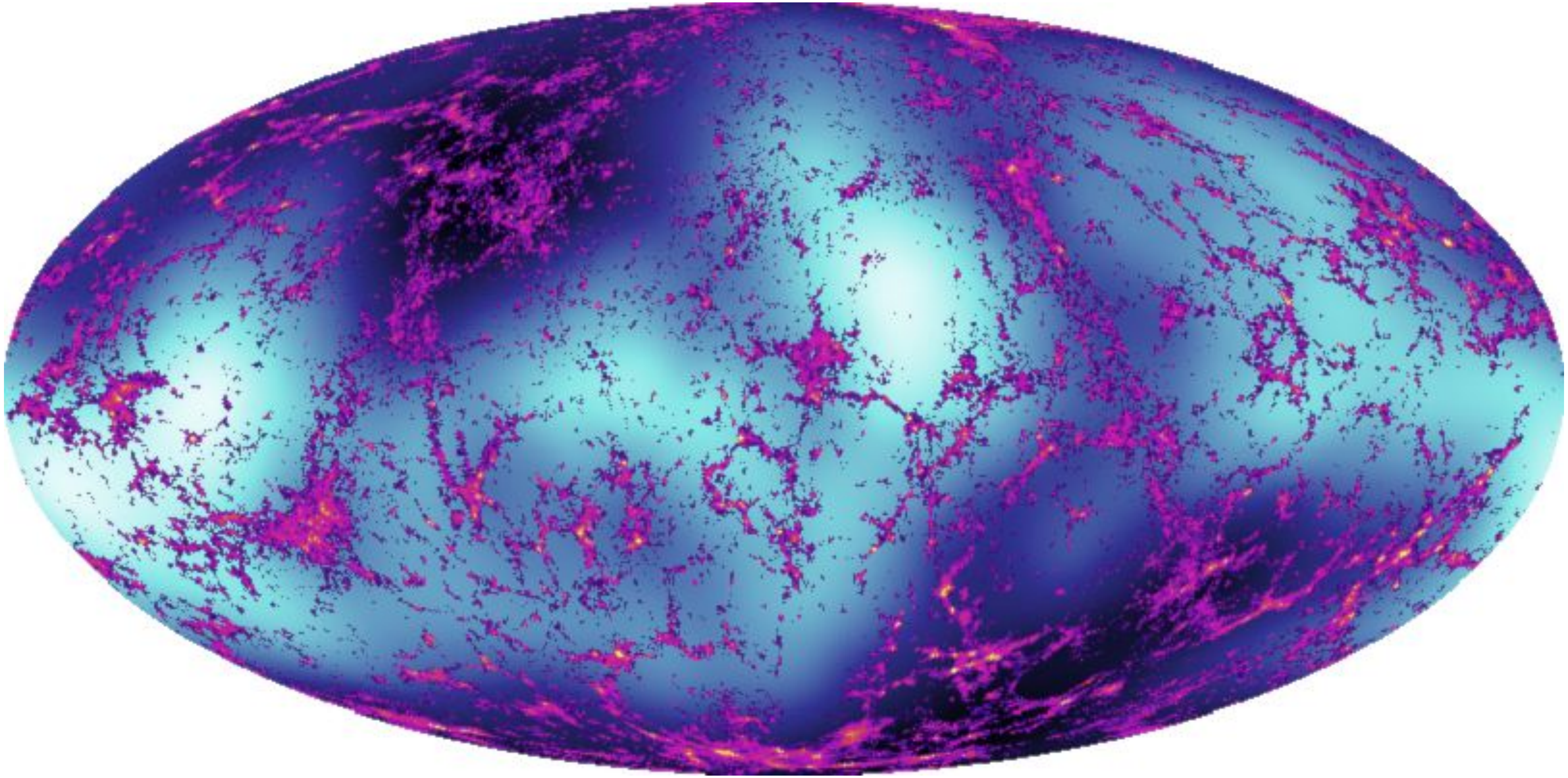
Reconstruction of anisotropies in the neutrino sky (Willem Elbers)



