Constrained simulations of the local Universe as a laboratory for precision cosmology and astrophysics



Richard Stiskalek

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2026

Acknowledgements

This thesis reflects many years of learning, covering only a selection of the topics pursued. Throughout, many people have shared their time and expertise. I am grateful to my supervisors, Harry Desmond, Julien Devriendt, and Adrianne Slyz. I first worked with Harry during an undergraduate summer project, and his feedback, critical and constructive in equal measure (with the odd moment of destruction), has pushed me forward. Julien and Adrianne have been equally supportive, granting me the freedom to pursue my own topics while sharing their expertise in galaxy-formation simulations.

Guilhem Lavaux has taught me a great deal on forward modelling. I thank Pedro Ferreira for his invaluable feedback and advice on navigating cosmology. I am also grateful to Shy Genel for opening new research directions. Earlier, several mentors sparked my interest in research: John Veitch, Chris Messenger, and Collin Capano during my undergraduate years, and Miguel Zumalacárregui and Marius Oancea during my master's. I am grateful to several collaborators, beginning with Jens Jasche, whose discussions on `BORG` have been invaluable. Stuart McAlpine's heroic effort brought the `Manticore` initial-condition reconstructions to fruition. I also thank Tariq Yasin, Deaglan Bartlett, Mike Hudson, Nick Choustikov, Sebastian von Hausegger, Francisco Villaescusa-Navarro, Lucia Perez, Teresa Huang, and Gabriele Montefalcone for their insights and contributions.

I gratefully acknowledge the STFC for funding my studentship, and Balliol College for the Snell Exhibition. I thank my family, especially my parents Bronislav and Irena, for their support throughout. Finally, I thank Maryna for sharing this journey — from Holywell Manor to Summertown, Warsaw, and London — and for her patience, encouragement, and the life we are building together.

Abstract

Much of modern cosmology rests on the assumption that the Universe is well-described by the ΛCDM model, and considerable effort goes into inferring its parameters and identifying tensions between experiments that, if real, would point to new physics. More fundamentally, cosmology is an inference problem: we seek the model, its parameters, and the initial conditions that produced the Universe we observe. The statistics of the initial conditions are now tightly constrained by observations, but their specific realisation is not. Standard analyses sidestep this realisation by compressing the data into summary statistics, which are lossy and discard realisation-level information. Field-level inference retains that information by jointly inferring the cosmological parameters and the initial-condition field, sharpening parameter constraints while also reconstructing the local density and velocity fields, anchoring studies of the distance ladder and galaxy formation to our local Universe. I test and apply the *Bayesian Origin Reconstruction from Galaxies* (BORG), which infers the distribution of initial conditions conditioned on a model, its parameters, and the observed local Universe. I focus on three questions: (i) the precision of the reconstruction of initial conditions, via the consistency of halo properties across posterior samples, (ii) its accuracy against independent data, and (iii) the utility of simulations that use such constrained initial conditions for the local distance ladder, motivated by the ongoing Hubble tension. I argue that these reconstructions of the local Universe are now reliable enough for routine use, and continue to improve rapidly. They address questions that depend on our specific realisation and allow novel tests of astrophysical and cosmological models in the nearby Universe, where the data are richest, before applying them further afield.

Contents

1 Introduction **1**

1.1 The standard cosmological model 2

1.1.1 Expansion history and energy content 2

1.1.2 Primordial perturbations and the power spectrum 4

1.1.3 Observational foundations 5

1.1.4 Dark energy and cosmological tensions 7

1.2 Structure formation 8

1.2.1 Linear perturbation theory 8

1.2.2 The peculiar velocity field 10

1.2.3 Dark matter haloes and galaxy bias 11

1.2.4 Lagrangian perturbation theory and initial conditions 13

1.2.5 Cosmological simulations 14

1.3 Bayesian inference 17

1.3.1 Bayes' theorem and the posterior 17

1.3.2 Model comparison 18

1.3.3 Markov chain Monte Carlo 20

1.3.4 Hamiltonian Monte Carlo 21

1.3.5 The No-U-Turn Sampler 22

1.3.6 Automatic differentiation 22

1.4 The initial conditions problem and BORG 23

1.4.1 The inverse problem 23

1.4.2 Early approaches and the constrained-simulation landscape 25

1.4.3 Bayesian forward modelling . . . 26
1.4.4 The 2M++ galaxy catalogue . . . 27
1.4.5 The BORG algorithm . . . 28
1.4.6 Constrained Simulations in BORG: CSiBORG . . . 30
1.5 Peculiar velocities . . . 33
1.5.1 Distance indicators . . . 35
1.5.2 Peculiar velocity surveys . . . 40
1.5.3 The peculiar velocity inference framework . . . 41
1.5.4 Selection effects . . . 46
1.5.5 The cosmic distance ladder and the H_0 tension . . . 49
1.6 Thesis outline . . . 52

2 Evaluating the variance of individual halo properties in constrained cosmological simulations 54

2.1 Introduction . . . 54
2.2 Simulated data . . . 56
2.2.1 CSiBORG . . . 56
2.2.2 Quijote . . . 57
2.2.3 Halo catalogues . . . 58
2.3 Methodology . . . 59
2.4 Results . . . 64
2.4.1 Overlaps . . . 66
2.4.2 Halo properties of overlapping haloes . . . 70
2.5 Discussion . . . 76
2.5.1 Interpretation of the results . . . 76
2.5.2 Comparison with the literature . . . 78
2.6 Conclusion . . . 81

3 The Velocity Field Olympics **83**

3.1 Introduction . . . 83

3.2 Data . . . 84

3.2.1 Local Universe reconstructions . . . 84

3.2.2 Peculiar velocity samples . . . 91

3.3 Methodology . . . 95

3.3.1 Distance measures . . . 97

3.3.2 Joint flow & distance calibration . . . 98

3.3.3 Evidence calculation . . . 104

3.4 Results . . . 107

3.4.1 Evidence comparison . . . 110

3.4.2 Magnitude of redshift residuals . . . 111

3.4.3 Consistency of the velocity scaling . . . 113

3.4.4 Magnitude of the external velocity . . . 114

3.4.5 Bulk flow curve and Local Group velocity . . . 115

3.4.6 Growth rate measurement . . . 119

3.5 Discussion . . . 123

3.5.1 Summary of results . . . 123

3.5.2 Interpretation of results . . . 127

3.5.3 Caveats . . . 130

3.5.4 Comparison with literature . . . 132

3.5.5 Future work . . . 133

3.6 Conclusion . . . 134

3.A SN flow model . . . 136

4 Measurement of the Hubble constant from Cepheids alone **140**

4.1 Introduction . . . 140

4.2 Data . . . 143

4.3 Methodology . . . 149
4.3.1 Bayesian forward modelling framework . . . 149
4.3.2 Peculiar velocity modelling . . . 153
4.3.3 Selection function modelling . . . 154
4.4 Results . . . 161
4.4.1 Host distances . . . 161
4.4.2 Hubble constant inference . . . 163
4.5 Discussion . . . 168
4.5.1 Importance of selection function modelling . . . 172
4.5.2 Comparison to Kenworthy et al. 2022 . . . 173
4.5.3 Ramifications of the results and future work . . . 180
4.6 Conclusion . . . 183
4.A Mock data bias tests . . . 186
4.B Local Universe reconstructions . . . 188
4.B.1 The `Manticore-Local` reconstruction . . . 189
4.B.2 The Carrick et al. 2015 reconstruction . . . 191
4.C Theory peculiar velocity covariance . . . 193
4.D Cepheid host properties . . . 194

5 Conclusions **196**

Bibliography **203**

List of Figures

2.1 FOF halo mass function in CSiBORG and Quijote. 58

2.2 Lagrangian patches at $z = 69$ for two haloes in CSiBORG. 61

2.3 Lagrangian patch overlaps in CSiBORG and Quijote. 65

2.4 Fraction of CSiBORG haloes exceeding Quijote overlap thresholds. 67

2.5 Symmetric maximum overlap fraction $f_{\rm sym}$ in CSiBORG. 68

2.6 Probability of halo matching across IC realisations. 69

2.7 Summed versus maximum Lagrangian patch overlap in CSiBORG. 70

2.8 Overlap-weighted mean separation of matched haloes. 71

2.9 Distributions of matched halo mass, spin, and concentration. 72

2.10 Most likely mass of matched haloes in CSiBORG. 73

2.11 Peculiar velocity alignment of matched haloes in CSiBORG. 74

2.12 Comparison of matching procedures with Hutt et al. (2022). 79

3.1 Density field slices for local Universe reconstructions. 92

3.2 Reconstructed LOS velocities towards the Virgo cluster. 93

3.3 Redshift distributions of peculiar velocity samples. 93

3.4 Bayesian evidence comparison of velocity field reconstructions. 109

3.5 Effect of smoothing on evidence and σ_v. 111

3.6 Inferred σ_v across reconstructions and peculiar velocity samples. 112

3.7 Inferred velocity field scaling parameter β. 112

3.8 Inferred external flow $|\boldsymbol{V}_{\rm ext}|$. 114

3.9 Simulation-frame bulk flow as a function of radius. 114

3.10 CMB-frame bulk flow magnitude and direction. 117
3.11 Predicted Local Group velocity in the CMB frame. 119
3.12 Non-linear to linear σ_8 mapping. 121
3.13 Velocity field scaling $\beta^\star$ of Carrick et al. (2015). 122
3.14 Inferred S_8 compared to literature. 124
3.15 Per-galaxy likelihood differences between CSiBORG2 and Courtois et al. (2023). 129
3.16 Comparison of full and approximate SN likelihood posteriors. 139

4.1 SN magnitude and redshift distributions for 35 Cepheid hosts. 148
4.2 Directed acyclic graph of the Cepheid distance ladder model. 152
4.3 ΛCDM peculiar velocity correlation matrix for Cepheid hosts. 154
4.4 Inferred anchor distances under different priors and selection models. . . . 163
4.5 Distance modulus shifts for 35 hosts under different priors and selections. . 164
4.6 Hubble diagram for 35 Cepheid host galaxies. 164
4.7 Inferred H_0 as a function of the SN-to-redshift selection split. 169
4.8 H_0 posteriors under SN magnitude and redshift selection. 169
4.9 Parameter constraints from `Manticore-Local` under SN and redshift selection. 170
4.10 Stacked H_0 comparison with literature values. 171
4.11 Peculiar velocity correlation matrix for $k > 0.5\,h\,\mathrm{Mpc}^{-1}$ modes only. 179
4.12 Mock H_0 bias test under SN magnitude and redshift selection. 187
4.13 LOS peculiar velocities from `Manticore-Local` and Carrick et al. (2015). 192

1

Introduction

The Λ-cold dark matter (ΛCDM) model describes the evolution of the Universe from its earliest moments to the present day, and its future, with a handful of parameters (Planck Collaboration et al., 2020b). These parameters govern the background cosmology and the primordial power spectrum, but they determine only the statistics of structure formation, not the specific outcome: the cosmic web of filaments, walls, and voids observed in our neighbourhood is one realisation drawn from an ensemble of statistically equivalent possibilities. Understanding it in detail requires reconstructing the specific initial conditions that gave rise to it (Hoffman & Ribak, 1991).

Over the past century, galaxy surveys have mapped this structure in progressively greater detail (for reviews see e.g. Peebles, 1980; Springel et al., 2006). The Harvard–Smithsonian Center for Astrophysics redshift survey first revealed this cosmic web observationally (de Lapparent et al., 1986), demonstrating the imprint of gravitational instability acting on primordial density perturbations (Bond et al., 1996; Davis et al., 1985), now known to be nearly Gaussian (Planck Collaboration et al., 2020c). Subsequent observations, including the 2dF Galaxy Redshift Survey (Colless et al., 2001), the Sloan Digital Sky Survey (SDSS; York et al., 2000), and the 2MASS Redshift Survey (Huchra et al., 2012), extended this census to hundreds of thousands of galaxies across the sky. The next generation, notably DESI (DESI Collaboration et al., 2024), the Vera C. Rubin Observatory (Ivezić et al., 2019), Euclid (Euclid Collaboration et al., 2025), and Roman (Akeson et al., 2019), will increase

sample sizes by orders of magnitude.

In our local neighbourhood, galaxy redshift surveys, compiled into all-sky catalogues such as 2M++ (Lavaux & Hudson, 2011), and peculiar velocity programmes such as Cosmicflows (Tully et al., 2013, 2023) have now mapped the local density and velocity fields in considerable detail. These data have conventionally been analysed through summary statistics (power spectra, correlation functions, number counts), but this compression is lossy for the non-Gaussian, non-linear late-time field: higher-order correlations and phase information are discarded (Rimes & Hamilton, 2005; Carron & Neyrinck, 2012; Schmidt, 2025). Field-level inference operates on the initial conditions and evolves them forward, avoiding this loss and enabling questions beyond summary statistics, such as reconstructing the formation history of individual observed structures. Recent algorithmic advances, including sampling of high-dimensional fields and GPU-accelerated N-body solvers, now make this approach computationally tractable for the local Universe.

Constrained simulations, constructed by inferring the posterior of the initial conditions and evolving samples forward with N-body simulations, produce high-resolution digital twins of the local Universe (Hoffman & Ribak, 1991; Lavaux & Jasche, 2016; Jasche & Lavaux, 2019; McAlpine et al., 2025). The reconstructed density and velocity fields enable object-by-object tests that summary statistics cannot capture, from predicting the properties of individual observed clusters to accounting for the local peculiar velocity field in H_0 inference.

1.1 The standard cosmological model

1.1.1 Expansion history and energy content

The standard cosmological model rests on the cosmological principle: the Universe is homogeneous and isotropic on sufficiently large scales. Observations of the galaxy distribution and the cosmic microwave background (CMB) support this assumption on scales above

~100 Mpc (e.g. Hogg et al., 2005; Scrimgeour et al., 2012). The cosmological principle leads to the Friedmann–Lemaître–Robertson–Walker (FLRW) metric,

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\,\mathrm{d}\Omega^2\right], \tag{1.1}$$

where $a(t)$ is the scale factor normalised to unity today, k the spatial curvature, positive, zero, and negative for closed, flat, and open geometries, respectively, and $\mathrm{d}\Omega^2$ the solid-angle element. With the comoving coordinate r in units of length and a dimensionless, dimensional consistency then requires k to carry units of inverse length squared. Keeping k dimensionless instead requires a scale factor with units of length, since $\mathrm{d}s^2$ has dimensions of length squared. Observations are consistent with spatial flatness ($k = 0$; Planck Collaboration et al., 2020b), and the ΛCDM model adopts this as a baseline. Within ΛCDM, the energy content of the Universe comprises radiation (Ω_r), baryonic matter (Ω_b), cold dark matter (DM; Ω_c), and a cosmological constant Λ (Ω_Λ). Here $\Omega_i \equiv \rho_i/\rho_\mathrm{crit}$ is the density parameter normalised by the critical density $\rho_\mathrm{crit} = 3H^2/(8\pi G)$, and the total matter density is $\Omega_\mathrm{m} = \Omega_\mathrm{b} + \Omega_\mathrm{c}$. Baryons contribute ~5% of the present-day energy density, DM ~26%, and Λ the remaining ~69%, driving the late-time accelerating expansion (Planck Collaboration et al., 2020b).

Applying Einstein's field equations to the FLRW metric yields two independent Friedmann equations, which relate the expansion rate $H(t) \equiv \dot{a}/a$ and its time derivative to the energy content. Here, an overdot denotes differentiation with respect to cosmic time. The first Friedmann equation reduces to the constraint $\sum_i \Omega_i + \Omega_k = 1$, where $\Omega_k \equiv -kc^2/(a^2H^2)$ is the curvature density parameter. The expansion rate as a function of redshift then follows as $H(z) = H_0\sqrt{\Omega_\mathrm{r}(1+z)^4 + \Omega_\mathrm{m}(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}$. The present-day expansion rate is conventionally parametrised as $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, where h is the dimensionless Hubble parameter. Distances and masses throughout this thesis are quoted in units of $h^{-1}\,\mathrm{Mpc}$ and $h^{-1}\,\mathrm{M}_\odot$ to factor out the H_0 dependence. Each component dilutes with

the expansion at a rate set by its equation of state $w_i = p_i/(\rho_i c^2)$: $\rho_{\rm r} \propto a^{-4}$, $\rho_{\rm m} \propto a^{-3}$, $\rho_\Lambda = \text{const}$, giving rise to successive epochs of radiation domination ($z \gtrsim 3400$), matter domination, and dark-energy domination ($z \lesssim 0.3$).

1.1.2 Primordial perturbations and the power spectrum

Quantum fluctuations generated during an early inflationary epoch seed density perturbations (Mukhanov & Chibisov, 1981; Mukhanov, 1985) characterised by the density contrast $\delta(\boldsymbol{x}) \equiv (\rho(\boldsymbol{x}) - \bar{\rho})/\bar{\rho}$. Its Fourier transform,

$$\tilde{\delta}(\boldsymbol{k}) = \int \delta(\boldsymbol{x})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\, \mathrm{d}^3\boldsymbol{x}\,, \tag{1.2}$$

defines the power spectrum $P(k)$, the fundamental two-point statistic,

$$(2\pi)^3\, \delta_{\rm D}(\boldsymbol{k} - \boldsymbol{k}')\, P(k) \equiv \langle \tilde{\delta}(\boldsymbol{k})\, \tilde{\delta}^*(\boldsymbol{k}') \rangle\,, \tag{1.3}$$

where $\langle\cdot\rangle$ denotes the ensemble average over realisations of the random field and $\delta_{\rm D}$ is the Dirac delta function. Inflation predicts a nearly scale-invariant dimensionless primordial curvature power spectrum $\mathcal{P}_{\mathcal{R}}(k) \propto k^{n_s - 1}$, evaluated at the end of the inflationary epoch. The spectral index n_s quantifies the tilt of this spectrum, and the *Planck* team reports $n_s = 0.965 \pm 0.004$ (Planck Collaboration et al., 2020b). The amplitude of the evolved matter fluctuations is conventionally characterised by σ_8, the root-mean-square linear density contrast smoothed on spheres of radius $8\ h^{-1}$ Mpc,

$$\sigma_8^2 = \int \frac{\mathrm{d}k\, k^2}{2\pi^2}\, P(k)\ |W(kR_8)|^2\ , \tag{1.4}$$

where $W(kR_8)$ is the Fourier-space top-hat window function at $R_8 = 8\ h^{-1}\,\text{Mpc}$. Six parameters specify the base ΛCDM model: $\{\Omega_{\rm b}h^2,\ \Omega_{\rm c}h^2,\ \theta_*,\ \tau,\ \ln(10^{10}A_s),\ n_s\}$, from

which all others, including σ_8 and H_0, are derived.[1] Here θ_* is the angular size of the sound horizon at last scattering, i.e. the ratio of $r_s(z_*)$ to the angular diameter distance $d_A(z_*)$. The scalar amplitude A_s sets the overall normalisation of the primordial power spectrum, $\mathcal{P}_\mathcal{R}(k) = A_s\,(k/k_{\rm pivot})^{n_s-1}$ at a pivot scale $k_{\rm pivot} = 0.05\,\mathrm{Mpc}^{-1}$, and thus determines the amplitude of all subsequent density fluctuations. The optical depth to reionisation τ is degenerate with A_s because a fraction $1 - e^{-\tau}$ of CMB photons scatter off the free electrons released at reionisation, damping the primary temperature anisotropies, and large-scale E-mode polarisation is needed to break this degeneracy (Zaldarriaga & Seljak, 1997). These six parameters determine the chronology of cosmic epochs: BBN, matter–radiation equality at $z \simeq 3400$, recombination at $z \simeq 1100$, and reionisation by $z \sim 6$ (Peebles, 1980; Weinberg, 2008; Dodelson & Schmidt, 2020).

1.1.3 Observational foundations

The FLRW metric, combined with the expansion history $H(z)$, determines the distance–redshift relations that underpin cosmological observations. The comoving distance to a source at redshift z is

$$\chi(z) = c \int_0^z \frac{\mathrm{d}z'}{H(z')}\,, \tag{1.5}$$

from which the angular diameter distance

$$d_A = \frac{\chi}{1+z} \tag{1.6}$$

and luminosity distance

$$d_L = (1+z)\,\chi \tag{1.7}$$

[1] In the expansion rate of a flat universe, $H(z) = (100\,\mathrm{km\,s^{-1}\,Mpc^{-1}})\sqrt{\Omega_\mathrm{r}h^2(1+z)^4 + \Omega_\mathrm{m}h^2(1+z)^3 + \Omega_\Lambda h^2}$, it is the combinations $\Omega_\mathrm{b}h^2$ and $\Omega_\mathrm{c}h^2$ that enter; the explicit factor $(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}})$ carries the units, and the radiation density is fixed by the measured CMB temperature $T_0 = 2.7255\,\mathrm{K}$ through $\rho_\gamma \propto T_0^4$. At early times the Ω_Λ term is negligible, so these physical densities alone determine $H(z)$, and hence the comoving sound horizon $r_s(z_*) = \int_{z_*}^{\infty} (c_s/H)\,\mathrm{d}z$. The distance to last scattering $d_A(z_*)$ does depend on h through the late-time Ω_Λ term, and it is through this dependence that the measured angular scale $\theta_* = r_s/d_A(z_*)$ determines H_0 in $\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

follow. In practice, Type Ia supernovae and the distance ladder measure d_L through the distance modulus $\mu = 5\log(d_L/10\,\mathrm{pc})$, while baryon acoustic oscillations constrain $d_A(z)$ and $H(z)$ separately through the transverse and radial acoustic scales. These relations, together with the physics of the early Universe, underpin four independent observational probes that constrain the ΛCDM parameters at the per-cent level (see e.g. Weinberg, 2008; Dodelson & Schmidt, 2020, for comprehensive treatments). The primordial element abundances produced by Big Bang nucleosynthesis depend on $\Omega_\mathrm{b}h^2$ alone, providing an independent constraint on the baryon density (Schöneberg, 2024). The CMB temperature and polarisation anisotropies, measured by *Planck*, constrain the baryon and DM densities, the acoustic scale, and the optical depth, yielding $H_0 = 67.4 \pm 0.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.315 \pm 0.007$, and $\sigma_8 = 0.811 \pm 0.006$ (Planck Collaboration et al., 2020b). Complementary ground-based experiments now surpass *Planck* at small angular scales: ACT DR6 provides significantly more sensitive polarisation measurements (Louis et al., 2025), and SPT-3G delivers the most precise high-multipole TE and EE spectra to date (Camphuis et al., 2026), both yielding ΛCDM parameters consistent with *Planck*. Baryon acoustic oscillations, imprinted at the sound horizon $r_\mathrm{d} \simeq 147\,\mathrm{Mpc}$, provide a standard ruler in the galaxy distribution that constrains the distance–redshift relation at intermediate redshifts, first detected by Eisenstein et al. (2005) and now measured to sub-per-cent precision by DESI (Adame et al., 2025a; Abdul Karim et al., 2025). Type Ia supernovae, standardised via the width–colour–luminosity relation (Phillips, 1993; Tripp, 1998), map the luminosity distance out to $z \sim 2$ and originally established the accelerating expansion (Riess et al., 1998; Perlmutter et al., 1999). Current compilations, including Pantheon+ (Scolnic et al., 2022; Brout et al., 2022), the DES five-year sample (DES Collaboration et al., 2024), and the Union3 analysis (Rubin et al., 2025), now include thousands of supernovae and constrain the dark energy equation of state at the few-per-cent level, with the Vera C. Rubin Observatory's LSST expected to increase the sample by over an order of magnitude (Ivezić et al., 2019).

1.1.4 Dark energy and cosmological tensions

The origin of Λ, however, is unknown. The energy density associated with the cosmological constant is now measured to per-cent-level precision, yet the observed value is some 120 orders of magnitude smaller than naive quantum field theory expectations, a discrepancy known as the cosmological constant problem (Weinberg, 1989). More generally, the dark energy equation of state need not equal $w = -1$ exactly, nor remain constant. The Chevallier–Polarski–Linder (CPL) parametrisation $w(a) = w_0 + w_a(1 - a)$ (Chevallier & Polarski, 2001; Linder, 2003) provides a minimal extension in which w_0 captures the present-day value and w_a its evolution. Recent DESI baryon acoustic oscillation measurements, combined with CMB and supernova data, suggest a 3–4σ preference for $w_0 > -1$ and $w_a < 0$ over a cosmological constant (Abdul Karim et al., 2025), though this significance is sensitive to the supernova calibration adopted and drops to $\sim 3\sigma$ with revised photometric calibrations (Popovic et al., 2026). Bayesian model comparison further softens this claim, finding no significant preference for dynamical dark energy over ΛCDM (Ong et al., 2025, 2026). The Hubble tension, a $\sim 5\sigma$ discrepancy between local and CMB-inferred values of H_0, is discussed in detail in Section 1.5.5. Another, though less statistically significant, discrepancy appears in the amplitude of matter fluctuations: weak-lensing and galaxy-clustering measurements of $S_8 \equiv \sigma_8\sqrt{\Omega_\mathrm{m}/0.3}$ fall 2–3σ below the CMB prediction (Heymans et al., 2021; Abbott et al., 2022; Wright et al., 2025). A potentially more severe anomaly is the cosmic dipole: the number-count dipole of distant quasars and radio galaxies exceeds the kinematic expectation from the CMB at $\gtrsim 5\sigma$, challenging the assumed isotropy of the large-scale matter distribution (Secrest et al., 2021, 2022; Wagenveld et al., 2023; Secrest et al., 2025; von Hausegger et al., 2026; Oayda & Lewis, 2026). On smaller, galactic scales, the standard model faces several challenges (see Bullock & Boylan-Kolchin, 2017, for a review). Empirical scaling relations between baryonic and dynamical properties of galaxies, such as the Tully–Fisher relation (Tully & Fisher, 1977),

its baryonic extension (McGaugh et al., 2000), and the radial acceleration relation (McGaugh et al., 2016), are tight and regular, yet ΛCDM hydrodynamical simulations reproduce them only imperfectly (Desmond et al., 2017; Keller & Wadsley, 2017; Desmond, 2023; Stiskalek & Desmond, 2023), though some simulations suggest that these relations emerge naturally (Ludlow et al., 2017; Navarro et al., 2017).

1.2 Structure formation

1.2.1 Linear perturbation theory

The nearly scale-invariant primordial spectrum introduced in Section 1.1.2 is modified as modes enter the Hubble horizon at different epochs and experience different growth histories (Mukhanov & Chibisov, 1981; Hawking, 1982; Starobinsky, 1982; Guth & Pi, 1982; Bardeen et al., 1986). Gravitational instability amplifies the primordial perturbations, eventually forming the filaments, walls, and nodes of the cosmic web (Bertschinger & Jain, 1994; Bertschinger, 1998; Bond et al., 1996).

Modes that enter the Hubble horizon during radiation domination grow only logarithmically, because the radiation-dominated expansion is so rapid that the subdominant matter cannot collapse under its own gravity within a Hubble time, whereas those entering during matter domination grow unimpeded, imprinting a scale-dependent transfer of power. The transfer function $T(k)$ encodes this scale dependence, defined as the ratio of the late-time linear amplitude of a mode to its primordial value. It is obtained by solving the coupled Boltzmann and Einstein equations for photons, baryons, dark matter, and neutrinos through horizon entry (Ma & Bertschinger, 1995; Eisenstein & Hu, 1998), and is normalised so that $T(k) \to 1$ as $k \to 0$. The linear matter power spectrum then takes the form

$$P(k, z) \propto k^{n_s}\, T^2(k)\, D^2(z)\,, \tag{1.8}$$

where $D(z)$ is the linear growth factor defined below and the normalisation is set by σ_8. The scaling k^{n_s} rather than k^{n_s-1} arises because the Poisson equation relates $\delta \propto k^2\Phi$ in Fourier space, contributing k^4 to the power spectrum, while converting the dimensionless primordial spectrum $\mathcal{P}_\mathcal{R}(k) \propto k^{n_s-1}$ to the dimensional $P(k)$ introduces a factor of k^{-3}, giving a net additional power of k. This suppressed growth of sub-horizon modes during radiation domination (the Mészáros effect; Meszaros 1974) produces a characteristic turnover at $k_{\rm eq} \simeq 0.01\,h\,\mathrm{Mpc}^{-1}$. Before recombination, photon pressure in the tightly coupled photon–baryon plasma drives acoustic oscillations that propagate outward from each initial overdensity. At decoupling ($z \simeq 1100$), these sound waves freeze out at the sound horizon, imprinting a characteristic scale in both the CMB anisotropies and the matter distribution. The baryon acoustic oscillations, already introduced as an observational probe in Section 1.1.3, imprint per-cent-level wiggles on $P(k)$ at $k \sim 0.05$–$0.3\ h\,\mathrm{Mpc}^{-1}$. In practice, modern analyses compute $T(k)$ by numerically integrating the coupled Boltzmann–Einstein hierarchy using codes such as CAMB (Lewis et al., 2000) and CLASS (Blas et al., 2011). The analytic approximation of Eisenstein & Hu (1998) remains accurate to a few per cent.

The density contrast obeys a second-order growth equation. In the linear regime ($|\delta| \ll 1$) it reads as (Peebles, 1980)

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3}{2}\Omega_\mathrm{m}(a)\,H^2\delta = 0\,, \tag{1.9}$$

whose growing-mode solution $D(a)$, conventionally normalised to unity at $a = 1$, factorises the spatial and temporal dependence: $\delta(\boldsymbol{x}, a) = D(a)\,\delta_0(\boldsymbol{x})$. The Hubble drag term $2H\dot{\delta}$ governs the epoch dependence: during matter domination $D \propto a$, while during Λ-domination the accelerating expansion suppresses growth. The dimensionless growth rate $f \equiv \mathrm{d}\ln D/\mathrm{d}\ln a$ is well approximated by $f \approx \Omega_\mathrm{m}(z)^\gamma$ with growth index $\gamma \simeq 0.55$ for ΛCDM (Wang & Steinhardt, 1998; Linder, 2005), while modified-gravity theories predict different values of γ (Dvali et al., 2000; Linder & Cahn, 2007).

1.2.2 The peculiar velocity field

Mass conservation in an expanding universe is expressed by the continuity equation,

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot [(1+\delta)\,\boldsymbol{v}] = 0\,, \tag{1.10}$$

where $\boldsymbol{v}$ is the peculiar velocity field. Linearising and combining with the growing-mode solution yields (Peebles, 1980)

$$\boldsymbol{v}(\boldsymbol{k}) = \frac{iaHf}{k^2}\,\boldsymbol{k}\,\delta(\boldsymbol{k})\,, \tag{1.11}$$

so that $\boldsymbol{v}$ is irrotational and fully determined by δ in the linear regime. In real space, this corresponds to

$$\boldsymbol{v}(\boldsymbol{r}) = \frac{Hf}{4\pi}\int \mathrm{d}^3\boldsymbol{r}'\,\delta(\boldsymbol{r}')\,\frac{\boldsymbol{r}'-\boldsymbol{r}}{|\boldsymbol{r}'-\boldsymbol{r}|^3}\,, \tag{1.12}$$

where each position contributes with a $1/r^2$ gravitational kernel. The $1/k^2$ kernel in Eq. (1.11) suppresses small-scale power, so the velocity field is dominated by large-scale modes and remains more linear than δ at a given scale. Because galaxy redshifts probe both the Hubble flow and the radial peculiar velocity, $cz \approx H_0 r + v_r$, the peculiar velocity field distorts the mapping from redshift to distance. Under the plane-parallel approximation, a galaxy at true distance r is displaced in redshift space to $s = r + v_r/(aH)$, compressing structures that are collapsing (the Kaiser effect on large scales; Kaiser, 1987) and elongating virialised clusters into "Fingers of God" on small scales (Jackson, 1972). On linear scales, the Kaiser effect modifies the power spectrum in a characteristic, angle-dependent way. Because the line-of-sight displacement is proportional to the velocity divergence, which in turn traces $f\delta$, the redshift-space density contrast acquires an angular dependence on $\mu_k \equiv \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}$, the cosine of the angle between the wavevector and the line of sight. The

redshift-space power spectrum becomes (Kaiser, 1987)

$$P^s(k, \mu_k) = (1 + \beta\, \mu_k^2)^2\, P(k)\,, \tag{1.13}$$

where $\beta = f/b$, b is the linear galaxy bias relating the galaxy overdensity to the matter overdensity via $\delta_g = b\,\delta$, and $P(k)$ is the real-space galaxy power spectrum. The μ_k^2 term boosts power along the line of sight, making the two-dimensional correlation function appear "squashed" on large scales. Measuring this anisotropy constrains β and hence $f\sigma_8$, making redshift-space distortions (RSDs) a primary probe of the rate of structure growth (Guzzo et al., 2008; Blake et al., 2011). Equations (1.11) and (1.12) are, however, strictly linear-theory results: they assume $|\delta| \ll 1$ and neglect the coupling between density and velocity that drives shell-crossing, virialisation, and the generation of vorticity. In the non-linear regime, the full continuity equation couples the velocity field to the non-linear density contrast; it depends on the full history of surrounding structures, and it acquires a rotational component sourced by multi-streaming in collapsed regions. Modelling these effects requires N-body simulations (Section 1.2.5), which follow particle trajectories through shell-crossing and into the virialised regime, producing the fully non-linear velocity field as the mass-weighted mean of particle velocities.

