

(My) Methods for Cosmic Web Analysis

Bayesian inference, Decision Theory, Information Theory

Florent Leclercq

www.florent-leclercq.eu

Imperial Centre for Inference and Cosmology
Imperial College London

In collaboration with

Guilhem Lavaux (IAP), Jens Jasche (Stockholm),
Benjamin Wandelt (IAP/CCA), Will Percival (Portsmouth/Waterloo)

and the Aquila Consortium
www.aquila-consortium.org

October 29th, 2018

The BORG inference framework

Bayesian Origin Reconstruction from Galaxies

- A Bayesian Hierarchical Model:

$$\mathcal{P}(\hat{\delta}) \propto \exp\left(-\frac{1}{2} \sum_k \frac{|\hat{\delta}_k|^2}{P_k}\right) \quad \text{initial conditions}$$

$$\rho_m = \mathcal{F}(\delta) \quad \text{total evolved matter density}$$

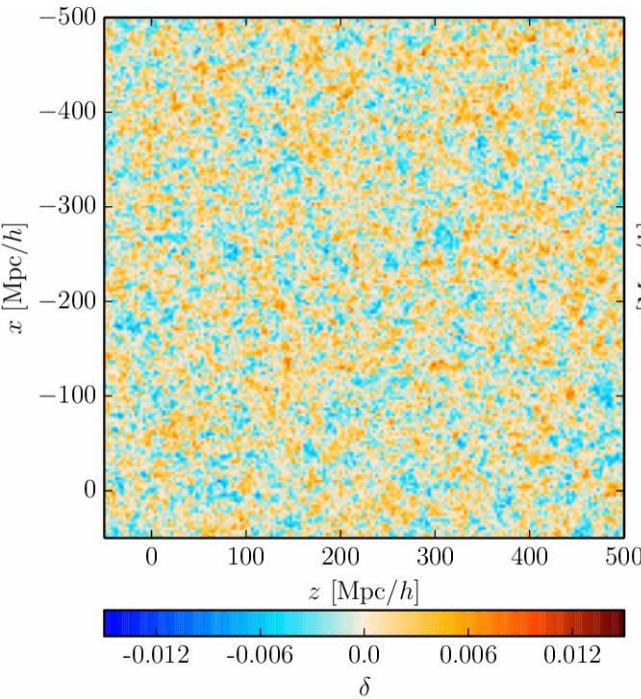
$$\rho_g \propto \rho_m^\alpha \quad \text{biased galaxy distribution}$$

$$\rho_g^s(\vec{x}) = S(\vec{x})\rho_g(\vec{x}) \quad \text{selected sample}$$

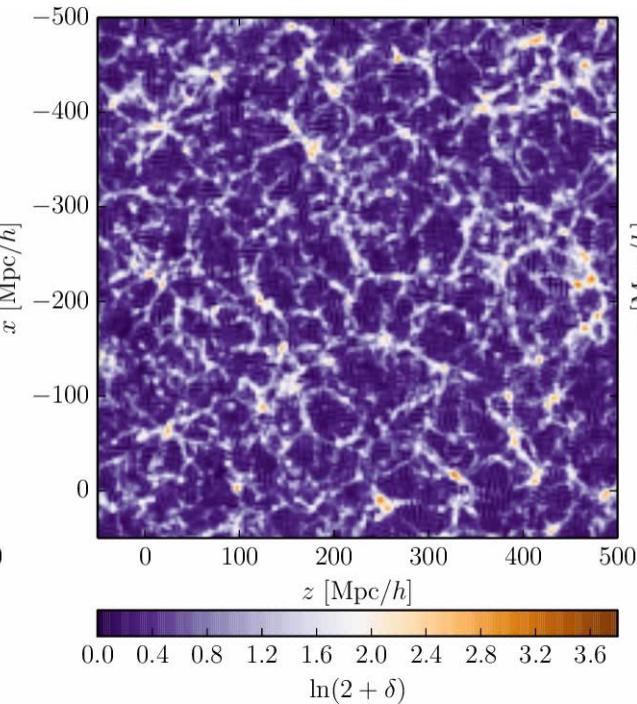
$$N_g \sim \mathcal{P}(N_g | \rho_g^s) \quad \text{galaxy number count:
random extraction (Poisson)}$$

- The multi-million dimensional posterior distribution is sampled via **Hamiltonian Monte Carlo**.

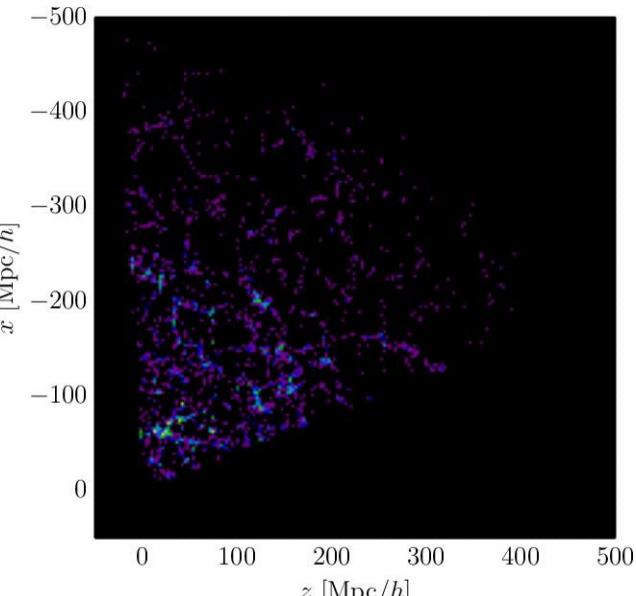
The BORG SDSS analysis



Initial conditions



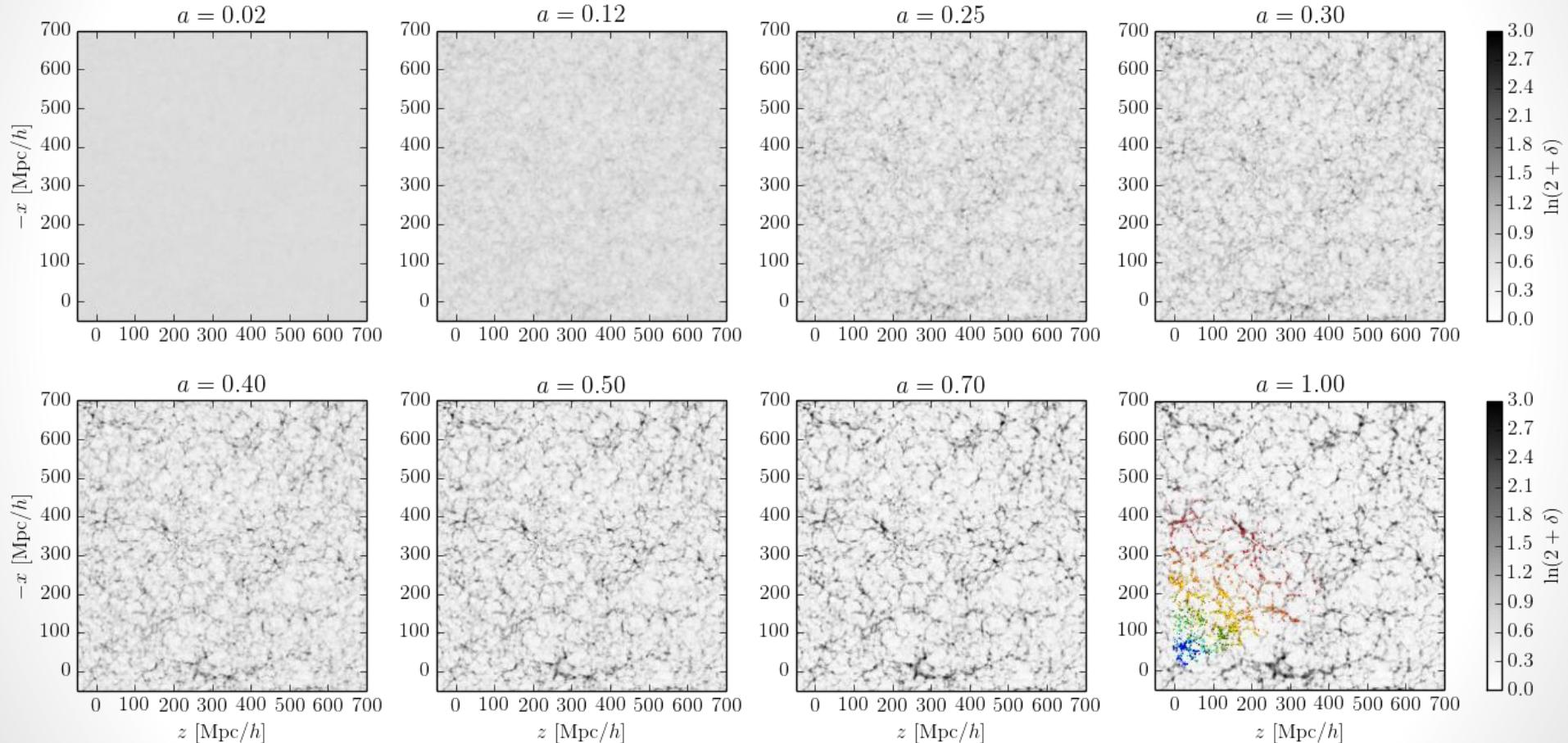
Final conditions



Observations

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

BORG infers the evolution of cosmic structure



BORG infers Lagrangian dynamics in real data

$$\vec{x} = \vec{q} + \vec{\Psi}(\vec{x})$$

- The dark matter phase-space sheet has been studied so far mostly in simulations

