The cosmological information of the cosmic web

Tony Bonnaire

École Normale Supérieure, Physics Department & Center for data sciences, Paris

N. Aghanim

J. Kuruvilla

A. Decelle

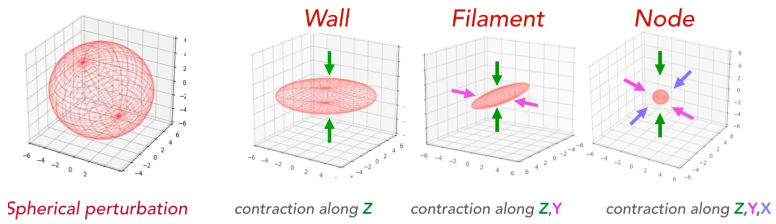
Talk based on

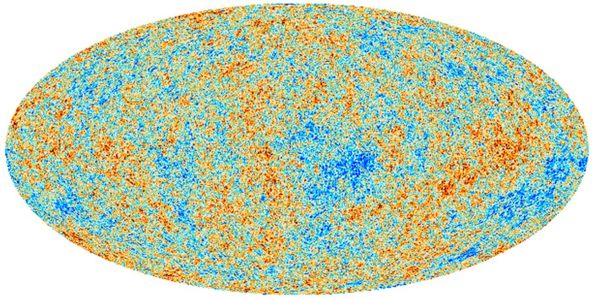
Bonnaire et al., Cosmology with cosmic web environments I. & II., A&A, 2022, 2023

Image from the Illustris collaboration

The cosmic web

The spatial arrangement of the large-scale matter distribution, commonly called the Cosmic Web, falls into 4 main types of structures: Nodes, Filaments, Sheets or walls, Voids

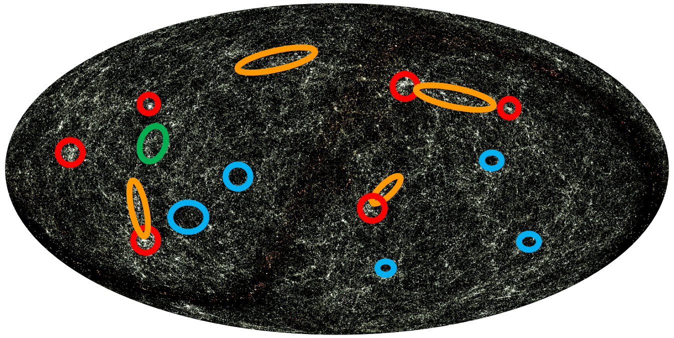




Initial density fluctuations, Gaussian field, $t = 380\,000$ years
[Planck collaboration, 2015]

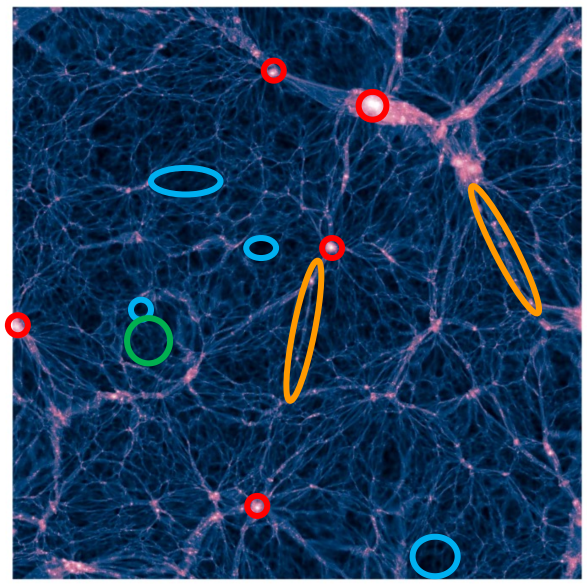
Gravity

Observation



Observed local galaxy distribution from 2MASS
[Skrutskie+06]

Simulation

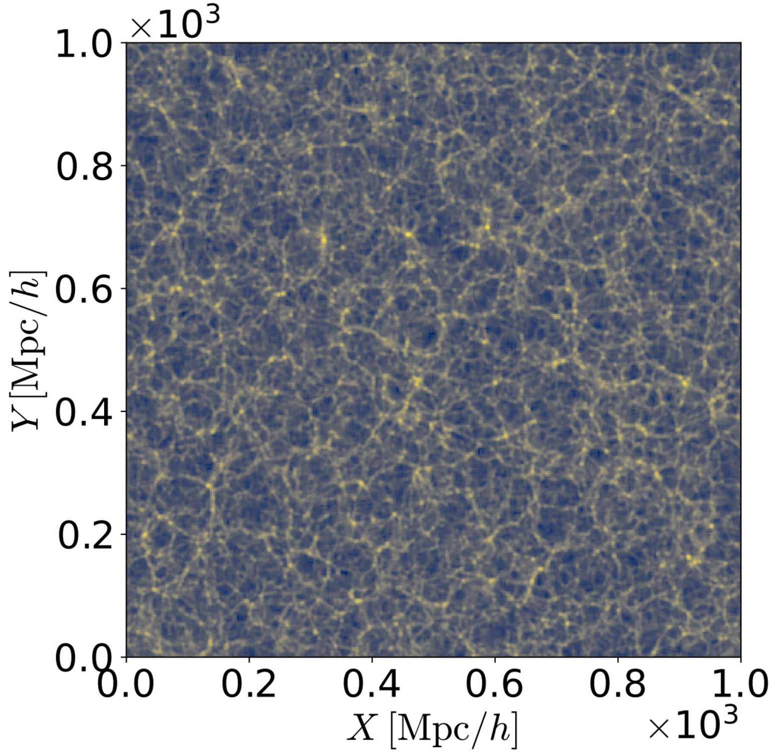


Dark matter density field in the Illustris simulation
[Vogelsberger+14]

Constraining cosmology: Context

- Statistical estimators of the spatial distribution of matter are needed = **summary statistics**
- The natural way of describing centred fields $< \delta_m > = 0$ is to use $< \delta_{\rm m} \delta_{\rm m} >$ which defines the **matter power spectrum** in Fourier space, $P^{\rm mm}$

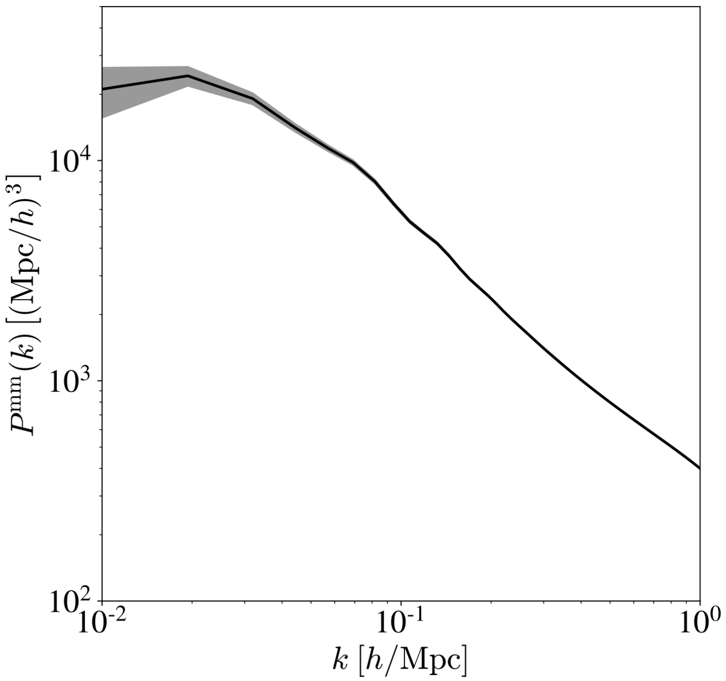
Dark matter density field from the Quijote simulation [Villaescusa-Navarro+18]



