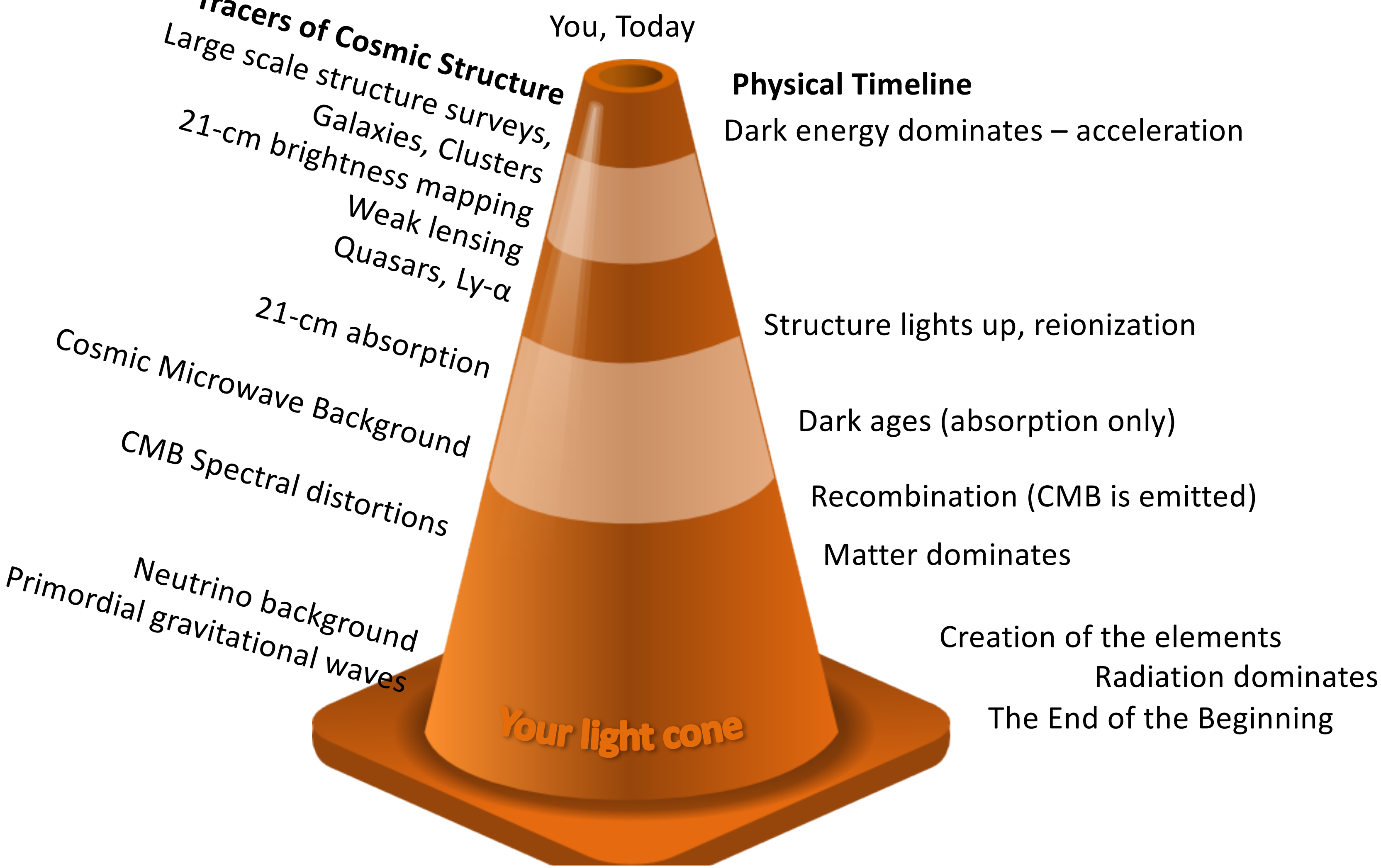
Physics meets ML to solve cosmological inference

Benjamin D. Wandelt

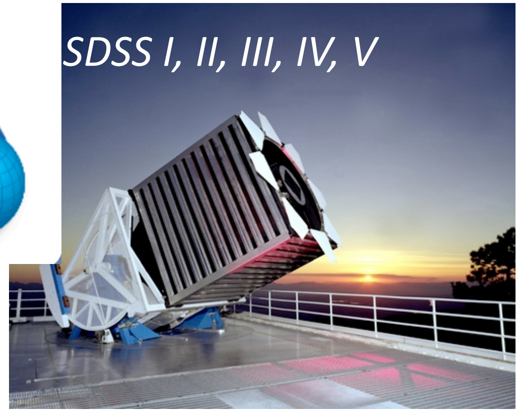
with Suvodip Mukherjee, Tom Charnock, Doogesh Kodi Ramanah, Justin Alsing, Marius Millea, Guilhem Lavaux, Stephen Feeney, Ethan Anderes, Jens Jasche,….

The big picture



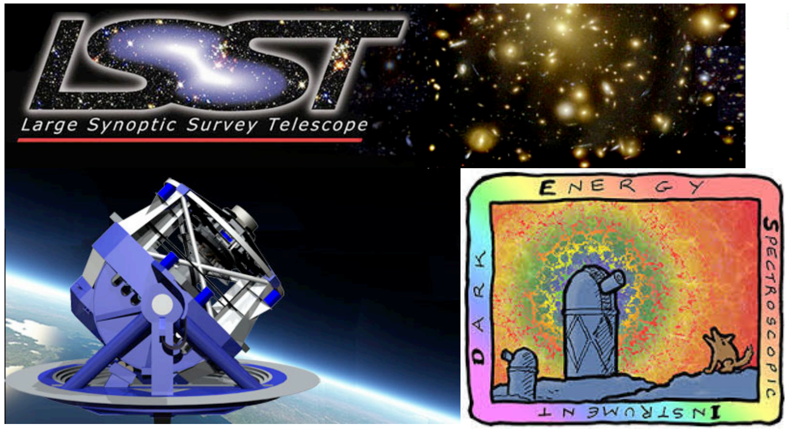
We live in the era of cosmological data

PLANCK
esa
EUCLID
Mapping the Geometry of the Dark Universe
SKA
SQUARE KILOMETRE ARRAY



MegaPrime
CFH
MegaCam
Legacy Survey
CFHTLS

CMB-S4
Next Generation CMB Experiment



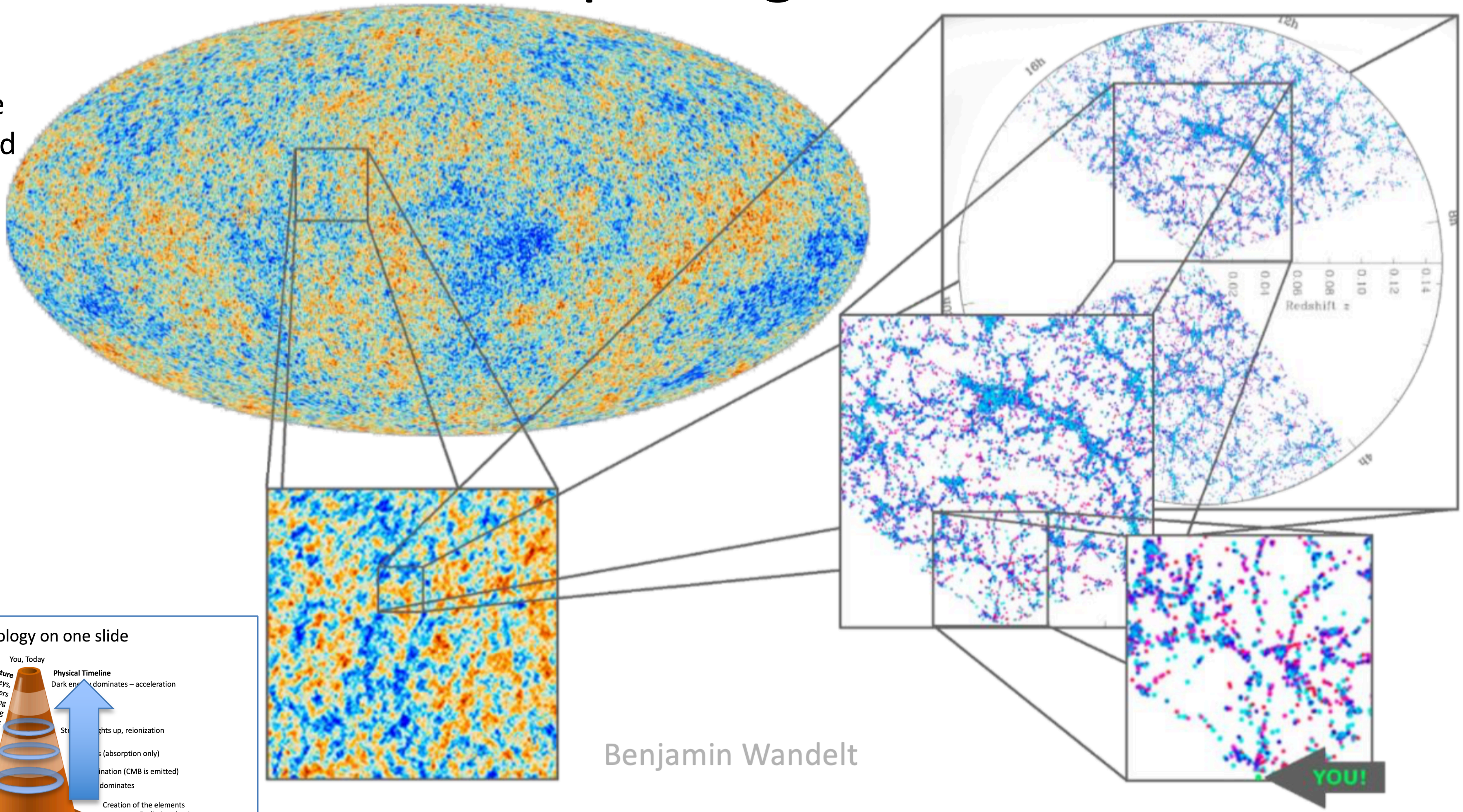
THE DARK ENERGY SURVEY
HYPER SUPRIME-CAM

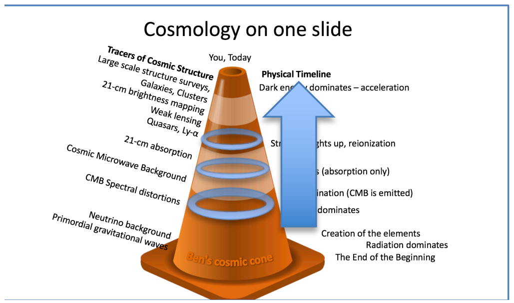
(Your favorite survey here)