Elbers et al 2023

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

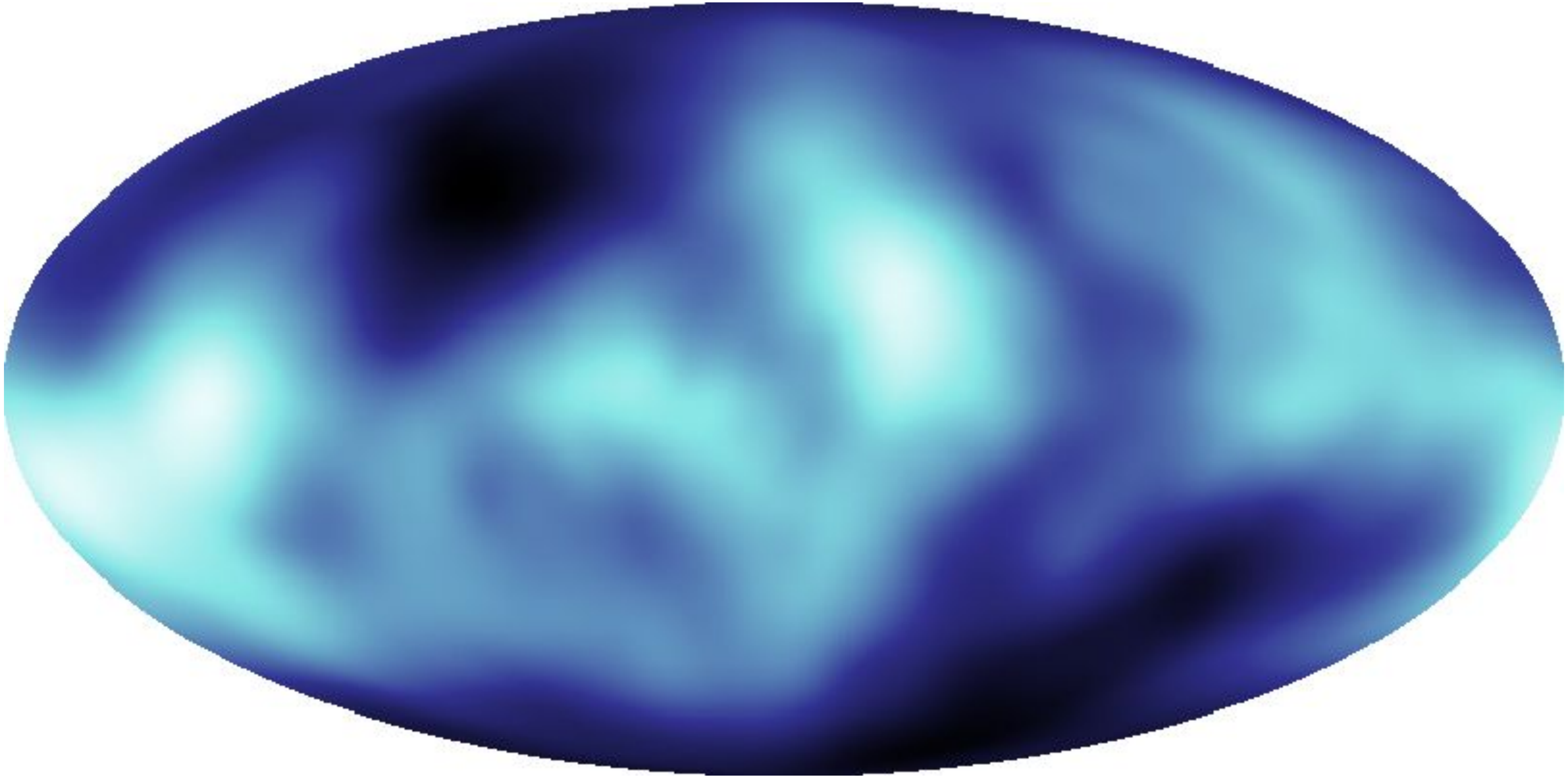
m_ν = 0.01 eV



Elbers et al 2023

