

Constraining cosmological simulations with peculiar velocities: A forward-modeling approach

A. Valade^{1,2,3}, N. Libeskind², D. Pomarède⁴, R. Stiskalek⁵, Y. Hoffman³, S. Gottlöber² and R. B. Tully⁶

¹ Aix Marseille Université, CNRS/IN2P3, CPPM, Marseille, France e-mail: avalade@aip.de

² Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany

³ Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel

⁴ Institut de Recherche sur les Lois Fondamentales de l'Univers, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

⁵ Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK

⁶ Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

Received Month XX, 20XX; accepted Month XX, 20XX

ABSTRACT

Context. Numerical simulations are a key tool for deciphering the dynamics of gravitation. Yet, they fail to spatially reproduce the Universe we observe, limiting the comparison between observations and simulations to a statistical level. This is highly problematic for rare, faint, or well-studied nearby objects that are observed in a single environment. The computational cost of recovering this environment in random simulations is prohibitive for most realistic applications.

Aims. We present HAMLET-PM, a method that enables the constraining of initial conditions for cosmological simulations so as to produce evolved numerical universes that can be directly compared to observations of the local Universe. Such simulations are known as constrained simulations.

Methods. Our method implements the field-level forward modeling of the early-time density field from sparse and noisy measurements of late-time peculiar velocities. The dynamics are integrated with a coarse particle-mesh gravity solver, probing the mildly nonlinear regime. We applied our code to the Cosmicflows-4 compilation of peculiar velocities up to $z \sim 0.05$ ($160 h^{-1} \text{Mpc}$). The resulting constrained simulations were analyzed under the framework of the local Universe model as well as an amended version that captures systematic spatial shifts in the scattering of simulated counterparts.

Results. A series of a hundred dark-matter only cosmological simulations with a resolution of 512^3 particles in a $500^3 [h^{-1} \text{Mpc}]^3$ box are presented. These spatially reproduce the observed large-scale structure in the constrained volume. Special attention is given to twelve prominent nearby galaxy clusters, whose simulated counterparts are matched on criteria of mass and separation. We provide a mass estimate constrained by the dynamical environment for each cluster.

Conclusions. Field-level forward modeling of the initial conditions can be applied to actual observations so as to produce highly constrained cosmological simulations. In its current form, this method already overtakes in quality the pipeline currently in use in the peculiar-velocity community, although systematic biases still need to be addressed. Furthermore, improving the model can be done easily thanks to the inherent flexibility of the Bayesian approach.

Key words. Cosmology, large-scale structure, cosmological simulations

1. Introduction

The large-scale distribution of matter in the Universe is understood to be the result of initial quantum fluctuations imprinted during the inflationary stage of the Big Bang followed by billions of years of gravitational instability. Although complex, it is a numerically tractable problem given the initial over- and under-density field, and the relatively simple law of gravitational attraction. The nature of the primordial perturbations in the density and velocity field is well established via observations of the temperature fluctuations in the cosmic microwave background (CMB). The CMB (e.g., Planck Collaboration et al. 2020) describes these perturbations statistically, allowing the large-scale structure to be simulated in a statistical sense, namely with the goal of producing representative volumes of the Universe that are consistent with having grown out of the primordial power spectrum of fluctuations.

Such simulations are indispensable tools for understanding the framework within which galaxies form, and they have been used extensively. Indeed a large variety of codes have been de-

veloped to model the complex hydrodynamics, star formation, and feedback processes, with enormous success (for a recent review, see Vogelsberger et al. 2019). Yet, while important in modeling the galaxy population as a whole, such simulations cannot be as easily compared with the single individual realization of the universe that we inhabit, namely the “local Universe.” To constrain models from observations, cosmologists then resort to summary statistics, such as the two-point correlation (e.g., Adame et al. 2025) function or the Alcock-Paczynski test (e.g., Cuceu et al. 2025).

This paper is part of the effort to use low redshift observations of the local Universe to produce spatially constrained initial conditions, which, once evolved, result in constrained simulations that can be directly compared to the observations (Kravtsov et al. 2002; Klypin et al. 2003). One of the main uses of such simulations so far is the study of galaxy clusters such as Virgo, Coma, and Perseus. Quantities of interest include their mass (e.g., Pfeifer et al. 2023, hereafter P23), merging history and formation (e.g., Sorce et al. 2016, 2020), connectivity (McAlpine et al. 2025, hereafter MA25), dynamical sig-

nature along the line of sight (Sorce et al. 2024), abundance (Dolag et al. 2023), and various other observables (X-Ray luminosity, temperature and SZ effect; Hernández-Martínez et al. 2024, hereafter HM24). Constrained simulations have also been utilized to constrain the galaxy bias (Hoffman et al. 2018).

Two main approaches to construct such initial conditions have been investigated over the last decade, each relying on different data sources, namely redshift data and peculiar velocity data. Redshift surveys have paved the way for the discovery and exploration of the large-scale structure, from the 1980s (de Lapparent et al. 1986) until today (e.g., DESI Collaboration et al. 2025). Modern surveys are dense, precise (with redshift errors of $\sigma_z \leq 10^{-4}$), and tend toward completeness in increasing volumes. Yet, their use for the problem of reconstructing the underlying density field is made difficult by galaxy bias (Desjacques et al. 2018, for a review). The modeling of this relationship on nonlinear scales is complex, and thus has to be approximated by empirical laws to maintain computational tractability in the context of field reconstruction (Jasche & Lavaux 2019; McAlpine et al. 2025).

The redshift contains information about both the expansion of the Universe and the local movement of a galaxy in its environment, known as the peculiar velocity. If measurements of physical distances are available, individual peculiar velocities can be extracted from the redshift data. This allows for the circumvention of galaxy bias, since peculiar velocities are unbiased tracers of the matter field at leading order. Yet, measurements of distances are difficult to obtain and catalogs thereof are thus scarcer, noisier, and prone to interpretation biases.

The Constrained Local UniversE Simulation consortium (CLUES; Yepes et al. 2009; Doumler et al. 2013b) has pioneered the use of peculiar velocities for field reconstructions and constraining simulations (e.g., Sorce et al. 2014, 2020, 2024; Hoffman et al. 2018; Libeskind et al. 2020; Pfeifer et al. 2023; Dolag et al. 2023). This pipeline unfolds in three steps: (1) the reconstruction of the density and velocity fields from observational data with a Wiener filter (WF; Zaroubi et al. 1995); (2) the shifting of the constraints back in time with a reverse Zel-dovich approximation (RZA; Doumler et al. 2013c); and (3) the creation of initial conditions for constrained simulations (CRs; Hoffman & Ribak 1992; Doumler et al. 2013a). This approach is somewhat limited by the inherent approximations of each step. The Wiener filter assumes a linear relationship between the density and peculiar velocity field. While this is an acceptable assumption on large scales (beyond a few megaparsecs), it breaks down in the nonlinear regime. The method also assumes that the uncertainties on the velocity are normally distributed. The latter hypothesis is at odds with the reality of velocity data, which motivated the development of “un-biasing” techniques (e.g., Sorce 2015; Hoffman et al. 2021; Sorce et al. 2023).

The application of field-level forward modeling to peculiar velocities to reconstruct the velocity field in the late-time Universe was first proposed by Lavaux (2016) and then later applied to the Cosmicflows-3 catalog (Tully et al. 2016) by Graziani et al. (2019). Boruah et al. (2020) and Valade et al. (2022, HAMLET) independently improved on the method presented in Graziani et al. (2019) by replacing the Gibbs-sampling with a Hamiltonian Monte Carlo (HMC) technique and taking advantage of the modern GPU-compatible Python computation library JAX (Bradbury et al. 2018), accelerating the

numerical exploration by orders of magnitude¹. An application of the HAMLET methodology to the Cosmicflows-4 catalog Tully et al. (2023, CF4) is presented in Valade et al. (2024), while Boruah et al. (2020) discuss the application of their method to a composite dataset. We note that a third separate single-CPU implementation of the HMC version of Graziani et al. (2019) has been developed and applied to Cosmicflows-4 by Courtois et al. (2023).

The peculiar-velocity-based reconstruction methods discussed above are based on the linear relationship between density and velocity and are thus practically limited to where peculiar velocities are accessible: the late-time Universe. The first fully Bayesian reconstruction of initial conditions from velocity data was proposed by Prideaux-Ghee et al. (2022). Here, the dynamics of particles are modeled with the Zeldovich approximation (Zel’dovich 1970) and the late-time velocity field is computed from the reconstructed cloud-in-cell (CiC) density field (Birdsall & Fuss 1969) within the linear theory. This approach was further improved on by Bayer et al. (2023) by replacing the Zeldovich approximation with a particle-mesh gravity solver (FlowPM; Modi et al. 2021) and by avoiding the use of linear theory in the late-time Universe.

In this work, we present HAMLET-PM, an extension of the HAMLET method (Valade et al. 2022) that enables one to directly constrain the initial conditions from $z = 0$ velocity data. The linear model of the former is replaced with a differentiable particle-mesh gravity solver (PMWD; Li et al. 2022). To tackle the issue of empty cells, the $z = 0$ velocity field is constructed using a hybrid approach mixing a CiC technique in dense regions and linear theory. The initial conditions are then augmented with random small-scale power and run to $z = 0$ with the AREPO code (Springel 2010), resulting in 100 simulations of the local Universe with 512^3 particles in a box of size $500 h^{-1}$ Mpc constrained up to $z \sim 0.05$ ($160 h^{-1}$ Mpc).

Throughout this paper we use $\Omega_m = 0.3$, $\Omega_b = 0.0471$ and $H_0 = 74.6$ km/s/Mpc. Consistent with other local measurements, the high value of H_0 is fit to the Cosmicflows-4 data so as to ensure no global radial in- or outflow (Tully et al. 2023; Duangchan et al. 2025).

2. Data

Peculiar velocities catalogs are built by complementing redshift surveys with physical distance measurements, so as to isolate the radial component of each galaxy’s peculiar velocity (v_r) from its Hubble expansion (Hubble 1929). To first order we write:

$$v_r = cz - H_0 d, \quad (1)$$

where H_0 is the Hubble constant and c the velocity of light. z and d are observed measurements of redshift and physical distance.

In this work, we employ the grouped version of the Cosmicflows-4 (CF4; Tully et al. 2023), the largest compilation of peculiar velocities to date, with about 56 000 galaxies gathered in more than 38 000 groups². The grouping offers two advantages over the raw data, as it mitigates the finger-of-God effect in massive galaxy clusters (Kaiser 1987) and improves the quality of the constraints by averaging distance measurements.

¹ We note that the original implementation of Valade et al. (2022) is based on TENSORFLOW (Dillon et al. 2017) and has since been ported to JAX (Bradbury et al. 2018).

² Note that we use the term “group” generically, even for groups composed of a single galaxy.

All in all, the sample is rather isotropic up to $z \approx 0.055$ but reaches to $z = 0.1$ in the SDSS footprint (York et al. 2000).

The CF4 catalog is a cross-calibrated compilation of numerous peculiar velocity surveys. The bulge of the estimated distances relies on scaling relationships such as the fundamental-plane relationship (Djorgovski & Davis 1987) with about 42 000 entries principally extracted from the SDSS-PV and 6dFGRSv surveys (Howlett et al. 2022; Magoulas et al. 2016), supplemented by about 12 000 Tully-Fisher measurements Tully & Fisher (1977). These yield distances with a 20% relative uncertainty. Type Ia supernovae constitute the third most numerous type of constraints, with about a thousand extracted from the SH0ES and PantheonPlus catalogs (Brout et al. 2022; Riess et al. 2022). The supernova component of CF4 provides more precise velocities (a 10% relative error on distances) on the entire volume probed. Lastly, high-precision estimators such as Cepheids or red giant stars yield distances with 5% relative errors in the vicinity of the Milky Way, up to $z \approx 0.02$.

Particular attention must be paid to the distance measurements. At the first order, the distances (d) are derived from distance moduli (μ):

$$d = 10^{\mu/5-5} \text{ Mpc}. \quad (2)$$

Distance moduli measurements are assumed to have normally distributed errors: $\mathcal{N}(\mu, \sigma_\mu)$. The scatter, σ_μ , depends on the distance estimator and on the grouping and ranges from 0.05 mag for some entries to 0.5 mag for the large majority of the catalog. It follows from eqs. (1) and (2) that distances and velocities are log-normally distributed³. The uncertainty thus scales with distance, resulting in a very poor signal-to-noise ratio for velocities, with a median value of 10% compared to the Λ CDM predicted value of $\sigma_v = 300$ km/s. Furthermore, if the data are naively interpreted, the log-normality of the error leads to a spurious infall (see Hoffman et al. 2021). Last but not least, it is to be noted that uncertainties stemming from the cross-calibration may result in large-scale flows (Whitford et al. 2023, fig. 10).

The computational cost of the method described below scales steeply with the reconstructed volume. We thus limit the CF4 data to a redshift velocity of $z_{\text{max}} = 0.055$, which only removes the lopsided extrusion of the SDSS-PV catalog.

3. Method

This work builds on the previously presented HAMLET code, whose details can be found in Valade et al. (2022). A quick summary is given below, followed by a thorough explanation of how modifications have been made to produce initial conditions.

3.1. Bayesian framework

The proposed method is written in a Bayesian context. The aim is to derive and estimate the posterior probability distribution of a set of free parameters given some fixed observations. In our case the free parameters are the distances of the constraints $\mathcal{D} = \{d_i\}$ for $i \leq n_{\text{obs}}$ and the Fourier modes of the high redshift over-density field projected on a grid, $\Delta = \{\delta_k\}$, associated with the accessible modes $\{k_j\}$ for $j < n_{\text{modes}}$. We denote the observations⁴ by their distance moduli, $\mathcal{M} = \{\mu_i^{\text{obs}}\}$, and their

³ We consider the error on the redshifts, $\sigma_{cz} \approx 50$ km/s, to be negligible.