1.2.3 Dark matter haloes and galaxy bias

On scales where $\delta \gtrsim 1$, linear theory breaks down and perturbations collapse into virialised DM haloes. Haloes have no sharp physical boundary: halo "mass" is defined by the choice of boundary. The most common convention draws a sphere of radius R_Δ centred on the density peak, within which the mean enclosed density equals Δ times a reference density, either the critical density $\rho_{\rm crit}$ or the mean matter density $\bar{\rho}$. Standard choices include $\Delta = 200$ with respect to $\rho_{\rm crit}$ (giving M_{200c} and R_{200c}), $\Delta = 500$, and the redshift-dependent virial overdensity predicted by the top-hat spherical collapse model (Bryan & Norman,

1998). The spherical collapse model predicts collapse when the linearly extrapolated density contrast reaches $\delta_c \simeq 1.686$ in an Einstein–de Sitter universe, with only a weak dependence on cosmology in ΛCDM (Gunn & Gott, 1972; Nakamura & Suto, 1997). Combining this threshold with Gaussian statistics yields the Press & Schechter (1974) halo mass function, later refined by ellipsoidal-collapse corrections (Sheth et al., 2001) and simulation calibrations (Tinker et al., 2008).

The internal structure of cold DM haloes is well described by the Navarro–Frenk–White (NFW) density profile (Navarro et al., 1996),

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2}\,, \tag{1.14}$$

where r_s is the scale radius at which the logarithmic slope equals -2 and ρ_s is a characteristic density. The NFW profile is a two-parameter family: specifying the halo mass M_{200c} (which fixes R_{200c}) and the concentration $c \equiv R_{200c}/r_s$ determine the profile completely. The concentration correlates inversely with halo mass and positively with formation time, with earlier-forming haloes being more concentrated (see Wechsler & Tinker, 2018, for a review). Beyond mass and concentration, a second fundamental halo property is the dimensionless spin parameter (Bullock et al., 2001),

$$\lambda = \frac{J_{200c}}{\sqrt{2}\, M_{200c}\, V_{200c}\, R_{200c}}\,, \tag{1.15}$$

where J_{200c} is the angular momentum of the mass enclosed within R_{200c} and $V_{200c}^2 = GM_{200c}/R_{200c}$ is the circular velocity at that radius. The spin quantifies how rotationally supported a halo is and follows an approximately log-normal distribution with median $\lambda \approx 0.035$, largely independent of mass, redshift, and cosmology (Bullock et al., 2001). Mass and position are determined primarily by the initial density peak and large-scale flows. Spin is seeded in the linear regime by tidal torques from the misalignment between the proto-halo's inertia tensor and the surrounding tidal field (Peebles, 1969; White, 1984), but is

substantially modified by subsequent mergers. Concentration correlates with formation time (Navarro et al., 1996), which depends on the stochastic accretion history and is thus sensitive to non-linear evolution.

Baryons cool within halo potential wells and form galaxies, whose spatial distribution traces the underlying matter field up to a bias factor b: on sufficiently large scales $\delta_{\rm g} \simeq b\,\delta$ (Kaiser, 1984), with extensions to non-linear and stochastic bias described through perturbative frameworks (Desjacques et al., 2018), or calibrated empirically through halo occupation distribution or subhalo abundance matching (Wechsler & Tinker, 2018). The bias is not only non-linear but also non-local: the halo field cannot be adequately predicted from the local matter density alone (Bartlett et al., 2024b). The bias depends on luminosity, colour, and environment, and at the level of two-point statistics, galaxy clustering constrains only the degenerate combination $b\sigma_8$, since the observed galaxy power spectrum scales as $b^2 P(k)$. Peculiar velocity analyses partially sidestep this limitation: the velocity field is sourced by the total matter distribution and constrains $f\sigma_8$ independently of b. Peculiar velocities are not observed directly but inferred by comparing a distance-indicator distance to the measured redshift, which still requires an assumption about galaxy bias.

1.2.4 Lagrangian perturbation theory and initial conditions

The Zel'dovich approximation (Zel'dovich, 1970) provides an analytic bridge between linear theory and the onset of non-linear collapse. It maps each mass element from its initial (Lagrangian) position $\boldsymbol{q}$ to its (Eulerian) position via

$$\boldsymbol{x}(t) = \boldsymbol{q} + D(t)\,\boldsymbol{\Psi}(\boldsymbol{q})\,, \tag{1.16}$$

where $\boldsymbol{\Psi}(\boldsymbol{q}) = -\nabla_{\boldsymbol{q}}\Phi(\boldsymbol{q})$ is the displacement field determined by the initial gravitational potential. This first-order Lagrangian perturbation theory is accurate until shell-crossing, the moment at which particle trajectories intersect, so that the Zel'dovich map from $\boldsymbol{q}$ to

$\boldsymbol{x}$ in Eq. (1.16) becomes many-to-one, and the density formally diverges. Collapse occurs preferentially along the axis of greatest compression, producing flattened structures known as Zel'dovich pancakes. Subsequent collapse along the second and third axes forms the filaments and nodes of the cosmic web. Second-order Lagrangian perturbation theory (2LPT; Bouchet et al., 1995; Crocce et al., 2006) extends the Zel'dovich approximation by accounting for the tidal deformation of collapsing regions, yielding more accurate displacements and velocities that are often used to initialise N-body simulations at lower redshifts. In practice, initial conditions are constructed by sampling Fourier modes from $P(k)$: each mode is a complex Gaussian with real and imaginary parts drawn independently from $\mathcal{N}(0, P(k)/2)$, so that phases $\phi(\boldsymbol{k})$ are uniform on $[0, 2\pi)$ and $\langle|\tilde{\delta}(\boldsymbol{k})|^2\rangle = P(k)$. The Zel'dovich or 2LPT displacement then converts these modes into particle positions and velocities. The phases encode the spatial arrangement of structures, determining where filaments, clusters, and voids form, while the amplitudes set their contrast. A standard simulation draws both amplitudes and phases at random, producing a realisation that is statistically consistent with ΛCDM but does not reproduce any particular observed structure. Ensemble statistics such as the two-point correlation function or the halo mass function converge over many realisations and do not require knowledge of the phases. Studying the formation history of a specific observed structure, however, requires a simulation that reproduces the observed large-scale density field. This motivates *constrained* simulations, in which the phases are inferred from observational data rather than drawn at random.

1.2.5 Cosmological simulations

Beyond shell-crossing, analytic methods break down and N-body simulations are required to follow the subsequent non-linear evolution (see Dehnen & Read, 2011; Angulo & Hahn, 2022, for reviews). Shell-crossing also marks the onset of practical irreversibility within collapsed structures: repeated phase-space folding drives violent relaxation (Lynden-Bell, 1967) and phase mixing, whereby particles on slightly different orbits gradually drift apart

over many orbital periods, scrambling the coarse-grained phase-space distribution inside haloes and rendering direct time-reversal numerically intractable due to the exponential divergence of nearby trajectories. The large-scale density field, however, retains the imprint of the initial conditions. Cosmological N-body simulations discretise the dark matter distribution into N particles and follow their trajectories under gravity from the initial conditions to the present day. The gravitational force is computed using one of several approaches, differing by the trade-off between speed and force resolution. At one extreme, the particle-mesh (PM) method (Efstathiou et al., 1985) computes forces on a fixed grid and is fast but limited in resolution; at the other, tree-based (Barnes & Hut, 1986; Springel, 2005) and adaptive-mesh (Kravtsov et al., 1997; Teyssier, 2002) codes resolve forces down to a softening scale orders of magnitude smaller than the grid spacing, at greater computational cost. The choice of gravity solver determines which physical scales can be modelled: PM simulations capture the large-scale density and velocity fields but not the internal structure of haloes, whereas full N-body simulations resolve halo profiles, substructure, and virialised dynamics.

With few time steps, PM cannot accurately track the full particle displacement, which grows monotonically with time. COLA (Tassev et al., 2013) circumvents this by solving not for the full particle trajectories but for the residual displacement relative to an analytic LPT solution. The LPT component, typically second-order, accounts for the bulk of the large-scale displacement analytically, so the integrator need only evolve the non-linear residual. This separation allows COLA to produce a more accurate density and velocity field with fewer time steps than a standard PM simulation would require.

PM forces are resolved only down to the grid spacing Δx; while halo structure can in principle be captured with sufficiently fine grids, the computational cost of doing so typically makes PM impractical for resolving halo interiors. COLA shares this spatial limitation, and its coarse time-stepping further prevents accurate modelling of virialised structures. Full N-body simulations overcome this by resolving gravitational interactions well below Δx. Tree

algorithms (Barnes & Hut, 1986) group distant particles into cells of a hierarchical oct-tree and approximate their collective gravitational effect via multipole expansions, evaluating the force to arbitrary accuracy by controlling the opening angle that determines when a cell is sufficiently distant to be treated as a single source. This yields a force resolution limited only by the gravitational softening length ϵ, typically a small fraction of the mean interparticle separation. Hybrid tree-PM schemes combine both approaches: the PM grid handles the long-range component of the force via fast Fourier transforms (FFTs), while the tree computes the short-range residual, achieving both the speed of PM on large scales and the resolution of the tree on small scales (Springel, 2005). The time-stepping is also adaptive: particles in dense, rapidly evolving regions (e.g. halo cores) are advanced with shorter time steps than those in the low-density field. Widely used codes include GADGET-2 (Springel, 2005) and its successor GADGET-4 (Springel et al., 2021), which employ the tree-PM approach, and RAMSES (Teyssier, 2002), which uses adaptive mesh refinement to recursively refine the grid in high-density regions, achieving comparable adaptive resolution through a different strategy. Both families of codes also support hydrodynamics (GADGET via smoothed-particle hydrodynamics and RAMSES via its grid-based Godunov solver), but in this thesis only their gravity solvers are used. Hydrodynamical simulations solve the equations of gas dynamics and incorporate models for baryonic processes such as radiative cooling, star formation, and feedback from supernovae and active galactic nuclei (e.g. Teyssier, 2002; Springel, 2010; Bryan et al., 2014; Hopkins, 2015).

While N-body simulations provide the machinery to evolve any realisation forward, the *initial conditions problem* consists of identifying which of the infinitely many possible realisations corresponds to our local Universe. For example, the peculiar velocities generated by nearby large-scale structure are a leading systematic in the local distance ladder, biasing the inference of H_0 whenever they are improperly modelled (Riess et al., 2022a; Kenworthy et al., 2022; Stiskalek et al., 2026b), and the reconstructed DM distribution can be used to constrain DM self-interactions and annihilation signals in the local volume (Kostić et al.,

2026; Bartlett et al., 2022). Constructing such a constrained simulation requires, however, inferring the initial phases from observational data.

1.3 Bayesian inference

1.3.1 Bayes' theorem and the posterior

Cosmology is, at its core, an inference problem: extracting the parameters $\boldsymbol{\theta}$ of a physical model from data $\boldsymbol{d}$, and determining which model best describes the data. The Bayesian formalism (see for example Jaynes, 2003) encodes this in Bayes' theorem, written for a model $\mathcal{M}$,

$$\mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d}, \mathcal{M}) = \frac{\mathcal{L}(\boldsymbol{d} \mid \boldsymbol{\theta}, \mathcal{M})\, \pi(\boldsymbol{\theta} \mid \mathcal{M})}{\mathcal{Z}(\mathcal{M})}, \tag{1.17}$$

where $\mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d}, \mathcal{M})$ is the posterior, $\mathcal{L}(\boldsymbol{d} \mid \boldsymbol{\theta}, \mathcal{M})$ the likelihood, $\pi(\boldsymbol{\theta} \mid \mathcal{M})$ the prior, and $\mathcal{Z}(\mathcal{M})$ the evidence (marginal likelihood). The evidence acts as a normalisation constant, independent of $\boldsymbol{\theta}$ but dependent on the model,

$$\mathcal{Z}(\mathcal{M}) = \int \mathcal{L}(\boldsymbol{d} \mid \boldsymbol{\theta}, \mathcal{M})\, \pi(\boldsymbol{\theta} \mid \mathcal{M})\, \mathrm{d}\boldsymbol{\theta}\,. \tag{1.18}$$

$\mathcal{L}$ encodes the noise model and the physics of the observation; π encodes prior knowledge or assumptions about the parameters. Evaluating $\mathcal{L}$ requires a forward model: a generative process that, given parameters $\boldsymbol{\theta}$, predicts observable data that can be compared to the actual observations under a specified noise model.

Most cosmological models have nuisance parameters alongside the quantities of interest. Marginalisation "removes" the nuisance parameters by integrating them out of the joint posterior,

$$\mathcal{P}(\boldsymbol{\theta}_1 \mid \boldsymbol{d}) = \int \mathcal{P}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \mid \boldsymbol{d})\, \mathrm{d}\boldsymbol{\theta}_2\,, \tag{1.19}$$

where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ denote the parameters of interest and nuisance parameters, respectively.

This propagates the uncertainty in $\boldsymbol{\theta}_2$ into the constraints on $\boldsymbol{\theta}_1$, avoiding the bias introduced by fixing nuisance parameters to point estimates.

Hierarchical models extend this framework to populations of objects. Each observed object i has latent parameters $\boldsymbol{\theta}_i$, drawn from a population-level model with hyperparameters $\boldsymbol{\phi}$. The joint posterior then factorises as

$$\mathcal{P}(\boldsymbol{\phi}, \{\boldsymbol{\theta}_i\} \mid \{\boldsymbol{d}_i\}) \propto \pi(\boldsymbol{\phi}) \prod_i \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\theta}_i, \boldsymbol{\phi})\, \pi(\boldsymbol{\theta}_i \mid \boldsymbol{\phi})\,. \tag{1.20}$$

The per-object latent variables can then be marginalised to obtain constraints on the population-level hyperparameters. This structure is central to the peculiar velocity and H_0 analyses in Chapters 3 and 4.

1.3.2 Model comparison

The evidence $\mathcal{Z}$ plays no role in parameter estimation but is central to model comparison. The Bayes factor between two competing models $\mathcal{M}_1$ and $\mathcal{M}_2$,

$$B_{12} = \frac{\mathcal{Z}(\mathcal{M}_1)}{\mathcal{Z}(\mathcal{M}_2)}\,, \tag{1.21}$$

quantifies the relative support for $\mathcal{M}_1$ over $\mathcal{M}_2$, assuming that both models are equally probable a priori. On the Jeffreys scale, $\ln B_{12} > 1$ constitutes substantial evidence, $\ln B_{12} > 2.5$ strong evidence, and $\ln B_{12} > 5$ decisive evidence for $\mathcal{M}_1$ (Jeffreys, 1939). This is an arbitrary scale, so the precise interpretation of a given $\ln B_{12}$ carries no fundamental significance (Trotta, 2008; Nesseris & García-Bellido, 2013). The evidence penalises wasted prior volume, a built-in Occam's razor: a model with superfluous parameters spreads probability over regions the data do not support, driving $\mathcal{Z}$ down unless the extra freedom is genuinely warranted. The evidence depends on the prior, so it has no absolute meaning and must be read together with the model and prior it assumes (Gelman & Yao, 2021). For example, doubling the prior range of one parameter dilutes $\ln \mathcal{Z}$ by $\ln 2$, enough

to cross the 'substantial' threshold on the Jeffreys scale. Computing $\mathcal{Z}$ requires integrating over the entire prior volume, achievable by nested sampling (Skilling, 2006) or by machine-learning approaches that estimate the evidence directly from posterior samples (McEwen et al., 2021; Polanska et al., 2025; Srinivasan et al., 2024). For nested models, the Savage–Dickey density ratio (Dickey, 1971; Verdinelli & Wasserman, 1995) offers a simpler route requiring only posterior samples from the more complex model, provided the posterior has sufficient support in the region of parameter space corresponding to the simpler model.

A frequently raised objection is the prior dependence of Bayesian inference (Trotta, 2008). Prior dependence is intrinsic to the framework: Bayes' theorem updates a prior state of information in light of new data, and the posterior necessarily carries an imprint of both. Uninformative priors, which need not be uniform, isolate the constraining power of a given experiment. This is a valid choice for that purpose, but claiming zero prior knowledge is rarely realistic. Informative priors that encode existing knowledge are more faithful to the Bayesian paradigm but conflate new and old information. In astronomy, strictly adopting the Bayesian information-update approach would implicitly assume the fidelity of previous experiments, an assumption difficult to defend as both experimental and statistical methods evolve.

This tension motivates the use of uninformative priors, but the notion of "uninformative" itself depends on parametrisation: a uniform prior in a variable differs from one uniform in its logarithm, and in cosmological inference this can shift the posteriors substantially (Hergt et al., 2021). Jeffreys' priors (Jeffreys, 1939) or reference priors (Bernardo, 1979) offer a more principled alternative but are not always straightforward to compute or apply. Prior dependence also propagates directly into Bayesian model comparison, since the evidence is the likelihood weighted by the prior and integrated over the parameter space. The choice of prior can therefore influence which model is favoured (Gelman & Yao, 2021).

Every analysis also involves a choice of what to model—raw pixel data, compressed catalogues, or summary statistics. Forward modelling makes this choice explicit: the

generative model specifies precisely which observables are predicted and at what level of realism. The right level of compression is problem-dependent, and its uncertainties do not propagate cleanly.

1.3.3 Markov chain Monte Carlo

Closed-form posteriors require assumptions (e.g. Gaussian likelihoods, conjugate priors, and a linear relation between model parameters and predicted data) that rarely hold in practice. Markov chain Monte Carlo (MCMC; Metropolis et al., 1953; Hastings, 1970) constructs a chain of samples of $\boldsymbol{\theta}$ asymptotically converging to $\mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d})$. A sufficient condition for convergence is detailed balance,

$$\mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d})\, T(\boldsymbol{\theta}' \mid \boldsymbol{\theta}) = \mathcal{P}(\boldsymbol{\theta}' \mid \boldsymbol{d})\, T(\boldsymbol{\theta} \mid \boldsymbol{\theta}')\,, \tag{1.22}$$

where $T(\boldsymbol{\theta}' \mid \boldsymbol{\theta})$ is the transition kernel, i.e. the probability of moving to $\boldsymbol{\theta}'$ given the current state $\boldsymbol{\theta}$. This states that, in equilibrium, the rate at which the chain transitions from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}'$ equals the rate of the reverse transition. The Metropolis–Hastings algorithm satisfies detailed balance by drawing a candidate $\boldsymbol{\theta}'$ from a proposal distribution $q(\boldsymbol{\theta}' \mid \boldsymbol{\theta})$ and accepting it with probability

$$\alpha = \min\left(1,\, \frac{\mathcal{P}(\boldsymbol{\theta}' \mid \boldsymbol{d})\, q(\boldsymbol{\theta} \mid \boldsymbol{\theta}')}{\mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d})\, q(\boldsymbol{\theta}' \mid \boldsymbol{\theta})}\right). \tag{1.23}$$

The simplest choice is a symmetric Gaussian proposal, $q(\boldsymbol{\theta}' \mid \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}' \mid \boldsymbol{\theta}, \Sigma)$, for which the proposal ratio cancels and only the posterior ratio matters. The proposal covariance Σ approximates the posterior shape and is typically adapted during warm-up. Convergence is assessed by running multiple independent chains from random starting points and comparing the between-chain and within-chain variance via the Gelman–Rubin diagnostic $\hat{R}$ (Gelman & Rubin, 1992).

1.3.4 Hamiltonian Monte Carlo

Random-walk MCMC scales as $O(n_d^2)$ evaluations per independent sample, where n_d is the number of parameters, becoming impractical in high dimensions. Hamiltonian Monte Carlo (HMC; Duane et al., 1987; Neal, 2011) sidesteps this bottleneck by recasting sampling as Hamiltonian dynamics on the posterior surface. It promotes $\boldsymbol{\theta}$ to a dynamical variable, augments it with auxiliary momenta $\boldsymbol{p}$, and identifies the negative log-posterior with a potential energy $U(\boldsymbol{\theta}) = -\ln \mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d})$. The total Hamiltonian,

$$H(\boldsymbol{\theta}, \boldsymbol{p}) = U(\boldsymbol{\theta}) + \tfrac{1}{2}\boldsymbol{p}^T M^{-1} \boldsymbol{p}\,, \tag{1.24}$$

defines a joint distribution $p(\boldsymbol{\theta}, \boldsymbol{p}) \propto \exp(-H)$ that factorises into the target posterior times a Gaussian in $\boldsymbol{p}$; marginalising over $\boldsymbol{p}$ therefore recovers $\mathcal{P}(\boldsymbol{\theta} \mid \boldsymbol{d})$. At each iteration, momenta are resampled from $\mathcal{N}(\mathbf{0}, M)$ and H is evolved under Hamilton's equations,

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}t} &= \frac{\partial H}{\partial \boldsymbol{p}} = M^{-1}\boldsymbol{p}\,, \\ \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \boldsymbol{\theta}} = -\nabla_{\boldsymbol{\theta}} U\,, \end{aligned} \tag{1.25}$$

using the symplectic leapfrog integrator for L steps of size $\varepsilon \equiv \Delta t$. The leapfrog conserves H to $O(\varepsilon^2)$, so the endpoint is accepted with probability $\alpha_{\mathrm{HMC}} = \min(1, \exp(-\Delta H))$, yielding high acceptance rates when the step size and trajectory length are well tuned to the posterior geometry. If H were conserved exactly, every proposal would be accepted; the Metropolis step corrects for the leapfrog integration error. The resulting trajectory explores the typical set, suppressing random-walk behaviour: the number of gradient evaluations per independent sample scales as $O(n_d^{1/4})$, more favourably than the $O(n_d^2)$ of random-walk MCMC. Each leapfrog step requires one evaluation of the gradient $\nabla_{\boldsymbol{\theta}} U$, computed either analytically or via automatic differentiation (Section 1.3.6), so a single HMC proposal costs L gradient evaluations compared to one likelihood evaluation for random-walk Metropolis–

Hastings. Choosing L is itself a trade-off: too few steps and the proposal remains close to its starting point, recovering random-walk behaviour; too many and the trajectory doubles back toward its origin, wasting computation.

1.3.5 The No-U-Turn Sampler

Standard HMC requires two tuning parameters: the step size ε and the number of leapfrog steps L. The No-U-Turn Sampler (NUTS; Hoffman & Gelman, 2011) eliminates this tuning by doubling the leapfrog trajectory until it begins to double back on itself; schematically, a U-turn is detected when the momentum at either end of the trajectory points back toward the other end, $\boldsymbol{p}^{\pm} \cdot (\boldsymbol{\theta}^{+} - \boldsymbol{\theta}^{-}) < 0$, where + and − label the forward and backward endpoints. A proposal is selected from the resulting trajectory by multinomial sampling weighted by $\exp(-H)$, preserving detailed balance, and the step size ε is tuned during warm-up to target an acceptance rate of $\sim$0.8. NUTS thereby achieves the efficiency of well-tuned HMC without manual specification of L or ε. Throughout this thesis, posterior sampling is performed with the NUTS implementation in NumPyro (Phan et al., 2019).

1.3.6 Automatic differentiation

Both HMC and NUTS require the gradient $\nabla_{\boldsymbol{\theta}} U$ at every leapfrog step. Automatic differentiation (AD; Gunes Baydin et al., 2015) computes exact derivatives (to machine precision) by applying the chain rule to the elementary operations composing an algorithm,

$$\frac{\partial f}{\partial \boldsymbol{\theta}} = \frac{\partial f_n}{\partial \boldsymbol{z}_{n-1}} \cdot \frac{\partial f_{n-1}}{\partial \boldsymbol{z}_{n-2}} \cdots \frac{\partial f_2}{\partial \boldsymbol{z}_1} \cdot \frac{\partial f_1}{\partial \boldsymbol{\theta}}, \tag{1.26}$$

where $f = f_n \circ f_{n-1} \circ \cdots \circ f_1$ and $\boldsymbol{z}_k = f_k(\boldsymbol{z}_{k-1})$ are intermediate variables. In *forward mode*, this product is accumulated left to right, costing n_d passes for the full gradient. *Reverse mode* (backpropagation) works from the scalar output backward, propagating sensitivities through every operation to all inputs in a single pass, so the full $\nabla_{\boldsymbol{\theta}} f$ costs roughly one

additional forward evaluation regardless of n_d. In PYTHON, JAX (Bradbury et al., 2021) provides composable transformations that differentiate, compile, and vectorise numerical code on CPUs, GPUs, and TPUs, using reverse mode by default; PYTORCH (Paszke et al., 2019) offers a similar capability. NUTS with JAX-based AD enables efficient posterior sampling in high-dimensional spaces.

1.4 The initial conditions problem and BORG

1.4.1 The inverse problem

The local Universe, within ~200 h^{-1} Mpc of the Milky Way, is the only volume in which individual haloes, peculiar velocities, and resolved galaxy properties are simultaneously accessible. It contains rich and well-characterised structures—the Virgo, Coma, and Perseus clusters, the Great Attractor region, the Shapley supercluster, and the Local Void—mapped by galaxy surveys (2MRS, 6dF, SDSS) and by peculiar velocity catalogues (e.g. Cosmicflows-4) that independently probe the gravitational field. Much of observational cosmology is calibrated against local measurements, from the distance ladder to galaxy scaling relations, so a reliable model of this neighbourhood is essential. An unconstrained simulation samples a random realisation of the density field, so any comparison to the observed local structures is limited by cosmic variance and must rely on statistical calibration of scaling relations. Reproducing the specific observed structures enables object-by-object comparison, but it requires reconstructing the initial conditions from which they grew.

Recovering those initial conditions from present-day observations is an inverse problem that becomes ill-posed on scales where shell-crossing occurs, rendering the mapping from initial to final conditions many-to-one and erasing information about pre-collapse trajectories. Observations probe the galaxy field, not the matter field, and the uncertain galaxy bias complicates the inversion further. RSDs compound the difficulty, since the observed redshift $cz \approx H_0 r + v_{\rm pec}$ mixes Hubble flow with peculiar velocity, and separating the two

requires the very density field one is trying to reconstruct. Angular masks, radial selection functions, and the Zone of Avoidance present further complications.

Gravity is time-reversible in principle, and in the linear regime the Zel'dovich displacement can be algebraically inverted to map the present-day field back to its Lagrangian configuration. Step-by-step backward integration of the full equations of motion is less tractable: the present-day data are incomplete (masked regions, selection effects, sparse sampling) and any noise is amplified exponentially when time is reversed. Beyond shell-crossing, phase mixing buries the information needed for reversal on scales below any achievable resolution. The earliest reconstructions therefore relied on linear statistical estimators. The Wiener filter is the linear minimum-variance estimator of a Gaussian random field given noisy data (Zaroubi et al., 1995, 1999). Let the signal be $\boldsymbol{s}$, drawn from a zero-mean Gaussian prior with covariance $S = \langle \boldsymbol{s}\boldsymbol{s}^\dagger \rangle$, and suppose the data are a linear transformation of the signal plus additive noise,

$$\boldsymbol{d} = R\boldsymbol{s} + \boldsymbol{n}\,, \tag{1.27}$$

where R is the response operator encoding the survey geometry (angular mask, radial selection, or any smoothing) and $\boldsymbol{n} \hookleftarrow \mathcal{N}(\boldsymbol{0}, N)$ is the zero-mean Gaussian noise. In the cosmological context, $\boldsymbol{s}$ is the three-dimensional density or velocity field on a grid, R maps this field to the observable galaxy counts or peculiar velocities, and N accounts for shot noise and measurement error. The Gaussian likelihood is $\mathcal{L}(\boldsymbol{d} \mid \boldsymbol{s}) = \mathcal{N}(\boldsymbol{d};\, R\boldsymbol{s},\, N)$ and the Gaussian prior is $\pi(\boldsymbol{s}) = \mathcal{N}(\boldsymbol{s};\, \boldsymbol{0},\, S)$. Since both are Gaussian, the posterior is also Gaussian,

$$\mathcal{P}(\boldsymbol{s} \mid \boldsymbol{d}) = \mathcal{N}(\boldsymbol{s};\, \hat{\boldsymbol{s}}_{\mathrm{WF}},\, D)\,, \tag{1.28}$$

with posterior covariance $D = (S^{-1} + R^\dagger N^{-1} R)^{-1}$ obtained by completing the square in the exponent of the likelihood and prior product. The posterior mean is the Wiener filter

estimate,

$$\hat{\boldsymbol{s}}_{\mathrm{WF}} = D\, R^{\dagger} N^{-1} \boldsymbol{d} = S R^{\dagger} (R S R^{\dagger} + N)^{-1} \boldsymbol{d}\,, \tag{1.29}$$

where the second equality follows from the Woodbury identity. On scales where the signal dominates ($RSR^{\dagger} \gg N$), the filter approaches the inverse mapping $R^{-1}\boldsymbol{d}$; where noise dominates ($N \gg RSR^{\dagger}$), it suppresses the reconstruction towards zero, the prior mean. Because the Wiener filter recovers a smooth, linear field, the Zel'dovich displacement can be inverted to map it back to initial conditions. Constrained realisations, introduced by Hoffman & Ribak (1991), supplement the Wiener filter mean with a random component drawn from the residual covariance, producing samples consistent with both the data and the prior. This decomposition underpins most subsequent Wiener-filter-based constrained-simulation programmes, though not the Bayesian forward-modelling approach discussed below. The idea of sampling Gaussian random fields subject to linear constraints was introduced by Bertschinger (1987) using path-integral methods. Ganon & Hoffman (1993) first applied the Hoffman–Ribak algorithm to peculiar velocity data, recovering the Great Attractor-like structures and other features of the local density field, and Kravtsov et al. (2002); Klypin et al. (2003) produced some of the earliest constrained N-body simulations of the local supercluster using galaxy group catalogues and peculiar velocities as constraints.

1.4.2 Early approaches and the constrained-simulation landscape

The constrained-simulation programme has since expanded considerably. `CLUES` applied the Wiener filter and Hoffman–Ribak algorithm to Cosmicflows-2 peculiar velocities, reproducing nearby structures including the Virgo cluster (Gottloeber et al., 2010; Sorce et al., 2014, 2016a). `SLOW` extended this to magnetohydrodynamics, enabling predictions of X-ray luminosity, the Sunyaev–Zel'dovich signal, and temperature profiles at cluster scales (Dolag et al., 2023). `ELUCID` took a different approach entirely, coupling an HMC sampler with a particle-mesh forward model to reconstruct the initial density field from

the SDSS DR7 galaxy group catalogue and demonstrating halo-by-halo correspondence above $\sim 10^{14}\ h^{-1}\,\mathrm{M}_\odot$ (Wang et al., 2014, 2016, 2018; Yang et al., 2018; Zhang et al., 2022). The Wiener-filter-based programmes share one limitation: the filter is linear and cannot capture the non-Gaussian structure induced by non-linear gravitational evolution. It also treats galaxies as a continuous field with Gaussian noise, which is unphysical in under-dense regions where the expected count per voxel approaches zero: a Gaussian noise model allows negative densities, whereas galaxy counts are non-negative integers drawn from a Poisson-like process. A further limitation is that generating multiple constrained realisations by varying the random seed for unconstrained modes does not constitute posterior sampling (e.g. Sorce et al., 2016b, 2020). The constrained large-scale modes are fixed to the Wiener-filter mean rather than drawn from their posterior, so that only the small-scale modes are randomised and the uncertainty on large scales is neglected. Such ensembles have no well-defined probabilistic interpretation.

1.4.3 Bayesian forward modelling

Rather than replacing the non-linear field with a linear approximation, forward modelling embeds a physical structure formation simulator in the inference: one proposes candidate initial conditions, evolves them forward through a gravity solver of arbitrary complexity, and compares the predicted observables to the data. The target is the full posterior $\mathcal{P}(\delta^{\mathrm{ic}} \mid \boldsymbol{d})$ over the primordial density field, not merely its mean. Following Bayes' theorem, this posterior is simply

$$\mathcal{P}(\delta^{\mathrm{ic}} \mid \boldsymbol{d}) \propto \mathcal{L}(\boldsymbol{d} \mid \delta^{\mathrm{ic}})\, \pi(\delta^{\mathrm{ic}})\,. \tag{1.30}$$

The advantage of inferring the initial conditions rather than the late-time field is that the primordial density field is Gaussian (to excellent approximation), as predicted by inflation and confirmed by *Planck* (Planck Collaboration et al., 2020c), whereas the late-time field is highly non-Gaussian. *Bayesian Origin Reconstruction from Galaxies* (BORG) exploits

this Gaussianity by parametrising the initial field as $\tilde{\delta}^{\rm ic}(\boldsymbol{k}) = \sqrt{P(k)}\,\tilde{\boldsymbol{s}}(\boldsymbol{k})$, where $\sqrt{P(k)}$ encodes the primordial power spectrum and $\boldsymbol{s}(\boldsymbol{q})$ is a white-noise field whose Fourier modes are independently drawn from a standard normal distribution (Jasche & Wandelt, 2013; Lavaux et al., 2019; McAlpine et al., 2025). In Fourier space, the prior on $\boldsymbol{s}$ is then simply $\pi(\boldsymbol{s}) \propto \exp\left(-\frac{1}{2}\boldsymbol{s}^\dagger\boldsymbol{s}\right)$. In this parametrisation, the cosmological information is encoded entirely in the deterministic function $\sqrt{P(k)}$, while the stochastic degrees of freedom, both amplitudes and phases of the initial density field, reside in the white-noise field $\boldsymbol{s}$. What remain to be specified are the data $\boldsymbol{d}$ and the likelihood quantifying the agreement between the forward model's predictions and the observations. In the current generation of BORG, the data are typically galaxy positions in redshift space, though both weak-lensing and peculiar-velocity likelihoods have been explored (Porqueres et al., 2021; Boruah et al., 2022; Prideaux-Ghee et al., 2023).