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

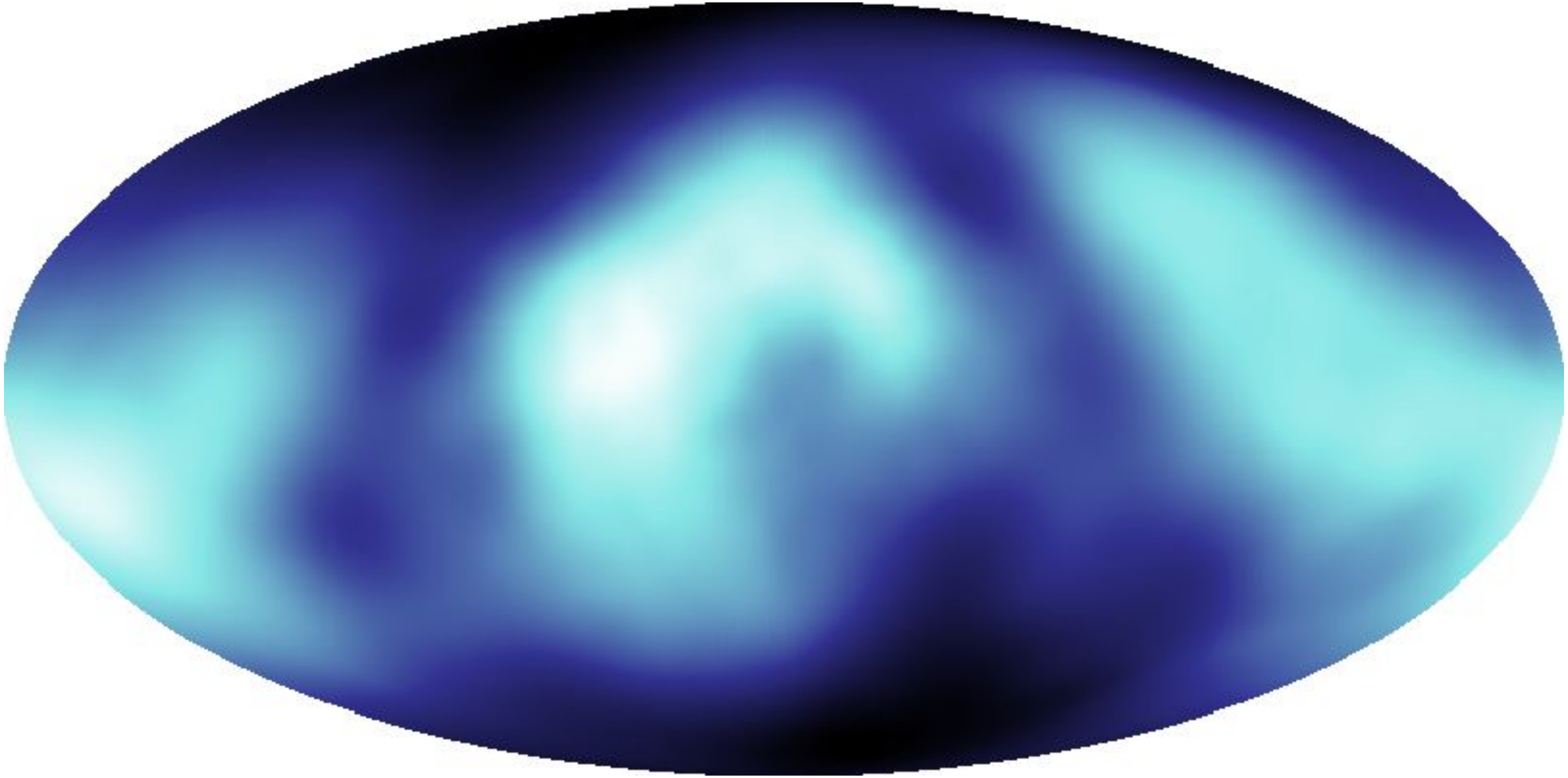
m_ν = 0.01 eV



Elbers et al 2023

