CONSTRAINED COSMOLOGICAL SIMULATIONS APPLIED TO HIGH-Z SPECTROSCOPIC SURVEYS

METIN ATA
(KAVLI IPMU →OSKAR KLEIN CENTRE, STOCKHOLM)

Collaborators: K.G. Lee (IPMU); O. Cucciati (INAF); B. Lamaux (U.C. Davis); D. Kashino (Nagoya U.) C. D-Vecchia; F. Kitaura (IAC), Thomas Müller (MPIA)

"FROM GALAXIES TO COSMOLOGY WITH DEEP SPECTROSCOPIC SURVEYS", 2022 JULY 6

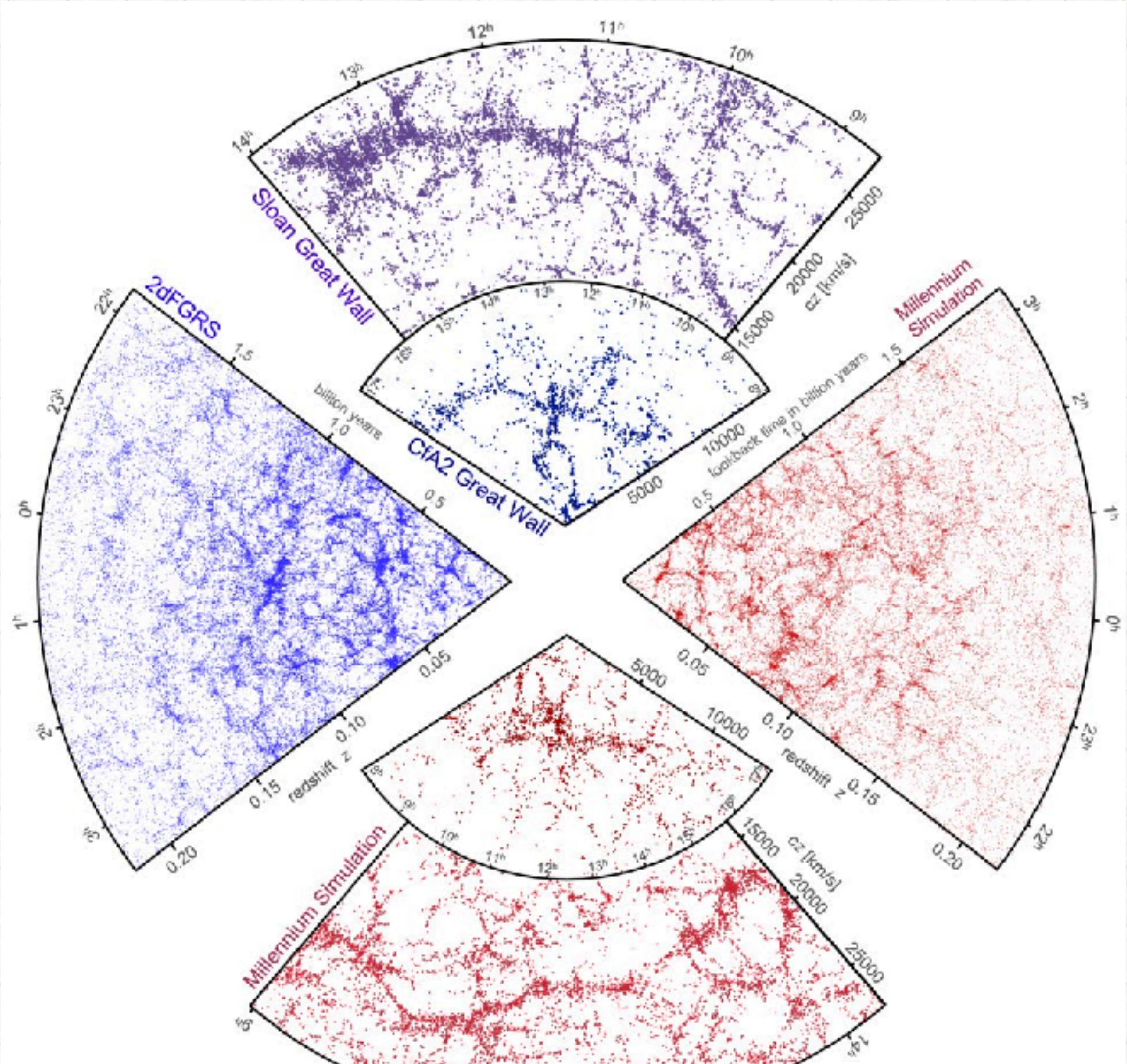
The University of Tokyo 東京大学

科研費 KAKENHI

KAVLI IPMU 東京大学 国際高等研究所 カブリ数物連携宇宙研究機構 KAVLI INSTITUTE FOR THE PHYSICS AND MATHEMATICS OF THE UNIVERSE

GALAXIES AS TRACERS OF LARGE-SCALE STRUCTURE

- Galaxies coalesce from the large-scale structure that manifests into a ‘cosmic web’ at ~Mpc scales
- The large-scale structure is a successful prediction of ΛCDM cosmology
- Question: How does the cosmic web influence galaxy formation and evolution across cosmic time?

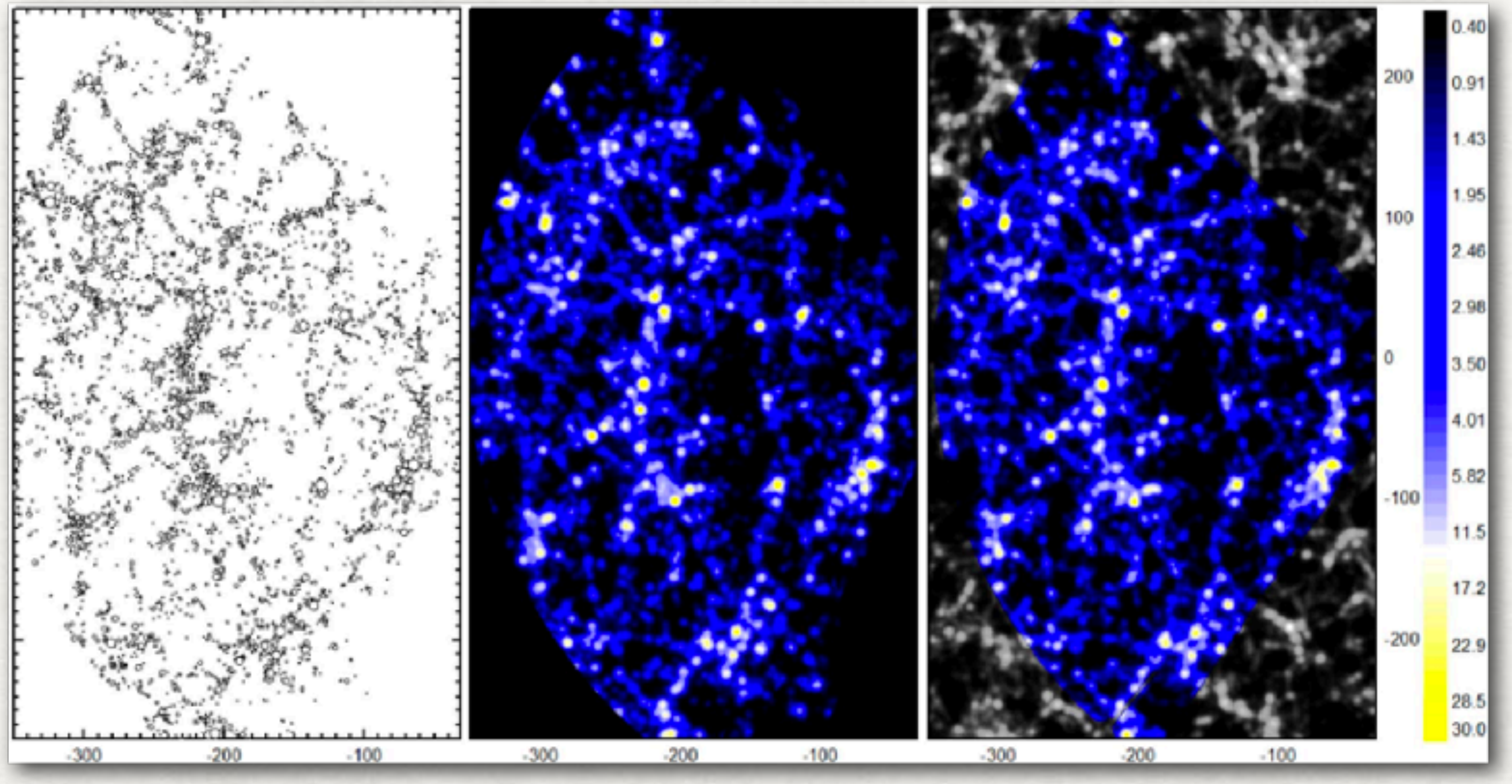


Credit: Millenium Simulation, CfA Redshift Survey, SDSS Survey, 2dF Survey

CONSTRAINED SIMULATIONS

- Density reconstruction: Construct the initial conditions so that matter density field is consistent with the distribution of tracers after forward modelling (typically using semi-numerical techniques, e.g. 2LPT)
  - Initial conditions: (nearly) Gaussian random fields at high redshifts (z~100) from which the large-scale structures/cosmic web emerged
  - Uncertainties (data & theoretical model) lead to an ensemble of correlated realisations that are consistent with the data
- Derived initial conditions can then be used to seed N-body simulations (or even hydro!)
- Application to SDSS Main Galaxy Survey (z~0.1) in the ELUCID Project (Wang+2014, 2016). See also Heß+2013 (2MASS)

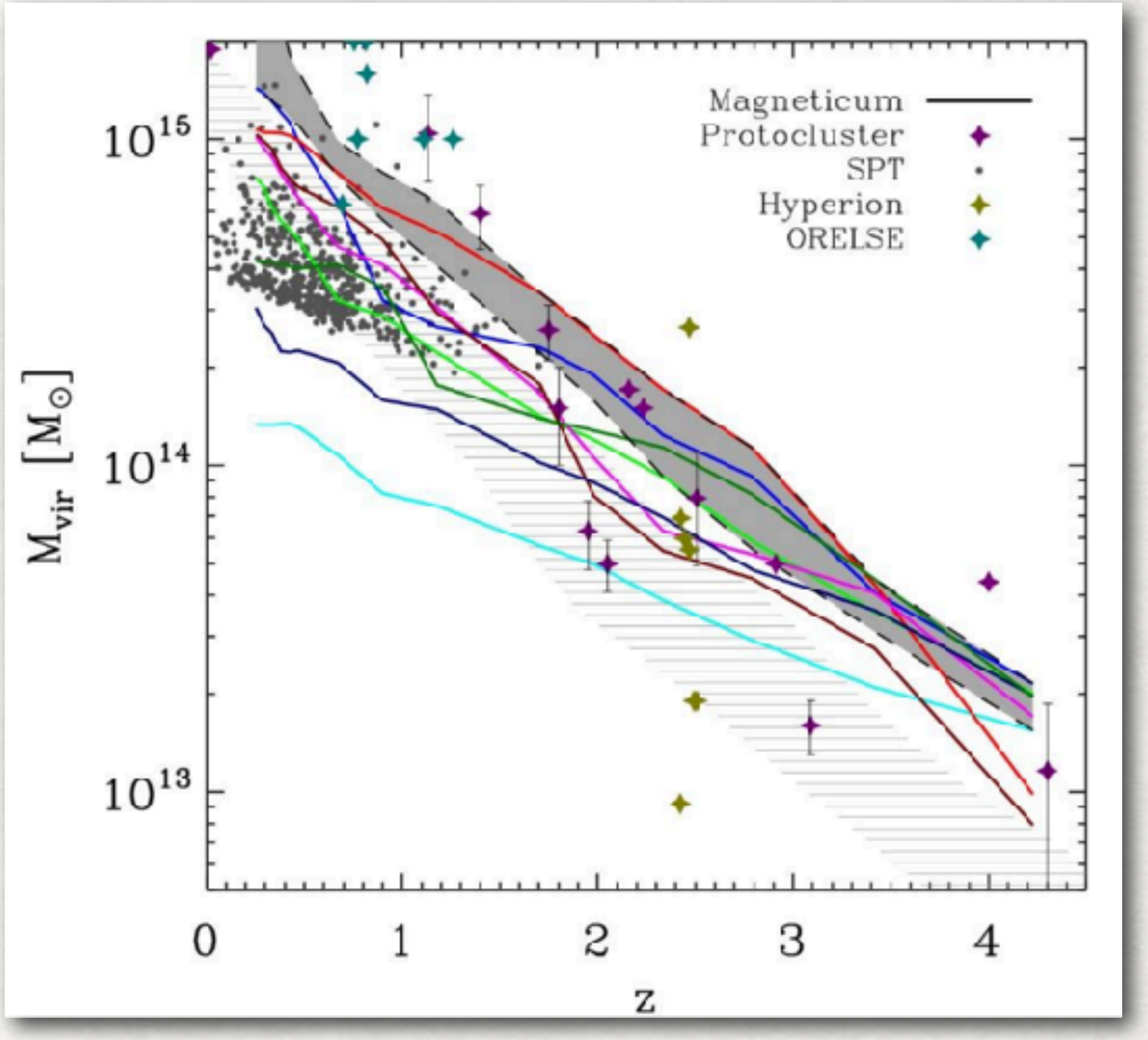
SDSS Data *Reconstructed Density* *N-body Sim*



Wang+2016

APPLICATION TO HIGH REDSHIFT GALAXY SURVEYS?