⁴ The superscript “obs” refers to observational data so as to avoid confusion with the similar quantities computed by the model.

redshifts, $\mathcal{Z} = \{z_i^{\text{obs}}\}$. Note that distance moduli are in fact not direct observations, though we consider them as such in this model for the sake of simplicity. A better modeling, such as that of Stiskalek et al. (2025), may be explored in further works.

Following Bayes theorem, the posterior probability can be written as

$$P(\Delta, \mathcal{D} | \mathcal{M}, \mathcal{Z}) \propto P(\mathcal{M}, \mathcal{Z} | \Delta, \mathcal{D})P(\Delta, \mathcal{D}), \quad (3)$$

where the denominator $P(\mathcal{M}, \mathcal{Z})$ has been omitted as it is not a function of the free parameters. The two terms on the right-hand side are named, respectively, the likelihood and the prior distribution. Implicit to these equations is the underlying mathematical model linking the free parameters and the observations. Section 3.2 details the description of the observations and section 3.3 focuses on the modeling of the density and velocity fields.

3.2. Observational model

The mathematical derivation of the model is detailed here, while a visual summary is presented in fig. A.1 under the form of a directed acyclic graph (DAG).

3.2.1. Likelihood

The forward model is composed of two parts. First, the observed quantities (μ, z) need to be computed from the model parameters (d, Δ) for each constraint, alongside with a few intermediaries:

$$v_r = \mathbf{v}(d\hat{\mathbf{r}}) \cdot \hat{\mathbf{r}}, \quad (4)$$

$$z_c = \frac{2}{3\Omega_m} \left(1 - \sqrt{1 - \frac{3\Omega_m H_0}{c} d} \right), \quad (5)$$

$$z = (1 + z_c)(1 + v_r/c) - 1, \quad (6)$$

$$d_L = (1 + z + (\mathbf{v}_{\text{CMB}} \cdot \hat{\mathbf{r}})/c)d, \quad (7)$$

$$\mu = 5 \log_{10}(d_L) + 25, \quad (8)$$

where $\mathbf{v}$ is the 3D peculiar velocity field, v_r is the radial peculiar velocity, $\hat{\mathbf{r}}$ is the unit direction vector, z_c and z are, respectively, the cosmological and full redshifts, d_L is the luminosity distance, $\mathbf{v}_{\text{CMB}}$ is the 3D velocity of the Milky Way with respect to the CMB, and μ is the distance modulus. The form of eq. (7) is given by Calcino & Davis (2017).

Once the observables have been derived, their probabilities can be written. The likelihood $P(\mathcal{M}, \mathcal{Z} | \Delta, \mathcal{D})$ reads

$$P(\mathcal{M}, \mathcal{Z} | \Delta, \mathcal{D}) = \prod_i \mathcal{N}[\mu_i^{\text{obs}}, \sigma_{\mu,i}](\mu_i) \times \mathcal{N}[z_i^{\text{obs}}, \sigma_z](z_i), \quad (9)$$

where $\mathcal{N}[m, s](x)$ denotes the Gaussian distribution with a mean, m , and a standard deviation, s , evaluated at x . Note that, in reality, uncertainties are correlated through (i) the fitting of the distance-indicator relations and (ii) the cross-calibration process, both of which propagate uncertainties shared across multiple constraints. For simplicity, however, the CF4 catalog reports only individual errors.

3.2.2. Selection function and prior distribution

Observational constraints (*e.g.*, flux-limitation) restrict the number of observed objects. While a proper description of these effects is important for an unbiased Bayesian analysis, such as that detailed in Kelly et al. (2008), the composite nature of the Cosmicflows-4 catalog makes the modeling of the selection

function an arduous task. We then opt for a simplified approach, detailed in appendix A, by considering an ad hoc prior function,

$$P(\mathcal{D}) = \prod_i \hat{\mathcal{H}}_{d_i}(d_i), \quad (10)$$

where $\hat{\mathcal{H}}_{d_i}$ is the distribution of observed redshift distances computed from the observed redshifts, smoothed with a $3 h^{-1}$ Mpc Gaussian kernel. Note that alternative approximations exist, such as Stiskalek et al. (2025).

3.3. Velocity field model

To reconstruct the late time velocity field, we employed the FastPM n-body gravity particle-mesh solver (Feng et al. 2016), an improved particle-mesh solver that enforces the correct linear displacement and yields slightly better results than COLA (Tassev et al. 2013) when used at low resolution. This work uses the PMWD implementation (Li et al. 2022), which is written in JAX and can thus be readily plugged into our HAMLET. These are gravity-only simulations, and thus do not include any baryonic effects. To keep the inference loop tractable, we limited ourselves to 128^3 dark matter particles in a box of $500^3 (h^{-1} \text{ Mpc})^3$, resulting in a mass resolution of $4.9 \times 10^{12} h^{-1} \text{ M}_\odot$. To resolve the gravitational potential, particles were gridded on a 128^3 mesh, leading to a force resolution of $3.9 h^{-1} \text{ Mpc}$.

To enforce the power spectrum, we set the following prior on the Fourier modes of the initial density field:

$$P(\Delta) = \prod_k \frac{|\hat{\delta}_k|^2}{\mathcal{P}(|k|)}, \quad (11)$$

where $\mathcal{P}$ is the matter-matter power-spectrum. This power spectrum was generated within PMWD using an amplitude of $A_s = 2 \times 10^{-9}$ and a spectral number of $n_s = 0.96$. The ICs were then brought to a redshift of $z = z_{\text{PM}}$ with a second-order Lagrangian perturbation theory (2-LPT; Buchert & Ehlers 1993) and further evolved to $z = 0$ with n_{PM} PM steps.

A first order CiC (Birdsall & Fuss 1969) algorithm was used to obtain a field representation of the density field (ρ^{CiC}) and velocity field ($\mathbf{v}^{\text{CiC}}$) at redshift $z = 0$. We set the CiC grid resolution to that of the particles' grid, i.e., a cell size of $3.9 h^{-1} \text{ Mpc}$.

Constructing the velocity field from an N -body simulation can lead to issues in both dense and empty regions (Hahn et al. 2015). Notably, the CiC of the velocity cannot be correctly estimated in void regions, where empty cells are found. To rectify this issue, we employed a linear model of the velocity field in nearly empty regions. Linear theory is accepted to be accurate in the early Universe or at low redshift on large scales (above $5 - 10 h^{-1} \text{ Mpc}$). The linear relationship between density and velocity can still be applied in an evolved universe as a simplifying approximation, to derive the velocity field from the nonlinear over-density field:

$$-H_0 f \delta^{\text{CiC}} = \nabla \cdot \mathbf{v}^{\text{lin}}, \quad (12)$$

where f is the linear growth factor and $\delta = (\rho - \bar{\rho})/\bar{\rho}$ is the density contrast. This approximation is also made by Prideaux-Ghee et al. (2022). Equation (12) tends to remain accurate at all scales in void regions and is well defined even in empty cells; however, it does overestimate the velocity in dense regions, leading to spurious dynamics in and around dense structures.

The velocity field used for the inference is thus a hybrid of these two fields, using the CiC representation in populated cells and the linear representation in emptier regions:

$$\mathbf{v} = w(\rho^{\text{CiC}}) \mathbf{v}^{\text{CiC}} + (1 - w(\rho^{\text{CiC}})) \mathbf{v}^{\text{lin}} \quad (13)$$

$$w(\rho^{\text{CiC}}) = \frac{1}{2} \left(1 + \text{erf} \left(\frac{\log_{10}(\rho^{\text{CiC}}) - \alpha}{\beta} \right) \right), \quad (14)$$

where ρ^{CiC} is the overdensity in a cell. Note that empty cells (i.e., $\rho^{\text{CiC}} = 0$) have null weights (i.e., $w(\rho) = 0$). In this work, we use $z_{\text{PM}} = 19$, $n_{\text{PM}} = 5$, $\alpha = -0.3$ and $\beta = 0.1$. These values are based on preliminary tests and will be more deeply discussed in further studies.

3.4. Running the constrained simulations

The posterior function described above was constructed with our banana framework⁵ and explored with a HMC algorithm (Neal 2011) implemented by the Python library blackjax⁶. This exploration is complicated by nonlinearities, necessitating significantly more computational power than the linear model presented in Valade et al. (2022, 2024). To reduce self-correlation in the random sample, ten chains of 6000 steps were run. For each chain, the first 1000 steps were discarded as warm-up and then only one sample every 500 steps was considered, leading to a total of $10 \times 10 = 100$ constrained initial conditions. Convergence was assessed visually, and we verified that the internal correlation on each Monte Carlo chain was minimal (see appendix B). Each Monte Carlo chain took about 10 hours to run on an NVIDIA-V100 GPU, meaning we were able to create statistically independent ICs every hour of computation.

The initial conditions inferred from the data according to the method described above have a relatively low resolution of 128^3 cells. To construct higher-resolution initial conditions, random small-scale power is injected so as to reach a resolution of 512^3 particles. For this simple purpose, we developed our own routine. The resulting particle mass is $7.8 \times 10^{10} h^{-1} \text{ M}_\odot$. The initial conditions were brought to $z = 63$ with the 2-LPT and then evolved to $z = 0$ with the AREPO gravity solver (Springel 2010). Dark matter halos were produced on the fly in AREPO according to a friend-of-friend algorithm, using the standard $l = 0.2$ linking length. Only halos with more than 32 particles were considered, leading to a minimum halo mass of $2.5 \times 10^{12} h^{-1} \text{ M}_\odot$.

4. Quality assessment of the constrained simulations

In the last few years, assessing the quality of such simulations has moved from a qualitative “visual” approach to a more quantitative one. So far, two metrics have emerged that compare constrained simulations to observations; namely, the local Universe model (LUM, P23) focusing on galaxy clusters, and the velocity field olympics (VFO; Stiskalek et al. 2025) studying the velocity field. Going even further, MA25 suggested a three-level verification program, from essential to more ambitious, which we summarize as follows: (1) verifying that the simulation is compatible with the underlying cosmological model; (2) quantifying the quality of the constrained volume via rigorous metrics; and (3) reproducing more observables (e.g., tSZ; Stiskalek & Desmond 2026), e.g., through hydro-dynamical re-simulations (e.g., Hernández-Martínez et al. 2024). Joining the

⁵ <https://github.com/anvvalade/banana>

⁶ <https://github.com/blackjax-devs/blackjax>

350 movement toward standardization, this manuscript endeavors to
fulfill tiers (1) and (2), leaving tier (3), however, to later studies.

4.1. Comparison to the literature

The advent of quantitative metrics enables a fair comparison between constrained simulation suites, as well as “direct” reconstructions of the local Universe. Here, we give a quick overview of the methods appearing in the result and discussion sections.

Two sets of velocity-based constrained simulations are discussed. First are the nine CLONES simulations (Sorce 2018) alongside with an accompanying high-resolution baryonic version of one of the CLONES simulations (Dolag et al. 2023; Hernández-Martínez et al. 2024) dubbed SLOW. Second, the 100 constrained simulations of P23 are considered. Both of them apply the CLUES pipeline (WF/RZA/CRs) but they differ in their debiasing methodologies, presented, respectively, in (Sorce 2018) and Hoffman et al. (2021), as well as in their data sources, Cosmicflows-2 (Tully et al. 2013) for the former and Cosmicflows-3 for the latter (Tully et al. 2016). The direct reconstructions of the local Universe of Courtois et al. (2023) and Valade et al. (2024) are also discussed, both of which apply a forward-modeling method to the Cosmicflows-4 dataset.

We also consider two sets of redshift-based constrained simulations, namely CSiBORG1 (Bartlett et al. 2021) and Manticore-Local (McAlpine et al. 2025). Both simulation suites stem from the Aquila consortium and apply a nonlinear forward modeling to augmented versions of the 2MASS catalog (Huchra et al. 2012). The core difference between CSiBORG1 and Manticore-Local is the replacement of a particle-mesh gravity solver by COLA (Tassev et al. 2013), as discussed in Stopyra et al. (2024), alongside the addition of new constraints to ensure that the initial conditions remain normally distributed. Lastly, two direct reconstruction methods are discussed, namely those of Carrick et al. (2015) and Lilow et al. (2024), both working with the 2MASS catalog of redshifts. The former is a relatively naive estimation of the density field from direct galaxy count, while the latter utilizes neural networks trained on full n -body simulations.

4.2. Recovering galaxy clusters

Massive galaxy clusters are found at the nodes of the large-scale structure. Such high-mass objects are relatively rare. Their abundance, properties, and environment in a given volume are subject to cosmic variance. Recovering them in a controlled environment with constrained simulations thus holds a significant scientific advantage.

4.2.1. The local Universe model (LUM)

The LUM is a framework proposed by P23 that aims to quantify the deviation from randomness of a constrained simulation, namely how successful a given attempt is. In essence, it looks at massive simulated halos in the neighborhood of the position of an observed cluster and quantifies the probability of the null-hypothesis: “this halo stems from a random simulation.” Details on our implementation of the LUM are found in appendix D.

4.2.2. Optimizing the local Universe model (opt-LUM)

The LUM introduces a relevant metric but overlooks several aspects of the problem it aims to solve. Indeed, it is unable to disen-

Table 1. Selected clusters in the local Universe.

