Farewell talk

Florent Leclercq

http://icg.port.ac.uk/~leclercq/

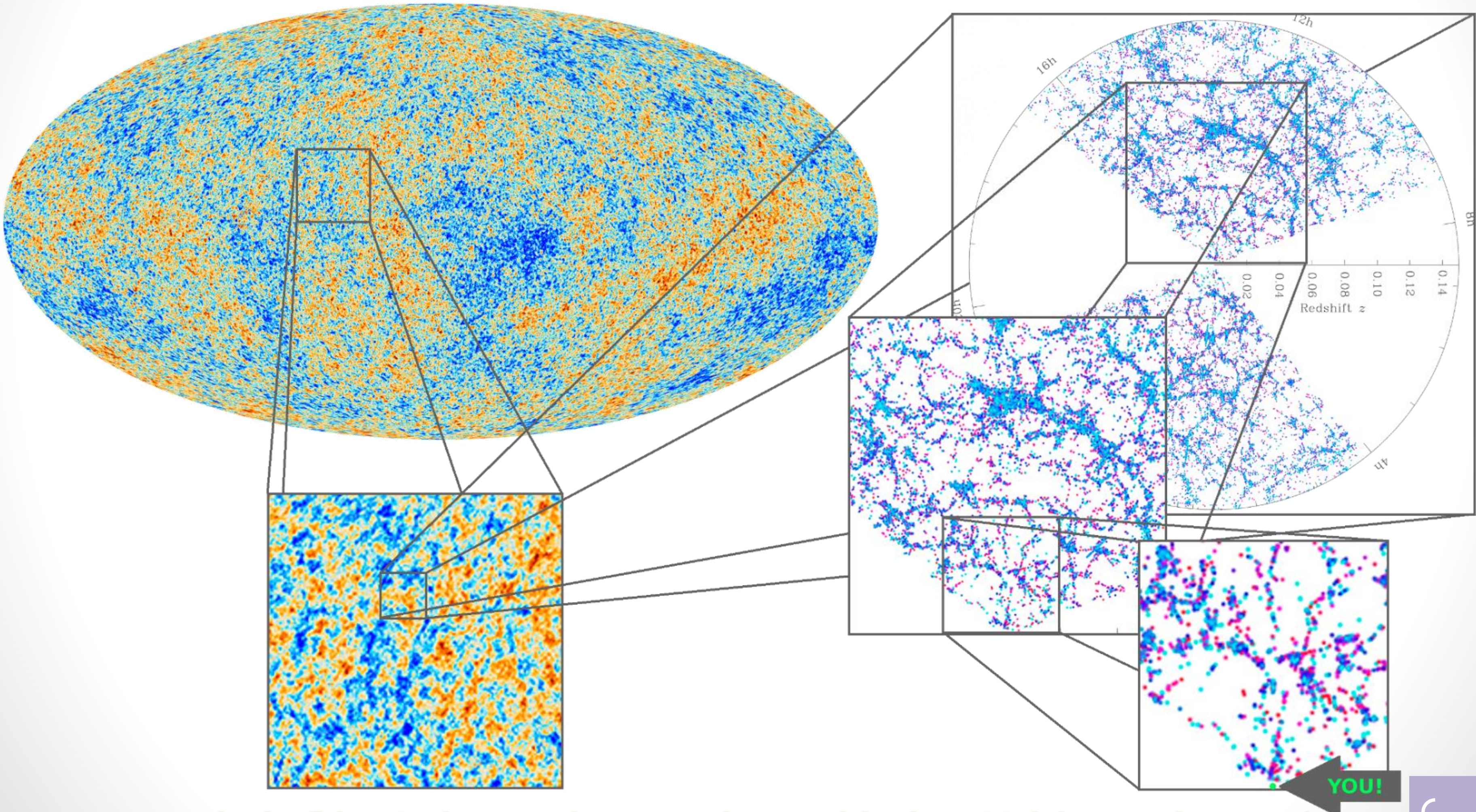
Institute of Cosmology and Gravitation, University of Portsmouth → Imperial Centre for Inference and Cosmology, Imperial College, London

May 16th, 2017

Wolfgang Enzi (MPA, Garching), Baptiste Faure (École polytechnique & ICG),
Jens Jasche (ExC Universe, Garching), Guilhem Lavaux (IAP), Will Percival (ICG),
Benjamin Wandelt (IAP), Matías Zaldarriaga (IAS, Princeton)

The big picture: the Universe is highly structured

You are here. Make the best of it…



Planck collaboration (2013-2015)

M. Blanton and the Sloan Digital Sky Survey (2010-2013)

How did structure appear in the Universe?

A joint problem!

- **How did the Universe begin?**
  - What are the statistical properties of the initial conditions?

- **How did the large-scale structure take shape?**
  - What is the physics of dark matter and dark energy?

Testing cosmological models with the LSS

J. Cham – PhD comics

Redshift range	Volume (Gpc^3)	k_{max} $(Mpc/h)^{-1}$	N_{modes}
0-1	50	0.15	10^7
1-2	140	0.5	5×10^8
2-3	160	1.3	10^{10}

M. Zaldarriaga

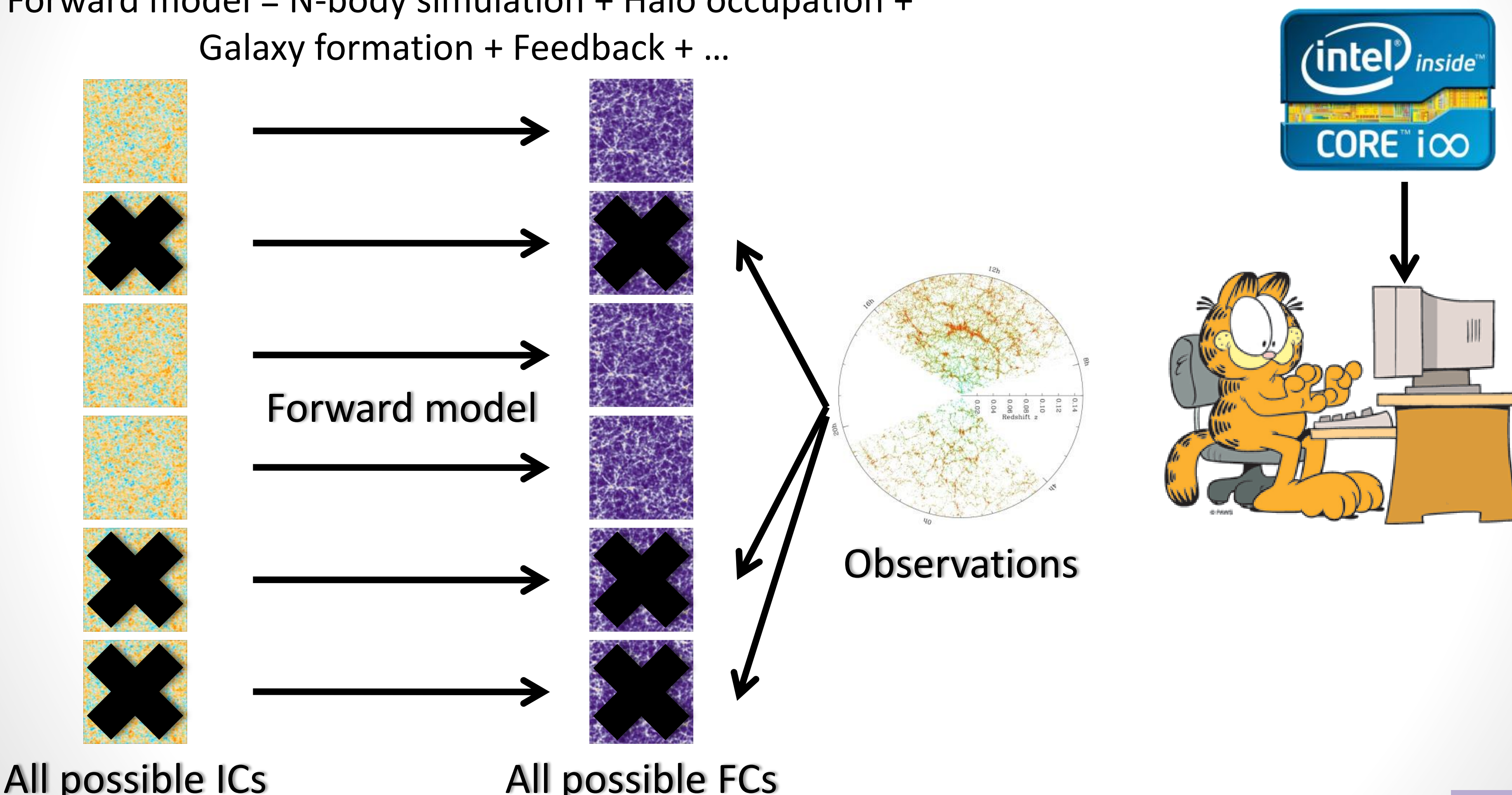
- Precise tests require many modes.
- In 3D galaxy surveys, the number of modes usable scales as $k_{\max}^3$.