1.4.4 The 2M++ galaxy catalogue

The BORG reconstructions used in this thesis are conditioned on the 2M++ galaxy catalogue, which is a whole-sky redshift compilation of 69,160 galaxies (Lavaux & Hudson, 2011). Photometry is drawn from the Two Micron All Sky Survey Extended Source Catalogue, with K-band apparent magnitudes corrected for Galactic extinction (Schlegel et al., 1998), k-corrections, and passive luminosity evolution. Redshifts are compiled from three overlapping surveys: the 2MASS Redshift Survey, which provides dense coverage to $K < 11.5$ over 91% of the sky (Huchra et al., 2012); the 6dF Galaxy Redshift Survey, covering the southern hemisphere to $K < 12.5$ (Jones et al., 2009); and SDSS Data Release 7, extending coverage in the northern Galactic cap to the same depth (Abazajian et al., 2009). The catalogue is flux-limited, so the observed number density decreases with distance as intrinsically fainter galaxies drop below the magnitude threshold. The radial selection function at each sky position is computed from a Schechter luminosity function inferred from the sample, giving the fraction of the luminosity function observable above the local apparent

magnitude limit. The Galactic plane obscures ∼10% of the sky. The catalogue has a high completeness within a spherical volume of radius ∼155 h^{-1} Mpc: this volume defines the 2M++ "constrained region" used by most BORG reconstructions based on redshift-space data. Deeper reconstructions covering the SDSS footprint have been pursued (Leclercq et al., 2017; Lavaux et al., 2019) and are now being revisited with newer generations of the BORG algorithm.

1.4.5 The BORG algorithm

The forward model starts from a candidate white-noise field $\boldsymbol{s}$ on a 256^3 grid, deterministically related to the initial density field at $z = 1000$ via the power spectrum as described above, and evolves it to $z = 0$ using the COLA integrator (Tassev et al., 2013; Stopyra et al., 2024a). The resulting real-space density field is then mapped to redshift space using the predicted velocity field and converted to predicted galaxy number counts through a galaxy bias model and the survey selection function. Because the forward model predicts both the density and velocity fields from the same initial conditions, the circularity between distance and velocity is resolved within the inference. These predictions are compared to the observed 2M++ galaxy catalogue via a Poisson likelihood. Specifically, the bias model maps the local redshift-space dark matter density ρ_i to a predicted galaxy count per voxel,

$$\lambda_i = \bar{N}\, R_i\, \frac{B(\rho_i)}{\langle B \rangle}\,, \tag{1.31}$$

where $\bar{N}$ is the mean galaxy count over the survey, R_i is the response operator encoding the angular mask and radial selection function, and $B(\rho)$ is the bias function normalised by its volume average $\langle B \rangle$ to enforce $\langle \delta_{\rm g} \rangle = 0$. The simplest choice is a power law, $B(\rho) \propto \rho^{\alpha}$, with α inferred jointly with the initial conditions (Jasche & Lavaux, 2019). The `Manticore` extension generalises this to a sigmoid-truncated double power law that suppresses galaxy formation below a density threshold and allows a different high-density scaling (McAlpine

et al., 2025). The flexibility of the bias model introduces a degeneracy with the density field, since a range of density–galaxy bias combinations can produce similar predicted counts, though the Gaussian prior on the initial conditions partially breaks this degeneracy. The predicted counts λ_i are compared to the observed galaxy counts N_i in each voxel via a Poisson likelihood,

$$\mathcal{L} = \prod_{\text{voxels}} \frac{\lambda_i^{N_i} \, e^{-\lambda_i}}{N_i!} \,, \tag{1.32}$$

which accounts for the discreteness of galaxy counts and avoids the negative densities allowed by Gaussian noise models. Parallel Bayesian approaches to the initial-conditions problem include `KIGEN` (Kitaura, 2013; Heß et al., 2013) and `HAMLET` (Valade et al., 2022), the latter reconstructing the linear density and velocity fields from peculiar velocity surveys using HMC, with its newer generation `HAMLET-PM` upgrading to a full PM gravity model (Valade et al., 2026).

The resulting `BORG` posterior spans 256^3 white-noise field parameters, along with the bias model parameters. The current generation of `BORG` reconstructions assumes a fixed cosmology. This is a practical limitation driven by the difficulty of modelling galaxy bias jointly with cosmological parameters; future `BORG` generations will overcome it. `BORG` samples this posterior using HMC, computing the gradient $\nabla_s \ln \mathcal{P}$ by differentiating through the entire forward model—gravity solver, bias prescription, selection function, and likelihood—which requires each component to be differentiable with respect to the initial conditions. Since the galaxy bias parameters are inferred jointly with the initial conditions, their uncertainty is accounted for within the reconstruction rather than fixing them to fiducial values. The reconstruction operates in a box of side length ∼1 Gpc centred on the Milky Way with a grid spacing of ∼3.9 Mpc. The constrained 2M++ volume occupies the central region; the surrounding volume is drawn from the ΛCDM prior and supplies the large-scale tidal forces acting on the local structures. The ∼3.9 Mpc grid spacing makes the `BORG` reconstruction effectively a low-pass-filtered version of the true density field.

Because the forward model evolves the density field under gravitational dynamics,

unobserved regions, including the Zone of Avoidance, are filled self-consistently rather than through ad hoc methods such as cloning galaxies from adjacent strips (Lavaux & Hudson, 2011). Each BORG posterior sample is a realisation of the primordial density field consistent with the late-time observations. Unlike standard cosmological analyses, which rely on large volumes to average over cosmic variance, BORG *conditions* on the specific local density field, which removes cosmic variance from the uncertainty budget. The degree of suppression varies across the posterior samples: in regions well mapped by observational data, such as around the Virgo, Coma, and Perseus clusters, different samples produce concordant structures, while near survey edges or behind the Galactic plane, the structures vary significantly. This constraint is also scale-dependent: Fourier modes well sampled by the galaxy distribution and within the reach of the forward model are well constrained by the data, while modes below the grid resolution or in sparsely observed regions are effectively drawn from the ΛCDM prior.

1.4.6 Constrained Simulations in BORG: CSiBORG

BORG must balance the physical fidelity of its forward model against the computational cost of coupling it to the inference. To keep the gravity model differentiable and the forward evaluation fast, it operates on a 256^3 initial condition grid spanning a $\sim(1\ \mathrm{Gpc})^3$ volume, sufficient to capture the large-scale structure of the cosmic web but unable to resolve haloes below $\sim 10^{14}\ h^{-1}\,\mathrm{M}_\odot$ or halo internal dynamics. The BORG posterior samples can, however, be resimulated at higher resolution with standard N-body simulations, producing constrained realisations that resolve individual haloes while inheriting the large-scale constraint from the data. The Constrained Simulations in BORG (CSiBORG) programme realises this (Bartlett et al., 2021; Desmond et al., 2022; Hutt et al., 2022).

Each CSiBORG realisation starts from a single BORG posterior sample. Fourier modes at wavenumbers $k > k_{\mathrm{BORG}} = \pi/(2.65\ h^{-1}\,\mathrm{Mpc})$ are unconstrained by the data, so independent random phases with ΛCDM power spectrum amplitudes are drawn for these modes, ensuring

that the total power spectrum matches the target $P(k)$ at all scales while the large-scale modes remain fixed by the posterior sample. The augmented initial conditions are then evolved to $z = 0$ with an N-body code (RAMSES or GADGET-4). These small-scale modes have negligible impact on the position, mass, or velocity of the most massive clusters, whose large initial Lagrangian patches are well constrained, but seed additional low-mass haloes that populate the simulation with the full hierarchy of structures expected in ΛCDM.

Three generations of CSiBORG exist, corresponding to successive improvements in the BORG initial conditions. CSiBORG1 comprises 101 DM-only simulations evolved from an earlier generation BORG posterior (10-step particle-mesh gravity solver; Jasche & Lavaux, 2019) using RAMSES (Teyssier, 2002), with adaptive mesh refinement reaching a spatial resolution of 2.6 h^{-1} kpc in the densest regions and a particle mass of 3.09×10^9 $h^{-1}\,\mathrm{M}_\odot$ within the constrained volume. CSiBORG2 comprises 20 realisations evolved from a later generation BORG posterior (20-step COLA gravity solver; Stopyra et al., 2024a) using GADGET-4 (Springel et al., 2021), adopting updated *Planck* 2018 cosmological parameters (Planck Collaboration et al., 2020b). The improved COLA gravity model corrects the halo mass function at the high-mass end and yields better agreement of the most massive cluster positions with observations. CSiBORG1 was introduced by Bartlett et al. (2021) and CSiBORG2 by Stiskalek et al. (2026c). The third generation, CSiBORG-Manticore, follows the same simulation setup as CSiBORG2 but comprises 50 realisations drawn from the Manticore-Local posterior (McAlpine et al., 2025), which employs an improved bias model and updated cosmological parameters.

DM haloes are identified in each snapshot using the friends-of-friends (FOF) algorithm (Davis et al., 1985). The algorithm links every pair of particles separated by less than a linking length $\ell = b\,\bar{d}$, where $\bar{d} = L_{\mathrm{box}}/N_{\mathrm{part}}^{1/3}$ is the mean interparticle separation and $b = 0.2$ is the dimensionless linking parameter adopted throughout this thesis. A halo is then defined as the set of all particles connected by such links; the relation is transitive, so if particle A is linked to B and B to C, all three belong to the same group. Unlike

spherical overdensity definitions, the FOF boundary traces the actual isodensity surface and naturally captures the triaxial shapes of haloes. Artificial bridging merges physically distinct haloes when a thin particle bridge along a filament connects them into a single FOF group (Eisenstein & Hut, 1998; Warren et al., 2006; Lukić et al., 2009), but this primarily affects close pairs and is not corrected for in this work. A minimum of 100 particles per halo is imposed to ensure numerical convergence of bulk properties (mass, centre of mass, bulk velocity), corresponding to a minimum halo mass of $3.09 \times 10^{11}\ h^{-1}\,\mathrm{M}_{\odot}$. The present-day density and velocity fields are constructed from the particle distribution using smoothed-particle hydrodynamics interpolation (Monaghan, 1992), producing continuous, adaptive-resolution fields that are evaluated on regular grids of $0.7\ h^{-1}$ Mpc spacing for the velocity field analyses of Chapters 3 and 4. Finally, in CSiBORG1 each halo's Lagrangian patch, the region in the initial conditions from which its constituent particles originate, is traced back to the initial snapshot, enabling cross-identification of haloes between different CSiBORG realisations via the overlap of their Lagrangian volumes (Chapter 2).

Variation across the suite of simulations quantifies reconstruction precision; accuracy would require comparison to the true density field, which is unavailable. A halo whose mass, Lagrangian patch, and position agree across many CSiBORG realisations is precisely reconstructed, but correspondence to observed clusters requires separate validation against galaxy catalogues. From the BORG 2M++ posterior, Hutt et al. (2022) quantified the reduction in cosmic variance on the halo mass function and developed a method to assess individual halo reconstruction consistency. Chapter 2 develops a framework to quantify this halo-level precision systematically; an alternative approach was recently introduced by McAlpine (2026). For the most massive nearby clusters, matching reconstructed haloes to observed counterparts is straightforward (Stopyra et al., 2024a; McAlpine et al., 2025); extending this to lower-mass objects would require a probabilistic framework. The BORG framework and its constrained realisations have enabled a broad programme of scientific applications, including constrained hydrodynamical simulations of the Local Group (Sawala

et al., 2022; McAlpine et al., 2022), void catalogues (Desmond et al., 2022; Stopyra et al., 2024b; Malandrino et al., 2026), searches for modified gravity (Desmond & Ferreira, 2020; Bartlett et al., 2021) and DM annihilation/decay (Bartlett et al., 2022; Kostić et al., 2026), tests of local–ΛCDM compatibility (Stopyra et al., 2021, 2024a), studies of the Great Attractor and the Local Group's fate (Stiskalek et al., 2026a), thermal Sunyaev–Zel'dovich analyses (Stiskalek & Desmond, 2026), and the galaxy–environment connection (Gallagher et al., 2026). On the data side, `BORG` has been extended to cosmic shear (Porqueres et al., 2021) and peculiar velocity (Prideaux-Ghee et al., 2023) likelihoods, though the latter remains in early stages. This thesis contributes to the `CSiBORG` programme by developing the halo-level precision framework of Chapter 2, validating the reconstructed velocity field against independent observations (Chapter 3), and exploiting it for a novel two-rung measurement of H_0 (Chapter 4).

1.5 Peculiar velocities

The peculiar velocity field encodes the time-integrated gravitational acceleration of galaxies, and comparing observed motions to predictions from density-field reconstructions constrains the growth rate $f\sigma_8$ and the matter distribution largely independently of galaxy bias (Yahil et al., 1991; Pike & Hudson, 2005; Davis et al., 2011; Carrick et al., 2015; Adams & Blake, 2017; Dupuy et al., 2019; Nusser, 2017; Huterer et al., 2017; Boruah et al., 2020; Adams & Blake, 2020; Said et al., 2020). Peculiar velocities also enter directly into local measurements of H_0: SH0ES (Riess et al., 2011, 2022a) applies corrections derived by Peterson et al. (2022); Carr et al. (2022) based on the Carrick et al. (2015) local Universe reconstruction, while LIGO–Virgo–KAGRA (Abbott et al., 2018, 2021) relies on the GLADE+ compilation (Dálya et al., 2022), with peculiar velocity corrections based on an older generation of `BORG` (Jasche & Lavaux, 2019). Because $f\sigma_8 \sim S_8$, peculiar velocity measurements can also bear on the S_8 tension between the CMB and low-redshift

probes (Heymans et al., 2021; Abbott et al., 2022; Amon et al., 2022; Madhavacheril et al., 2024).

Peculiar velocity data can either constrain the initial conditions directly—through Wiener filtering (Sorce et al., 2014, 2016a) or Bayesian forward models (Valade et al., 2022, 2023; Prideaux-Ghee et al., 2023; Courtois et al., 2023; Valade et al., 2026)—or be forward-modelled given a fixed reconstruction to infer other quantities. This thesis adopts the latter approach: the initial conditions (and hence the density and velocity fields) are taken as given from an existing reconstruction such as `CSiBORG` or Carrick et al. (2015), and peculiar velocity data are forward-modelled to infer distance-indicator calibration, $f\sigma_8$, bulk flows, or H_0. Though the reconstruction enters as a fixed input, not as a quantity to be inferred, an overall velocity scaling or a galaxy bias model may be calibrated against the data. This section develops the forward-modelling framework for this fixed-reconstruction approach, as used in Chapters 3 and 4.

At low redshift, peculiar velocities of $\sim$300 km s^{-1} constitute a $\sim$5% perturbation to the Hubble flow at $z = 0.02$ and must be accounted for in precision cosmology. At $z \gtrsim 0.1$, individual peculiar velocities become subdominant, but their statistical imprint on galaxy clustering, through the Kaiser effect (Kaiser, 1987) and Fingers of God (Jackson, 1972), remains detectable as RSDs. One imprint of the local peculiar velocity field is the CMB dipole, which implies that the Local Group moves at $\sim$627 km s^{-1} relative to the CMB rest frame (Planck Collaboration et al., 2020a). Within ΛCDM, this motion must be generated by the gravitational pull of surrounding large-scale structure, and a reconstruction of the local velocity field should recover it self-consistently from the reconstructed density field. However, the reconstruction volume required to source this dipole is uncertain and may extend beyond the 2M++ volume (Stiskalek et al., 2026a).

As discussed in Section 1.2.3, peculiar velocities constrain $f\sigma_8$ (Carrick et al., 2015; Boruah et al., 2020; Boubel et al., 2024). Joint analysis of 6dFGS and SDSS peculiar velocities yields $f\sigma_8 = 0.338 \pm 0.027$ at effective redshift $z = 0.035$ (Said et al., 2020),

more than 3σ below the *Planck* ΛCDM prediction of $f\sigma_8 \approx 0.43$, indicating tension with ΛCDM. However, an analysis of TFR, FP, and SN Ia distances with the framework developed in Chapter 3 yields $S_8 = 0.793 \pm 0.035$ (or equivalently $f\sigma_8 = 0.391 \pm 0.022$), consistent with, though mildly lower than, *Planck*, and disfavouring the extremely low values reported by some earlier peculiar velocity analyses (Stiskalek, 2025).

The SH0ES distance ladder calibrates SNe Ia in the Hubble flow at $0.023 < z < 0.15$, where peculiar velocities are a per-cent-level correction to the recession velocity but still contribute a systematic uncertainty to H_0. Peculiar velocity corrections applied to SN Ia hosts, typically derived from density-field reconstructions such as Carrick et al. (2015) based on the 2M++, contribute ~ 0.3 km s^{-1} Mpc^{-1} systematic uncertainty to H_0 (Peterson et al., 2022; Hollinger & Hudson, 2025), subdominant but not negligible. Gravitational-wave standard sirens likewise require peculiar velocity corrections to convert the observed host-galaxy redshifts to cosmological redshifts, though the statistical uncertainty on the luminosity distance currently dominates (Dálya et al., 2022; Gray et al., 2023). Improving the fidelity of the local velocity field model, the central aim of Chapters 3 and 4, tightens H_0 constraints and enables sparser but well-characterised local distance samples.

1.5.1 Distance indicators

Peculiar velocities cannot be measured directly. The observable is a galaxy's redshift, which combines the cosmological recession velocity, a deterministic function of distance, with the peculiar velocity. Disentangling the two therefore requires an independent distance estimate, i.e. a distance indicator. These typically relate an observable property of the galaxy to its absolute magnitude M, yielding a distance modulus

$$\mu \equiv m - M = 5 \log \frac{d_\mathrm{L}}{\mathrm{Mpc}} + 25\,, \tag{1.33}$$

where m is the apparent magnitude and d_{L} the luminosity distance.[2]

The Tully–Fisher relation (TFR; Tully & Fisher, 1977) correlates the luminosity of a spiral galaxy with its rotational velocity. In ΛCDM, the relation arises from the interplay between dark matter halo mass and disc formation (e.g. Mo et al., 1998; Steinmetz & Navarro, 1999), while in modified Newtonian dynamics it follows directly that $M_{\mathrm{b}} \propto V^4$, where M_{b} is the baryonic mass. A varying mass-to-light ratio with luminosity or velocity means that the observed L–V relation need not have a slope of four (Milgrom, 1983; McGaugh et al., 2000). The observed slope and scatter depend on the photometric band: near-infrared bands, which trace stellar mass more faithfully, yield slopes closer to $L \propto V^{3.5}$–V^4, while bluer bands give shallower slopes owing to sensitivity to star formation history and dust (Aaronson et al., 1982; Masters et al., 2008). Following the (effective) parametrisation adopted in Chapter 3, the predicted absolute magnitude may be written as

$$M(\eta) = \begin{cases} a_{\mathrm{TFR}} + b_{\mathrm{TFR}}\,\eta + c_{\mathrm{TFR}}\,\eta^2 & \text{if } \eta > 0\,, \\ a_{\mathrm{TFR}} + b_{\mathrm{TFR}}\,\eta & \text{otherwise}\,, \end{cases} \tag{1.34}$$

where $\eta \equiv \log\!\left(W/\,\mathrm{km\,s^{-1}}\right) - 2.5$, W is the HI 21-cm line width corrected for inclination (a proxy for twice the maximum rotation velocity), and a_{TFR}, b_{TFR}, c_{TFR} are the zero-point, slope, and curvature. The quadratic term accounts for the observed steepening of the relation at high rotational velocities. The observed apparent magnitude m_{obs} is compared to the predicted value $m_{\mathrm{pred}} = \mu(r) + M(\eta_{\mathrm{true}})$, where $\mu(r)$ is the distance modulus, thereby constraining the distance r. The TFR parameters and the intrinsic scatter σ_{int} are unknown and must be inferred jointly with the other model parameters. The TFR delivers ∼20% distance errors per object and is the dominant contributor to current peculiar velocity compilations (Tully et al., 2023). The baryonic Tully–Fisher relation (BTFR; McGaugh et al., 2000) replaces luminosity with total baryonic mass $M_{\mathrm{bar}} = M_\star + M_{\mathrm{gas}}$, where $M_\star$ is the stellar mass and M_{gas} the gas mass, suppressing the band-dependent scatter and

[2]Throughout this thesis, log denotes the base-10 logarithm and ln the natural logarithm.

producing a tighter correlation, particularly for gas-dominated dwarfs.

While the TFR targets rotationally supported, late-type galaxies, the Fundamental Plane (FP; Djorgovski & Davis, 1987; Dressler et al., 1987) provides the analogous distance indicator for pressure-supported, early-type galaxies, generalising the earlier Faber–Jackson relation $L \propto \sigma^4$ (Faber & Jackson, 1976) by incorporating surface brightness as a third parameter. It relates three physical properties: the effective radius R_e, the mean surface brightness $\langle I \rangle_e$ within R_e, and the central velocity dispersion σ_0. Typically, the relation takes the form $\log R_e = a_{\rm FP} \log \sigma_0 + b_{\rm FP} \log \langle I \rangle_e + c_{\rm FP}$, motivated by the virial theorem for a pressure-supported system. Of these, $\langle I \rangle_e$ and σ_0 are directly observable and effectively distance-independent, while R_e is not: only the effective angular radius θ_e is measured. The FP predicts R_e from the other two observables, and comparing with the measured θ_e yields the angular diameter distance to the galaxy. The FP yields $\sim$25% distance scatter per object, but the abundance of early-type galaxies compensates through large sample sizes (Springob et al., 2014; Howlett et al., 2022).

Distance indicators can also exploit individual stars or transients within galaxies rather than integrated galaxy properties. Type Ia supernovae (SNe Ia), thermonuclear explosions of carbon–oxygen white dwarfs accreting from a binary companion, are precise, standardisable candles. The linear Tripp standardisation (Tripp, 1998) leads to a standardised apparent magnitude

$$m_{\rm standard} = m_{\rm obs} + \mathcal{A}\, x_1 - \mathcal{B}\, c \,. \tag{1.35}$$

The stretch x_1 and colour c are per-supernova parameters derived by fitting the SALT2 spectral-energy-distribution model (Guy et al., 2007) to the multi-band light curve of each SN. By contrast, the coefficients $\mathcal{A}$ and $\mathcal{B}$ in Eq. (1.35) are global standardisation parameters inferred from the population. Both the dependence on stretch and colour and the linear form of Eq. (1.35) are empirical choices: the latter amounts to a leading-order Taylor expansion of an underlying, more complex magnitude–stretch–colour relation, assumed to hold over the narrow range of x_1 and c spanned by supernova cosmology samples. The stretch term

encodes the broader-brighter width–luminosity relation, while the colour term conflates the intrinsic colour–luminosity correlation with host-galaxy dust reddening, so $\mathcal{B}$ is an effective slope interpolating between the intrinsic value and the dust law (Mandel et al., 2017). Over wider ranges this assumption breaks down, and Ginolin et al. (2025) report a significantly non-linear stretch–luminosity relation in the volume-limited ZTF DR2 sample. Some alternative standardisations adopt higher-order terms, for example the second-order colour–stretch dependence of SNooPy (Burns et al., 2018). The standardised apparent magnitude is then related to the distance modulus $\mu(r)$ and the peak standardised absolute magnitude $M_{\rm SN}$ by $m_{\rm standard} = \mu(r) + M_{\rm SN}$. After standardisation, SNe Ia achieve ∼10% distance precision per event and extend to much greater distances than the TFR or FP.

The Cepheid period–luminosity relation measured for individual stars in the Milky Way and nearby galaxies links the pulsation period of a classical Cepheid to its absolute magnitude (Leavitt & Pickering, 1912a). The predicted Wesenheit apparent magnitude of a Cepheid at distance modulus μ is

$$m_W = \mu + M_W + b_W \log\left(\frac{P}{\hat{P}}\right) + Z_W\,[\mathrm{O/H}]\,, \tag{1.36}$$

where P is the pulsation period, [O/H] the metallicity, M_W the fiducial absolute Wesenheit magnitude at a pivot period $\hat{P} = 10$ days and solar metallicity, b_W the period–luminosity slope, and Z_W the metallicity coefficient. The Wesenheit magnitude $m_W \equiv m_H - R(m_V - m_I)$ is a reddening-free combination that minimises extinction systematics (Madore, 1982). It is reddening-free by construction: R is set to the ratio of total-to-selective extinction $A_H/(A_V - A_I)$ of an assumed reddening law, so that the extinction A_H added to m_H is exactly cancelled by the colour term $R\,(A_V - A_I)$, leaving m_W unchanged. This holds strictly only for a universal reddening law with the adopted R and for Cepheids of negligible intrinsic colour width at fixed period; otherwise the fixed coefficient conflates intrinsic colour–luminosity variation with extinction, the same single-slope degeneracy that Mandel

et al. (2017) quantify for the linear colour standardisation of SNe Ia. The negligible intrinsic colour width is an assumption, but one supported by the tight Cepheid period–luminosity relation in the Wesenheit band, whose intrinsic scatter is only about 0.06 mag. Geometric anchors, namely Milky Way parallaxes, detached eclipsing binaries in the Large Magellanic Cloud (LMC), and the NGC 4258 megamaser, calibrate this scaling relation. The calibrated Leavitt relation is in turn used to calibrate SN Ia absolute magnitudes through a distance ladder, as discussed in Section 1.5.5.

Many other distance indicators exist. Water megamasers in the accretion discs of active galactic nuclei emit at 22 GHz and can be spatially resolved with radio interferometry. Mapping the positions, line-of-sight velocities, and accelerations of individual maser spots traces the Keplerian rotation curve of the disc, from which the physical radius and hence a single-step geometric distance can be derived (Pesce et al., 2020b). The technique was pioneered for the megamaser disc in NGC 4258 (Miyoshi et al., 1995; Herrnstein et al., 1999; Humphreys et al., 2013; Reid et al., 2019) and later extended to galaxies in the Hubble flow by the Megamaser Cosmology Project (Reid et al., 2009; Braatz et al., 2010; Reid et al., 2013; Kuo et al., 2013; Gao et al., 2016; Pesce et al., 2020a,b). Surface Brightness Fluctuations (SBF; Tonry & Schneider, 1988) measure distances to early-type galaxies by exploiting the Poisson variance in the number of stars per resolution element: at greater distances, more stars contribute to each pixel, reducing the pixel-to-pixel flux variance in a predictable way. SBF reaches ∼100 Mpc with ∼5–7% distance precision in the near-infrared (Jensen et al., 2021). The Tip of the Red Giant Branch (TRGB) exploits the sharp luminosity cutoff at the tip of the red giant branch, set by the helium core flash in the degenerate cores of low-mass stars, whose I-band absolute magnitude is nearly insensitive to age and metallicity for old, metal-poor populations (Da Costa & Armandroff, 1990; Freedman et al., 2019). TRGB distances reach ∼20 Mpc and serve as an independent second-rung calibrator in the distance ladder, complementing Cepheids and providing a cross-check on the SN Ia calibration (Freedman, 2021).

1.5.2 Peculiar velocity surveys

These distance indicators underpin five decades of peculiar velocity surveys (see Strauss & Willick, 1995, for an older but excellent review). The field began with Aaronson et al. (1982), who used the infrared TFR to detect Virgo infall at ~150–400 $\mathrm{km\,s^{-1}}$. Later, the "Seven Samurai" used FP distances to ~400 ellipticals to discover coherent streaming towards the Great Attractor (Dressler et al., 1987; Lynden-Bell et al., 1988). In the following decades, sample sizes grew rapidly: SFI++ compiled ~5,000 spiral TFR distances (Masters et al., 2006; Springob et al., 2007), 2MTF added ~2,000 with uniform all-sky coverage (Masters et al., 2008; Hong et al., 2019), and 6dFGS measured ~8,900 FP distances in the southern hemisphere (Springob et al., 2014). The SDSS Peculiar Velocity Survey (Howlett et al., 2022) and DESI DR1 (Douglass et al., 2025; Said et al., 2025) now provide the largest homogeneous TFR and FP samples. On the SN Ia side, Foundation (Foley et al., 2018; Jones et al., 2019) and LOSS (Li et al., 2000; Ganeshalingam et al., 2013) are the principal samples used in the peculiar velocity analyses of Chapter 3, complemented by the larger Pantheon+ (Scolnic et al., 2022; Brout et al., 2022) and DES 5-year (DES Collaboration et al., 2024) compilations. CosmicFlows-4 (CF4; Tully et al., 2023) is the largest current distance tracer compilation, comprising ~56,000 galaxies in ~38,000 groups from eight methodologies (TFR, FP, SNe Ia, TRGB, Cepheids, SBF, SNe II, masers) with absolute calibration anchored to geometric distances. CF4 is, however, a compilation rather than a homogeneous survey: it merges dozens of programmes with different selection functions, zero-point calibrations, photometric bands, and sky masks, making a robust characterisation of its overall selection function and systematics impractical. Next-generation surveys—the full DESI footprint, the 4MOST Hemisphere Survey (de Jong et al., 2019), and WALLABY (Koribalski et al., 2020)—will deliver $\gtrsim 10^5$ peculiar velocities, enabling precision tests of structure growth at $z < 0.1$.

Other than the measurements of growth rate, one standard diagnostic extracted from

peculiar velocity surveys is the bulk flow: the volume-averaged peculiar velocity within a sphere of radius R,

$$\boldsymbol{U}(R) = \frac{1}{V} \int_{|\boldsymbol{r}|<R} \boldsymbol{v}(\boldsymbol{r})\, \mathrm{d}^3\boldsymbol{r}\,, \tag{1.37}$$

which measures the net gravitational pull on the survey volume from the surrounding matter distribution. In ΛCDM, the expected bulk flow amplitude decreases with distance as the velocity field averages over increasingly many independent modes. The CF4 bulk flow measurements suggest a rising amplitude with distance: $\sim$400 km s^{-1} at 200 h^{-1} Mpc, roughly three times the ΛCDM expectation, corresponding to 4.8σ tension (Watkins et al., 2023). This excess may reflect systematic biases in the minimum-variance estimator when applied to inhomogeneous survey geometries, though Whitford et al. (2023) argue that the estimator itself is unbiased. Their analysis does not, however, test whether the underlying distances are well calibrated; a zero-point miscalibration in one or more distance indicators would propagate directly into the inferred bulk flow. The forward-modelling approaches developed in this thesis find no evidence for anomalously large bulk flows, whether applied to TFR data (Stiskalek et al., 2026d) or galaxy cluster scaling relations (Yasin et al., 2026), though neither analysis uses the same data as Watkins et al. (2023), who measure the bulk flow at large distances using the SDSS FP sample.

1.5.3 The peculiar velocity inference framework

Inferring the growth rate or the bulk flow requires a forward model that, given a set of parameters (such as distances, velocities, calibration coefficients), predicts the distribution of the observables (such as redshifts, apparent magnitudes, linewidths). In the CMB frame, the observed redshift of a galaxy decomposes into contributions from the cosmic expansion and the peculiar velocity:

$$(1 + z_{\mathrm{CMB}}) = (1 + z_{\mathrm{cosmo}})(1 + z_{\mathrm{pec}}). \tag{1.38}$$

The cosmological redshift $z_{\rm cosmo}$ maps to a unique comoving distance $r(z_{\rm cosmo})$ via the assumed background cosmology, and the peculiar velocity redshift is $z_{\rm pec} = v_r/c$, where v_r is the line-of-sight component of the peculiar velocity (assumed non-relativistic). The bulk flow is a summary statistic of the peculiar velocities of individual galaxies, computed *a posteriori*.

As with any forward model, this takes the hierarchical form of Eq. (1.20): the distance indicator calibration and, potentially, cosmological parameters are the population-level hyperparameters $\boldsymbol{\phi}$, while the true distances $\{r_i\}$ along with other per-object observables such as linewidth are latent variables marginalised analytically or numerically. We now derive the per-object likelihood, illustrating it on the TFR. The SN Ia formulation is nearly identical, replacing linewidth with stretch and colour, and a similar approach applies to other distance indicators.

For a TFR galaxy, the observables are its redshift $z_{\rm obs}$, apparent magnitude $m_{\rm obs}$, and linewidth $\eta_{\rm obs}$. Assuming conditional independence between the three observables, the likelihood factorises as

$$\mathcal{L}(z_{\rm obs},\, m_{\rm obs},\, \eta_{\rm obs} \mid \boldsymbol{\phi},\, r,\, \eta_{\rm true}) = \mathcal{L}(z_{\rm obs} \mid \boldsymbol{\phi},\, r)\, \mathcal{L}(m_{\rm obs} \mid \boldsymbol{\phi},\, r,\, \eta_{\rm true})\, \mathcal{L}(\eta_{\rm obs} \mid \eta_{\rm true}). \tag{1.39}$$

Here $\eta_{\rm true}$ is the galaxy's true linewidth, whereas $\eta_{\rm obs}$ is a noisy realisation of $\eta_{\rm true}$. The redshift likelihood follows from introducing the predicted redshift $z_{\rm pred}$, a deterministic function of r and the velocity model via Eq. (1.38), and a true redshift $z_{\rm true}$ that differs from $z_{\rm pred}$ due to unresolved small-scale motions. With both the measurement $\mathcal{L}(z_{\rm obs} \mid z_{\rm true}) = \mathcal{N}(cz_{\rm obs};\, cz_{\rm true},\, \sigma_{cz}^2)$ and the residual scatter $p(z_{\rm true} \mid z_{\rm pred}) = \mathcal{N}(cz_{\rm true};\, cz_{\rm pred},\, \sigma_v^2)$ modelled as Gaussians, the convolution over $z_{\rm true}$ evaluates analytically,

$$\mathcal{L}(z_{\rm obs} \mid \boldsymbol{\phi},\, r) = \mathcal{N}\Big(cz_{\rm obs};\, cz_{\rm pred},\, \sigma_v^2 + \sigma_{cz}^2\Big), \tag{1.40}$$

where σ_{cz} is the redshift measurement uncertainty, which for spectroscopic surveys is

typically subdominant to $\sigma_v \sim 250$ km s^{-1} (Carrick et al., 2015; Stiskalek et al., 2026c).