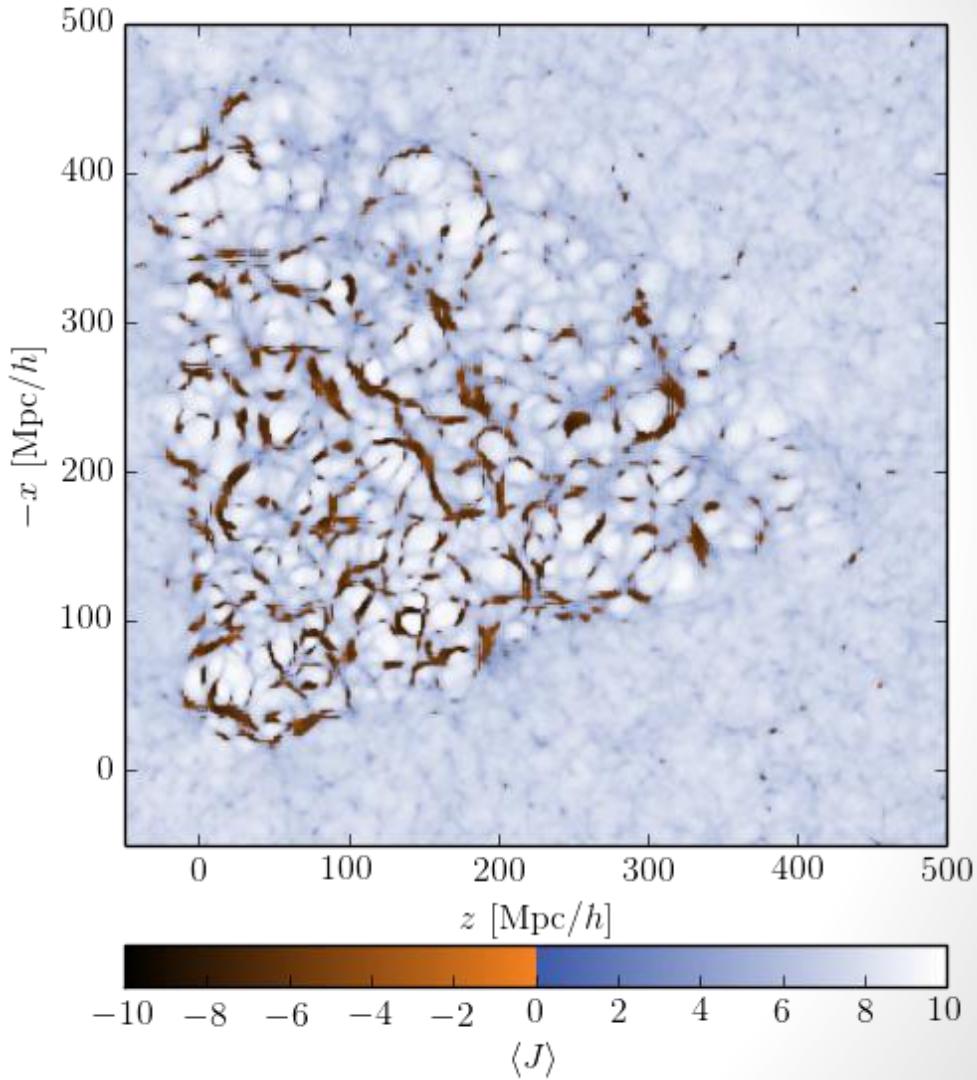
e.g. [Abel, Hahn & Kaehler 2012, arXiv:1111.3944](#)

[Shandarin, Habib & Heitmann 2012, arXiv:1111.2366](#)

[Neyrinck 2012, arXiv:1202.3364](#)

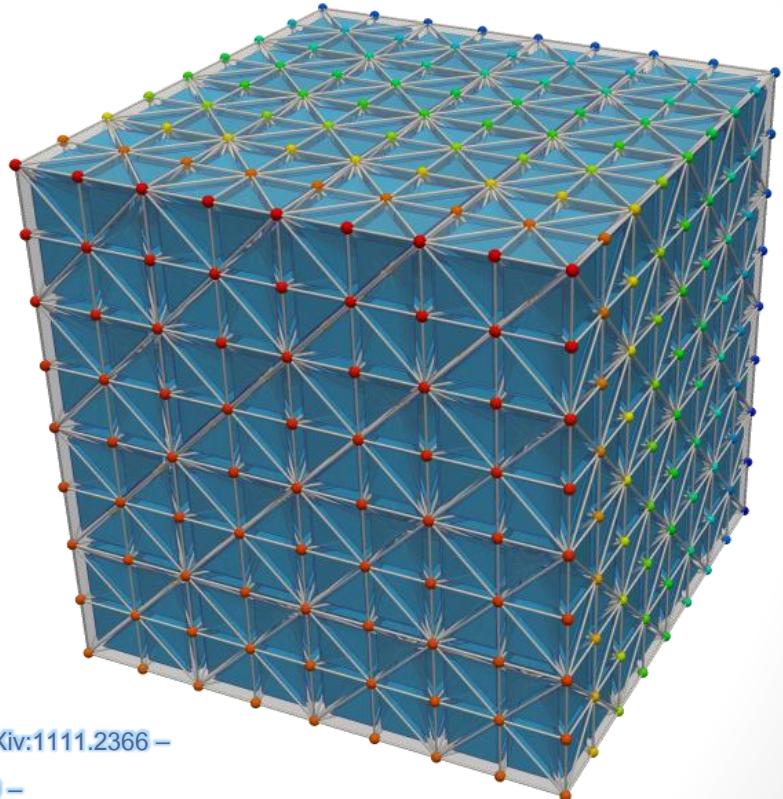
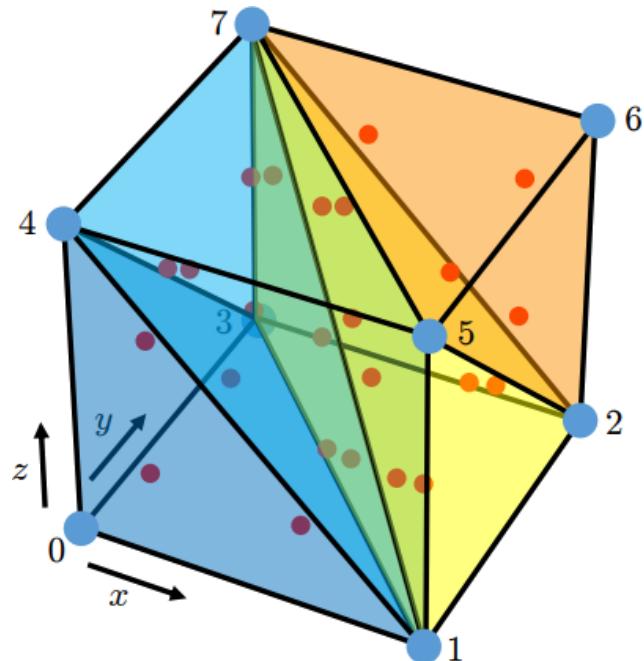
[Sousbie & Colombi 2016, arXiv:1509.07720](#)

- BORG infers Lagrangian dynamics in real data
- Identified structures have a direct physical interpretation