Cosmological data covers a hierarchy of scales on the past light cone

Cosmic Microwave Background

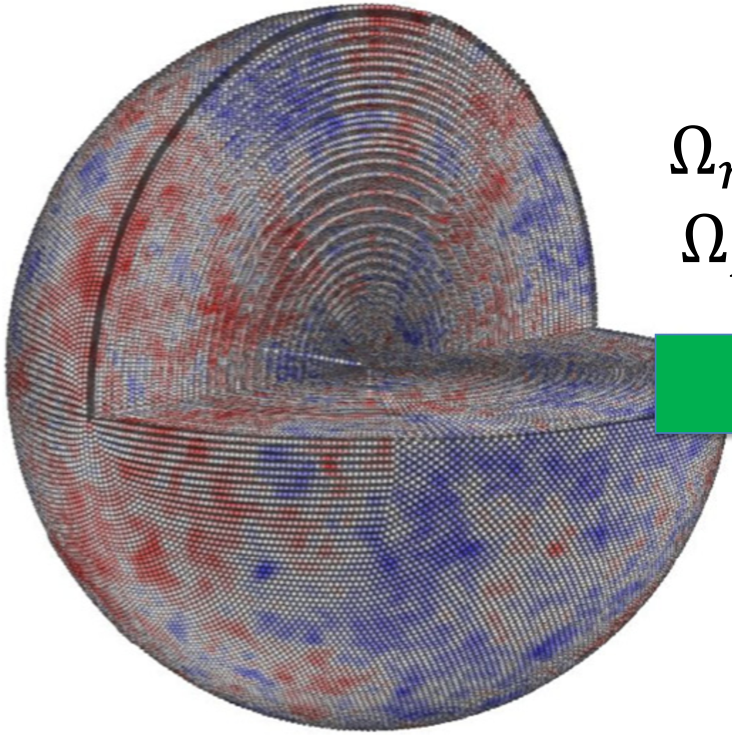
Galaxy surveys





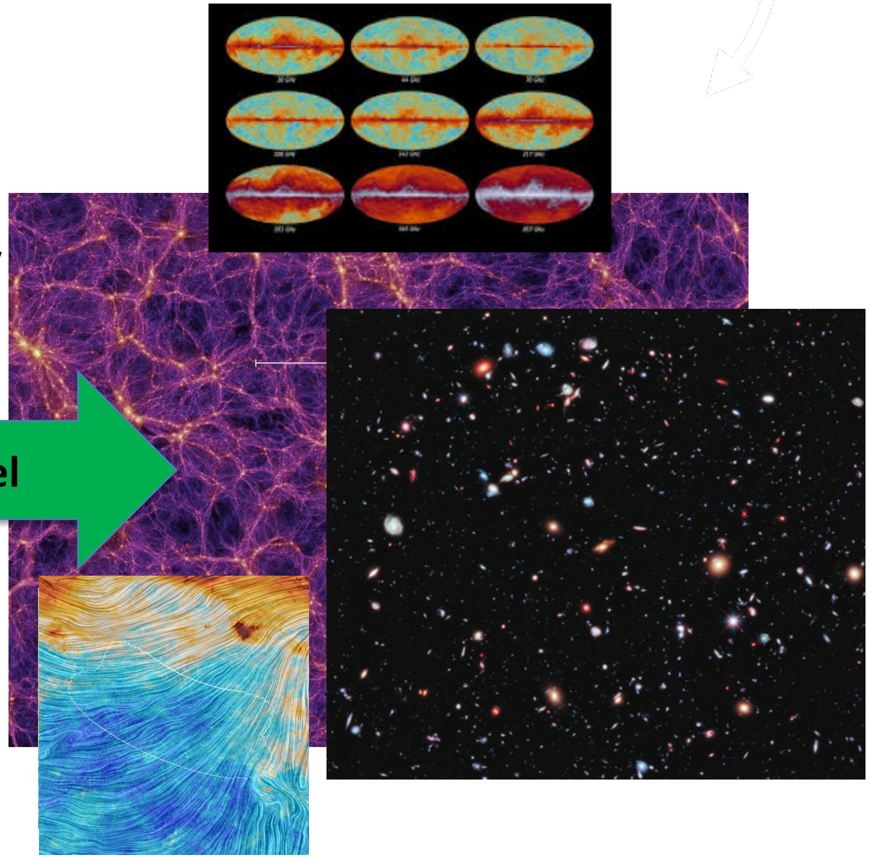
Cosmology

$A_s, n_s, r, f_{nl},...$



$\Omega_m, \Omega_b, m_\nu, ...$
$\Omega_\Lambda, w_0, w_a, ...$

Forward model



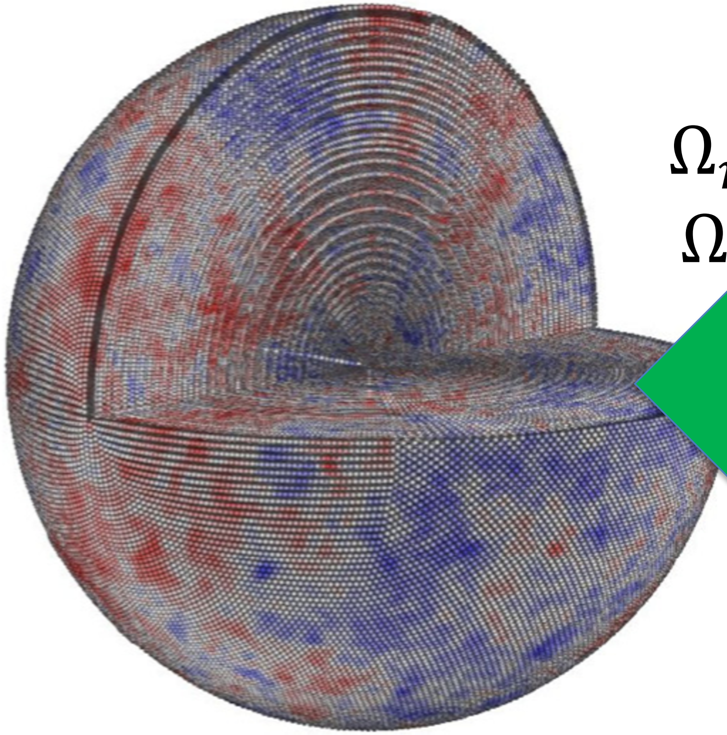
Initial conditions of the universe

The observed universe

Benjamin Wandelt

Cosmological inference: solving the non-linear inverse problem

$A_s, n_s, r, f_{nl}, \ldots$

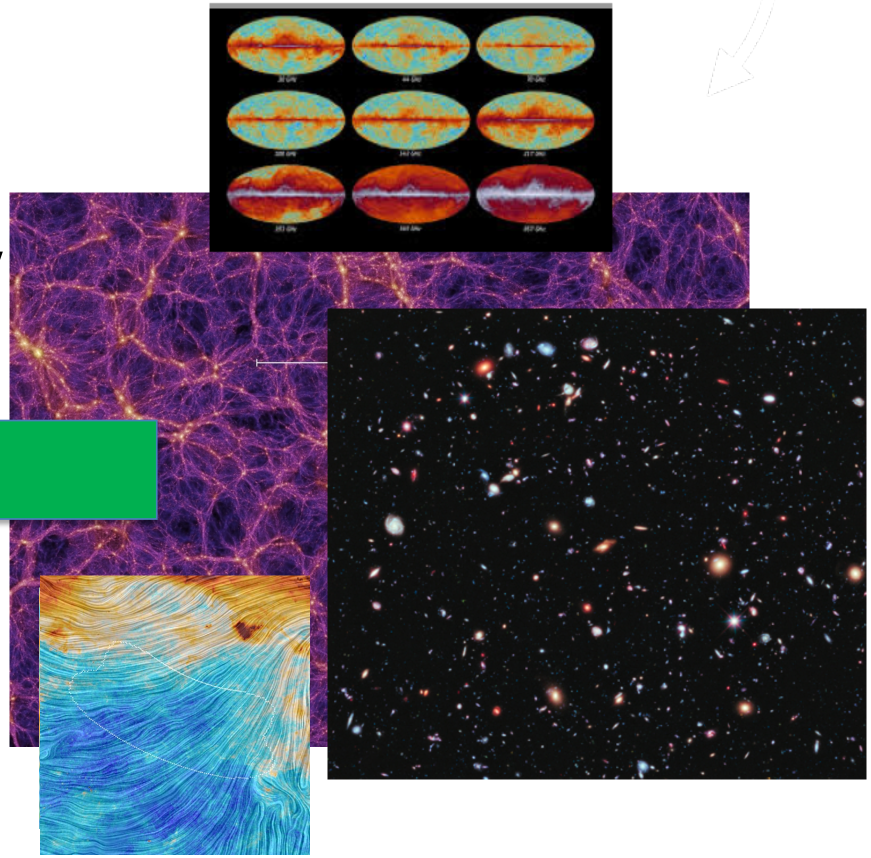


Initial conditions

$\Omega_m, \Omega_b, m_\nu, \ldots$

$\Omega_\Lambda, w_0, w_a, \ldots$

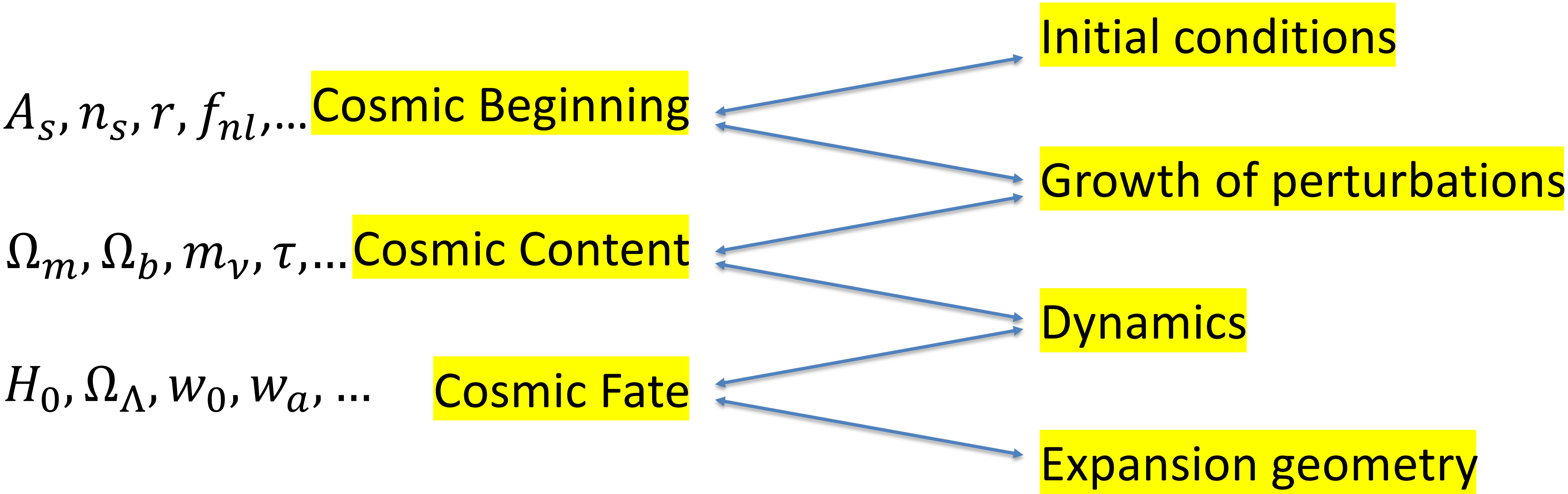
Inference



The observed universe

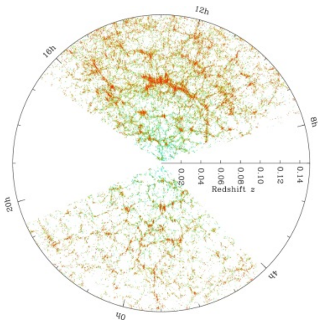
Benjamin Wandelt

The goals of cosmological inference

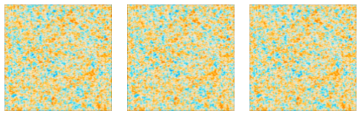


Now have a deep tool set to solve cosmological inference

Full, detailed, likelihood-based forward model



BORG →

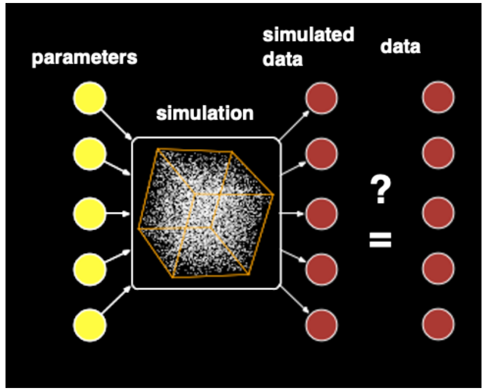


Initial condition reconstruction
Non-Gaussian, field-based AP

Simulation-based inference/ Likelihood-free inference

Freedom to model

Tools:
pyDELFI
Moment Networks

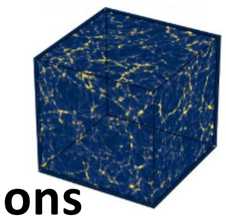


Global generalization of Alcock Paczynski test

A new multi-tracer Global AP test!

$$\Xi(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-g}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \\ \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}^T_{\text{g-sn}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{sn-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \end{pmatrix}$$

A new Global AP-test in D_L-space



Physical Insight

New, robust cosmological tests:
Field-based Alcock-Paczynski (AP) on light-cone
Multi-probe cross-correlation AP

Simulation image from Illustris-TNG

Machine learning innovations

Neural Physical Engines
Neural Density Estimation
Information Maximizing Neural Networks
Wasserstein Generative Adversarial Networks

Simulation innovations

Perfectly parallel n-body sims
Super-resolution of n-body sims
Variance reduction (CARPool)
Quijote and CAMELS projects

Let’s start with the initial conditions

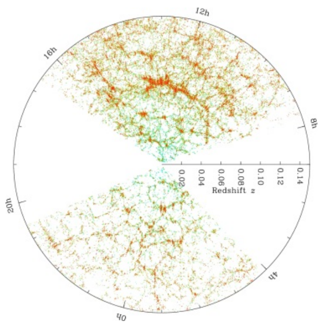
Initial condition reconstruction using a fully generative *probabilistic* forward model of galaxy surveys

BORG: *Bayesian Origin Reconstruction from Galaxies*

- Gaussian prior + **Gravity** + likelihood for galaxies

  (includes particle-mesh or LPT gravity solver, survey model, bias model, automatic noise level calibration, selection function, mask, ...)

- Hamiltonian Markov Chain **with >10^7 parameters**

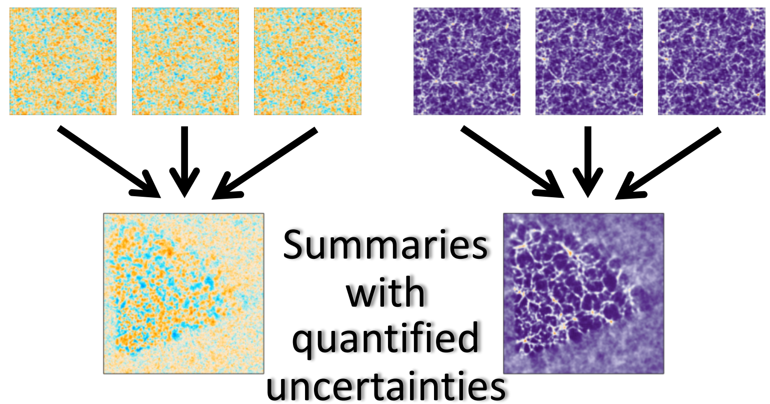


Observations

(galaxy catalog + meta-data: selection functions, completeness...)

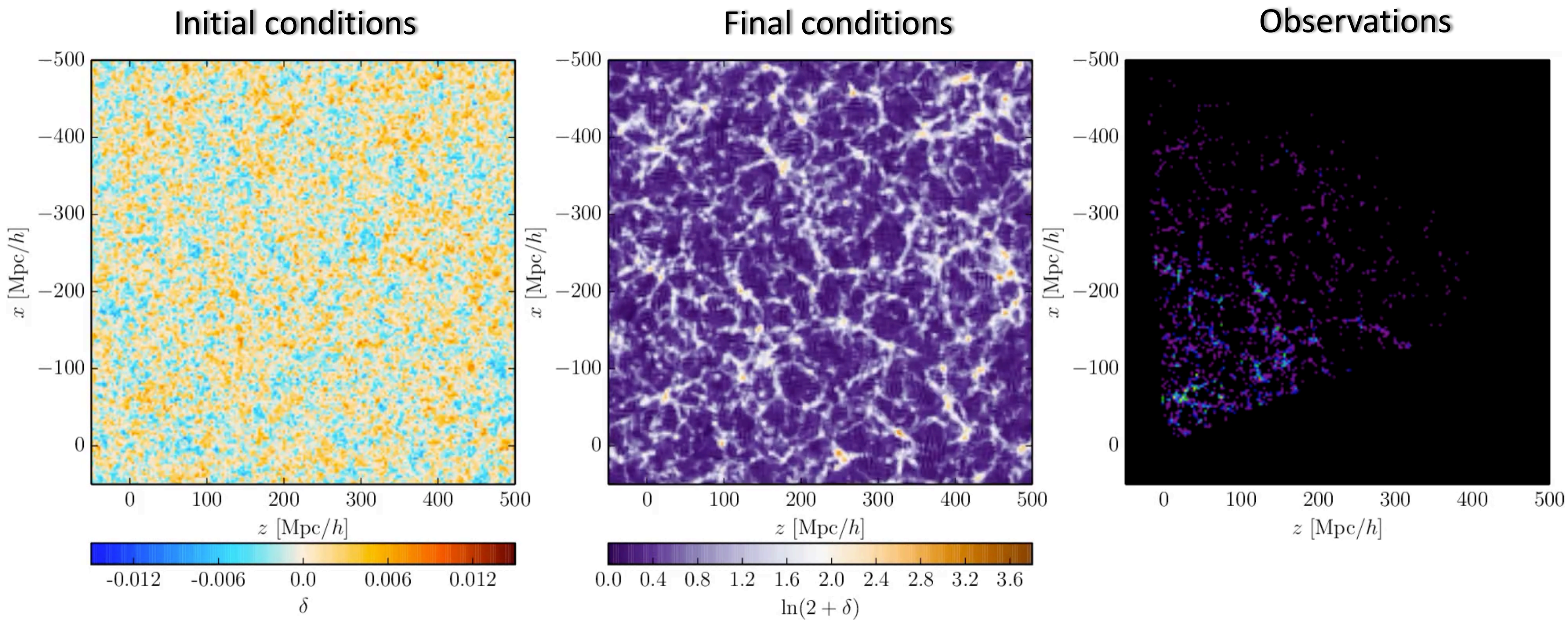
→ BORG → Initial conditions, and inferred dark matter densities

z=100 z=0



Jasche & Wandelt 2013, arXiv:1203.3639
Jasche, Leclercq & Wandelt 2015, arXiv:1409.6308
Lavaux & Jasche 2017...