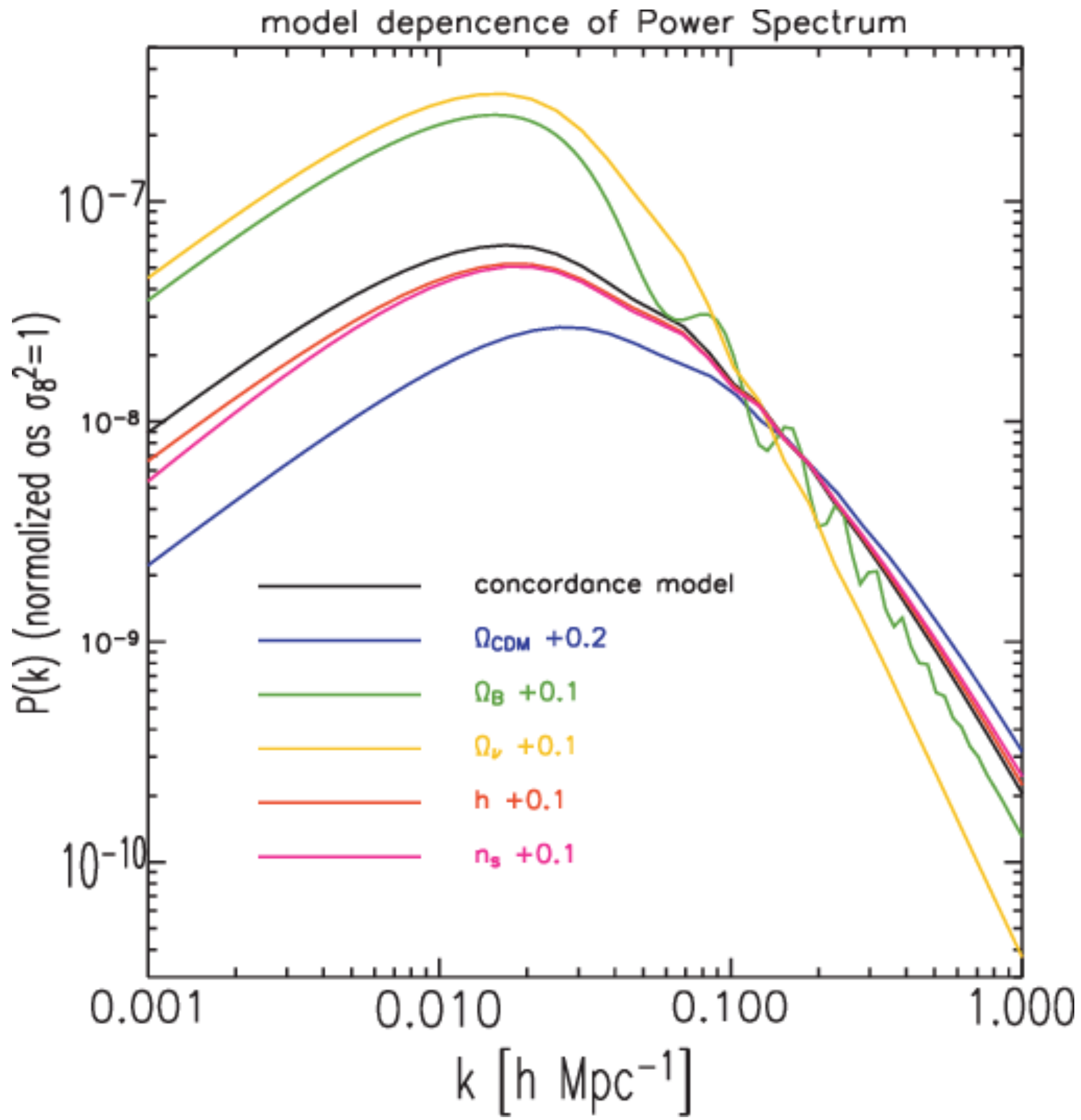
Summarised as

$$P^{\rm mm}(k)\delta_{\rm D}(\boldsymbol{k}+\boldsymbol{k}') = \frac{1}{(2\pi)^3} < \tilde{\delta}^{\rm m}(\boldsymbol{k})\tilde{\delta}^{\rm m}(\boldsymbol{k}') >$$

Matter power spectrum



Constraining cosmology: Context



Plot from Matsubara+06

Ω_{m}: matter density

Ω_{b}: baryon density

h: Hubble parameter at $z = 0$

n_s: spectral index

M_ν: summed neutrino mass

σ_8: amplitude of the linear power spectrum at a scale of 8 Mpc/h

Matter power spectrum: sensitive to cosmology

→ Constrains the model and its parameters

How to fight the degeneracies?

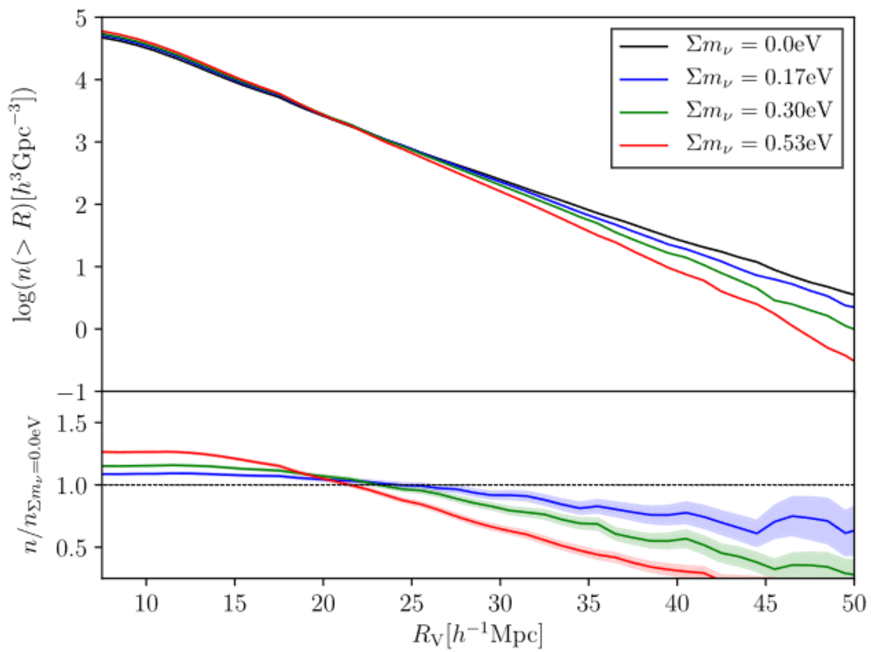
Include higher-order information directly or indirectly:

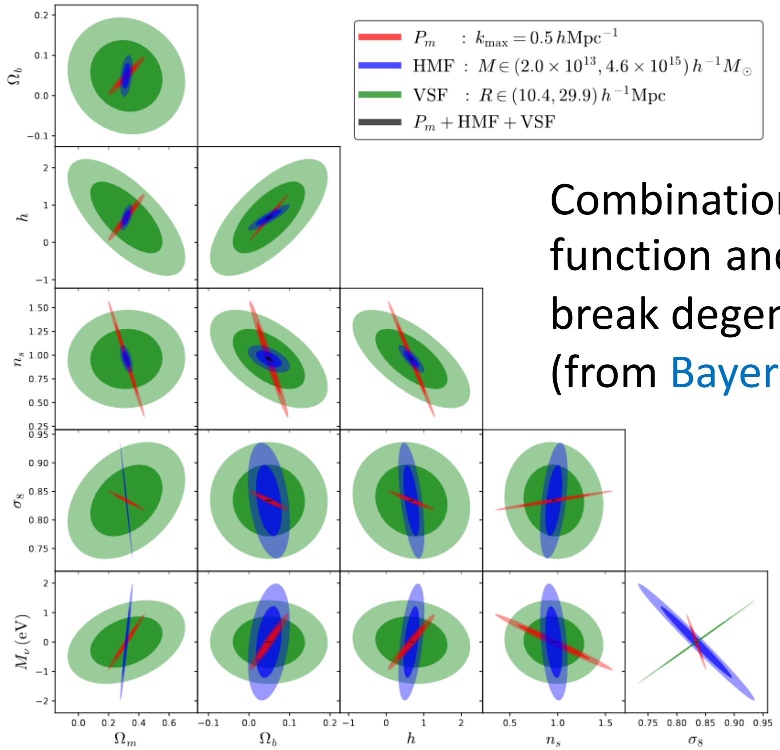
- Direct higher-orders [Yankelevich+19, Hahn+21, Agarwal+21, Gualdi+21]
- Velocity information [Mueller+15, Kuruvilla+21]
- Marked power spectrum [Beisbart+00, Stheth+06, White+16, Massara+20]
- Neural networks [Ribli+19]
- Wavelet scattering transform [Mallat+12, Allys+19/20, Cheng & Menard+20, Valogiannis+22/23]
- **Environments information** [Kreisch+19, Bayer+21, Bonnaire+22/23]
- Density splits [Uhleman+19, Paillas+20]
- MST information [Naidoo+19, Naidoo+21]

Cosmology with voids and nodes

- Massive nodes are used through their distribution of counts, shapes, etc. to break degeneracies [Bachal+97, Holder+01]
- Voids are pristine environments perfect for the study of dark energy and to constrain neutrino mass [e.g. Pisani+15, Massara+15]

Void abundance sensitive to neutrino mass (from Kreisch+19)





Combination of halo mass function and void size function break degeneracies (from Bayer+21)

Question addressed

What is the theoretical cosmological information contained in the cosmic web environments?

What is the theoretical cosmological information contained in the cosmic web environments?

Probing large, linear and small, non-linear scales using simulations

Constraints on 6 cosmological parameters $(\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, h, n_{\mathrm{s}}, \sigma_8, M_\nu)$

Statistical estimator in environments

Quijote simulations

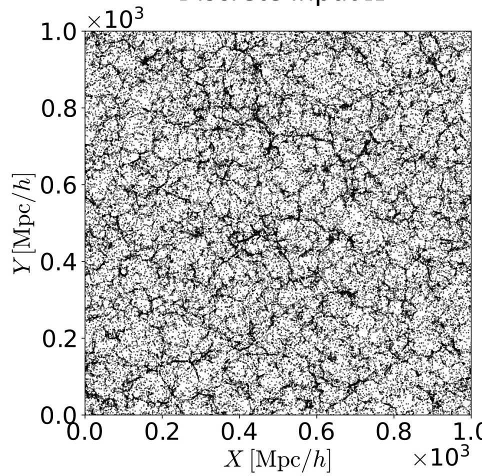
- Quijote [Villaescusa-Navarro+20] = large suite of 44,100 simulations spanning thousands of cosmological models
- Fiducial cosmology consistent with the Planck15 cosmology

Name	Ω_m	Ω_b	h	n_s	σ_8	M_ν	ICs	# of real.
Fiducial	0.3175	0.049	0.6711	0.9624	0.834	0	2LPT	15000
Ω_m^+	0.3275	0.049	0.6711	0.9624	0.834	0	2LPT	500
Ω_m^-	0.3075	0.049	0.6711	0.9624	0.834	0	2LPT	500
Ω_b^+	0.3175	0.051	0.6711	0.9624	0.834	0	2LPT	500
Ω_b^-	0.3175	0.047	0.6711	0.9624	0.834	0	2LPT	500
h^+	0.3175	0.049	0.6911	0.9624	0.834	0	2LPT	500
h^-	0.3175	0.049	0.6511	0.9624	0.834	0	2LPT	500
n_s^+	0.3175	0.049	0.6711	0.9824	0.834	0	2LPT	500
n_s^-	0.3175	0.049	0.6711	0.9424	0.834	0	2LPT	500
σ_8^+	0.3175	0.049	0.6711	0.9624	0.849	0	2LPT	500
σ_8^-	0.3175	0.049	0.6711	0.9624	0.819	0	2LPT	500
M_ν^0	0.3175	0.049	0.6711	0.9624	0.834	0	ZA	500
M_ν^+	0.3175	0.049	0.6711	0.9624	0.834	0.1	ZA	500
M_ν^{++}	0.3175	0.049	0.6711	0.9624	0.834	0.2	ZA	500
M_ν^{+++}	0.3175	0.049	0.6711	0.9624	0.834	0.4	ZA	500

$$L_\mathrm{box} = 1\ \mathrm{Gpc}/h$$

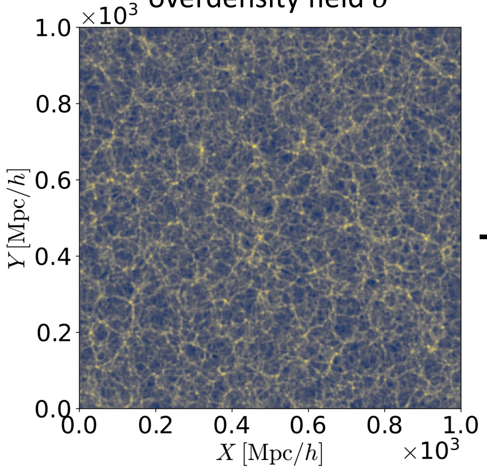
$$N_\mathrm{part} = 512^3\ \text{DM (and neutrinos if any)}$$

Discrete input $\boldsymbol{X}$



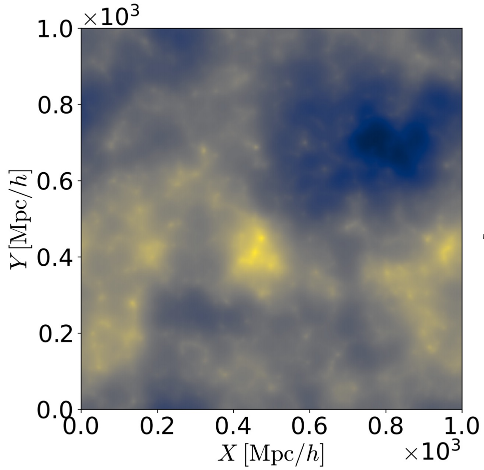
1 — B-spline interpolation of order 4

Continuous overdensity field δ



2 — Solve Poisson equation over the $N_g = 360^3$ cells $\Delta\Phi_r = \delta$

Reduced gravitational potential Φ_r



Physical classification from tidal tensor

T-Web formalism

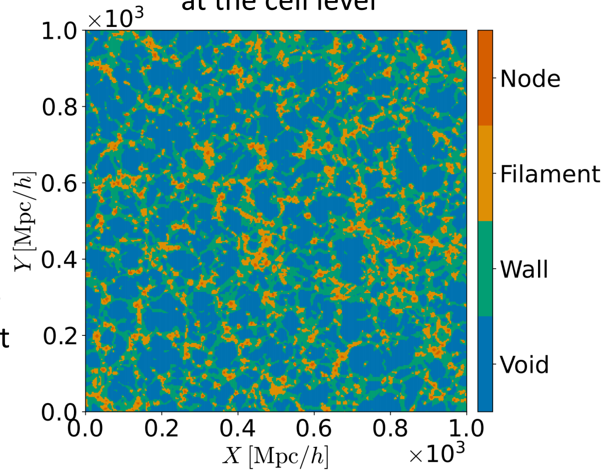
[Hahn+08, Forero-Romero+08]