- The challenge: non-linear evolution at **small scales** and **late times**.
- The strategy:
  - Inferring the initial conditions from galaxy positions
  - Pushing down the smallest scale usable for cosmological analysis

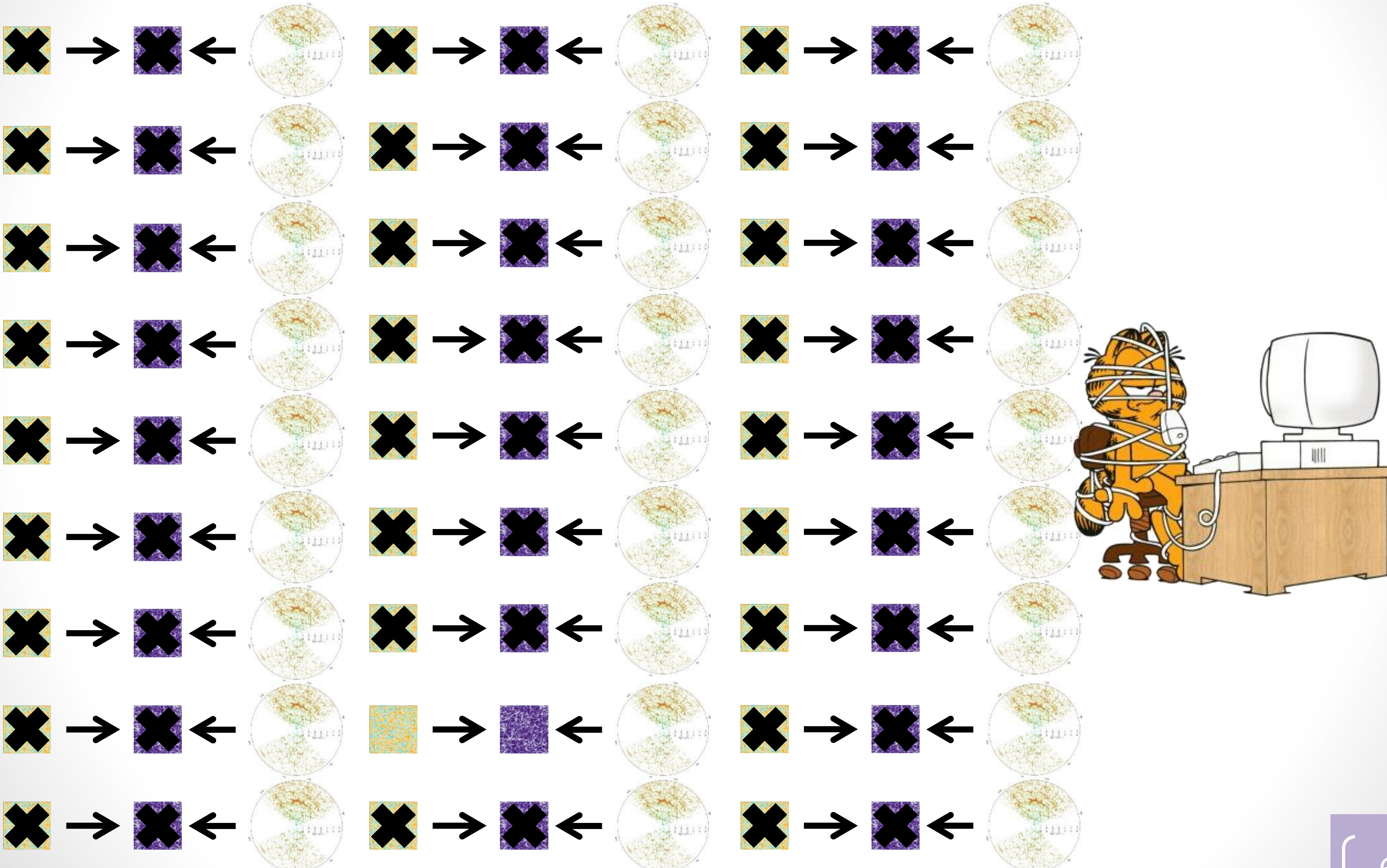
In other words: go beyond the **linear** and **static** analysis of the LSS.

Bayesian forward modeling: the ideal scenario

Forward model = N-body simulation + Halo occupation +
Galaxy formation + Feedback + …