The magnitude likelihood compares the observed apparent magnitude to its predicted value $m_{\rm pred} = M(\eta_{\rm true}) + \mu(r)$, where $M(\eta_{\rm true})$ is the absolute magnitude from the TFR calibration, r is the comoving distance, and $\mu(r) = 5\log(d_{\rm L}(r)/{\rm Mpc}) + 25$ is the distance modulus with $d_{\rm L}(r)$ the luminosity distance at comoving distance r,

$$\mathcal{L}(m_{\rm obs} \mid \boldsymbol{\phi}, r, \eta_{\rm true}) = \mathcal{N}\left(m_{\rm obs};\, m_{\rm pred},\, \sigma_{\rm int}^2 + \sigma_m^2\right), \tag{1.41}$$

combining the intrinsic scatter $\sigma_{\rm int}$ of the distance indicator with the photometric measurement error σ_m in quadrature (analogously to the σ_v and σ_{cz} combination in the redshift likelihood). Finally, the linewidth likelihood is

$$\mathcal{L}(\eta_{\rm obs} \mid \eta_{\rm true}) = \mathcal{N}(\eta_{\rm obs};\, \eta_{\rm true},\, \sigma_\eta^2), \tag{1.42}$$

assuming a Gaussian measurement error σ_η.

Assuming independence, the prior on the latent variables factorises as $\pi(r, \eta_{\rm true} \mid \boldsymbol{\phi}) = \pi(r \mid \boldsymbol{\phi})\, \pi(\eta_{\rm true} \mid \boldsymbol{\phi})$. The distance prior encodes both the geometric volume element and the local source density:

$$\pi(r \mid \boldsymbol{\phi}) \propto r^2\, n(r, \boldsymbol{\phi}), \tag{1.43}$$

where $n(r, \boldsymbol{\phi})$ is the galaxy number density along the line of sight, computed from the assumed density field via a galaxy-bias model to account for inhomogeneous Malmquist bias (Lynden-Bell et al., 1988; Hudson & Ostro, 1994; Strauss & Willick, 1995).[3] For example, adopting a uniform prior in distance modulus μ rather than a uniform-in-volume

[3]The term "Malmquist bias" refers to several distinct effects. In its classical sense (Malmquist, 1922), it is a selection effect: a flux-limited sample preferentially contains the intrinsically most luminous objects at large distances, so the mean absolute magnitude of the detected population is brighter than that of the underlying luminosity function; in our framework this effect is captured by the selection treatment of Section 1.5.4. Here the term denotes the distance-inference effect encoded in Eq. (1.43): a uniform-in-volume distribution of sources implies an r^2 prior on distance (the *homogeneous* Malmquist bias), while replacing the uniform density with the local source density $n(r)$ captures the preference of galaxies to reside in overdensities (the *inhomogeneous* Malmquist bias).

prior ($\pi(r) \propto r^2$) biases H_0 high, because the flat-in-μ prior upweights small distances relative to the true volume element (Desmond et al., 2026; Stiskalek et al., 2026e), though the modelling of selection effects adds additional layers of complexity. We assume that the true linewidth has a Gaussian hyperprior,

$$\pi(\eta_{\rm true} \mid \boldsymbol{\phi}) = \mathcal{N}(\eta_{\rm true};\, \hat{\eta},\, w_\eta^2), \tag{1.44}$$

with mean $\hat{\eta}$ and width w_η inferred jointly, following the marginalised normal regression framework of Bartlett & Desmond (2023). This simplification has negligible effect when the measurement uncertainty is small. In principle, one should adopt the linewidth mass function, akin to the HI mass function for HI linewidths.

Combining the likelihood and prior and marginalising over the latent variables yields the posterior for model parameters $\boldsymbol{\phi}$:

$$\begin{aligned} \mathcal{P}(\boldsymbol{\phi} \mid z_{\rm obs},\, m_{\rm obs},\, \eta_{\rm obs}) \propto \pi(\boldsymbol{\phi}) \iint \mathrm{d}r\, \mathrm{d}\eta_{\rm true}\, & \mathcal{N}\Big(cz_{\rm obs};\, cz_{\rm pred},\, \sigma_v^2 + \sigma_{cz}^2\Big) \\ & \times \mathcal{N}\Big(m_{\rm obs};\, m_{\rm pred},\, \sigma_{\rm int}^2 + \sigma_m^2\Big)\, \mathcal{N}(\eta_{\rm obs};\, \eta_{\rm true},\, \sigma_\eta^2) \\ & \times \pi(r \mid \boldsymbol{\phi})\, \mathcal{N}(\eta_{\rm true};\, \hat{\eta},\, w_\eta^2), \end{aligned} \tag{1.45}$$

though we note that selection must be modelled jointly, as discussed in Section 1.5.4. We evaluate the integral over r and $\eta_{\rm true}$ numerically at each MCMC step to retain a low-dimensional parameter space. For independent observations, the posterior is the product over all galaxies. When the reconstruction provides an ensemble of density and velocity fields (e.g. the `CSiBORG` posterior samples), the single-realisation likelihood is averaged over realisations, propagating reconstruction uncertainty into the final constraints. The above derivation assumes a relatively simple distance indicator relation within the forward model. The framework extends to more sophisticated generative models, such as BayeSN (Mandel et al., 2022; Grayling et al., 2024).

The peculiar velocity model enters the likelihood through the predicted radial velocity.

We consider four models of increasing complexity. The simplest assumes all peculiar velocities vanish on average, modelled entirely through a free parameter σ_v that absorbs all unmodelled velocity structure as residual scatter, leading to a diagonal covariance $C_{ij} = \sigma_v^2 \delta_{ij}$. The next level adds a coherent external bulk flow, $v_{r_i} = \hat{\boldsymbol{r}}_i \cdot \boldsymbol{V}_{\rm ext}$, constant to first order across the survey volume, with σ_v capturing residual scatter. The three Cartesian components of $\boldsymbol{V}_{\rm ext}$ describe the net motion of the survey volume.

A more conservative approach replaces the diagonal covariance assumption with the full ΛCDM velocity covariance matrix. The line-of-sight (LOS) velocity covariance between galaxies i and j is (Kaiser, 1988; Hui & Greene, 2006; Davis et al., 2011)

$$C_{ij} = \left\langle (\boldsymbol{v}_i \cdot \hat{\boldsymbol{r}}_i)(\boldsymbol{v}_j \cdot \hat{\boldsymbol{r}}_j) \right\rangle = \frac{\dot{D}_i \dot{D}_j}{2\pi^2} \int \mathrm{d}k\, P(k)\, F(kr_i,\, kr_j,\, \hat{\boldsymbol{r}}_i \cdot \hat{\boldsymbol{r}}_j), \tag{1.46}$$

where $\dot{D}_i \equiv \mathrm{d}D_i/\mathrm{d}\tau$ is the conformal-time derivative of the linear growth factor at the redshift of galaxy i, $P(k)$ is the matter power spectrum, and the geometric kernel

$$F(u, v, \cos\theta) = \sum_{\ell=0}^{\infty} (2\ell + 1)\, j_\ell'(u)\, j_\ell'(v)\, P_\ell(\cos\theta) \tag{1.47}$$

encodes the projection onto the lines of sight, with $j_\ell'(x)$ the derivative of the spherical Bessel function of order ℓ and P_ℓ the Legendre polynomial of order ℓ. An additional diagonal term $\sigma_v^2 \delta_{ij}$ can be used to account for small-scale virialised motions not captured by the linear power spectrum, e.g. for galaxies in clusters. This approach effectively marginalises over all realisations of the (linear) velocity field, yielding conservative constraints on $f\sigma_8$ and H_0 that assume only the ΛCDM power spectrum. This approach is conservative but not optimal: constrained realisations provide information about the specific velocity field, obviating marginalisation over all possibilities. This is the baseline velocity model in both Chapters 3 and 4.

The most informative and detailed model uses local Universe reconstructions, which

predict the velocity field $\boldsymbol{v}(\boldsymbol{r})$ at each position. The radial peculiar velocity is

$$v_{r_i} = \beta\, \hat{\boldsymbol{r}}_i \cdot \boldsymbol{v}(\boldsymbol{r}_i) + \hat{\boldsymbol{r}}_i \cdot \boldsymbol{V}_{\rm ext}, \tag{1.48}$$

with σ_v capturing residual scatter not accounted for by the reconstruction. For linear-theory reconstructions such as Carrick et al. (2015), the scaling parameter $\beta = f/b$ absorbs the growth rate and galaxy bias, and the velocity field is derived from the 2M++ luminosity-weighted galaxy density field by iteratively solving the linearised continuity equation to map the field from redshift- to real-space. For N-body resimulations such as `CSiBORG`, which assume a fixed cosmology, the velocity field is modelled directly from the particle distribution, and β serves as a consistency check: a value of unity indicates that the reconstruction predicts the correct amplitude of the velocity field and is not significantly misspecified. When $\boldsymbol{v}$ is derived from a reconstruction, the external bulk flow $\boldsymbol{V}_{\rm ext}$ models the gravitational pull of structures beyond the reconstruction volume.

Several independent reconstructions of the local density and velocity field are now available, broadly split by the data that constrain them. Reconstructions constrained by galaxy redshift surveys include the linear-theory approach of Carrick et al. (2015), the neural-network method of Lilow et al. (2024) trained on N-body simulations and applied to 2MRS, and the aforementioned `CSiBORG` programme and `Manticore-Local`. Reconstructions constrained instead by peculiar velocity catalogues include the Wiener-filter method of Sorce & Tempel (2018) applied to Cosmicflows-2 and the HMC-based forward model of Courtois et al. (2023) applied to Cosmicflows-4, or the HAMLET method of Valade et al. (2022, 2026).

1.5.4 Selection effects

The per-galaxy likelihood above neglected selection. Magnitude- or flux-limited surveys preferentially include brighter or nearer objects, so an explicit selection term in the likelihood

is required to avoid biased inference (Kelly et al., 2008).

Let $\boldsymbol{d}_i$ denote the observed data vector for the ith source and $\boldsymbol{\phi}$ the population-level hyperparameters as in Eq. (1.20). We define the detection indicator $S_i = 1$ if the source is detected and $S_i = 0$ otherwise. Let $\mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\phi})$ denote the *untruncated* population likelihood, the probability of the data under the full population model, not conditioned on detection. For n detected sources drawn from a total population of N, the joint likelihood of all data—observed and unobserved—is the product over all N sources weighted by the binomial coefficient $\binom{N}{n}$ counting the number of ways of assigning the $N - n$ undetected sources:

$$\mathcal{L}(\boldsymbol{d} \mid \boldsymbol{\phi}, N) = \binom{N}{n} \left[\prod_{i \in \mathcal{A}_{\text{obs}}} p(S_i = 1 \mid \boldsymbol{d}_i)\, \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\phi}) \right] \left[\prod_{j \in \mathcal{A}_{\text{mis}}} p(S_j = 0 \mid \boldsymbol{d}_j)\, \mathcal{L}(\boldsymbol{d}_j \mid \boldsymbol{\phi}) \right], \tag{1.49}$$

where $\mathcal{A}_{\text{obs}}$ and $\mathcal{A}_{\text{mis}}$ are the sets of observed and unobserved sources, respectively. Marginalising over the unknown data $\boldsymbol{d}_j$ of all unobserved sources collapses each factor in the second product into a source-independent non-detection probability, giving

$$\mathcal{L}(\boldsymbol{d}_{\text{obs}} \mid \boldsymbol{\phi}, N) \propto \binom{N}{n} \left[p(S = 0 \mid \boldsymbol{\phi}) \right]^{N-n} \prod_{i \in \mathcal{A}_{\text{obs}}} p(S_i = 1 \mid \boldsymbol{d}_i)\, \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\phi}), \tag{1.50}$$

where

$$p(S = 0 \mid \boldsymbol{\phi}) \equiv \int \mathrm{d}\boldsymbol{d}\; p(S = 0 \mid \boldsymbol{d})\, \mathcal{L}(\boldsymbol{d} \mid \boldsymbol{\phi}) \tag{1.51}$$

is the marginal population non-detection probability, satisfying

$$p(S = 0 \mid \boldsymbol{\phi}) + p(S = 1 \mid \boldsymbol{\phi}) = 1. \tag{1.52}$$

The total number N is an unknown nuisance parameter, with the joint posterior of $\boldsymbol{\phi}$ and N being

$$\mathcal{P}(\boldsymbol{\phi}, N \mid \boldsymbol{d}_{\text{obs}}) \propto \pi(\boldsymbol{\phi})\, \pi(N) \binom{N}{n} \left[p(S = 0 \mid \boldsymbol{\phi}) \right]^{N-n} \prod_{i \in \mathcal{A}_{\text{obs}}} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\phi}). \tag{1.53}$$

With some foresight, and motivated by scale invariance, we may adopt a log-uniform prior $\pi(N) \propto 1/N$ and marginalise by summing over all $N \geq n$. Multiplying and dividing by $[p(S=1 \mid \boldsymbol{\phi})]^n$ and using $\binom{N}{n} = \binom{N-1}{n-1} N/n$,

$$\begin{aligned} \mathcal{P}(\boldsymbol{\phi} \mid \boldsymbol{d}_{\rm obs}) \propto \pi(\boldsymbol{\phi})\, [p(S=1 \mid \boldsymbol{\phi})]^{-n} \prod_{i \in \mathcal{A}_{\rm obs}} p(S_i=1 \mid \boldsymbol{d}_i)\, \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\phi}) \\ \times \sum_{N=n}^{\infty} \binom{N-1}{n-1} [p(S=1 \mid \boldsymbol{\phi})]^n\, [p(S=0 \mid \boldsymbol{\phi})]^{N-n}. \end{aligned} \tag{1.54}$$

The sum on the second line is the normalisation of the negative binomial distribution in N and evaluates to unity, yielding

$$\mathcal{P}(\boldsymbol{\phi} \mid \boldsymbol{d}_{\rm obs}) \propto \pi(\boldsymbol{\phi})\, [p(S=1 \mid \boldsymbol{\phi})]^{-n} \prod_{i \in \mathcal{A}_{\rm obs}} p(S_i=1 \mid \boldsymbol{d}_i)\, \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\phi}). \tag{1.55}$$

Selection is thus entirely encapsulated in the factor $[p(S=1 \mid \boldsymbol{\phi})]^{-n}$, which reweights the posterior to compensate for the fraction of the population that falls below the detection threshold. The remaining factor $p(S_i = 1 \mid \boldsymbol{d}_i)$ is the per-source detection probability: unity for a sharp selection cut, or a smooth function of the data if the selection boundary is probabilistic. If the posterior on N is desired, it follows from the conditional

$$p(N \mid n, \boldsymbol{\phi}) = \binom{N-1}{n-1} [p(S=1 \mid \boldsymbol{\phi})]^n\, [p(S=0 \mid \boldsymbol{\phi})]^{N-n}, \tag{1.56}$$

which can be further marginalised over $\boldsymbol{\phi}$ using the posterior samples. This is relevant when constraining the number density of objects through the stellar mass function or luminosity function. The detection probability is obtained by integrating the selection indicator over the observable and latent-variable space:

$$p(S = 1 \mid \boldsymbol{\phi}) \equiv \int \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{\theta}\; p(S=1 \mid \boldsymbol{x})\, \mathcal{L}(\boldsymbol{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi})\, \pi(\boldsymbol{\theta} \mid \boldsymbol{\phi}), \tag{1.57}$$

where $\boldsymbol{x}$ denotes the observables (e.g. apparent magnitude, redshift) and $\boldsymbol{\theta}$ the per-source

latent parameters (e.g. distance). This is the expected fraction of the population that is detected.

The derivation above assumed an explicit likelihood and a selection term $p(S = 1 \mid \boldsymbol{\phi})$ cheap enough to evaluate—analytically or numerically—at each MCMC step. This is straightforward when the generative model admits a closed-form likelihood, but need not hold in general. When the forward model involves expensive or non-differentiable simulations, the selection term can be emulated with a surrogate model. Boyd et al. (2026) took this further in the context of supernova cosmological inference, training a normalising flow on simulations to learn the likelihood of the observed data given model parameters, thereby absorbing selection effects implicitly. A second assumption is a scale-invariant prior on the total population size, $\pi(N) \propto 1/N$ (uniform in $\log N$), which permits analytic marginalisation over N. This is the scale-invariant Jeffreys prior for a Poisson population size, the natural non-informative choice for a positive count whose order of magnitude is unknown a priori. We verify on simulated mock catalogues that the resulting inference is unbiased, as demonstrated explicitly in Chapter 4. Other choices would require sampling N explicitly, jointly with $\boldsymbol{\phi}$ under an arbitrary prior $\pi(N)$, and evaluating Eq. (1.50) directly. Chapter 4 applies this framework to the Cepheid host-galaxy rung of the distance ladder.

1.5.5 The cosmic distance ladder and the H_0 tension

With a single distance indicator, H_0 and the zero-point are perfectly degenerate: $\mu = 5\log(d_L/\mathrm{Mpc}) + 25$ and $d_L \propto H_0^{-1}$, so a shift in H_0 is absorbed by the zero-point. For studies of large-scale flows this poses no difficulty: one can work in units of $h^{-1}\,\mathrm{Mpc}$, so that the inferred zero-point absorbs the unknown value of h. To break the degeneracy and measure H_0, one must calibrate the distance indicator against another with an absolute distance scale. That calibrator may itself require calibration, forming a ladder of rungs ultimately anchored to a geometric distance.

The best-known realisation is the SH0ES distance ladder (Riess et al., 2022a), which

comprises three rungs. The third (uppermost) rung is a sample of a few hundred SNe Ia in the Hubble flow, sufficiently distant for peculiar velocities to be subdominant to the cosmological redshift. Measuring H_0 from these supernovae requires their absolute magnitudes, which demands non-trivial standardisation. The second rung provides this absolute calibration: galaxies that host both a SN Ia and an independent distance indicator, in the SH0ES case, Cepheid variable stars observed with the Hubble Space Telescope (HST) out to ~30 Mpc (42 SNe Ia in 37 Cepheid host galaxies), though other indicators, e.g. the TRGB, can serve the same role (Freedman et al., 2019). The first rung then calibrates the Cepheids themselves using geometric anchors. Milky Way (MW) Cepheids provide one geometric anchor, under the assumption that they are drawn from the same population as Cepheids in the nearby host galaxies, with trigonometric parallaxes measured by *Gaia* (Gaia Collaboration et al., 2023). Two further anchors cross-check this calibration and provide further information: Cepheids in the LMC, whose distance is measured via detached eclipsing binaries (Pietrzyński et al., 2019), and Cepheids in NGC 4258, whose geometric distance comes from mapping Keplerian water megamaser orbits (Reid et al., 2019). Together, these three rungs yield a calibrated Hubble diagram from which H_0 can be measured. Other cosmological parameters that depend on the shape of the diagram rather than its overall normalisation do not require an absolute distance scale (e.g. $\Omega_{\rm m}$).

The measured H_0 is sensitive to systematics at every rung, including photometric calibration (Riess et al., 2022a,b), the Cepheid metallicity dependence (Romaniello et al., 2022; Ripepi et al., 2026), crowding and blending in HST photometry (Riess et al., 2022a, 2023; Bhardwaj et al., 2023), peculiar velocities (Carreres et al., 2025; Tsaprazi & Heavens, 2025), and the Cepheid period–luminosity calibration, yet none accounts for the full discrepancy. James Webb Space Telescope (JWST) Cepheid photometry in SNe Ia hosts confirms HST-based distances, with minor crowding and blending effects (Riess et al., 2024). Because the ladder is a chain of observed samples, the selection function at each rung must be modelled. Chapter 4 develops a Bayesian framework that explicitly models the selection of SH0ES

Cepheid host galaxies. The SH0ES team instead applies bias corrections based on the BEAMS with Bias Corrections (BBC) method (Kessler & Scolnic, 2017), which acts as an additive correction to the observed magnitudes to yield an unbiased Hubble diagram.

The Cepheid–SN Ia route is not the only path through the ladder (H0DN Collaboration et al., 2026). The TRGB method exploits the luminosity cutoff at the tip of the red giant branch (Section 1.5.1, Freedman et al. 2019). The J-region Asymptotic Giant Branch (JAGB) method uses carbon stars as standard candles in the near-infrared J-band (Freedman & Madore, 2020; Lee et al., 2025). Both can replace Cepheids as the second rung, providing an independent cross-check on SH0ES, though they rely on the same geometric anchors for calibration. Water megamasers offer a complementary single-step route: very long baseline interferometric mapping of Keplerian disc orbits yields geometric distances directly in the Hubble flow, bypassing both the second and third rungs entirely (Pesce et al., 2020b).

The third rung can also be bypassed entirely: if a Cepheid host galaxy has a spectroscopic redshift, the Cepheid-calibrated distance modulus can be compared directly to the redshift to infer H_0, with no reliance on SNe Ia. This two-rung ladder trades sample size and redshift range for independence from SN Ia standardisation systematics. Kenworthy et al. (2022) first pursued this approach, finding no tension with the SN-calibrated H_0. In Stiskalek et al. (2026e), we placed it on a more rigorous statistical footing, including improved peculiar velocity modelling, as presented in Chapter 4.

This distance ladder formulation leads to the most statistically significant anomaly in modern cosmology: the ∼9% discrepancy between early- and late-Universe measurements of H_0. *Planck* 2018 yields $H_0 = 67.4 \pm 0.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ assuming flat ΛCDM (Planck Collaboration et al., 2020b), while SH0ES measures $H_0 = 73.04 \pm 1.04\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ from Cepheid-calibrated SNe Ia (Riess et al., 2022a); the tension rises to 5.8σ with the latest calibration (Breuval et al., 2024). An updated *Planck* analysis by Tristram et al. (2024) yields a consistent $H_0 = 67.49 \pm 0.53\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, offering no softening of the tension. Other early-Universe probes tell the same story: ACT DR6 (Louis et al., 2025)

and SPT-3G (Camphuis et al., 2026) both recover $H_0 \approx 67$–$68\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ without using *Planck* data, and DESI DR2 BAO combined with *Planck* and ACT yields $H_0 = 67.97 \pm 0.38\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Abdul Karim et al., 2025).

On the late-Universe side, the Chicago–Carnegie Hubble Program (Freedman et al., 2025) calibrates SNe Ia using the TRGB and JAGB methods in place of Cepheids, reporting $H_0 = 69.85 \pm 1.75\,(\mathrm{stat}) \pm 1.54\,(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $H_0 = 67.96 \pm 2.00\,(\mathrm{stat}) \pm 1.31\,(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, respectively, both between the *Planck* and SH0ES values but with uncertainties too large to discriminate. Strong lensing time delays yield $H_0 = 73.3^{+1.7}_{-1.8}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Wong et al., 2020), though relaxing the lens mass profile weakens this to $H_0 = 74.5^{+5.6}_{-6.1}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Birrer et al., 2020). Other late-Universe methods, including SBF (Blakeslee et al., 2021; Jensen et al., 2025) and Type II supernovae (de Jaeger et al., 2022; Vogl et al., 2025), independently prefer $H_0 \gtrsim 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Gravitational-wave sirens (Abbott et al., 2017, 2023) offer a clean, single-step measurement but currently lack the precision to discriminate between the two values.

No proposed resolution simultaneously satisfies all observational constraints (see Abdalla et al., 2022; Di Valentino et al., 2021; Schöneberg et al., 2022; Di Valentino et al., 2025, for reviews). Known systematics in the distance ladder—photometric calibration, Cepheid metallicity dependence, SN Ia selection effects, peculiar velocity corrections—each contribute at the ∼1% level; none accounts for the full ∼9% discrepancy. The BORG programme's reconstruction of local large-scale flows provides an improved peculiar velocity model for nearby distance indicators, enabling supernova-free, one-rung or two-rung inference of H_0 (Stiskalek et al., 2026e).

1.6 Thesis outline

The three chapters are self-contained and are based on published papers. They share a common theme: establishing how well the local Universe can be reconstructed, from

individual haloes (Chapter 2) to large-scale flows (Chapter 3), and then leveraging these reconstructions for precise H_0 inference (Chapter 4). These contributions point towards two longer-term goals: BORG reconstructions constrained jointly by redshift and peculiar velocity data, and a Bayesian hierarchical model of the three-rung distance ladder. With the first and second rungs modelled in this thesis, only the third rung remains, though it is the most challenging rung.

Throughout, $\mathcal{N}(x \mid \mu, \sigma^2)$ denotes the Gaussian (normal) probability density function with mean μ and variance σ^2, with $\mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ its multivariate generalisation; $x \hookleftarrow \mathcal{N}(\mu, \sigma^2)$ indicates a draw from the corresponding distribution. Φ denotes the standard normal cumulative distribution function (CDF), and all logarithms are base-10 unless indicated otherwise.

2

Evaluating the variance of individual halo properties in constrained cosmological simulations

BORG infers the large-scale density field of the local Universe, but quantifying its precision at the halo level has not previously been established. In this chapter, I introduce a Lagrangian-patch overlap metric that tracks the consistency of halo mass, position, velocity, spin, and concentration across the posterior.

Based on Stiskalek et al. (2024), published in Monthly Notices of the Royal Astronomical Society.

2.1 Introduction

Constrained simulations of the local Universe, in which the phases of the primordial density field are inferred from observational data rather than drawn at random, allow object-by-object comparison between simulation and observation. The *Bayesian Origin Reconstruction from Galaxies* (BORG) algorithm (Jasche & Wandelt, 2013; Jasche & Lavaux, 2019) infers these phases within a Bayesian forward model, producing posterior samples of the initial conditions, typically constrained by the 2M++ galaxy catalogue (Lavaux & Hud-

son, 2011). The *Constrained Simulations in* BORG (CSiBORG) suite (Bartlett et al., 2021; Desmond et al., 2022) evolves each posterior sample with a full N-body simulation, so that variation across the ensemble quantifies the reconstruction uncertainty while offering high-resolution predictions of the dark matter distribution and halo properties in the local Universe. Alternative constrained-simulation programmes, including CLUES (Gottloeber et al., 2010; Sorce et al., 2016a) and SLOW (Dolag et al., 2023), rely on the Wiener filter, which reconstructs the posterior mean but cannot sample the full posterior and therefore does not provide a probabilistic uncertainty estimate.

The large-scale density field reconstructed by BORG has been validated through its power spectrum, halo mass function, and bulk velocity statistics (Jasche & Lavaux, 2019; Hutt et al., 2022). A question that has not been addressed, however, is how well these large-scale constraints percolate down to the properties of individual haloes. The BORG posterior is constrained on scales above the grid resolution ($\sim$2.65 Mpc/h); below this scale, the CSiBORG initial conditions are drawn from the Λ-cold dark matter (ΛCDM) prior. Whether the constrained large-scale modes suffice to pin down the mass, position, velocity, spin, or concentration of a given halo depends on the size of its Lagrangian patch relative to the grid: massive haloes ($M \gtrsim 10^{14}\ h^{-1}\,\mathrm{M}_\odot$) originate from patches spanning $\gtrsim$100 BORG cells and are therefore expected to be tightly constrained, whereas lower-mass haloes occupy only a few cells and may vary substantially across the posterior. Quantifying this mass-dependent reconstruction precision is the objective of this chapter.

We develop a framework to assess whether a halo present in one CSiBORG realisation is also present in another, and if so, how consistently its properties are recovered. The method is based on the overlap of haloes' initial Lagrangian patches: two haloes in different realisations are identified as counterparts if the volumes from which their constituent particles originate overlap significantly in the initial conditions. This approach does not require a consistent particle ordering between runs, unlike traditional halo-matching methods (Butsky et al., 2016; Desmond et al., 2017; Mitchell et al., 2018; Cataldi et al., 2021). We quantify the

consistency of the halo reconstruction, its *precision*, as a function of halo mass, and establish the significance of the results in contrast with the unconstrained `Quijote` suite (Villaescusa-Navarro et al., 2020). We emphasise that precision (internal consistency across posterior samples) is distinct from accuracy (agreement with the real local Universe), which would require external observational validation.

The structure of this chapter is as follows. Section 2.2 introduces the two sets of simulations employed in our work. Section 2.3 presents the overlap metric and its interpretation. Section 2.4 contains our results, and Section 2.5 interprets them.

2.2 Simulated data

In this section, we describe the two sets of simulations that we use, `CSiBORG` and `Quijote`, and their halo catalogues.

2.2.1 CSiBORG

The `CSiBORG` suite (Section 1.4.6; Bartlett et al. 2021) consists of 101 dark matter (DM)-only N-body simulations in a 677.7 Mpc/h box centred on the Milky Way (MW), with initial conditions (ICs) sampled from the `BORG` 2M++ posterior (Lavaux & Jasche, 2016; Jasche & Lavaux, 2019). The `BORG` density field is constrained within a spherical volume of radius $\sim$155 Mpc/h, where the catalogue has high completeness, on a 256^3 grid with 2.65 Mpc/h spacing. Each `BORG` posterior sample is propagated linearly to $z = 69$ and augmented with white noise on a 2048^3 grid in the central region, yielding a spatial resolution of 0.33 Mpc/h and a particle mass of 3.09×10^9 $\mathrm{M}_\odot/h$. A buffer region of $\sim$10 Mpc/h ensures a smooth transition to the outer box. The adopted cosmological parameters are from the Planck Collaboration et al. (2014) best-fit results with the Wilkinson Microwave Anisotropy Probe (WMAP) polarisation, high-multipole, and BAO data, except $H_0 = 70.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ from the 5-year WMAP combined analysis (Hinshaw et al., 2009) ($\Omega_\mathrm{m} = 0.307$, $\Omega_\Lambda = 0.693$,

$\Omega_\mathrm{b} = 0.04825$, $\sigma_8 = 0.8288$, $n_s = 0.9611$). The density field is evolved to $z = 0$ with RAMSES (Teyssier, 2002), refining only the central region to level 18 (2.6 kpc/h resolution).

As we will be studying the reconstruction of individual objects, whose initial Lagrangian patches are constrained in BORG, it is illustrative to consider how many BORG cells constitute the Lagrangian patch of a halo. We find this to be approximately $N \approx 7\ M_\mathrm{tot}/(10^{13} M_\odot/h)$. BORG constrains the average field value in each cell with physical size 2.65 Mpc/h, and we do not a priori expect haloes with Lagrangian patches spanning only a few BORG cells to be consistently reconstructed in CSiBORG since such haloes likely vary strongly across the BORG posterior. However, haloes above $\sim 10^{14}\ M_\odot/h$ comprise initially $\gtrsim 100$ cells and are therefore likely well constrained.

2.2.2 Quijote

We compare the CSiBORG results to the publicly available Quijote simulations[1] (Villaescusa-Navarro et al., 2020). Quijote is a suite of unconstrained simulations evolved from $z = 127$ to $z = 0$ using the GADGET-III code (Springel et al., 2008). We use 10 realisations of the Quijote DM-only simulations with randomly drawn IC phases, each with a volume of $(1\ \mathrm{Gpc}/h)^3$ and a particle mass of $8.72 \times 10^{10}\ \mathrm{M}_\odot/h$ in a fiducial cosmology: $\Omega_\mathrm{m} = 0.3175$, $\Omega_\Lambda = 0.6825$, $\Omega_\mathrm{b} = 0.049$, $H_0 = 67.11\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, $\sigma_8 = 0.834$ and $n_s = 0.9624$. Besides the constraints, the most significant difference with CSiBORG is the volume, so for an approximately fair comparison when calculating any extrinsic quantity we mimic the CSiBORG high-resolution region by splitting each Quijote box into 27 non-overlapping spherical sub-volumes of radius 155 Mpc/h centred at $n \times 155$ Mpc/h for $n = 1, 3, 5$ along each axis.

[1] https://quijote-simulations.readthedocs.io/

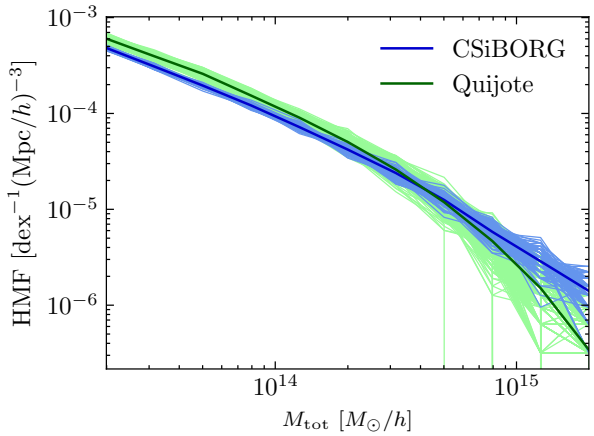


Figure 2.1: The friends-of-friends (FOF) halo mass function (HMF) in CSiBORG and Quijote. Thin lines show individual realisations; bold lines show the mean. CSiBORG undershoots and overshoots the Quijote halo mass function (HMF) at low and high masses, respectively, an offset driven by the BORG gravity model rather than deviations of the local Universe from the cosmic average (Section 2.2.3).

2.2.3 Halo catalogues

Haloes are identified in both CSiBORG and Quijote using the friends-of-friends algorithm (FOF; Davis et al. 1985; see also Section 1.4.6) with linking length $b = 0.2$. FOF can artificially bridge nearby structures along filaments (Eisenstein & Hut, 1998); Warren et al. (2006) and Lukić et al. (2009) proposed corrections based on particle number and concentration, but we do not apply them here as high-precision masses are not required.

To reduce numerical resolution errors of recovered haloes and their properties, the authors have suggested that haloes must contain at least $50 - 100$ particles (Springel et al., 2008; Onions et al., 2012; Knebe et al., 2013; van den Bosch & Jiang, 2016; Griffen et al., 2016), though Diemer & Kravtsov (2015) suggested stricter criteria for measuring e.g. concentration to avoid having only a few particles in the inner parts of the halo, especially if the inner region spans only a few force resolution lengths. Recently, van den Bosch & Ogiya (2018) proposed a more stringent criterion to determine the numerical convergence of infalling subhaloes. However, in this work, we are not concerned with substructure or halo profiles and therefore adopt the simpler and less restrictive criterion of 100 particles for all haloes, corresponding to a minimum halo mass of 3.09×10^{11} $\mathrm{M}_\odot/h$ and 8.72×10^{12} $\mathrm{M}_\odot/h$ for CSiBORG and Quijote, respectively. In CSiBORG we identify haloes only inside the high-resolution region. As CSiBORG has a mass resolution nearly two orders of magnitude better than Quijote, when comparing the two simulations, we only consider haloes above

the mass resolution of Quijote.