Environment $\mathcal{E}(\mathbf{x})$	Condition
Void	$\lambda_1, \lambda_2, \lambda_3 < \lambda_{th}$
Wall	$\lambda_1, \lambda_2 < \lambda_{th}, \lambda_3 > \lambda_{th}$
Filament	$\lambda_1 < \lambda_{th}, \lambda_3, \lambda_2 > \lambda_{th}$
Node	$\lambda_1, \lambda_2, \lambda_3 > \lambda_{th}$

*Threshold λ_{th} and smoothing scale σ^2 are embedded as **nuisance parameters.**

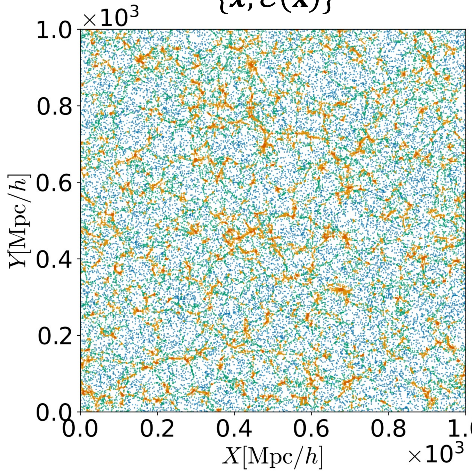
3

Cosmic web classification at the cell level



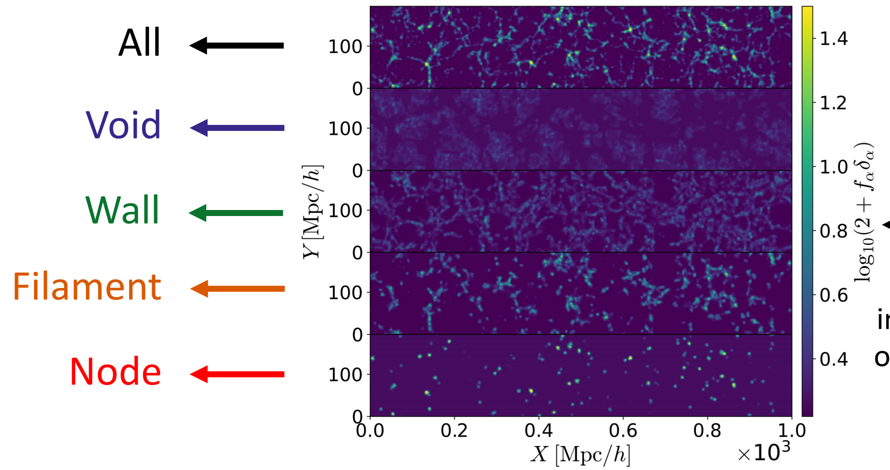
4 — Propagate the classification at the particle level

Particles classification $\{\boldsymbol{x}, \mathcal{E}(\mathbf{x})\}$



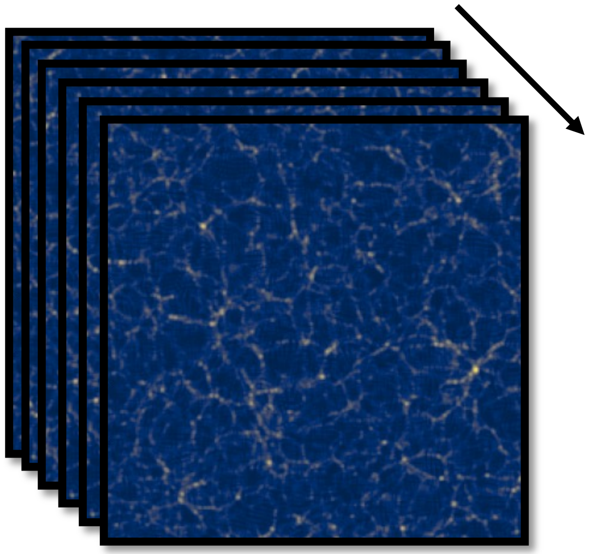
5 — B-spline interpolation of order 4 for each environment

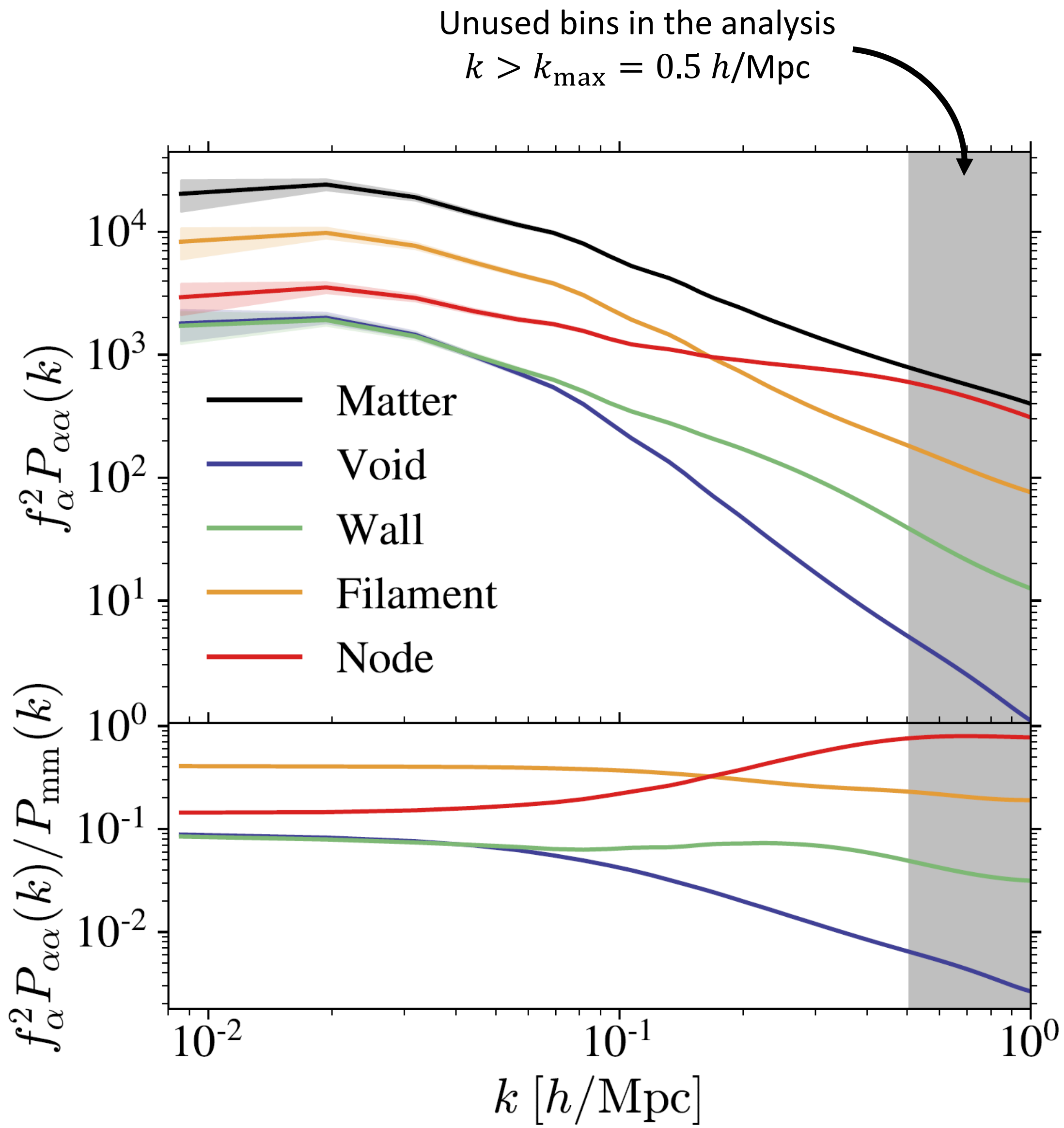
Overdensity fields in environments



Matter distribution in environments

7,000 realisations of fiducial cosmology from Quijote





Fisher forecast

The Fisher formalism allows to derive the (best possible) **marginalised errors on the parameters** based on two ingredients