Tracing the dark matter phase-space sheet

- Delaunay tessellation of elementary Lagrangian cubes

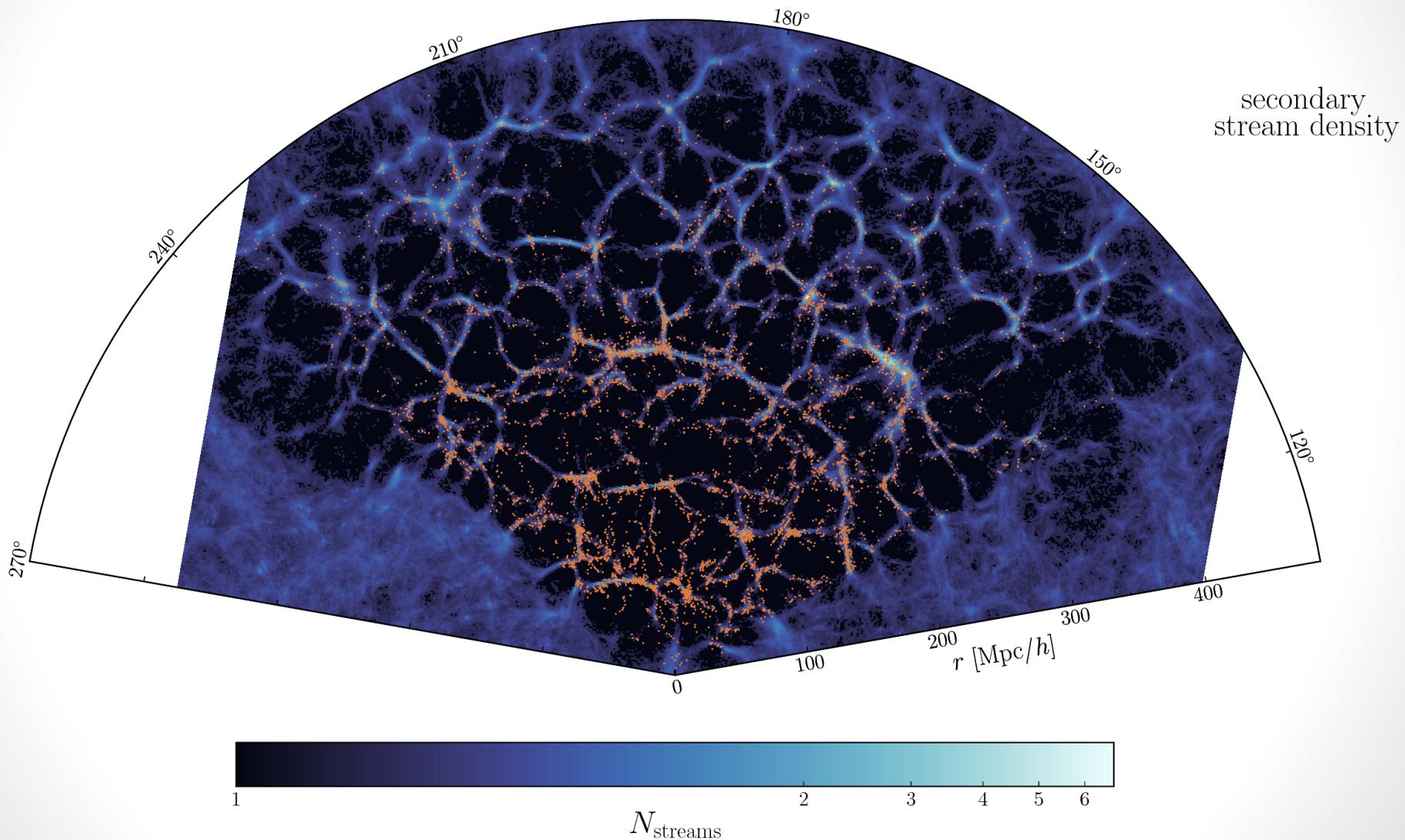


Abel, Hahn & Kaehler 2012, arXiv:1111.3944 – Shandarin, Habib & Heitmann 2012, arXiv:1111.2366 –

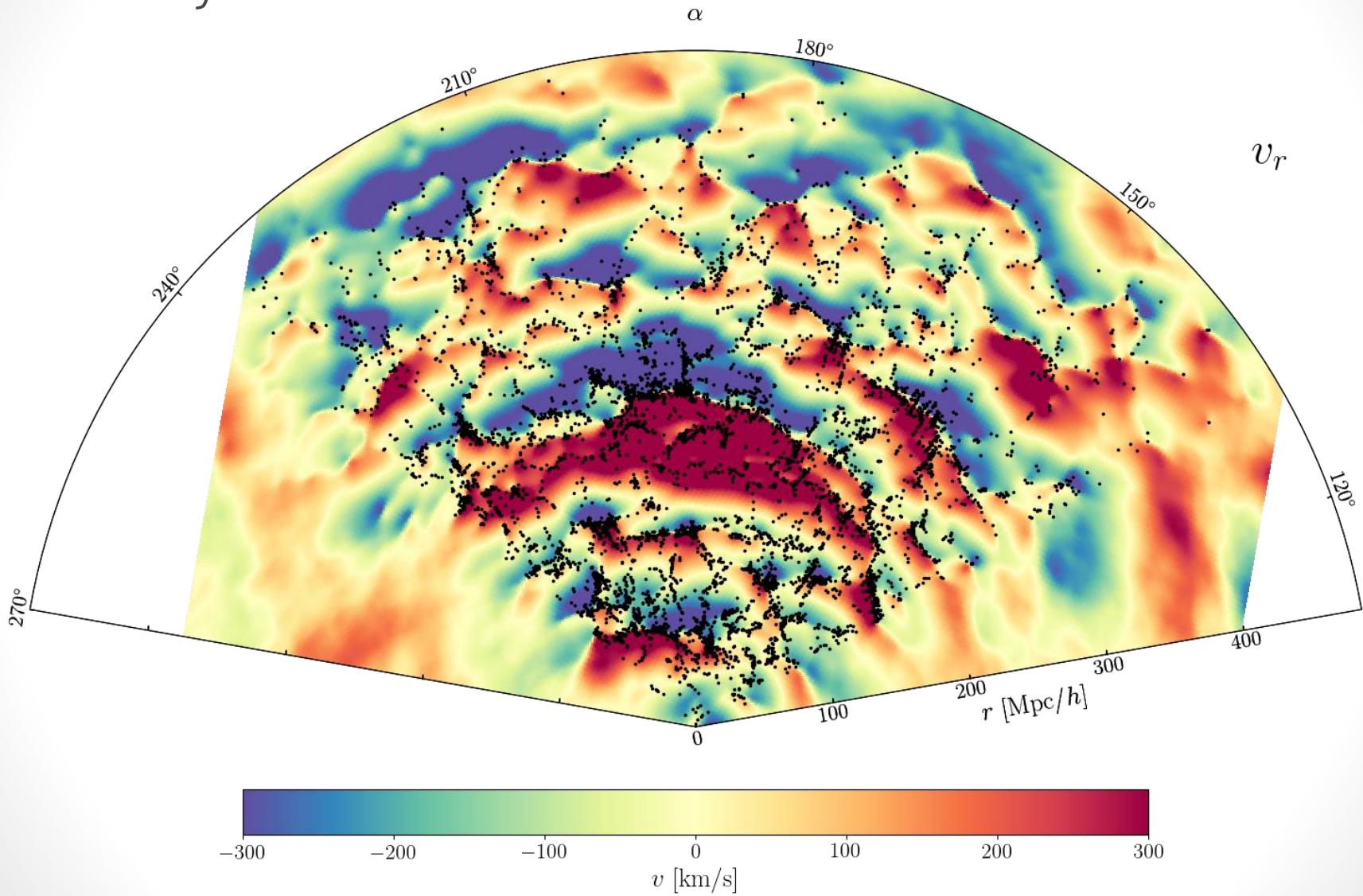
Hahn, Abel & Kaehler 2013, arXiv:1210.6652 – Hahn & Angulo 2016, arXiv:1501.01959 –