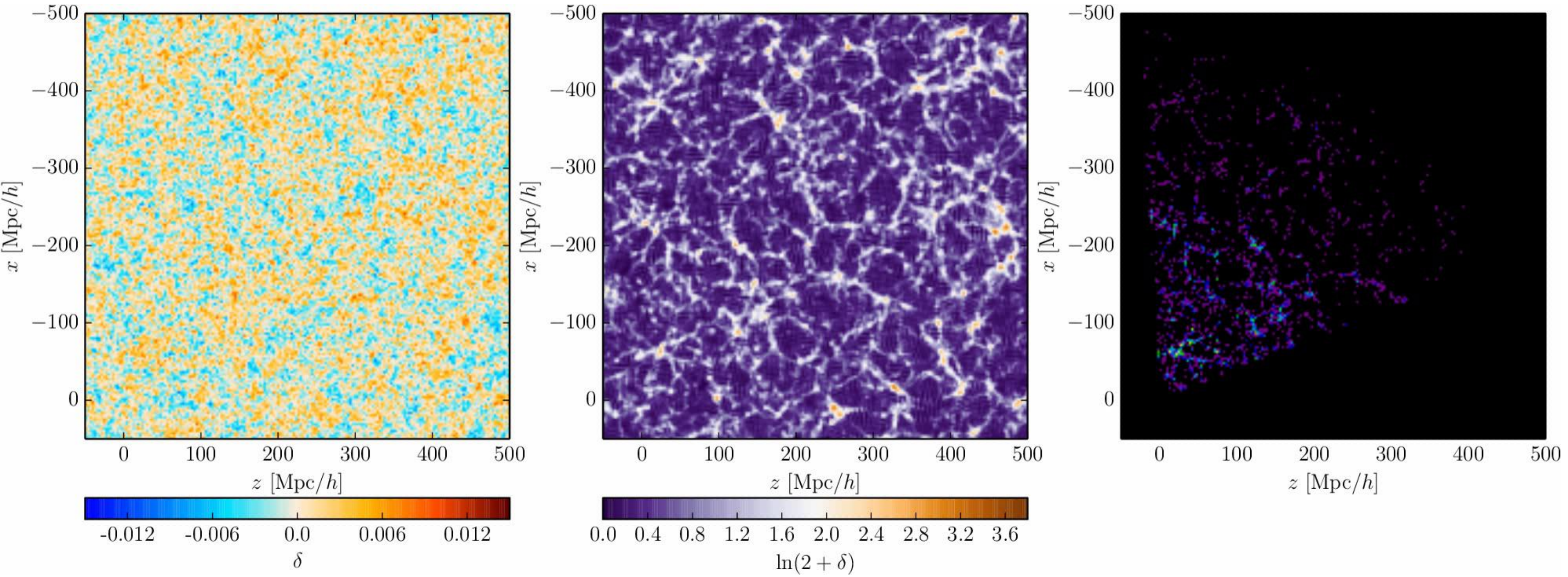


Bayesian forward modeling: the ideal scenario



Likelihood-Based Solution: BORG

Likelihood-based solution: BORG



Initial conditions

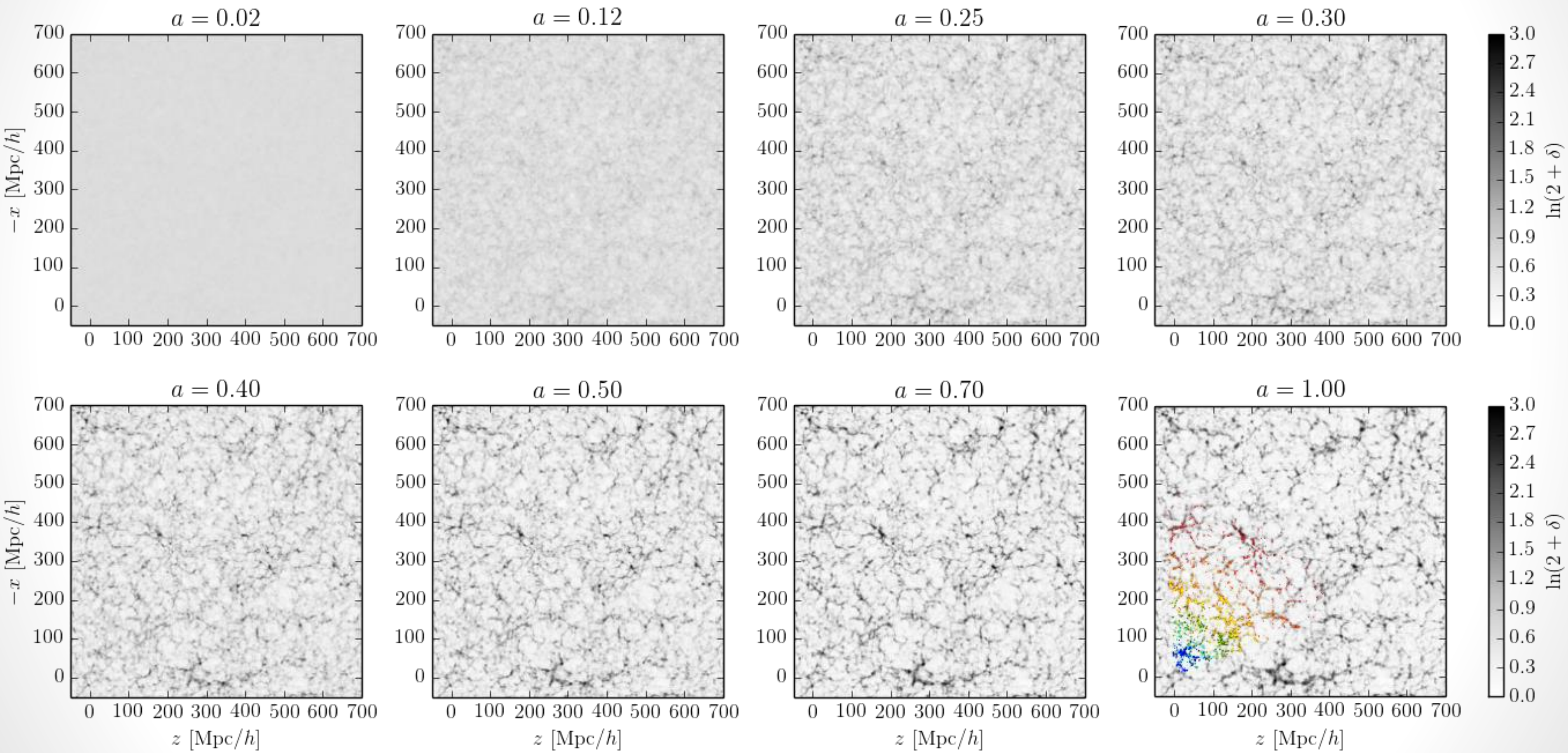
Final conditions

Observations

334,074 galaxies, ≈ 17 millions parameters, 3 TB of primary data products,
12,000 samples, ≈ 250,000 data model evaluations, 10 months on 32 cores

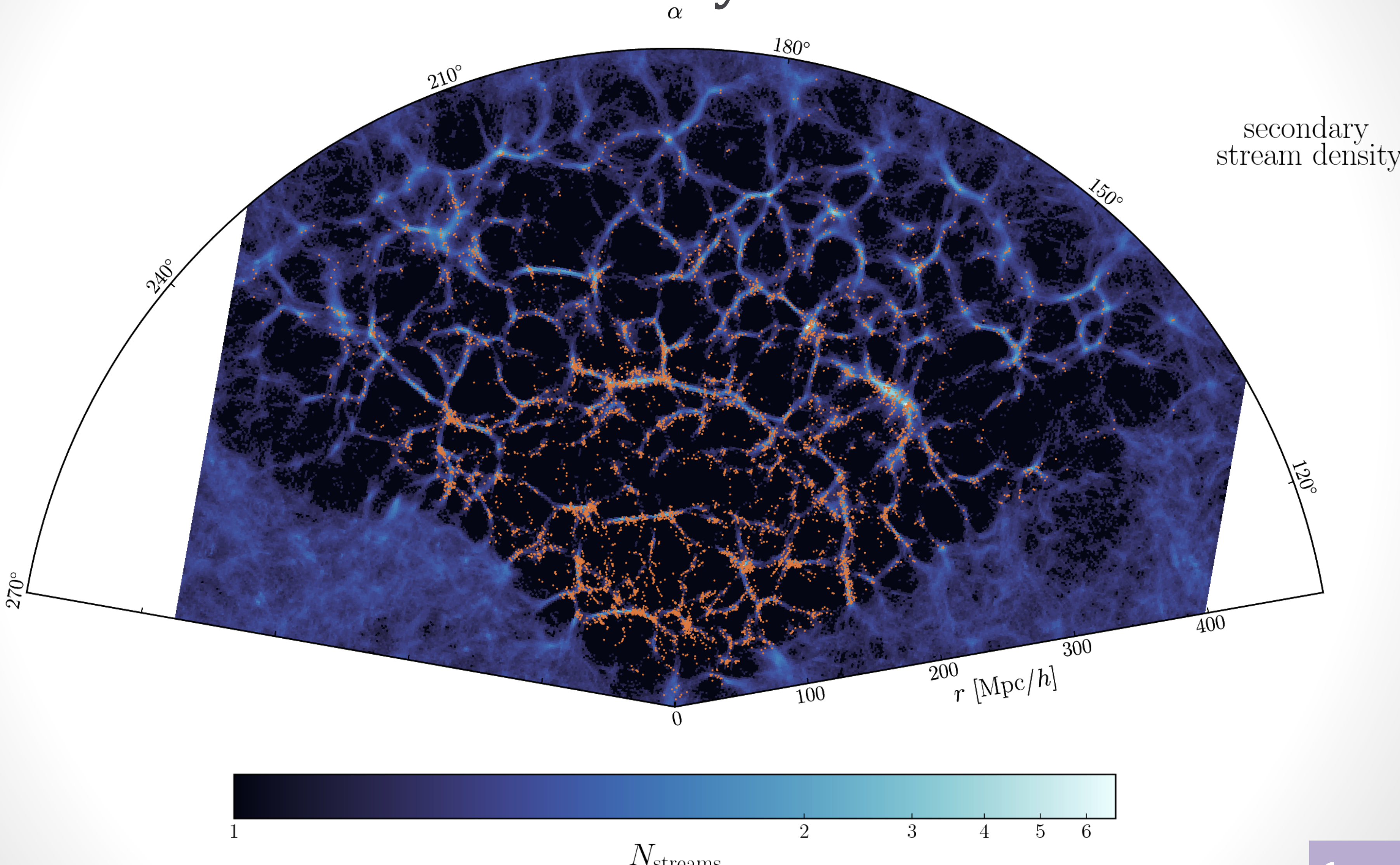
Jasche, FL & Wandelt 2015, arXiv:1409.6308

Evolution of cosmic structure



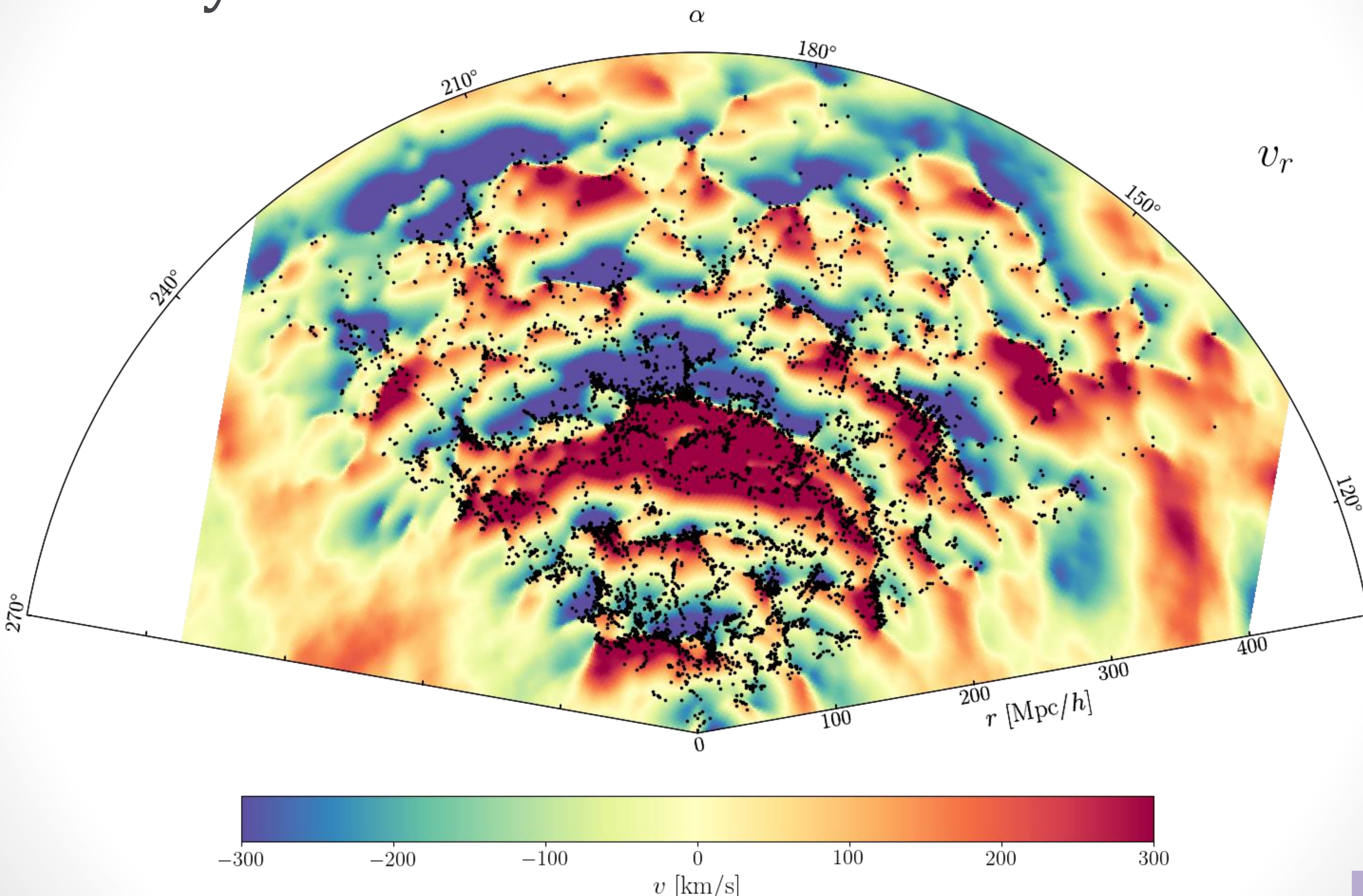
Jasche, FL & Wandelt 2015, arXiv:1409.6308