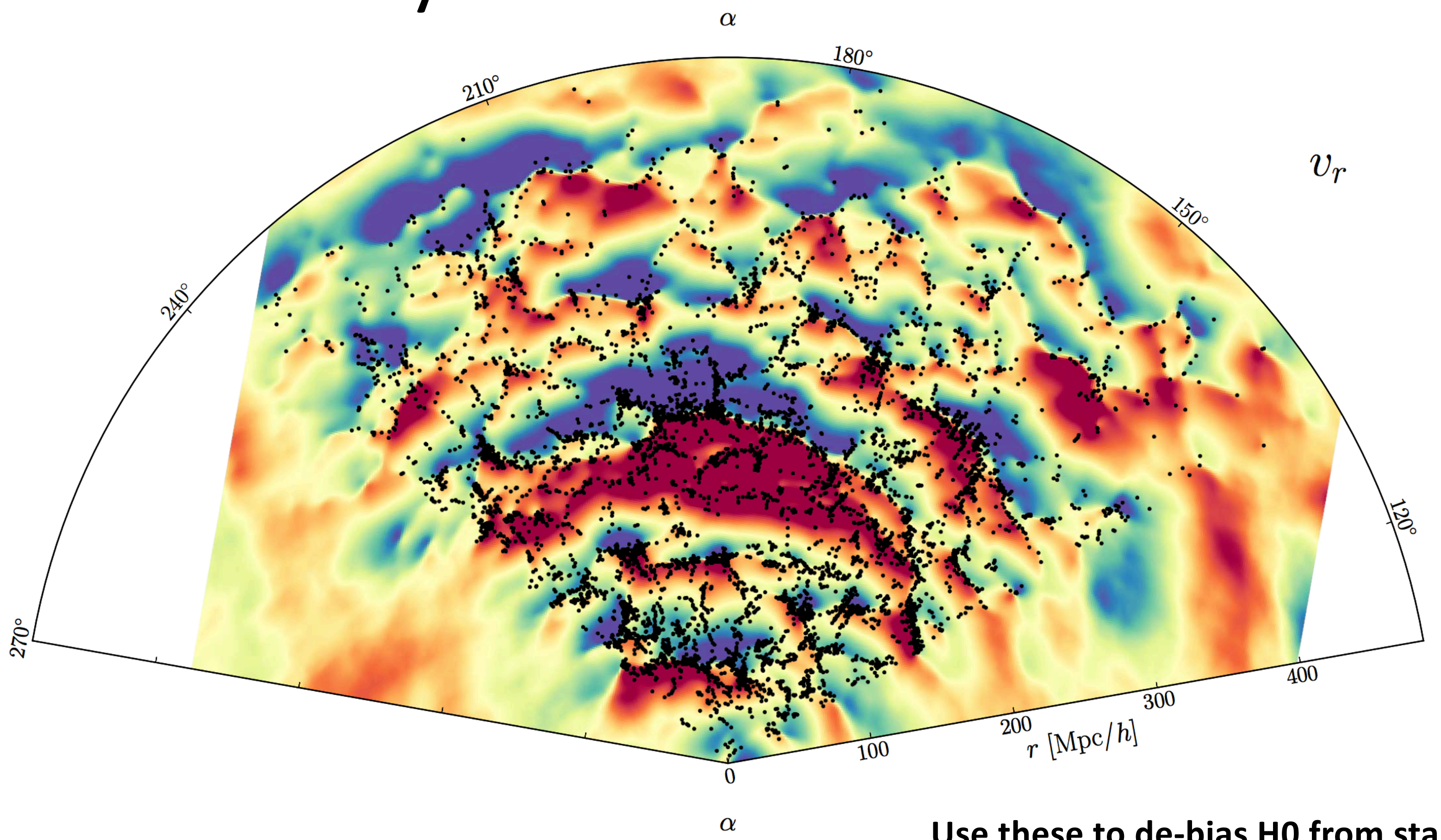
Bayesian reconstruction of cosmological initial conditions with BORG



from Jasche, Leclercq & Wandelt 2014, arXiv:1409.6308

See full bibliography and current status at **http://aquila-consortium.org**

Example Bayesian LCDM results: dynamical velocities



Leclercq et al. 2017

Use these to de-bias H0 from standard sirens: Mukherjee et al arXiv:1909.08627

So is that it – are we done? Problem solved?

- These runs are for fixed fiducial cosmology.
- The full statistical power even of current data is so enormous that even with a very detailed data model, the inference of cosmological parameters is still dominated by systematics.
- Need:
  - more reality in the data model; and
  - better ability to project/cut/mask the data for cosmological inference to become insensitive to remaining model error

Why is this so hard?

The challenges of learning from large scale structure

Limited information – only one universe! Need careful treatment of “cosmic variance” uncertainties

Non-linearity – affects most of the modes in the late universe

Bias – we observe tracers, not the matter

Non-Gaussianity – signal, noise, foregrounds

Large data sets – observational rather than experimental and often indirect

Systematics – astrophysical “contaminants,” instrumental and observational effects

Machine learning to the rescue

Fully *ab initio,* physics-based models like BORG allow the tightest possible confrontation of models and data.

But is it really practical to write down a likelihood that includes ***everything?***

Principled used of Machine Learning (ML) can help in connecting physical models to data.

Neural Physical Engines: Modeling **bias** with ML

- We designed a new type of neural network to learn (cosmological) physics
- The network encodes relevant approximate physical symmetries/constraints
  - Translation invariance
  - Local rotational invariance
  - Locality
- A neural network with only **17 parameters**!

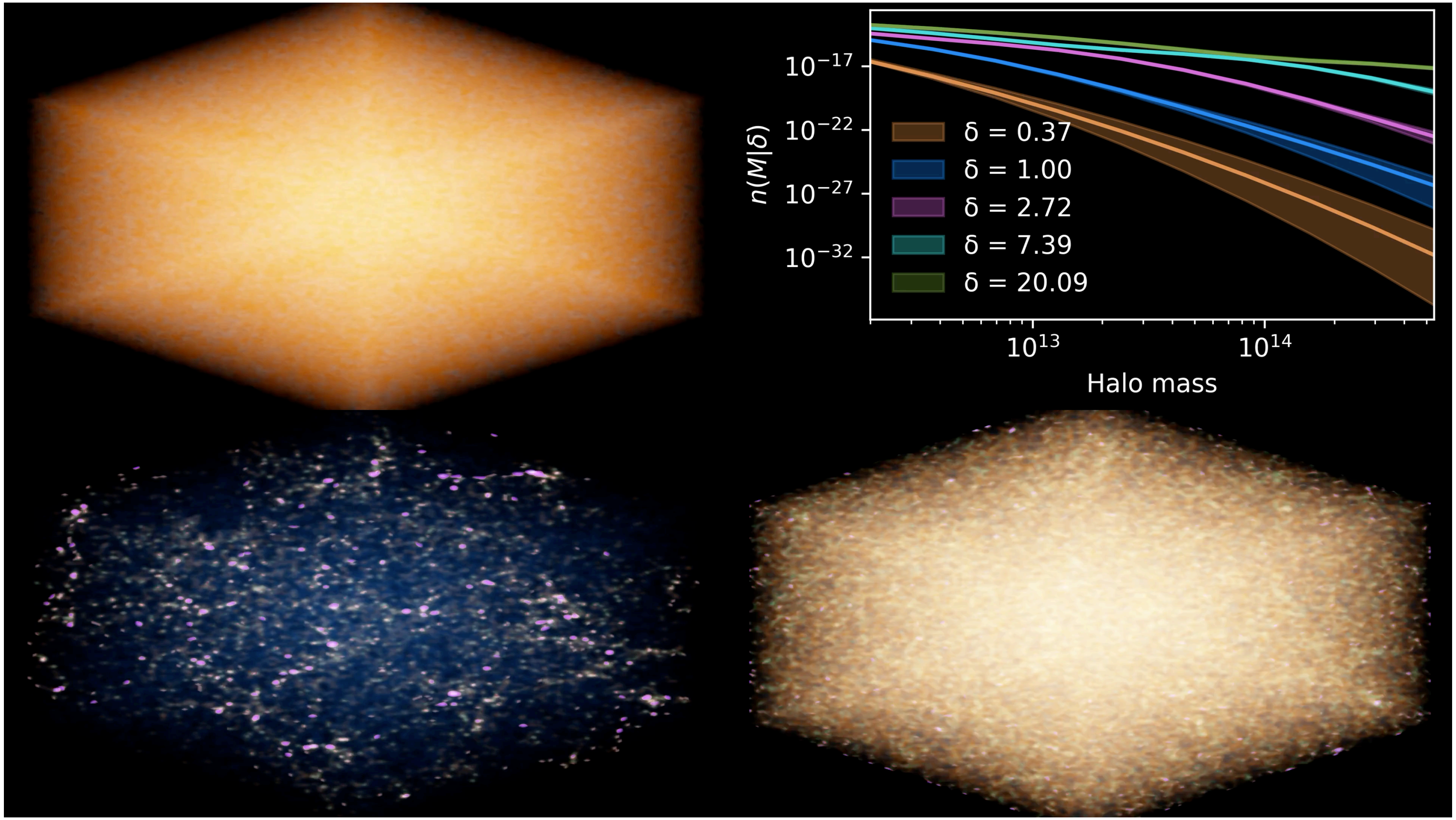
- Use it in the BORG framework **as a bias layer to map DM density to halos**
  Charnock, Lavaux, Wandelt, Boruah, Jasche, Hudson (arXiv:1909.06379)

- *This allows **zero shot learning**:* needs *no* training data!
- A fully Bayesian neural network with data-driven MCMC inference of network parameters and cosmological initial conditions

Neural physical engines for inferring the halo mass distribution function
Charnock, Lavaux, Wandelt, Boruah, Jasche, Hudson (arXiv:1909.06379)

DM reconstruction (shown at z=0) and Initial conditions (not shown)

Simulated data: halo distribution



Neural forward model of halo distribution within BORG

Too much, too fast?
Let's relax and focus on geometrical tests

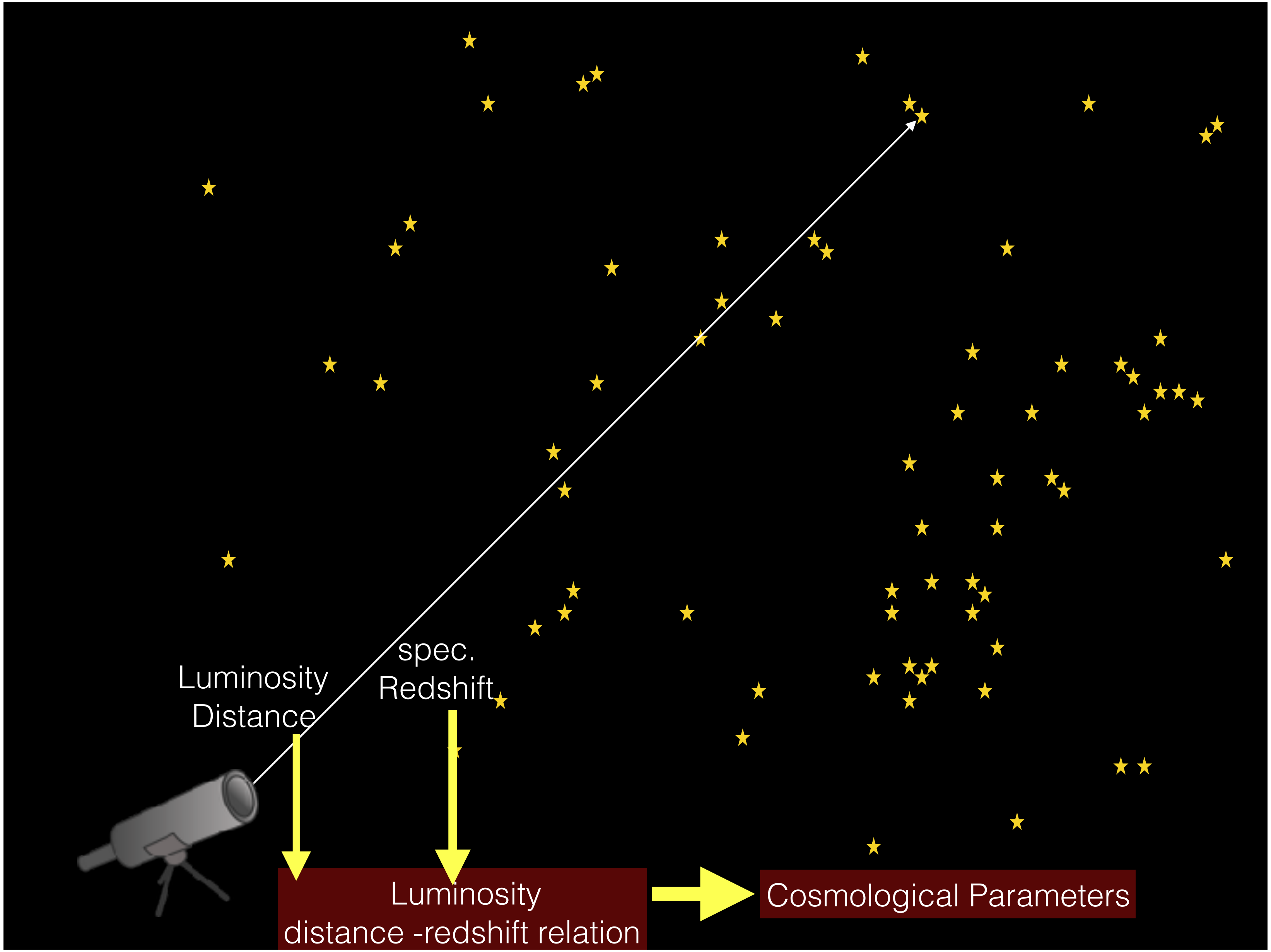
Cosmographic/geometric tests probe aspects of the data that are robust to model misspecification

This avoids having to model the full complexity of the data.

Back to Cosmology 101

- Chapter 1: Homogeneous and isotropic universe
  - 1.1 FLRW metric
  - 1.2 RW equation
  - …
- Chapter 2: Classical cosmological tests
  - 2.1 Luminosity-distance redshift test

    "Observe an object's luminosity distance and redshift and plot them against each other"


Luminosity
Distance
spec.
Redshift
Luminosity
distance -redshift relation
Cosmological Parameters

Cosmology 101

- Chapter 1: Homogeneous and isotropic universe
  - 1.1 FLRW metric
  - 1.2 RW equation
  - …
- Chapter 2: Classical cosmological tests
  - 2.1 Luminosity-distance redshift test

    "Observe an object's luminosity distance and redshift and plot them against each other"

    But there are no objects in a homogeneous and isotropic universe!

    Clearly need to consider *structure*.