Cluster	SGX	SGY	SGZ	MCXCII name
	[Mpc/h]			
Virgo	-2.68	11.70	-0.48	Virgo
Fornax	-1.47	-11.24	-10.23	Fornax
Centaurus	-30.63	13.35	-6.76	A3526
Hydra	-23.91	20.52	-24.23	A1060
Norma	-46.10	-6.49	5.71	A3627
Perseus	50.65	-10.88	-13.16	A426
Leo	-2.50	62.72	-11.65	A1367
Coma	0.41	68.17	9.94	Coma
Ophiuchus	-59.99	7.18	57.50	Ophiuchus
Hercules	-38.75	73.81	64.77	A2063
Shapley (A)	-102.02	55.41	4.26	A3571
Shapley (B)	-121.81	73.54	-3.40	A3558

Notes. Positions were derived from the MCXCII X-Ray catalog (Sadibekova et al. 2024). Redshift distances have been transformed into comoving distances using our cosmology.

tangle systematic positional errors in the halos distribution from random scatter. This systematic shift may be a shortcoming of the method or may stem from an observational effect, such as uncertainties on the observed cluster position, whether this uncertainty stems from redshift space distortion or direct distance estimations. Such shifts systematically degrade the p value. We introduce here the opt-LUM metric, which is designed to capture such a tight gathering of massive halos – provided there is one in the neighborhood of the observed cluster – while disregarding less massive and less densely scattered halos.

The opt-LUM is a simple algorithmic amendment to the LUM. In essence, for each cluster, the method iteratively shifts the target position until it is at the center of mass of the matched halos. The procedure unfolds as follows. At the initial time, the targets are set at the observed clusters positions. Then, (1) halos are matched with the targets using the LUM approach described above and (2) for each cluster, the new target position is set to the “center of mass” of the matched halos:

$$\mathbf{r}_{\text{target}} = \frac{\sum_i \mathbf{r}_i w_i}{\sum_i w_i}, \quad w_i = \log_{10} \left(\frac{M_i}{\min M_i} \right), \quad (15)$$

where i iterates over the matched halos. Steps (1) and (2) are repeated until $\max(p_{\text{val}}^{(s+1)} - p_{\text{val}}^{(s)}) < 10^{-4}$, where p_{val} is the p value, (s) denotes the s -th iteration step, and the maximum is taken over all considered clusters. The above threshold is reached in fewer than ten iterations.

By construction, the opt-LUM gives higher detection rates than those of the pure LUM, yet the meanings of these two values slightly differ. The p values yielded by the opt-LUM metric are with respect to the center of mass of the cloud of halos, while the p values of the LUM are with respect to the observational cluster. The opt-LUM is not meant to replace the LUM metric, but rather to augment it.

4.2.3. Selected clusters in the local Universe

There is so far no consensus in the constrained simulation community on a benchmark catalog of clusters to be reproduced, in order to validate one’s bid at a constrained simulation. Indeed, the building of such a wish list depends on the data used – mostly

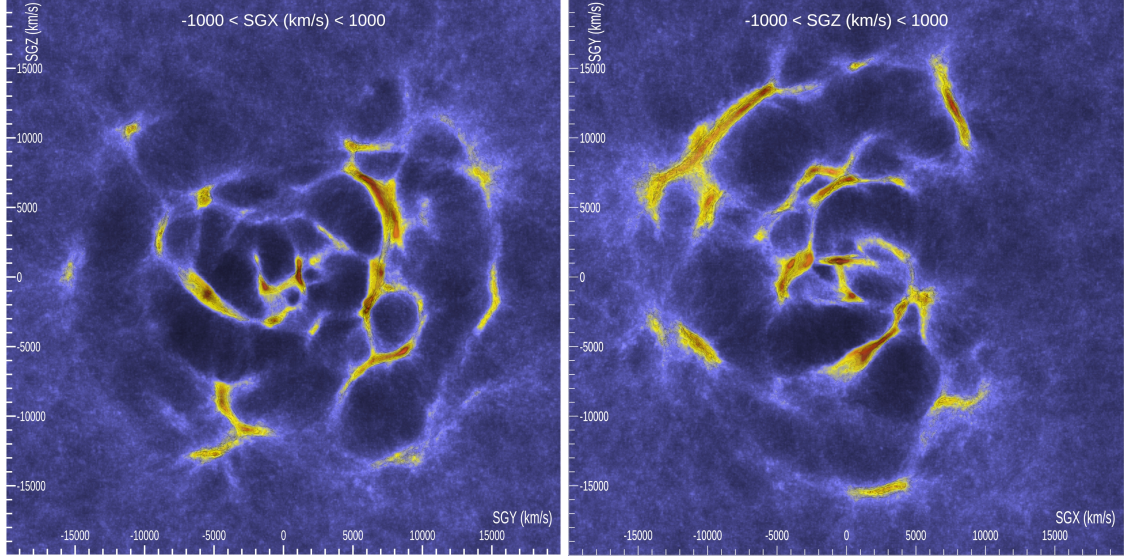


Fig. 1. Slices of the mean density field through the supergalactic planes $SGX=0$ and $SGZ=0$, cropped to the constrained region. The mean density was computed as a geometrical mean, *i.e.*, $\bar{\rho} = (\prod_i \rho_i)^{1/n}$. This approach increases the readability and enables a comparison with Hoffman et al. (2018) and Pfeifer et al. (2023).

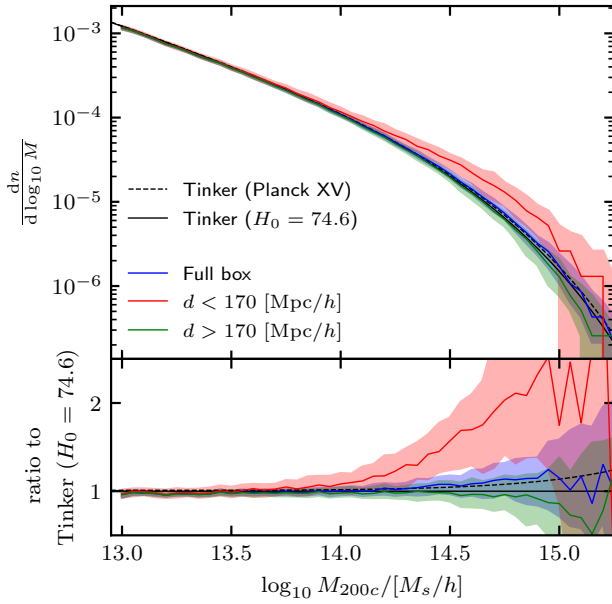


Fig. 2. Halo mass function of the entire volume, in the constrained volume and in the remaining, unconstrained volume. The shades indicate the 1σ confidence level derived from the 100 constrained simulations. The empirical halo mass function of Tinker et al. (2008) is shown twice: once with the cosmology of the simulation suite and once with the standard Planck 15 values (Planck Collaboration et al. 2016). The lower panel shows the ratio to Tinker et al. (2008) with the cosmology of the simulations suite.

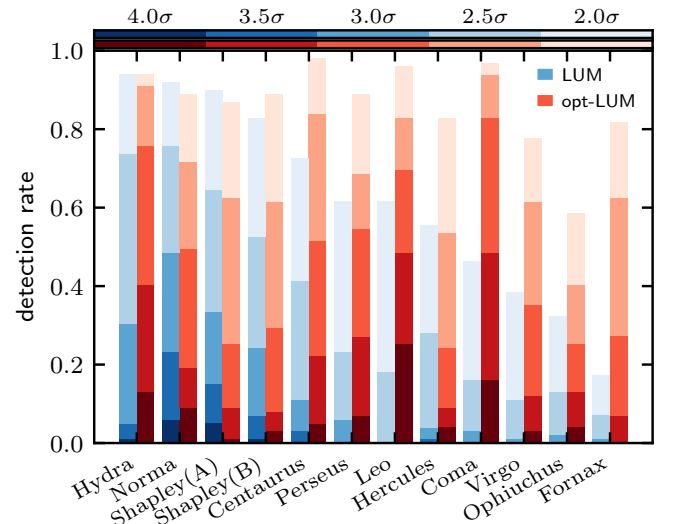


Fig. 3. Detection rate of major (or nearby) clusters of the local Universe, defined in P23 as the rate of simulations meeting the upper p value threshold for each given cluster. Results for the LUM approach are given in blue and results for the opt-LUM algorithm (see eq. (15)) in red.

sky coverage and depth – and also on the constraining power of both the data and the reconstruction method, as less massive halos are harder to recover than more massive ones. The choice of studied clusters seems, then, to have grown organically in each group.

In this work, 12 clusters are studied. These have been selected from the largest clusters, so as to probe different

well-reconstructed regions, while maximizing the overlap with the preexisting literature. The positions of the clusters have been extracted from the MCXC-II catalog of X-ray clusters (Sadibekova et al. 2024). Estimations of mass have been extracted from MCXC-II; the literature review from HM24 (Table A2), which includes mass estimations derived from X-ray, Sunyaev-Zeldovich, and dynamical probes; and finally mass measurements from Tully (2015) comprising values from an empirical luminosity-mass relationship and from dynamical probes.

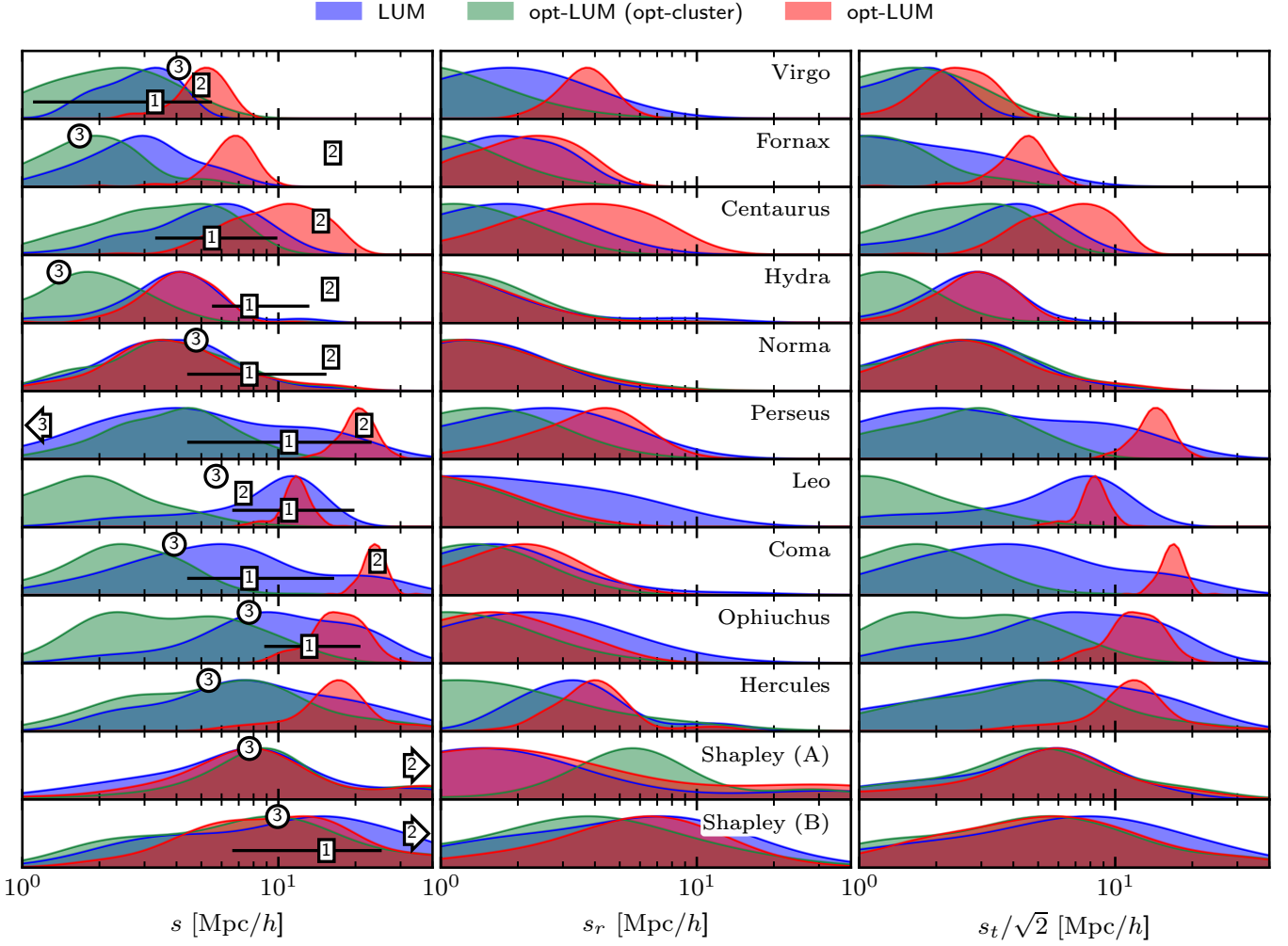


Fig. 4. Posterior separation of the matched halos to the observational cluster. The left, central, and right columns represent, respectively, the 3D, radial, and ortho-radial separations. The blue and red curves correspond, respectively, to halos matched with LUM and opt-LUM methods. Values in green indicate the separation between the halos matched by the opt-LUM method and the associated shifted target. The ortho-radial component is normalized by a factor of $\sqrt{2}$ so as to enable a comparison with the radial component (for an isotropic scatter, $|s_t|/|s_r| = \sqrt{2}$). Clusters are sorted by increasing distance. Values from the literature include P23 (1), HM24 (2), Tully (2015, 2MASS, 3). The separations presented in MA25, all below $1 h^{-1}$ Mpc, are not plotted here.

4.3. Assessing the velocity field

The assessment described above focuses on measuring how well a constrained simulation reproduces the local cluster positions; namely, the main sources of gravitational attraction. Stiskalek et al. (2025) offer the first quantitative benchmark of constrained simulations based on the reconstructed velocity field, known as the velocity field olympics (VFO).

In this framework, the reconstructed velocity fields are compared to observational catalogs of peculiar velocities. Accordingly, Stiskalek et al. (2025) primarily consider the Tully-Fischer datasets SFI++, 2MTF, CF4-i, and CF4-W1 (resp. Masters et al. 2006, 2008; Kourkchi et al. 2020a,b), while the analysis is extended to other catalogs in the appendix. For each set of data, an elaborate forward model is built so as to reproduce the linewidth, as well as biases (e.g., in-homogeneous Malmquist bias; Strauss & Willick 1995). In the absence of absolute calibration, the outcome is independent of the assumed H_0 constant. However, it does depend on smoothing of the reconstructed

fields and on the overlap between the data used to constrain the simulations and the test data.