CSiBORG and Quijote assume different cosmological parameters, yielding a difference in both the large-scale structure and halo population. Hutt et al. (2022) study how this affects the halo mass function (HMF), and show in their Fig. 1 no significant disparities due to the different cosmologies. They do, however, quantify the reduced spread of the HMFs of individual CSiBORG realisations relative to Quijote due to the suppression of cosmic variance by the IC constraints. We show the HMFs in Fig. 2.1. It is found that the CSiBORG and Quijote HMFs have some systematic disagreement, with the former predicting higher high-mass-end abundances (see also McAlpine et al. 2022; Desmond et al. 2022; Hutt et al. 2022). However, this becomes statistically significant only above $M \sim 10^{15} M_\odot/h$. On the other hand, CSiBORG systematically undershoots the Quijote HMF below $\sim 10^{14.2}\ M_\odot/h$. These discrepancies are corrected by the upgraded 20-step COLA gravity solver adopted in the second-generation BORG posterior (Stopyra et al., 2024a; Tassev et al., 2013).

2.3 Methodology

We aim to develop a metric to quantify the precision with which haloes are robustly reconstructed between boxes, e.g. across the simulations of constrained suites. We do so by evaluating the similarity of proto-haloes at high redshift, measured by their initial Lagrangian patches. This is a priori a sensible approach as it is the initial conditions that are constrained: focusing on the initial snapshot facilitates the establishment of a more causally coherent framework, circumventing reliance on interpretations deduced purely from the final conditions as in e.g. (Hutt et al., 2022). Most of our work is intended to interpret and show that it is useful a posteriori as well.

Similar approaches are used to associate haloes between DM-only and hydrodynamical simulations sharing the same ICs (e.g. Butsky et al. 2016; Desmond et al. 2017; Mitchell et al. 2018; Cataldi et al. 2021). For instance, Desmond et al. leverage the ability to match

DM particle IDs between DM-only and hydrodynamical runs to match haloes as follows: If halo a from the DM-only run shares the most particles with halo b from the hydrodynamical run, and conversely, b shares most particles with a, they are identified as a match. This bears a strong resemblance to the method we develop because the particle IDs are assigned based on their position in the initial snapshot. In `CSiBORG` the particle IDs are not consistent between realisations, so this method cannot be used directly.

We cross-correlate two IC realisations as follows. We denote the first simulation as the "reference" and the second as the "crossing" simulation and calculate the overlap of each halo in the reference simulation with all haloes in the crossing simulation. We identify `FOF` haloes in the final snapshot of both the reference $\mathcal{A}^{\text{th}}$ and the crossing $\mathcal{B}^{\text{th}}$ IC realisation and trace their constituent particles back to the initial snapshot. To calculate the intersecting mass of a halo $a \in \mathcal{A}$ and $b \in \mathcal{B}$, we use the nearest grid point (NGP) scheme to assign the halo particles to a 2048^3 grid in the initial snapshot, matching the initial refinement of the high-resolution region. We denote the mass assigned to the nth cell as $m_a^{(n)}$ and $m_b^{(n)}$, respectively, for the two haloes. On the same grid, we also calculate the background mass field of all particles assigned to a halo in the final snapshot in the respective simulation, $\widehat{m}_{\mathcal{A}}^{(n)}$ and $\widehat{m}_{\mathcal{B}}^{(n)}$.

In the NGP scheme, a particle located at $x_i \in [0, 1]$ along the ith axis is assigned to the $\lfloor x_i N_{\text{cells}} \rfloor^{\text{th}}$ cell, where N_{cells} is the total number of cells along an axis. We then apply a Gaussian smoothing to the NGP field using the kernel width of one cell and define the intersecting mass of haloes a and b as

$$X_{ab} \equiv \sum_n \frac{2 m_a^{(n)} m_b^{(n)}}{\widehat{m}_{\mathcal{A}}^{(n)} + \widehat{m}_{\mathcal{B}}^{(n)}}, \tag{2.1}$$

where $n = 1, \ldots, 2048^3$ is the grid index. This definition accounts for the fact that particles

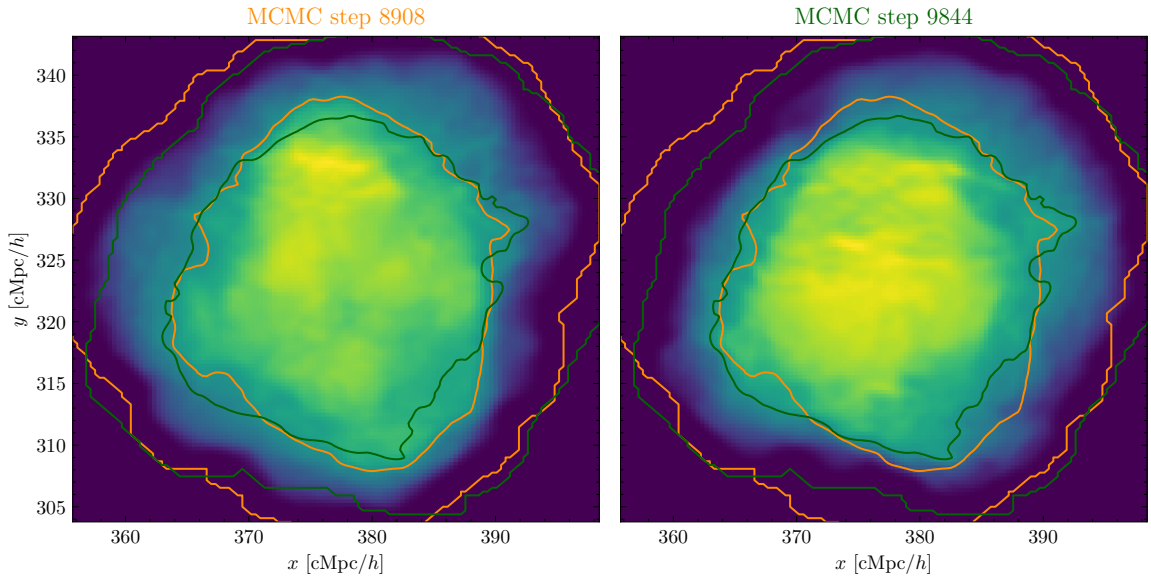


Figure 2.2: Example of Lagrangian patches at $z = 69$ for two haloes identified at $z = 0$ from CSiBORG at two distinct Markov Chain Monte Carlo steps of the BORG posterior. Both haloes have a mass of $10^{15}\ M_\odot/h$ at $z = 0$ and a mutual Lagrangian patch overlap of approximately 75%. The plot shows the projected mass density of the Lagrangian patches in the x-y plane, along with density contours corresponding to the two haloes. The contour colours match that of the labels above the figures. Notably, the size of each Lagrangian patch of such haloes is around 30 cMpc$/h$. Both panels share the same x and y axes.

of more than one halo may contribute to a single cell and may be motivated as

$$X_{ab} = \sum_n \left[\left(\frac{m_b^{(n)}}{\widehat{m}_{\mathcal{A}}^{(n)} + \widehat{m}_{\mathcal{B}}^{(n)}} \right) m_a^{(n)} + \left(\frac{m_a^{(n)}}{\widehat{m}_{\mathcal{A}}^{(n)} + \widehat{m}_{\mathcal{B}}^{(n)}} \right) m_b^{(n)} \right], \tag{2.2}$$

i.e. the contribution of a is weighted by the mass of b in that cell normalized by the total mass in that cell, and vice versa. Next, we define the IC overlap between the haloes a and b as

$$O_{ab} = \frac{X_{ab}}{M_a + M_b - X_{ab}}, \tag{2.3}$$

where $M_a = \sum_n m_a^{(n)}$ is the total particle mass of the ath halo, and similarly for the bth halo. This ensures that $O_{ab} \in [0, 1]$ and can be interpreted simply as the mass of the ath halo that overlaps with the bth halo normalised by their total mass. The definition of Eq. (2.1) accounts for the fact that some cells of a halo $a \in \mathcal{A}$ may overlap simultaneously with haloes $b_1, b_2 \in \mathcal{B}$. Due to the presence of b_2, the intersecting mass X_{ab_1} is appropriately reduced (denominator of Eq. (2.1)) to avoid double-counting. For further illustration, in Fig. 2.2, we present an example of Lagrangian patches from two haloes in `CSiBORG` with a 75% overlap. Despite being drawn from distinct steps of the `BORG` posterior, the two Lagrangian patches are located at the same comoving position and lead to a collapsed structure at $z = 0$ with a mass of approximately $10^{15}\ M_\odot/h$, also at a similar location.

We can illustrate the overlap on a toy example of two haloes. If $m_a^{(n)} = m_b^{(n)} = \widehat{m}_{\mathcal{A}}^{(n)} = \widehat{m}_{\mathcal{B}}^{(n)} = m$ in all cells, then X_{ab} is simply m times the number of cells occupied simultaneously by both a and b. Moreover, if the ath halo is completely enclosed within the bth halo, then the overlap can be expressed as $O_{ab} = M_a/M_b$, i.e. the ratio of their masses. In the case of perfect overlap $M_a = M_b$ and thus $O_{ab} = 1$.

The overlap measures the similarity of two haloes in their Lagrangian patches. However, a halo may overlap with many smaller haloes, producing a large set of small overlaps with none of the overlapping haloes being similar to the reference halo in mass or size. Therefore, we also calculate the maximum overlap that a halo a has with any halo in the crossing

simulation $\max_{b\in\mathcal{B}} O_{ab}$. If this quantity is sufficiently high across all crossing simulations, it implies that the halo is consistently reconstructed across the IC realisations. While the overlap between a pair of haloes is symmetric, if a halo $a_0 \in \mathcal{A}$ has a maximum overlap $O_{a_0 b_0} = \max_{b\in\mathcal{B}} O_{a_0 b}$ with some halo $b_0 \in \mathcal{B}$, this does not imply that b_0 also has a maximum overlap with a_0 since its maximum overlap is defined as $\max_{a\in\mathcal{A}} O_{b_0 a}$ and the two sets of haloes over which the overlaps are maximised are not the same.

The properties of the overlap O_{ab} lend it a natural interpretation as the probability of a match between haloes in two simulations. That a reference halo can have a non-zero overlap with multiple haloes that themselves may overlap in their initial Lagrangian patches is already accounted for in the definition of the overlap through the denominator of the intersecting mass in Eq. (2.1), which implies that the overlap of a pair of haloes is modified by the presence of other overlapping haloes. Therefore, we consider that the probability of a reference halo *a* being matched to *some* halo in the crossing simulation is simply the sum of the overlaps with all haloes in the crossing simulation:

$$P(\text{a} \in \mathcal{A} \text{ matched in } \mathcal{B}) = \sum_{b\in\mathcal{B}} O_{ab}, \tag{2.4}$$

and the probability of not being matched to any halo in the crossing simulation is

$$P(\text{a} \in \mathcal{A} \text{ not matched in } \mathcal{B}) = 1 - \sum_{b\in\mathcal{B}} O_{ab}. \tag{2.5}$$

The definitions of the intersecting mass and overlap in Eqs. (2.1) and (2.3) ensure that not only the overlap between a pair of haloes is always ≤ 1, but also that the sum of overlaps with a reference halo such as in Eq. (2.4) is always ≤ 1. In the denominator of Eq. (2.1), the intersecting mass is weighted by the fraction of a cell occupied by the pair of haloes, taking into account the contributions from all haloes in both simulations. This ensures that the intersecting mass is not counted multiple times when summing over crossing haloes.

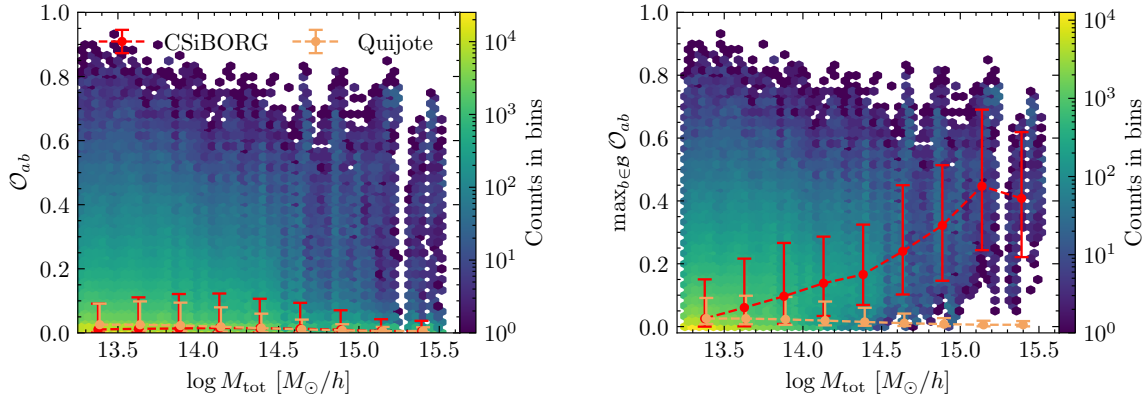
We calculate the most likely value of a matched halo property $\mathcal{H}$ (for instance, mass)

based on haloes that overlap with the reference halo. For each crossing simulation indexed i we select $\mathcal{H}_i$ of the halo with the highest overlap w_i with the reference halo. We then calculate the most likely matched property as the mode of the distribution of $\mathcal{H}_i$ weighted by w_i. To identify it, we employ the shrinking sphere method, commonly utilised for finding the centre of DM haloes (Power et al., 2003). We iteratively shrink the search radius around the weighted average of the enclosed $\mathcal{H}_i$ samples until fewer than 5 samples are enclosed, at which point we take their average as the mode of the distribution. We verify that 5 samples are sufficient and our conclusions are not affected by increasing it.

2.4 Results

We showcase the framework on the `CSiBORG` suite. We first calculate the overlaps of halo initial Lagrangian patches defined in Eq. (2.3). We then calculate the halo mass, spin, and concentration of overlapping haloes and compare them to the reference halo. At each step, we compare our findings to the unconstrained `Quijote` suite to assess the significance of the results. We calculate the overlaps for haloes whose total mass exceeds $10^{13.25}\ M_\odot/h$ and pairs of haloes closer in mass than 2 dex. The latter is simply to reduce computational expense, since halo pairs with larger differences in mass have negligible overlaps ($\lesssim$ 1 per cent).

Although the DM particle mass differs between `CSiBORG` and `Quijote`, this does not prevent us from comparing trends across the two simulation suites. This discrepancy would only be problematic if we were matching a `CSiBORG` box to `Quijote`, which we do not. The overlap metric itself does not directly depend on the particle mass, as the particles are deposited onto a regular grid: it is primarily influenced by the shot noise resulting from sampling the density field with a finite number of particles. However, we do not expect this to significantly impact our results. A halo with a mass of $10^{13.25}\ M_\odot/h$ in `Quijote` already consists of 200 particles, with the shot noise decreasing further at higher halo masses.



(a) Concatenated pair overlaps O_{ab} between a reference halo and all haloes from all remaining simulations.

(b) Maximum pair overlaps of a reference halo $\max_{b\in\mathcal{B}} O_{ab}$ with a single simulation, concatenated across all remaining simulations.

Figure 2.3: Comparison of Lagrangian patch overlaps in `CSiBORG` and `Quijote` simulations. In both cases, a single reference simulation is assumed, and for each halo a list of overlaps is calculated per crossing simulation. The x-axis denotes the mass of the reference halo, the hex bins illustrate `CSiBORG` overlaps, with the lines delineating overlaps in `CSiBORG` and `Quijote`, each accompanied by 1σ error bars showing the spread among points. The difference in overlap behaviour between the two simulations becomes evident at high mass, where the constraints of the `CSiBORG` suite become significant. The concatenated pair overlaps show no trend because they are dominated by frequent chance overlaps. On the other hand, there is a clear trend in the maximum overlaps towards higher halo masses.

2.4.1 Overlaps

We begin by using a single reference simulation and assessing how consistently its haloes are present in the remaining IC realisations. We calculate the *non-zero* overlaps between its haloes and those of another simulation following the approach of Section 2.3. This yields for each halo in the reference simulation a set of overlaps (which could potentially be empty), but its sum will never exceed one because of the denominator of Eq. (2.1). Retaining the same reference, the process is reiterated across the remaining IC realisations.

2.4.1.1 Pair overlap

We present the overlaps between the haloes in Fig. 2.3a where, for every reference halo, we concatenate all sets of overlaps with the remaining IC realisations. For comparison, in Fig. 2.3a we also show the overlaps in `Quijote` where we again use a single reference simulation. The hex-bin shows the `CSiBORG` results and the lines indicate binned medians and 2σ spread. It is evident that in both `CSiBORG` and `Quijote` the most likely overlap value is close to zero, which follows from the predominance of non-matched haloes in both simulation suites. However, a notable distinction emerges in `CSiBORG`: there is a pronounced tail of high overlaps, especially evident for haloes above $10^{14}\ M_\odot/h$. These massive haloes have counterparts that occupy the same part of the initial snapshot in the majority of the IC realisations. This is not the case in `Quijote`, where in fact the high-overlap tail becomes less significant for more massive haloes due to their rarity.

2.4.1.2 Maximum pair overlap

Next, we keep the same reference simulation, but instead calculate the maximum overlap of a reference halo in each other IC realisation. In Fig. 2.3b we present the median of the maximum overlaps for each reference halo across the remaining IC realisations. Unlike before, in `CSiBORG` the mean trend of the maximum overlaps as a function of mass is no

longer close to zero. On the other hand, in Quijote the mean trend of maximum overlaps remains close to zero. This indicates that there are haloes in any pair of simulations that match well in CSiBORG, especially at high mass, while there are not in Quijote. We verify that this conclusion holds irrespective of the choice of reference simulation.

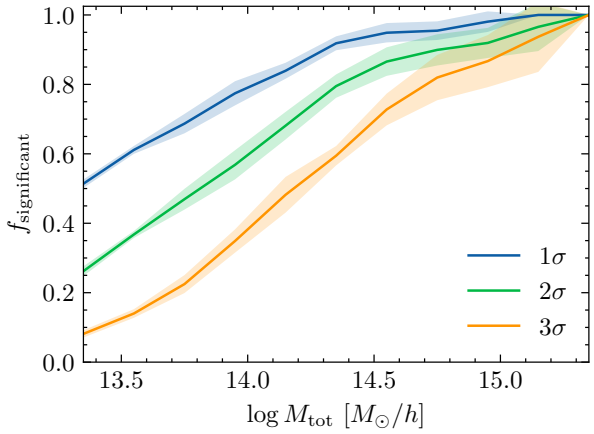
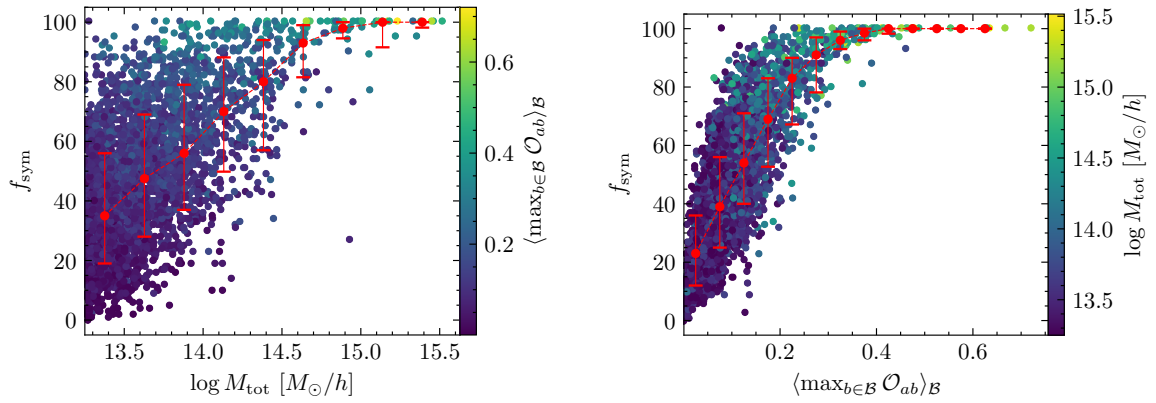


Figure 2.4: Fraction of CSiBORG reference haloes, as a function of $\log M_{\rm tot}$, with a mean maximum overlap exceeding the 1, 2, 3σ thresholds in Quijote (84.1, 97.7, 99.9 per cent; Fig. 2.3). Bands show the 1σ spread among CSiBORG realisations. Even at the lower mass threshold, a large fraction of CSiBORG haloes exceeds the Quijote overlaps.

In Fig. 2.4 we show the fraction of CSiBORG haloes in a reference simulation that have a median maximum overlap with other simulations over 1, 2, 3σ-level thresholds calculated in Quijote (84.1, 97.7, 99.9 per cent). This figure again highlights that more massive haloes are more clearly constrained: already at $10^{14} M_\odot/h$ about 75 per cent of CSiBORG haloes have median maximum overlaps more significant than the 1σ level in Quijote.

In Fig. 2.5, we show in CSiBORG for every halo from a single reference simulation the number of simulations in which a reference halo a and a halo b from a crossing simulation *both* have maximum overlaps with each other ($f_{\rm sym}$). We plot $f_{\rm sym}$ against both the reference halo mass and the average maximum overlap of a halo. On average, haloes around $\sim 10^{14} M_\odot/h$ have symmetric maximum overlaps in 50 per cent of the IC realisations. In contrast, haloes around $\sim 10^{15} M_\odot/h$ have symmetric maximum overlaps in nearly all IC realisations. On the other hand, the relationship between $f_{\rm sym}$ and the average maximum overlap has less scatter than its relationship with mass. For example, haloes with an average maximum overlap of ~ 0.2 have symmetric maximum overlaps in approximately 50 per cent of the realisations. The results of Fig. 2.4 and Fig. 2.5

demonstrate that the most massive haloes exhibit reduced variance in CSiBORG compared to a suite of unconstrained simulations, highlighting the robustness of our method in identifying consistently constrained haloes across different simulations. However, this does not imply that these haloes are accurately reconstructed relative to real observations; rather, it reflects the internal consistency of the simulation suite.



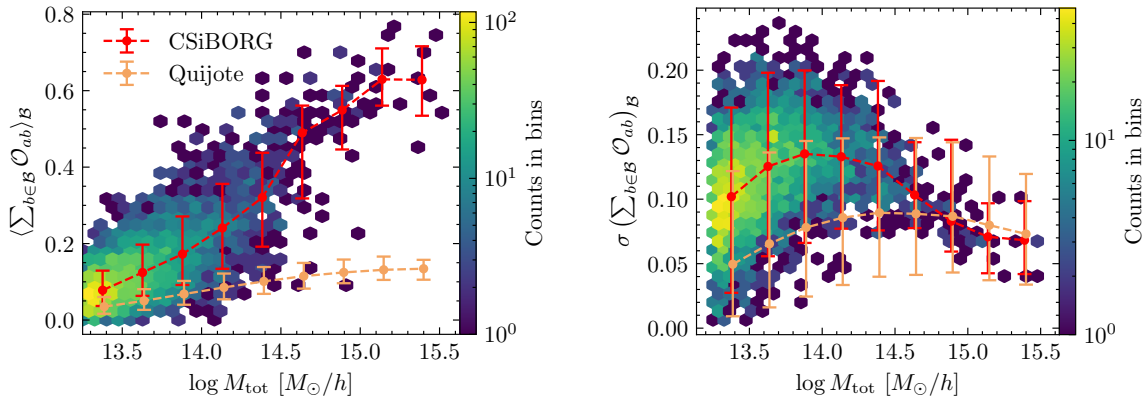
(a) $f_{\rm sym}$ against the reference halo mass.

(b) $f_{\rm sym}$ against the average maximum overlap of a halo.

Figure 2.5: Number of simulations in which a reference halo a and a halo b from a crossing simulation *both* have maximum overlaps with each other, $f_{\rm sym}$, shown for every halo from a single reference simulation in CSiBORG. The red line is an arithmetic average in a bin along with 1σ spread. The highest mass haloes have symmetric maximum overlaps in nearly all IC realisations.

2.4.1.3 Probability of being matched

Lastly, we calculate the probability that a reference halo has a counterpart in *any* other IC realisation, as given by Eq. (2.4). In Fig. 2.6 we plot the mean and standard deviation of this probability, averaged over the remaining IC realisations. If a halo has no potential match in another simulation, then the sum of its overlaps in that simulation is simply zero. Mirroring previous results, there is a clear distinction between CSiBORG and Quijote above $\sim 10^{14}\ M_\odot/h$, in that the majority of CSiBORG haloes are matched while the majority of Quijote haloes are not. Below this mass, the distinction weakens, though it remains significant. The uncertainty on the probability of a match shown in Fig. 2.6b peaks at



(a) Mean probability of a match averaged over the remaining IC realisations.

(b) Standard deviation of the probability of a match averaged over the remaining IC realisations.

Figure 2.6: Probability of a reference halo having a match in the remaining IC realisations calculated as the sum of overlaps. The hex bins show the CSiBORG data and the lines are the mean trends in CSiBORG and Quijote, with 1σ error bars characterising the spread among points. The probability of a match is typically significantly higher in CSiBORG except for the lower mass threshold where the constraints weaken. The truncation at low masses in Fig. 2.6a, and the reduction in uncertainty toward very low masses in Fig. 2.6b, are due to the lower mass threshold of $10^{13.25} M_\odot/h$ when matching haloes.

$10^{14}\, M_\odot/h$. Below this mass, and particularly near the lower mass threshold of $10^{13.25} M_\odot/h$, there are only haloes with a mass above this threshold to overlap with and not below it, thus underestimating the uncertainty.

This is complementary to the results of Fig. 2.3b, which was, instead, sensitive to the maximum overlap a reference halo has with another simulation. On the other hand, in Fig. 2.6 a high probability of a match can also be due to adding many small overlaps.

In both CSiBORG and Quijote there is a trend that more massive haloes have higher sum of overlaps, although this is more pronounced in CSiBORG. However, this is not surprising, since the most massive haloes have initial Lagrangian patches of $\sim$ 10 Mpc$/h$ and thus naturally overlap with more objects. For comparison, see the trend of Fig. 2.3b where the mean trend of the maximum overlaps in Quijote never rises—while the large haloes have overlaps with many small ones such that the sum of overlaps may increase, the maximum overlap a given reference halo has with any one halo in a crossing simulation does not

increase with mass.

In Fig. 2.7, we show in CSiBORG the relation between the probability of being matched and the maximum overlap averaged over the IC realisations. The two quantities are strongly correlated, though the most massive haloes deviate from the $1-1$ line, as smaller overlaps start to contribute more significantly to the sum over overlaps. Nevertheless, even for the most massive haloes, this shows that sum over overlaps is typically dominated by one object that has a large overlap with the reference halo.

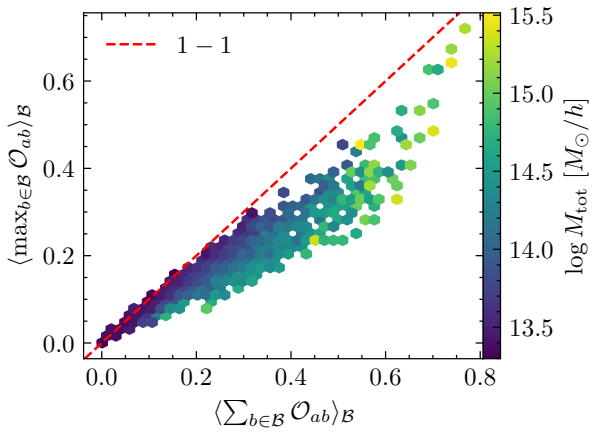


Figure 2.7: Relation between the mean summed overlap and mean maximum overlap of reference haloes, both averaged over the remaining CSiBORG IC realisations. The most massive haloes deviate from the 1–1 line due to contributions of smaller overlaps, but the summed overlaps are typically dominated by a single object, making the match relatively unambiguous.

2.4.2 Halo properties of overlapping haloes

Now that we have explored the properties of the overlap statistic and its behaviour across the CSiBORG suite, we turn our attention to investigating the properties of haloes that it matches. This includes their separation (Section 2.4.2.1), mass (Section 2.4.2.2), peculiar velocity (Section 2.4.2.3) and concentration and spin (Section 2.4.2.4).

2.4.2.1 Final snapshot separation of overlapping haloes

For a reference halo that has a significant overlap with a halo in another IC realisation, our anticipation is that their proximity in the ICs will make the pair unusually close in the final snapshot as well. Because there are nearby haloes between boxes even in the absence of any IC constraints, we juxtapose our results with those from the unconstrained Quijote

suite. Any suppression in distance observed in CSiBORG relative to Quijote can then be ascribed to the constraints.

We show the results in Fig. 2.8. We cross-match all IC realisations to a single reference simulation and compute $\langle \Delta R \rangle$, the overlap-weighted average separation of haloes in the final snapshot averaged over all IC realisations. We calculate the separation as $\Delta R = \|(\boldsymbol{x}_a - \boldsymbol{x}_b)\|$, where $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ are the position of the reference and matched halo, respectively, and $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}$. In CSiBORG haloes are consistently more likely to remain close in the final snapshot if they originate from the same Lagrangian patch. The lowest mass matched objects in CSiBORG show a variation of ~ 10 Mpc$/h$. For an individual reference halo and its maximum-overlap matches from the remaining IC realisations we find, as expected, a strong negative correlation between the overlap value and their $z = 0$ separation.

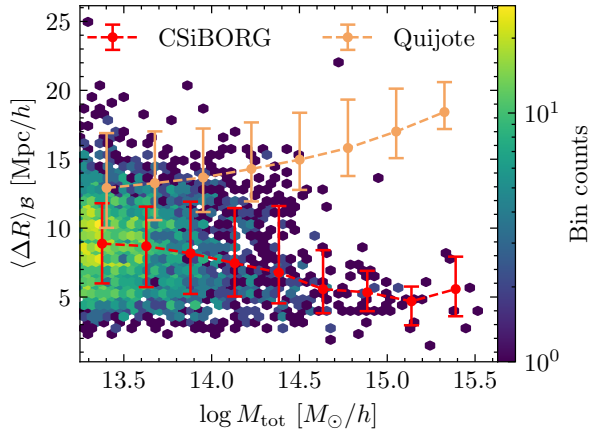


Figure 2.8: The overlap-weighted mean separation ΔR of haloes in the final snapshot of CSiBORG realisation 7444, averaged over all remaining IC realisations. Hex bins show CSiBORG overlaps; lines show the CSiBORG and Quijote mean trends with 1σ spread. Halo positions are constrained across the entire mass range relative to Quijote, but with significant scatter at low masses.

2.4.2.2 Mass of matched haloes

The next quantity that we look at is the most probable mass of the matched haloes. We calculate this following the approach outlined at the end of Section 2.3. From each crossing simulation, we find the mass of the maximally overlapping halo and construct a weighted histogram, where the weights are the maximum overlaps. We then define the most probable mass as the mode of this distribution, with its uncertainty given by the square root of the overlap-weighted average square of residuals around the mode. We plot an example of

this in the left panel of Fig. 2.9 for the most massive halo in a single CSiBORG realisation, finding the most likely mass to be within 0.2 dex of the reference halo mass. This is in part by construction, since the overlap itself is preferentially higher for haloes that have a similar mass. However, while the overlap is preferentially higher for haloes of a similar mass, it does not guarantee the objects being a "good" match. If that were the case, then we would find that even in Quijote the mass of the matched haloes is on average close to the mass of the reference halo, which it is not.

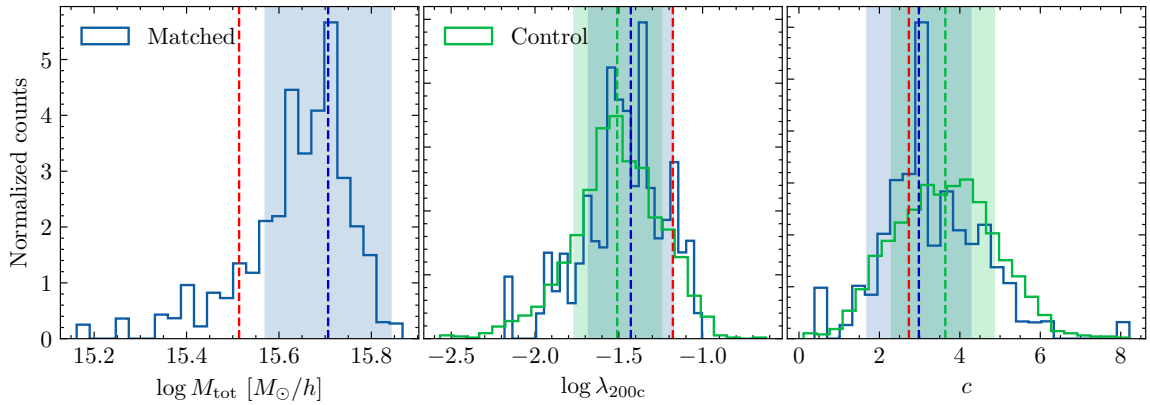


Figure 2.9: Distributions of most likely matched haloes' properties ($\log M_{\rm tot}$, $\log \lambda_{\rm 200c}$, c) with the most massive high-resolution halo in one IC realisation (7444) of CSiBORG shown as the blue histogram. The green "control" histogram is the spin and concentration of haloes with a similar mass from the remaining IC realisations: from each we take the 10 haloes closest in mass to the reference halo. The blue and green vertical lines are the mode of the "matched" and "control" histograms, respectively, and the corresponding shaded bands are 1σ uncertainties. The red line is the reference halo property. A comparison to the control distribution reveals that neither the spin nor the concentration is constrained. The matched concentration of this halo has a sharp peak near the reference concentration, however a similar trend is *not* observed for other massive haloes.