A new way to think about cosmological tests

- Consider two types of tracers
  - A luminosity distance tracer *sn*
  - A redshift tracer *g*
- Let's write down the simplest possible model for the *fields* they trace:
  - Gaussian random field

$$-2\mathcal{L}_{\mathrm{full}}(\boldsymbol{\delta}_g, \boldsymbol{\delta}_{sn}|\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\delta}_g \\ \boldsymbol{\delta}_{sn} \end{pmatrix}^T \boldsymbol{\Xi}^{-1} \begin{pmatrix} \boldsymbol{\delta}_g \\ \boldsymbol{\delta}_{sn} \end{pmatrix} + \ln|\boldsymbol{\Xi}|$$

A new way to think about cosmological tests

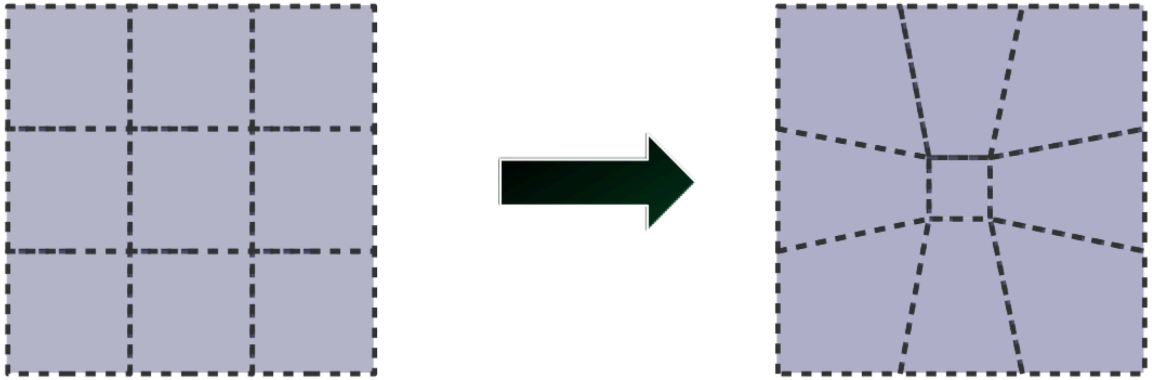
$$-2\mathcal{L}_{\rm full}(\boldsymbol{\delta}_g, \boldsymbol{\delta}_{sn}|\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\delta}_g \\ \boldsymbol{\delta}_{sn} \end{pmatrix}^T \boldsymbol{\Xi}^{-1} \begin{pmatrix} \boldsymbol{\delta}_g \\ \boldsymbol{\delta}_{sn} \end{pmatrix} + \ln|\boldsymbol{\Xi}|$$

Assume both cluster and are mapped from comoving coordinates into luminosity distance and redshift space. Then

$$\boldsymbol{\Xi}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-g}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \\ \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}^T_{\text{g-sn}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{sn-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \end{pmatrix}$$

What are the ***Z*** and ***D***?

The transformation matrix ***Z***



Comoving coordinates

$\vec{x}$

Scaled redshift coordinates

$$\vec{\mathfrak{z}}_i = \frac{c}{H_0} z_i \hat{u}_i$$

D is the analogous transformation to distance space.

A new way to think about cosmological tests

$$-2\mathcal{L}_{\rm full}(\boldsymbol{\delta}_g, \boldsymbol{\delta}_{sn}|\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\delta}_g \\ \boldsymbol{\delta}_{sn} \end{pmatrix}^T \boldsymbol{\Xi}^{-1} \begin{pmatrix} \boldsymbol{\delta}_g \\ \boldsymbol{\delta}_{sn} \end{pmatrix} + \ln|\boldsymbol{\Xi}|$$

Assume both cluster and are mapped from comoving coordinates into luminosity distance and redshift space. Then

Global generalization of Alcock Paczynski test

A new multi-tracer Global AP test!

$$\boldsymbol{\Xi}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-g}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \\ \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}^T_{\text{g-sn}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{sn-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \end{pmatrix}$$

A new Global AP-test in D_L-space

Re-Discover the classic luminosity distance-redshift test as a special case of this multi-probe AP test!

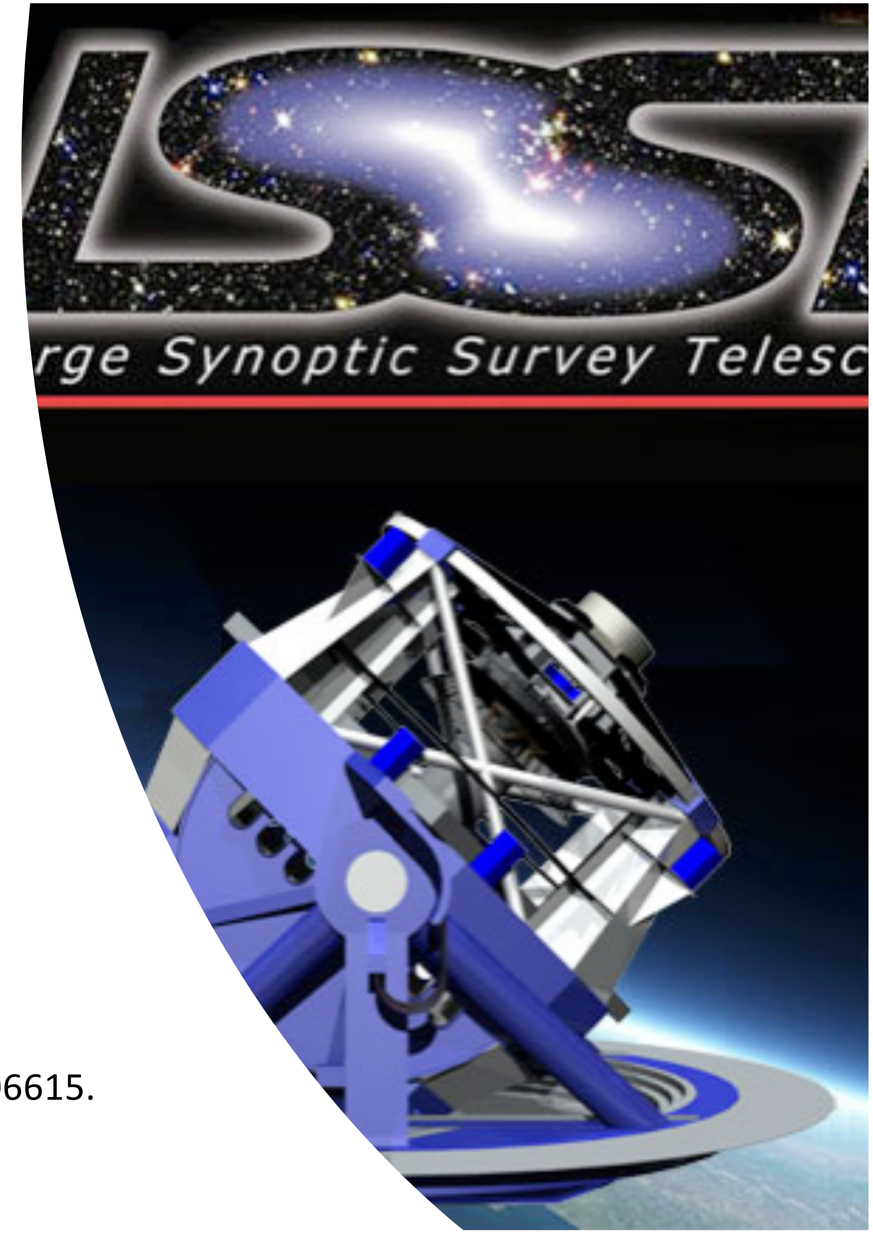
$$\boldsymbol{\Xi}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-g}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \\ \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-sn}}^T\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{sn-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \end{pmatrix}$$

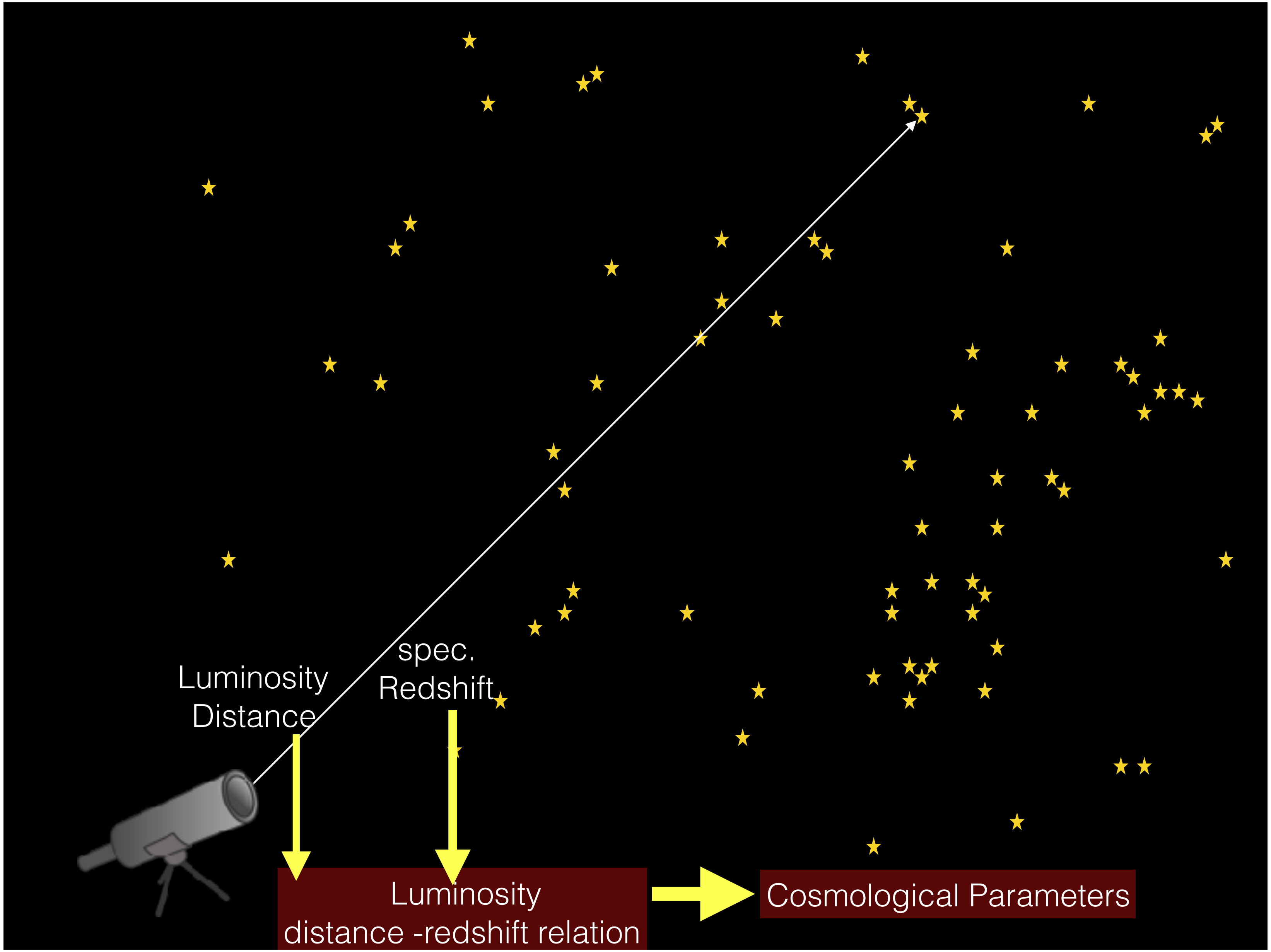
- The multi-probe AP test involves summing over all pairs of distance and redshift tracers.
- If we (incorrectly) ignore spatial clustering by forcing the covariance to be diagonal, we get a single sum with those objects that trace both D_L and z. This is the D_L-z test!
- But galaxies *are* clustered so we can use all pairs -> can get better performance
- *Can exploit this to solve major problems in SN and GW cosmology for the next decade.*

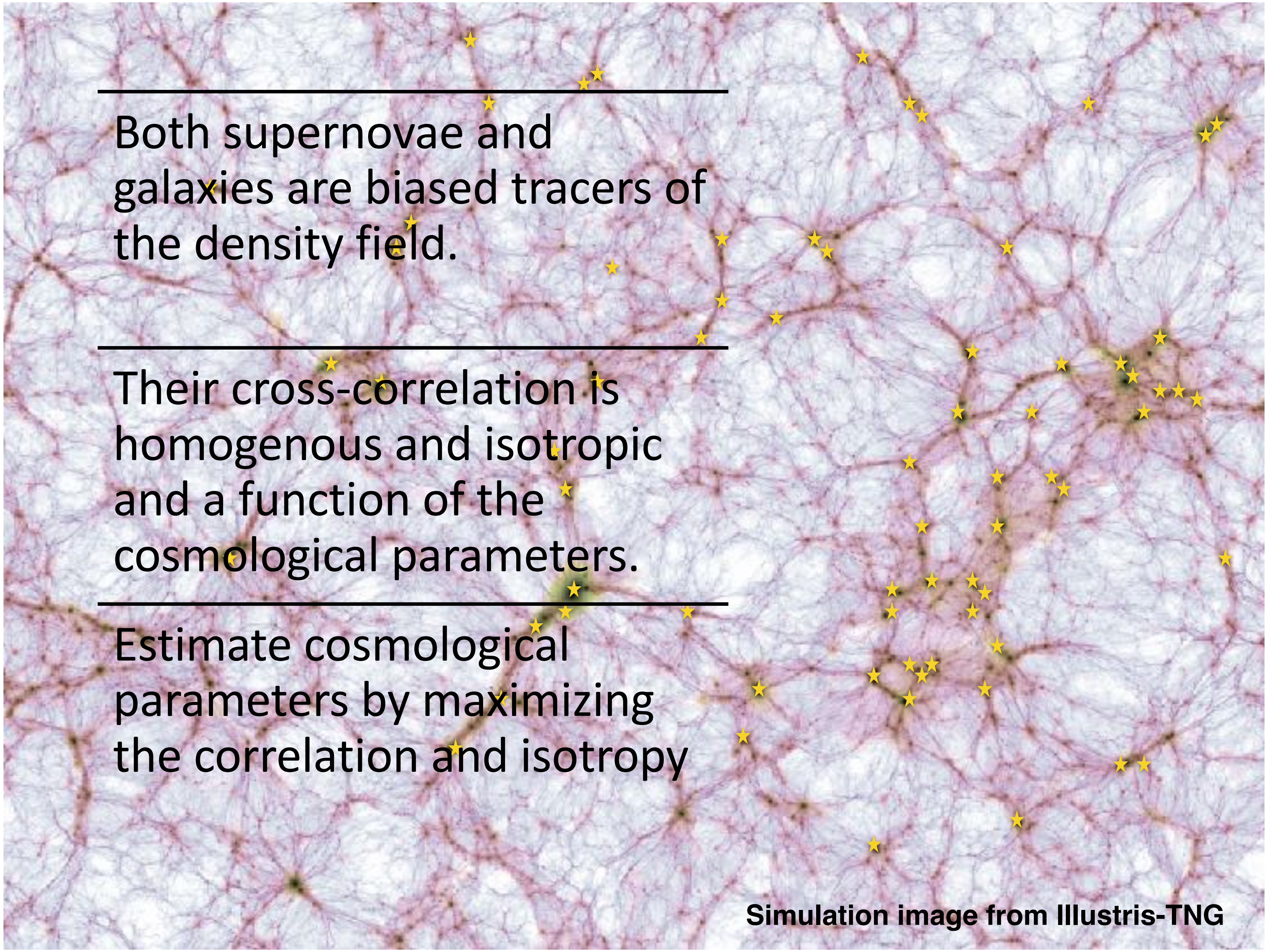
Next-decade supernova cosmology

- Upcoming surveys will have tens of thousands of supernovae
- Far too many to follow all of them up spectroscopically!
- Photometric information leads to type contamination and photo-z systematics.

S. Mukherjee & B. Wandelt, arXiv: 1808:06615.




Luminosity
Distance
spec.
Redshift
Luminosity
distance -redshift relation
Cosmological Parameters



Both supernovae and galaxies are biased tracers of the density field.

Their cross-correlation is homogenous and isotropic and a function of the cosmological parameters.