The results of the VFO consist of Bayesian evidence for each test dataset, which quantifies the ability of a given model to match a dataset. The higher the Bayesian evidence, the more compatible the reconstructed field is with the test data. Note that the evidence scales with the size of the considered data. The approach also fits for the global parameter β , which we interpret in this work as an amplification parameter of the velocity field with respect to the data. A β factor superior (resp. inferior) to unity means that the tested method underestimates (resp. overestimates) the contrast of the observed velocities.

5. Results

5.1. Density field

To give an insight into the constraining level of the method, fig. 1 shows two slices through the density field averaged over the 100 constrained simulations. Outside the constrained volume,

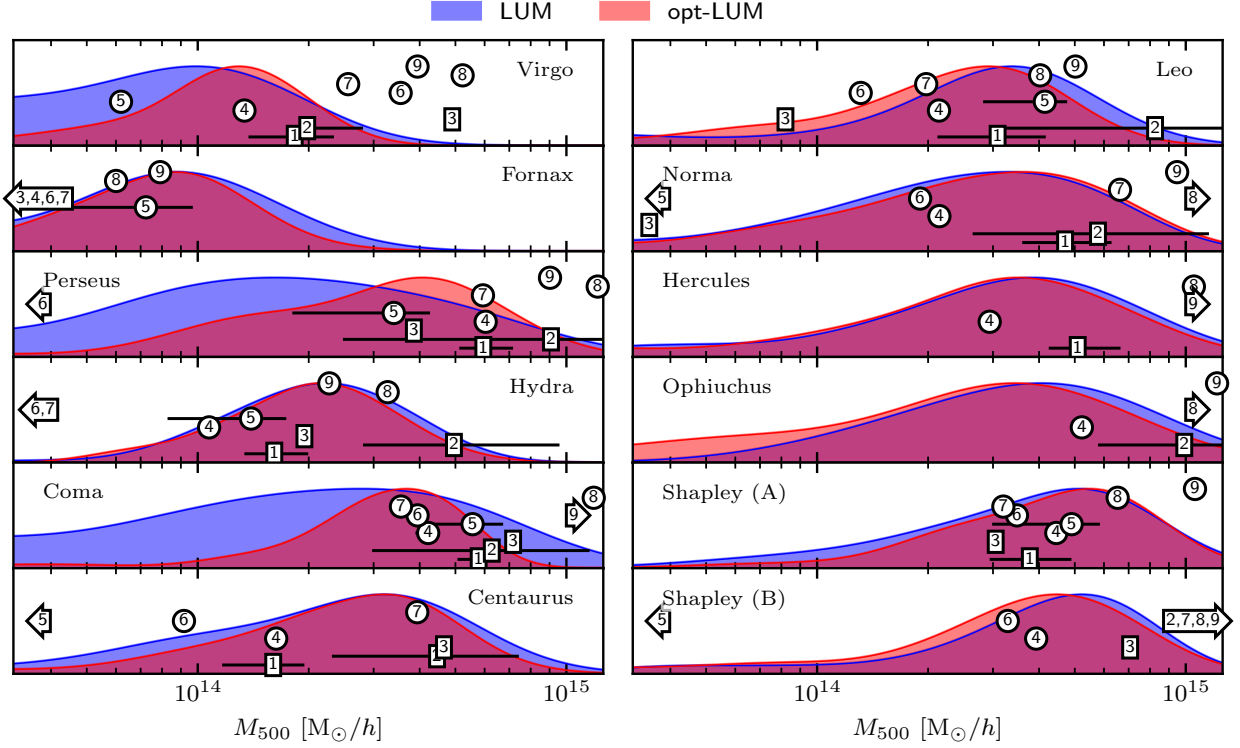


Fig. 5. Posterior mass distributions of major (or nearby) clusters of the local Universe. To facilitate a comparison with mass estimates in the literature, the M_{500c} is used; that is, the mass in the inner region of the halo such that $\rho > 500\rho_c$. Values of P23, provided in M_{200c} have been divided by an empirical factor of 3/2. Values from the literature in squares are taken from constrained simulations, while circled values are observational. Arrows pointing to the left (*resp.* right) indicate that the value is out of the plotted range. Clusters are sorted by increasing posterior mean mass. Values from the literature include MA25 (1), P23 (2), HM24 (3: SLOW, 5: X-Ray, 6: SZ, 7: dyn), Sadibekova et al. (2024, MCXC, 4), Tully (2015, 2MASS, BW: 8, M/L: 9).

the density field converges on the cosmic average. This is the expected behavior within the adopted Bayesian framework and seems to indicate that the Monte Carlo chains converge, although it cannot be excluded that the convergence inside and outside the constrained volume are different.

On the other hand, structures within the constrained volume are persistent, meaning they are “constrained.” Prominent voids are visible as well as a system of filaments – or cut walls – linking denser nodes, as expected from observations of the large-scale structure (Bond et al. 1996; Libeskind et al. 2018). Nonlinear structures appear much more pronounced than in linear reconstructions of the local Universe (*e.g.*, Graziani et al. 2019; Sorce et al. 2023; Courtois et al. 2023; Hoffman et al. 2024; Valade et al. 2024). A similar conclusion is drawn when comparing to previously published mean density fields of velocity-based constrained simulations (Hoffman et al. 2018; Pfeifer et al. 2023), despite the post-reconstruction Gaussian smoothing of $4 h^{-1}$ Mpc and $5 h^{-1}$ Mpc, respectively, applied in those works limiting the precision of the comparison. We note that, as in previous forward-modeling reconstructions from radial peculiar velocities (Graziani et al. (2019); Valade et al. (2024), transverse structures appear to be much better captured by the model and the data than radial structures, leading to a “bullseye” visual impression.

The halo mass function is shown in fig. 2 for the full box, the constrained region ($d < 165 h^{-1}$ Mpc), and in the remaining volume. Two reference halo mass functions are given, one for the “standard” Planck-XV cosmology and one for the cosmol-

ogy used in this work, both computed with the empirical model of Tinker et al. (2008). The current cosmology is expected to produce less massive clusters than the standard Planck-XV cosmology, reaching about 20% under-abundance of $10^{15} h^{-1} M_\odot$ halos. The halo mass distribution in the full box is consistent with Λ CDM. The scatter between realizations includes the expected distributions of both Planck-XV and custom cosmologies. Yet, a marked over-abundance of massive clusters is found in the constrained regions: starting at $10^{14} h^{-1} M_\odot$ we produce $\sim 5\%$ more halos than expected, whereas at $10^{15} h^{-1} M_\odot$ this increases to roughly double the expected number of halos. This over-abundance is in strong tension with the underlying cosmological model and at odds with observations of the local Universe (Böhlinger et al. 2020; Dolag et al. 2023; McAlpine et al. 2025), indicating a shortcoming of the method. This over-abundance stems in all likelihood from the mismatch between the low-resolution of the particle mesh solver used in the inference loop and the high-resolution n-body solver used in the posterior simulations. This issue has already been flagged by Stopyra et al. (2024) and is further discussed in section 6. The LUM has been recalibrated to take this spurious over-abundance into account (see appendix D for details). The power spectra of the initial conditions and evolved fields can be seen in appendix C.

5.2. Massive galaxy clusters

Detection rate The detection rate for each cluster is shown in fig. 3, for both the LUM and opt-LUM algorithms. Introduced

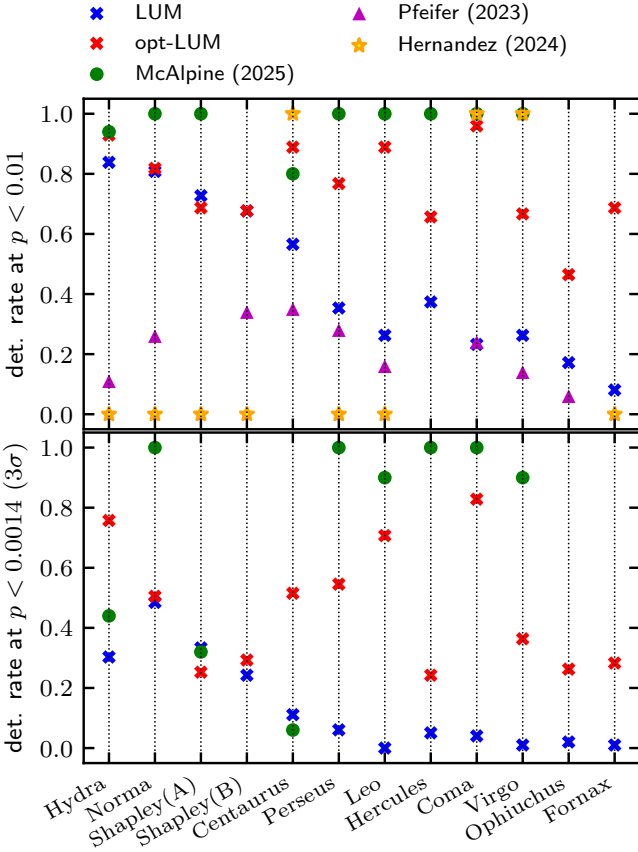


Fig. 6. Comparison of the detection rates of several simulation suites. For compatibility with P23 and MA25, the upper panel gives the detection rate at $p = 0.01$, while the lower panel shows the detection rate at $p = 0.00014$ to improve the comparison and avoid saturated values at $p = 0.01$. To report results from HM24, we set a value of one if the cluster is detected with at least a p value lower than the threshold, and zero otherwise. This, however, is not a proper detection rate.

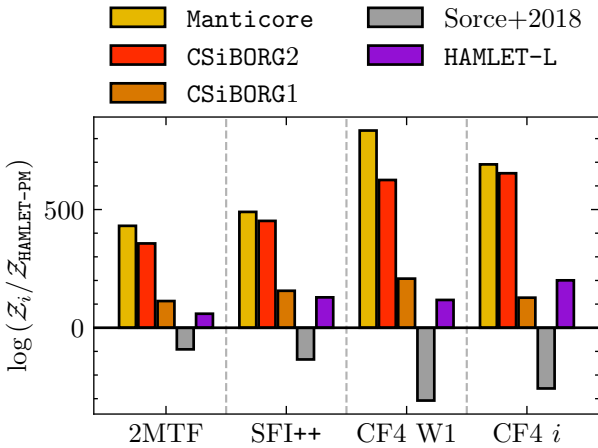


Fig. 7. Log of the evidence⁷ of several constrained simulation suites as well as the linear reconstruction Valade et al. (HAMLET-L; 2024) with respect to this work. Manticore-Local (McAlpine et al. 2025), CSiBORG2 (Stopyra et al. 2024; Stiskalek et al. 2025), and CSiBORG1 (Bartlett et al. 2021) are constrained with redshift data, while CLONES (Sorce 2018) is constructed on velocity data (CF2; Tully et al. 2013).

in P23, this quantifies the rate at which a match below a given p value is found. The p value thresholds cover the range from 2σ to 4σ , following the convention of MA25.

We first discuss the results of the LUM algorithm. More than half of the studied clusters (7/12) are found in more than 60% of the simulations with (at least) 2σ deviation from randomness. Four of them reach a 3σ deviation in at least 20% of the simulations. Marginally, the 4σ threshold is achieved for Shapley (A) and Norma. Fornax, a less massive but very nearby cluster, is only recovered in fewer than 20% of the simulations. Ophiuchus, close to the Zone of Avoidance and thus poorly constrained, also performs below average, with a 30% detection rate at 2σ . Lastly, Hercules, Leo, Perseus, and Coma display a detection rate of between 60% to 40% at 2σ , with no detection at 3σ , which is unexpected given their respective observed masses and the good data coverage in their respective surroundings.

The opt-LUM shows significantly improved detection rates with respect to the standard LUM. All clusters have a detection rate higher than 80% at 2σ , except Virgo, which scores slightly below this threshold, and Ophiuchus, reaching only 60%. The remarkable increase in detection rate for Coma, Perseus, Hercules, and Leo is the subject of appendix F. Half of the selected clusters achieve a 3σ deviation in more than 20% of the simulations and most clusters (10/12) have a non-null 4σ detection rate.

Separation The separation between the clusters and their simulated counterparts is presented in fig. 4. This separation was computed in redshift space, to keep consistency with the matching procedure. Three quantities are presented: the separation in 3D (s), its projection on the line of sight (s_r), and the remaining transverse distance (s_t). This parametrization is motivated by the symmetry of the problem, as the restriction to the line of sight of the velocity data as well as the redshift space distortion necessitate a peculiar treatment of the radial direction. For each quantity, the separation between three pairs of objects were computed: LUM matched halos to observed clusters (blue), opt-LUM matched halos to observed clusters (red), and opt-LUM matched halos to their optimized target (green).

We first discuss the results of the LUM algorithm. The 3D separation roughly grows with distance, from less than $5 h^{-1}$ Mpc for Virgo and Fornax up to more than $20 h^{-1}$ Mpc for Hercules or Shapley (B). For the majority of the clusters, the scatter is more pronounced in the plane of the sky than in the radial direction (i.e., $s_t/s_r > \sqrt{2}$), an effect that we attribute to the relatively low sensitivity of radial velocities to the transversal positions of objects due to projection effects. Our separations are consistent with P23, yielding a slightly tighter scatter for clusters within $50 h^{-1}$ Mpc. For Perseus, Coma, Leo, and Ophiuchus, the recovered scatter is greater than that of P23. Some of the reconstructed clusters align better with the 2MASS galaxy groups of Tully (2015) than with the MCXC-II positions; this is the case for Virgo, Norma, Coma, Ophiuchus, and Shapley (A, B). This is expected, as the grouping of the CF4 velocity data is based on that of Tully (2015).