By analysing histograms like Fig. 2.9, we calculate and show in Fig. 2.10 the most likely mass of all reference haloes in CSiBORG above $10^{13.25}$ $M_\odot/h$. In this figure, we show three panels: comparison between the reference and most likely halo mass, the uncertainty of the most likely mass as a function of the reference mass and the ratio of the most likely mass to the reference mass as a function of the median probability of being matched.

The reference and most likely matched halo masses agree well for the most massive

objects, with deviations from the $1-1$ line increasing towards lower mass. However, these objects also have a consistently lower probability of being matched (right panel of Fig. 2.10) and, therefore, it is not unexpected. Below the scale of $\sim 10^{14}\ M_\odot/h$ the matching is less robust: we only cross-match haloes with mass above $10^{13.25}\ M_\odot/h$ since below this scale the IC constraints weaken as indicated by Fig. 2.3b and Fig. 2.6a. Therefore, a reference halo near the threshold will only have potential matches that are above this threshold, thus biasing the matching procedure.

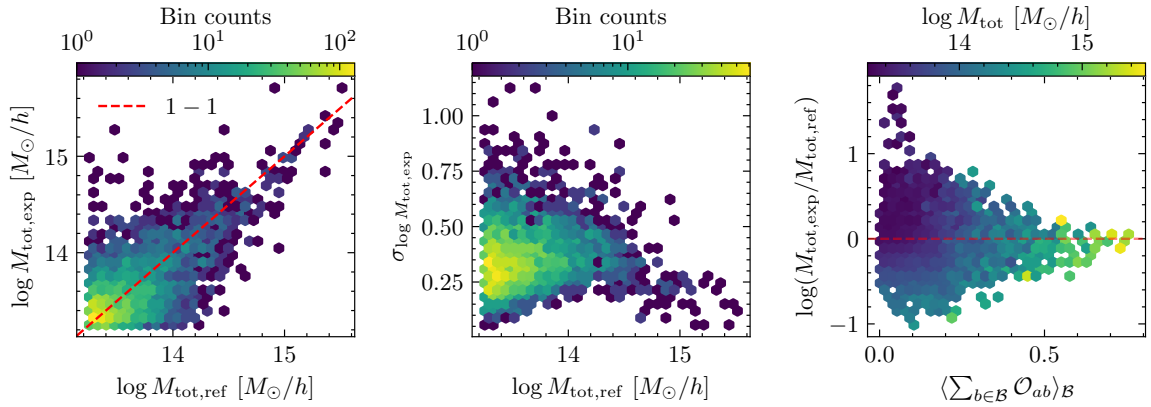


Figure 2.10: Most likely mass of haloes matched to a single realisation of CSiBORG averaged over all crossing simulations and defined as the mode of a histogram such as Fig. 2.9. The *left* panel shows the reference-to-expected halo mass relation, the *middle* panel shows the uncertainty of the expected mass and the *right* panel shows the ratio of the expected mass to the reference mass as a function of probability of being matched defined in Eq. (2.4). The agreement between the reference and matched mass is strongly correlated with the matching probability.

2.4.2.3 Peculiar velocity alignment

In Fig. 2.11, we show the alignment and ratio of the magnitudes of the peculiar velocities in the final snapshot of CSiBORG. We define the alignment angle θ between the peculiar velocity vectors of haloes a and b with peculiar velocities $\boldsymbol{v}_a$ and $\boldsymbol{v}_b$, respectively, as:

$$\cos\theta = \frac{\boldsymbol{v}_a \cdot \boldsymbol{v}_b}{|\boldsymbol{v}_a|\,|\boldsymbol{v}_b|}. \tag{2.6}$$

For each reference halo we calculate the alignment of its peculiar velocity vector with the corresponding vector of the highest overlapping halo from another IC realisation. We find that the alignment of a reference halo and a single maximum overlap crossing halo is strongly correlated with the magnitude of this overlap across the IC realisations. The alignment and ratio of the magnitudes plotted in Fig. 2.11 is averaged over the crossing IC realisations. While most matched cluster-mass haloes show strong alignment in the final snapshot, there are exceptions. These often coincide with a lower mean maximum overlap. Nevertheless, even plotting the alignment as a function of the mean maximum overlap, there remains a significant scatter. We also compare the magnitudes of peculiar velocities, finding their ratios to be on average ~ 1, however, with significantly larger scatter towards smaller overlaps.

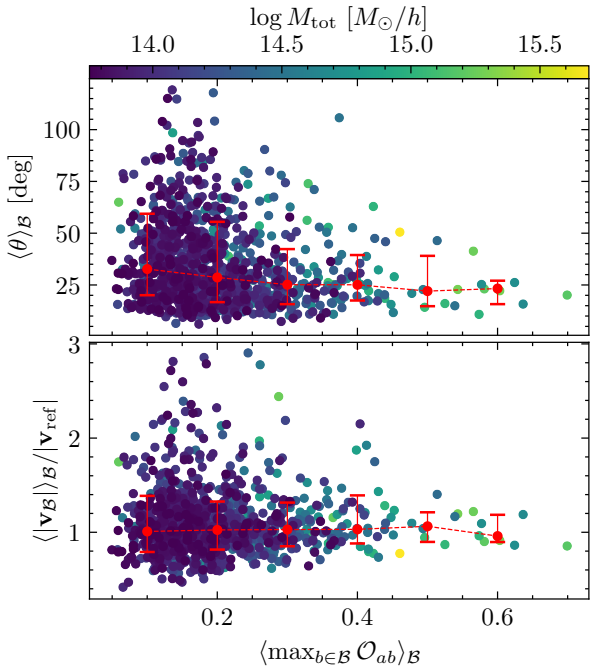


Figure 2.11: Mean alignment angle (*top*) and ratio of magnitudes (*bottom*) of peculiar velocities of reference haloes with the remaining IC realisations in `CSiBORG`, as a function of the mean maximum overlap. The red line shows the mean trend with 1σ spread. Higher-overlap haloes have better-aligned peculiar velocities in the final snapshot.

2.4.2.4 Spin and concentration constraint

Lastly, we turn our attention to the spin and concentration of these haloes. We use the Bullock spin definition,

$$\lambda_{200c} = \frac{J_{200c}}{\sqrt{2} M_{200c} V_{200c} R_{200c}}, \tag{2.7}$$

where J_{200c} is the angular momentum magnitude of particles within R_{200c} and $V_{200c}^2 = GM_{200c}/R_{200c}$ (Bullock et al., 2001). We define concentration as the ratio of the virial radius to the scale radius of the Navarro-Frenk-White (NFW) profile, $c = R_{200c}/R_s$ (Navarro et al., 1996).

We calculate the most likely spin and concentration of the matched haloes following the approach outlined in Section 2.3. In the middle panel of Fig. 2.9 we show the comparison of the spins of overlapping haloes to the spin of a reference halo, which we take to be the most massive halo in one IC realisation of `CSiBORG`. However, unlike previously, we do not find any preference for the spin of overlapping haloes to be similar to the reference halo spin, regardless of whether we weight the matched haloes by overlap or not. The matched distribution in Fig. 2.9 has a secondary peak near the reference spin; however, it is not statistically significant. In fact, the distribution of the matched spins is in good agreement with the simulation average, which is approximately a mass-independent Gaussian distribution in $\log \lambda_{200c}$. We find similar conclusions to hold for all haloes, regardless of their mass.

Next, we investigate the concentration of the matched haloes. In the right panel of Fig. 2.9 we show the comparison of the concentrations of overlapping haloes to the concentration of a reference halo, which we again take to be the most massive halo in one IC realisation of `CSiBORG`. We find that in this particular example, the mode of the weighted distribution agrees well with the reference halo concentration, but the width of this distribution is still similar to the expectation from the simulated mass–concentration relation. To delve deeper, we assess the matched concentrations for all haloes with a mass exceeding $10^{15}\ M_\odot/h$,

which we previously identified as being consistently reconstructed. We find no discernible correlation between the most likely and reference concentrations. We also find that the concentration of the highest overlap halo from each IC realisation is within 10 (30) per cent of the reference value in only 20 (50) per cent of realisations, without any clear dependence on the reference halo mass. The agreement seen in Fig. 2.9 is thus coincidental or specific to the halo in question. However, even in such instances, the significance is questionable as there is no notable improvement over the mass–concentration relation.

2.5 Discussion

2.5.1 Interpretation of the results

Constrained simulations enable an object-by-object comparison of the local Universe with theory. If a direct, or at least probabilistic, correspondence between a simulated halo and an observed structure can be established, constrained simulations potentially allow inference of the properties and assembly histories of nearby objects and hence pose rigorous tests for galaxy formation models and cosmology. In this work, we outline and test a method for assessing whether a halo is robustly reconstructed across a set of simulations, differing for example in ICs drawn from the posterior of a preceding inference. This allows us to compare the properties of the "same" haloes across the simulations. By "robustly reconstructed," we refer specifically to the suppression of variance in halo properties across the sampled ICs, meaning the haloes originate from the same Lagrangian patches and have similar properties at $z = 0$. An important additional aspect in assessing the quality of any local Universe reconstruction is how faithfully it reproduces the observed Universe. However, we do not test this aspect in the present work.

We illustrate our framework by applying it to the `CSiBORG` suite of 101 constrained simulations of the local ~ 155 Mpc$/h$ Universe with ICs on a grid of spacing 2.65 Mpc$/h$ derived from the `BORG` inference of the 2M++ catalogue. We find that cluster-mass haloes

($M \gtrsim 10^{14}\ M_\odot/h$) are consistently reconstructed across the suite and originate from the same Lagrangian region. Haloes of this mass are distributed over ~ 70 cells in the initial snapshot at $z = 69$. Assuming that the comoving cell density in the initial snapshot is $\Omega_\mathrm{m}\rho_\mathrm{c}$, where ρ_c is the current critical density of the Universe, we can approximate the spatial resolution, L, required to distribute a mass M over N cells as:

$$L \approx 2.6\ \mathrm{Mpc}/h \left(\frac{M}{10^{14}\ M_\odot/h} \frac{70}{N} \right)^{1/3}. \tag{2.8}$$

Assuming the criterion of 70 resolution elements across the initial Lagrangian patch to be universal, this suggests that for haloes of mass 10^{15}, 10^{14}, $10^{13}\ M_\odot/h$ to be robustly reconstructed one must use IC constraints at the scales 5.5, 2.6 and 1.2 Mpc/h, respectively.

While high-mass haloes are consistently reconstructed and have counterparts of similar mass and peculiar velocity across all IC realisations, the overlapping haloes originating from the same Lagrangian regions are similar in neither their spin nor concentration. Despite the BORG reconstruction employed in this work not utilising a peculiar velocity catalogue as a constraint—relying solely on galaxy positions in the 2M++ catalogue—it is reassuring to find that the highest mass haloes indeed have aligned velocities, since to some extent galaxy positions are complementary to the peculiar velocity field information. In fact, BORG has been used in the past to reconstruct the local peculiar velocity field (Jasche & Lavaux, 2019).

Cadiou et al. (2021) demonstrated that the angular momentum of haloes can be accurately predicted directly from their Lagrangian patches, but that it exhibits chaotic behaviour under small changes to the patch boundary. Although the high-mass CSiBORG haloes are present in all IC realisations, their overlaps never reach unity and so their Lagrangian patches are not exactly aligned. Therefore, given the findings of Cadiou et al., the lack of constraint on spin is not surprising. The halo concentration depends on both the Lagrangian patch configuration and subsequent accretion history (Rey et al., 2019). While we find that on average the halo

concentration in CSiBORG is not constrained, we leave a detailed examination of specific observed clusters for future work along with comparing the mass accretion histories of haloes that are initially strongly overlapping. This will help tease out the properties of haloes and their constituent particles that are responsible for setting their concentrations.

For certain haloes, we observe that those originating from consecutive CSiBORG simulations tend to have higher maximum overlaps. CSiBORG resimulates every 24th step of the BORG chain, yet the autocorrelation length of the chain is ~ 100. CSiBORG thus effectively oversamples the BORG posterior—even though it independently varies the unconstrained small-scale modes at each step—which could lead to an overestimation of confidence in the BORG constraints if relying only on a few consecutive CSiBORG samples. To mitigate this effect, we have averaged across all remaining 100 IC realisations for each reference realisation, the majority of which are fully decorrelated from it. A fully satisfactory solution would be to resimulate only decorrelated IC realisations, however it is challenging to derive a large number of these due to the high computational cost of the BORG inference.

2.5.2 Comparison with the literature

Hutt et al. (2022) introduce an algorithm to assess whether "twins" of a single halo can be identified in all IC realisations of a constrained simulation. This is done by selecting a reference simulation and cataloguing the positions of all haloes. A halo is then chosen from the reference simulation, and its size is calculated as $R_{200\text{c}} = [(3M_{200\text{c}})/(4\pi \cdot 200 \cdot \rho_{\text{crit}})]^{1/3}$. In the other simulations, the haloes within a designated "search radius" of the reference halo are then identified and the one most similar in mass selected. If no haloes are found within the search radius, the reference halo is discarded as not present within the crossing simulation. This approach differs from ours in several important ways:

1. we match haloes directly in the ICs, which are constrained by BORG, instead of relying on the forward-modelled realisations of the final snapshot,

2. we do not require the matched haloes to be within a fixed radius of one another, but instead calculate the overlap of their initial Lagrangian patches,

3. while the procedure of Hutt et al. is inherently binary, ours has a continuous interpretation in terms of probabilities.

It is nevertheless instructive to compare our results. In Fig. 2.12 we show the fraction of matches identified via the approach of Hutt et al. that are also the maximum-overlap pairs for several choices of search radii in `CSiBORG`. Due to the stringent requirement of Hutt et al. that a halo must have a "twin" in *all* IC realisations, there is a strong agreement between the two approaches. We calculate $f_{\mathrm{agreement}}$, the fraction of matches identified by Hutt et al. that are also the maximum overlap pairs as a function of a lower mass threshold. For the fiducial search radius used by Hutt et al. the agreement is ~ 90 per cent regardless of the mass threshold. However, this comes at the cost of only a small number of haloes being identified as matches by Hutt et al.: for example, at $10^{14}\ M_\odot/h$ this is only ~ 10 per cent of the haloes. Our method has the advantage of providing useful information even in the regime where the matching across all realisations is partial or even weak.

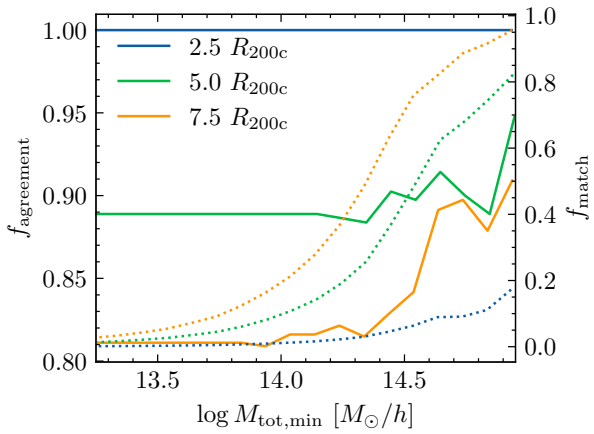


Figure 2.12: Comparison of our matching procedure to that of Hutt et al. (2022), who identify "twin" haloes in the final snapshots of `CSiBORG`. Solid lines show $f_{\mathrm{agreement}}$, the fractional agreement between the Hutt et al. matches and the maximum-overlap pairs above mass $M_{\mathrm{tot,min}}$. Dotted lines show f_{match}, the fraction of haloes identified as matches by Hutt et al.. The high $f_{\mathrm{agreement}}$ indicates excellent agreement regardless of search radius, but only a small fraction of haloes are identified as matches by Hutt et al.; our method assigns a match probability even to uncertain matches.

Another methodology with similarities to ours is that of Pfeifer et al. (2023), in which the goal is to identify the simulation most representative of the local Universe. However, a

crucial difference between our work and that of Pfeifer et al. is that we focus on studying the consistency of constrained haloes across the posterior samples of an IC inference, rather than assessing the reliability of matching these haloes to the local Universe. They use a Wiener filter-based reconstruction that must be supplied with random small-scale/unconstrained modes. This strategy can alternatively be conceptualised as an intensive fine-tuning process to pinpoint the optimal random seed on a very fine grid level, which can then be re-simulated (McAlpine et al., 2022). Pfeifer et al. find the realisation that best resembles the local Universe by minimising the sum of p-values of simulated cluster-mass haloes being at the locations of observed clusters. However, while they find the most representative constrained simulations, their approach does not provide any information about the consistency with which the matched simulated haloes are reconstructed. We believe that this is crucial information to determine the confidence with which we can assert the properties of the dark matter distribution of the local Universe, given the quality and quantity of data used to set the constraints.

A similar technique to match observed galaxies to haloes from constrained cosmological simulations has been applied by Zhang et al. (2022); Xu et al. (2024). Zhang et al. develop a "neighbourhood" subhalo abundance matching (SHAM), a SHAM model specifically tailored for constrained cosmological simulations which ranks haloes based on both their peak mass and closeness in position and velocity to the observed galaxy (Yang et al., 2018). This statistical approach enables them to connect haloes in their single constrained simulation with observed galaxies, thereby facilitating the study of, e.g., the galaxy-to-halo size relation. Our method would allow folding in the uncertainties associated with the reconstruction.

2.6 Conclusion

We have investigated the extent to which haloes are robustly reconstructed across multiple cosmological simulations that sample the IC posterior of a previous inference. While this question is particularly pertinent to simulation suites constrained (with uncertainty) to match the local Universe, other applications include the matching of haloes between DM-only simulations and their hydrodynamical counterparts, or simulations of varying cosmology such as ΛCDM vs modified gravity.

We argue that for a halo to be consistently reconstructed, its Lagrangian patch in the initial conditions must strongly overlap with a halo in most other realisations of the suite, a condition implying similarity in both mass and location. This is a stricter and more causal measure than similarity in the final snapshot, providing a clearer condition for two haloes to be "the same". In the future, an even stricter measure of similarity may be introduced by measuring the haloes' initial overlap in the 6D phase space directly, instead of only the 3D position space.

We apply the method to `CSiBORG`, a suite of constrained simulations with initial conditions sampled from the posterior of the `BORG` algorithm applied to the 2M++ galaxy number density field. We establish the significance of the results by comparing to the unconstrained `Quijote` suite. Based on the criteria mentioned above, we find that cluster-mass haloes ($M \gtrsim 10^{14} M_\odot/h$) are consistently reconstructed in `CSiBORG` in position, mass, and peculiar velocity, with higher mass haloes typically more strongly constrained. For haloes below this mass threshold, the constraints diminish, and haloes do not consistently originate from the same Lagrangian patches. Regarding secondary halo properties such as spin and concentration, even consistently matched, high-mass haloes display variations across the IC realisations. The absence of constraints on concentration beyond the mass-concentration relation is surprising given its strong dependence on mass assembly history. This might imply that the assembly history of even the most massive haloes in `CSiBORG` remains uncon-

strained; however, we defer a comprehensive study of the mass assembly history to future work. In the future, we will also investigate the extent to which halo overlap correlates with reliability of match to an observed object in the local Universe. In sum, our framework provides a step towards the goal of identifying the ICs that led to the observed Universe and the objects it contains.

3

The Velocity Field Olympics

The previous chapter quantified BORG's internal precision; external accuracy requires an independent test. I run a posterior predictive check of the reconstructed velocity field against peculiar velocity catalogues, benchmarking BORG against five alternative reconstructions in a unified Bayesian framework.

Based on Stiskalek et al. (2026c), published in Monthly Notices of the Royal Astronomical Society.

3.1 Introduction

Spectroscopic analysis of the sky, particularly of galaxies, provides fundamental insights into the nature of our Universe. The spectroscopic data are primarily encoded in redshift measurements, which combine two components: the uniform cosmic expansion and the peculiar motions of galaxies relative to this expansion. The peculiar velocity field constrains the growth rate $f\sigma_8$ (Nusser, 2017; Huterer et al., 2017; Boruah et al., 2020; Said et al., 2020) and enters directly into local H_0 determinations such as SH0ES (Riess et al., 2022a) and gravitational-wave standard sirens (Abbott et al., 2021), yet the accuracy of the underlying velocity field reconstructions has never been systematically compared.

This chapter introduces a self-consistent Bayesian framework to benchmark the local

Universe velocity field reconstructions currently available in the literature. The framework accounts for homogeneous and inhomogeneous Malmquist biases, distance-tracer selection effects, and joint calibration of the distance-indicator relation. We apply it to six reconstructions spanning linear inverse methods (Carrick et al., 2015), machine-learning approaches (Lilow et al., 2024), Wiener-filter constrained realisations (Sorce & Tempel, 2018; Sorce, 2020), forward modelling of peculiar velocities (Courtois et al., 2023), and the *Bayesian Origin Reconstruction from Galaxies* (BORG) forward model (Jasche & Lavaux, 2019; Stopyra et al., 2024a), calibrating each against Tully–Fisher and Type Ia supernova distance indicators to determine which local Universe reconstructions provide the best explanation of the peculiar velocity data, and to identify potential systematic biases in each model.

The chapter is structured as follows. In Section 3.2, we describe the peculiar velocity field models derived from reconstructions and distance data used in this work. In Section 3.3, we introduce our statistical and validation framework. In Section 3.4, we present the results of our analysis, which include the comparison of the evidence for the different velocity models and consistency parameters with cosmology. In Section 3.5, we discuss the implications of our findings. Finally, in Section 3.6, we summarise our findings.

3.2 Data

In this section, we discuss first the different kinds of velocity model reconstructions we use in this work (Section 3.2.1), and then the distance data we use to validate these models (Section 3.2.2).

3.2.1 Local Universe reconstructions

We use several reconstructions of the local Universe, constrained by either redshift-space galaxy positions or peculiar velocities. The former approach typically employs a counts-

per-voxel likelihood and requires a galaxy bias to predict the number of galaxies in a cell based on the cell's matter density, after which it is limited mainly by shot noise. The latter relies on a (Gaussian) likelihood defined in the projected peculiar velocity, but is limited by the large noise in peculiar velocity measurements. We summarise these reconstructions in Table 3.1, and describe them one-by-one here.

3.2.1.1 Carrick et al.

In Carrick et al. (2015, hereafter C15), the luminosity-weighted density field is derived from the redshift-space positions of galaxies in the 2M++ catalogue, using the iterative scheme of Yahil et al. (1991). In this scheme each galaxy is weighted by its K-band luminosity rather than counted equally, so the resulting galaxy overdensity is a mass-tracing proxy rather than a pure number-count field. The weights are further corrected for the survey's magnitude limit, compensating for the fainter galaxies that go undetected with increasing distance, and the resulting field is normalised by a luminosity-dependent bias before it is used to predict the linear-theory velocity field. 2M++ is a whole-sky redshift compilation of 69, 160 galaxies (Lavaux & Hudson, 2011), derived from photometry from the Two-Micron-All-Sky Extended Source Catalog (Skrutskie et al., 2006) and redshifts from the 2MASS Redshift Survey (2MRS, Huchra et al., 2012), the 6dF Galaxy Redshift Survey (Jones et al., 2009), and the Sloan Digital Sky Survey (SDSS) Data Release 7 (Abazajian et al., 2009). The 2M++ K-band apparent magnitudes are corrected for Galactic extinction, k-corrections, evolution, and surface brightness dimming. The catalogue is apparent magnitude limited to $K < 11.5$ in the regions of sky covered by 2MRS and to $K < 12.5$ within the areas covered by 6dF and SDSS.

C15 first fits a luminosity function to the 2M++ catalogue to assign luminosity weights to each galaxy. Then, the luminosity-dependent bias of Westover (2007) is applied to normalise the density contrast across all radii. Galaxy positions are moved from the redshift-space to real-space using an iterative approach. Within the inner region of approximately

125 h^{-1} Mpc the Zone of Avoidance is masked in the 2M++ data and filled by "cloning" galaxies above and below the Galactic plane to reconstruct the density field in this region. The resulting velocity field is calculated from the density field assuming linear perturbation theory. The velocity field is independent of the dimensionless growth factor, and thus must be scaled to match the peculiar velocities (with a factor $\beta^\star$ in the notation of C15). This single field is generated on a 256^3 grid with a box size of 400 h^{-1} Mpc, a voxel size of 1.6 h^{-1} Mpc, and assuming $\Omega_\mathrm{m} = 0.3$.

The parameter $\beta^\star$ is defined by

$$\beta^\star \equiv \frac{f\sigma_{8,\mathrm{NL}}}{\sigma_8^b} \tag{3.1}$$

where $f \approx \Omega_\mathrm{m}^{0.55}$ is the dimensionless growth rate (Section 1.2.1). The term σ_8^b refers to the typical fluctuation scale in the galaxy over-density field at a radius of 8 h^{-1} Mpc (the 'b' denotes that it is measured from a biased tracer, galaxies). $\sigma_{8,\mathrm{NL}}$ is the typical fluctuation scale of the non-linear matter field at a radius of 8 h^{-1} Mpc. The value of σ_8^b in the 2M++ data has been measured by Westover (2007) to be 0.98 ± 0.07 by fitting projected correlation functions to 2MRS galaxies (Huchra et al., 2012), whereas C15 measured $\sigma_8^b = 0.99 \pm 0.04$ by using the maximum-likelihood scheme of counts in cells within radial shells (Efstathiou et al., 1990). Thus, C15 together with peculiar velocity samples can be used to constrain the growth of structure or the S_8 parameter (see e.g. Boruah et al. 2020; Said et al. 2020). The value of $\beta^\star$ in the case of C15 calibrated against LOSS, Foundation, 2MTF, and SFI++, has been previously studied extensively in Boruah et al. (2020).

3.2.1.2 Lilow et al.

In Lilow et al. (2024, hereafter L24), the density and peculiar velocity fields are reconstructed using the 2MRS data, but with a machine learning approach, rather than linear theory. L24 employ an autoencoder with a U-Net architecture, previously tested in their earlier

work (Ganeshaiah Veena et al., 2023). The network was trained separately for density and velocity on `Quijote`-based N-body simulation mocks (Villaescusa-Navarro et al., 2020) including redshift-space distortions, galaxy bias, and the 2MRS selection function, with the velocity field constrained to be irrotational. Comparison to Wiener filter reconstructions on mock data showed substantially lower reconstruction errors and an accurate recovery of the non-linear velocity–density relation. Notably, the network captures information on large-scale bulk flows through redshift-space distortions, and when applied to 2MRS it recovers known clusters and a Local Group velocity closely aligned with the observed CMB dipole. The trained model is subsequently applied to the 2MRS data. L24 assume cosmological parameters from the fiducial `Quijote` simulations: $\Omega_m = 0.3175$, $\Omega_\mathrm{b} = 0.049$, $h = 0.6711$, $n_s = 0.9624$, $\sigma_8 = 0.834$. The single field is presented on a 128^3 grid with a box size of 400 h^{-1} Mpc and a voxel size of 3.1 h^{-1} Mpc.

3.2.1.3 `CSiBORG`

The *Constrained Simulations in* `BORG` (`CSiBORG`) suite (Section 1.4.6) resimulates posterior samples of the `BORG` 2M++ initial conditions as full N-body calculations, so that variation across the ensemble quantifies the reconstruction uncertainty. The first-generation suite, referred to as `CSiBORG` in Chapter 2 and hereafter `CSiBORG1`, is supplemented in this chapter by a second generation, `CSiBORG2`, based on an upgraded `BORG` posterior.

Here we employ two versions of the `CSiBORG` simulations, based on different versions of the `BORG` 2M++ initial conditions (ICs). The first suite, `CSiBORG1`, uses the `BORG` 2M++ ICs from Jasche & Lavaux (2019), while the second suite, `CSiBORG2`, is based on the updated `BORG` 2M++ ICs from Stopyra et al. (2024a). Both versions of the ICs are constrained within the 2M++ region, covering a spherical volume with a radius of approximately 155 h^{-1} Mpc centred on the Milky Way ($z \lesssim 0.06$), embedded in a box of 1000 Mpc. The regions outside the 2M++ volume are largely unconstrained due to catalogue selection, providing the expected long-range forces in a Λ-cold dark matter (ΛCDM) framework. The

BORG algorithm models the density field on a 256^3 grid, yielding a spatial resolution of 2.65 h^{-1} Mpc.

The CSiBORG1 suite is described in Section 2.2.1. For this work, we use only 20 of the 101 resimulations to reduce the computational cost, though we verify that our results are not significantly affected by using all 101. The CSiBORG2 suite follows a similar setup but uses newer BORG ICs (Stopyra et al., 2024a), which use a more accurate gravity model: a 20-step COLA integrator (Tassev et al., 2013) instead of the 10-step particle mesh solver of CSiBORG1. This was shown to improve the agreement of cluster masses with observational data. The cosmological parameters for CSiBORG2 are based on Planck Collaboration et al. (2020a), with lensing and baryon acoustic oscillations: $\Omega_m = 0.3111$, $\sigma_8 = 0.8102$, $H_0 = 67.66$ km s^{-1} Mpc^{-1}, $n_s = 0.9665$, $\Omega_\mathrm{b} = 0.049$. We resimulate 20 samples of the BORG posterior ourselves using Springel2021 (Springel et al., 2021). We note that Stopyra et al. (2024a) already resimulated 20 samples of the BORG posterior, but without the zoom-in technique: the CSiBORG setup allows us to reach ~200 times better mass resolution in the constrained region. In both CSiBORG suites, we construct the $z = 0$ density and velocity fields from the particle snapshots using the smoothed-particle hydrodynamics (SPH) technique (Monaghan, 1992; Colombi et al., 2007) and set the minimum number of points to smooth over to 32 (following Section IV.B.1 of Bartlett et al. 2022). We render the final fields in voxels of 0.7 h^{-1} Mpc.

Unlike C15, which is derived independently of the growth factor and consequently of σ_8, the derivations of the BORG-based reconstructions (and all others presented in this work) assume fixed values for Ω_m and σ_8. They naturally account for the $\beta^\star$ factor of Eq. (3.1), which is effectively pinned to 1. However, to search for systematics in the reconstructions leading to a mismatch with the peculiar velocity data, we allow the velocity fields other than C15 to be scaled by a free parameter β, with the fiducial value of 1, similarly to the A_L parameter in cosmic microwave background (CMB) lensing (Calabrese et al., 2008). We retain the notation of $\beta^\star$ for the scaling factor of C15, who measured $\beta^\star = 0.431 \pm 0.021$.

3.2.1.4 Sorce

The reconstructions presented in Sorce & Tempel (2018); Sorce (2020, hereafter S18) employ the technique of constrained realisations. Unlike redshift space density–constrained methods, these simulations use only radial peculiar velocities—drawn from the grouped CosmicFlows-2 (CF2) catalogue (Tully et al., 2013) of ∼5000 objects—as constraints. Grouping suppresses virial motions, while a bias–minimisation scheme iteratively restores Gaussianity to the velocity distribution, reducing spurious infall. The cosmic displacement field is reconstructed with a Wiener filter and constraints are relocated to progenitor positions using the reverse Zel'dovich approximation, with noisy radial velocities replaced by their Wiener–filtered 3D estimates. Constrained realisations then combine these fields with Gaussian random modes to statistically restore missing large- and small-scale power.

S18 thus produces ICs constrained solely by peculiar velocities, which are evolved to $z = 0$ with `Springel2005` (Springel, 2005) in a 500 h^{-1} Mpc box (512^3 particles with a particle mass of $8 \times 10^{10} h^{-1}\,\mathrm{M}_\odot$). Subsequent refinements introduced improved modelling of peculiar–velocity uncertainties, enabling stable simulacra of local clusters (Virgo, Centaurus, Hydra, Coma). Sorce (2020) incorporated the "paired fixed field" technique (Pontzen et al., 2016), demonstrating that about 80 per cent of the large–scale power is already constrained by local data and provided a method to estimate uncertainties without hundreds of realisations. We use a single such simulation, assuming Planck cosmology (Planck Collaboration et al., 2014), and construct SPH density and velocity fields on a 0.7 h^{-1} Mpc grid.

3.2.1.5 Courtois et al.

In Courtois et al. (2023, hereafter C23), the density and velocity fields of the local Universe below $z_{\mathrm{CMB}} = 0.08$ are reconstructed using the CosmicFlows-4 catalogue (CF4; Tully et al. 2023), which includes distances and redshifts for 38,000 groups. Similarly to S18, the reconstruction is constrained by peculiar velocities and employs a forward-modelling

procedure with a Hamiltonian Monte Carlo (HMC) algorithm (Graziani et al., 2019).

Given the reported CF4 redshifts and distances, they sample the "true" distance to each source, the density field along with cosmological and nuisance terms such as the growth rate or the non-linear velocity dispersion $\sigma_{\rm NL}$, while also marginalising over a zero-point parameter $h_{\rm eff}$ to absorb calibration differences between distance indicators. Under the assumption of linear theory, the velocity field is derived from the density field and compared with the observed peculiar velocities. The reconstruction is performed on grids of resolution from 64^3 to 256^3. We use the grouped CF4 version and adopt 100 posterior samples at 256^3 resolution. From these reconstructions, they obtain a growth rate $f\sigma_8 = 0.36 \pm 0.05$ (grouped CF4). The implied bulk flow is 230 ± 136 km s^{-1} at 300 h^{-1} Mpc, and distances are compatible with $H_0 \simeq 74.5 \pm 0.1$ km s^{-1} Mpc^{-1}.