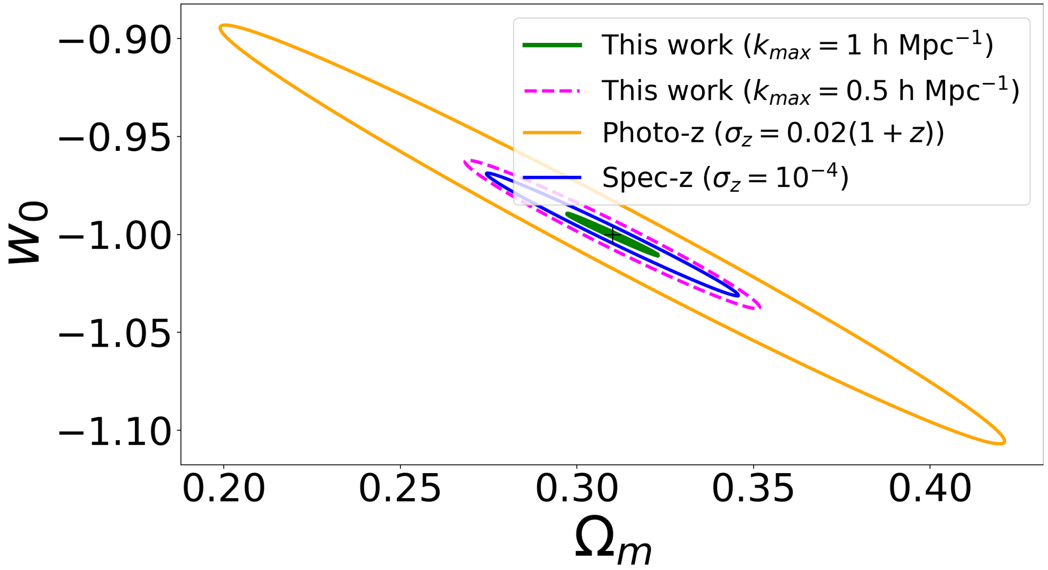
Estimate cosmological parameters by maximizing the correlation and isotropy

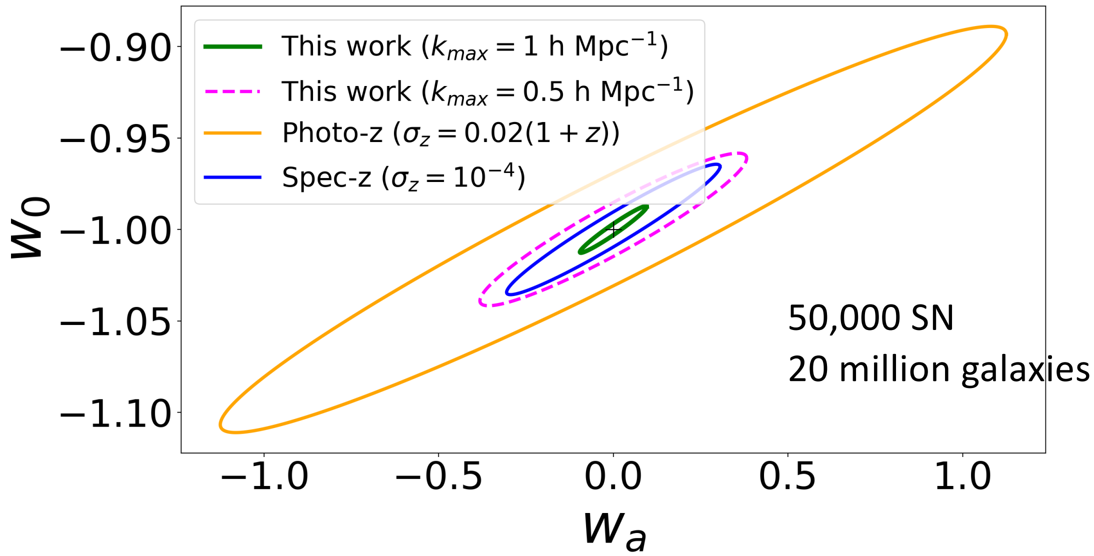
Simulation image from Illustris-TNG

Luminosity distance–redshift test using SN-galaxy cross correlations

- Robust to type contamination
- Insensitive to photo-z systematics
- *Suppression of cosmic variance comes from multi-tracer approach, as expected for background test!*

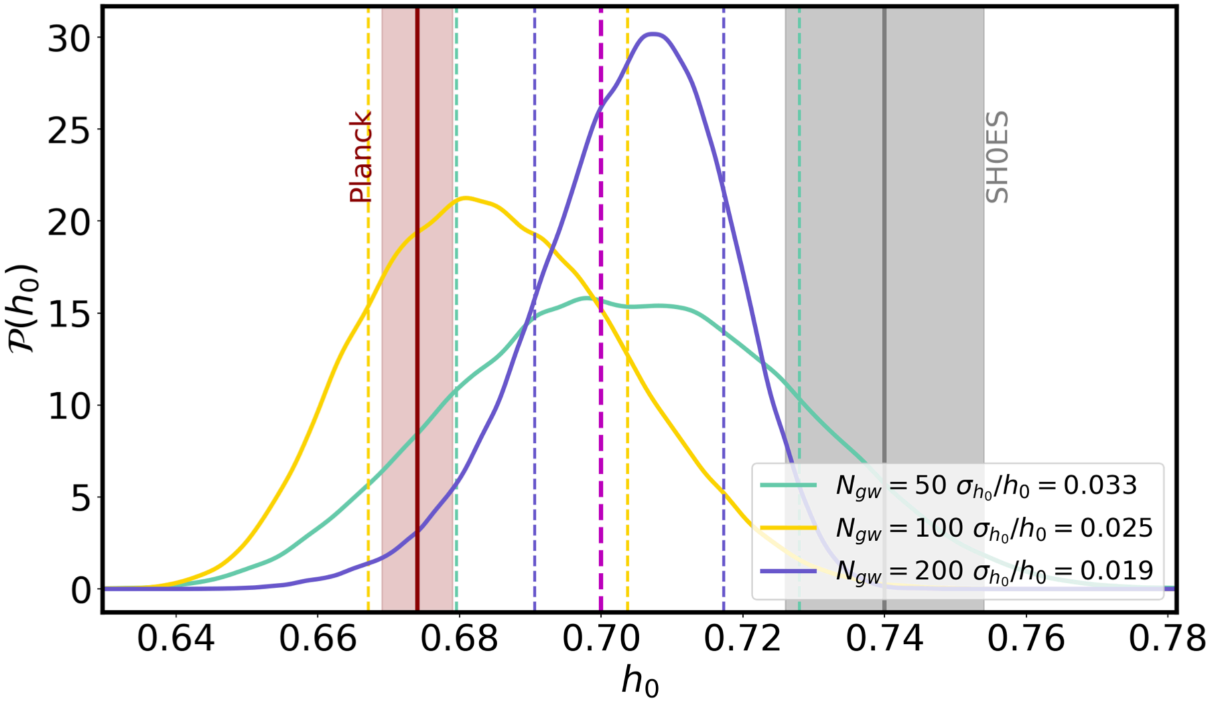
S. Mukherjee and B. Wandelt, arXiv: 1808:06615 .





Dark Sirens with multi-probe AP

- The technique applies to any distance tracer, including dark gravitational wave sirens.
- 200 GW events *without EM counterpart* suffice to reach the same precision on H0 as the SH0ES measurement



Mukherjee, BDW, Nissanke & Silvestri arXiv:2007.02943

Can we use this geometrical approach to do cosmological inference with BORG?

- Going to a geometric approach decouples the "bias" model from cosmological parameters
- By *only keeping the cosmological parameter dependence in coordinate mapping* we can use BORG to do a generalized, non-Gaussian, field-based "Alcock-Paczynski" on the light cone

Alcock Paczynski test

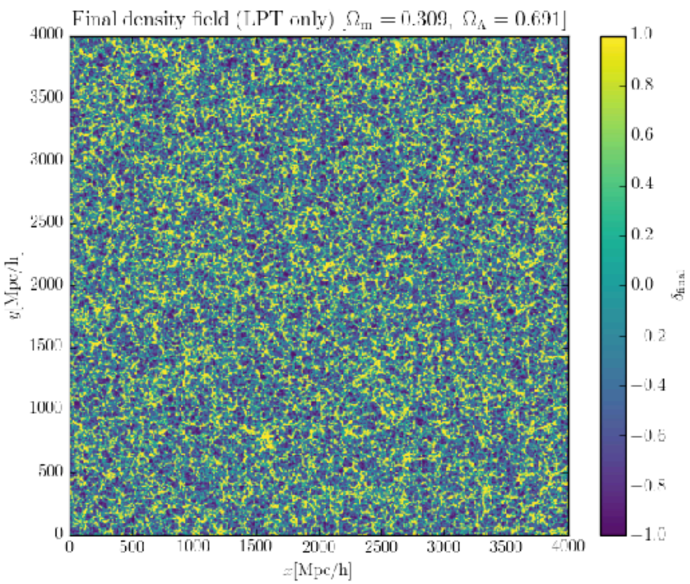
$$\boldsymbol{\Xi}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-g}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{Z}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{g-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \\ \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}^T_{\text{g-sn}}\boldsymbol{Z}(\boldsymbol{\theta}) & \boldsymbol{D}^T(\boldsymbol{\theta})\boldsymbol{\xi}_{\text{sn-sn}}\boldsymbol{D}(\boldsymbol{\theta}) \end{pmatrix}$$

A field-based AP test (not just 2-point stats!)

Coordinate Transformation

$(z, \vec{u})$ **Redshift space** $\xleftrightarrow{\{\Omega_m, \Omega_\Lambda, \Omega_k, \omega, H_0\}}$ $(d, \vec{u})$ **Comoving space**

comoving space



(Alcock & Paczyński 1979)

- Distortions due to assumption of incorrect cosmological parameters
- Structure: **Spherical → Ellipsoidal**
- Statistical distribution: **Isotropic → Anisotropic**

$$d = \int_{z_1}^{z_2} \frac{1}{cH(z)}$$

$$H(z) = H_0(\Omega_{\rm m}(1+z)^3 + \Omega_{\rm k}(1+z)^2 + \Omega_\Lambda)^{\frac{1}{2}}$$

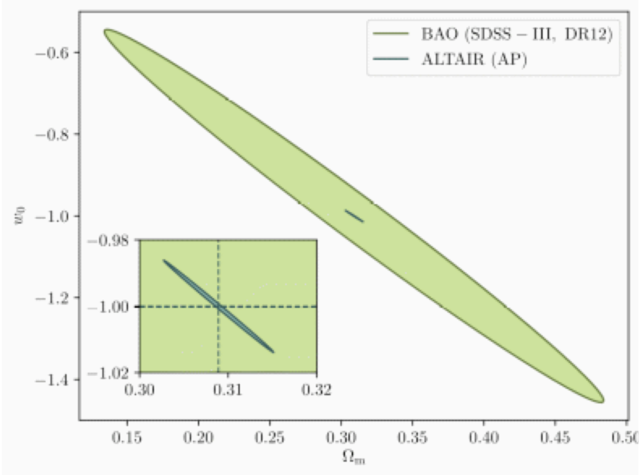
Kodi Ramanah et al., arXiv 1808.07496

High precision inferences

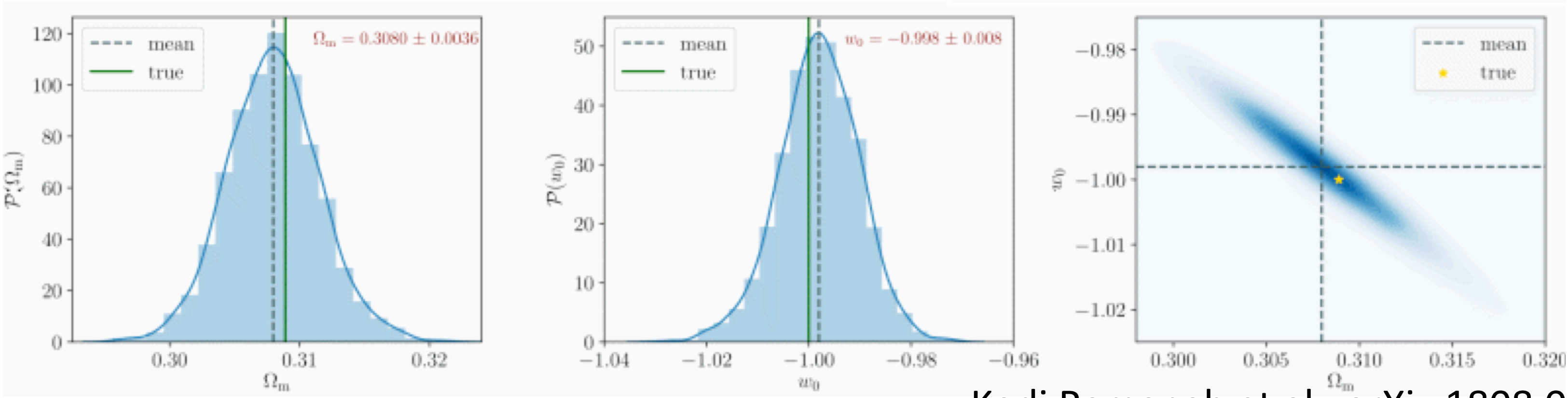
- Probing deep redshift range; geometric distortion due to cosmic expansion is highly informative

$$\{\Omega_{\rm m} = 0.3080 \pm 0.0036, \quad w_0 = -0.998 \pm 0.008\}$$

Comparison to standard BAO constraints

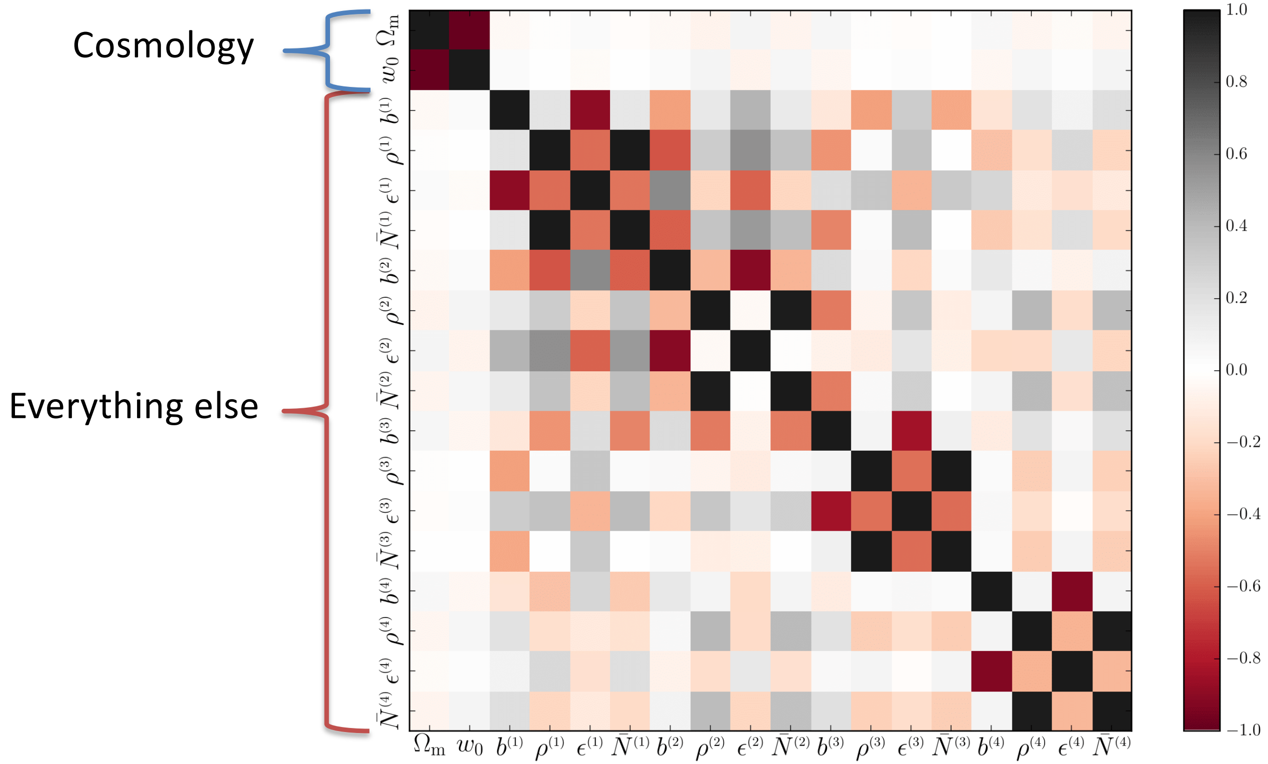


Marginal & joint posteriors



Kodi Ramanah et al., arXiv 1808.07496

Focusing on geometry works: Cosmology and bias parameters decouple!