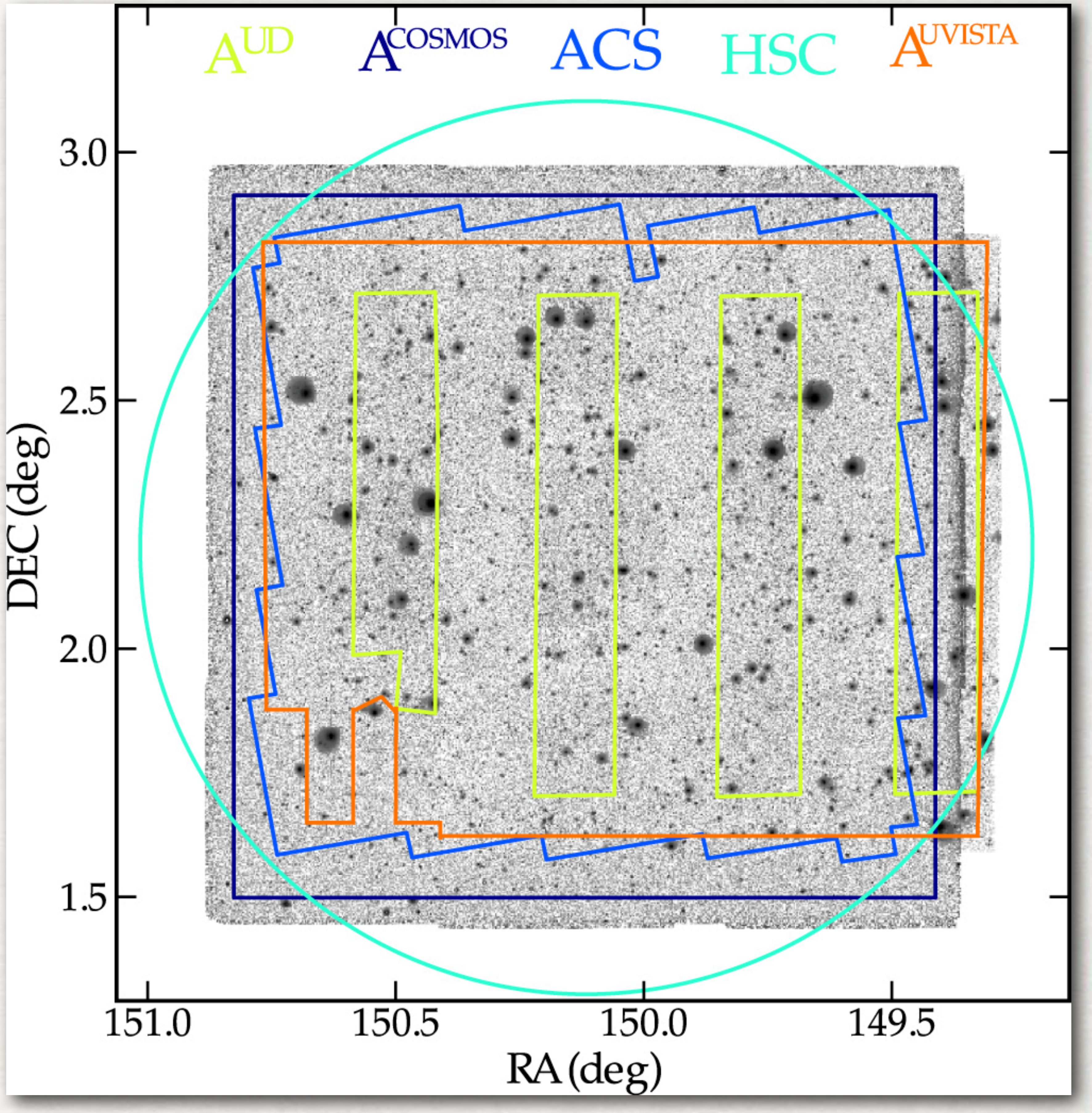
- Local Universe and near field galaxy surveys conducted with great success over the last 2 decades (2MRS, 6dFGS, GAMA, BOSS…)
- Next generation surveys (PFS, MOONS…) will cover high redshifts (z>1) with similar fidelity as the previous generations of low-redshift surveys
- Some advantages to applying constrained realization techniques to high redshift:
    - Gravitational evolution is still only mildly non-linear at z>1.5: shell-crossing has generally yet to occur
    - Directly provides matter density field as galaxy environment measure, with the observational selection function already applied
    - Can say interesting things about the growth of galaxy protoclusters into virialized clusters



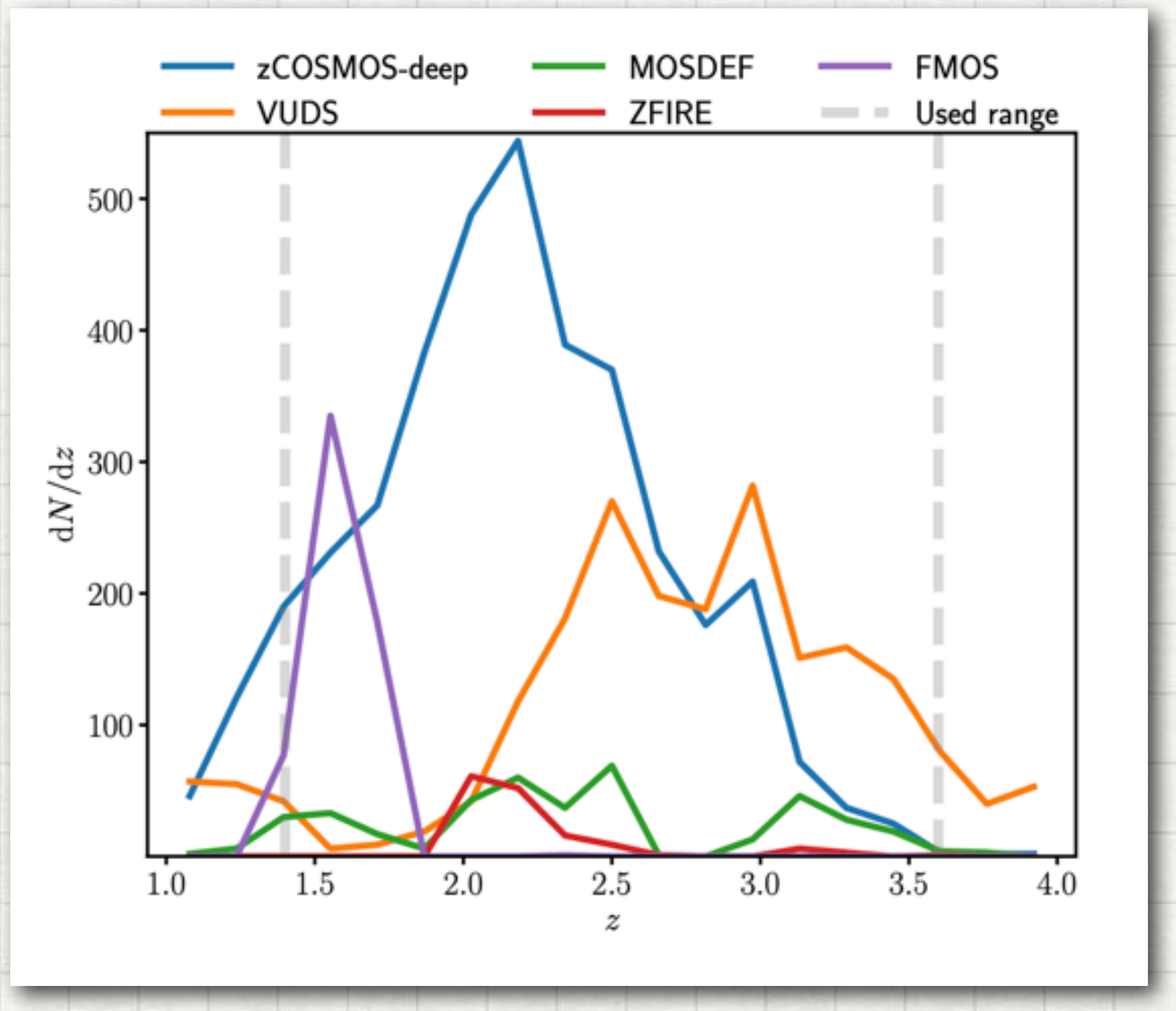
Curves: growth of galaxy protoclusters in Magneticum simulation; Credit: Rhea-Silvia Remus

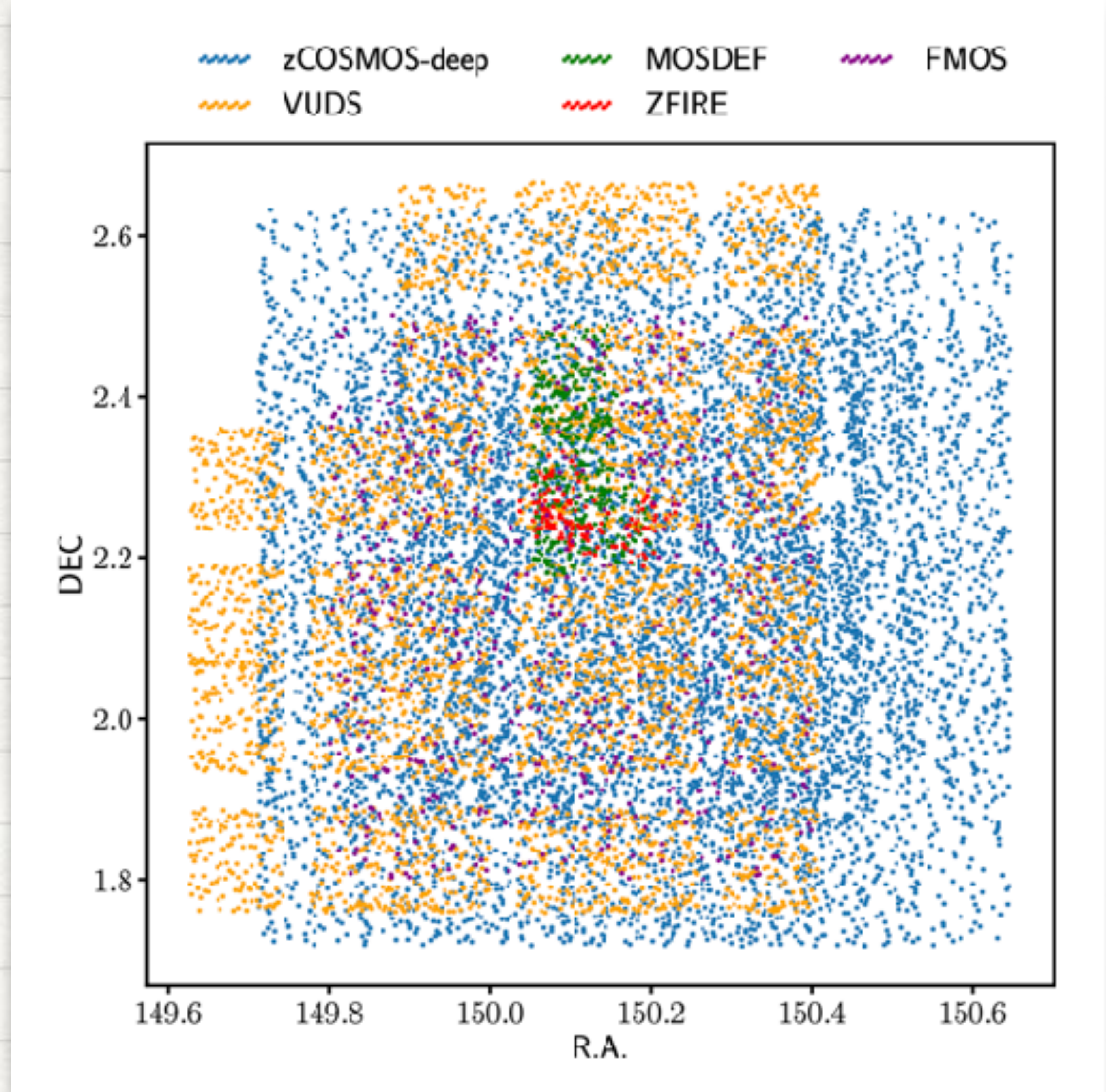
THE COSMOS FIELD

- COSMOS Survey (Capak+2007; Laigle+2016…)
- Photometric survey $(1 < z < 6)$ with 5×10^5 galaxies
- Covers the peak of star formation
- Ideal redshift range to to probe the cosmic web with “protoclusters”
- One of the most extensive spectroscopic campaigns in any high redshift deep field
- Angular footprint (~1 sq deg) also gives it sufficient transverse coverage (~100 cMpc @ z=2.5) to probe protoclusters scale (>20-30 cMpc)