3.2.1.6 Visual comparison

We show slices through the density field at $z = 0$ from the six reconstructions we consider in Fig. 3.1. C15 and L24 are qualitatively similar, but do not display significant cosmic web-like features. In contrast, the CSiBORG fields reveal a prominent cosmic web structure due to their use of non-linear modelling. This is true even after averaging the density field over 20 independent posterior samples, suggesting that it is not due to the unconstrained white noise of these simulations. The C23 field appears more diffuse despite having a voxel resolution of 3.9 h^{-1} Mpc (similar to C15), which suggests that despite the voxel resolution the field varies only at much larger scales. Lastly, while the field of S18 is also constrained at the level of linear theory, it was resimulated at a much higher resolution with additional small-scale white noise, resulting in a clear cosmic web.

Unlike the CSiBORG reconstructions, the S18 field is a single realisation and, thus, the cosmic web features are more likely due to unconstrained, small-scale modes, not matching the actual local Universe. We also show the line-of-sight (LOS) velocity in the direction of the Virgo cluster in Fig. 3.2 as an illustration of the differences between the fields, with the

bands indicating the standard deviation among the 20 realisations for the appropriate fields.

Alias	Constraint Data	Resolution
C15	2M++ redshift-space positions	4 h^{-1} Mpc
L24	2MRS redshift-space positions	3.1 h^{-1} Mpc
CSiBORG1	2M++ redshift-space positions	2.6 h^{-1} Mpc
CSiBORG2	2M++ redshift-space positions	2.6 h^{-1} Mpc
S18	CF2 peculiar velocities	1.9 h^{-1} Mpc
C23	CF4 peculiar velocities	3.9 h^{-1} Mpc

Table 3.1: Summary of local Universe reconstructions investigated in this work. C15: Carrick et al. (2015), linear inverse modelling of the luminosity-weighted galaxy density field. L24: Lilow et al. (2024), machine-learning inverse method trained on `Quijote` simulations. CSiBORG1: Jasche & Lavaux (2019), `BORG` forward model with simulations from Bartlett et al. (2021). CSiBORG2: Stopyra et al. (2024a), `BORG` forward model with an improved gravity model. S18: Sorce & Tempel (2018); Sorce (2020), constrained realisation with Wiener filter and reverse Zel'dovich approximation. C23: Courtois et al. (2023), HMC sampling of the density field at $z = 0$ constrained against peculiar velocities.

3.2.2 Peculiar velocity samples

We use the Tully–Fisher relation (TFR) and supernova (SN) samples to calibrate the flow models, which we now discuss in turn. We do not use the absolute calibration of the distance tracers, so the zero-point of the distance indicators is degenerate with the Hubble constant. However, since we work in distance units of h^{-1} Mpc, our results are independent of H_0 except for its influence on the power spectrum of the reconstructions. We show the samples' observed redshift distribution in the CMB frame in Fig. 3.3. Our method for extracting luminosity distance information from the TFR and SN data is described in Section 3.3.

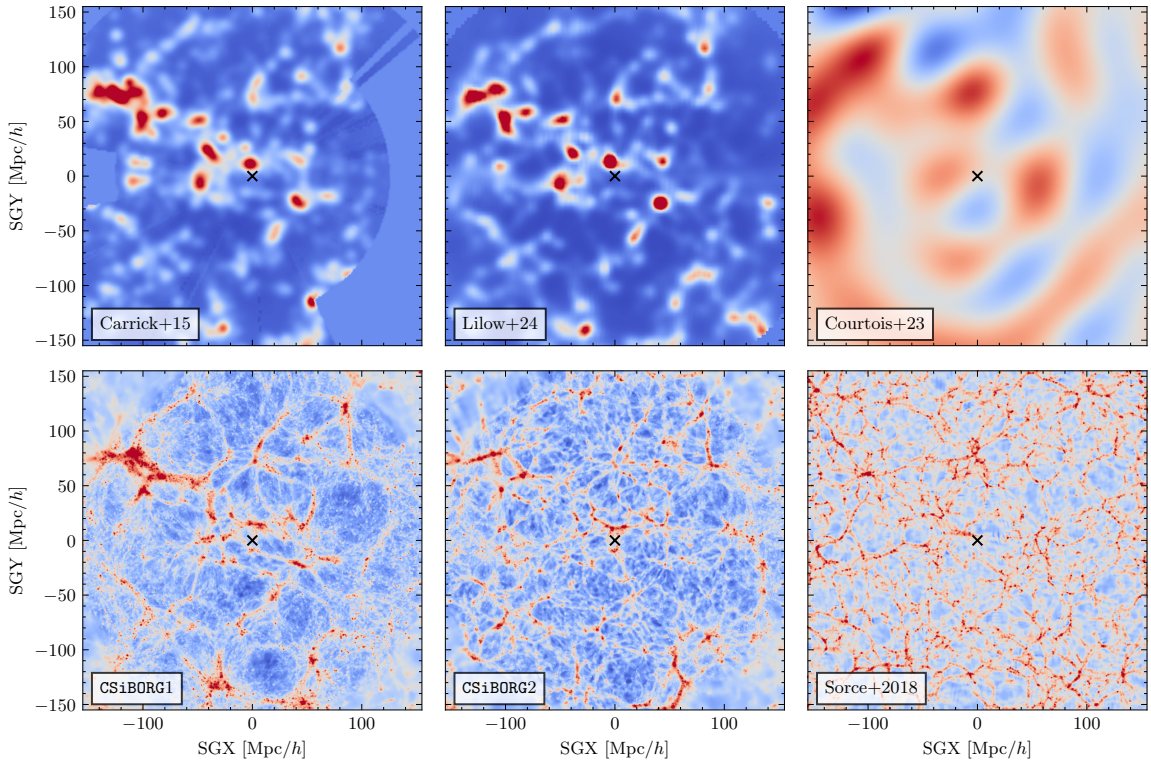


Figure 3.1: Slices of the density field in the SGX-SGY plane from $-155\ h^{-1}$ Mpc to $155\ h^{-1}$ Mpc in supergalactic coordinates at $z = 0$ for the reconstructions used in this work (Table 3.1). Redder (bluer) colours indicate overdensities (underdensities). No additional smoothing is applied. Bottom-row fields are resimulated at higher resolution. The CSiBORG suites and Courtois et al. (2023) are averaged over 20 posterior samples. The black cross marks the Local Group.

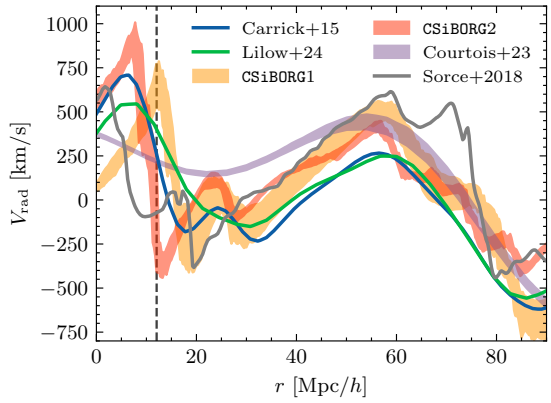


Figure 3.2: Reconstructed velocities along the LOS to the Virgo cluster, whose approximate distance is indicated by the dashed vertical line. Shaded regions show the 16$^{\rm th}$–84$^{\rm th}$ percentile range for reconstructions with multiple posterior samples. For Carrick et al. (2015) we assume $\beta^\star = 0.43$.

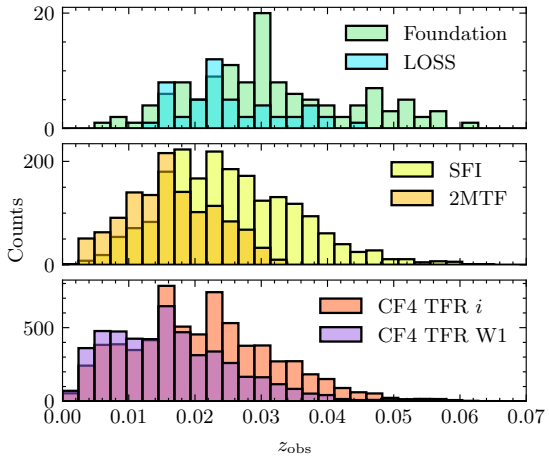


Figure 3.3: Distribution of observed CMB-frame redshifts ($z_{\rm CMB}$) for the peculiar velocity samples used in this work. Most samples lie within $z_{\rm CMB} \lesssim 0.05$; 2MTF is limited to $z_{\rm CMB} \lesssim 0.03$ by its magnitude cut. The samples are separated into three rows for clarity, sharing the x-scale but with independent y-scales.

3.2.2.1 Tully–Fisher samples

The TFR relates the HI linewidth W to the absolute magnitude M of spiral galaxies. We write the relation as

$$M(\eta) = \begin{cases} a_{\rm TFR} + b_{\rm TFR}\eta + c_{\rm TFR}\eta^2 & \text{if } \eta > 0 \\ a_{\rm TFR} + b_{\rm TFR}\eta & \text{otherwise} \end{cases} \tag{3.2}$$

where $a_{\rm TFR}$, $b_{\rm TFR}$, and $c_{\rm TFR}$ are the zero-point, slope, and curvature, respectively, though the latter term is commonly neglected as the relation is fairly linear. We reparameterise the linewidth to introduce a parameter η such that

$$\eta = \log \frac{W}{\mathrm{km\,s^{-1}}} - 2.5. \tag{3.3}$$

Henceforth, we refer to η as the linewidth. We model the curvature of the TFR in high-linewidth galaxies only, which is why the curvature term applies only to galaxies with $\eta > 0$. In the subsequent analysis we jointly infer the TFR calibration parameters along with the intrinsic scatter $\sigma_{\rm TFR}$. We use three TFR samples: 2MTF, SFI++ and CF4, which we now describe in turn.

2MTF: The 2MTF survey is an all-sky sample consisting of 2,062 galaxies with TFR data up to redshift $z_{\rm CMB} \approx 0.03$, with a K-band apparent magnitude limit of 11.25 (Masters et al., 2008; Hong et al., 2019). We utilise the 2MTF dataset compiled by Boruah et al. (2020), which removes duplicates from the SFI++ sample, uses only K-band magnitudes and retains only galaxies with linewidths within the range $-0.1 < \eta < 0.2$. This final sample contains 1,247 galaxies.

SFI++: The SFI++ catalogue is an all-sky sample consisting of 4,052 galaxies and 736 groups up to $z_{\rm CMB} \lesssim 0.05$, without a strict apparent magnitude threshold (Masters et al., 2006; Springob et al., 2007). Here we use the galaxy-only sample curated by Boruah et al. (2020) and adopt I-band magnitudes, though we do not apply their strict linewidth selection because we allow for TFR curvature. The final sample contains 2,010 objects.

CF4: The CF4 TFR survey consists of 9,792 galaxies up to $z_{\rm CMB} \lesssim 0.05$, without a strict apparent magnitude threshold (Kourkchi et al., 2020a,b), and is a subset of the CF4 sample (Tully et al., 2023). We use both the SDSS i-band and Wide-field Infrared Survey Explorer (WISE) W1 magnitudes: while the former is limited to the SDSS footprint, the latter covers the entire sky. We treat the two bands as separate samples, though they are not independent, as some objects have both SDSS and WISE photometry. We retain objects already present in either the 2MTF or SFI++ samples, but select only galaxies with $\eta > -0.3$, galactic latitude $|b| > 7.5°$, and quality flag 5 (best). The final SDSS i-band and WISE W1 samples contain 5,027 and 3,278 galaxies, respectively. When analysing both samples jointly, we preferentially retain the WISE photometry for galaxies with measurements in both bands. The publicly available CF4 catalogue does not list magnitude uncertainties,

thus we set them to a fiducial (conservative) value of 0.05 mag (Kourkchi et al., 2019).

3.2.2.2 Supernova samples

Type Ia SNe are standardised via the SALT2 light-curve model (Guy et al., 2007) and the Tripp formula (Tripp, 1998) (Section 1.5.1),

$$m_{\text{standard}} = m_{\text{obs}} + \mathcal{A}x_1 - \mathcal{B}c, \tag{3.4}$$

where m_{obs} is the observed apparent magnitude, x_1 the stretch, and c the colour. We take the absolute magnitude M_{SN}, $\mathcal{A}$, and $\mathcal{B}$ as global parameters inferred jointly per SN sample. We denote the standardisation coefficients $\mathcal{A}$ and $\mathcal{B}$ (rather than the conventional α and β) to avoid confusion with variables defined later.

We use two SN samples in this work: the Lick Observatory Supernova Search (LOSS; Li et al. 2000) and the Foundation Supernova Survey (Foundation; Foley et al. 2018; Jones et al. 2019). We adopt the curated samples from Boruah et al. (2020), where the LOSS sample is derived by Ganeshalingam et al. (2013). In this curation, SNe already present in LOSS were removed from Foundation. The final LOSS and Foundation samples contain 55 and 125 SNe, respectively. For discussion of our choice to use only these samples, and other samples that could potentially be used, see Section 3.5.5.

3.3 Methodology

In this section, we first explain the relationship between distance and redshift through peculiar velocities (Section 3.3.1). Next, we describe the probabilistic model used to infer the calibration parameters of the distance indicators and the velocity fields (Section 3.3.2). Finally, we outline the method for computing the evidence of each model (Section 3.3.3).

Sample	Type	Number of objects	Description
2MTF	Tully–Fisher	1,247	Full-sky sample curated by Boruah et al. (2020), retaining only linewidth range of $-0.1 < \eta < 0.2$ and with duplicates from the SFI++ sample removed (partially included in the CF4 sample)
SFI++	Tully–Fisher	2,010	Full-sky sample curated by Boruah et al. (2020). No strict linewidth selection applied (partially included in the CF4 sample).
CF4 (SDSS i-band)	Tully–Fisher	5,027	Part of the CF4 survey, limited to the SDSS footprint. We retain only $\eta > -0.3$ and galactic latitude $\lvert b \rvert > 7.5°$.
CF4 (WISE W1-band)	Tully–Fisher	3,278	Part of the CF4 survey with full-sky coverage (a subset of these has i-band photometry as well). Same selection criteria as the CF4 SDSS i-band sample.
Foundation	Type Ia SN	125	Based on Foley et al. (2018); curated by Boruah et al. (2020).
LOSS	Type Ia SN	55	Based on Ganeshalingam et al. (2013); curated by Boruah et al. (2020). Overlap with Foundation sample removed.

Table 3.2: Summary of the Tully–Fisher and SN peculiar velocity samples used to test the velocity field reconstructions.

3.3.1 Distance measures

The observed CMB-frame redshift decomposes into cosmological and peculiar contributions as shown in Section 1.5.3, where $z_{\rm pec} = V_{\rm pec}/c$ and $z_{\rm cosmo}$ maps to comoving distance r via the flat ΛCDM relation

$$r(z_{\rm cosmo}) = \frac{c}{H_0}\int_0^{z_{\rm cosmo}} \frac{{\rm d}z'}{\sqrt{\Omega_{\rm m}(1+z')^3 + 1 - \Omega_{\rm m}}}. \tag{3.5}$$

Within the single-flow approximation, $V_{\rm pec}$ at the position of the source $\boldsymbol{r}$ is

$$V_{\rm pec} = (\beta \boldsymbol{v}(\boldsymbol{r}) + \boldsymbol{V}_{\rm ext}) \cdot \hat{\boldsymbol{r}}, \tag{3.6}$$

where $\hat{\boldsymbol{r}}$ is the unit vector along the source's LOS and $\boldsymbol{v}(\boldsymbol{r})$ is the 3D velocity field evaluated at $\boldsymbol{r}$.

Eq. (3.6) introduces two new parameters for the velocity field, $\boldsymbol{V}_{\rm ext}$ which is the "external" velocity of the box, and β which is a scaling parameter. $\boldsymbol{V}_{\rm ext}$ is a constant vector accounting for the motion of the reconstruction box relative to the rest of the Universe, and is sourced by matter outside of this volume which cannot be picked up by the reconstructions themselves. β describes two effects. First, for the linear theory inverse reconstructions of C15 where velocities are computed from the galaxy density field, β is expected to be f/b where f is the dimensionless growth rate and b is the linear galaxy bias. In this case β contains important cosmological information. For the other reconstructions the velocities are derived from the matter field rather than the galaxy field, and hence this factor is not present and β is expected to be one. However, we keep it there as a measure of how reliably the reconstructions are matching the peculiar velocities implied by the data: if they are systematically over- or under-estimating these velocities, we will find $\beta < 1$ or $\beta > 1$, respectively. A third parameter we introduce for each velocity field is σ_v, a Gaussian uncertainty between the field's prediction and the "observed" value which is assumed to

be position-independent. This describes to first order the effect of small-scale motions that cannot be captured by the velocity reconstruction methods. However, we note that Hollinger & Hudson (2024) showed on C15-like mock reconstructions that the recovered σ_v is inflated for galaxies near the Zone of Avoidance (which is masked in our catalogues). We refer to $\{\boldsymbol{V}_{\rm ext}, \beta, \sigma_v\}$ collectively as "calibration" parameters of the fields. The distance modulus is $\mu = m - M = 5\log\left[(1 + z_{\rm cosmo})\, r/{\rm Mpc}\right] + 25$ (Section 1.5.3).

3.3.2 Joint flow & distance calibration

We simultaneously calibrate both the distance indicator relation and auxiliary parameters associated with the density and velocity field, which we now introduce. We either use the provided density and radial velocity fields (in the case of C15; C23; L24) or construct them ourselves using SPH from particle snapshots (for `CSiBORG1`, `CSiBORG2`, and S18). We evaluate these 3D fields along each tracer's LOS up to a radial distance of $200\,h^{-1}$ Mpc with a uniform spacing of 0.5 h^{-1} Mpc, which is much higher than both the typical reconstruction resolution and distance uncertainty. We use the linear regular grid interpolator `RegularGridInterpolator` implemented in `scipy`[1] (Virtanen et al., 2020).

We jointly infer the density and velocity field calibration ($\alpha_{\rm low}$, $\alpha_{\rm high}$, ρ_t or b_1, and β, σ_v, $\boldsymbol{V}_{\rm ext}$), the distance indicator calibration ($a_{\rm TFR}$, $b_{\rm TFR}$, $c_{\rm TFR}$ for the TFR or $M_{\rm SN}$, $\mathcal{A}$, $\mathcal{B}$ for SNe, and $\sigma_{\rm int}$), along with the Bayesian evidence for each model. We denote these model parameters as $\boldsymbol{\theta}$ and outline our model on the example of the TFR, though the SN calibration is near-identical and we derive it in Section 3.A. For a TFR catalogue, the observables are the galaxy redshift (converted to the CMB frame) $z_{\rm obs}$, apparent magnitude $m_{\rm obs}$, and linewidth $\eta_{\rm obs}$. We also introduce two more latent parameters per object: distance r and the "true" linewidth $\eta_{\rm true}$. In the following, we omit the variables that the probabilities do not explicitly depend on to simplify the notation.

The posterior for a single observation follows the hierarchical framework of Sec-

[1] `https://docs.scipy.org`

tion 1.5.3. For the TFR, the observables are $\{z_{\rm obs}, m_{\rm obs}, \eta_{\rm obs}\}$ and the latent variables are the true distance r and true linewidth $\eta_{\rm true}$. The likelihood factorises into independent redshift, magnitude, and linewidth terms (selection is discussed below):

$$\mathcal{P}(\boldsymbol{\theta}, r, \eta_{\rm true} \mid z_{\rm obs}, m_{\rm obs}, \eta_{\rm obs}) \propto \mathcal{L}(z_{\rm obs} \mid \boldsymbol{\theta},\, r)\, \mathcal{L}(m_{\rm obs} \mid \boldsymbol{\theta},\, r,\, \eta_{\rm true})\, \mathcal{L}(\eta_{\rm obs} \mid \eta_{\rm true})\, \pi(r, \eta_{\rm true}, \boldsymbol{\theta}). \tag{3.7}$$

Implicitly, the likelihood also depends on the underlying density and velocity field, over which we later marginalise in Eq. (3.24). Convolving the measurement uncertainty $\sigma_{cz_{\rm obs}}$ with the residual velocity scatter σ_v gives the redshift likelihood

$$\mathcal{L}(z_{\rm obs} \mid \boldsymbol{\theta},\, r) = \mathcal{N}\Big(cz_{\rm obs};\, cz_{\rm pred},\, \sigma_v^2 + \sigma_{cz_{\rm obs}}^2\Big), \tag{3.8}$$

where $z_{\rm pred}(r,\, \boldsymbol{\theta})$ is the predicted redshift from Eqs. (3.5) to (3.6). The magnitude likelihood compares the observed apparent magnitude to $m_{\rm pred} = M(\eta_{\rm true}) + \mu(r)$, with $M(\eta_{\rm true})$ from the TFR, Eq. (3.2),

$$\mathcal{L}(m_{\rm obs} \mid \boldsymbol{\theta},\, r,\, \eta_{\rm true}) = \mathcal{N}\Big(m_{\rm obs};\, m_{\rm pred},\, \sigma_{\rm int}^2 + \sigma_m^2\Big), \tag{3.9}$$

where $\sigma_{\rm int}$ is the intrinsic scatter and σ_m the measurement error. The third term is the likelihood of the observed linewidth,

$$\mathcal{L}(\eta_{\rm obs} \mid \eta_{\rm true}) = \frac{\mathcal{N}(\eta_{\rm obs};\, \eta_{\rm true},\, \sigma_\eta^2)}{p(S = 1 \mid \hat{\eta},\, w_\eta)}, \tag{3.10}$$

where σ_η is the linewidth measurement error and $p(S = 1 \mid \hat{\eta},\, w_\eta)$ accounts for any truncation in $\eta_{\rm obs}$ and is derived below in Eq. (3.19).

Next we consider the prior, which factorises as

$$\pi(r,\, \eta_{\rm true},\, \boldsymbol{\theta}) = \pi(r \mid \boldsymbol{\theta})\, \pi(\eta_{\rm true} \mid \boldsymbol{\theta})\, \pi(\boldsymbol{\theta}). \tag{3.11}$$

The first term is the distance prior, which accounts for both the homogeneous and inhomogeneous Malmquist biases (Section 1.5.3). We model it as

$$\pi(r \mid \boldsymbol{\theta}) = \frac{n(r, \boldsymbol{\theta})\, f(r, \boldsymbol{\theta})}{\int \mathrm{d}r'\, n(r', \boldsymbol{\theta})\, f(r', \boldsymbol{\theta})}, \tag{3.12}$$

where $n(r, \boldsymbol{\theta})$ denotes the number density, $f(r, \boldsymbol{\theta})$ is defined following Lavaux (2016) as

$$f(r, \boldsymbol{\theta}) = r^p \exp\left[-\left(\frac{r}{R}\right)^q\right], \tag{3.13}$$

with p, q, and R being free model parameters. In particular, it is expected that $p \approx 2$ to reproduce the homogeneous Malmquist bias, the exponential cut-off provides an effective description of selection effects, R sets the scale at which incompleteness becomes significant, and q controls the sharpness of this transition. We emphasise that this is a purely phenomenological approach to modelling selection. A rigorous treatment would require explicitly forward-modelling the survey selection, as outlined by Kelly et al. (2008), but this in turn requires a detailed knowledge of the selection function. Since the TFR samples have a complex selection determined by a combination of HI flux and optical magnitude cuts, we instead adopt this phenomenological model. In practice, this distance prior effectively accounts for distance-related selection effects such as cuts in observed redshift or apparent magnitude.

The choice of source number density modelling depends on the reconstruction method. For fields based on linear theory (C15, C23), we adopt a linear model,

$$n(r, b_1) = 1 + b_1 \delta(r), \tag{3.14}$$

where b_1 is a free parameter and $\delta(r)$ is the density contrast at distance r along the LOS. Since a linear-theory $\delta(r)$ can drop below $-1/b_1$, where Eq. (3.14) returns an unphysical

negative number density, we pass it through a smooth, strictly positive clip,

$$n(r,\, b_1) \to \tfrac{1}{2}\left[\, n + \sqrt{n^2 + \tau^2}\,\right], \quad \tau = 0.1, \tag{3.15}$$

which recovers $n \approx 1 + b_1\delta$ at typical densities and drives n smoothly to zero where the linear model turns negative. This regularisation primarily affects cells where the linear bias model is already unphysical, perturbing the number density only negligibly elsewhere, and varying τ within reasonable bounds makes no appreciable difference to the inferred parameters or evidence ratios. The direct distance tracer samples have different characteristics, which is why we treat b_1 as a free parameter. Similarly for C23 we sample b_1 under a uniform prior. For fields derived from N-body simulations (CSiBORG1, CSiBORG2, S18, and L24), we instead use a double-power law model

$$n(r,\, \rho_t,\, \alpha_{\rm low},\, \alpha_{\rm high}) = \left(\frac{\rho}{\rho_t}\right)^{\alpha_{\rm low}} \left(1 + \frac{\rho}{\rho_t}\right)^{\alpha_{\rm high} - \alpha_{\rm low}}, \tag{3.16}$$

where $\rho \equiv \rho(r)$ is the density at distance r along the LOS to the host, $\alpha_{\rm low}$ and $\alpha_{\rm high}$ are the two slopes, and ρ_t is the transition density, all of which we infer jointly.

Next, in Eq. (3.11) we specify the hyperprior on $\eta_{\rm true}$, for which we adopt a Gaussian distribution,

$$\pi(\eta_{\rm true} \mid \boldsymbol{\theta}) = \mathcal{N}\left(\eta_{\rm true};\, \hat{\eta},\, w_\eta^2\right), \tag{3.17}$$

where $\hat{\eta}$ and w_η denote its mean and width, respectively, and are jointly inferred. Assigning a Gaussian hyperprior to the true (latent) parameter, distinct from the observed quantity, follows the Marginalised Normal Regression framework (Bartlett & Desmond, 2023), which has been shown to yield unbiased regression, and is consistent with approaches used in SN cosmology (March et al., 2011, 2014; Rubin et al., 2015; March et al., 2018; Rubin et al., 2025). For the remaining parameters, we adopt a reference (scale-invariant) prior $\pi(\sigma_v) \propto 1/\sigma_v$, applied also to $\sigma_{\rm int}$, uniform priors on the TFR parameters ($a_{\rm TFR}$, $b_{\rm TFR}$,

$c_{\rm TFR}$), a reference prior on w_η, and a uniform prior on $\hat{\eta}$. The external velocity $\boldsymbol{V}_{\rm ext}$ is taken to be uniform in both magnitude and direction.

For n independent observations, the selection-corrected posterior follows from the framework derived in Section 1.5.4 (see also Kelly et al. 2008). We further account for truncation in the observed linewidth. The posterior is

$$\mathcal{P}(\boldsymbol{\theta} \mid \{\boldsymbol{d}\}) \propto \frac{\pi(\boldsymbol{\theta}) \prod_{i=1}^{n} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\theta})}{[p(S=1 \mid \boldsymbol{\theta})]^n}, \tag{3.18}$$

where $\pi(\boldsymbol{\theta})$ denotes the prior on the model parameters, $\boldsymbol{d}_i$ are the data for the $i^{\rm th}$ event, and $\mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\theta})$ is the likelihood of $\boldsymbol{d}_i$ given parameters $\boldsymbol{\theta}$. The factor $p(S=1 \mid \boldsymbol{\theta})$ is the fraction of sources that are observed given some total population, and accounts for such "missing" data. In the case of the TFR model above and truncation in $\eta_{\rm obs}$, this probability is given by

$$p(S=1 \mid \hat{\eta},\, w_\eta) = \iint \mathrm{d}\eta_{\rm obs}\, \mathrm{d}\eta_{\rm true}\; p(S=1 \mid \eta_{\rm obs}) \mathcal{L}(\eta_{\rm obs} \mid \eta_{\rm true})\, \pi(\eta_{\rm true} \mid \hat{\eta},\, w_\eta), \tag{3.19}$$

where $p(S=1 \mid \eta_{\rm obs})$ is a binary truncation indicator between $\eta_{\rm min}$ and $\eta_{\rm max}$,

$$p(S=1 \mid \eta_{\rm obs}) = \begin{cases} 1 & \text{if} \quad \eta_{\rm min} < \eta_{\rm obs} < \eta_{\rm max}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.20}$$

$\mathcal{L}(\eta_{\rm obs} \mid \eta_{\rm true})$ denotes the Gaussian likelihood of the observed given the true linewidth, and $\pi(\eta_{\rm true} \mid \hat{\eta},\, w_\eta)$ is the Gaussian prior on the true linewidth. Given these assumptions, it can be shown that

$$p(S=1 \mid \hat{\eta},\, w_\eta) = \Phi\left(\frac{\eta_{\rm max} - \hat{\eta}}{\sqrt{\sigma_\eta^2 + w_\eta^2}}\right) - \Phi\left(\frac{\eta_{\rm min} - \hat{\eta}}{\sqrt{\sigma_\eta^2 + w_\eta^2}}\right), \tag{3.21}$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution,

defined as

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t. \tag{3.22}$$

For this reason, we already included this $p(S = 1 \mid \hat{\eta}, w_\eta)$ term in Eq. (3.10). In principle, the same approach should be applied to model selection in optical and HI flux. For instance, the latter would require inferring the HI flux from the HI galaxy mass. However, to avoid this complexity, we instead follow the approach outlined above: selection effects in optical magnitude or HI flux are effectively absorbed into the phenomenological distance prior, while truncation in the linewidth distribution is treated separately. This does not constitute double-counting, as the distance-prior parameterisation is inferred and would therefore favour no cutoff if all selection effects are already accounted for. Furthermore, linewidth selection has a negligible impact on the results presented here.

Combining the above terms and numerically marginalising over r and $\eta_{\rm true}$ at each Markov chain Monte Carlo (MCMC) step yields the full posterior

$$\begin{aligned} \mathcal{P}(\boldsymbol{\theta} \mid z_{\rm obs}, m_{\rm obs}, \eta_{\rm obs}) \propto & \iint \mathrm{d}r\, \mathrm{d}\eta_{\rm true}\, \mathcal{N}\left(cz_{\rm obs};\, cz_{\rm pred},\, \sigma_v^2 + \sigma_{cz_{\rm obs}}^2\right) \\ & \times \mathcal{N}\left(m_{\rm obs};\, m_{\rm pred},\, \sigma_{\rm int}^2 + \sigma_m^2\right) \frac{\mathcal{N}(\eta_{\rm obs};\, \eta_{\rm true},\, \sigma_\eta^2)}{p(S = 1 \mid \hat{\eta}, w_\eta)} \\ & \times \frac{n(r, \boldsymbol{\theta})\, f(r, \boldsymbol{\theta})}{\int \mathrm{d}r'\, n(r', \boldsymbol{\theta})\, f(r', \boldsymbol{\theta})}\, \mathcal{N}\left(\eta_{\rm true};\, \hat{\eta},\, w_\eta^2\right) \pi(\boldsymbol{\theta}). \end{aligned} \tag{3.23}$$

For the radial distance grid, we adopt uniform spacing in comoving distance with step size 0.5 h^{-1} Mpc, converting to cosmological redshift or distance modulus using `Astropy`[2] (Astropy Collaboration et al., 2013, 2018, 2022). For the $\eta_{\rm true}$ grid, we use 51 uniformly spaced points spanning a range centred on $\eta_{\rm obs}$, with the width defined as ten times σ_η. We explicitly marginalise over both r and $\eta_{\rm true}$, rather than sampling them with HMC, so that the resulting posterior remains relatively low-dimensional and amenable to evidence calculation.

For reconstructions providing multiple realisations, we marginalise over the field by

[2] https://www.astropy.org

averaging the posterior over available samples,

$$\mathcal{P}(\boldsymbol{\theta} \mid z_{\rm obs}, m_{\rm obs}, \eta_{\rm obs}) \approx \frac{1}{N} \sum_{i=1}^{N} \mathcal{P}(\boldsymbol{\theta} \mid z_{\rm obs}, m_{\rm obs}, \eta_{\rm obs}, \boldsymbol{v}_i, \rho_i), \tag{3.24}$$

where ρ_i and $\boldsymbol{v}_i$ represent the $i^{\rm th}$ density and velocity field realisation from the reconstruction posterior (e.g. `BORG`). In the case of a SN calibration, the three independent variables (m, x_1, c) are treated analogously to the TFR observables, except we marginalise numerically only over r and explicitly sample the true x_1 and c to avoid computing a 3D numerical integral.[3]

We sample the posterior using the NUTS variant of HMC (Section 1.3.4) as implemented in `NumPyro`[4] (Phan et al., 2019), running chains until $\hat{R} - 1 \leq 10^{-3}$ (Gelman & Rubin, 1992). All model parameters, along with their priors, are summarised in Table 3.3.

3.3.3 Evidence calculation

We compare velocity field reconstructions using the Bayesian model evidence $\mathcal{Z}$ and the Bayes factor,

$$\mathcal{Z} = \int \mathrm{d}\boldsymbol{\theta}\, \mathcal{L}(D \mid \boldsymbol{\theta})\, \pi(\boldsymbol{\theta}). \tag{3.25}$$

The TFR model introduced in Section 3.3.2 includes, in addition to the model parameters, two (three) latent variables per galaxy in the case of a TFR (SN) catalogue. This increases the dimensionality of the parameter space to $O(1000)$, making evidence calculation infeasible using standard methods such as nested sampling (Skilling, 2006). Recently, machine learning-based methods for evidence calculation, such as `harmonic` (McEwen et al., 2021; Polanska et al., 2025) and *floZ* (Srinivasan et al., 2024), have been developed. While these methods have shown success in $O(100)$-dimensional parameter spaces, they have not

[3] Though not realised at the time of writing this work, since x_1 and c enter linearly through the Tripp formula, it is trivially possible to analytically marginalise over these parameters, unlike the TFR, which has a quadratic dependence on $\eta_{\rm true}$ through the curvature term.