The opt-LUM approach always leads to separations of the observed clusters that are higher than (or at least as high as) those of the LUM. This is an expected result, as this method allows for the anchor to which the halos are matched to move. Three different behaviors emerge from the application of the opt-LUM algorithm. First, for Virgo, Hydra, Norma, and Shapley (A & B), the results are very similar to that of the LUM, indicating that there is no ambiguity about the group of halos that can be identified as counterparts. Second is the case of Fornax and Centaurus, for

which the opt-LUM slightly shifts the target both in the plane of the sky. These differences of a few comoving Megaparsecs can be attributed to errors in the positions of the clusters and static biases in the resimulations, be they positional (on d) or dynamical (on v_r) as the redshift is a combination of both ($cz = v_r + H_0 d$).

Last, it appears that for Perseus, Leo, Coma, Ophiuchus, and Hercules, the target is moved between $10 - 20 h^{-1}$ Mpc away from the cluster, almost uniquely in the transversal direction. In these five cases, the target converges on a cloud of massive halos scattered by less than $2 - 3 h^{-1}$ Mpc for Coma and Leo, $5 - 10 h^{-1}$ Mpc for Perseus and Ophiuchus, and more than $10 h^{-1}$ Mpc for Hercules. This configuration is confirmed by visualization of the halos in 3D⁸. Furthermore, the simulated counterparts of Coma, Perseus, and Hercules selected in SLOW as well as the so-called “best match” of P23 (not reported in fig. 4) appear to be part of these offset clouds. This result is further discussed in section 6.2.

Mass The posterior mass distribution of the identified counterparts is shown in fig. 5 for both LUM and opt-LUM methods. Mass estimates from other constrained simulations as well as from observations are overplotted. We note that a large scatter is found in the literature, across methods and sources. For instance, estimations from Tully (2015) are systematically higher than others, indicating that the estimated mass may not be comparable with the M_{500c} convention adopted in this paper.

The posterior mass distribution of halos selected by the LUM metric is consistent with the literature, with the notable exception of Perseus as well as Coma and Virgo to a lesser extent, which are less massive in this work than in others. This issue is solved when considering the masses of the opt-LUM counterparts; then, the mass estimations of all clusters are consistent with the literature.

5.3. Velocity field

The application of the VFO metric on our data is presented in fig. 7, using the current work as base for comparison. An extended version of this figure is found in appendix appendix G. For consistency with the results published in Stiskalek et al. (2025), the VFO was applied to only 20 out of the 100 simulations. We selected the first and last state of each chain, so as to ensure that the tested simulations are not correlated (see appendix B).

While the numerical values differ between test catalogs, the ranking of the methods remain the same. HAMLET-PM is outperformed by both redshift-based constrained simulations, namely CSiBORG1 and Manticore-Local. Our method, however, scores better than the CLONES simulations, based on the Cosmicflows-2 data (Tully et al. 2013) with the WF/RZA/CRs pipeline of the CLUES.

Interestingly, the particle-mesh model does not grade as well as the linear one. This is confirmed by the fit β parameter, which indicates that the particle-mesh model overestimates the amplitude of the velocity field by a factor two ($\beta^{-1} = 2.18 \pm 0.21$), while the linear model does not show this effect ($\beta^{-1} = 1.03 \pm 0.03$). As is further discussed in section 6, this is in all likelihood due to the systematic differences between the coarse grav-

ity solver used for the inference and the higher-resolution gravity solver used for the final simulations (Stopyra et al. 2024).

6. Discussion

This work extends the forward-modeling HAMLET framework presented in Valade et al. (2022, 2023, 2024) to produce a novel method, dubbed HAMLET-PM, aimed at constraining initial conditions for cosmological simulations directly from velocity data. All in all, HAMLET-PM appears as a leap forward for the analysis of peculiar velocity data at field-level, compared to the traditional heterogeneous approach. While the traditional CLUES pipeline designed to produce such initial conditions is heterogeneous and limited by (1) linear theory and (2) the tension between the modeling of the observations by the Wiener filter and their actual form (Doumler et al. 2013b; Pfeifer et al. 2023; Dolag et al. 2023), our code HAMLET-PM is able to properly model nonlinearities in the velocity fields with a coarse gravity solver (Li et al. 2022) as well as complex observational effects. Moreover, HAMLET-PM is entirely Bayesian, and thus allows for a self-consistent approach to the problem of generating constrained initial conditions. The full forward-modeling framework is flexible and can be easily extended or corrected. This opens the doors to further improvements of our method.

In parallel, this work contributes to the effort of quantifying and standardizing tests for constrained simulations with the introduction of the opt-LUM, an iterative simulated counterparts matching algorithm built on the methodology of the LUM that captures static biases in the posterior distribution of counterparts. Furthermore, the opt-LUM appears to yield better mass estimations than the simple LUM.

6.1. Tensions with the underlying model

The treatment of peculiar velocity data is admittedly challenging, as measurements are scarce, noisy, and prone to interpretation biases. While HAMLET-PM paves the way to the building of a realistic observational model for such data, shortcomings are still to be acknowledged. HAMLET-PM overestimates the number of massive clusters above $10^{14} h^{-1} M_\odot$ in the local Universe, reaching +150% for $10^{15} h^{-1} M_\odot$ halos, which is not only at odds with Λ CDM but also local Universe estimates (Böhringer et al. 2020; Dolag et al. 2023; McAlpine et al. 2025). Compared to Λ CDM predictions, more particles appear to be found in halos as there is no under-abundance of low-mass halos. Furthermore, the power spectrum of the initial conditions appears to be slightly overestimated across the probed frequency range (by 5% to 10%), as does the amplitude of the evolved velocity field (by almost 220%).

This tension with the underlying model comes from the shortcoming of the coarse gravity solver used in the inference loop. This effect is first presented in Feng et al. (2016), who demonstrated that a five-step FastPM underestimates the halo density mass function toward the high end by 60%. Furthermore, very similar results have been described by Stopyra et al. (2024) in the context of constraining simulations from redshift data with a ten-step particle-mesh solver Jasche & Lavaux (2019). Our result here is slightly different than that of Stopyra et al. (2024), as we reconstruct the velocity field instead of the density field. Yet, it is reasonable to expect the effect to be the same. It appears that the inference exaggerates the contrast in the ICs in order to fit the observed galaxy density so as to palliate the inability of the coarse particle-mesh solver to reproduce accurately densely collapsed objects. Once resimulated with an accurate gravity solver,

⁷ Note the slight calligraphic difference between the notation for Bayesian evidence following Stiskalek et al. (2025) used in figs. 7 and G.1 that of the ensemble of observed redshifts discussed in section 3.

⁸ This visualization is available online [under this link](#).

the overestimation of the contrast in the ICs transforms into over-collapsed halos, *i.e.*, halos that are too massive. To mitigate this issue, Stopyra et al. (2024) suggests replacing the 10-step particle-mesh solver with a 20-step COLA solver. As this work relies on FastPM, which already enforces correct motion on the large scales as COLA. Thus, increasing the number of time steps while reducing the force-resolution of the coarse gravity solver in the inference loop would alleviate the issue (Feng et al. 2016). We leave the investigation of an optimal set of hyper-parameters to a future work. We note that other effects may also contribute to the overestimation of the contrast, such as the inhomogeneous Malmquist bias (Boruah et al. 2022) or the mishandling of the selection function (see fig. 9 of Carreres et al. 2023).

The over-abundance of massive halos may be problematic for the LUM-related metrics as an overestimation of the mass of the halos matched with observational clusters would artificially inflate our performance. Yet, it is noted that the masses of the studied clusters are in good agreement with the literature, indicating that our LUM-results appear to not be boosted by an overestimation of the mass.

6.2. Comparison with other works

The product of our work is twofold, as it offers (1) a picture of the density and velocity fields in the late-time local Universe and (2) a series of cosmological simulations tailored to reproduce velocity observations. We thus compare our results not only to other constrained simulations but also to “direct” reconstructions of the local Universe. We first start the discussion by a rapid qualitative overview, moving toward a more rigorous quantitative analysis.

Qualitative comparison Compared to other velocity-based reconstructions, the improvements appear in the increased sharpness of the reconstructed large-scale structure with respect to previously published works (*e.g.*, Hoffman et al. 2018; Pfeifer et al. 2023; Sorce et al. 2023; Courtois et al. 2023; Valade et al. 2024). We note the clear emergence of nonlinear structures, such as thin walls and filaments. A similar conclusion arises from the comparison with direct reconstructions of the density field from redshift data, such as Carrick et al. (2015) and Lilow et al. (2024). The most recent suite of redshift-based constrained simulations, Manticore-Local, yields a sharper picture of the cosmic web (see Fig. 9 and D3 of McAlpine et al. 2025), although we highlight that visual inspection alone cannot assess the accuracy of the reconstructed fields.

Quantitative comparison To complement the limitations of qualitative comparison, two metrics have recently been proposed. Both highlight the superiority of our approach over other velocity-based pipelines, while also demonstrating its shortcomings when compared against redshift-based works.

The VFO enables a comparison between constrained simulations and direct reconstructions from the angle of the resulting velocity field(s). The results, presented in figs. 7 and G.1, show that the VFO ranks both our linear and nonlinear reconstructions above the other velocity-based methods tested, considerably narrowing the gap to redshift-based approaches, while the most recent Manticore-Local suite retains a comfortable lead. On the CF4-*i* dataset, our linear reconstruction overtakes CSiBORG1 – a former release of simulations constrained by the BORG algorithm, now superseded by Manticore-Local – to

match Carrick et al. (2015); Lilow et al. (2024). We note, however, a degree of circularity in this last assessment: the data used for evaluation shares entries with the data on which the reconstructions are based.

A similar conclusion is drawn from the application of the LUM metric to various (suites of) constrained simulations. Presented in fig. 6 for two detection rates, the results evidence that HAMLET-PM recovers massive clusters of galaxies at a higher rate than P23 but a lower detection rate than Manticore-Local for most clusters. We note that Perseus and Coma are found at the same detection rate by P23, which is further discussed in appendix F.

A comparison with the SLOW simulation is made difficult by the absence of statistical ensemble, as only one set of ICs was selected out of nine pre-run constrained simulations (Sorce 2018; Dolag et al. 2023). Yet, we tested that 96% (*resp.* 20%) of our simulations have more than three counterparts with p value < 0.01 (*resp.* 0.0014), ensuring that at least the large majority of our simulations would achieve a similar result. Six of our simulations find eight counterparts out of the twelve targeted clusters.

Conclusion The quantitative comparison of velocity-based reconstructions, supported by a qualitative assessment, shows that our method outdoes previous works. This tends to indicate that the full forward modeling is able to put tighter constraints on galaxy clusters than the standard CLUES pipeline. A sensitive improvement may also stem from the use of a larger dataset; namely, CF4 versus CF2 for SLOW and CF3 for P23.

Yet, our method is largely outperformed by Manticore-Local. We propose two explanations for this effect. First, upgrading the gravity solver used in BORG has largely improved their results, while reducing biases (Stopyra et al. 2024; McAlpine et al. 2025). Second, Manticore-Local has a higher grid resolution (3.9 Mpc vs. 5.22 Mpc here). More importantly, the reconstruction of Manticore-Local builds on redshift data, which is denser (~69 000 galaxies vs. ~31 000 here) and orders of magnitude more precise (a median signal to noise of less than 1% vs. more than 180% here). This observational comparative advantage of redshift data over velocity data should hold with upcoming surveys.

Last but not least, reconstruction from redshift data has a clear advantage from the viewpoint of cluster-based metrics such as the LUM. Indeed, the exact position of massive clusters is very well defined in the redshift data, while it is much less clear in the velocity data. From a pure geometrical point of view, radial velocities are weakly sensitive to transverse shifts in the positions of the galaxy clusters. The flattening of the counterparts’ scatter around their associated observational clusters is symptomatic of this limitation.

In parallel to the application to upcoming large new datasets, future works should focus on better handling observational systematics, such as the selection function or a better modeling of uncertainties stemming from the cross-calibration process, as well as improving the modeling of the density and velocity fields so as to probe deeper the nonlinear regime.

Acknowledgements. AV thanks S. McAlpine as well as E. Hernandez-Martinez for sharing results of their respective works. AV has been supported by the Agence Nationale de la Recherche of the French government through the program ANR-21-CE31-0016-03. RS acknowledges financial support from STFC Grant No. ST/X508664/1, the Snell Exhibition of Balliol College, Oxford.