SPECTROSCOPIC DATA





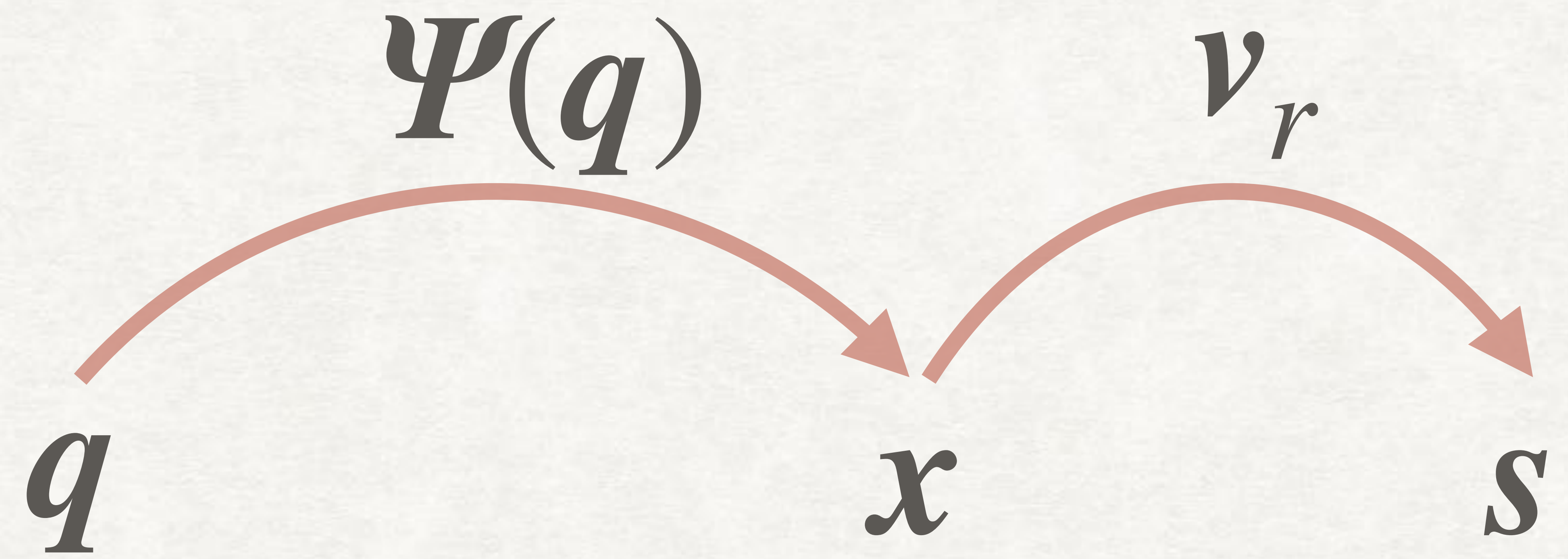
PART I: DENSITY RECONSTRUCTION WITH "COSMIC-BIRTH"

Ata et al 2021, MNRAS, 500, 3, 3194

- Mapping initial (Lagrangian) $\boldsymbol{q}$ to final (Eulerian) $\boldsymbol{x}$ positions using the cosmological displacement field $\boldsymbol{x} = \boldsymbol{\Psi}(\boldsymbol{q}) + \boldsymbol{q}$

- $\boldsymbol{\Psi}(\boldsymbol{q})$ is given by the forward model (X-LPT, Nbody solver, particle mesh…) and is a function of the initial field (e.g. Zel'dovich Approx. $\boldsymbol{\Psi} = \frac{d^3k}{(2\pi)^{3/2}} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{q}} \frac{i\boldsymbol{k}}{k^2} \hat{\delta}(\boldsymbol{k})$)

- We need to solve this problem iteratively

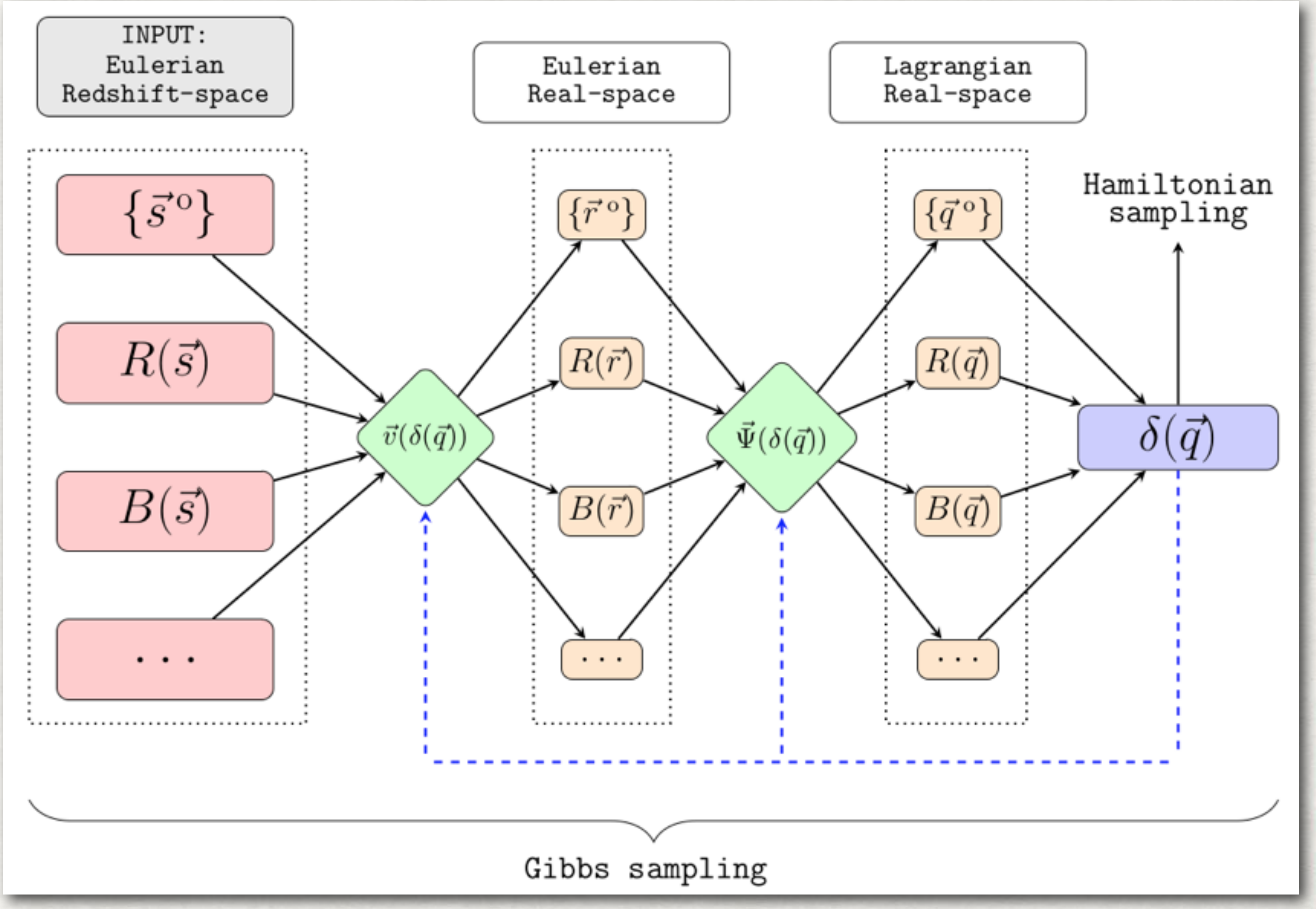


COSMIC BIRTH

- COSMIC BIRTH algorithm (Kitaura, Ata+2021, arXiv:1911.00284)

- Move galaxies back to initial positions and infer the Gaussian density field at Lagrangian frame $\delta(\boldsymbol{q})$ with Hamiltonian Monte-Carlo (HMC)

- Then use nested Gibbs sampling
  Peculiar velocities,
  Displacements,
  Response function of the survey
  Bias parameters

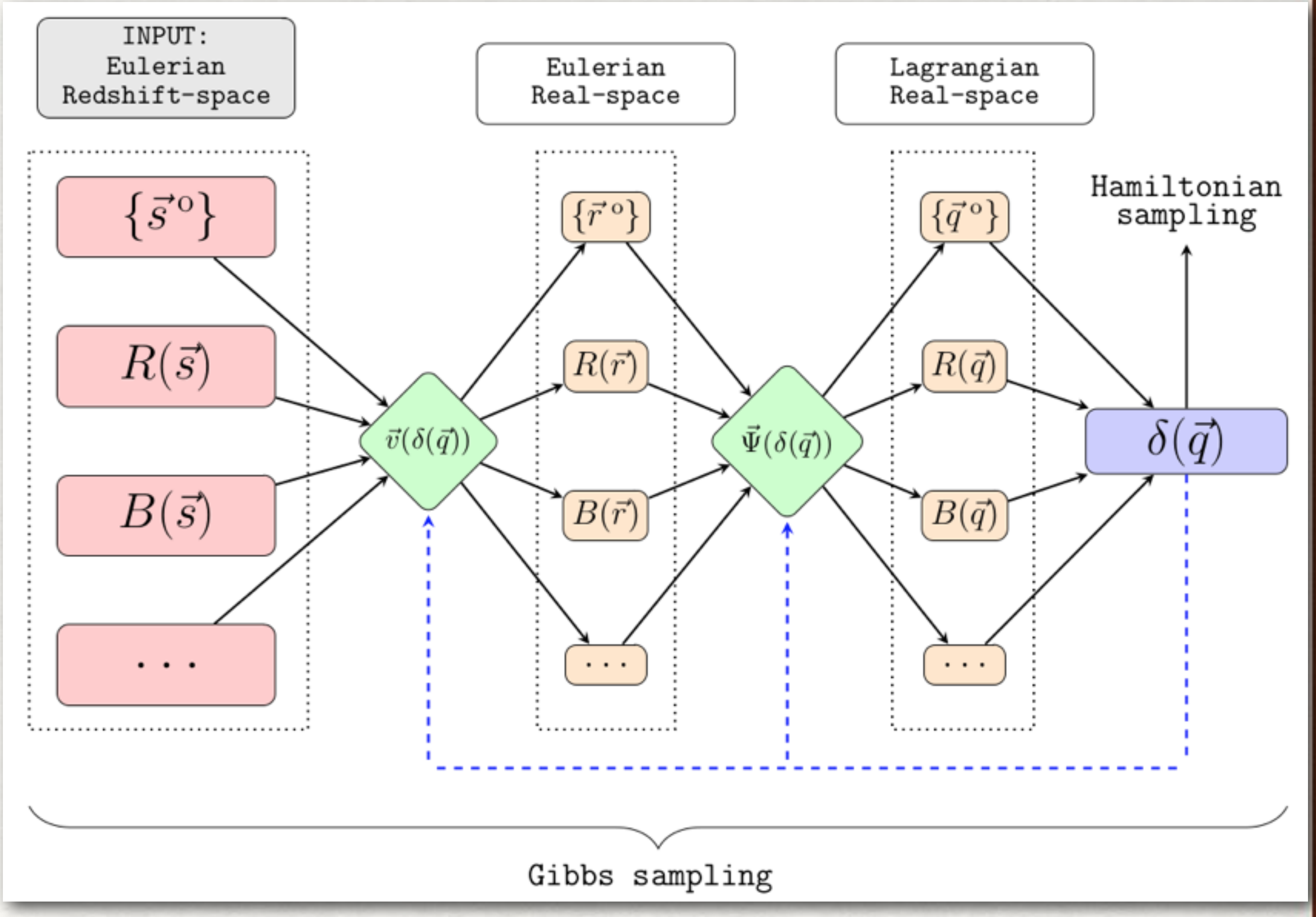


COSMIC BIRTH

- $\mathbf{R}$ is the survey response operator $\boldsymbol{\delta}_{\rm obs} = \mathbf{R}\,\boldsymbol{\delta}$