Dark matter stream density



FL, Jasche, Lavaux, Wandelt & Percival 2017, JCAP in press

Velocity field

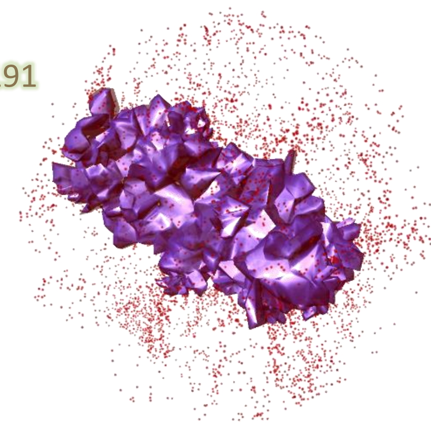


FL, Jasche, Lavaux, Wandelt & Percival 2017, JCAP in press

Cosmic web elements: some algorithms

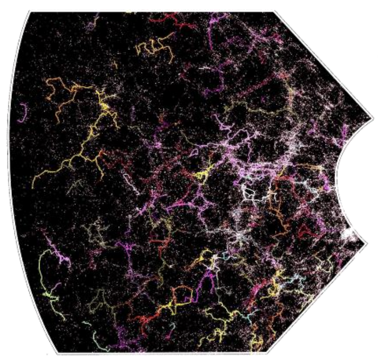
- "**Structure finders**" focus on one element at a time
  - **ZOBOV**/**VIDE**

Neyrinck 2008, arXiv:0712.3049
Sutter *et al*. 2015, arXiv:1406.1191



  - **DisPerSE**

Sousbie 2011, arXiv:1009.4015
Sousbie *et al*. 2011, arXiv:1009.4014



- "**Classifiers**" dissect the cosmic web all at once
  - The **T-web** (tidal field tensor)

Hahn *et al*. 2007, arXiv:astro-ph/0610280

  - **DIVA** (Lagrangian displacement field, potential structures)

Lavaux & Wandelt 2010, arXiv:0906.4101

  - **ORIGAMI** (particle crossings)

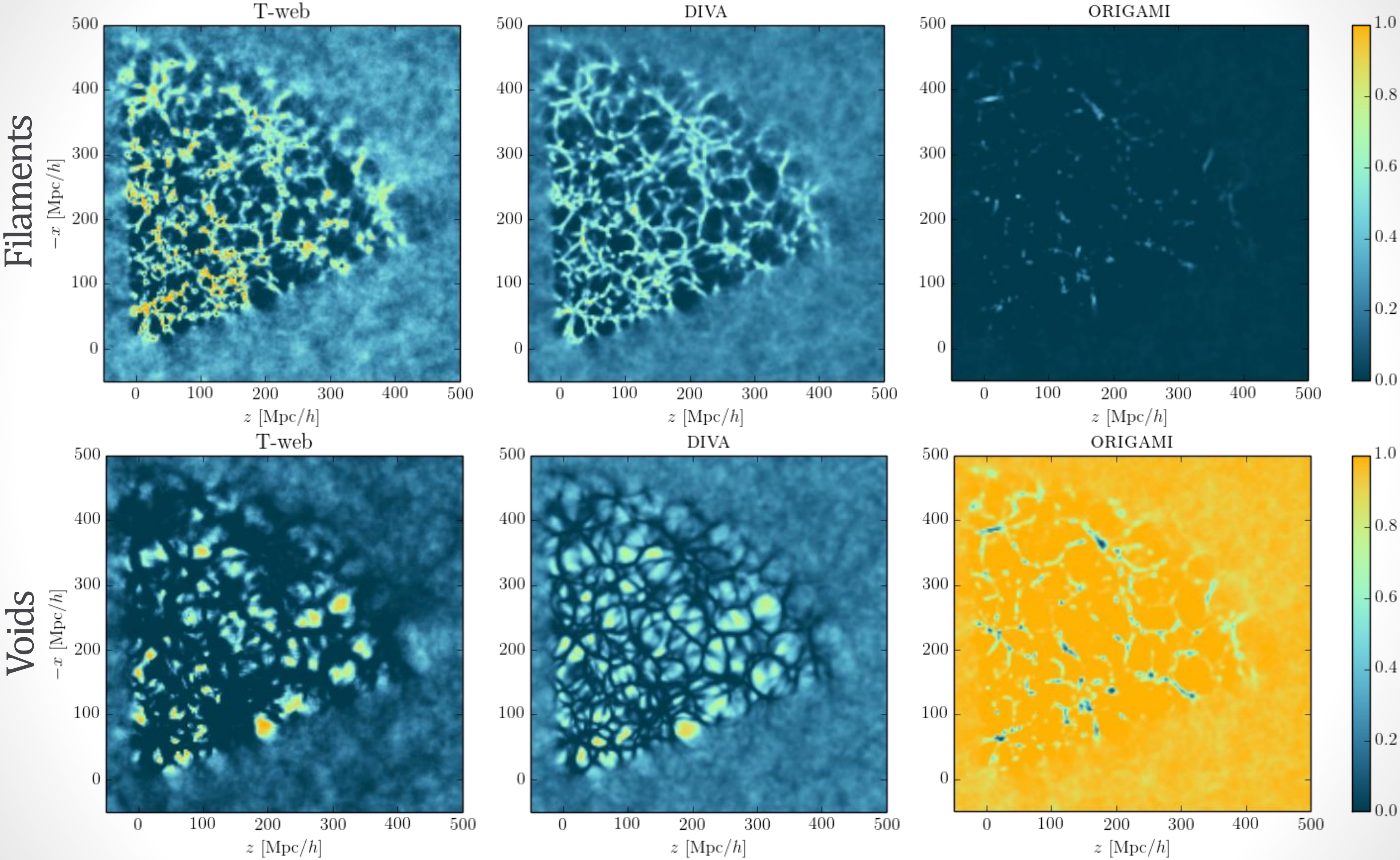
Falck, Neyrinck & Szalay 2012, arXiv:1201.2353

  - **LICH** (Lagrangian displacement field, potential and vortical structures)

FL, Jasche, Lavaux, Wandelt & Percival 2017

and many others…

Comparing classifiers

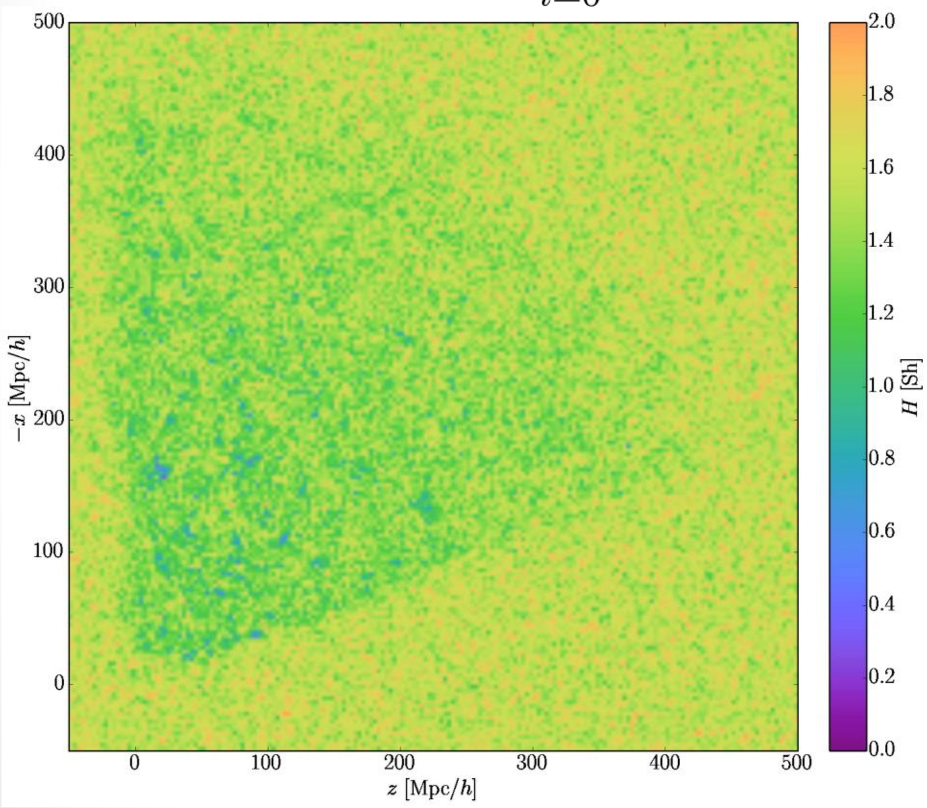


FL, Jasche & Wandelt 2015a, arXiv:1502.02690
FL, Lavaux, Jasche & Wandelt 2016, arXiv:1606.06758

How is information propagated?