Kodi Ramanah et al., arXiv 1808.07496

Relaxed?

Good! Let's get back to solving the full problem!

Good! Let's get back to solving the full problem!

The full problem

To succeed we need more freedom than a traditional likelihood approach can provide:

- FREEDOM to make our physical model anything we want
- FREEDOM to project/summarize/cut/mask our data any way we want

Simulating data is **much easier** than deriving an accurate likelihood.

Can we analyse data if all we can do is simulate it?

Simulations are draws from the likelihood

$$P(\boldsymbol{\theta}|\mathbf{d}) = \frac{P(\mathbf{d}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{d})}$$

$$\mathbf{d}^* \leftarrow \text{simulation}(\mathbf{d}^*|\boldsymbol{\theta})$$

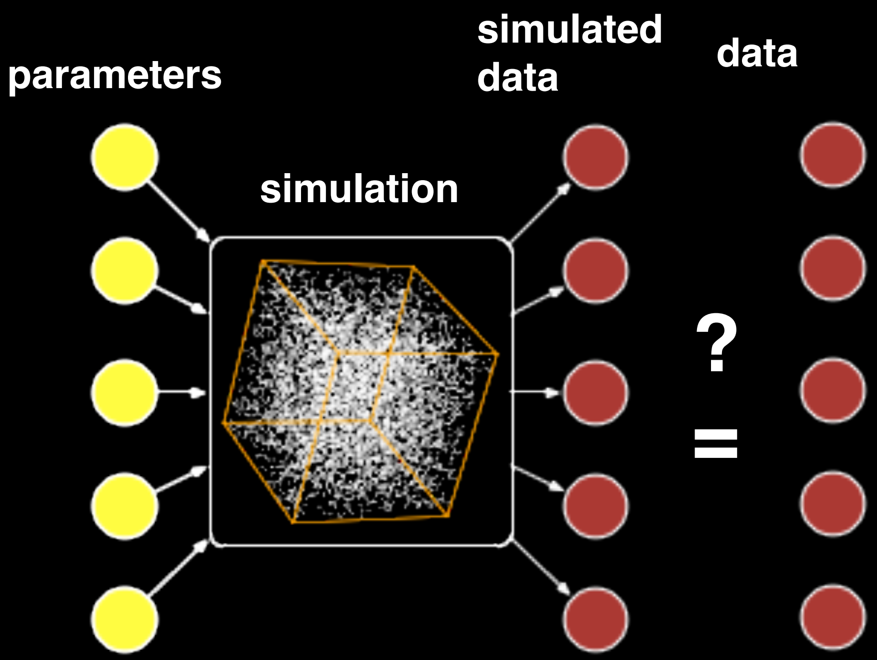
Simulation based (likelihood-free) inference for realistic data models

with Justin Alsing, Tom Charnock, Stephen Feeney, Francisco Villaescusa, Guilhem Lavaux




Benjamin Wandelt

Likelihood-free inference 101



Draw from prior:

$$\boldsymbol{\theta} \leftarrow P(\boldsymbol{\theta})$$

Simulate data:

$$\mathbf{d}^* \leftarrow P(\mathbf{d}^*|\boldsymbol{\theta})$$

If $\rho(\mathbf{d}^*, \mathbf{d}) < \epsilon$

accept;

else:

reject;

In the limit $\epsilon \rightarrow 0, \ \{\boldsymbol{\theta}\} \leftarrow P(\boldsymbol{\theta}|\mathbf{d})$

Benjamin Wandelt

Likelihood-free inference 101

How to reduce data-space?

Draw from prior:

$\theta \leftarrow P(\theta)$

Simulate data:

$\mathbf{D}^* \leftarrow P(\mathbf{D}^*|\theta)$

If $\rho(\mathbf{D}^*, \mathbf{D}) < \epsilon$:

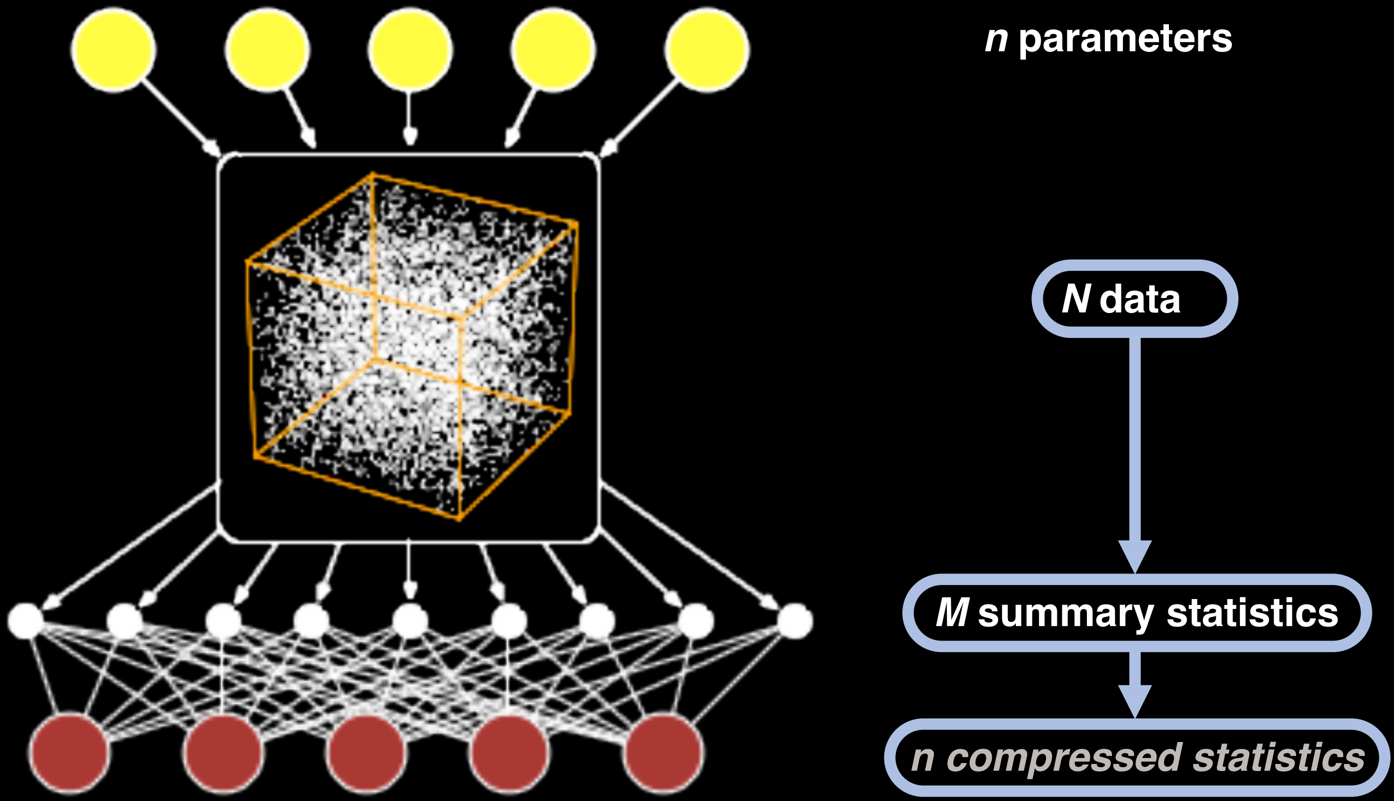
accept;

else:

reject;

In the limit $\epsilon \rightarrow 0, \quad \{\theta\} \leftarrow P(\theta|\mathbf{D})$

Reducing data space: massive data compression



Score compression: Alsing & Wandelt arXiv:1712.00012; Heavens, Jimenez & Lahav 2000

Likelihood-free inference 101

How to reduce data-space? ✅

How to explore parameter-space?

In the limit $\epsilon \to 0, \quad \{\boldsymbol{\theta}\} \leftarrow P(\boldsymbol{\theta}|\mathbf{D})$

Machine Learning to the rescue!

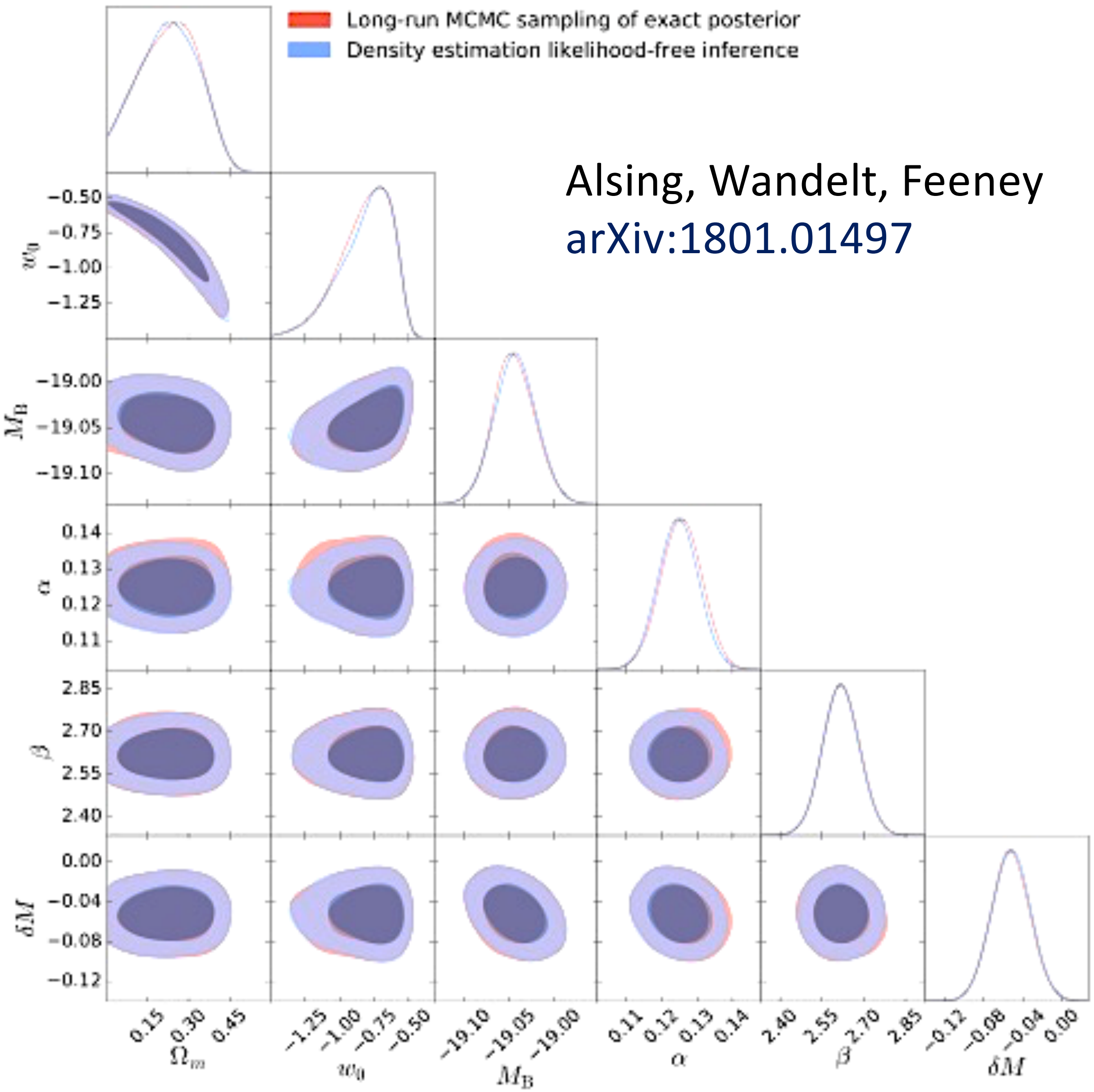
Density estimation Likelihood free inference (DELFI):

Directly learn probability density of parameters and compressed data

Alsing, Feeney & Wandelt arXiv: 1801.01497

DELFI
Posterior
inference
works…

and it's
much
faster than
MCMC!



Alsing, Wandelt, Feeney
arXiv:1801.01497

(O(1000) simulations)

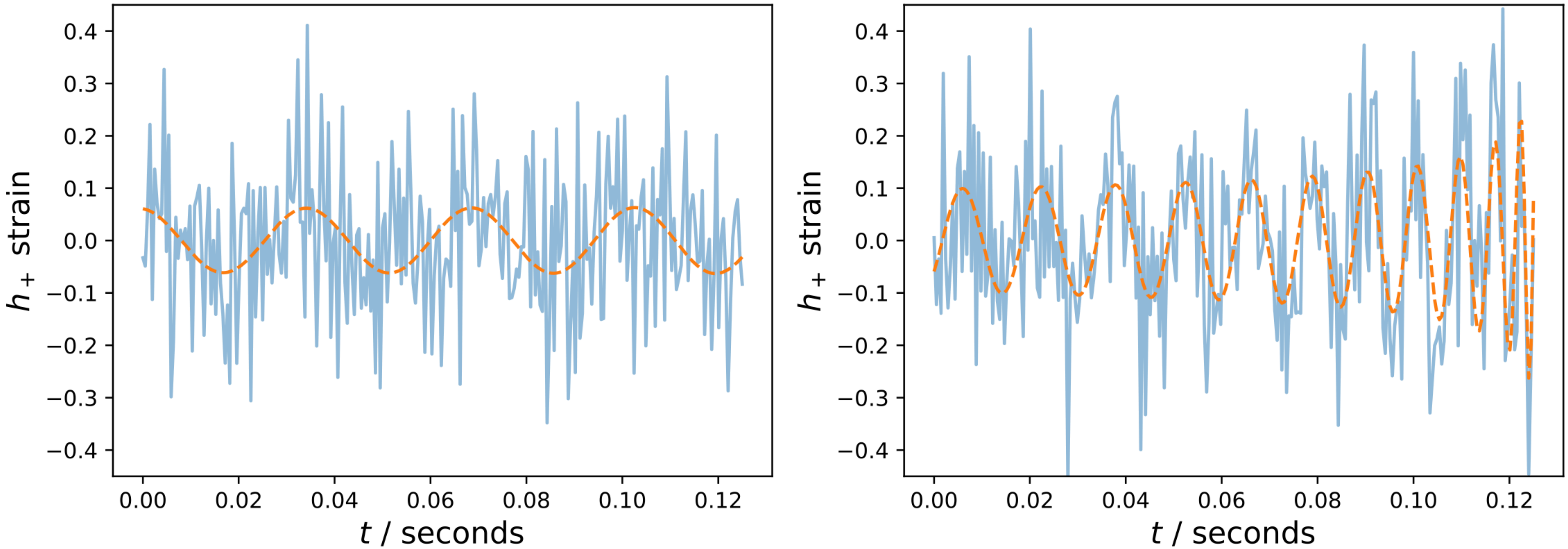
Latest news in DELFI

- New *nuisance-hardened* compression greatly reduces required number of simulations and allows many more parameters (Alsing & Wandelt arXiv:1903.01473).