- $\mathbf{B}(\boldsymbol{\delta})$ is the bias function, connecting number of galaxies to DM per cell i $\langle \rho_{\rm Gal} \rangle_i = \bar{N} R_{ii} B(\delta_i)$: a nonlinear, nonlocal and stochastic function

- Real and redshift-space connected via peculiar velocity: $\boldsymbol{s} = \boldsymbol{r} + \dfrac{\boldsymbol{v}_{\rm r}}{H(a)\,a}$ $\quad \boldsymbol{v}_{\rm r} = (\boldsymbol{v} \cdot \boldsymbol{r})\boldsymbol{e}_r$



- $\boldsymbol{\psi}(\boldsymbol{q})$: displacement field using 2LPT

$$\boldsymbol{\psi} = -D_1 \nabla_q \phi^{(1)} + D_2 \nabla_q \phi^{(2)} \qquad \nabla_q^2 \phi^{(1,2)}(\boldsymbol{q}) = \delta^{(1,2)}(\boldsymbol{q}) \qquad \delta^{(2)}(\boldsymbol{q}) = \sum_{i>j} \left(\phi^{(1)}_{,ii}\phi^{(1)}_{,jj} - [\phi^{(1)}_{,ij}]^2\right)$$

DENSITY SAMPLING

- Step I: Assume tracers are in Lagrangian space $\boldsymbol{q}$ and sample density with HMC:

$$\delta(\boldsymbol{q}) \curvearrowleft \mathscr{P}(\delta(\boldsymbol{q})\,|\,\boldsymbol{q}, B(\boldsymbol{q}), \mathbf{R})$$

- Step II: Use the density field to infer Lagrangian real-space positions:

$$\boldsymbol{q} \curvearrowleft \mathscr{P}(\boldsymbol{q}\,|\,\delta(\boldsymbol{q}), \boldsymbol{s}, \boldsymbol{\psi})$$

DENSITY SAMPLING

- Use Bayesian approach to define posterior of parameter Θ given the data d: $\mathscr{P}(\boldsymbol{\Theta}|\boldsymbol{d}) \propto \pi(\boldsymbol{\Theta}) \times \mathscr{L}(\boldsymbol{d}|\boldsymbol{\Theta})$, where data d is the galaxy counts per cell and parameters Θ are density values
- $$\pi(\boldsymbol{\delta}_{\mathrm{L}}) = \frac{1}{\sqrt{(2\pi)^{N_c}\det(\mathbf{C}_{\mathrm{L}})}} \exp\left(-\frac{1}{2}\boldsymbol{\delta}_{\mathrm{L}}^{\mathrm{T}}\mathbf{C}_{\mathrm{L}}^{-1}\boldsymbol{\delta}_{\mathrm{L}}\right) \quad \boldsymbol{\delta}_{\mathrm{L}} = \log(1+\boldsymbol{\delta}) - \boldsymbol{\mu}\,,$$
- $\mathscr{L}(\boldsymbol{N}_{\mathrm{G}}|\boldsymbol{\lambda}) = \prod_i \frac{\lambda_i^{N_{\mathrm{G}i}} \mathrm{e}^{-\lambda_i}}{N_{\mathrm{G}i}!}$ in each cell i

HAMILTONIAN MONTE CARLO

- We use Hamilton dynamics for our probabilistic system $\mathscr{H}(\boldsymbol{q},\boldsymbol{p}) = \mathscr{U}(\boldsymbol{q}) + \mathscr{K}(\boldsymbol{p})$
- Canonical posterior probability defined as: $\mathscr{P}(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{Z}\,\mathrm{e}^{-\mathscr{H}(\boldsymbol{q},\boldsymbol{p})}$
- Kinetic term artificial to calculate Hamiltonian EoM: $\mathscr{P}(\boldsymbol{p}) \propto \mathrm{e}^{-\mathscr{K}} = \mathrm{e}^{-\frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{p}}$
- Variable of interest $\boldsymbol{q}$ in potential: $-\ln \mathscr{P}(\boldsymbol{q}) = \mathscr{U}(\boldsymbol{q})$
- Link potential term to Bayesian posterior $-\ln \mathscr{P}(\boldsymbol{\delta}_L \,|\, \boldsymbol{N}_G) \propto -\ln \pi(\boldsymbol{\delta}_L) - \ln \mathscr{L}(\boldsymbol{N}_G \,|\, \boldsymbol{\delta}_L)$

HAMILTON EQUATIONS OF MOTION

- Evolve HMC system according to Hamiltonian mechanics:

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{dt}} = \frac{\partial \mathscr{H}(\boldsymbol{q},\boldsymbol{p})}{\partial \boldsymbol{p}} = \mathbf{M}^{-1}\boldsymbol{p} \qquad \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{dt}} = -\frac{\partial \mathscr{H}(\boldsymbol{q},\boldsymbol{p})}{\partial \boldsymbol{q}} = -\frac{\partial \mathscr{U}(\boldsymbol{q})}{\partial \boldsymbol{q}}$$

- For a pseudo time step $\Delta t = \epsilon$ we use 2nd order: $\mathscr{T}_2(\epsilon) = \mathscr{T}_p(\epsilon/2)\mathscr{T}_q(\epsilon)\mathscr{T}_p(\epsilon/2)$

$$\boldsymbol{p}\left(t + \frac{\epsilon}{2}\right) = \boldsymbol{p}(t) - \frac{\epsilon}{2}\frac{\partial \mathscr{U}}{\partial \boldsymbol{q}}(\boldsymbol{q}(t)),$$

$$\boldsymbol{q}(t+\epsilon) = \boldsymbol{q}(t) + \epsilon\,\mathbf{M}^{-1}\boldsymbol{p}\left(t + \frac{\epsilon}{2}\right),$$

$$\boldsymbol{p}(t+\epsilon) = \boldsymbol{p}\left(t + \frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\frac{\partial \mathscr{U}}{\partial \boldsymbol{q}}(\boldsymbol{q}(t+\epsilon))$$

HAMILTON EQUATIONS OF MOTION

- Evolve HMC system according to Hamiltonian mechanics:

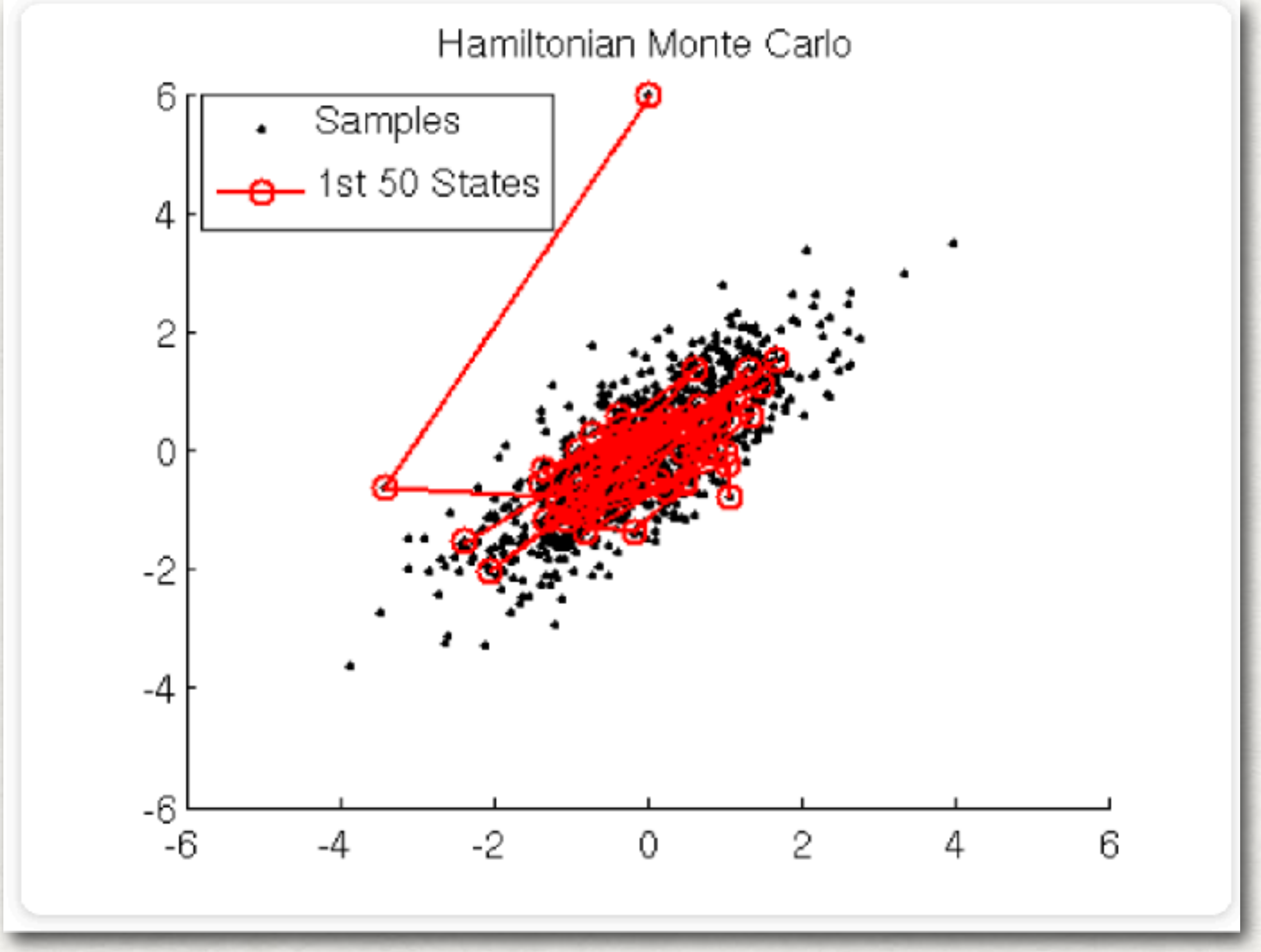
$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{dt}} = \frac{\partial \mathscr{H}(\boldsymbol{q},\boldsymbol{p})}{\partial \boldsymbol{p}} = \mathbf{M}^{-1}\boldsymbol{p} \qquad \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{dt}} = -\frac{\partial \mathscr{H}(\boldsymbol{q},\boldsymbol{p})}{\partial \boldsymbol{q}} = -\frac{\partial \mathscr{U}(\boldsymbol{q})}{\partial \boldsymbol{q}}$$