References

- Adame, A. G., Aguilar, J., Ahlen, S., et al. 2025, JCAP, 2025, 008
- Angulo, R. E. & Pontzen, A. 2016, MNRAS, 462, L1
- Bartlett, D. J., Desmond, H., & Ferreira, P. G. 2021, Phys. Rev. D, 103, 023523
- Bayer, A. E., Modi, C., & Ferraro, S. 2023, JCAP, 2023, 046
- Birdsall, C. K. & Fuss, D. 1969, Journal of Computational Physics, 3, 494
- Böhringer, H., Chon, G., & Collins, C. A. 2020, A&A, 633, A19
- Bond, J. R., Kofman, L., & Pogosyan, D. 1996, 380, 603
- Boruah, S. S., Hudson, M. J., & Lavaux, G. 2020, 498, 2703
- Boruah, S. S., Lavaux, G., & Hudson, M. J. 2022, MNRAS, 517, 4529
- Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: Composable Transformations of Python+NumPy Programs
- Brout, D., Scolnic, D., Popovic, B., et al. 2022, ApJ, 938, 110
- Buchert, T. & Ehlers, J. 1993, 264, 375
- Calcino, J. & Davis, T. 2017, 2017, 038
- Carreres, B., Bautista, J. E., Feinstein, F., et al. 2023, A&A, 674, A197
- Carrick, J., Turnbull, S. J., Lavaux, G., & Hudson, M. J. 2015, 450, 317
- Courtois, H. M., Dupuy, A., Guinet, D., et al. 2023, A&A, 670, L15
- Cuceu, A., Herrera-Alcantar, H. K., Gordon, C., et al. 2025, DESI DR1 Ly α Forest: 3D Full-Shape Analysis and Cosmological Constraints
- de Lapparent, V., Geller, M. J., & Huchra, J. P. 1986, 302, L1
- DESI Collaboration, Abdul-Karim, M., Adame, A. G., et al. 2025, Data Release 1 of the Dark Energy Spectroscopic Instrument
- Desjacques, V., Jeong, D., & Schmidt, F. 2018, Physics Reports, 733, 1
- Dillon, J. V., Langmore, I., Tran, D., et al. 2017, TensorFlow Distributions
- Djorgovski, S. & Davis, M. 1987, 313, 59
- Dolag, K., Sorce, J. G., Pilipenko, S., et al. 2023, 677, A169
- Doumler, T., Courtois, H., Gottlöber, S., & Hoffman, Y. 2013a, MNRAS, 430, 902
- Doumler, T., Gottlöber, S., Hoffman, Y., & Courtois, H. 2013b, MNRAS, 430, 912
- Doumler, T., Gottlöber, S., Hoffman, Y., & Courtois, H. 2013c, 430, 912
- Duangchan, C., Valade, A., Libeskind, N. I., & Steinmetz, M. 2025, A&A, 708, A119
- Feng, Y., Chu, M.-Y., Seljak, U., & McDonald, P. 2016, MNRAS, 463, 2273
- Graziani, R., Courtois, H. M., Lavaux, G., et al. 2019, MNRAS, 488, 5438
- Hahn, O., Angulo, R. E., & Abel, T. 2015, MNRAS, 454, 3920
- Hernández-Martínez, E., Dolag, K., Seidel, B., et al. 2024, A&A, 687, A253
- Hoffman, Y., Carlesi, E., Pomarède, D., et al. 2018, Nat. Astron., 2, 680
- Hoffman, Y., Nusser, A., Valade, A., Libeskind, N. I., & Tully, R. B. 2021, 505, 3380
- Hoffman, Y. & Ribak, E. 1992, ApJ, 384, 448
- Hoffman, Y., Valade, A., Libeskind, N. I., et al. 2024, MNRAS, 527, 3788
- Howlett, C., Said, K., Lucey, J. R., et al. 2022, 515, 953
- Hubble, E. 1929, Proceedings of the National Academy of Sciences, 15, 168
- Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26
- Jasche, J. & Lavaux, G. 2019, 625, A64
- Kaiser, N. 1987, 227, 1
- Kelly, B. C., Fan, X., & Vestergaard, M. 2008, ApJ, 682, 874
- Klypin, A., Hoffman, Y., Kravtsov, A. V., & Gottlöber, S. 2003, ApJ, 596, 19
- Kourkchi, E., Tully, R. B., Anand, G. S., et al. 2020a, ApJ, 896, 3
- Kourkchi, E., Tully, R. B., Eftekhazadeh, S., et al. 2020b, 902, 145
- Kravtsov, A. V., Klypin, A., & Hoffman, Y. 2002, ApJ, 571, 563
- Kuhn, H. W. 1955, Naval Research Logistics Quarterly, 2, 83
- Lavaux, G. 2016, MNRAS, 457, 172
- Li, Y., Lu, L., Modi, C., et al. 2022, Pmwd: A Differentiable Cosmological Particle-Mesh N^2S -Body Library
- Libeskind, N. I., Carlesi, E., Grand, R. J. J., et al. 2020, MNRAS, 498, 2968
- Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, 473, 1195
- Lilow, R., Ganeshaiah Veena, P., & Nusser, A. 2024, A&A, 689, A226
- Magoulas, C., Springob, C., Colless, M., et al. 2016, in The Zeldovich Universe: Genesis and Growth of the Cosmic Web, ed. R. van de Weygaert, S. Shandarin, E. Saar, & J. Einasto, Vol. 308, 336–339
- Masters, K. L., Springob, C. M., Haynes, M. P., & Giovanelli, R. 2006, 653, 861
- Masters, K. L., Springob, C. M., & Huchra, J. P. 2008, AJ, 135, 1738
- McAlpine, S., Jasche, J., Ata, M., et al. 2025, MNRAS, 540, 716
- Modi, C., Lanusse, F., & Seljak, U. 2021, A&C, 37, 100505
- Neal, R. 2011, in Handbook of Markov Chain Monte Carlo, 113–162
- Pfeifer, S., Valade, A., Gottlöber, S., et al. 2023, MNRAS, 523, 5985
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, 594, A13
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
- Prideaux-Ghee, J., Leclercq, F., Lavaux, G., Heavens, A., & Jasche, J. 2022, Field-Based Physical Inference From Peculiar Velocity Tracers
- Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJ, 934, L7
- Sadibekova, T., Arnaud, M., Pratt, G. W., Tarrío, P., & Melin, J. B. 2024, A&A, 688, A187
- Sawala, T., Jenkins, A., McAlpine, S., et al. 2021, 501, 4759
- Sorce, J. G. 2015, 450, 2644
- Sorce, J. G. 2018, MNRAS, 478, 5199
- Sorce, J. G., Courtois, H. M., Gottlöber, S., Hoffman, Y., & Tully, R. B. 2014, 437, 3586
- Sorce, J. G., Gottlöber, S., Hoffman, Y., & Yepes, G. 2016, 460, 2015
- Sorce, J. G., Gottlöber, S., & Yepes, G. 2020, MNRAS, 496, 5139
- Sorce, J. G., Mohayaee, R., Aghanim, N., Dolag, K., & Malavasi, N. 2024, A&A, 687, A85
- Sorce, J. G., Stoica, R. S., & Tempel, E. 2023, 679, A1
- Springel, V. 2010, MNRAS, 401, 791
- Stiskalek, R. & Desmond, H. 2026, Enhancing Thermal Sunyaev-Zel'dovich Analyses with Digital Twins of the Local Universe
- Stiskalek, R., Desmond, H., Devriendt, J., et al. 2025, The Velocity Field Olympics: Assessing Velocity Field Reconstructions with Direct Distance Tracers
- Stopyra, S., Peiris, H. V., Pontzen, A., Jasche, J., & Lavaux, G. 2024, MNRAS, 527, 1244
- Strauss, M. A. & Willick, J. A. 1995, 261, 271
- Tassev, S., Zaldarriaga, M., & Eisenstein, D. J. 2013, JCAP, 2013, 036
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
- Tully, R. B. 2015, 149, 171
- Tully, R. B., Courtois, H. M., Dolphin, A. E., et al. 2013, 146, 86
- Tully, R. B., Courtois, H. M., & Sorce, J. G. 2016, AJ, 152, 50
- Tully, R. B. & Fisher, J. R. 1977, A&A, 500, 105
- Tully, R. B., Kourkchi, E., Courtois, H. M., et al. 2023, 944, 94
- Valade, A., Hoffman, Y., Libeskind, N. I., & Graziani, R. 2022, 513, 5148
- Valade, A., Libeskind, N. I., Hoffman, Y., & Pfeifer, S. 2023, 519, 2981
- Valade, A., Libeskind, N. I., Pomarède, D., et al. 2024, Nat. Astron.
- Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2019, Cosmological Simulations of Galaxy Formation
- Whitford, A. M., Howlett, C., & Davis, T. M. 2023, MNRAS, 526, 3051
- Yepes, G., Martínez-Vaquero, L. A., Gottlöber, S., & Hoffman, Y. 2009, in American Institute of Physics Conference Series, Vol. 1178, 5th International Workshop on the Dark Side of the Universe, ed. C. Balazs & F. Wang, 64–75
- York, D. G., Adelman, J., Anderson, John E., Jr., et al. 2000, 120, 1579
- Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 449, 446
- Zel'dovich, Ya. B. 1970, 5, 84

Appendix A: More on the forward model

In a Bayesian framework, selections effects are written under the form $P(O \mid \text{obs})$, where O is the event that “is-observed” and obs is any observable quantity. Because Cosmicflows-4 is a compilation of several surveys, its selection function $P(O \mid \text{obs})$ is composite, and its description requires the modeling of other survey- and method-dependent observables, such as the apparent magnitude in different bands.

To circumvent the issue, we note that these observables obs are constructed from the free parameters, and that the selection functions can be seen as functions of the free parameters. In our case, the selection functions can thus be written as $P(O \mid d, \Delta)$, which we simplify to $P(O \mid d)$, assuming the effects of the velocity field on the selection is negligible.

We then note that the posterior depends on the product of the selection function $P(O \mid d)$ and the prior $P(d)$ such that

$$P(d, \Delta \mid \mathcal{M}, \mathcal{Z}) \propto P(O \mid d) \times P(d) \propto P(d \mid O). \quad (\text{A.1})$$

The right-hand side quantity, $P(d \mid O)$, is in fact the distribution of true distances in our observational catalog, which can be approximated by

$$P(d \mid O) = (\mathcal{N}(0, \sigma_{v,\Lambda\text{CDM}}) * \mathcal{H}_{d_z})(d) \quad (\text{A.2})$$

where $\sigma_{v,\Lambda\text{CDM}} \approx 300 \text{ km/s}$ is the expected distribution of peculiar velocities in ΛCDM , the asterisk $*$ denotes a convolution, and $\mathcal{H}_{d_z}$ is the distribution of redshift-derived distances

$$d_z = \frac{H_0}{c} \int_0^z \frac{dy}{\Omega_m(1+y)^3 + \Omega_\Lambda}. \quad (\text{A.3})$$

We apply separately this method to extract (1) groups with a majority of SDSS-PV measurements, (2) groups with a majority of 6dF-PV measurements and (3) other groups. For each sub-catalog, we describe the selection function as

$$P(O \mid d)P(d) = P(d \mid O) = \mathcal{H}_{d_z}(d) \quad (\text{A.4})$$

where $\mathcal{H}_{d_z}$ is the distribution of observed redshift distances computed from the observed redshifts.

Appendix B: Convergence of the Monte Carlo chain

Given a Bayesian model, a Monte Carlo technique aims to produce a sample of n (vectors of) parameters $\{\theta_i\}$ that asymptotically follows the posterior distribution, *i.e.*, $\theta_i \sim P(\theta_i \mid D)$ where D is the data. Ideally, consecutive states of the Monte Carlo chains would also be independent, $P(\theta_i \mid \theta_{i-1}, D) = P(\theta_i \mid D)$. In practice, because each state is constructed from the previous one, the samples remain correlated along the chain, and the number of effectively independent samples is lower than the chain length.

To avoid giving disproportionate weight to over-sampled regions of parameter space, the chain must be “thinned out” until consecutive states are approximately uncorrelated. By construction, separate Monte Carlo chains are independent of one another.

Figure B.1 presents the autocorrelation of the chain of the modes of the over-density field, binned in frequencies. It appears that the correlation between high-frequencies modes very rapidly decreases while low-frequencies modes ($k < 0.13 h \text{ Mpc}^{-1}$, *i.e.*, $\lambda > 50 h^{-1} \text{ Mpc}$) are correlated over about 100 Monte Carlo steps. Beyond the autocorrelation analysis, we report Gelman-Rubin values – not shown here – consistent with unity except in the lowest k -bin, where a value of 0.96 is found.

This suggests that individual chains do not fully explore the parameter space at low k , as expected, and motivated running several chains in parallel. We note, however, that the small number of chains strongly affects the precision of the Gelman-Rubin test, and that values slightly below unity remain consistent with convergence.

The initial conditions selected for the 100 simulations presented in this work are extracted from each of the 10 Monte Carlo chains with a thinning of 50 steps. This means that consecutive simulations may not be entirely independent on the large scales, yet, note that (1) the power added to increase the resolution of the initial conditions is entirely random and that (2) this paper essentially focuses on dark matter halos, whose properties are affected by high-frequency modes (Sawala et al. 2021).

Appendix C: Power Spectra

Figure C.1 presents the power spectra of the 100 constrained simulations. The power spectrum of the initial conditions appears to match well the ΛCDM prediction, while we note that the power appears to be slightly too high over all the probed range which is discussed alongside with other results in section 6. The evolved power spectra follows the ΛCDM linear power spectrum up to $k \approx 0.2 h \text{ Mpc}^{-1}$, frequency beyond which the evolved power spectra depart from linear predictions due nonlinearities in the field.

Appendix D: Local Universe model

In each simulation, each halo can be matched to only one observational cluster. The optimal configuration is computed by optimizing the sum of the p values through the “Hungarian algorithm” (Kuhn 1955). Matches with $p_{\text{val}} > 0.023$ (2σ) are discarded, meaning that all clusters may not have a counterpart in all simulations.

Contrary to P23 and MA25, we argue in favor of running the LUM in redshift space (*i.e.*, setting the halos at their redshift position). This choice is motivated by the use of redshift positions for the observational clusters, which constitutes the bulk of the values found in the literature.

The p values computed by the LUM are cosmology-dependent. This appears in the pseudo-density, mass-dependent parameter $\alpha(M)$ defined as

$$p_{\text{val}}(r, M) = \int_0^s P(s, M) ds \quad (\text{D.1})$$

$$P(r, M) = e^{-\frac{4}{3}\pi\alpha(M)r^3} - e^{-\frac{4}{3}\pi\alpha(M)(r+\Delta r)^3} \quad (\text{D.2})$$

where $P(r, M)$ encodes the probability to find a halo of mass $> M$ within a distance $[r, r + \Delta r]$. The form of $P(r, M)$ is exact for unclustered halos where $\alpha(M)$ is then simply the density of halos of mass $> M$. To account for clustering, the α parameter can be calibrated on cosmological simulations. The best practice is then to use an accompanying series of random simulations sharing the cosmology of the studied constrained simulation. This is the path chosen by P23 and MA25. Alternatively, one may calibrate the LUM by choosing random observers in the studied constrained simulation (HM24). As it creates the risk of biasing the random signal, this approach should be followed only after testing that the constrained simulation is statistically consistent with the underlying cosmology.