- New version of DELFI now released including neural density estimators to fit the likelihood (Alsing, Charnock, Feeney, Wandelt arXiv:1903.00007)
  - Also includes active learning for deciding where to run simulations

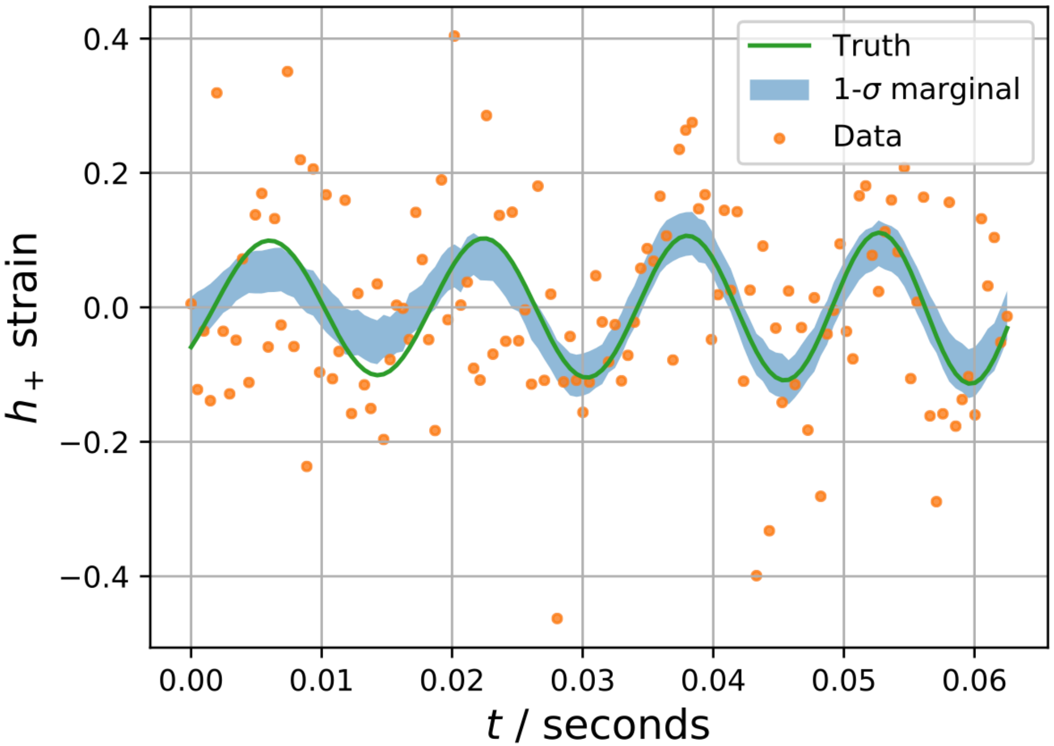
- Can go to much higher number of parameters through a combination of direct neural estimates of posterior moments and low-dimensional posterior marginals (Jeffrey & Wandelt arXiv:2011.05991, presented at NeurIPS 2020)

Moment networks: BBH mergers

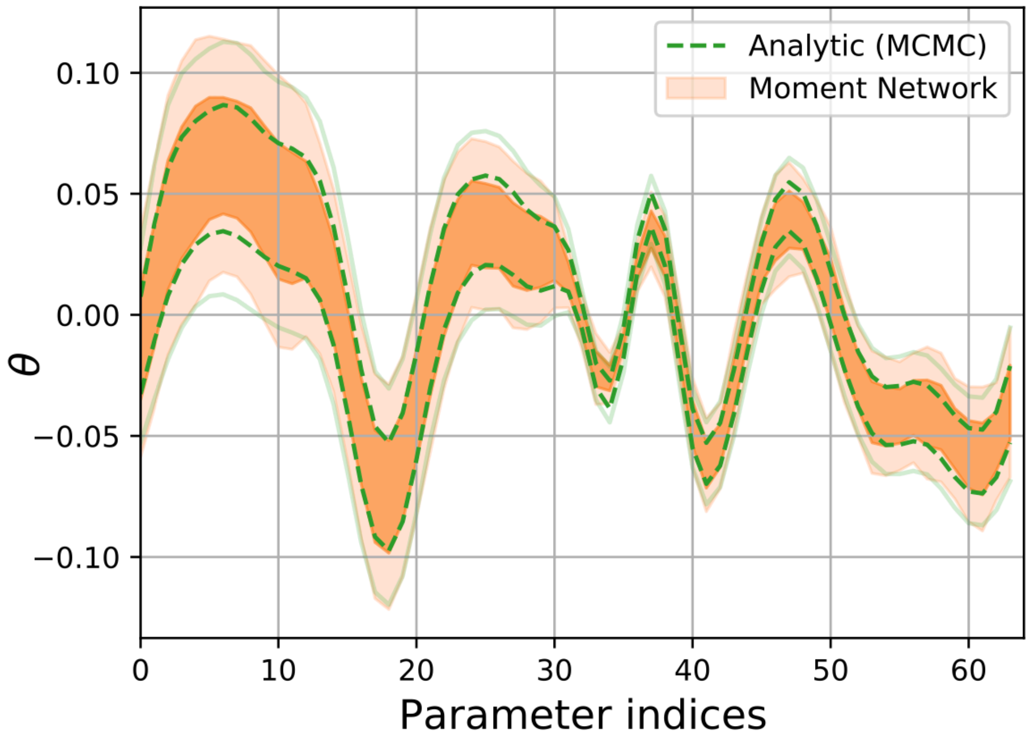


BBH merger simulations, LIGO noise

Moment networks: BBH mergers



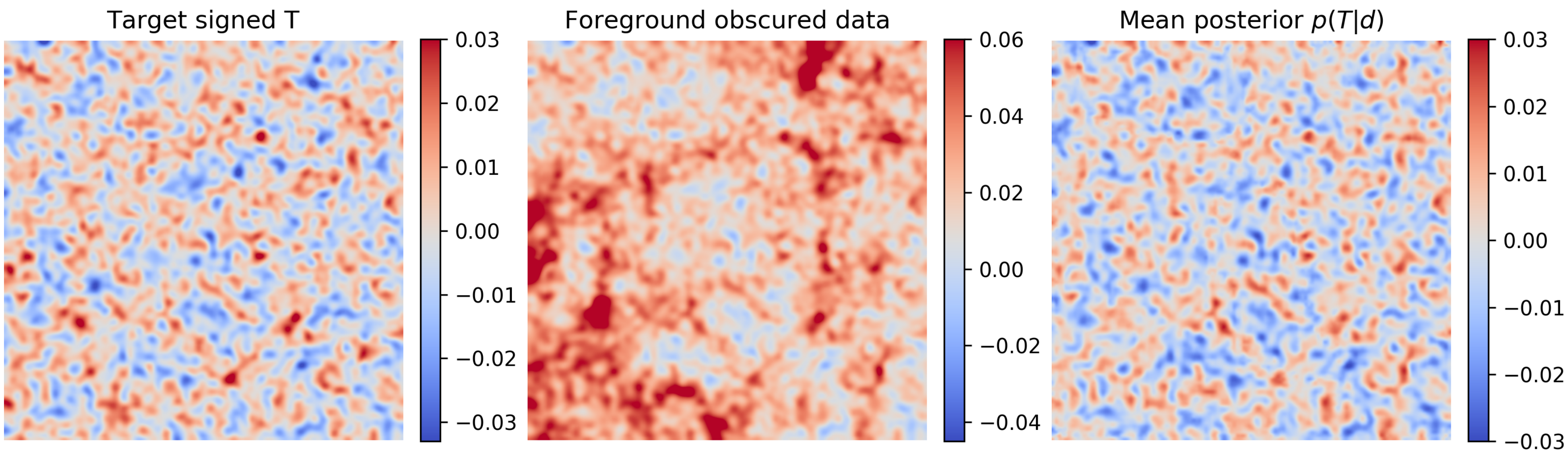
SBI reconstruction of BBH merger simulations, LIGO noise



Validation

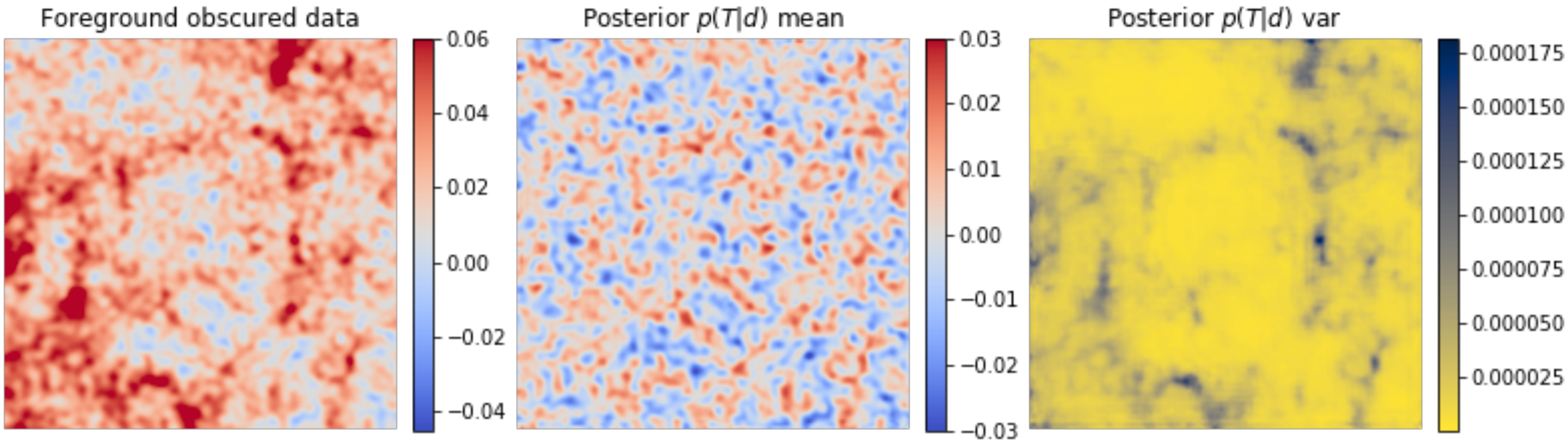
Jeffrey & Wandelt, arXiv:2011.05991

Moment networks: CMB foregrounds



Jeffrey & Wandelt, arXiv:2011.05991

Moment networks: CMB foregrounds



Jeffrey & Wandelt, arXiv:2011.05991

But what if you don't know how to compute informative summaries of your data?

Machine Learning to the rescue!

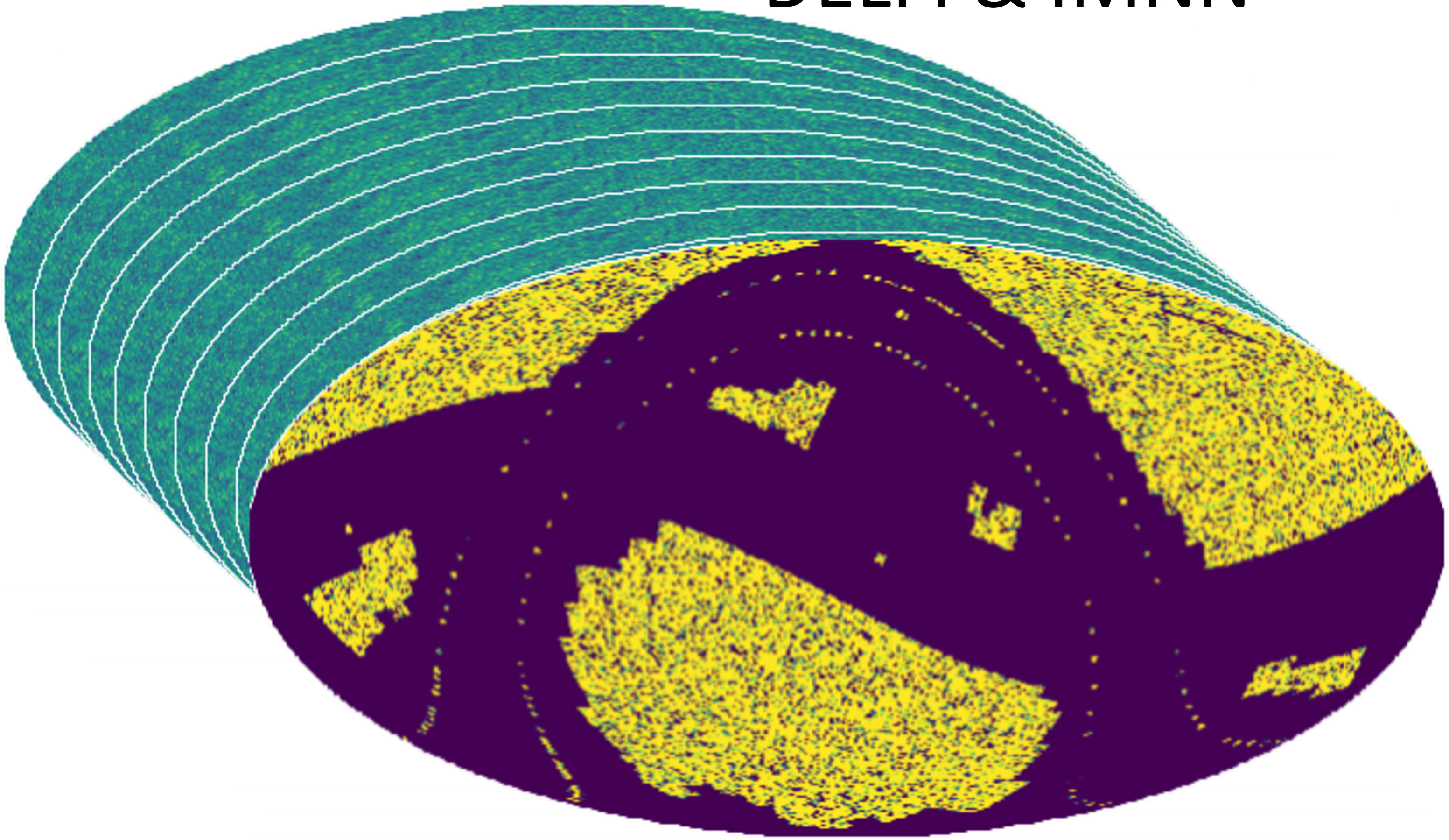
Automatic Physical Inference with Information Maximizing Neural Networks (IMNN)

Charnock, Lavaux, Wandelt (arXiv:1802:03537)

- Goal: remove the need to "guess" heuristic, informative summaries of the data
- Setup: design a neural network that maps the data into a small set of maximally informative *summaries*.