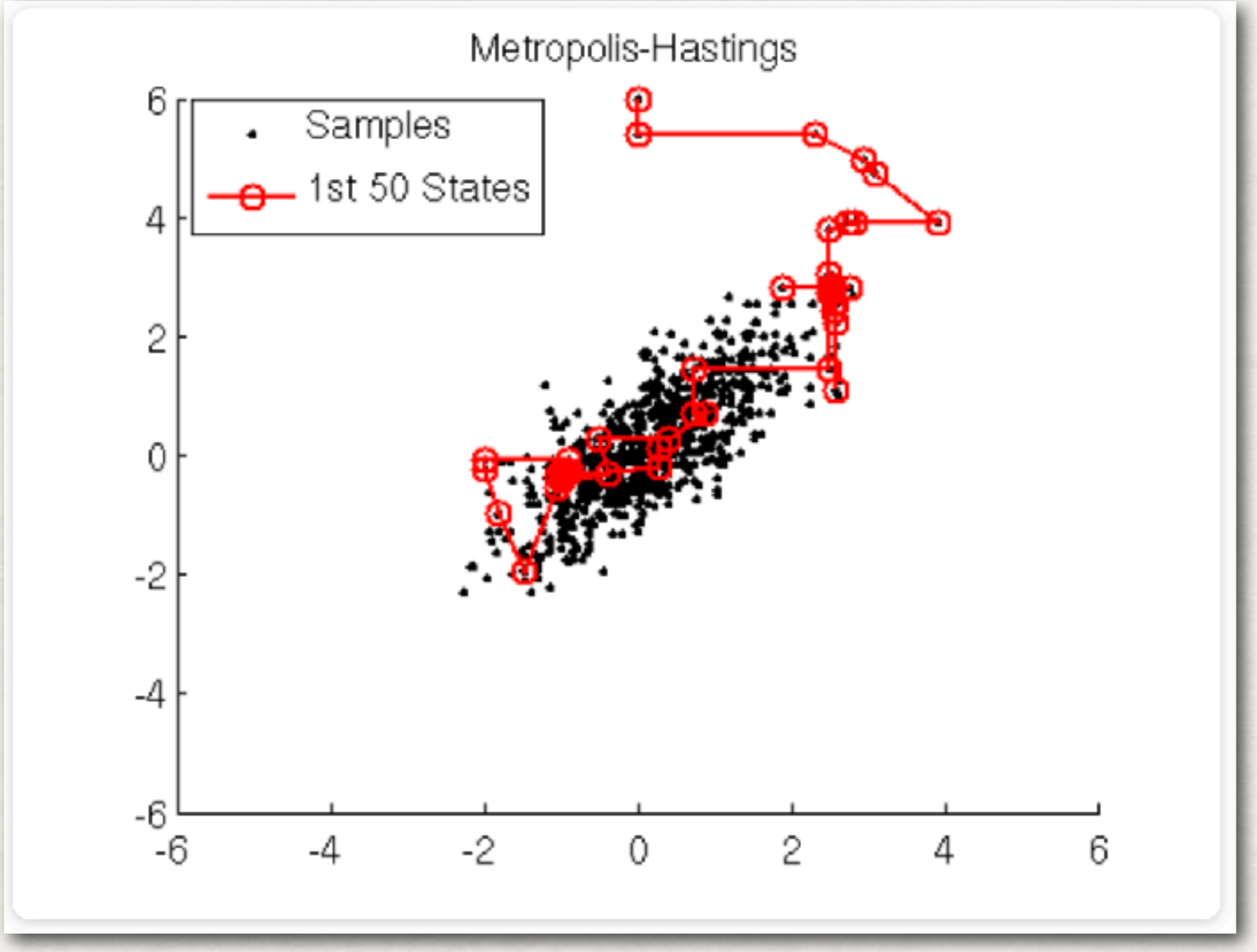
- Second Order: $\mathscr{T}_2(\epsilon) = \mathscr{T}_p(\epsilon/2)\mathscr{T}_q(\epsilon)\mathscr{T}_p(\epsilon/2)$

- Higher even order achieved by recursive application of second order leap-frog:
  $\mathscr{T}_{n+2}((2i-s)\epsilon) = \mathscr{T}_n^{\,i}(\epsilon)\mathscr{T}_n(-s\epsilon)\mathscr{T}_n^{\,i}(\epsilon)$, with:
  $\Delta\tau_{4\mathrm{th}} = (2i - s(n=2))\,\epsilon^{\mathrm{eff}} = (2i-(2i)^{1/(3)})\,\epsilon^{\mathrm{eff}}$

- i is a free parameter of the scheme, we chose $i = 4$

The dimensionality of the parameter space is in principle *every single grid cell* in the observed volume!





https://theclevermachine.wordpress.com/

PART II: CONSTRAINED SIMULATIONS

Ata et al 2022, Nature Astronomy, arXiv:2206.01115

- Environment of ~10 Mpc scales is decisive for the fate of a protocluster
- Previous works analysed singular protoclusters with simple methods (e.g. looking for analogous overdensities in N-body sims, simple analytic spherical overdensity estimate)

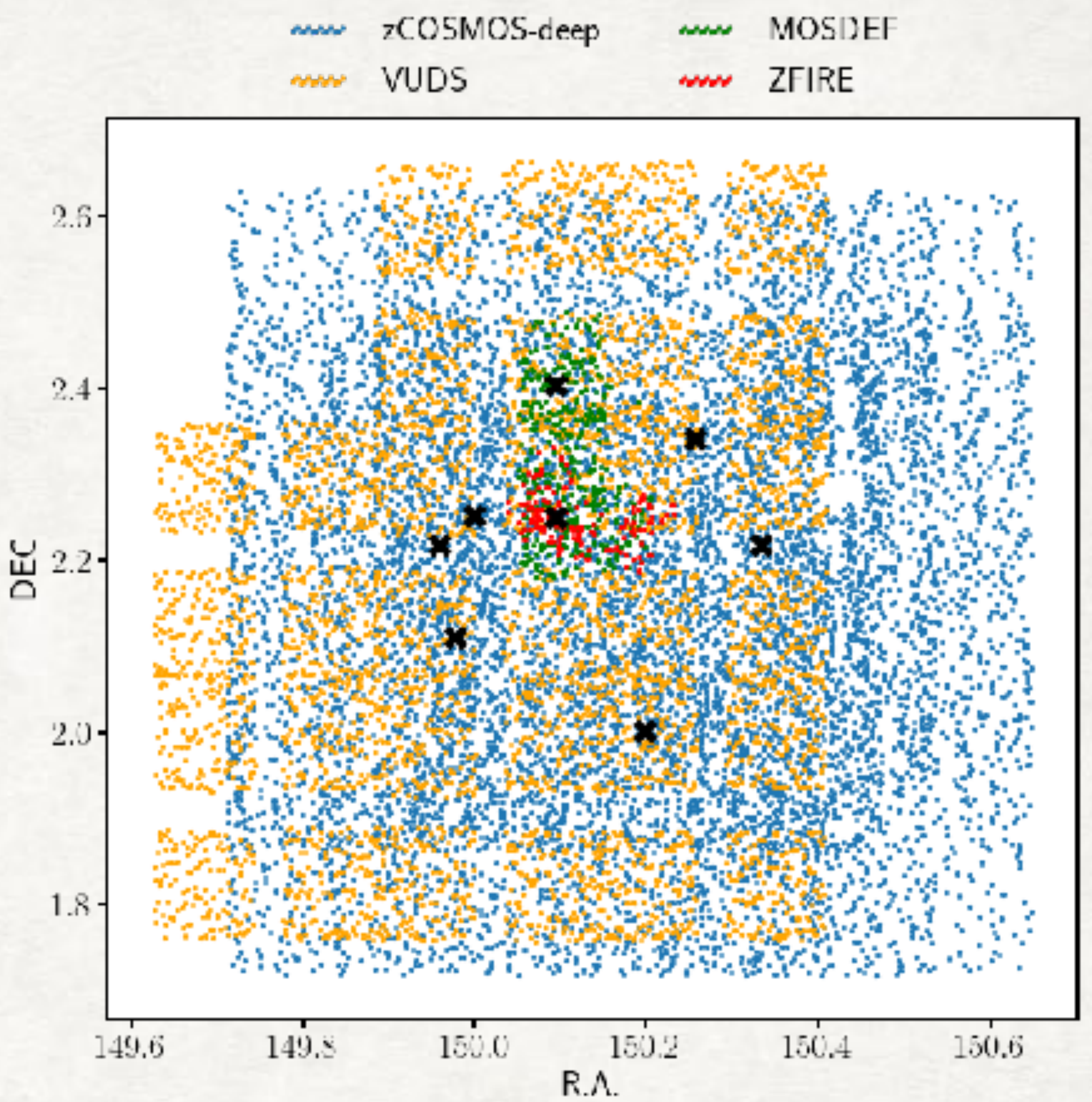
- At the right time,
- At the right position,
- Forms the right densities

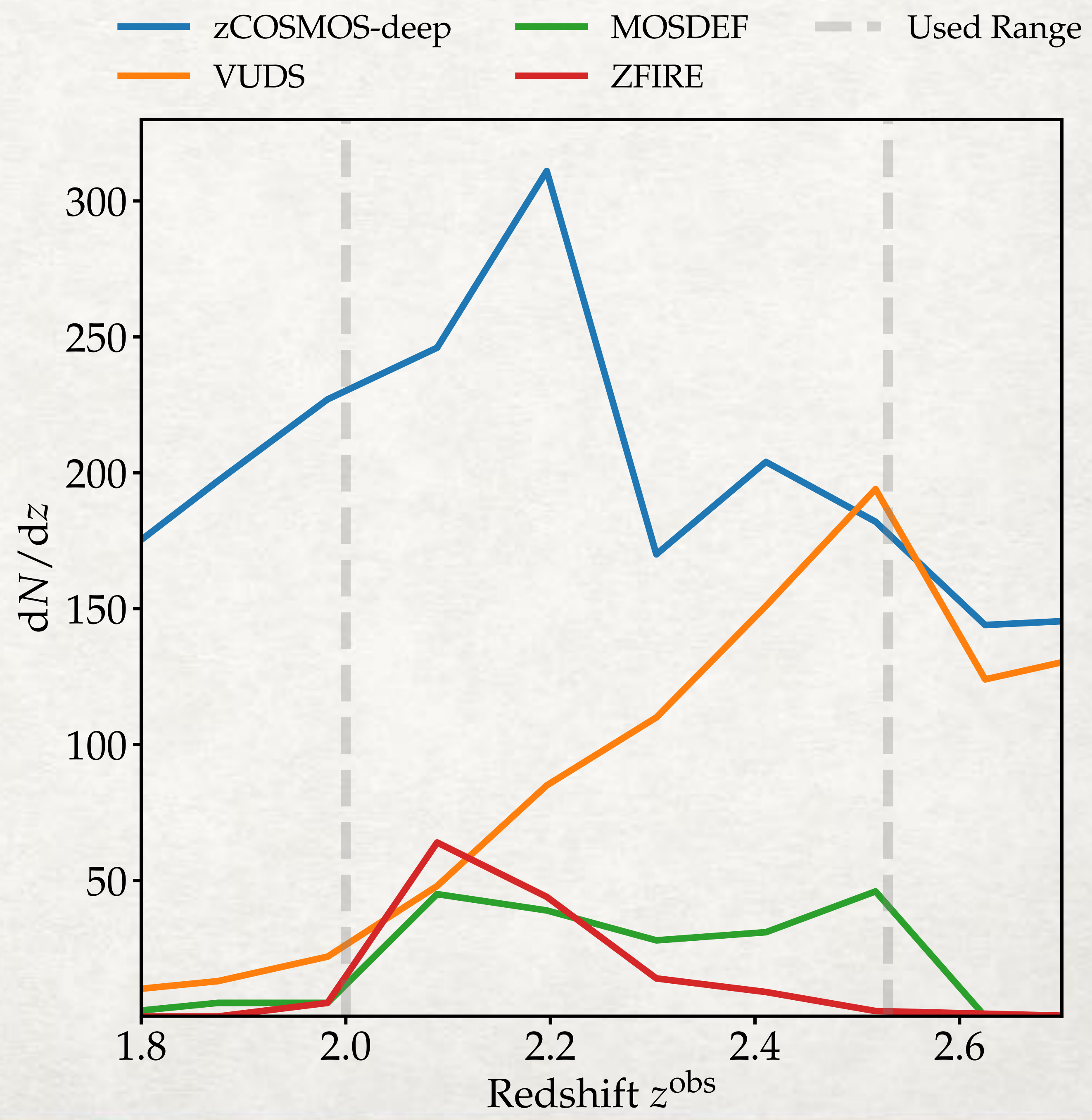
} **Constrained Simulations**

WELCOME TO COSTCO

(**CO**NSTRAINED **S**IMULATIONS OF **T**HE **CO**SMOS FIELD)