Shannon entropy

$$H\left[\mathcal{P}(\mathrm{T}(\vec{x}_k)|d)\right] \equiv - \sum_{i=0}^{3} \mathcal{P}(\mathrm{T}_i(\vec{x}_k)|d)\log_2(\mathcal{P}(\mathrm{T}_i(\vec{x}_k)|d))$$ in shannons (Sh)



More about cosmic web analysis:

FL, Jasche & Wandelt 2015a, arXiv:1502.02690
(T-web, entropy, relative entropy)
FL, Jasche & Wandelt 2015b, arXiv:1503.00730
(decision theory for structure classification)
FL, Lavaux, Jasche & Wandelt 2016, arXiv:1606.06758
(mutual information, classifier utilities)
FL, Jasche, Lavaux, Wandelt & Percival 2017
(phase-space structure of dark matter)

FL, Jasche & Wandelt 2015a, arXiv:1502.02690

Likelihood-Free Solution

Why is likelihood-free rejection so expensive?

1. It rejects most samples when ϵ is small

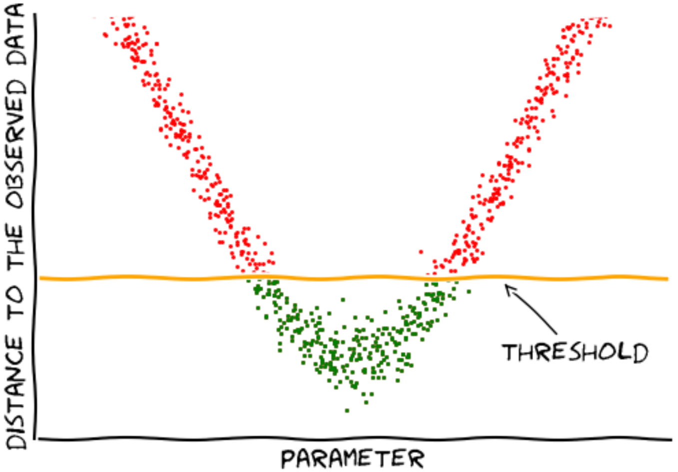
2. It does not make assumptions about the shape of $L(\theta)$

3. It uses only a fixed proposal distribution, not all information available

4. It aims at equal accuracy for all regions in parameter space

Effective likelihood approximation:

$$L(\theta) \approx \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}\left(\mathrm{d}(\tilde{d}(\theta), d) \leq \epsilon \right)$$



Proposed solution

Bayesian optimisation for likelihood-free inference (BOLFI)

1. It rejects most samples when ϵ is small

   ⇨ **Don't reject samples: learn from them!**

2. It does not make assumptions about the shape of $L(\theta)$

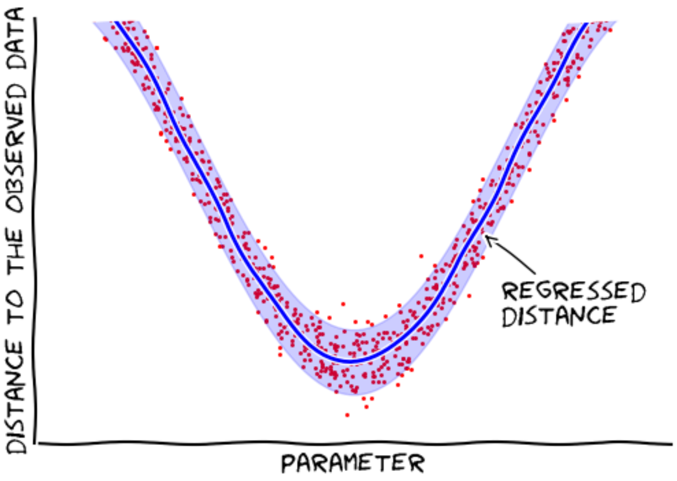
   ⇨ **Model the distances, assuming the average distance is smooth**

3. It uses only a fixed proposal distribution, not all information available

   ⇨ **Use Bayes' theorem to update the proposal of new points**

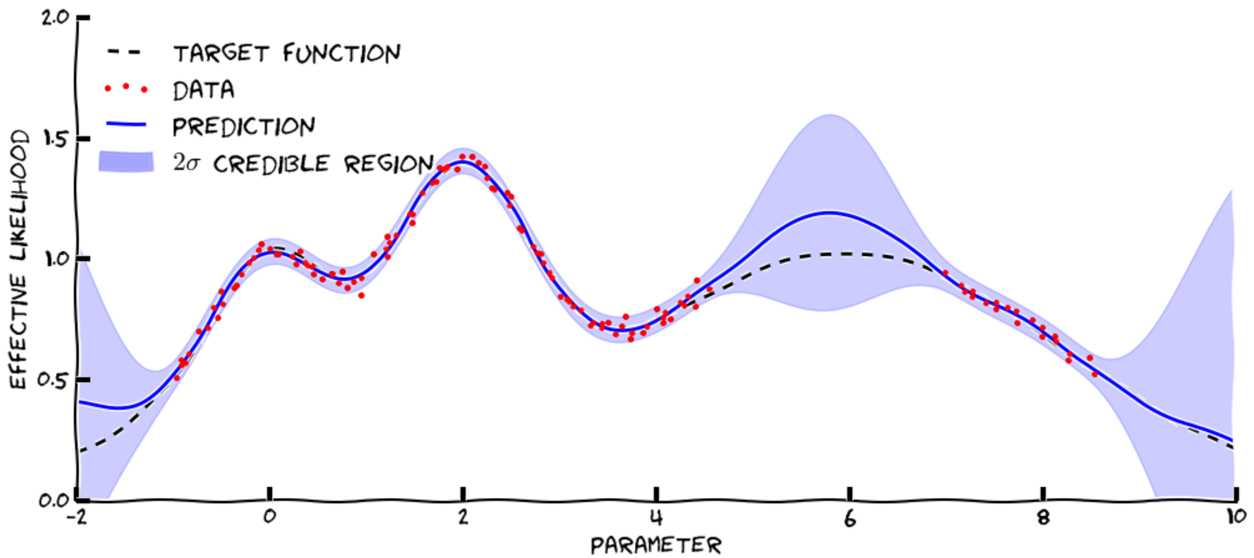
4. It aims at equal accuracy for all regions in parameter space

   ⇨ **Prioritize parameter regions with small distances to the observed data**



Gutmann & Corander JMLR 2016, arXiv:1501.03291

Regressing the effective likelihood (points 1 & 2)