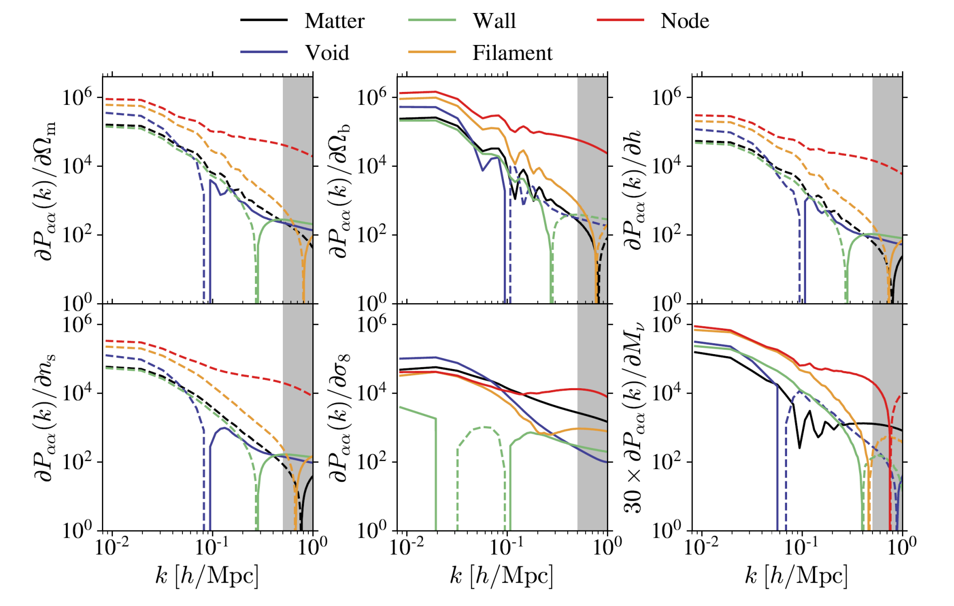
$$I(\boldsymbol{\theta})_{i,j} = \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta_i}\right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta_i}\right)$$

→ **Can be derived numerically with the Quijote simulations**

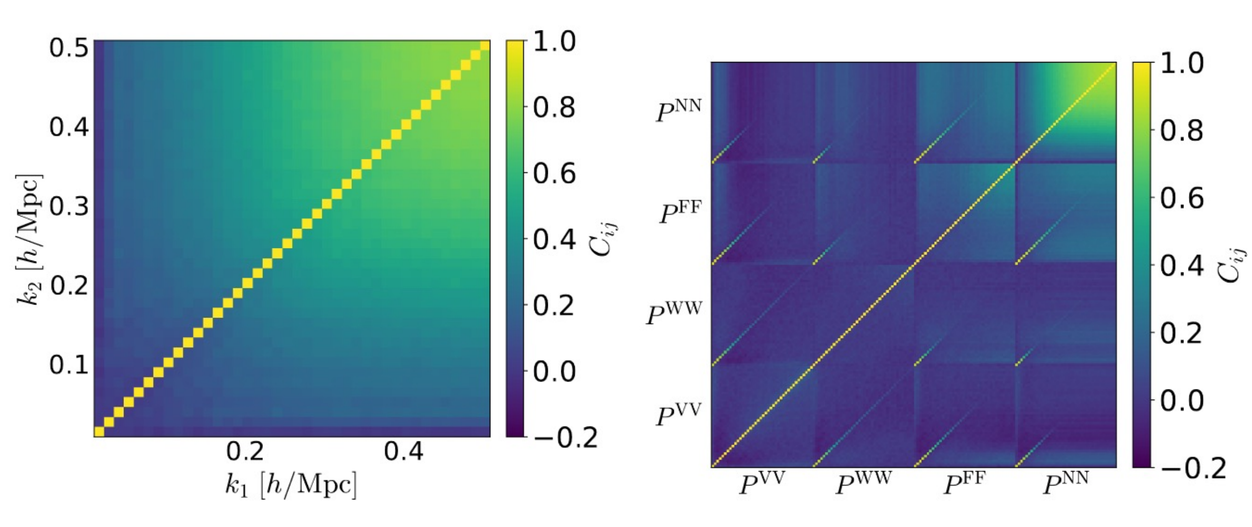
Derivatives

Covariance Σ

$N_{\mathrm{deriv}} = 500$



$N_{\mathrm{fid}} = 7000$



Fisher forecast in real-space

	Ω_m	Ω_b	h	n_s	σ_8	M_ν
Matter	0.0969	0.0413	0.5145	0.5019	0.0132	0.8749
Void	2.5	1.8	1.7	1.7	0.3	1.0
Wall	1.3	1.0	1.0	1.3	0.1	0.8
Filament	3.0	2.2	2.1	2.0	0.6	1.1
Node	1.0	0.9	0.8	0.8	0.1	0.5
Combination	**7.7**	**4.5**	**6.5**	**15.7**	**2.9**	**15.2**

Table of improvement factors

- **Individual environments are performing better** in some cases than the matter power spectrum
- In real space, the combination of auto-spectra in environments yields **2.9 to 15.7 improvement factors** over the matter power spectrum
- Some environments are providing **complementary information**

Fisher forecast in real-space

	Ω_m	Ω_b	h	n_s	σ_8	M_ν
Matter	0.0969	0.0413	0.5145	0.5019	0.0132	0.8749
Void	2.5	1.8	1.7	1.7	0.3	1.0
Wall	1.3	1.0	1.0	1.3	0.1	0.8
Filament	3.0	2.2	2.1	2.0	0.6	1.1
Node	1.0	0.9	0.8	0.8	0.1	0.5
Combination	**7.7**	**4.5**	**6.5**	**15.7**	**2.9**	**15.2**

Table of improvement factors

- **Individual environments are performing better** in some cases than the matter power spectrum
- In real space, the combination of auto-spectra in environments yields **2.9 to 15.7 improvement factors** over the matter power spectrum
- Some environments are providing **complementary information**

Fisher forecast in real-space

	Ω_m	Ω_b	h	n_s	σ_8	M_ν
Matter	0.0969	0.0413	0.5145	0.5019	0.0132	0.8749
Void	2.5	1.8	1.7	1.7	0.3	1.0
Wall	1.3	1.0	1.0	1.3	0.1	0.8
Filament	3.0	2.2	2.1	2.0	0.6	1.1
Node	1.0	0.9	0.8	0.8	0.1	0.5
Combination	**7.7**	**4.5**	**6.5**	**15.7**	**2.9**	**15.2**