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

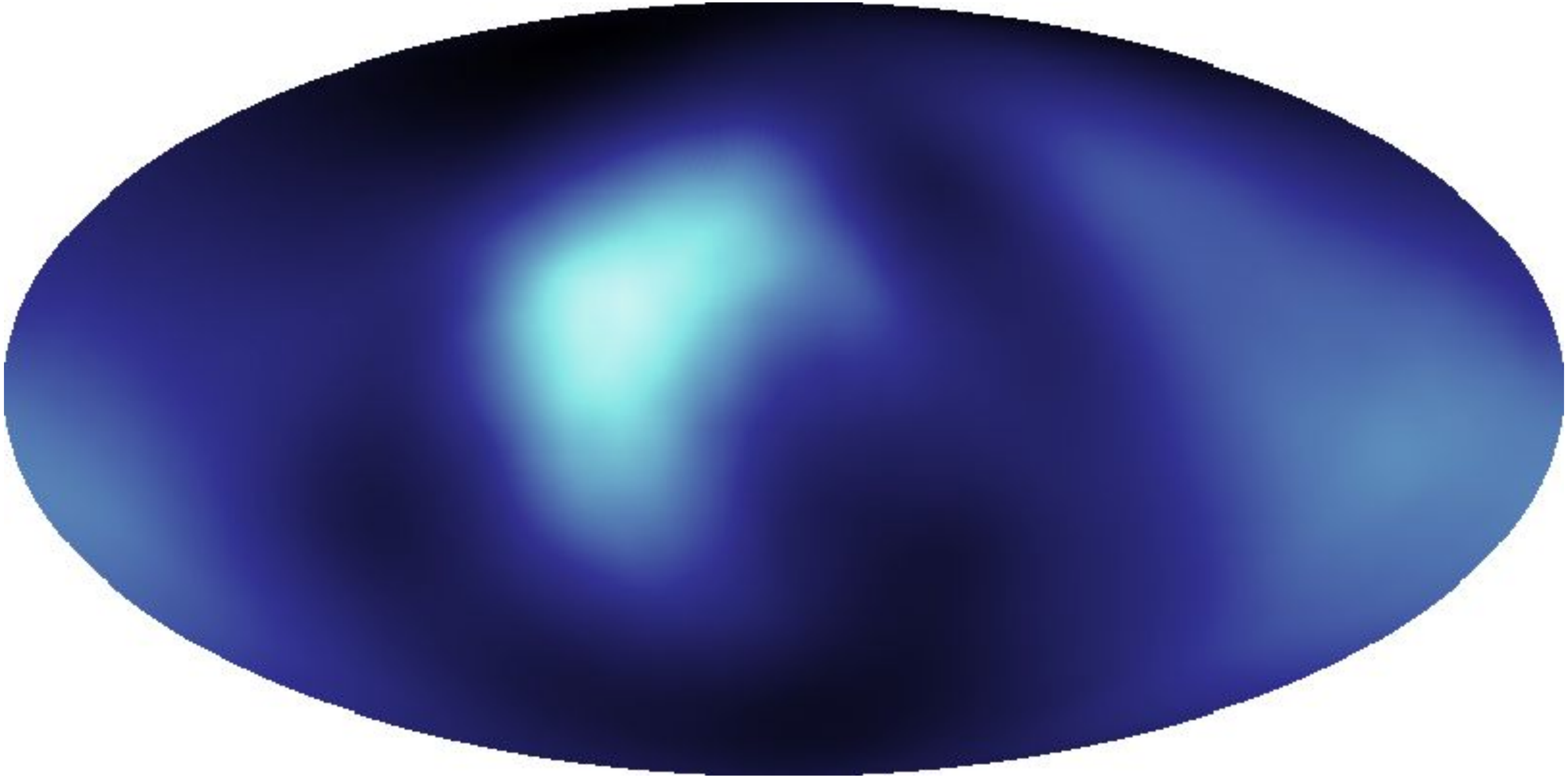
$m_\nu = 0.05$ eV



Elbers et al 2023

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