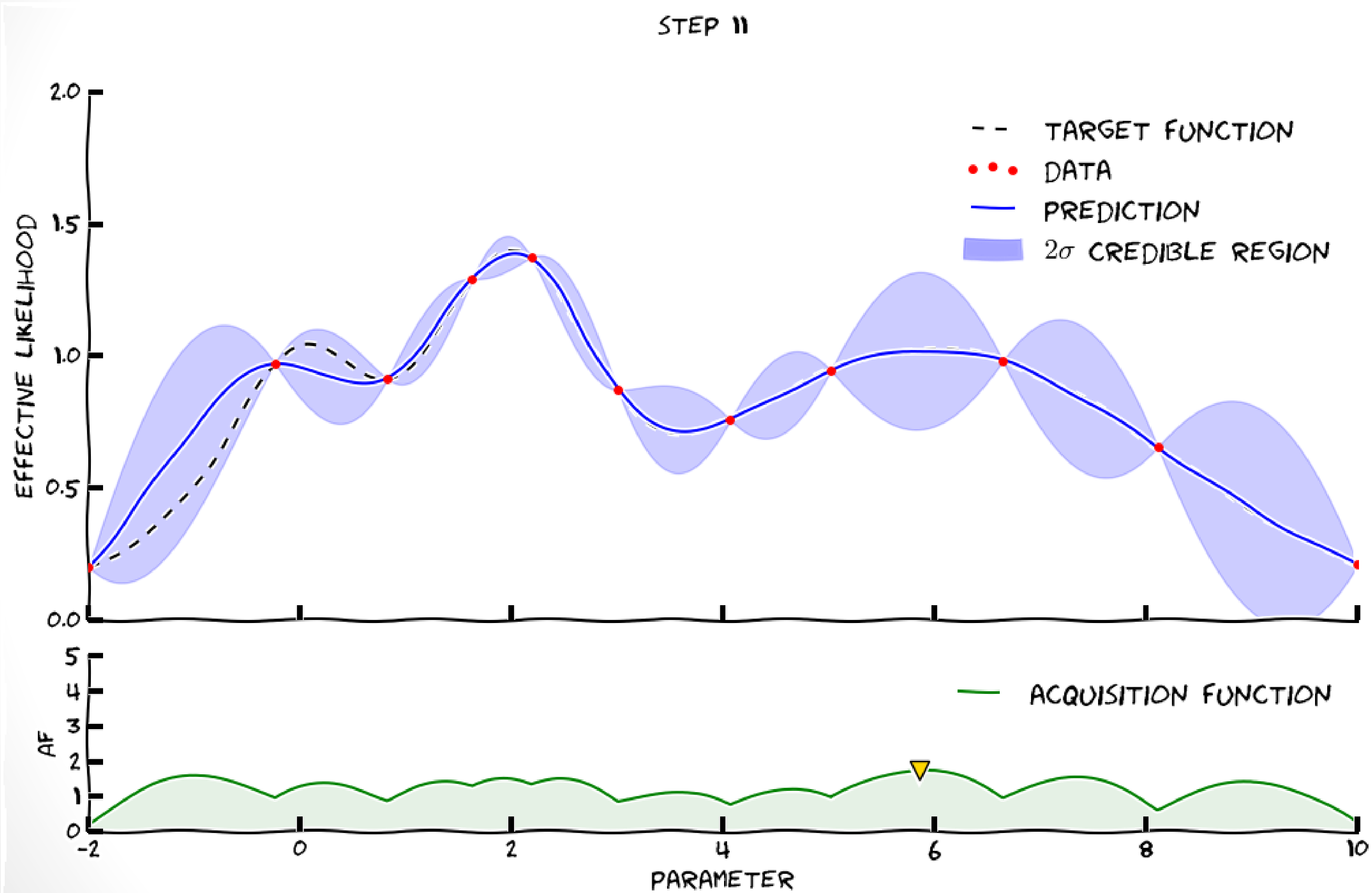
1. "It rejects most samples when ϵ is small"

- Keep all values (θ_i, d_i) $\mathrm{d}_i = \mathrm{d}(\tilde{d}(\theta_i), d)$

2. "It does not make assumptions about the shape of $L(\theta)$"

- Model the conditional distribution of distances given this training set

Data acquisition (points 3 & 4)



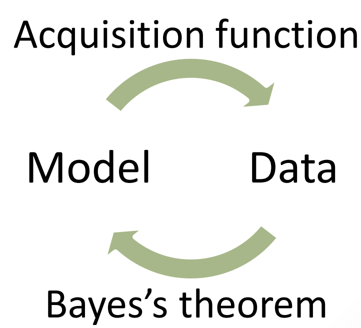
Data acquisition (points 3 & 4)

3. “It uses only a fixed proposal distribution, not all information available”

- Samples are obtained from sampling an **adaptively-constructed proposal distribution**, using the regressed effective likelihood

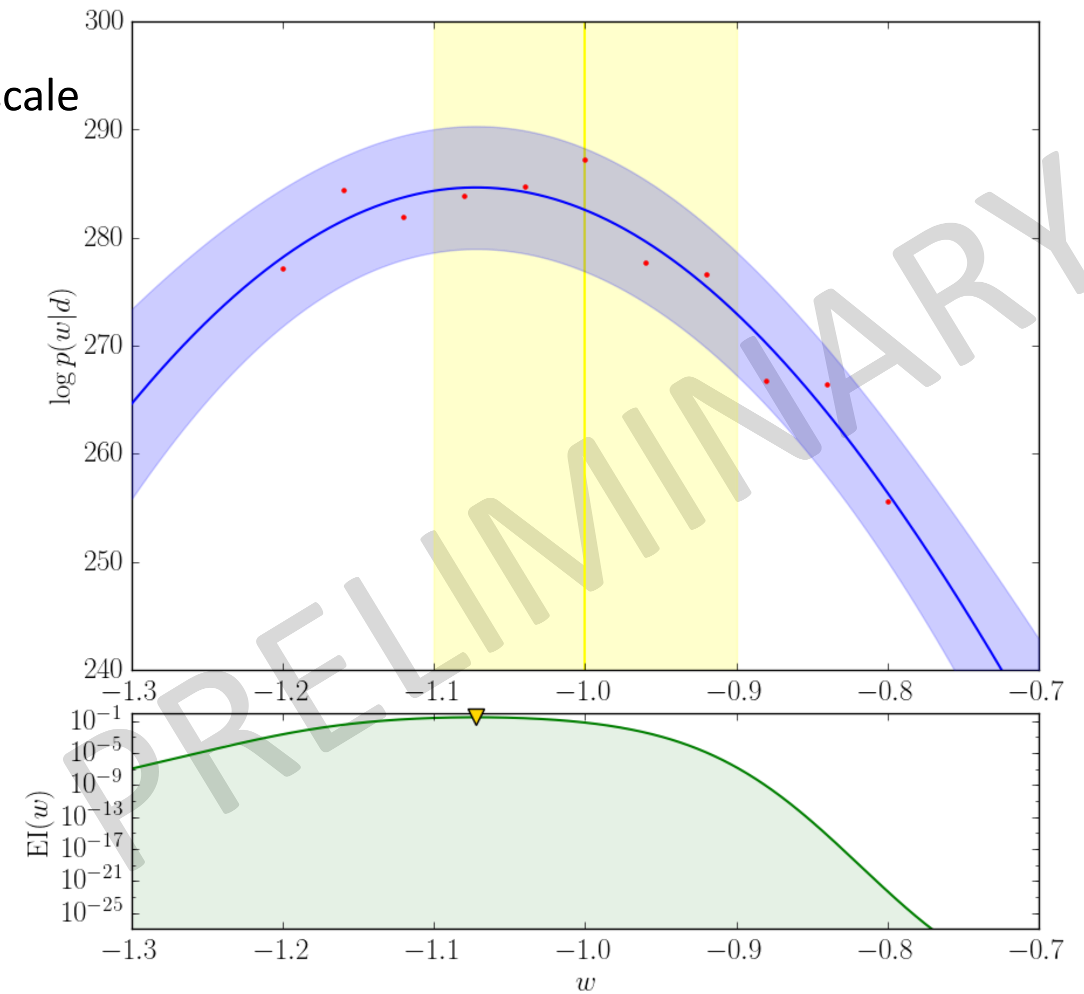
4. “It aims at equal accuracy for all regions in parameter space”

- The **acquisition function** finds a compromise between exploration (trying to find new high-likelihood regions) & exploitation (giving priority to regions where the distance to the observed data is already known to be small)
- **Bayesian optimisation** (decision making under uncertainty) can then be used