Sousbie & Colombi 2016, arXiv:1509.07720

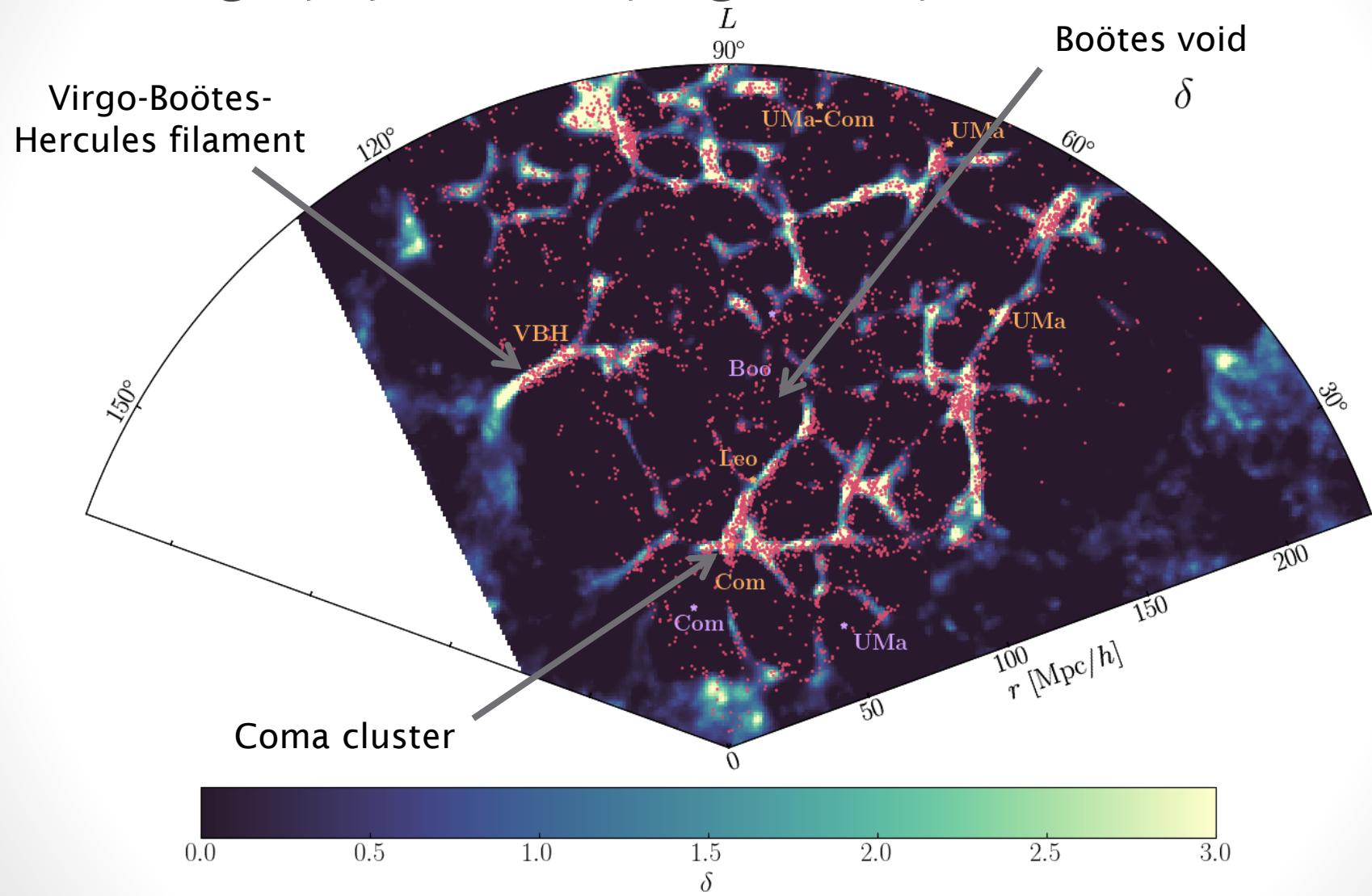
Dark matter stream density



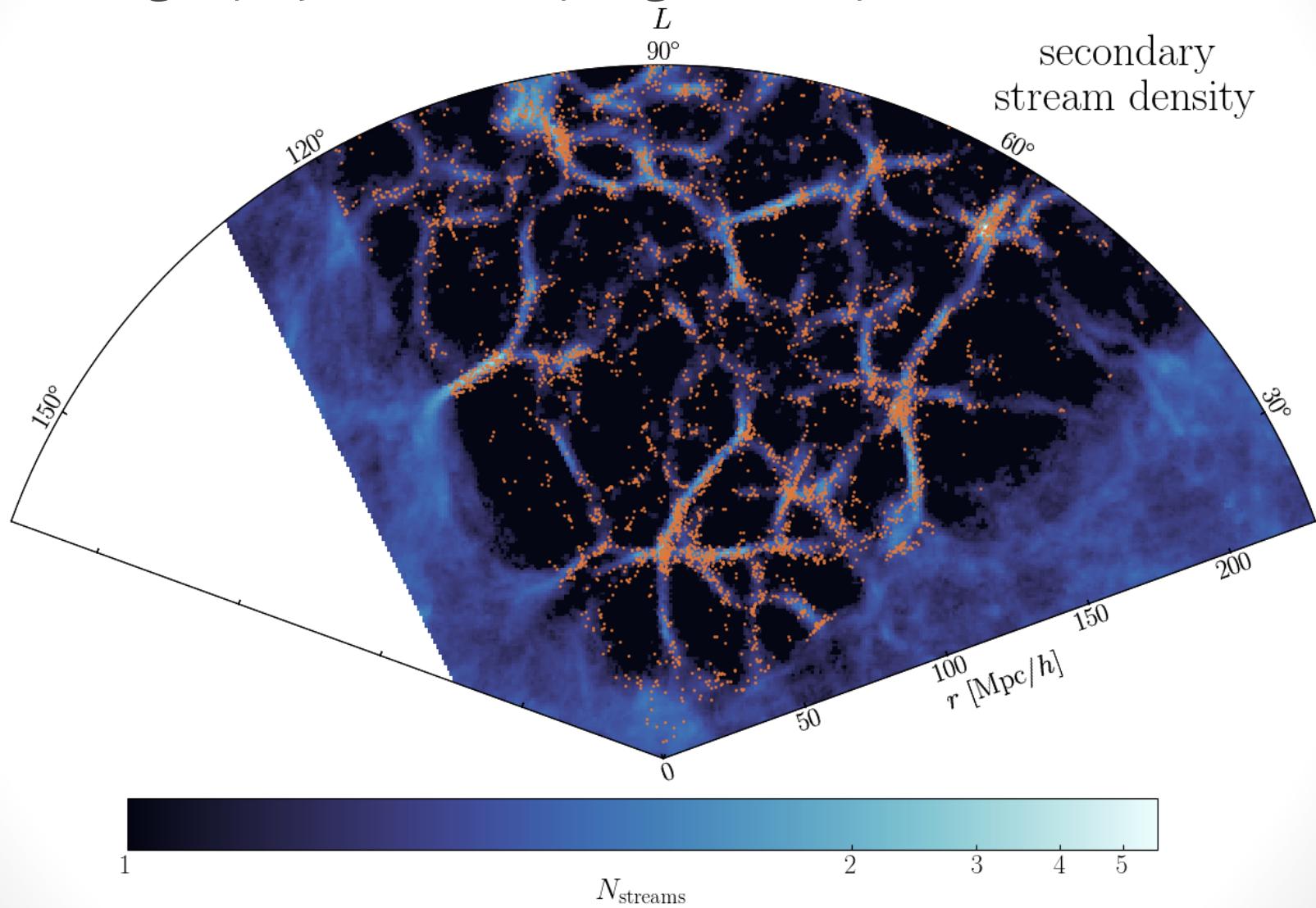
Velocity field



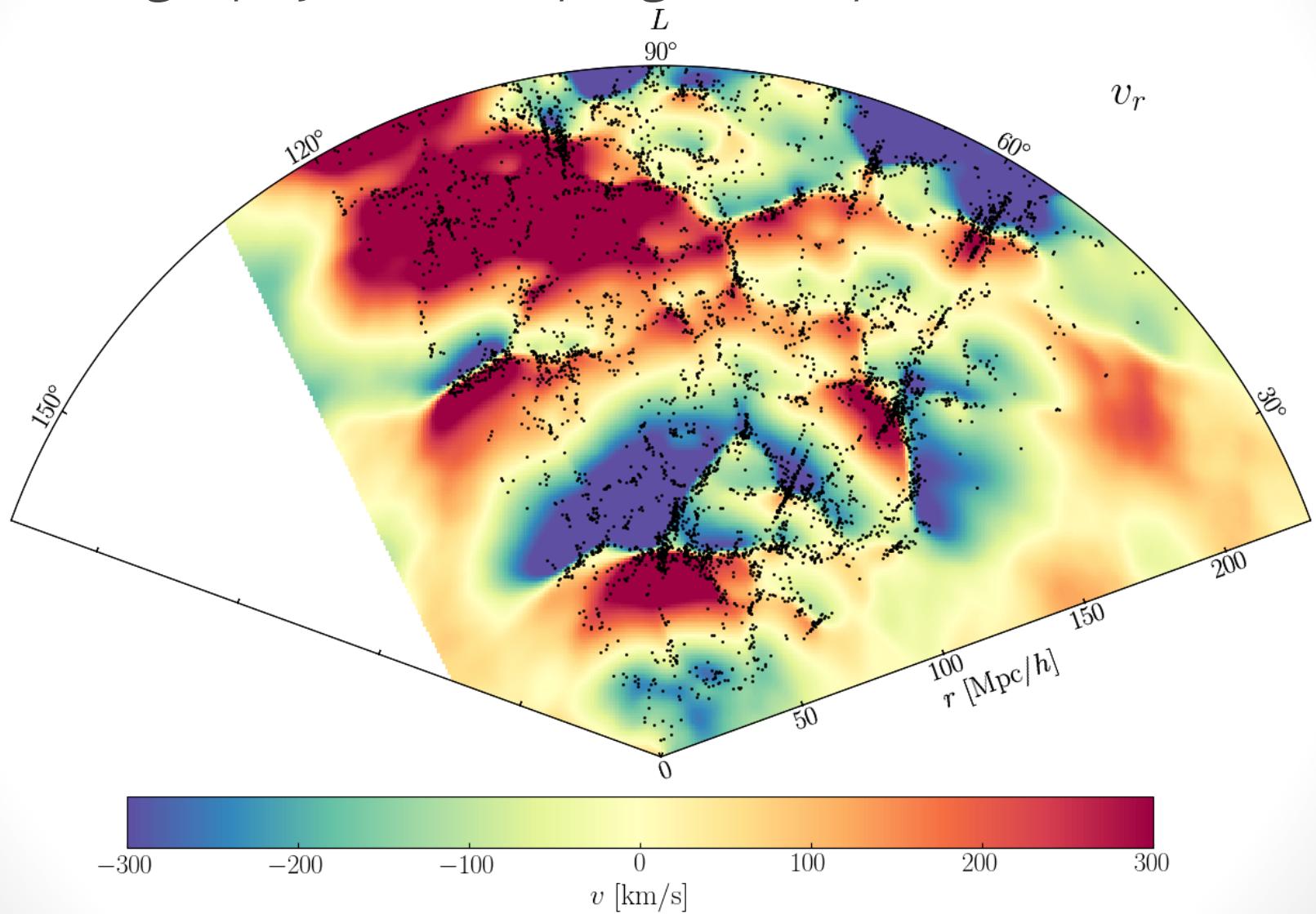
Cosmography in the supergalactic plane



Cosmography in the supergalactic plane



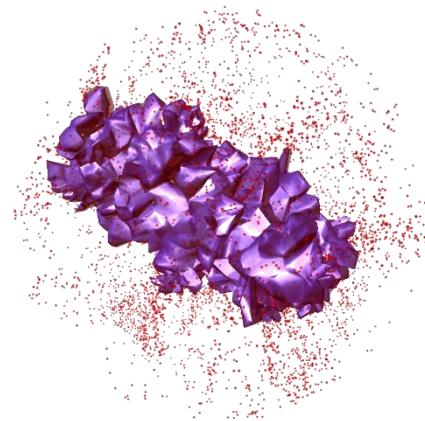
Cosmography in the supergalactic plane



Cosmic web elements: some algorithms

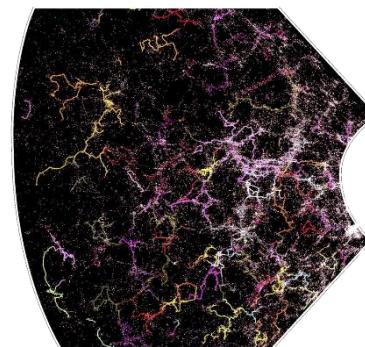
- “**Structure finders**” focus on one element at a time
 - **ZOBOV/VIDE**