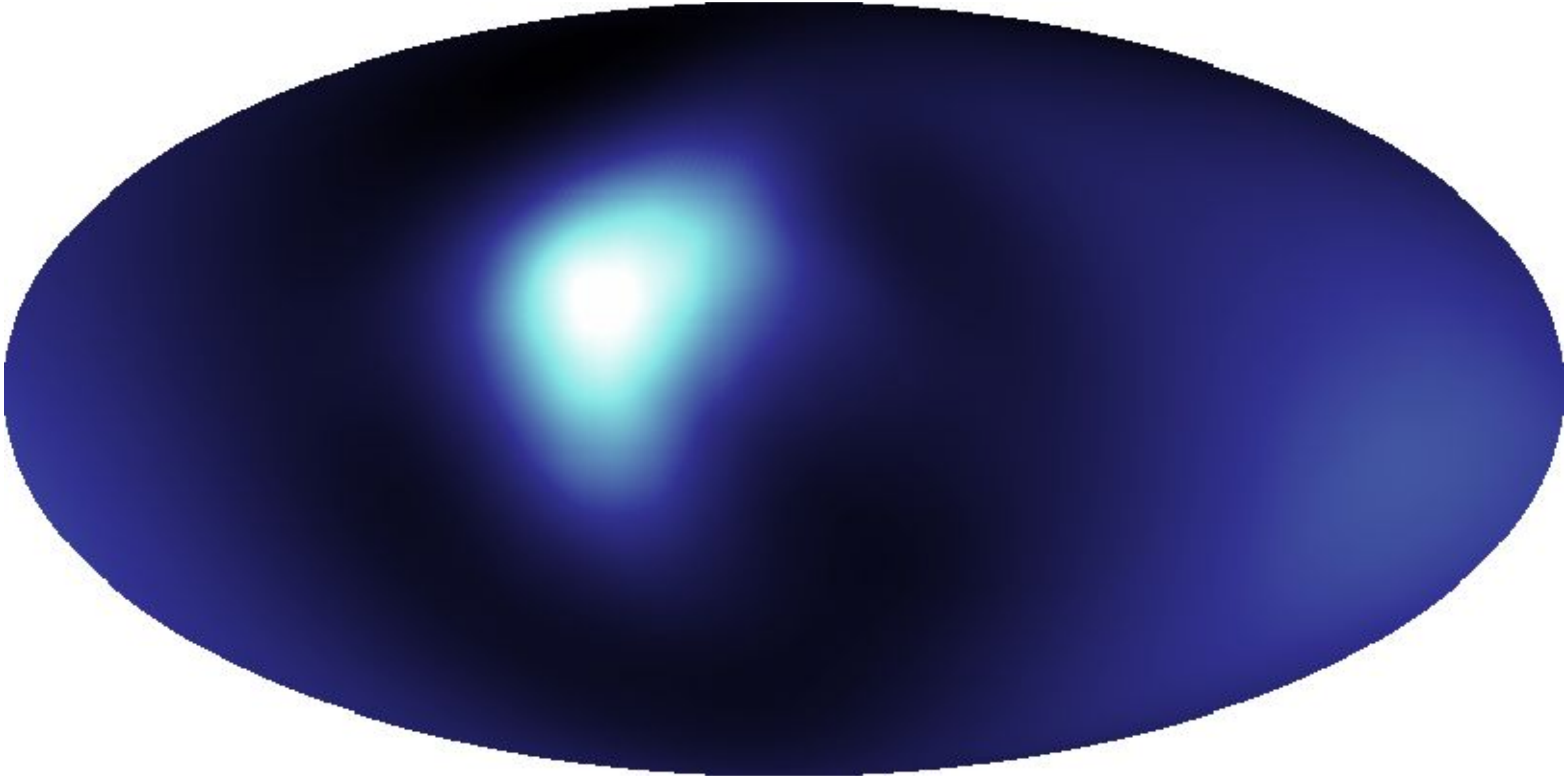
m_ν = 0.1 eV



Elbers et al 2023

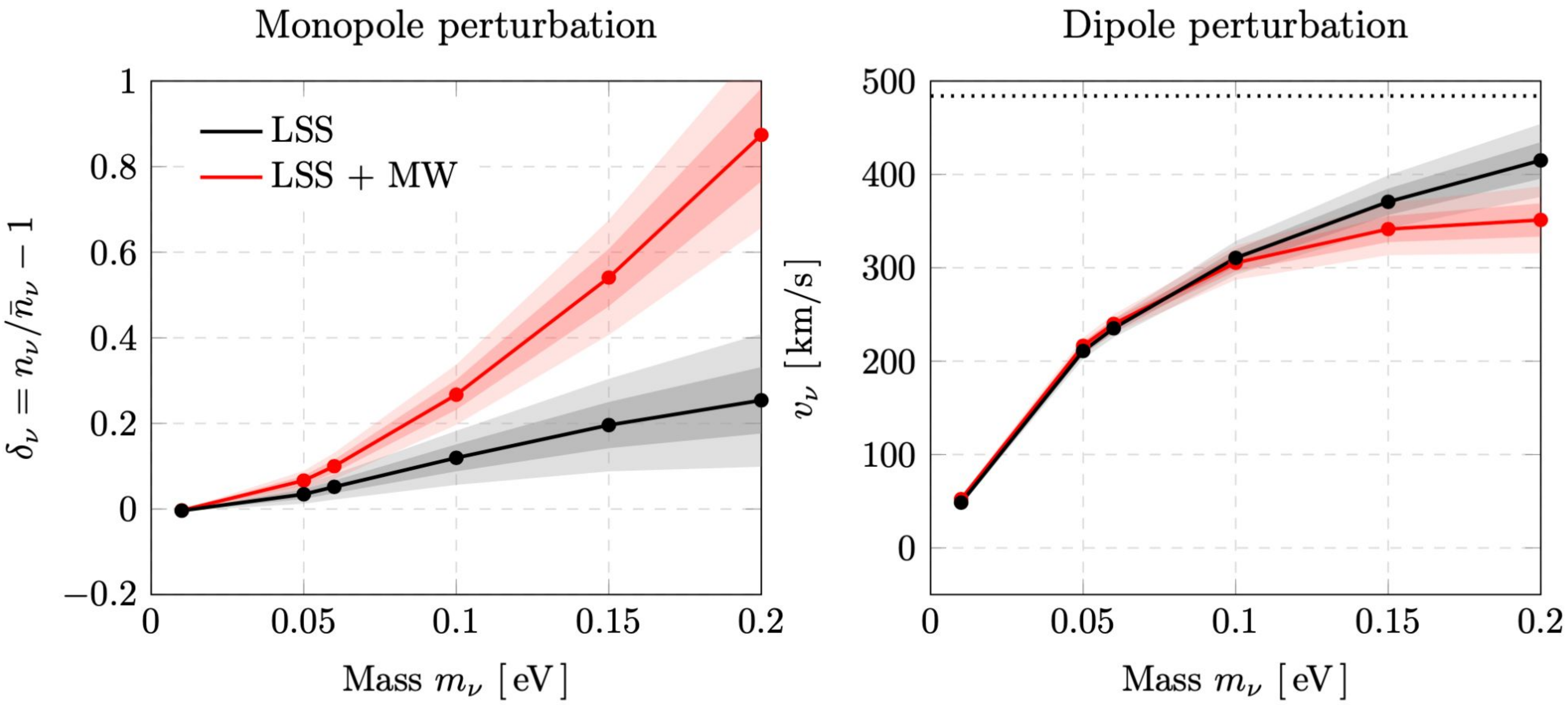
Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

$m_\nu = 0.2$ eV



Elbers et al 2023

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)