- COSTCO is a suite of 50 cosmological simulations run from HMC realisations of the COSMOS initial conditions $2 \leq z \leq 2.5$
- Cubic volume of 512 Mpc/h side length and 2 Mpc/h resolution —> $M_{\rm res} \sim 7 \times 10^{11} M_{\odot}/h$
- Identify massive halos at the $z = 0$ snapshot and trace back the member DM particles of each halo, defining the protocluster
- Match simulations to observations in position and mass

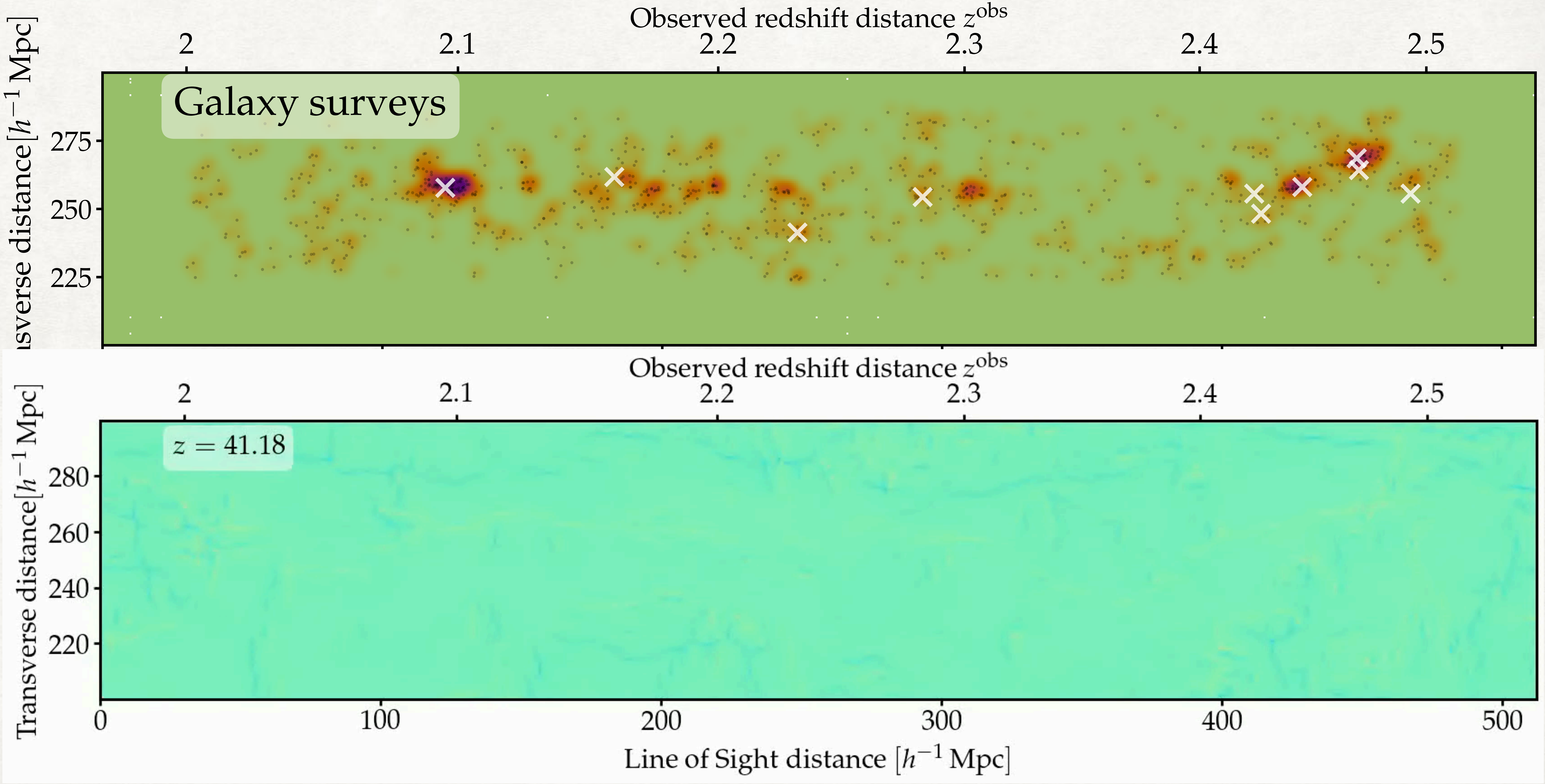




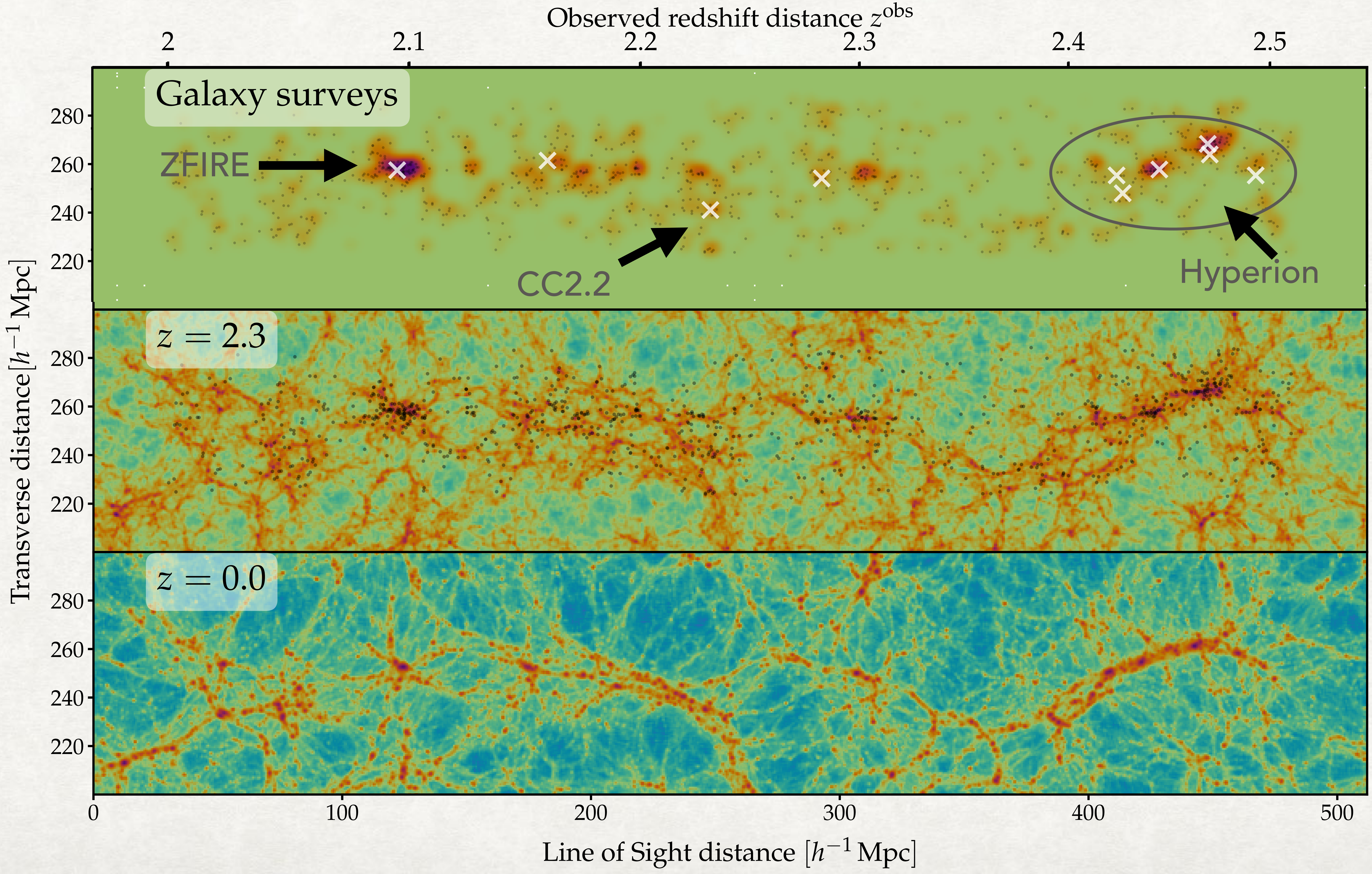
CONSTRAINED SIMULATIONS OF COSMOS FIELD

- ~50000 MCMC samples of z=100 initial conditions produced by COSMIC BIRTH
- 50 random selected (thinning, autocorrelation length…) samples for simulations
- 256^3 particles in a box with 512 Mpc/h side length (~7E11 Msun/h resolution)
- Simulations run from z=100 to z=0
- We write snapshots at z=2.5-2.0 in $z = 2 - 2.5 \quad \Delta z = 0.1$ and $z = 0$
- Consistently used Planck 2018 cosmology from Initial Conditions to final analysis