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

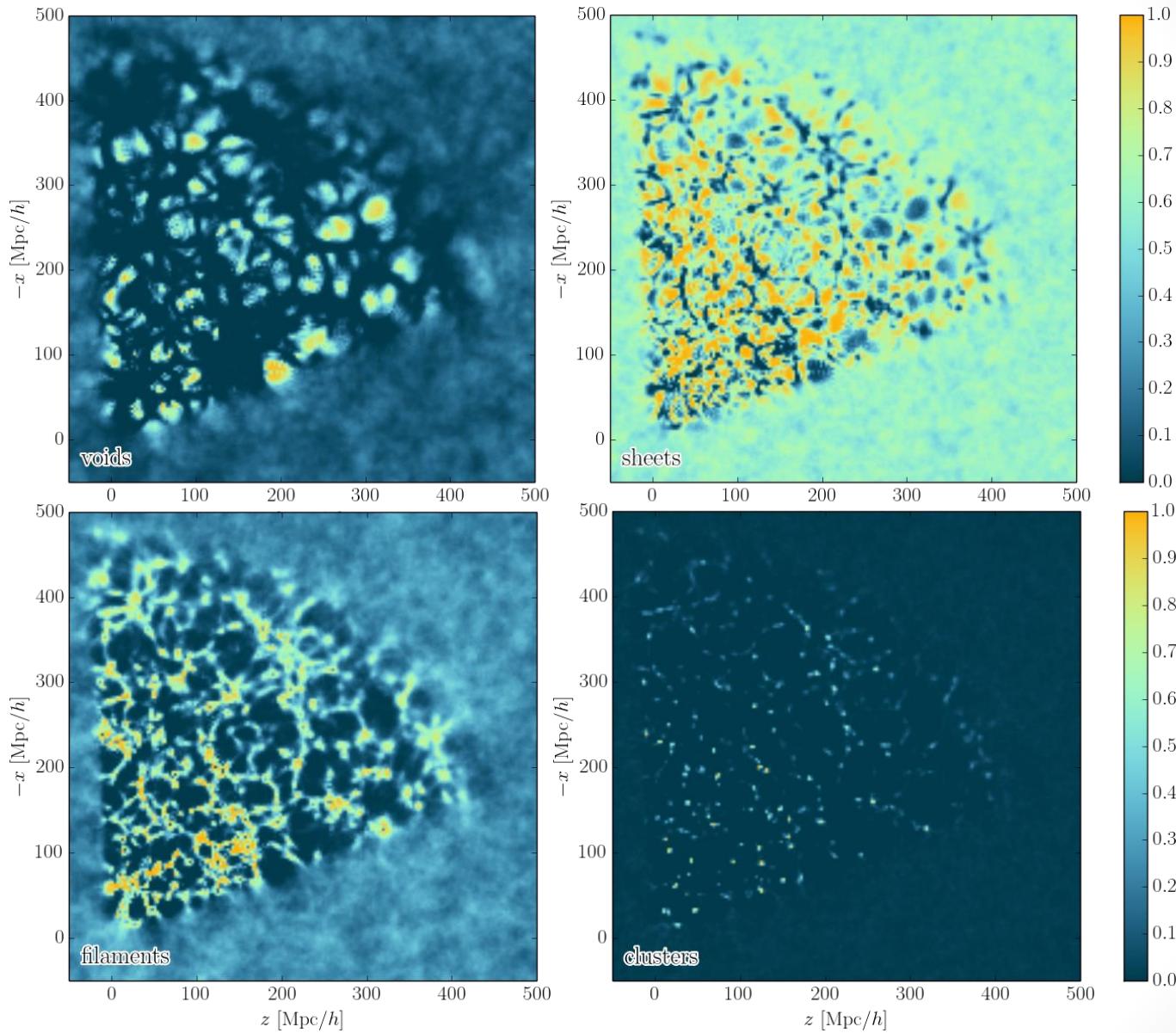


- “**Classifiers**” dissect the cosmic web all at once
 - The **T-web** (tidal field tensor)
[Hahn et al. 2007, arXiv:astro-ph/0610280](#)
 - **DIVA** (Lagrangian displacement field, potential structures)
[Lavaux & Wandelt 2010, arXiv:0906.4101](#)
 - **ORIGAMI** (particle crossings)
[Falck, Neyrinck & Szalay 2012, arXiv:1201.2353](#)
 - **LICH** (Lagrangian displacement field, potential and vortical structures)
[FL, Jasche, Lavaux, Wandelt & Percival 2017](#)

and many others...

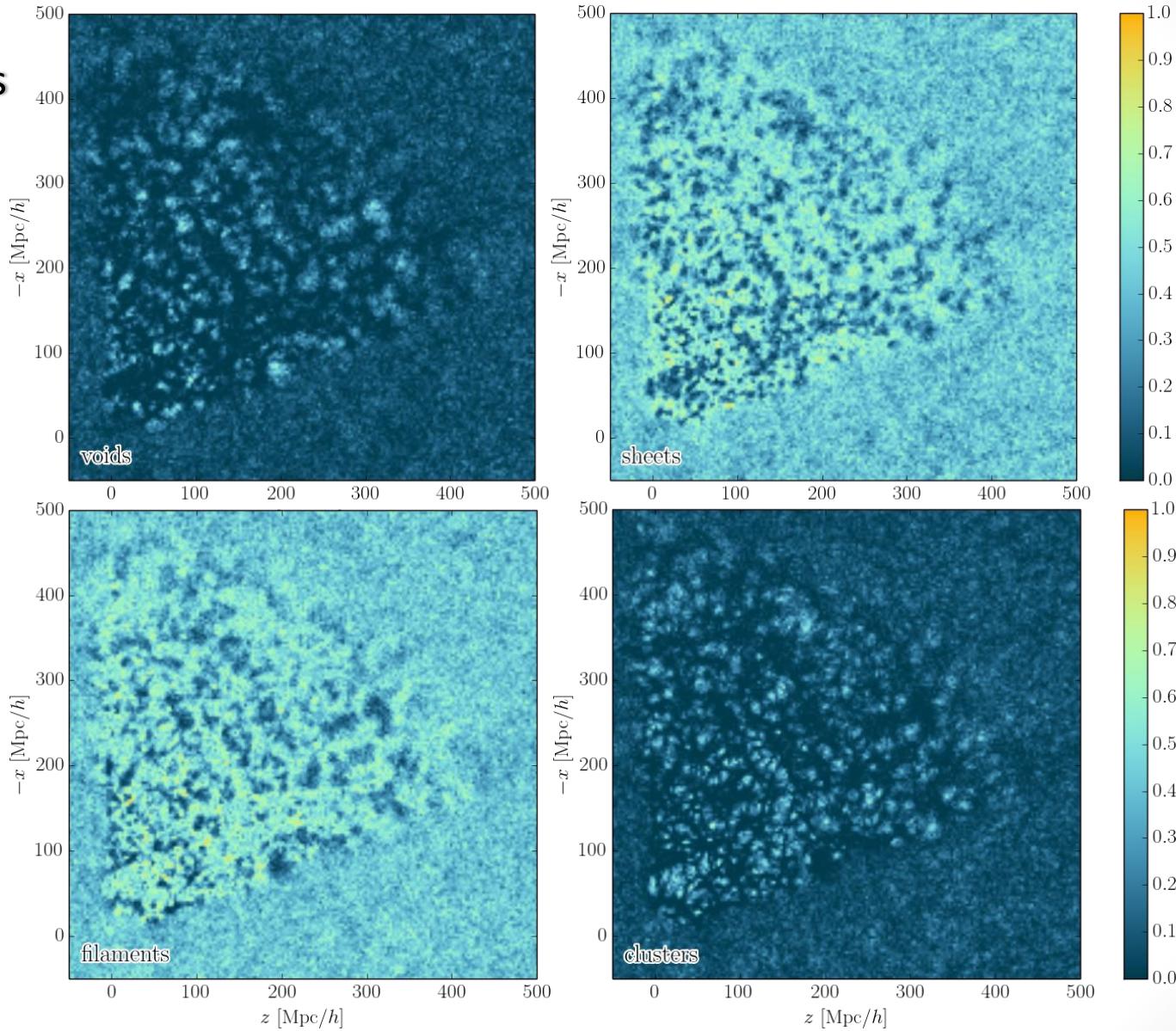
T-web structures inferred by BORG

Final conditions



T-web structures inferred by BORG

Initial conditions

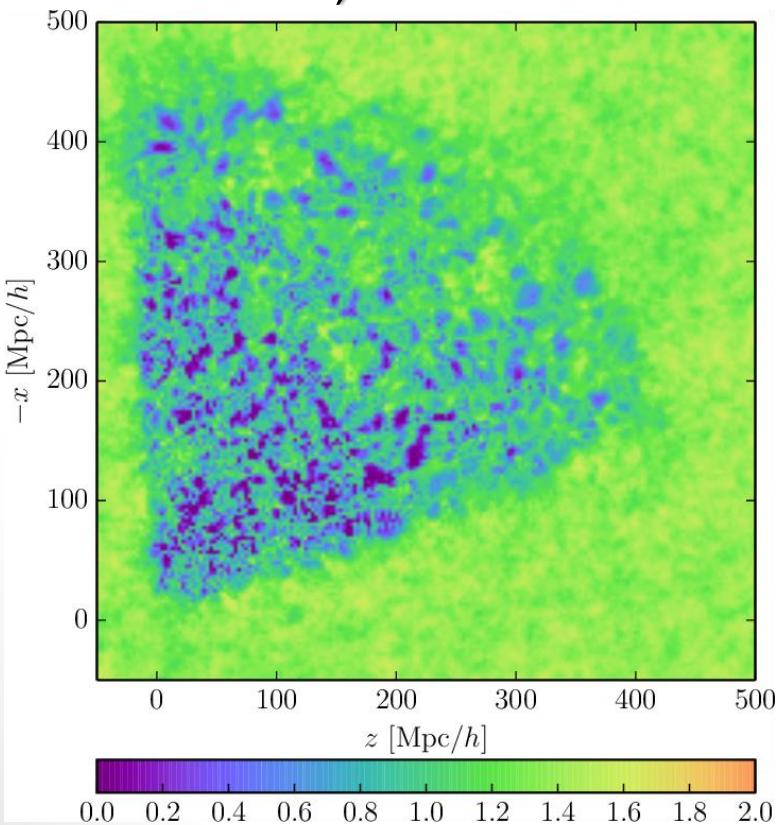


What is the information content of these maps?