Elbers et al 2023

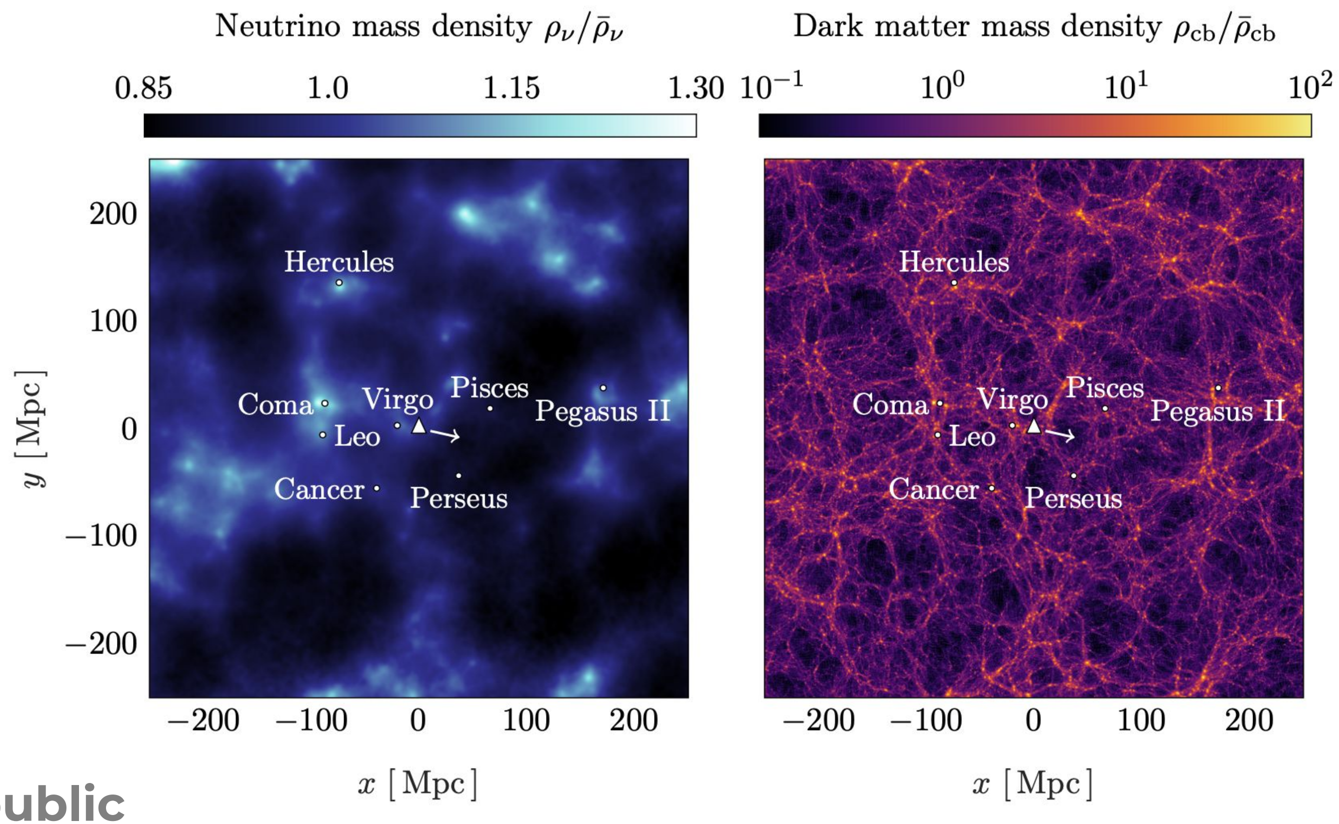
Predicted number of events per year for PTOLEMY

		(LSS)		(LSS + MW)	
$\sum m_\nu$	Ordering	Γ^D_{CNB}	Γ^M_{CNB}	Γ^D_{CNB}	Γ^M_{CNB}
0.06 eV	(NO)	6.86 ± 0.002	8.11 ± 0.004	6.87 ± 0.002	8.12 ± 0.004
0.10 eV	(IO)	4.33 ± 0.04	8.52 ± 0.11	4.45 ± 0.05	8.90 ± 0.12
0.15 eV	(D)	4.24 ± 0.05	8.53 ± 0.12	4.37 ± 0.05	8.92 ± 0.13
0.30 eV	(D)	4.56 ± 0.13	9.07 ± 0.26	5.16 ± 0.14	10.3 ± 0.29
0.45 eV	(D)	4.86 ± 0.22	9.70 ± 0.44	6.26 ± 0.27	12.5 ± 0.54
0.60 eV	(D)	5.10 ± 0.32	10.2 ± 0.63	7.61 ± 0.44	15.2 ± 0.88

Elbers et al 2023

Reconstruction of anisotropies in the neutrino sky (Willem Elbers)

Preliminary Work!



Data will be public

2 × 9 × 6 simulation files, corresponding

- 2 with and without Milky Way,
- 9 different posterior realizations
- 6 different neutrino masses.

Elbers et al 2023

06

Summary & Conclusion

Summary & Conclusion

1 Physical forward modeling and field-level inference

- Goes beyond summary statistics
- Utilizes all of the data
- Performs causal inferences

2 Improved inferences

- tight constraints of cosmological parameter
- Cluster mass measurements and large scale structure environments
- Detailed insights into galaxy formation
- Turning LSS into a physics lab

3 Outlook and work needed

- Small-scale galaxy biasing
- Improving inference speed using ML
- Scaling to next-generation galaxy surveys and data volumes
- A complete and joint characterization of the LSS phenomenology

THANKS

visit us at:
https://aquila-consortium.org/