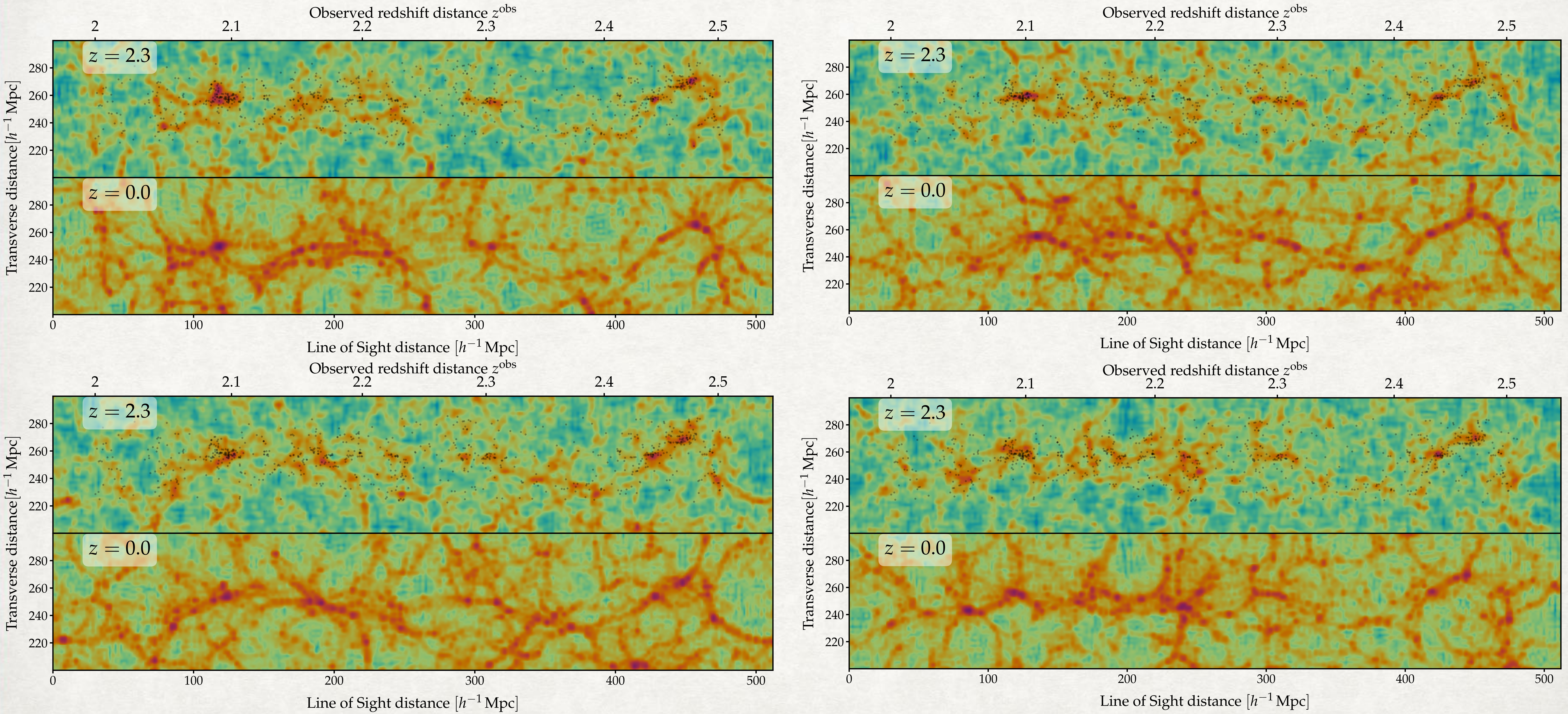
COSTCO TIME EVOLUTION



COSTCO TIME EVOLUTION



COSTCO REALISATIONS



MATCH KNOWN PROTOCLUSTERS IN COSTCO

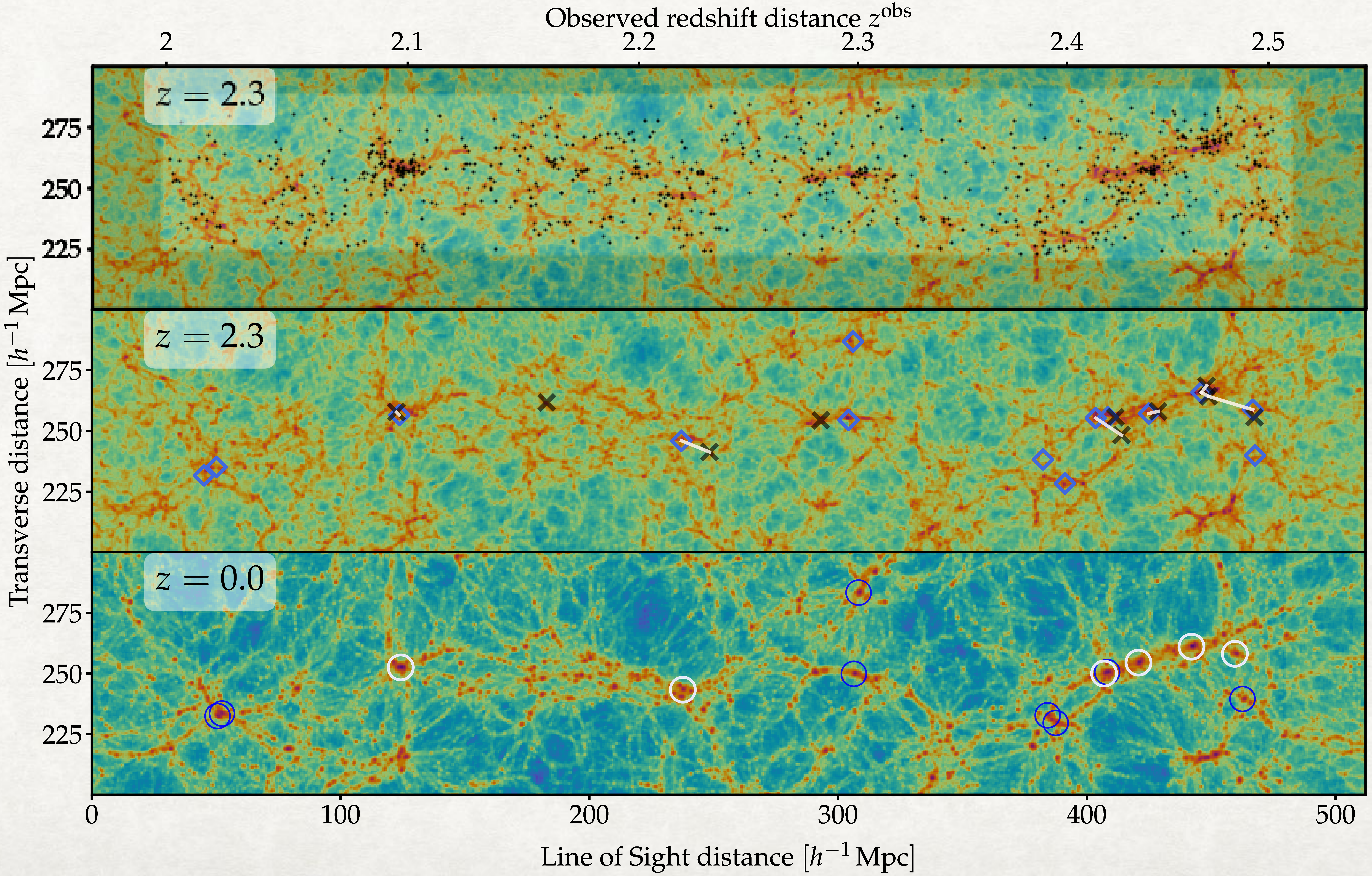
1. Identify massive halos at the $z = 0$ snapshot with $M_{\rm vir} \geq 2 \times 10^{14}\, h^{-1} M_{\odot}$

2. Trace the member particles back to $z = 2.3$ and compute center of mass position of their Lagrangian region

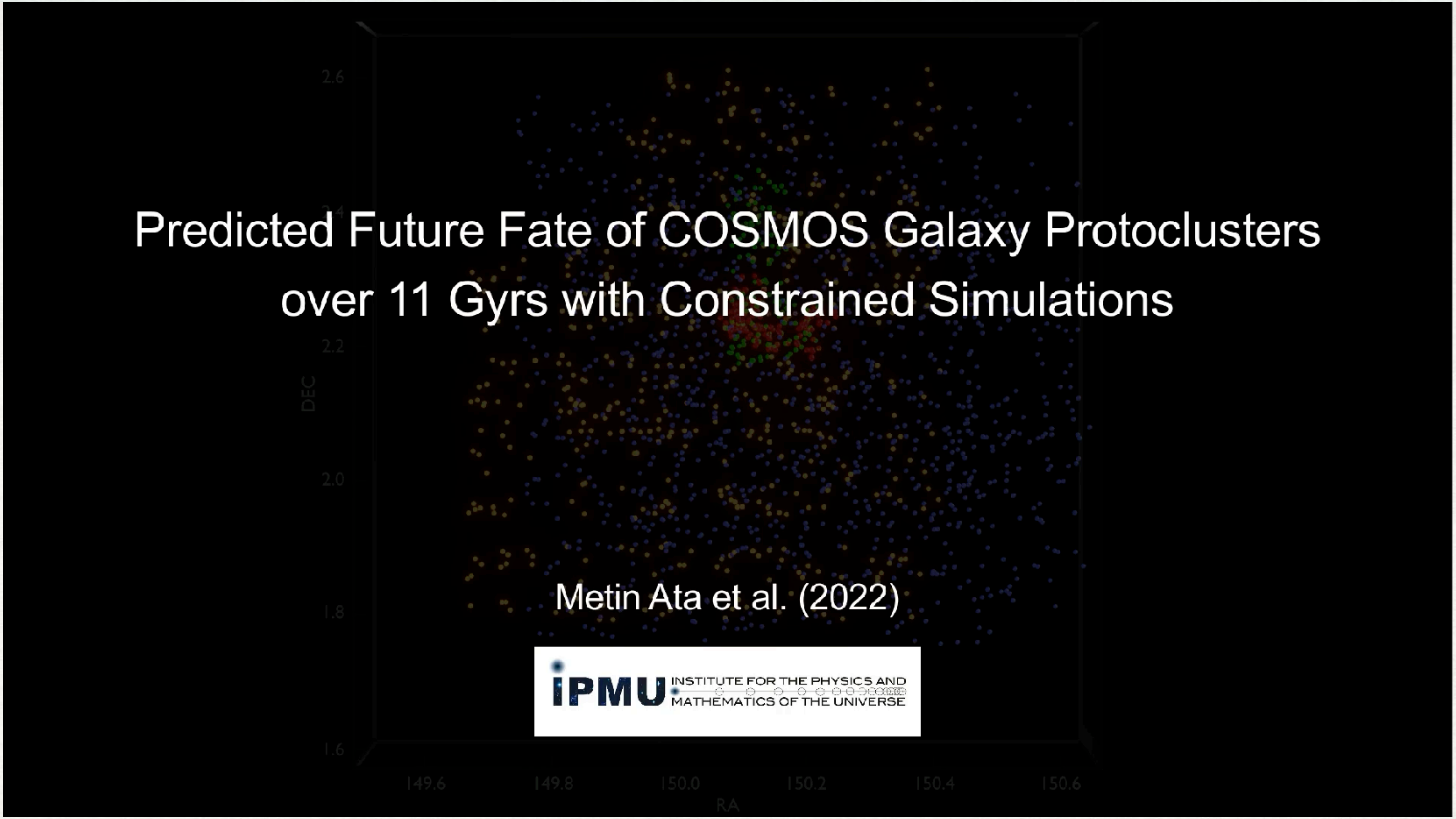
3. Search in a radius of $r \leq 15\, h^{-1}{\rm Mpc}$ for reported protoclusters (e.g. Chiang+2017)

4. Average the findings over 50 realisations

MATCH KNOWN PROTOCLUSTERS IN COSTCO



CLUSTER-PROTOCLUSTER



https://www.youtube.com/watch?v=IIqCZUb-OqQ

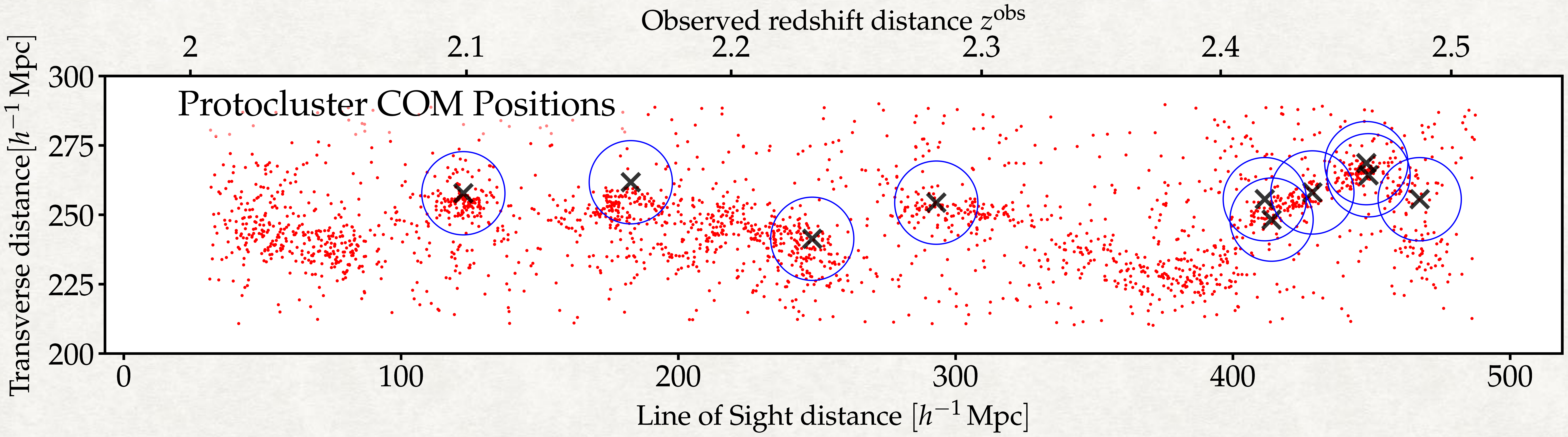
RESULTS

- ZFIRE and Hyperion appear in all realisations
    - While ZFIRE is an isolated structure of $M \sim 1.2 \times 10^{15}\, h^{-1} M_{\odot}$, Hyperion will likely become a very massive $M \sim 2.5 \times 10^{15}\, h^{-1} M_{\odot}$ filament with an spatial extent of $\sim 65\, h^{-1}\mathrm{Mpc}$
- Confirm CC2.2 (Darvish+2020)
- Disfavor CCPC-z22-006 (Franck+2016)
- Inconclusive on G237 (Polletta+2021)
- 5 new protoclusters discovered in COSMOS!