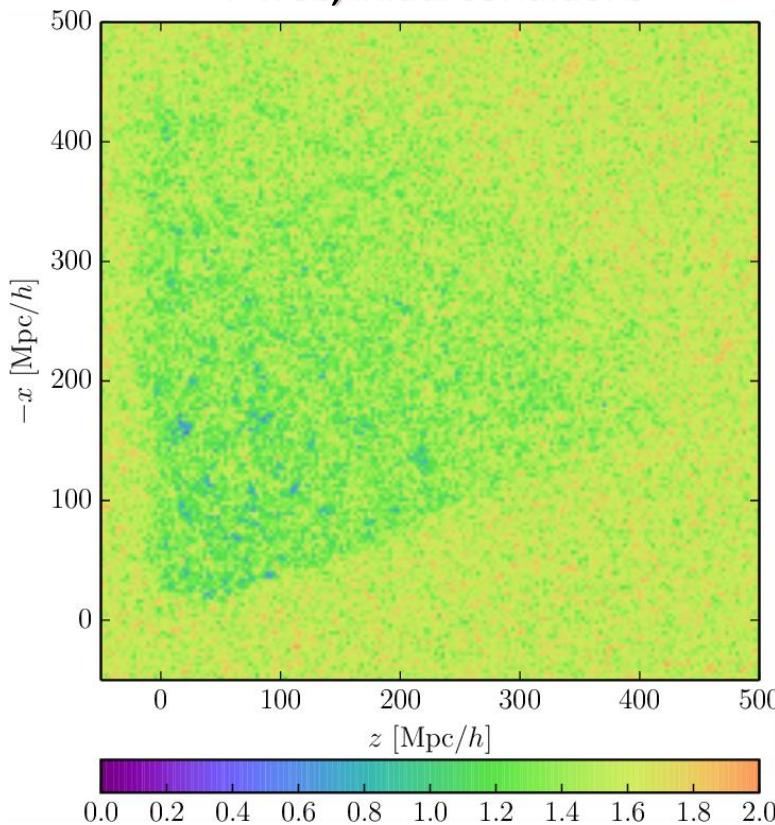
- Shannon entropy

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

T-web, final conditions



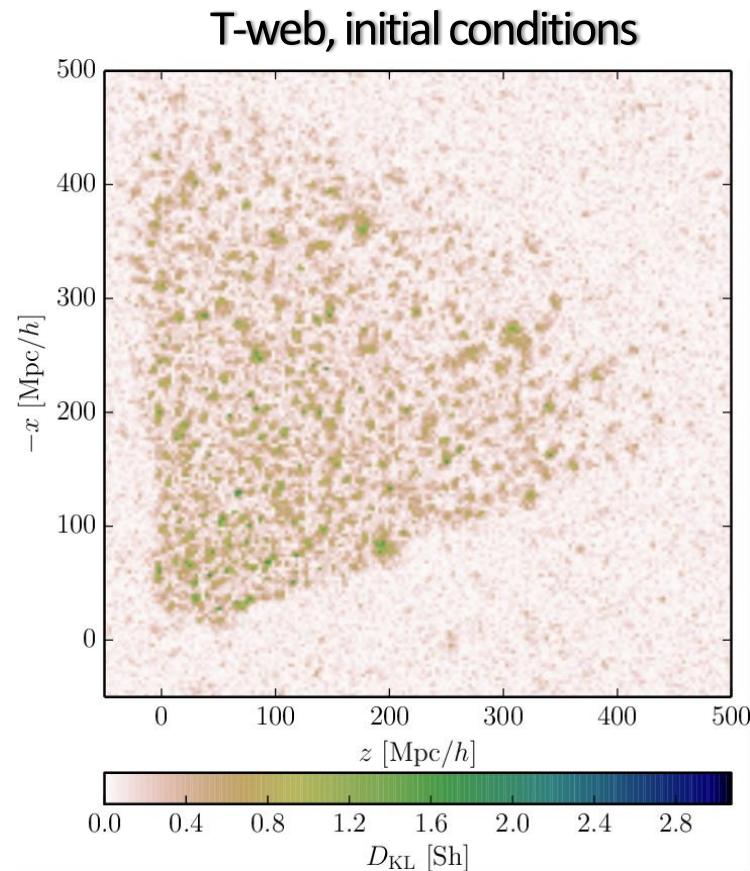
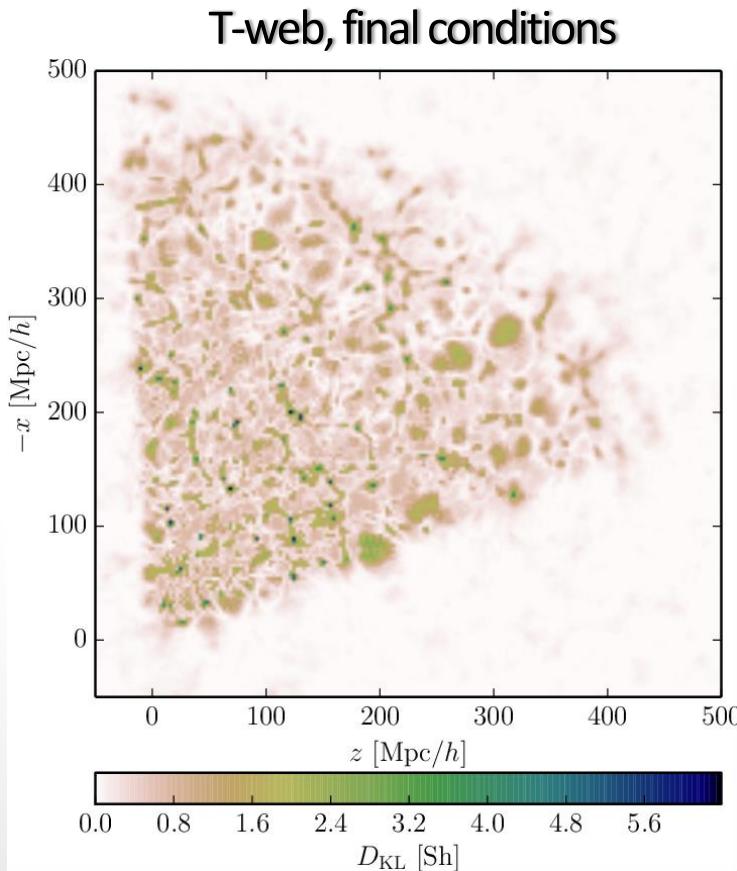
T-web, initial conditions



How much did the data surprise us?

- information gain a.k.a. relative entropy or Kullback-Leibler divergence

$$D_{\text{KL}} [\mathcal{P}(\mathbf{T}(\vec{x}_k)|d) || \mathcal{P}(\mathbf{T})] = \sum_i \mathcal{P}(\mathbf{T}_i(\vec{x}_k)|d) \log_2 \left(\frac{\mathcal{P}(\mathbf{T}_i(\vec{x}_k)|d)}{\mathcal{P}(\mathbf{T}_i)} \right) \quad \text{in Sh}$$



A decision rule for structure classification

- Space of “input features”:

$$\{T_0 = \text{void}, T_1 = \text{sheet}, T_2 = \text{filament}, T_3 = \text{cluster}\}$$

- Space of “actions”:

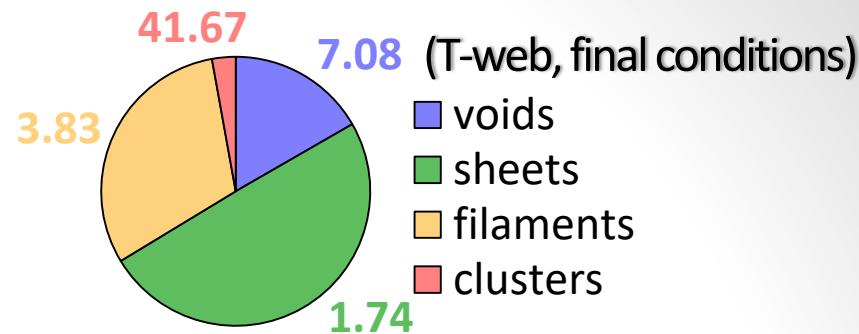
$$\{a_0 = \text{“decide void”}, a_1 = \text{“decide sheet”}, a_2 = \text{“decide filament”}, \\ a_3 = \text{“decide cluster”}, a_{-1} = \text{“do not decide”}\}$$

→ A problem of **Bayesian decision theory**:

one should take the action that maximizes the utility

$$U(a_j(\vec{x}_k)|d) = \sum_{i=0}^3 G(a_j|T_i) \mathcal{P}(T_i(\vec{x}_k)|d)$$

- How to write down the gain functions?