For this work, we have chosen to use the available computational power to build a large enough statistical sample of

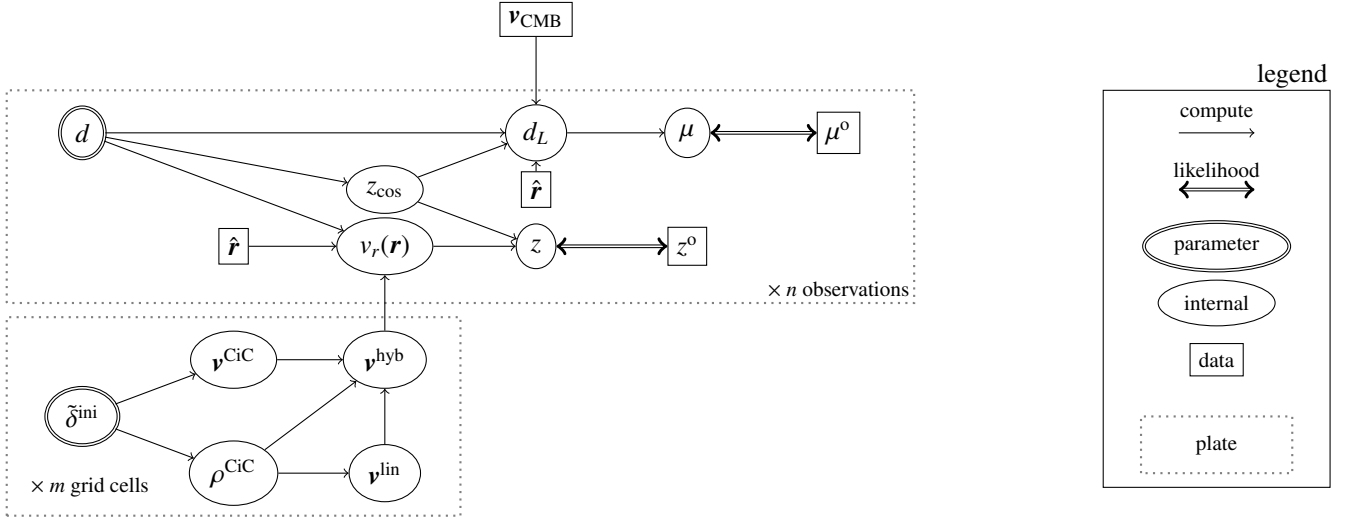


Fig. A.1. Directed acyclic graph (DAG) of the forward model. Each physical quantity is represented as node of the graph – either in an ellipse for modeled quantities or in a square for observational data. Free parameters of the posterior are on the left-most side of the diagram and highlighted by a double ellipse. Simple arrows indicate a relationship between quantities (*i.e.*, “A is computed from B”), while double arrows indicate a comparison (*i.e.*, a likelihood function). Note that prior probability functions do not appear in this representation. Dashed square boxes specify the dimensionality of the physical quantities. The definitions and relationships between quantities are found in section 3.2.

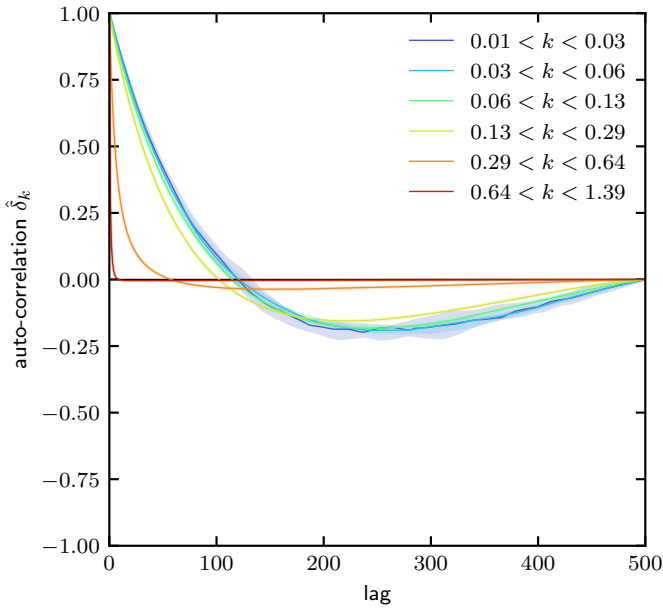


Fig. B.1. Autocorrelation of the (real part) of the Fourier modes of the initial density field, binned in frequencies. The shades represent the scatter alongside the ten run Monte Carlo chains.

constrained simulations. In absence of companion random simulations, the first approach to derive the LUM cannot be used. Furthermore, as presented below in fig. 2 and discussed in sections 5.1 and 6, our simulations deviate from the expected halo mass distribution, making the second approach inapplicable. To account for this effect, we use the calibration of P23 that we correct as such

$$\alpha(M) = \alpha_{P23}(M) \frac{\int_{m>M} \rho_{\text{cst}}(m) dm}{\int_{m>M} \rho_{P23}(m) dm} \quad (\text{D.3})$$

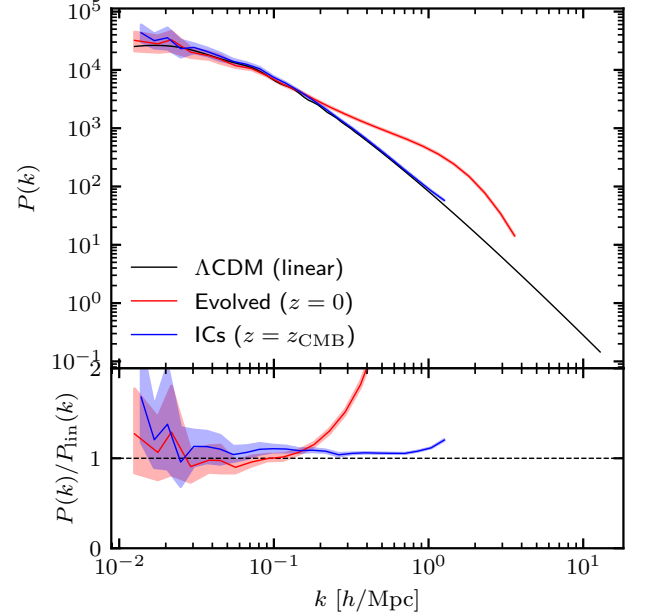


Fig. C.1. Power spectra of the initial as well as the evolved density fields compared with the linear power spectrum. For the reconstructed fields, the solid lines and the shades respectively represent the median and 68% interval around it over the 100 constrained simulations. The initial (*resp.* evolved) density fields have a resolution of 128^3 (*resp.* 512^3) nodes.

where $\rho_{\text{cst}}(M)$ is the halo mass function in the constrained volume ($d < 170 h^{-1} \text{ Mpc}$) and $\rho_{P23}(M)$ is the halo mass function predicted by Tinker et al. (2008) for the cosmology of P23. Physically, this correction only accounts for the variation of halo mass function with cosmology while neglecting the difference in clustering. In effect, it increases the pseudo-density for high-mass halos, from +5% at $M_{200c} = 10^{14} h^{-1} M_{\odot}$ up to +150% at

1070 $M_{200c} = 10^{15} h^{-1} M_{\odot}$. For a given halo-cluster configuration, our
cosmology predicts a higher p value than that of P23.

Appendix E: Influence of the hyperparameters of the gravity solver

1080 The bias in the velocity field and halo masses may stem from more than a single cause. We argue, however, that the principal cause is the inexpensive setup of the FastPM solver used in the inference loop. A direct proof would consist in rerunning the inference with a more expensive setup and performing posterior simulations. By the nature of the Monte Carlo, such a comparison cannot be made between individual states but rather across the entire chain; the associated computational cost precludes such an analysis within the scope of the present work. We instead provide an elaborate toy example that corroborates our interpretation.

Although the influence of the FastPM hyperparameters on simulation quality is extensively discussed in Feng et al. (2016) and Stoppyra et al. (2024) from the perspective of the density field and halo mass, the velocity field is not addressed. We propose a simplified numerical setup to highlight the amplitude of the deviation in the evolved velocity field induced by these hyperparameters. Random initial conditions are generated for 128^3 particles in a $500 h^{-1}$ Mpc box, comparable to the resolution used in the inference loop. Random power is then added on the high-frequency modes to produce 256^3 -particle initial conditions. We use a pair of such initial conditions to capture the variance introduced by the high-frequency random modes, following the sign-pair trick proposed by Angulo & Pontzen (2016). The high-resolution simulations are run with a force resolution $B = 2$, a starting redshift $z_{\text{start}} = 64$, and $n_{\text{step}} = 64$ steps; the density and velocity fields are computed via CiC on a 128^3 -cell grid. A series of low-resolution simulations is then run with various values of B , z_{start} , and n_{step} .

1100 The density and velocity profiles centered on the most massive high-resolution cell are shown in fig. E.1. The rightmost panels display the ratio of the amplitude of the coarse velocity profiles to that of the reference simulation. Although this quantity is not strictly equivalent to the β parameter of the VFO model, it is a reasonable proxy in our simple setup. Several conclusions can be drawn from these plots. First, the density peak is not well reproduced in the coarse setting: it is less dense and may be spatially shifted, as expected from the literature⁹. Second, the amplitude of the velocity field around massive halos is substantially underestimated—by a factor of ≈ 3 in this particular setting, a value we expect to depend on the environment. Although starting the coarse simulations at a higher redshift and resolving gravity with a finer force resolution and more time steps substantially improves the shape of the density field, the amplitude of the velocity field nonetheless remains underestimated in this particular environment.

1120 This toy example sheds light on the bias observed above. The coarse FastPM setup is unable to properly reproduce the nonlinear velocity profile around massive structures. Additional power is therefore injected into the initial conditions during the Monte Carlo walk to compensate for the missing amplitude. Once resimulated with a higher-precision gravity solver, this excess power in the initial conditions propagates directly into the evolved fields, producing an overall factor of $\beta^{-1} = 2$ and the enhanced accretion of matter onto massive halos. We emphasize

⁹ We note the influence of the high-frequency modes on top of the accuracy of the gravity solver (Sawala et al. 2021).

that this amplification is not spatially constant, so a simple renormalization of the initial conditions between the inference and the N -body simulations would not be satisfactory. While this point warrants a more thorough investigation, the example also suggests that targeting a higher resolution is more important than refining the solver setup at fixed resolution.

Appendix F: Shift of Coma and Perseus in velocity reconstructions

1140 The position on the sky of the simulated Coma and Perseus is very interesting. The opt-LUM algorithm proposed here enables us to properly quantify for the first time this shift, which seems to have been present in the literature for a decade. Indeed, all the previously studied velocity-based constrained simulations appear to systematically shift these two clusters by a solid $10 - 20 h^{-1}$ Mpc orthogonally to the line of sight. Furthermore, all linear reconstructions (Sorce et al. 2016; Graziani et al. 2019; Sorce et al. 2023; Hoffman et al. 2024; Valade et al. 2024) as well as the quasi linear fields from Hoffman et al. (2018) and P23, present a density peak about $10 h^{-1}$ Mpc from Perseus' position in +SGY direction, within the zone of avoidance. This is also the case for Coma in our reconstruction. However, the shift in peak density out of the super-galactic planes canonically used to visualize the density field, which prevents us from verifying easily if this is also the case in the literature. We highlight that such a shift appears not to be present in redshift-based reconstructions such as Manticore-Local (MA25), which is consistent with the fact that the redshift distribution is in this case the only source of information.

1160 Furthermore, none of the three studied simulation suites (P23, HM24 and HAMLET-PM) uses the same pipeline nor the same data source. While we note that these datasets build up on each other and thus share constraints, the addition of new, independent additional data should have mitigated systematic biases. The contribution of new data is significant in these regions, with the addition of the SDSS-PV sample for Coma and the CF4-TF sample for Perseus. Similarly, the inter-calibration between sources catalogs is rerun between successive Cosmicflows releases, making it unlikely to be the explanation for this shift.

The shifted position of Perseus is in the ZoA and appears to be close to 3C129.1, three times lighter than Perseus itself (Sadibekova et al. 2024). The shifted Coma is found in the CfA Great Wall, where no other massive cluster is found in the MCXC-II catalog. The density field in these regions is presented in appendix G.

If confirmed, this result may lead to two separate conclusions. Either velocity data is able to detect a major issue in our physical model, such as a strong decoupling between the presence of galaxies and the dark-matter density field or in the laws of gravitation. Alternatively, it may show that velocity-based reconstructions have to be checked carefully for spurious detections. We leave this open problem to the community.

Appendix G: Supplementary material

1180 Supplementary material is gathered in this section. It includes an extended visualization of the VFO results in fig. G.1, as well as the tables gathering the posterior information on massive galaxy clusters in table G.1 (LUM-matched) and table G.2 (opt-LUM-matched).