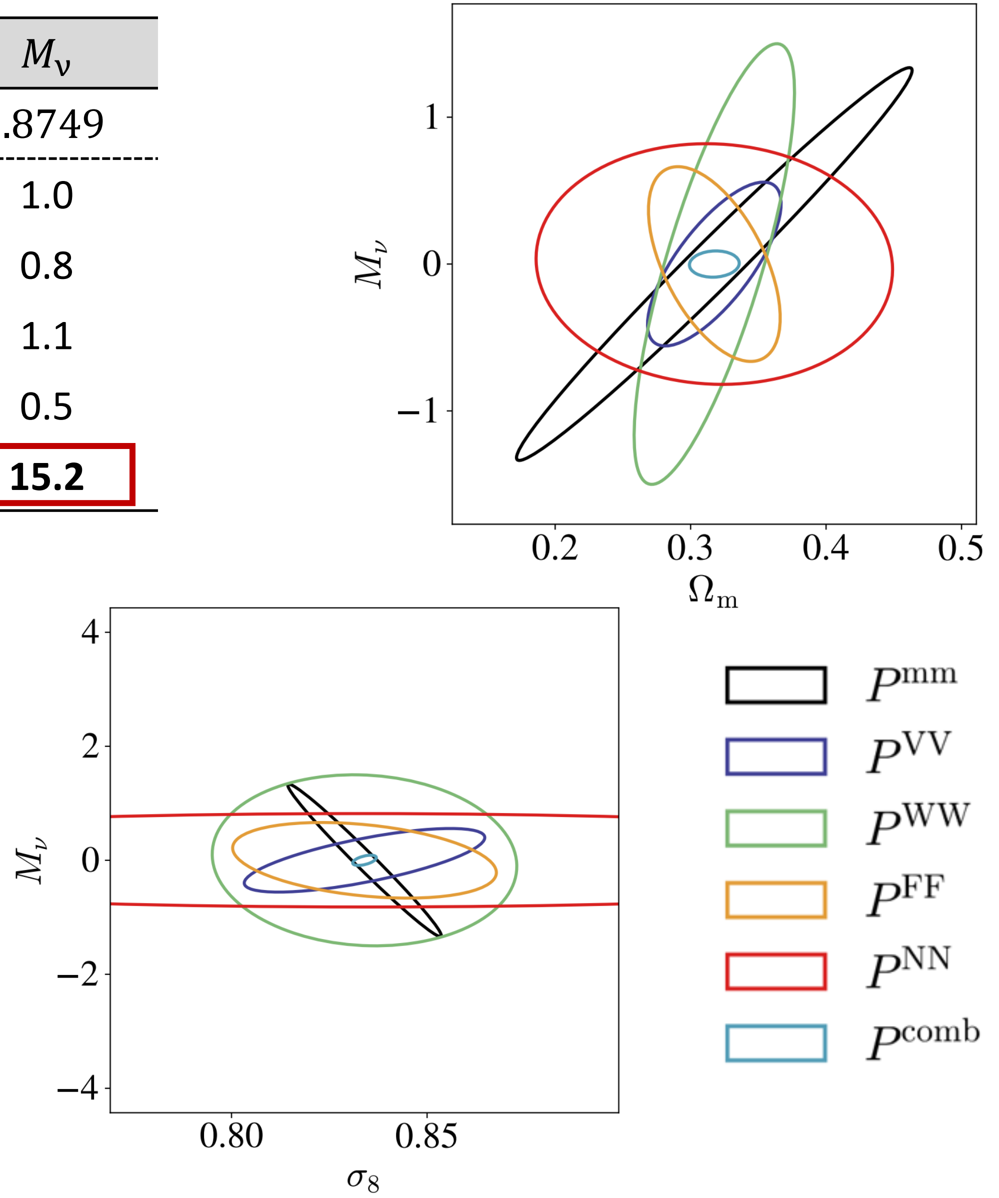
Table of improvement factors

- **Individual environments are performing better** in some cases than the matter power spectrum
- In real space, the combination of auto-spectra in environments yields **2.9 to 15.7 improvement factors** over the matter power spectrum
- Some environments are providing **complementary information**

Fisher forecast in real-space

	Ω_m	Ω_b	h	n_s	σ_8	M_ν
Matter	0.0969	0.0413	0.5145	0.5019	0.0132	0.8749
Void	2.5	1.8	1.7	1.7	0.3	1.0
Wall	1.3	1.0	1.0	1.3	0.1	0.8
Filament	3.0	2.2	2.1	2.0	0.6	1.1
Node	1.0	0.9	0.8	0.8	0.1	0.5
Combination	**7.7**	**4.5**	**6.5**	**15.7**	**2.9**	**15.2**

Table of improvement factors



- **Individual environments are performing better** in some cases than the matter power spectrum
- In real space, the combination of auto-spectra in environments yields **2.9 to 15.7 improvement factors** over the matter power spectrum
- Some environments are providing **complementary information**

A word on the classification parameters

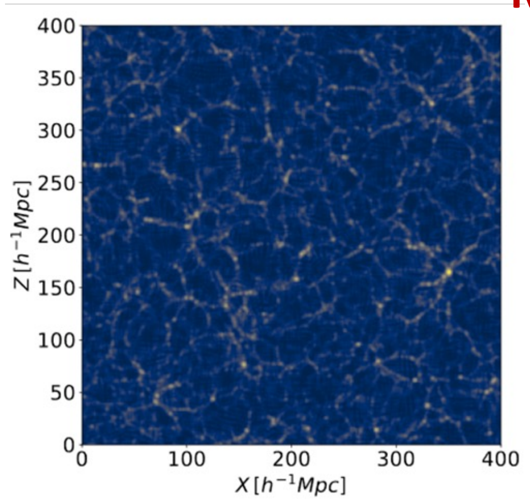
- Parameters of the classification (smoothing scale and eigenvalue threshold) = nuisance parameters
- They are well-constrained by the procedure and **have a limited impact on the obtained constraints**
- **Opens the possibility to apply different definitions and still obtain the same results!**

	Ω_{m}	Ω_{b}	h	n_{s}	σ_8	M_ν
Matter	0.0969	0.0413	0.5145	0.5019	0.0132	0.8749
Free Marginalised* over λ_{th} and σ_N	7.7	4.5	6.5	15.7	2.9	15.2
Fixed $\lambda_{\mathrm{th}} = 0.3$ and $\sigma_N = 2$ Mpc/h	7.9	4.5	6.6	16.4	7.2	24.3

Similar results (Ω_{m} to n_{s})

Most of the impact is on σ_8 and M_ν

*Derivatives taken with
$\lambda_{\mathrm{th}} = \{0.2, 0.3, 0.4\}$
$\sigma_N = \{1.5, 2, 2.5\}$ Mpc/h

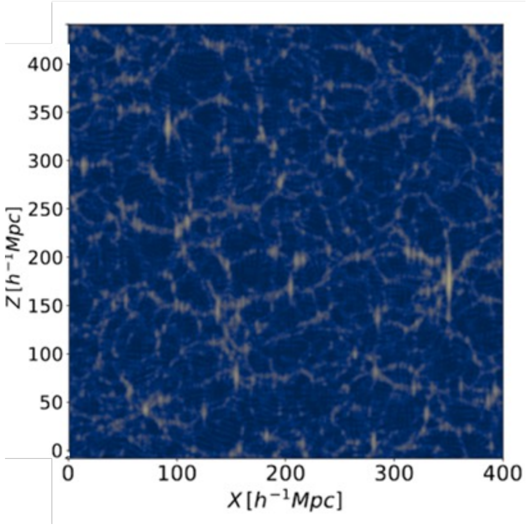


Real space

$$P^{\alpha\alpha}(k)\delta_{\mathrm{D}}(\boldsymbol{k}+\boldsymbol{k}') = \frac{1}{(2\pi)^3} < \tilde{\delta}^{\alpha}(\boldsymbol{k})\tilde{\delta}^{\alpha}(\boldsymbol{k}') >$$