Gambling with the Universe

- One proposal:

$$G(a_j | \mathbf{T}_i) = \begin{cases} \frac{1}{\mathcal{P}(\mathbf{T}_i)} - \alpha & \text{if } j \in \llbracket 0, 3 \rrbracket \text{ and } i = j \quad \text{"Winning"} \\ -\alpha & \text{if } j \in \llbracket 0, 3 \rrbracket \text{ and } i \neq j \quad \text{"Losing"} \\ 0 & \text{if } j = -1. \quad \text{"Not playing"} \end{cases}$$

- Without data, the expected utility is

$$U(a_j) = 1 - \alpha \quad \text{if } j \neq 1 \quad \text{"Playing the game"}$$

$$U(a_{-1}) = 0 \quad \text{"Not playing the game"}$$

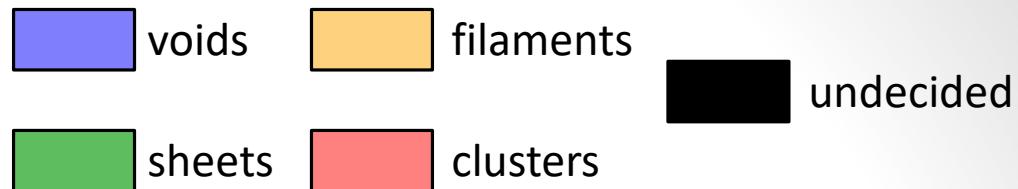
- With $\alpha = 1$, it's a *fair game* \rightarrow always play

\rightarrow "speculative map" of the LSS

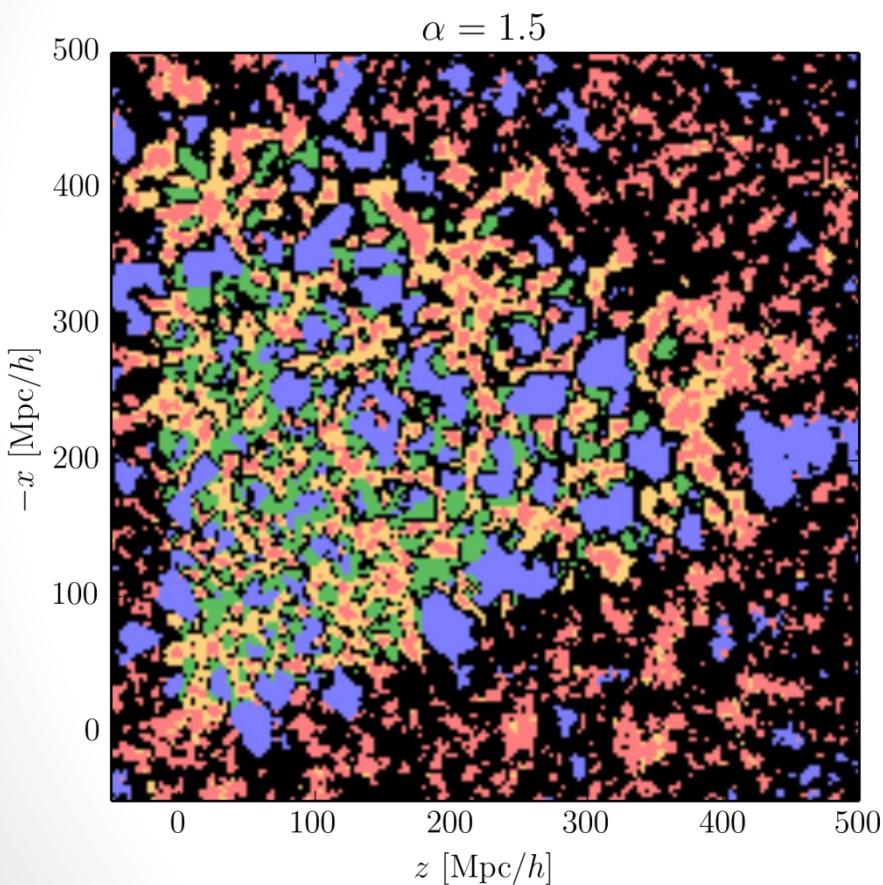
- Values $\alpha > 1$ represent an *aversion for risk*

\rightarrow increasingly "conservative maps" of the LSS

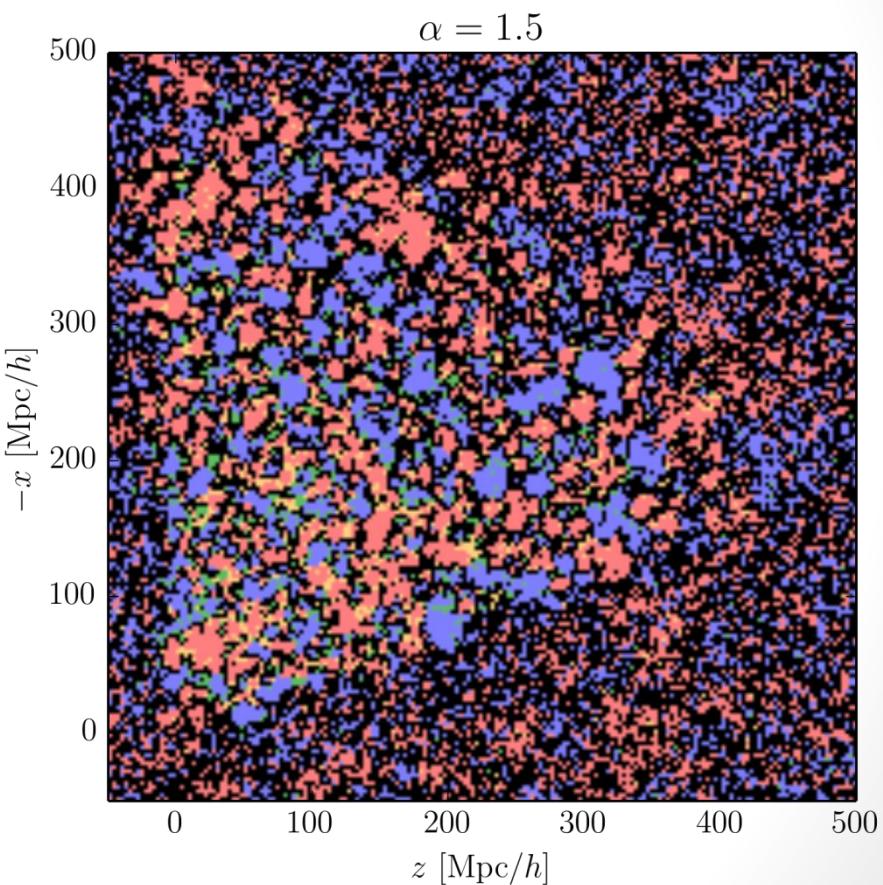
Playing the game...



Final conditions

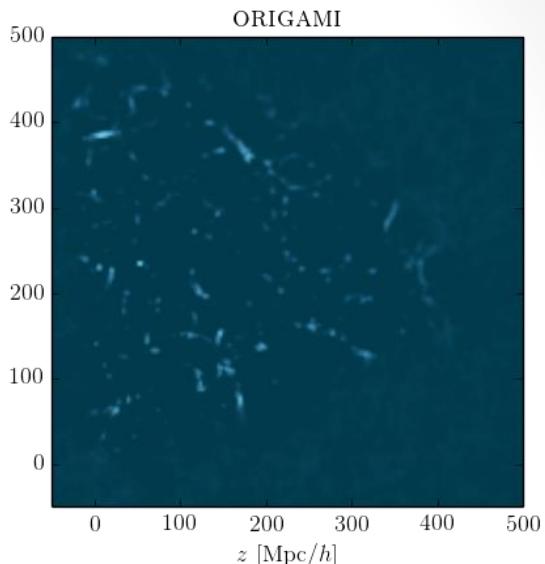
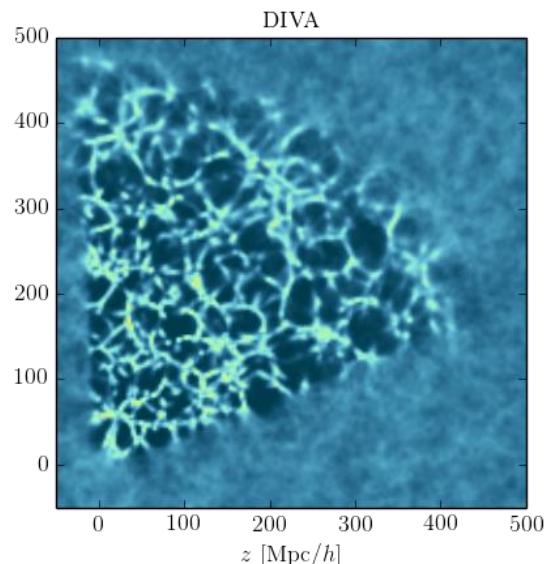
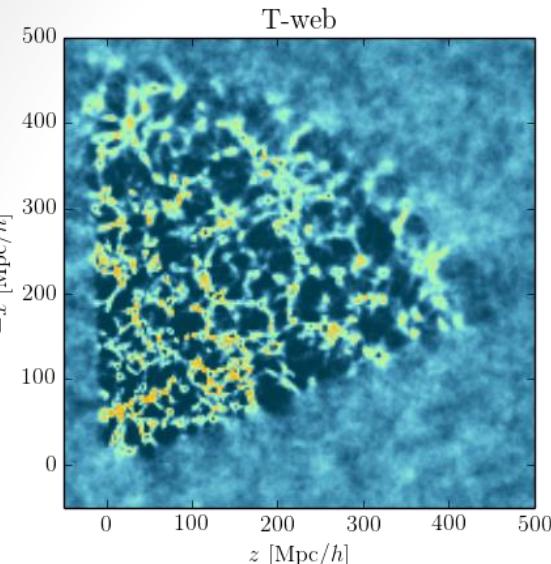