¹⁰ Note the slight calligraphic difference between the notation for Bayesian evidence following Stiskalek et al. (2025) used in figs. 7

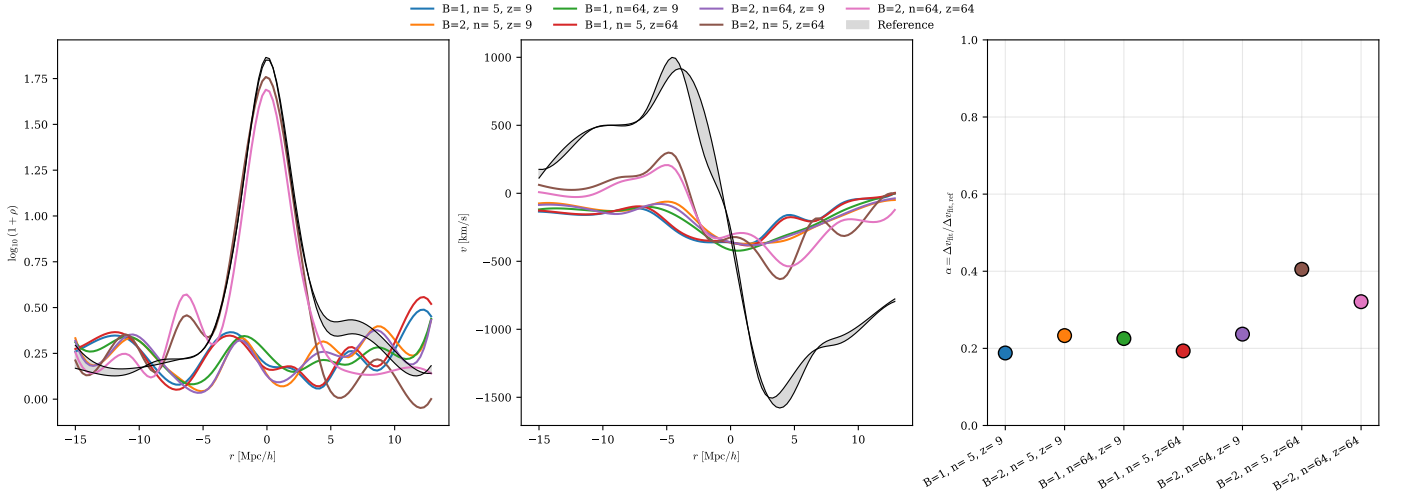


Fig. E.1. Density and velocity profiles for various FastPM setups. The profiles are centered on a the densest cell of the reference (higher-resolution) simulation.

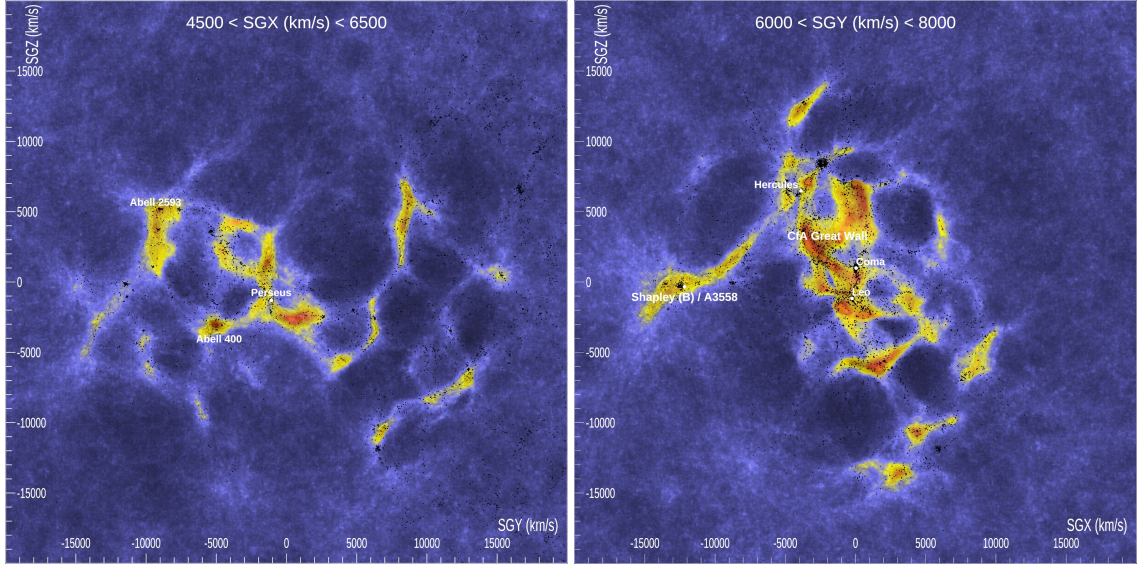


Fig. F.1. Density field around the regions of Coma and Perseus. These slices have been selected to show the observational positions of the clusters as well as the density peaks to which the opt-LUM counterparts belong.

and G.1 that of the ensemble of observed redshifts discussed in section 3.

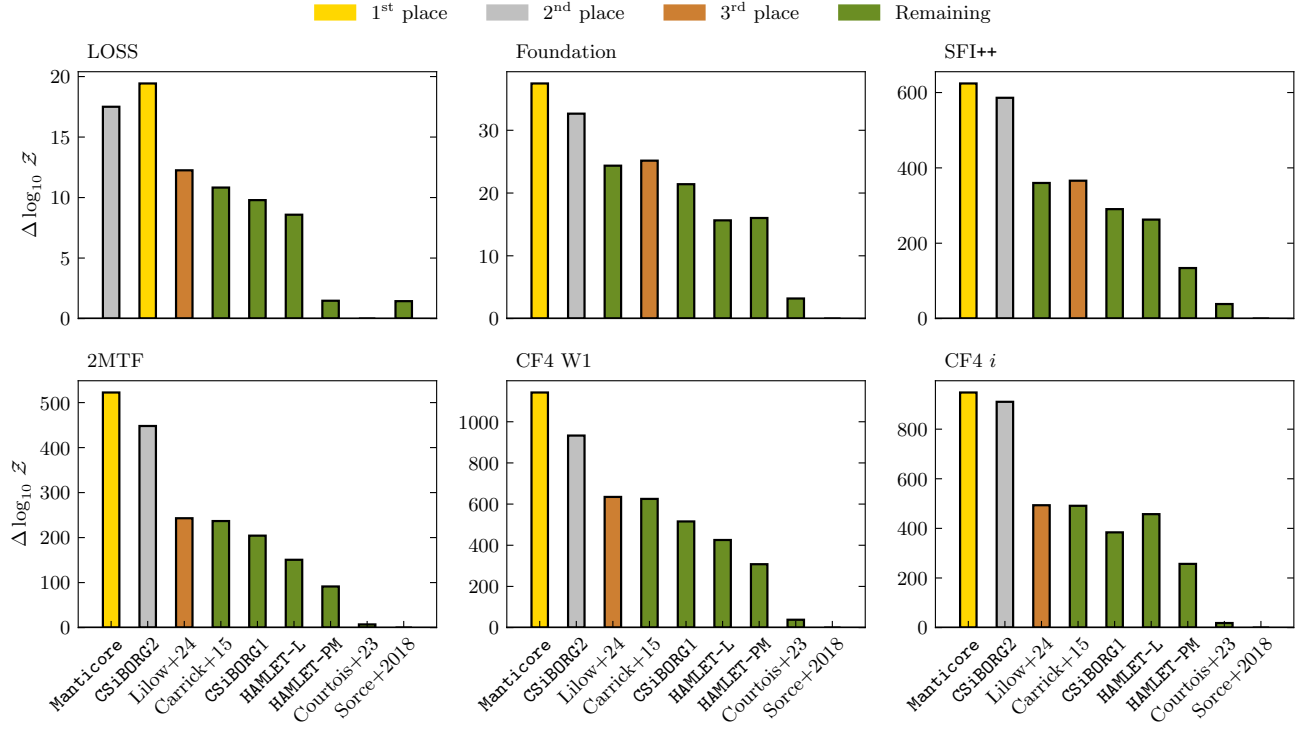


Fig. G.1. The results of the original Velocity Fields Olympics are presented here¹⁰, augmented by our linear and non linear reconstructions (Valade et al. 2024, this work), as well as the last constrained simulations of the AQUILA consortium, namely Manticore-Local (McAlpine et al. 2025). Our methods tighten the gap between velocity-based and redshift-based reconstructions by performing markedly better than all other tested velocity-based methods while remaining below the scores of redshift-based reconstructions.

Table G.1. Summary of the LUM matched halos

Cluster	SGX	SGY [Mpc/h]	SGZ	M_{200c} [$10^{14} M_{\odot}/h$]	M_{500c} [$10^{14} M_{\odot}/h$]	p value 10^{-3}	separation [Mpc/h]
Virgo	-2.3	12.3	0.2	$1.2^{+0.8}_{-0.6}$	$0.8^{+0.6}_{-0.5}$	$6.6^{+5.1}_{-2.9}$	$3.2^{+0.7}_{-0.7}$
Fornax	0.0	-10.1	-8.8	$1.3^{+0.5}_{-0.4}$	$0.9^{+0.5}_{-0.3}$	$8.6^{+8.7}_{-6.0}$	$3.2^{+2.1}_{-1.2}$
Centaurus	-28.2	13.0	-11.1	$4.1^{+2.2}_{-2.7}$	$2.6^{+1.5}_{-1.7}$	$4.0^{+6.1}_{-2.6}$	$5.7^{+2.9}_{-2.3}$
Hydra	-23.3	22.4	-22.0	$3.1^{+1.5}_{-1.2}$	$2.1^{+1.2}_{-0.9}$	$2.4^{+5.6}_{-1.9}$	$3.9^{+1.6}_{-1.4}$
Norma	-46.5	-5.6	6.6	$4.0^{+3.4}_{-2.2}$	$2.7^{+2.4}_{-1.5}$	$1.0^{+5.3}_{-0.9}$	$3.8^{+2.6}_{-1.5}$
Perseus	52.0	-10.0	-15.0	$2.5^{+3.8}_{-1.4}$	$1.7^{+2.6}_{-0.9}$	$8.8^{+9.0}_{-6.6}$	$4.8^{+8.0}_{-2.1}$
Leo	-7.3	63.4	-15.2	$4.7^{+2.0}_{-2.2}$	$3.2^{+1.2}_{-1.5}$	$10.9^{+5.7}_{-7.1}$	$10.9^{+1.3}_{-5.9}$
Coma	-2.8	69.7	8.6	$2.9^{+4.8}_{-1.7}$	$1.9^{+2.9}_{-1.1}$	$9.8^{+9.0}_{-5.5}$	$6.2^{+12.8}_{-3.0}$
Ophiuchus	-54.3	11.5	62.0	$4.9^{+4.1}_{-2.2}$	$3.4^{+2.7}_{-1.5}$	$10.4^{+8.7}_{-8.5}$	$9.4^{+12.2}_{-3.0}$
Hercules	10.9	54.9	74.5	$5.1^{+4.4}_{-3.1}$	$3.5^{+2.8}_{-2.1}$	$6.0^{+6.7}_{-3.9}$	$7.9^{+11.2}_{-4.1}$
Shapley (A)	-104.5	55.7	6.9	$6.4^{+3.3}_{-3.6}$	$4.4^{+2.0}_{-2.6}$	$2.7^{+8.2}_{-2.5}$	$7.5^{+5.4}_{-4.0}$
Shapley (B)	-124.8	73.7	-3.4	$7.6^{+2.6}_{-3.9}$	$4.8^{+1.9}_{-2.3}$	$3.3^{+9.4}_{-2.6}$	$13.6^{+16.4}_{-9.1}$

Table G.2. Summary of the opt-LUM matched halos

Cluster	SGX	SGY	SGZ	M_{200c}	M_{500c}	p value	separation
	[Mpc/h]			[$10^{14} M_{\odot}/h$]		10^{-3}	[Mpc/h]
Virgo	−5.8	13.5	−0.1	$1.7^{+0.7}_{-0.5}$	$1.2^{+0.5}_{-0.4}$	$1.5^{+5.6}_{-1.3}$	$5.3^{+0.9}_{-1.1}$
Fornax	3.5	−12.4	−6.3	$1.2^{+0.5}_{-0.5}$	$0.8^{+0.3}_{-0.3}$	$2.0^{+6.0}_{-1.5}$	$6.7^{+1.1}_{-1.1}$
Centaurus	−25.3	12.0	−15.1	$3.8^{+2.4}_{-1.8}$	$2.7^{+1.5}_{-1.3}$	$1.1^{+4.1}_{-1.0}$	$9.8^{+5.3}_{-3.4}$
Hydra	−23.5	23.2	−21.8	$2.9^{+1.6}_{-1.1}$	$2.1^{+1.1}_{-0.8}$	$0.3^{+1.8}_{-0.3}$	$4.0^{+1.5}_{-1.2}$
Norma	−46.4	−5.1	7.1	$4.2^{+3.6}_{-2.3}$	$2.8^{+2.5}_{-1.5}$	$1.1^{+6.7}_{-1.0}$	$4.0^{+3.2}_{-1.5}$
Perseus	54.6	5.9	−25.4	$5.3^{+2.6}_{-2.8}$	$3.5^{+1.9}_{-1.8}$	$0.6^{+6.5}_{-0.5}$	$21.6^{+3.1}_{-3.2}$
Leo	−11.6	60.9	−18.4	$3.6^{+2.0}_{-1.9}$	$2.5^{+1.3}_{-1.3}$	$0.2^{+2.6}_{-0.2}$	$11.7^{+1.2}_{-1.3}$
Coma	−22.6	65.6	15.7	$4.8^{+2.1}_{-1.5}$	$3.3^{+1.4}_{-1.1}$	$0.2^{+0.7}_{-0.2}$	$23.7^{+2.5}_{-2.2}$
Ophiuchus	−49.0	18.6	65.3	$4.1^{+4.4}_{-2.6}$	$2.9^{+2.9}_{-1.9}$	$1.7^{+10.7}_{-1.5}$	$18.3^{+3.6}_{-4.2}$
Hercules	−2.1	60.4	74.4	$4.3^{+5.0}_{-3.1}$	$3.0^{+3.0}_{-2.1}$	$3.1^{+8.4}_{-2.8}$	$19.5^{+3.9}_{-2.6}$
Shapley (A)	−106.3	57.5	7.1	$7.7^{+2.5}_{-3.5}$	$5.2^{+1.8}_{-2.7}$	$2.8^{+7.7}_{-2.4}$	$8.6^{+14.3}_{-3.2}$
Shapley (B)	−129.7	72.2	−4.3	$5.8^{+3.4}_{-4.0}$	$3.8^{+2.3}_{-2.8}$	$2.9^{+5.9}_{-2.5}$	$10.8^{+7.6}_{-4.2}$