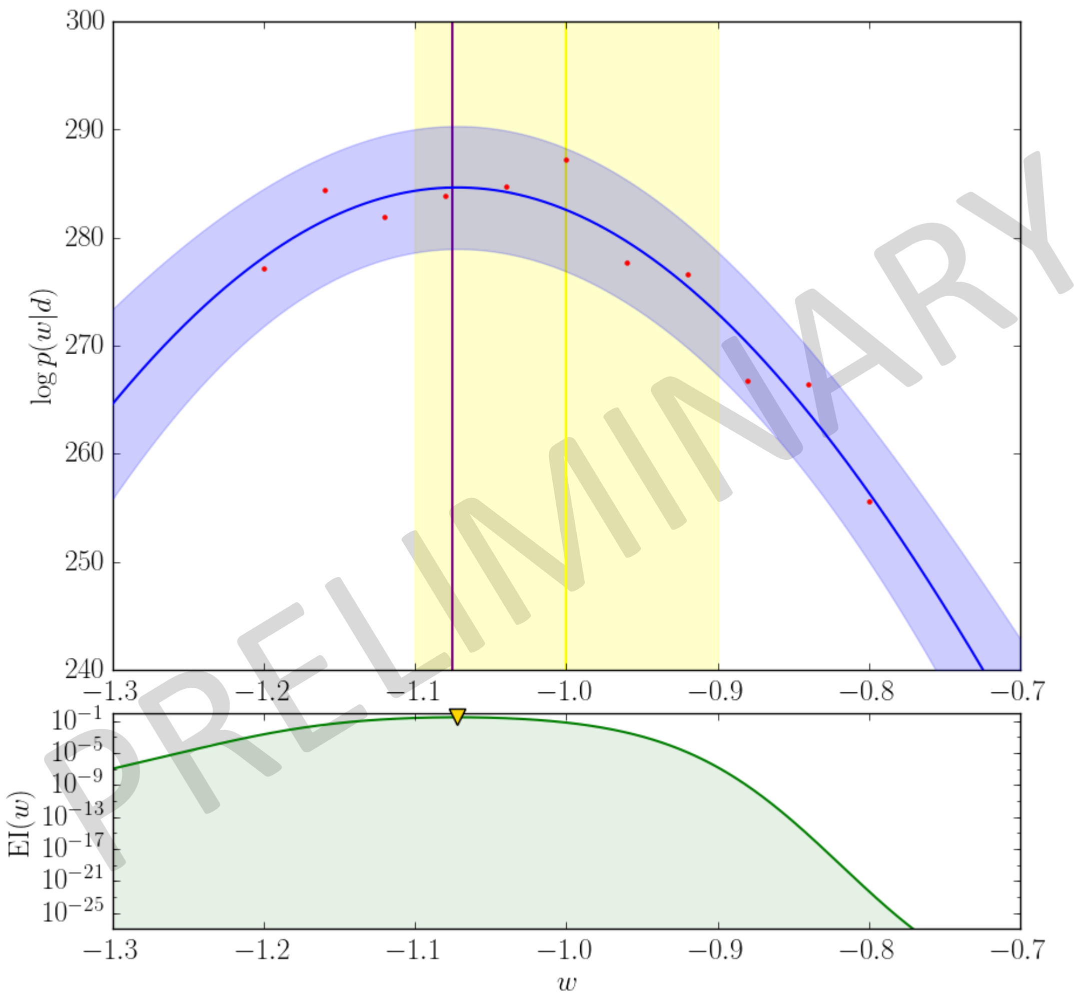
Likelihood-free large-scale structure inference

- 1100 large-scale structure simulations
- $\approx 10^7$ hidden variables



with W. Enzi & J. Jasche

Likelihood-free large-scale structure inference



This proof-of-concept has been performed completely blindly.

with W. Enzi & J. Jasche

Optimising the Data Model with sCOLA

tCOLA: *COmoving Lagrangian Acceleration (temporal domain)*

- Write the displacement vector as: $\mathbf{s} = \mathbf{s}_{\mathrm{LPT}} + \mathbf{s}_{\mathrm{MC}}$

  Tassev & Zaldarriaga 2012, arXiv:1203.5785

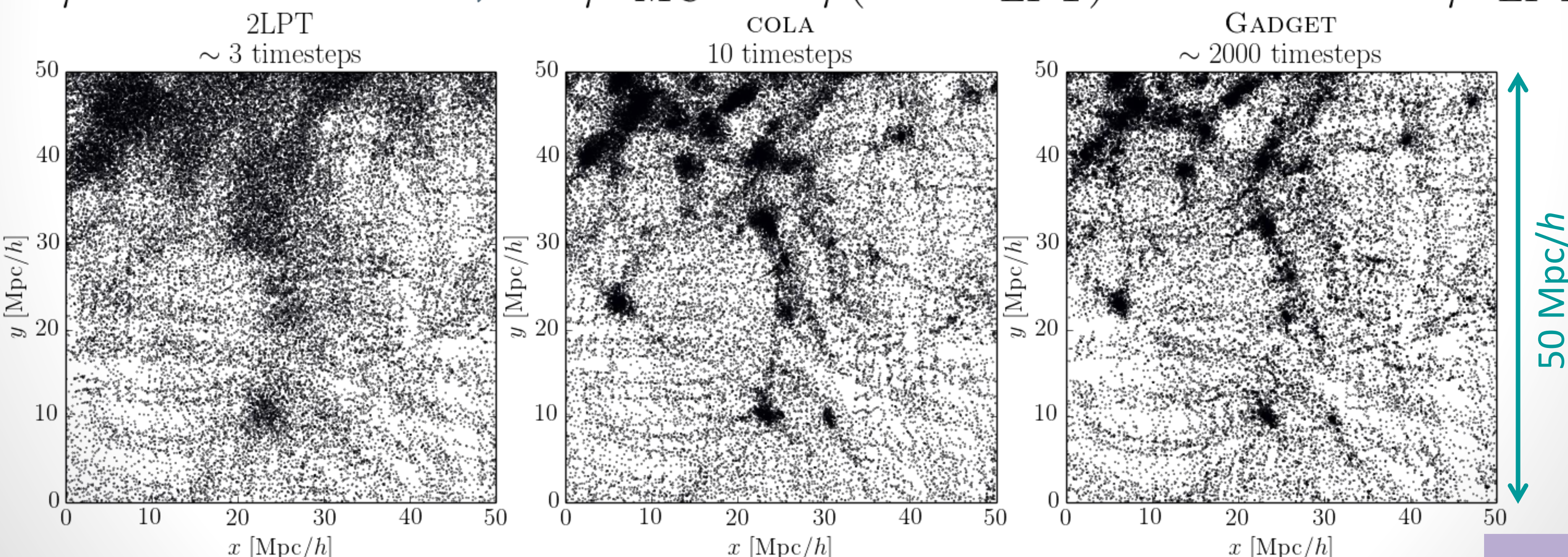
- Time-stepping (omitted constants and Hubble expansion):

Standard:

$$\partial_\tau^2 \mathbf{s} = -\nabla\Phi$$

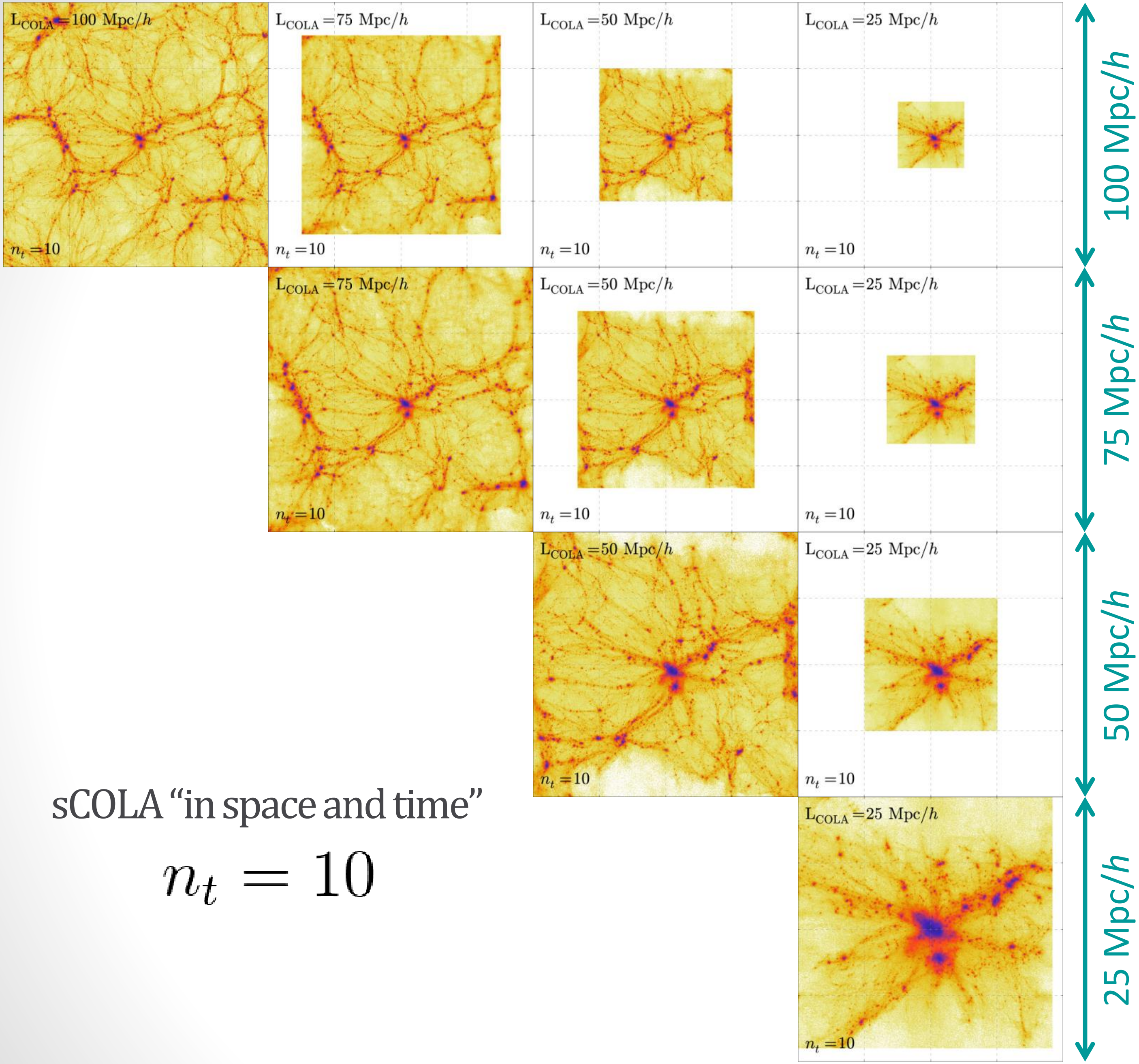
Modified:

$$\partial_\tau^2 \mathbf{s}_{\mathrm{MC}} = \partial_\tau^2(\mathbf{s} - \mathbf{s}_{\mathrm{LPT}}) = -\nabla\Phi - \partial_\tau^2 \mathbf{s}_{\mathrm{LPT}}$$