	Ω_{m}	Ω_{b}	h	n_{s}	σ_8	M_ν
Matter	0.0969	0.0413	0.5145	0.5019	0.0132	0.8749

Up to an order of magnitude improvement for the matter power spectrum

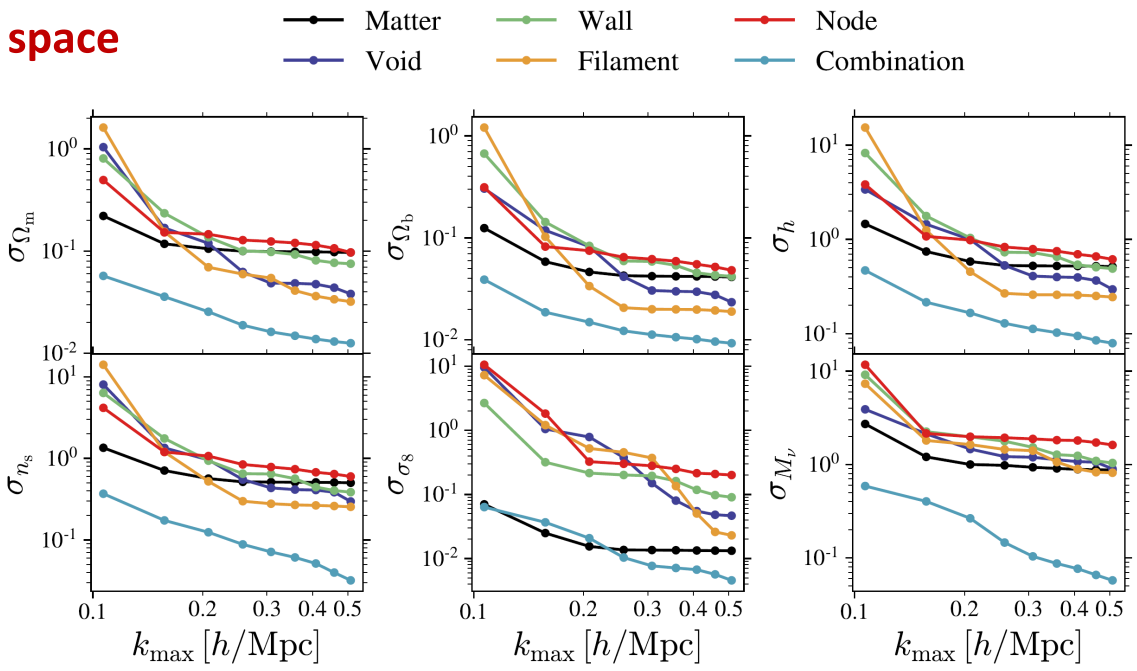


Redshift-space

$$P_\ell^{s,\alpha\alpha}(k) = \frac{2\ell+1}{2}\int_{-1}^{1} P^{s,\alpha\alpha}(k,\mu)\mathcal{L}_\ell(\mu)\mathrm{d}\mu$$

	Ω_{m}	Ω_{b}	h	n_{s}	σ_8	M_ν
Matter, $\ell = \{0,2\}$	0.0046	0.0133	0.1396	0.0719	0.0020	0.0834
Combined environments	1.7	1.4	1.8	2.4	1.0	2.7
Cross-spectra	2.3	1.6	2.2	2.9	1.1	3.4

Real space



- Quick **saturation** of the matter power spectrum information at scales $> 0.2\ h/\text{Mpc}$
- Combined environments => Improvement for all parameters seen at all considered scales

What is the theoretical cosmological information contained in the cosmic web environments?

Take-home messages:

- Splitting the particle set through the environments can bring **sizable gains** on constraints for all parameters **at all scales**
- Gains observed both in **real and redshift spaces**
- **Not too dependent** on the definition of the environments

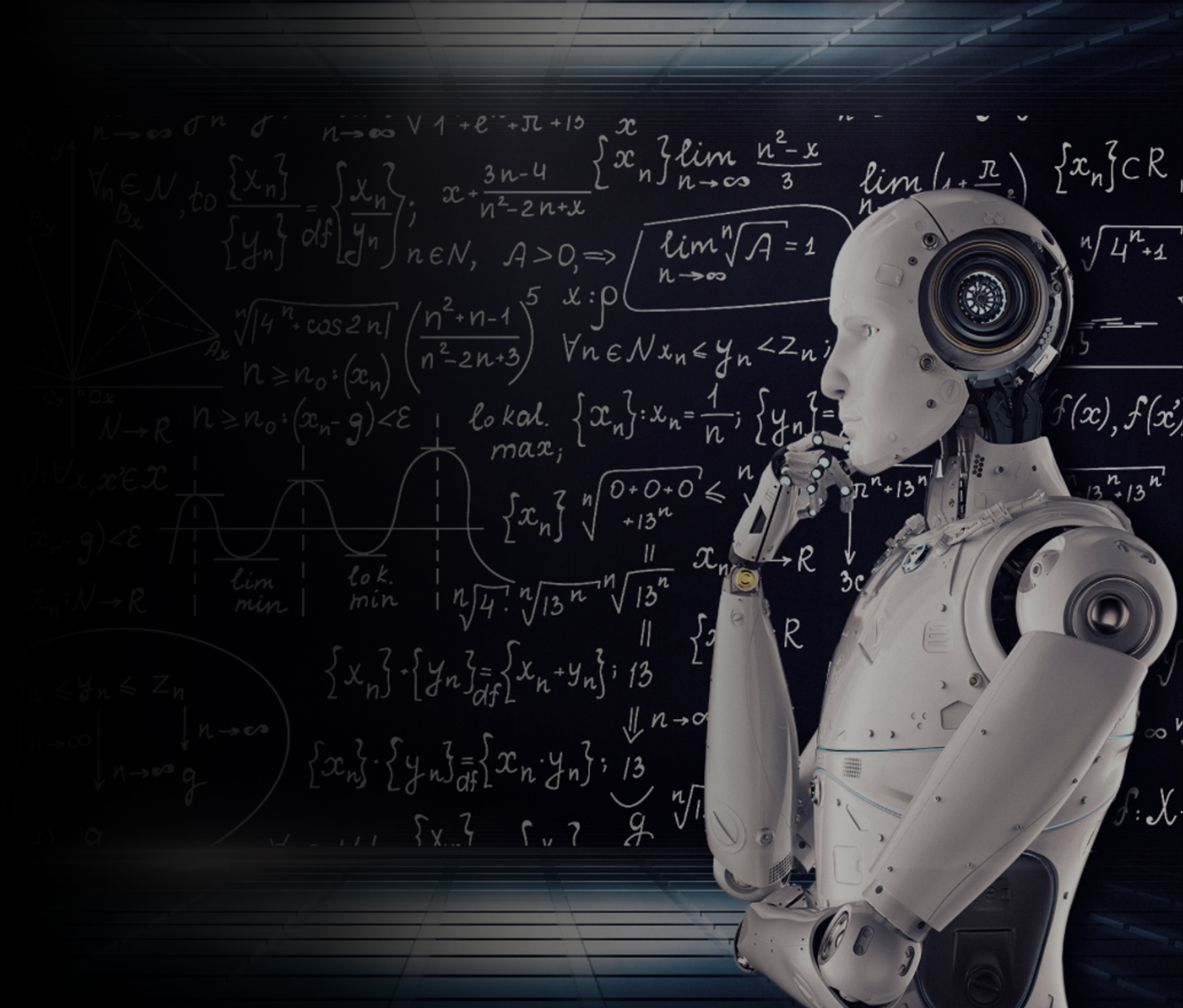
Some perspectives

All this was fun but... **The interesting questions start now**:

- What about matter tracers? (biases in the environments, mass threshold, shot noise, etc.)
- How to do cosmology with that? (build likelihood, accurate covariances, SBI, etc.)
- Basically, how to move from this idealised setup to a more realistic one?

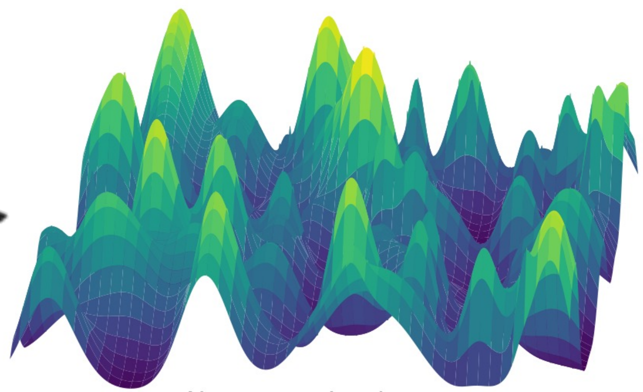
Ideas or suggestions? tony.bonnaire@ens.fr

Physics for Machine Learning



Non-convex optimization

Reasons of the success of gradient descent in high-dimensional and non-convex landscapes


Non-convex Landscape

Trivialization of the landscape?