- Training uses physical simulations of the model to maximize the information in the summaries about the parameters of the model.
- The achieved loss on a test set is meaningful – it's the information content of the data.
- Can prove that the IMNN computes the optimal (score) compression without knowing the likelihood (Wandelt et al., in prep)

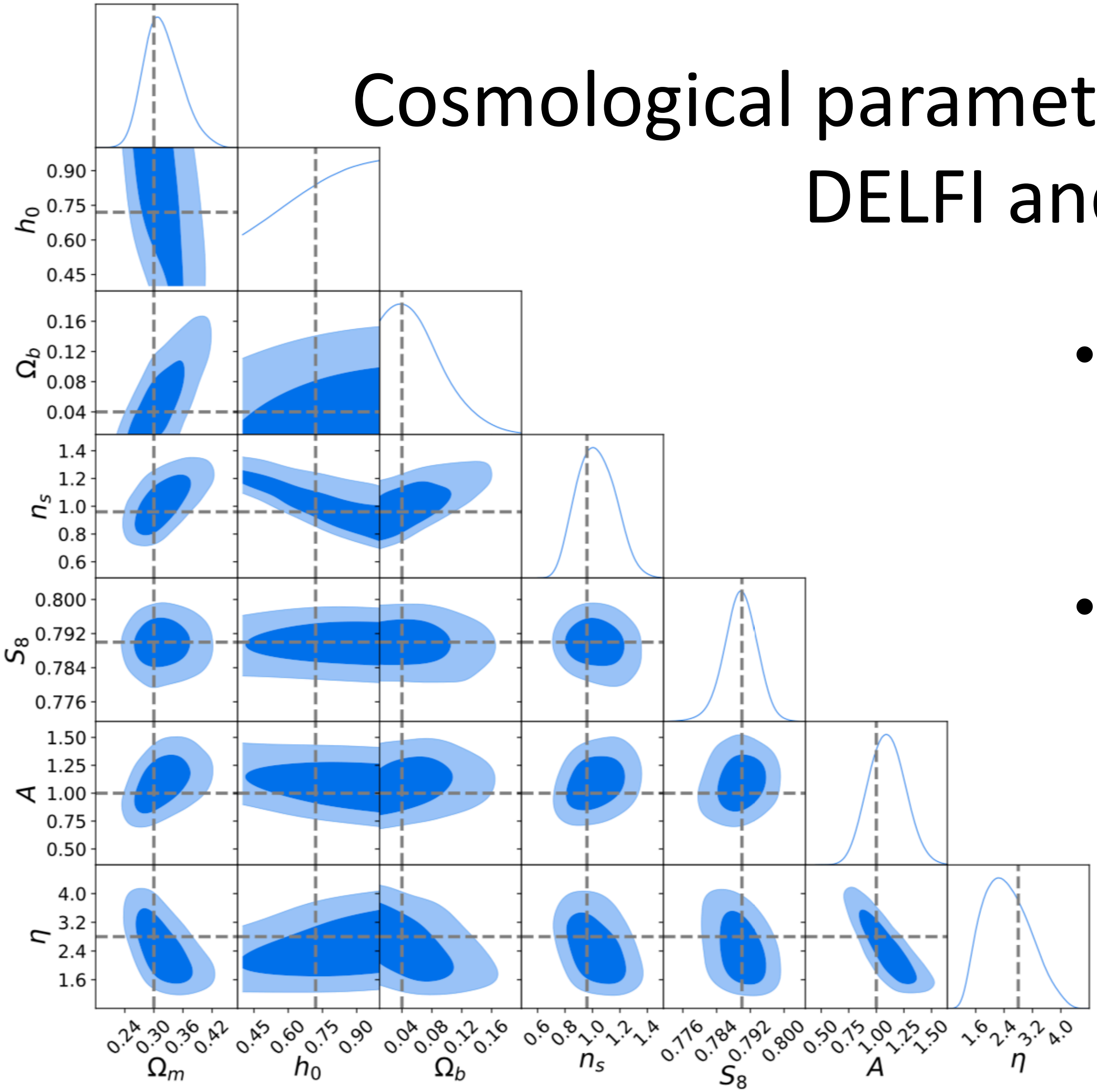
Example application: weak lensing tomography with DELFI & IMNN



Mask and noise

10 spherical shells of correlated signal simulations

Taylor, et al., arXiv: 1904.05364



Cosmological parameter inferences using DELFI and IMNN

- First Bayesian **weak lensing analysis with Non-Gaussian lensing potential**
- Enabled by DELFI and IMNN

Taylor, et al., arXiv: 1904.05364

(see also Diaz Rivero & Dvorkin arxiv:2007.05535)

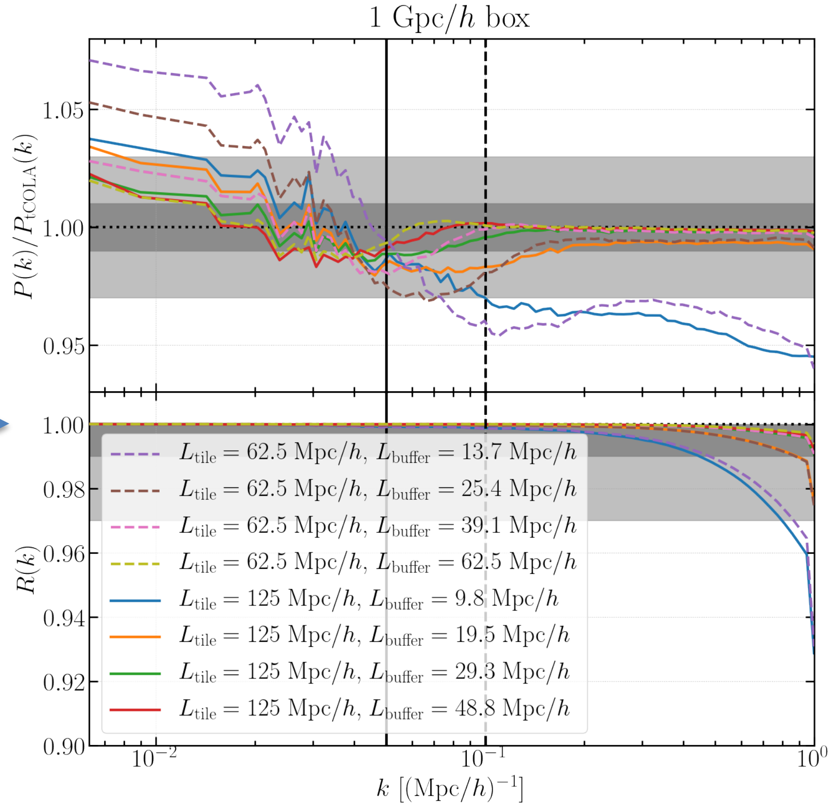
How do we get the all these simulations we need for LFI/SBI?

Using recent advances in computational physics, stats, and machine learning

Benjamin Wandelt

Leclercq et al: Simbelmynë: Perfectly Parallel n-body sims.

Opens up new ways to do larger and deeper n-body sims on a broad range of computational architectures



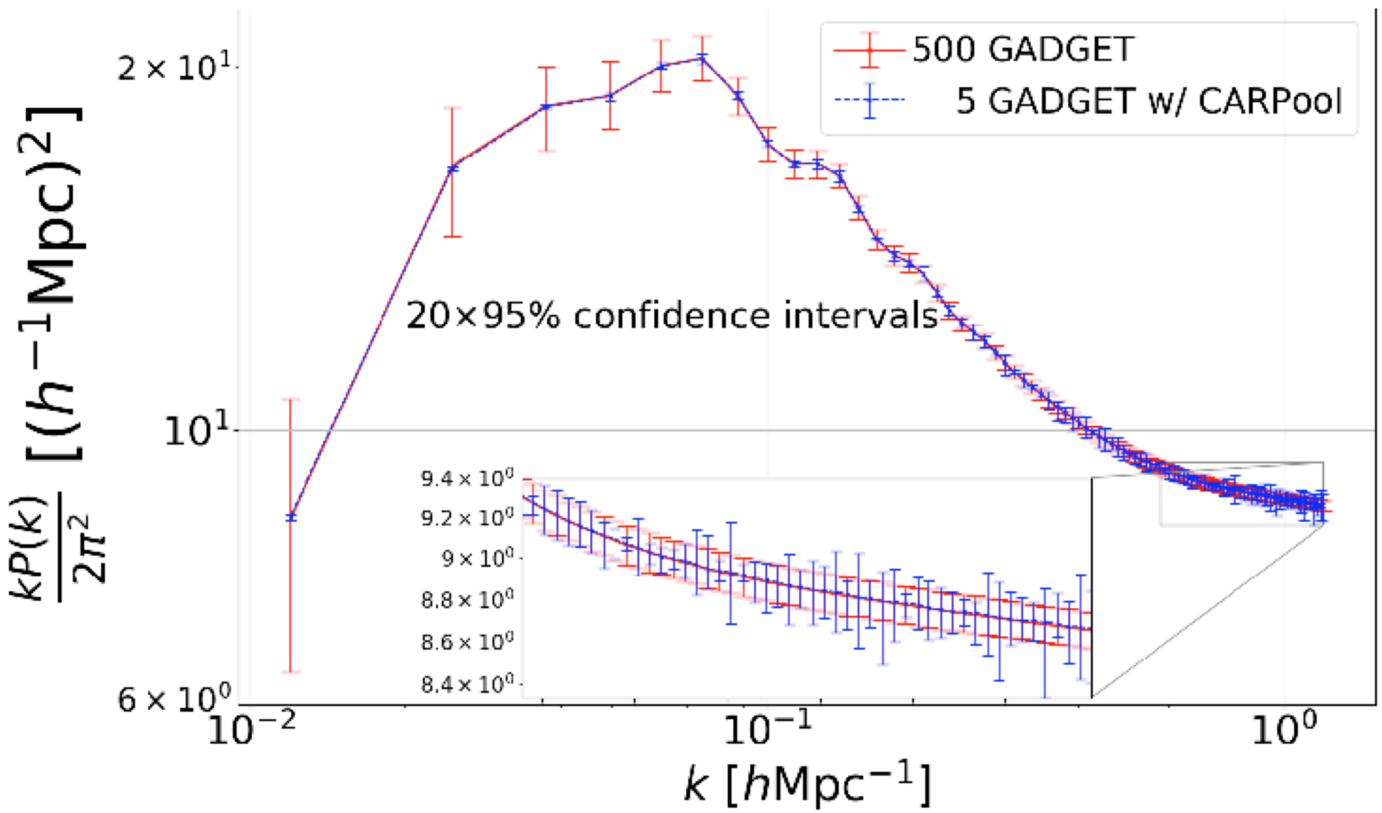
<1% error on P(k) small scales

100% correlation with serial simulation

Leclercq et al: arXiv:2003.04925

N. Chartiers et al: CARPool reduces the number of needed simulations by orders of magnitude

Convergence **A**cceleration by **R**egression and **Pool**ing) uses fast, approximate surrogates to give **unbiased, low-variance** estimates of full simulation results.



N. Chartiers et al: arXiv:2009.08970

F. Villaescusa-Navarro et al.: The QUIJOTE simulations to train machine learning surrogates

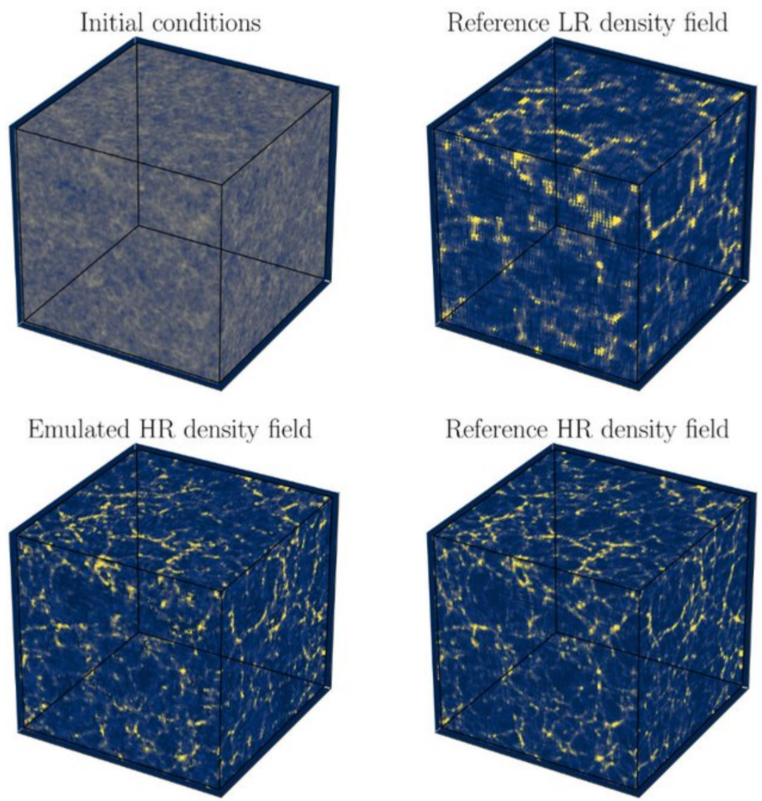
- Largest release of N-body simulation data to date
  - 43,100 full GADGET 3 simulations $(1\ \mathrm{Gpc})^3$, 512^3 or 1024^3 particles
  - ~1 PB of data
- Goal: quantify statistics information content of non-Gaussian non-linear density field about cosmological parameters
- Includes full dark matter snapshots, halo and void catalogues, and many pre-computed statistics.

Excellent tool for training machine learning surrogates.

Villaescusa-Navarro et al, arXiv:1909.05273

Kodi Ramanah et al: **neural super-resolution** of n-body simulations

Uses a Wasserstein-GAN to generate high-resolution n-body output from low-res result and high-res initial conditions with ~ 1% accuracy.



Kodi Ramanah et al, arXiv:2001.05519

CAMELS

Cosmology and Astrophysics with Machine Learning

Collaborative project to generate large suites of full, cosmological hydrosimulations as a function of cosmological parameters and astrophysics models with two different codes (AREPO/Illustris & GIZMO/SIMBA).

Use to train and validate machine learning surrogates, and likelihood-free, simulation-based inference.

F. Villaescusa-Navarro et al: arXiv:2010.00619

Summary

- Will be awash in data. Many advances in cosmology hinge on solving the cosmological inference problem. Let's solve it! We want:
  - The initial conditions of the universe to study the cosmic beginning
  - High-accuracy, high-precision inferences of expansion geometry and growth to lift the mysteries of dark matter and cosmic acceleration.

- We now have a tool set to attack this problem based on advances in physics, stats, and machine/deep learning
  - Full physical forward model inference such as BORG
  - Neural physical engine layer to model observations from dark matter
  - New, geometrical cosmological tests (field-based Alcock-Paczynski and multi-probe generalizations) that are robust to model misspecification
  - Likelihood-free, simulation-based inference

- New approaches to solve the simulation problem, e.g
  - Perfectly parallel sims with Simbelmynë
  - Unbiased variance reduction with CARPool
  - High-performance neural surrogates

Codes

BORG and related projects: aquila-consortium.org

IMNN: bitbucket.org/tomcharnock/imnn/

DELFI: github.com/justinalsing/pydelfi

The Quijote Simulations: github.com/franciscovillaescusa/Quijote-simulations

The Camels Simulations: camel-simulations.org

Simbelmynë perfectly parallel n-body code: simbelmyne.florent-leclercq.eu