Tassev, Zaldarriaga & Einsenstein 2013, arXiv:1301.0322

sCOLA:
Extension to the spatial domain

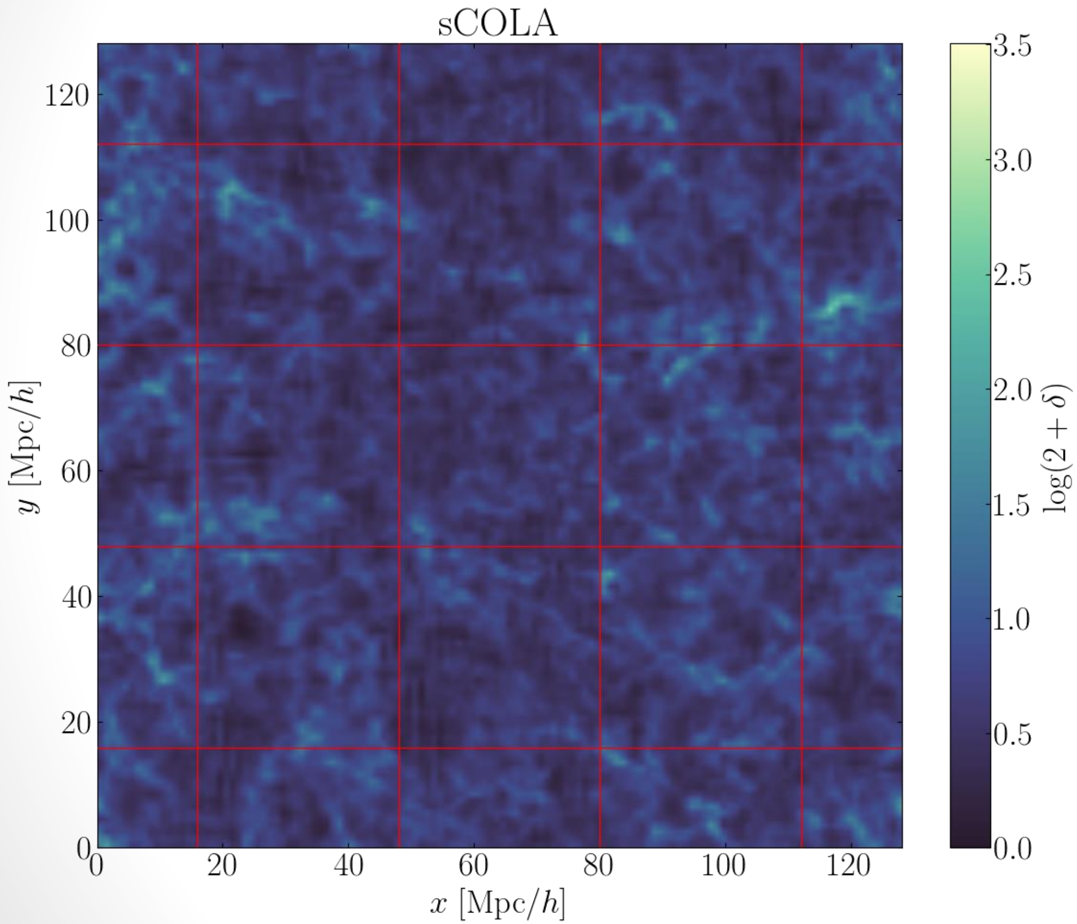


sCOLA "in space and time"

$$n_t = 10$$

Tassev, Eisenstein, Wandelt & Zaldarriaga 2015, arXiv:1502.07751

Using sCOLA to parallelize *N*-body sims



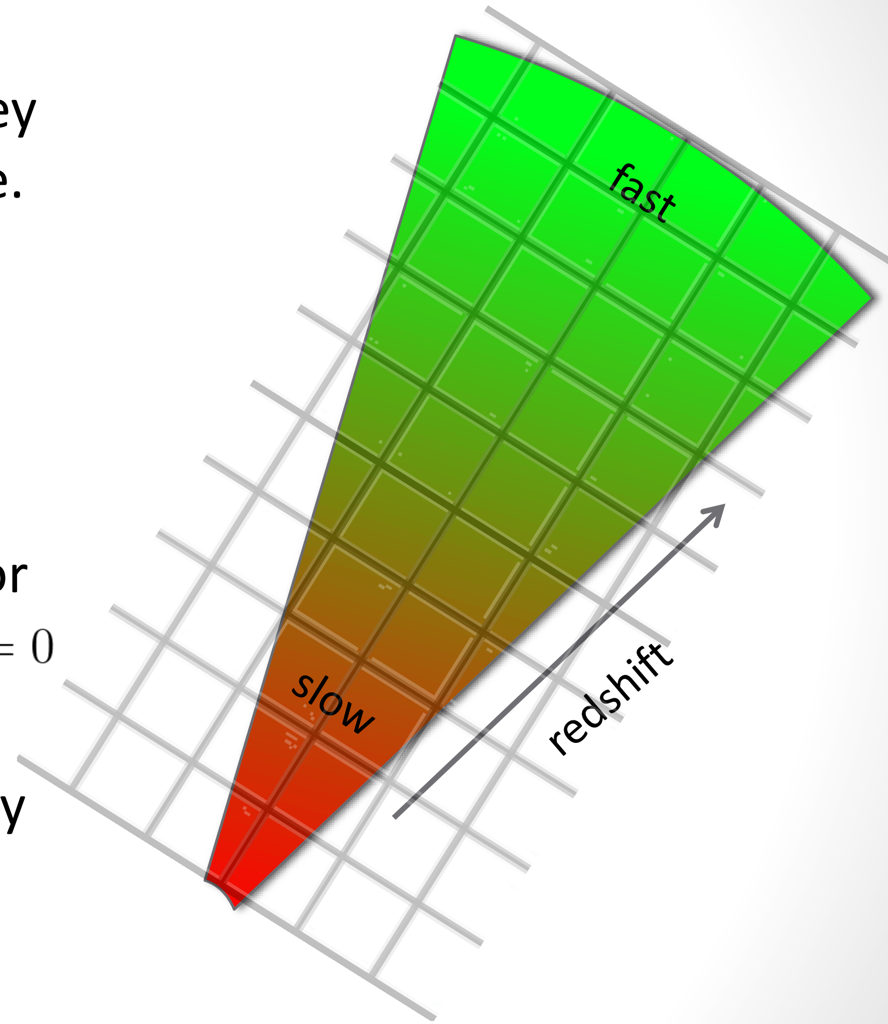
Parallelisation potential:

- Subvolumes…
  - do not need to communicate,
  - can even be run out of order!
- Factor ~ 8 overhead due to boundary regions.
- But $\sim 50\ \mathrm{Mpc}/h$ *N*-body sims can be done **in cache or on a GPU**.
  ⇨ speed-up of s
- Potential parallelisation speed-up:
  $$\frac{1}{8} \times s \times \left(\frac{10\ \mathrm{Gpc}/h}{50\ \mathrm{Mpc}/h}\right)^3 = s \times 10^6$$

with B. Faure (master project), B. Wandelt, W. Percival & M. Zaldarriaga

Constructing lightcones

- Subvolumes only need to run until they intersect the observer's past lightcone.
- Most of the high-z volume will be faster than $z = 0$.
- Many unobserved subvolumes do not even have to run!
- The wall-clock time limit is the time for running a single $\sim 50\ \mathrm{Mpc}/h$ box to $z = 0$ at the observer position.
- Leads to **further speed-up**, especially for deep surveys.



with B. Faure (master project), B. Wandelt, W. Percival & M. Zaldarriaga

Summary

- A likelihood-based method for principled analysis of galaxy surveys:
  Bayesian large-scale structure inference (BORG)
  - Simultaneous analysis of the morphology and formation history of the large-scale structure.
  - Characterization of the dynamic cosmic web underlying galaxies.
- A likelihood-free method for models where the likelihood is intractable but simulating is possible:
  Regression of the distance + Bayesian optimisation
  - Number of required simulations reduced by several orders of magnitude.
  - The approach will allow to **ask targeted questions to cosmological data**, including all relevant physical and observational effects.
- Optimisation of the data model using **tCOLA + sCOLA**
  - Enormous parallelisation potential for dark matter simulations.
  - Further speed-up expected for realistic synthetic observations.