Initial conditions

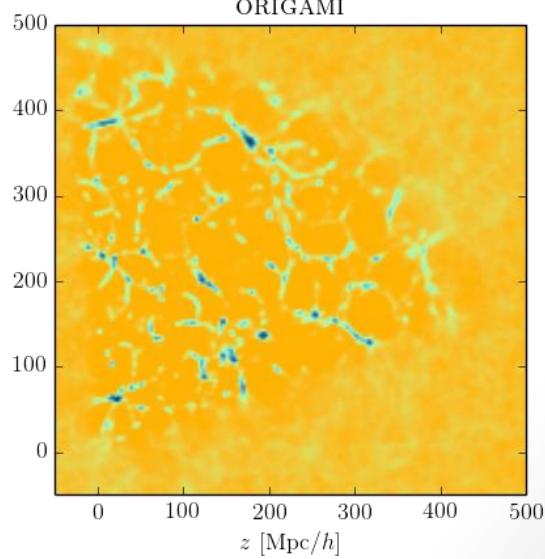
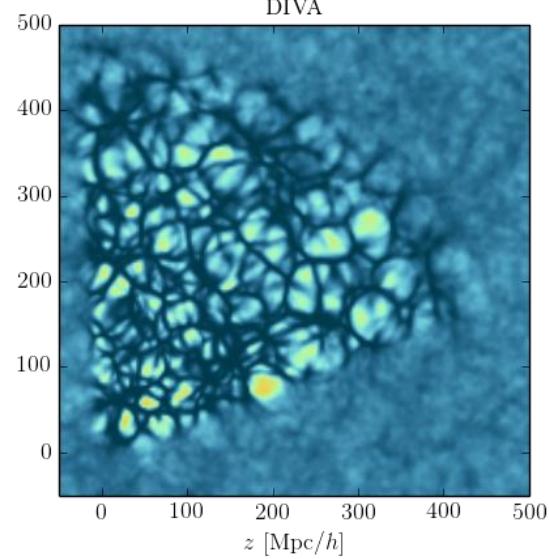
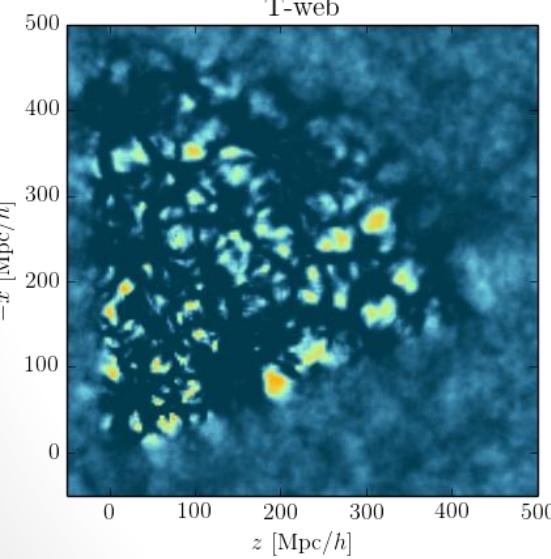


Comparing classifiers

Filaments



Voids



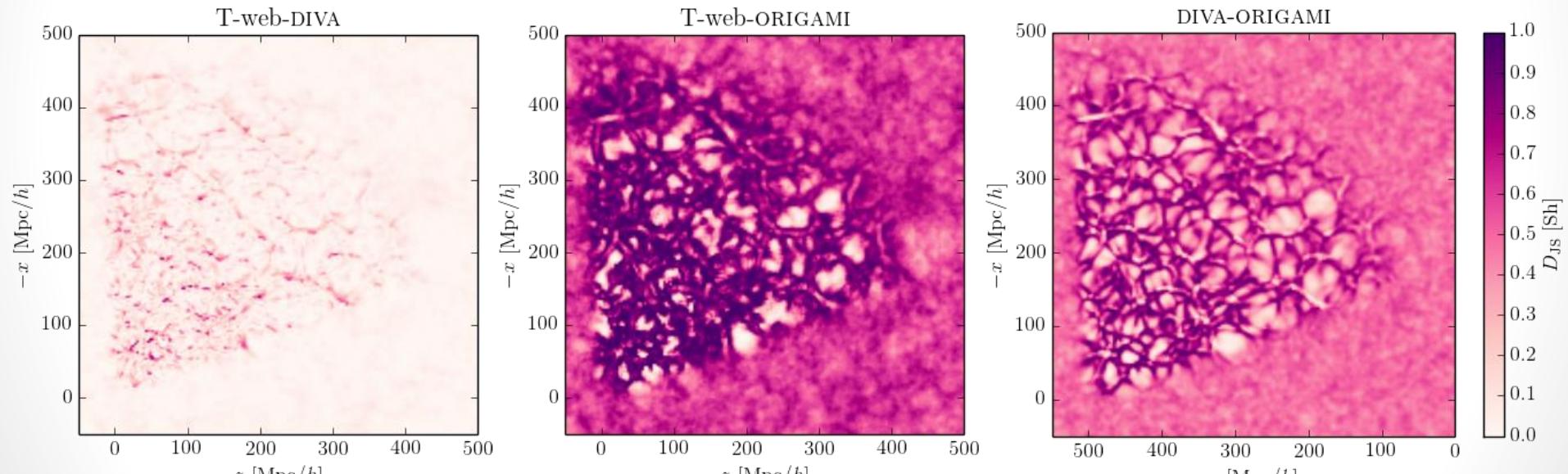
FL, Jasche & Wandelt 2015, arXiv:1502.02690

FL, Jasche, Lavaux, Wandelt & Percival 2017, arXiv:1601.00093

How similar are different classifications?

- Jensen-Shannon divergence

$$D_{\text{JS}}[\mathcal{P} : \mathcal{Q}] \equiv \frac{1}{2} D_{\text{KL}} \left[\mathcal{P} \middle\| \frac{\mathcal{P} + \mathcal{Q}}{2} \right] + \frac{1}{2} D_{\text{KL}} \left[\mathcal{Q} \middle\| \frac{\mathcal{P} + \mathcal{Q}}{2} \right]$$



(more about the Jensen-Shannon divergence later)

The BORG SDSS data release

- All data products are publicly available: [doi: 10.5281/zenodo.1455729](https://doi.org/10.5281/zenodo.1455729)

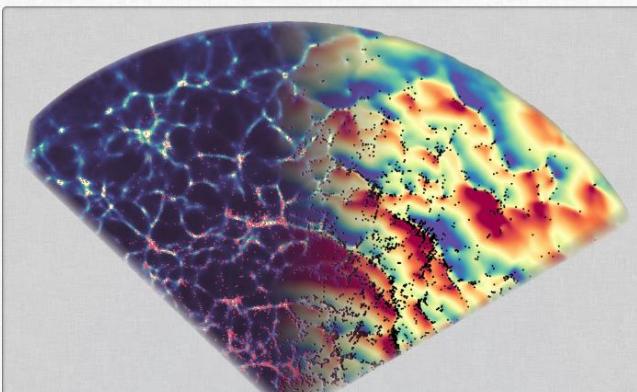
<http://data.florent-leclercq.eu/>

PUBLIC DATA AND SOFTWARE

The BORG SDSS data release

[GitHub v1.0](#) [DOI 10.5281/zenodo.1455729](#)

This is the repository for the public releases of data products that follow a **chrono-cosmographic analysis of the three-dimensional large-scale structure of the nearby Universe**, as traced by SDSS main sample galaxies. These data are the product of a collaboration of [Jens Jasche](#), [Florent Leclercq](#) and [Benjamin Wandelt](#), with additional contributions from [Nico Hamaus](#), [Guilhem Lavau](#) and [P.M. Sutter](#).



Slice through the cosmic density field (left side) and radial velocity field (right side) inferred by BORG from SDSS data and estimated using phase-space interpolation techniques. The slice is about 3 Mpc/h thick around the celestial equator. Figures by Florent Leclercq ([Leclercq et al. 2017](#)), montage by [Guilhem Lavau](#).

https://github.com/florent-leclercq/borg_sdss_data_release

```
In [4]: #Minimum and maximum position along the x-axis in Mpc/h
xmin=velocity['ranges'][0]
xmax=velocity['ranges'][1]

##minimum and maximum position along the y-axis in Mpc/h
ymin=velocity['ranges'][2]
ymax=velocity['ranges'][3]

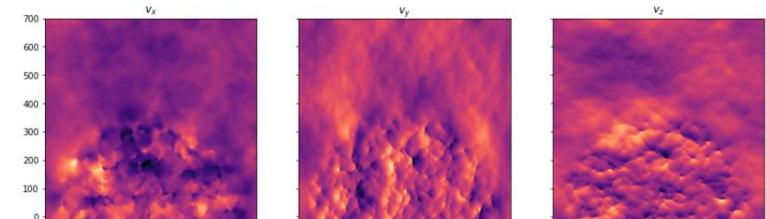
##minimum and maximum position along the z-axis in Mpc/h
zmin=velocity['ranges'][4]
zmax=velocity['ranges'][5]
```

Units are Mpc/h.

(Note that all the maps that are part of the BORG SDSS data products have consistent coordinate systems. The coordinate transform to change from Cartesian to spherical coordinates and vice versa is given in appendix B of [Jasche et al. 2015](#)).

Example plot: mean

```
In [5]: from matplotlib import pyplot as plt
%matplotlib inline
f, (ax1, ax2, ax3) = plt.subplots(1, 3, sharey=True, figsize=(15,5))
ax1.imshow(vx_mean[:,128], origin='lower', extent=[ymin,ymax,zmin,zmax], cmap="magma")
ax1.set_title('$_v_x$')
ax1.set_aspect('equal')
ax2.imshow(vy_mean[:,128], origin='lower', extent=[ymin,ymax,zmin,zmax], cmap="magma")
ax2.set_title('$_v_y$')
ax3.imshow(vz_mean[:,128], origin='lower', extent=[ymin,ymax,zmin,zmax], cmap="magma")
ax3.set_title('$_v_z$')
plt.show()
```



Conclusions

- **BORG** is a Bayesian inference engine allowing the analysis of the large-scale structure and its formation history.
- Thanks to **BORG**, the cosmic web can be described using various classifiers.
- A probabilistic analysis of the cosmic web yields a data-supported connection between cosmology and information theory.
- Decision theory offers a framework to choose between different classifiers, with utility functions depending on the desired use.