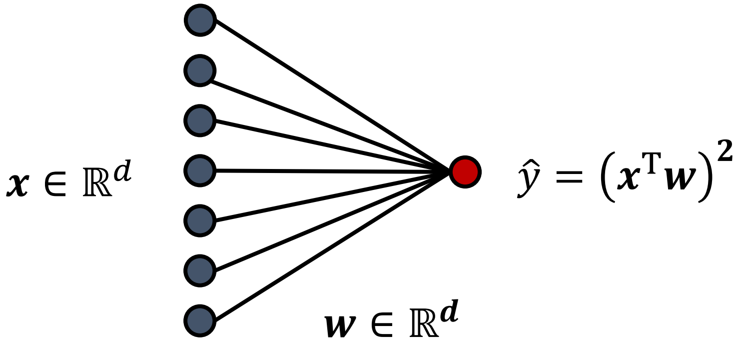
In which regime do we find an interesting minimum?

Statistical physics of spin-glasses:

- **Replica method**, dynamical mean field theory
- **Phase transitions** and finite size scaling analyses
- Kac-Rice analyses of the topology of the landscape

→ Requires the adaption and extension of these tools to data science (different energy functions, disorder is not Gaussian)

Phase retrieval: a prototypical example of single-layer neural network

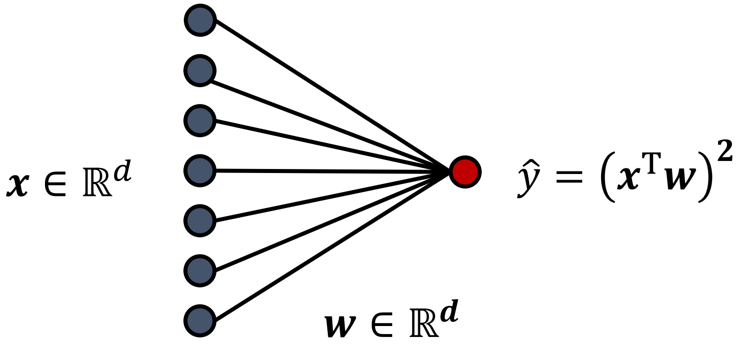


Setup:

- n Gaussian samples $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1}^n$
- Weights initialised randomly
- **Gradient descent** with fixed (vanishing) learning rate
- Thermodynamic limit: $d \to +\infty, n \to +\infty, \alpha = n/d \sim O(1)$
- Teacher-student: true labels comes from the same architecture with $\boldsymbol{w}_\star$

G. Biroli **C. Cammarota**

Phase retrieval: a prototypical example of single-layer neural network

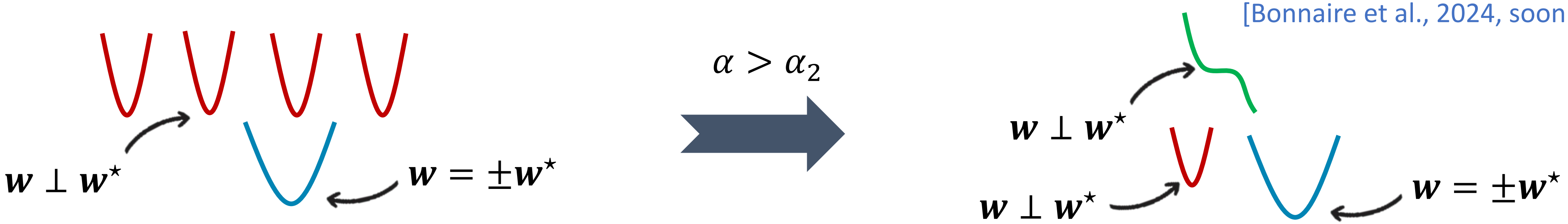


Setup:

- n Gaussian samples $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1}^n$
- Weights initialised randomly
- **Gradient descent** with fixed (vanishing) learning rate
- Thermodynamic limit: $d \to +\infty, n \to +\infty, \alpha = n/d \sim O(1)$
- Teacher-student: true labels comes from the same architecture with $\boldsymbol{w}_\star$

When $\boldsymbol{\alpha_1} < \boldsymbol{\alpha} < \boldsymbol{\alpha_2}$, the **initial curvature is informative** but gradient descent dynamics is stuck in **bad minima.**
When $\boldsymbol{\alpha} > \boldsymbol{\alpha_2}$, the « **bad minima** » hindering the recovery of the signal $\boldsymbol{w}^\star$ **turn into saddle-points** with exactly one direction pointing towards the **correct solution**.

[Bonnaire et al., 2024, soon on the arXiv!]

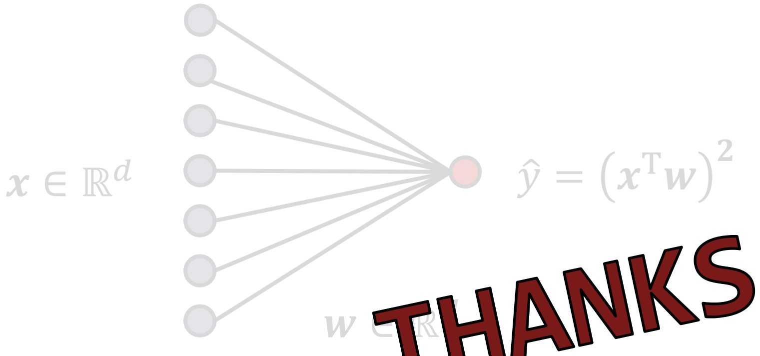


The phase retrieval case

G. Biroli 

C. Cammarota

Phase retrieval: a prototypical example of single-layer neural network

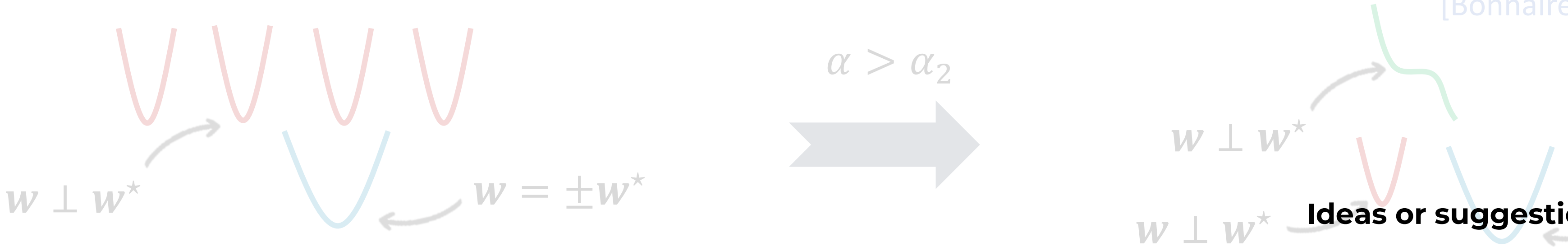


Setup:

- n Gaussian samples $X = \{x_i\}_{i=1}^n$
- Weights initialised randomly
- **Gradient descent** with fixed (vanishing) learning rate
- Thermodynamic limit: $d \to +\infty, n \to +\infty, \alpha = n/d \sim O(1)$
- Teacher-student: true labels comes from the same architecture with $w_\star$

When $\alpha_1 < \alpha < \alpha_2$, the **initial curvature is informative** but gradient descent dynamics is stuck in **bad minima.**
When $\alpha > \alpha_2$, the « **bad minima** » hindering the recovery of the signal $w^\star$ **turn into saddle-points** with exactly one direction pointing towards the **correct solution**.

[Bonnaire et al., 2024, soon on the arXiv!]



THANKS FOR YOUR ATTENTION!

Ideas or suggestions? ✉ **tony.bonnaire@ens.fr**