Name [Ref.]	R.A. [deg]	Dec [deg]	z^{obs}	Final Mass [$h^{-1}\, M_{\odot}$]
ZFOURGE/ZFIRE [9, 10]	150.094	2.251	2.095	$(1.2 \pm 0.3) \times 10^{15}$
G237* [19]	150.507	2.312	2.16	−−
CC2.2* [18]	150.197	2.003	2.232	$(4.2 \pm 1.9) \times 10^{14}$
CCPC-z22-006 [11]	149.930	2.200	2.283	−−
	150.093	2.404	2.468	
	149.976	2.112	2.426	
	149.999	2.253	2.444	
Hyperion [12–17]	150.255	2.342	2.469	$(2.5 \pm 0.5) \times 10^{15}$
	150.229	2.338	2.507	
	150.331	2.242	2.492	
	149.958	2.218	2.423	
COSTCO J100026.4+020940	150.110	2.161	2.298	$(4.6 \pm 2.2) \times 10^{14}$
COSTCO J095924.0+021435	149.871	2.229	2.047	$(6.1 \pm 2.5) \times 10^{14}$
COSTCO J100031.0+021630	150.129	2.275	2.160	$(5.3 \pm 2.6) \times 10^{14}$
COSTCO J095849.4+020126	149.706	2.024	2.391	$(6.6 \pm 2.3) \times 10^{14}$
COSTCO J095945.1+020528	149.938	2.091	2.283	$(4.3 \pm 2.4) \times 10^{14}$

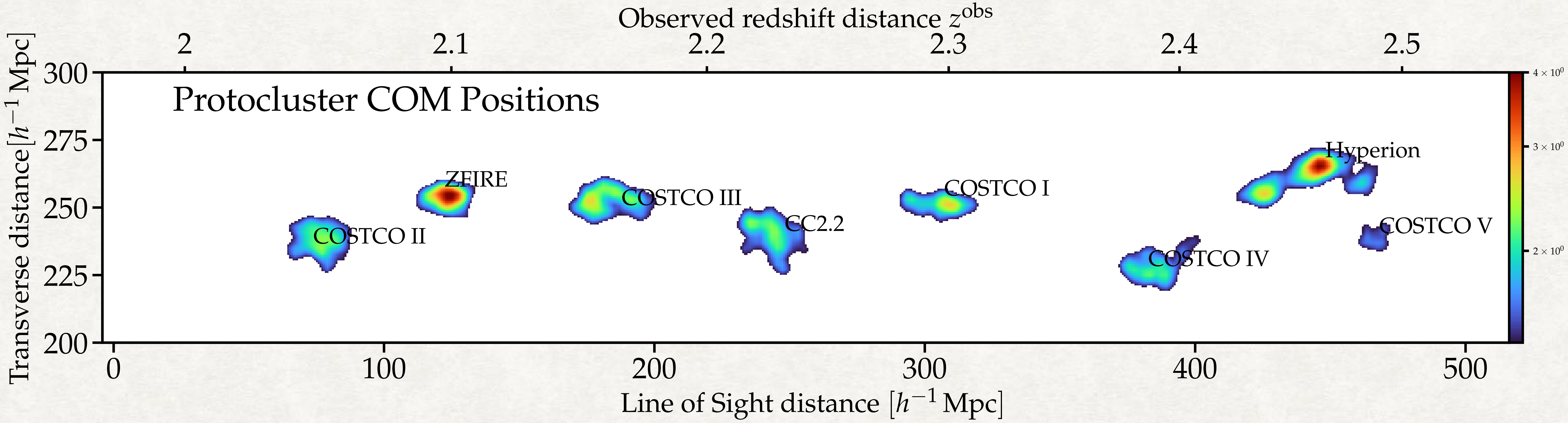
MIND THE ORPHANS

- Stack all 50 COSTCO N-body runs into one volume: The COSTCO Multiverse
- Identify protoclusters that are forming in significantly many realisations (28 of 50)
- DBSCAN multiverse clustering analysis

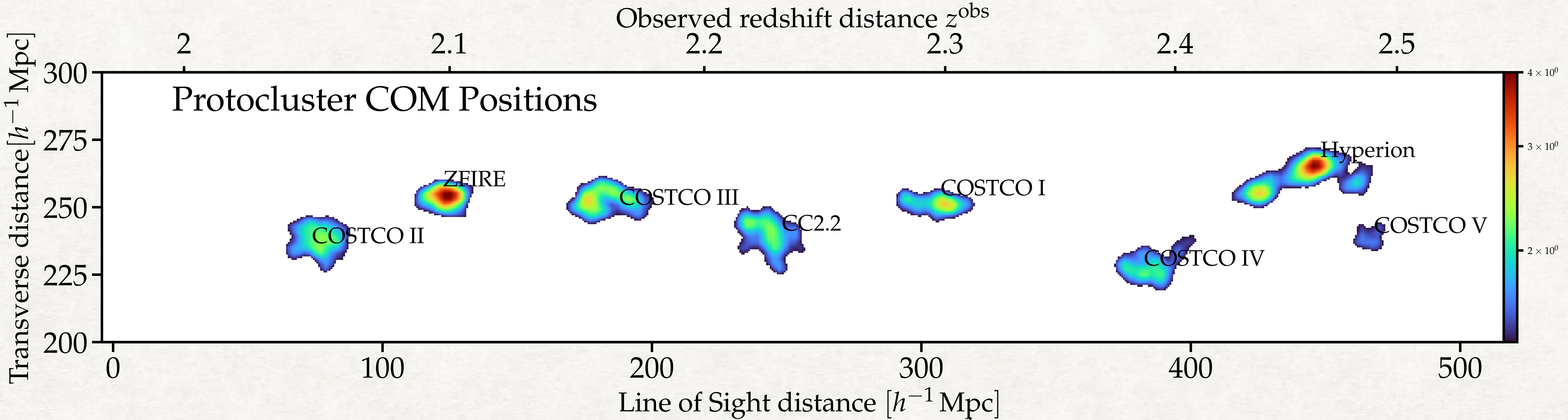


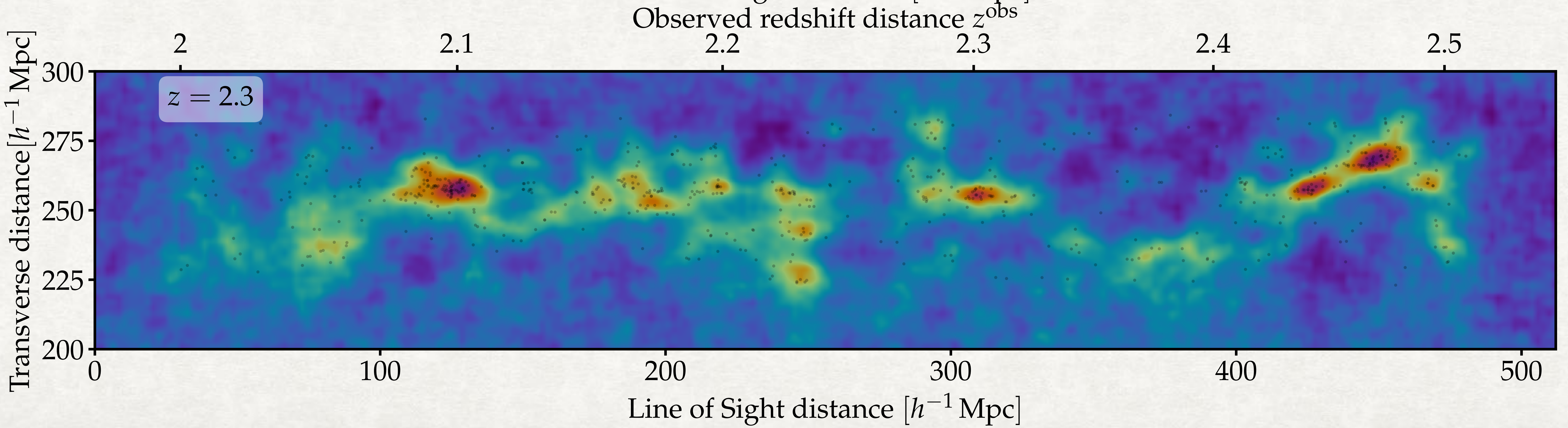
MIND THE ORPHANS

- Stack all 50 COSTCO runs into one volume: The COSTCO Multiverse
- Identify protoclusters that are forming in significantly many realisations (28 of 50)
- DBSCAN multiverse clustering analysis



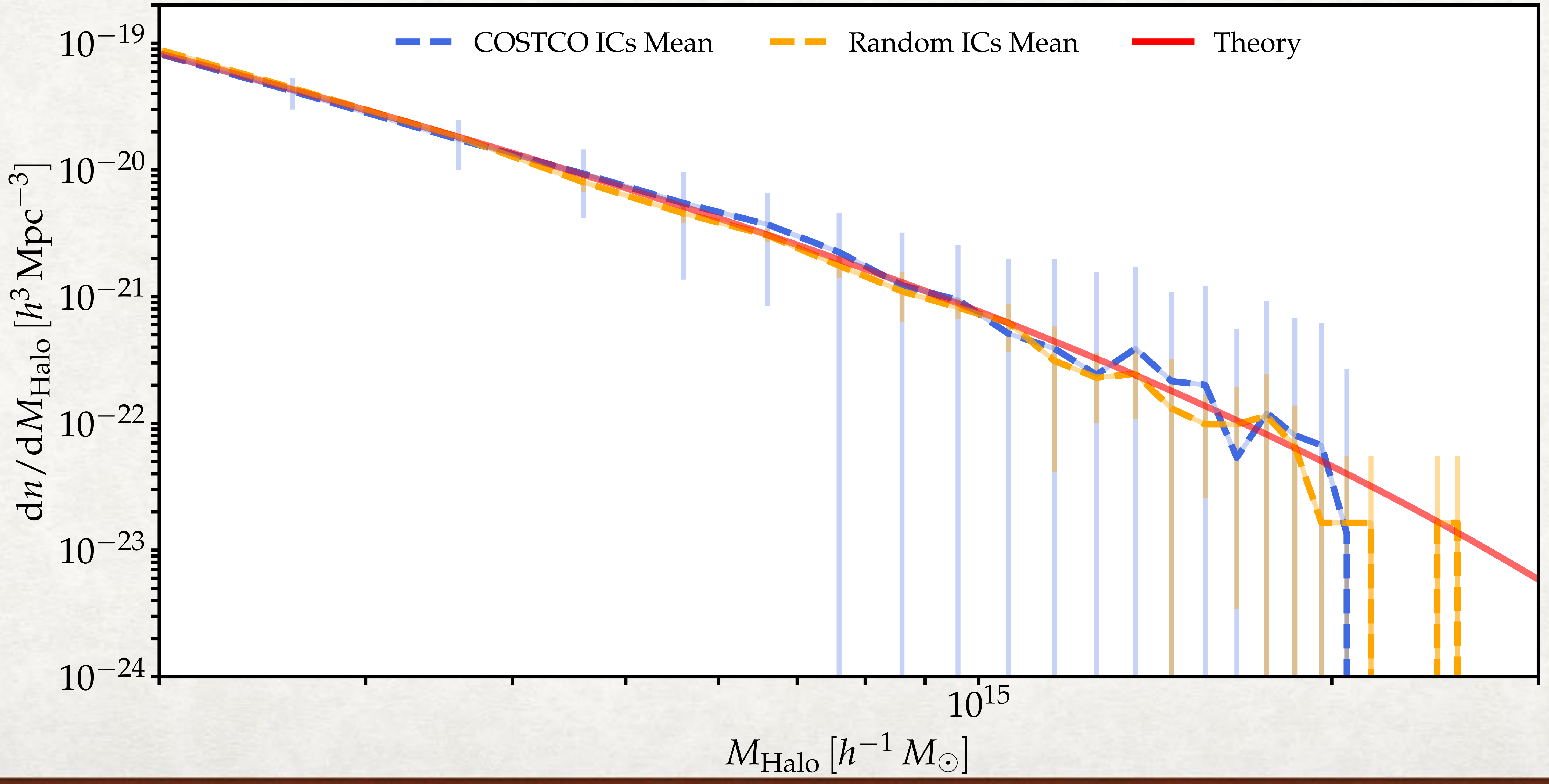
MATCHING THE ORPHANS





COSTCO HALO MASS FUNCTION

- Are there too many massive protocluster progenitors in COSMOS at 2<z<2.5?
- Compare the COSTCO halo mass function with a set of random simulation with same setup (box size, resolution, …)
- Number of ~10^{15} h^{-1} $M_{\odot}$ halos in COSTCO is consistent with global ΛCDM



SUMMARY

- Application of density reconstruction and constrained simulations to COSMOS spectroscopic data at $2.0<z<2.5$
- Constrained simulations now allow *direct* modeling of galaxy protoclusters:
  - "Fast forward" to $z=0$ to measure their final cluster masses
  - Hyperion will collapse into a giant supercluster filament (~100 cMpc long)
  - 5 new protoclusters discovered in existing data
- Follow-up projects:
  - Galaxy properties as a function of matter density field (Rieko Momose)
  - Quasar Lyman-alpha forest absorption as a function of matter density field (Ilya Khrykin)
  - Comparison with CLAMATO IGM Tomography (Chenze Dong)