[4] https://num.pyro.ai/en/latest/

Parameter	Description	Prior
Galaxy bias and flow model parameters		
$\boldsymbol{V}_{\rm ext}$	External velocity vector	Uniform in both magnitude and direction
σ_v	Scatter between the observed and predicted redshifts	$\pi(\sigma_v) \propto 1/\sigma_v$
β or $\beta^\star$	Velocity field scaling parameter	Uniform
b_1	Linear galaxy bias parameter (C15, C23)	$\pi(b_1) = \mathcal{N}(\beta^\star/f, 0.04)$ for C15, uniform for C23
$\alpha_{\rm low}$, $\alpha_{\rm high}$, $\ln \rho_t$	Double power-law galaxy bias parameters (CSiBORG1, CSiBORG2, S18, L24)	$\pi(\alpha_{\rm low}) = \mathcal{N}(1,1)$, $\pi(\alpha_{\rm high}) = \mathcal{N}(0.5,1)$ (both truncated below at zero) and $\pi(\ln \rho_t) = \mathcal{N}(0,2)$
Distance calibration parameters		
$a_{\rm TFR}$, $b_{\rm TFR}$, $c_{\rm TFR}$	TFR coefficients (zero-point, slope, and curvature)	Uniform
$M_{\rm SN}$, $\mathcal{A}$, $\mathcal{B}$	Supernova coefficients (absolute magnitude, stretch, and colour)	Uniform
$\sigma_{\rm int}$	Magnitude intrinsic scatter	$\pi(\sigma_{\rm int}) \propto 1/\sigma_{\rm int}$
R, q, p	Phenomenological distance prior	$\pi(R) = \mathcal{U}(0,100)$ Mpc, $\pi(p) = \mathcal{N}(2,0.1)$ and $\pi(q) = \mathcal{N}(1,30)$
Latent parameter hyperpriors		
$\eta_{\rm true}$	True TFR linewidth	$\mathcal{N}(\eta_{\rm true} \mid \hat{\eta}, w_\eta^2)$
$\hat{\eta}$	Mean of the Gaussian hyperprior on true linewidth	Uniform
w_η	Width of the Gaussian hyperprior on true linewidth	$\pi(w_\eta) \propto 1/w_\eta$
$x_{1,\rm true}$, $c_{\rm true}$	True SN stretch and colour parameters	Gaussian hyperpriors with free mean and width
$\boldsymbol{\mu}_{\rm SN}$	Mean of the Gaussian hyperprior on the SN stretch, and colour parameters	Uniform
$\mathbf{C}_{\rm SN}$	Covariance matrix of the Gaussian hyperprior on the SN stretch, and colour parameters	$\pi(\sigma_i) \propto 1/\sigma_i$ for standard deviations σ_i and uniform in the correlation coefficient

Table 3.3: Summary of the free parameters of our model and their priors.

yet been shown capable of handling the dimensionality that would be required for a full calculation here.

To circumvent this problem, in the TFR model we explicitly marginalise over both latent parameters (r and $\eta_{\rm true}$) at each MCMC step, thereby removing them from the posterior and yielding a low-dimensional distribution for which the evidence can be estimated. In contrast, the presence of a third latent parameter in the SN model makes explicit numerical marginalisation impractical.[5] Thus, for the SN model only we consider an alternative model in which we fix the true values of x_1 and c to their observed values, effectively imposing a Dirac distribution prior on them. The observational uncertainties are then propagated not by marginalising over the true values but instead through first-order error propagation, summing them in quadrature with $\sigma_{\rm int}$ (similar to the method of Boruah et al. 2020) as

$$\sigma_{\rm int} \to \sqrt{\sigma_{\rm int}^2 + \mathcal{A}^2\sigma_{x_1}^2 + \mathcal{B}^2\sigma_c^2}, \tag{3.26}$$

where σ_{x_1} and σ_c are the measurement uncertainties of the SALT2 SN light curve parameters. This approach reduces the dimensionality of the model to under twenty, and incurs only a small systematic error if the intrinsic scatters of the scaling relations dominate over the measurement uncertainties. In Section 3.A we use the Foundation sample and compare explicitly the posteriors that the two methods generate. The only parameter that is biased as a result of this assumption is $\mathcal{B}$, whereas all others remain practically unchanged. We emphasise however that this SN model approximation is only used for calculating evidences; all other results with the LOSS and Foundation samples use the full model where $x_{1,\rm true}$ and $c_{\rm true}$ are explicitly sampled (described in Section 3.A).

Now that the dimensionality of the parameter space is sufficiently low, we may apply the `harmonic`[6] package, which applies normalising flows and the harmonic estimator to calculate the model evidences directly from the HMC posterior samples (McEwen et al.,

[5] As noted above, x_1 and c can in fact be analytically marginalised. This was realised only after the original analysis.

[6] https://github.com/astro-informatics/harmonic

2021; Polanska et al., 2025). As can be seen from Eq. (3.25), the evidence depends on the chosen prior distributions (including the widths of uniform priors). However, in most cases we always compare models with the same parameters, in which case the widths of uniform priors are unimportant. The exception is the model of C15, which necessitates the additional velocity field scaling parameter $\beta^\star$. Expanding the uniform prior on β beyond the range permitted by the likelihood reduces the evidence. Given that the width of the posterior on β ranges from $\sim$0.5 in the worst case to 0.05 in the best, tightening the prior to encapsulate only the posterior would improve the evidence by up to $20 - 200\times$, or $1.3 - 2.3$ on a base-10 logarithm scale. This is not significant given the differences in evidence between reconstructions.

3.4 Results

We use the following methods to assess the relative goodness of the local Universe reconstructions summarised in Table 3.1:

1. Bayesian evidence comparison between the different reconstructions for each peculiar velocity sample (Section 3.4.1).

2. The magnitude of the inferred σ_v (Section 3.4.2).

3. Difference between the inferred β (see Eq. (3.6)) and unity in cases where the reconstructions are derived from the matter field (Section 3.4.3).

4. Comparison of the inferred $\boldsymbol{V}_{\rm ext}$ and the agreement of the reconstructed velocity at $r \to 0$ with the Local Group velocity in the CMB frame (Section 3.4.4 and Section 3.4.5).

The interpretation of these metrics is as follows. The evidence gives the relative probability of the data under the model, which as described above is the main goodness-of-fit statistic of Bayesian inference. Criterion 2 describes the magnitude of residuals

between the observed and predicted redshifts, which measures the accuracy of the model in the posterior region of parameter space. We will also show it to correlate well with the evidence. Criterion 3 also indicates whether the predicted velocities are in broad agreement with the observations: if $\beta < 1$ then the velocities are over-predicted on average while if $\beta > 1$ they are underpredicted, a systematic bias in either case. Lastly, criterion 4 measures the velocity at which the data requires the reconstruction box to move relative to the CMB. A smaller $\boldsymbol{V}_{\rm ext}$ is preferred, especially when the reconstruction volume is large, as this implies that the velocities within the box are sufficient to account for the data. However, $\boldsymbol{V}_{\rm ext}$ depends on both the reconstruction and the peculiar velocity sample used to constrain it, and thus it can only serve as a qualitative comparison. The dependence on the reconstruction volume arises partly from the fact that if the velocities within the box are sourced solely by internal matter, their average will be zero when the entire box is enclosed. Consequently, the average enclosed velocity in the CMB frame (the bulk flow) at such radii is determined exclusively by $\boldsymbol{V}_{\rm ext}$ which we assume to be constant. The bulk flow magnitude follows a Maxwell-Boltzmann distribution in the ΛCDM model (see e.g. Appendix B of Davis et al. 2011). In Section 3.4.5 we also present and compare the radial dependence of the predicted bulk flows and in Section 3.4.6 we report constraints on the growth-rate of large-scale structure using C15 and the CF4 TFR samples.

In addition to using the individual peculiar velocity samples outlined in Table 3.2, we also construct a single joint dataset by combining the LOSS, Foundation, and both CF4 TFR samples. For the latter, we prioritise SDSS photometry when available. We do not include 2MTF or SFI++ in this joint catalogue as these were used in constructing CF4; all objects in our joint dataset therefore contain just one measurement per galaxy. We select only objects with $z_{\rm CMB} < 0.05$ that are within the reconstructed volume.

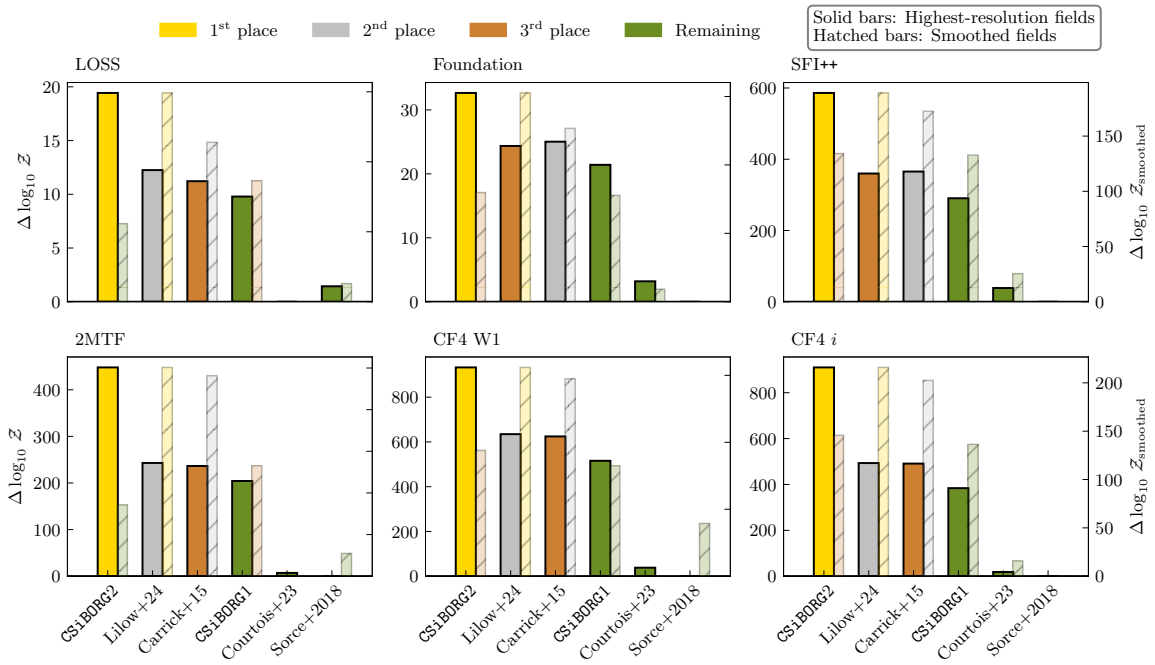


Figure 3.4: Differences in logarithmic evidences $\mathcal{Z}$ from our flow validation model for various local Universe reconstructions (shown on the x-axis, see Table 3.1), compared against peculiar velocity samples (individual panels, see Table 3.2). *Higher bars indicate a preferred model,* and a bar of zero height indicates the reference (least successful) model. The logarithmic evidences are normalised with respect to the reference model as only relative differences are meaningful. Solid bars show evidences using the highest available resolution for each model, while hatched bars show evidences when all fields are smoothed to the resolution of 7.8 h^{-1} Mpc, twice that of Courtois et al. (2023). Overall, CSiBORG2 is the preferred model while the CosmicFlows-based reconstructions (Sorce & Tempel 2018 and Courtois et al. 2023) are disfavoured. Upon smoothing, the reconstruction of Lilow et al. (2024) becomes marginally preferred.

3.4.1 Evidence comparison

In Fig. 3.4, we show the Bayesian evidences to compare the relative goodness-of-fit of reconstructions against the peculiar velocity samples. Since the absolute value of the evidence of a model is meaningless, we only show the differences of the logarithmic evidences $\Delta \log_{10} \mathcal{Z}$, relative to the model with the lowest evidence. The TFR samples have greater discriminating power (larger differences in $\log_{10} \mathcal{Z}$) compared to SN samples, since despite their smaller intrinsic scatter, the SN samples that we use contain significantly fewer objects. For all models except that of C15 we fix $\beta = 1$ for this calculation. We find `CSiBORG2` to be the best model across all peculiar velocity samples, with L24 typically ranking second and C15 third with only a small difference between these latter two. Relative to these, the two reconstructions based on CosmicFlows, S18 and C23, are strongly disfavoured. When testing the joint SN and TFR dataset, we find qualitatively similar results.

A possible driver of the evidence ratios is the varying resolutions of the reconstruction. To isolate this effect, we consider smoothing all fields to a resolution of 7.8 h^{-1} Mpc which is twice the voxel resolution of C23. This is achieved by applying a Gaussian kernel with a standard deviation of $\sqrt{(7.8\ h^{-1}\,\mathrm{Mpc})^2 - \sigma_0^2}$ to the densities and velocities, where σ_0 is the original resolution (voxel size) of the field. However, a Gaussian smoothing kernel with a standard deviation σ_0 is not equivalent to a field with voxel resolution σ_0, thus this "matching" of resolutions is only approximate. We then recalculate the evidences for the smoothed fields, showing the results as hatched bars in Fig. 3.4. This reduces the differences in $\log_{10} \mathcal{Z}$ between the reconstructions, and also causes L24 to become the preferred model although both `CSiBORG` fields continue to perform well. The preference for L24 upon smoothing might arise from accurately capturing large-scale flows, as indicated by the relatively small magnitude of $\boldsymbol{V}_{\mathrm{ext}}$ discussed in Section 3.4.4. Even after smoothing, the CosmicFlows-based reconstructions remain strongly disfavoured. This indicates that the finer-grained information in the other reconstructions contains accurate information, and

that the redshift-space positions of 2M++ galaxies afford better constraints than the peculiar velocities of CosmicFlows.

In Fig. 3.5 we investigate the effect of smoothing on C15 using the joint SNe and TFR dataset. We plot both $\Delta \log_{10} \mathcal{Z}$ and σ_v as functions of smoothing scale $R_{\rm smooth}$. As expected, the goodness-of-fit degrades with increasing smoothing scale. Notably, σ_v increases proportionally, suggesting it is a good indicator of goodness-of-fit. However, σ_v should not be compared between different peculiar velocity samples, as it depends on sample properties such as the relative prevalence of field and cluster galaxies. We infer from these results that the field of C15 would need to be smoothed with a scale of $R_{\rm smooth} \approx 22\ h^{-1}$ Mpc to match the evidence of C23, since we find a $\Delta \log_{10} \mathcal{Z}$ of 890 between C15 and C23.

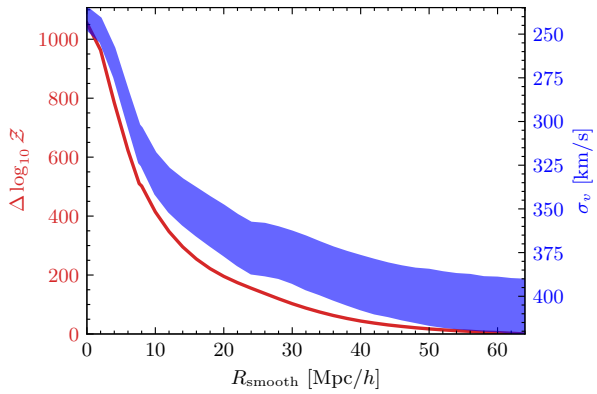


Figure 3.5: Differences in logarithmic evidence ($\Delta \log_{10} \mathcal{Z}$, higher is better) and inferred σ_v as functions of smoothing scale $R_{\rm smooth}$ for the joint SN and TFR dataset against Carrick et al. (2015). The σ_v axis is inverted. Both the evidence and σ_v degrade rapidly with increased smoothing, confirming that σ_v is an effective goodness-of-fit indicator.

3.4.2 Magnitude of redshift residuals

We now turn to σ_v, which measures the residuals between the observed and predicted redshifts. This contains information on how well the reconstruction reproduces the "observed" velocity field of the local Universe. σ_v depends on whether galaxy non-linear velocities were suppressed by grouping. While it is preferable for σ_v to be lower, there must be a lower floor (per sample) describing the effect of small-scale motions that the reconstructed fields cannot hope to capture. This floor would be raised by any unaccounted-for systematic effects or other forms of model misspecification.

In Fig. 3.6 we present the constraints on σ_v using each of the velocity fields. Generally,

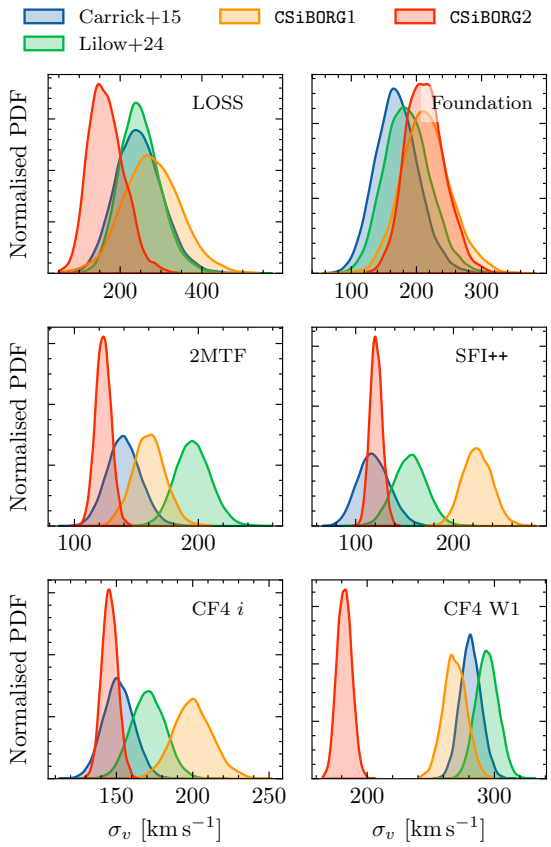


Figure 3.6: Posteriors on σ_v per reconstruction and dataset. Reconstructions with lower σ_v tend to be preferred in evidence (Fig. 3.4). CSiBORG2 and C15 yield the lowest σ_v on the CF4 datasets. Values for C23 and S18 are omitted ($\sigma_v \approx 500$ km s^{-1}).

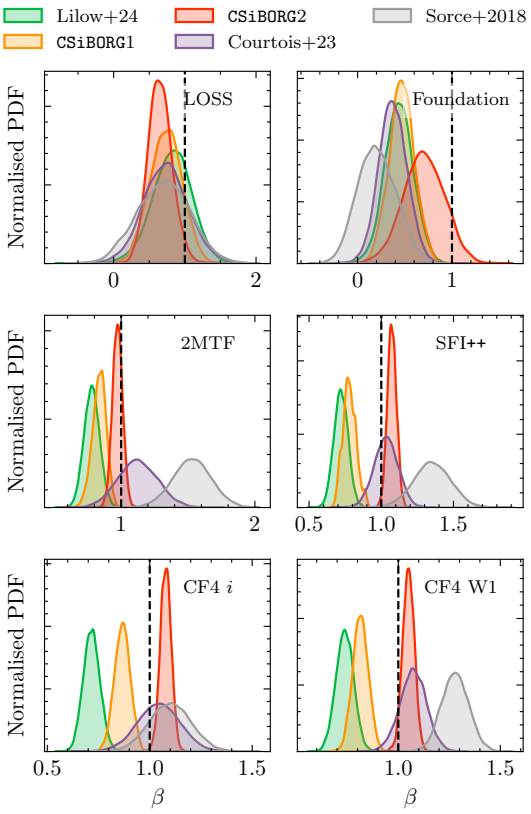


Figure 3.7: Inferred β, Eq. (3.6), with a uniform prior, for models deriving velocities from the matter field. The fiducial value $\beta = 1$ corresponds to no rescaling. Both L24 and CSiBORG1 systematically overpredict velocities ($\beta < 1$), whereas CSiBORG2 is consistent with unity except for CF4 i-band ($< 3\sigma$).

reconstructions with lower σ_v values also achieve higher evidence scores, while those with higher σ_v are less favoured. However, the two SN samples alone lack sufficient data to provide meaningful constraints on σ_v. For the TFR samples, it is notable that C15 yields marginally lower σ_v values compared to L24, despite the evidence comparison's weak preference for L24. This, together with the fact that the evidence comparison does not vastly prefer L24 over C15, appears to indicate that the "augmentation" to C15 provided by learning is not particularly successful. The CSiBORG2 velocities clearly produce the highest evidences and often require the lowest σ_v.

In contrast, reconstructions based on CosmicFlows data require the largest σ_v, even for the CF4 samples. Owing to their large values relative to the other models, we do not show them explicitly in Fig. 3.6. However, in the CF4 W1 sample we find $\sigma_v = 387 \pm 17$ km s^{-1} for C23 and 395 ± 15 km s^{-1} for S18. For the CF4 i-band sample and for 2MTF, σ_v is around 500 km s^{-1} in both cases. This outcome is surprising as these reconstructions are directly constrained by CosmicFlows data, of which the TFR samples are a subset.

3.4.3 Consistency of the velocity scaling

Next we examine β, which scales the velocity field before comparison to the data; see Eq. (3.6). Here we treat β as a free parameter with a uniform prior, with the expectation that $\beta = 1$ for reconstructions where the velocities are sourced by the matter field; a field is better behaved to the extent that this is preferred *a posteriori*.

The constraints on β are shown in Fig. 3.7. We find that the model preferred by the Bayesian evidence, CSiBORG2, somewhat underpredicts velocities only in the SFI++ and CF4 i-band samples, where $\beta = 1.13 \pm 0.03$ and 1.14 ± 0.02. However, even in these cases the deviations are 2.3 and 3σ, respectively, while all other samples are consistent with $\beta = 1$ to within 1.5σ. (We note that in our previous evidence comparison we had fixed $\beta = 1$.) In contrast, CSiBORG1 yields β values inconsistent with unity at more than 3σ in all cases except for the LOSS catalogue. Similarly, L24 consistently produces $\beta < 1$ on all TFR

samples, suggesting its reconstructed velocities are too high or otherwise systematically biased. On the other hand, S18 has β substantially higher ($> 2.5\sigma$) on all TFR samples except CF4 i-band. C23 is consistent with unity in all cases, except the Foundation sample. We leave a detailed discussion of the velocity scaling in C15 to Section 3.4.6, as in C15 it is directly related to the growth rate of large-scale structure.

3.4.4 Magnitude of the external velocity

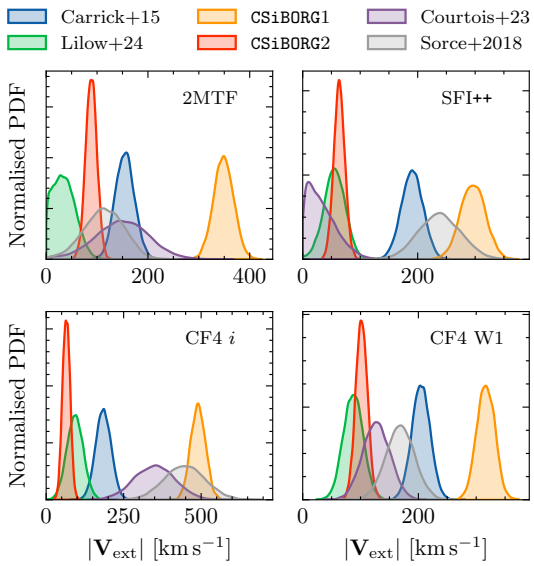


Figure 3.8: Inferred magnitude of $\boldsymbol{V}_{\rm ext}$; see Eq. (3.6). A higher magnitude indicates that the reconstruction box is moving faster relative to the rest of the Universe.

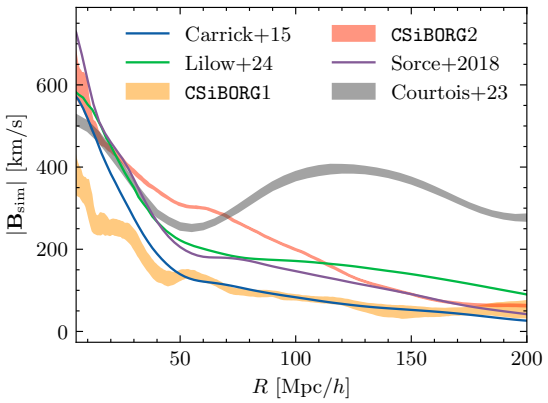


Figure 3.9: Bulk flow magnitude in the simulation rest frame (without $\boldsymbol{V}_{\rm ext}$) as a function of radius R. All reconstructions except Courtois et al. (2023) show the decreasing trend expected in ΛCDM. Bands indicate the 16$^{\rm th}$–84$^{\rm th}$ percentile range for reconstructions with posterior samples.

Next we study $|\boldsymbol{V}_{\rm ext}|$, the speed of the entire reconstruction box in the CMB frame sourced by matter outside it. This depends on both the reconstruction method and the peculiar velocity survey, which sets the effective probed volume.

The posteriors on $|\boldsymbol{V}_{\rm ext}|$ are shown in Fig. 3.8. We observe that although C15 and L24 yield similar goodness-of-fit values, the latter requires a significantly smaller external flow despite both reconstructions being constrained by (almost) the same data. The external flow magnitude in L24 is nearly consistent with zero for the 2MTF, SFI++, and CF4 TFR i-band

samples, with values of 35 ± 21, 55 ± 19, and 93 ± 26 km s^{-1}, respectively, since their reconstruction already includes $\boldsymbol{V}_{\rm ext}$-like contribution by construction. In contrast, C15 requires larger magnitudes of 154 ± 19, 190 ± 18, and 184 ± 24 km s^{-1} for those samples.

Among the other reconstructions, CSiBORG2 generally requires the smallest external flow magnitude (100 km s^{-1}), whereas CSiBORG1 typically requires a larger speed of approximately 400 km s^{-1}. For the CosmicFlows-based reconstructions, we find that their preferred flow magnitudes disagree significantly only in the case of SFI++, with S18 requiring a notably higher magnitude. Although in other cases they agree well, the magnitude required by S18 is systematically higher, potentially due to the shallower depth of the CosmicFlows-2 data. However, for both CF4 samples, S18 and C23 require a higher flow magnitude than CSiBORG2 (or L24).

3.4.5 Bulk flow curve and Local Group velocity

Finally, we calculate the bulk flows produced by the various reconstructions. The velocities in each reconstruction are generated by the internal matter distributions and thus are in the "simulation" rest frame, where the average velocity of the box is zero by definition. The exception to this is L24 whose neural network predicts both the density and velocity field, without explicitly computing the latter from the former and, thus, can learn "super-survey" scales encoded in the redshift-space distortions. This means that the L24 velocities already have $\boldsymbol{V}_{\rm ext}$ encoded in them, and hence the fiducial value for our free parameter would be zero. We denote the simulation bulk flow as $\boldsymbol{B}_{\rm sim}$ and define it according to the standard bulk flow definition (Watkins et al., 2009; Feldman et al., 2010; Nusser & Davis, 2011; Hoffman et al., 2015; Watkins & Feldman, 2015; Nusser, 2016; Scrimgeour et al., 2016; Feix et al., 2017; Hellwing et al., 2018; Peery et al., 2018; Whitford et al., 2023; Watkins et al., 2023)

$$\boldsymbol{B}_{\rm sim}(R) = \frac{\beta}{V} \int_V \boldsymbol{v}(\boldsymbol{r}) \mathrm{d}V, \tag{3.27}$$

where V is the volume enclosed within a sphere of radius R around the observer at the centre of each box. To calculate this we again set $\beta = 1$ for all reconstructions except C15, where we use the inferred posterior samples of β which we discuss in Section 3.4.6.

In Fig. 3.9 we show the simulation-frame $\boldsymbol{B}_{\rm sim}(R)$. All reconstructions except for C23 display a decreasing trend broadly consistent with the ΛCDM expectation. C15 assumes only the continuity equation to relate the galaxy density field to the velocity field, treating the proportionality constant ($\beta^\star$) as a free parameter. By contrast, other reconstructions (partially) assume ΛCDM during the process itself: L24 uses a ΛCDM training set, BORG uses a ΛCDM prior on the ICs and a ΛCDM gravity model, C23 uses a ΛCDM prior on the $z = 0$ density field and linear theory to calculate velocities, and S18 assumes ΛCDM when constraining the ICs with the reverse Zel'dovich approximation (and then to simulate the reconstruction). Most reconstructions show a bulk flow magnitude approaching approximately 600 $\mathrm{km\,s^{-1}}$ near the origin, except for CSiBORG1, which reaches only 400 $\mathrm{km\,s^{-1}}$.

Next we study the radial dependence of the bulk flow in the CMB frame, which we calculate as

$$\boldsymbol{B}(R) = \boldsymbol{B}_{\rm sim}(R) + \boldsymbol{V}_{\rm ext}. \tag{3.28}$$

Note that this exhibits radial variation only within the reconstruction box: beyond it $\boldsymbol{B}(R)$ equals the constant $\boldsymbol{V}_{\rm ext}$. In Fig. 3.10 we present the CMB frame bulk flow magnitude and direction in galactic coordinates. We overplot the velocity of the Local Group in the CMB frame (on which more below) as well as the 1σ expectation for the bulk flow magnitude in ΛCDM using a top-hat window function of radius R (see e.g. Sec. 6.1.2 of Boruah et al. 2020). We find that the bulk flow magnitude for most reconstructions aligns well with the ΛCDM expectation. This is not true for CSiBORG1 and C23, however, which overpredict the bulk flow at large R. In CSiBORG1 this is driven by the large $|\boldsymbol{V}_{\rm ext}|$ (Fig. 3.8), while for C23 it is driven by the large internal bulk flow (Fig. 3.9). Although this trend does not indicate significant tension with ΛCDM at the radii we probe ($\lesssim 150\ h^{-1}$ Mpc), Fig. 3 of Courtois et al. (2023) shows that their bulk flow appears to continue increasing beyond the radii we

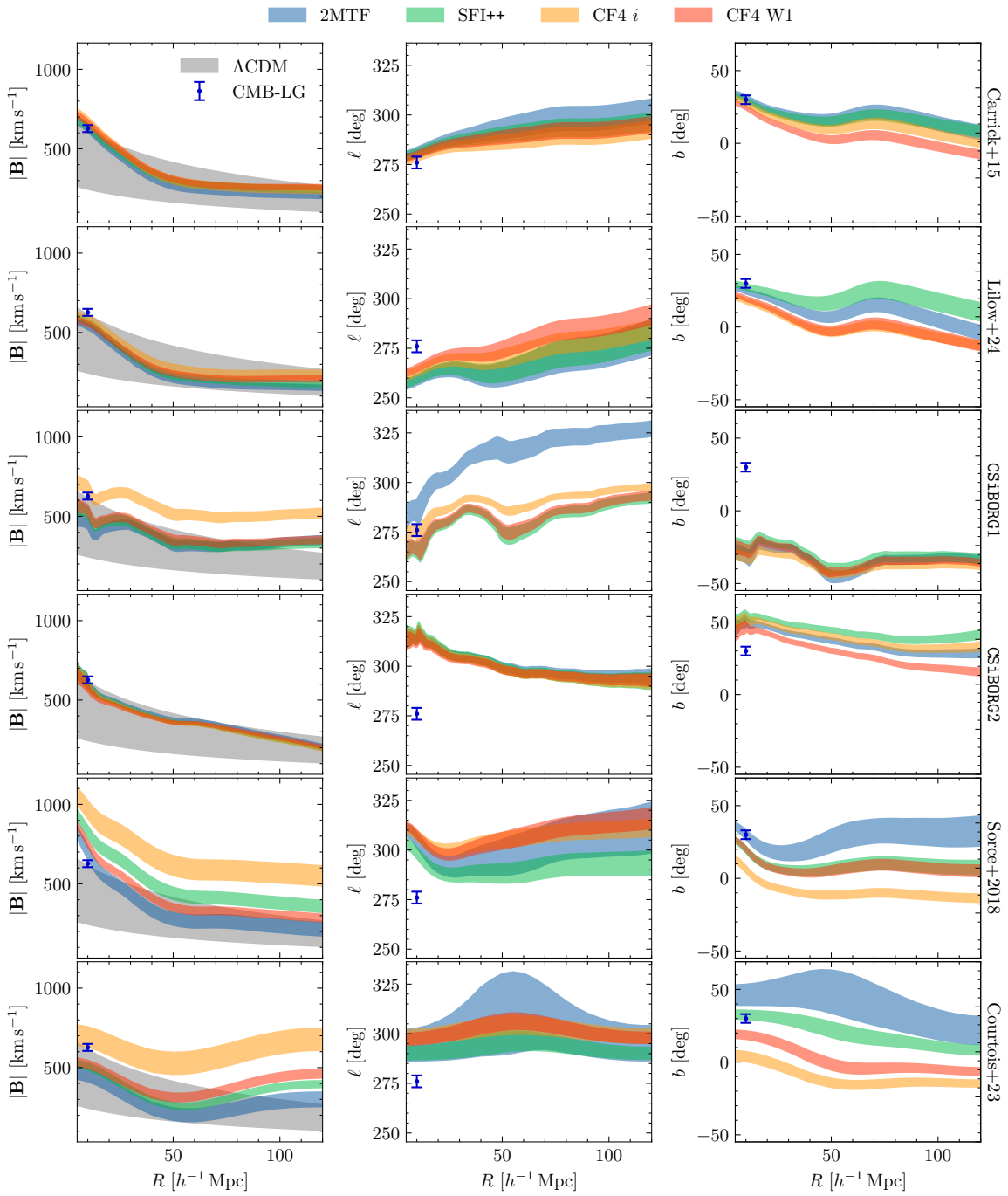


Figure 3.10: The bulk flow magnitude (left column) and direction in galactic longitude (ℓ; centre column) and latitude (b; right column) as a function of radial distance for each reconstruction (rows). The bulk flow is computed as $\boldsymbol{B}_{\rm sim} + \boldsymbol{V}_{\rm ext}$, where $\boldsymbol{B}_{\rm sim}$ is the bulk flow within each box (Fig. 3.9), and $\boldsymbol{V}_{\rm ext}$ is the external velocity constrained by the displayed peculiar velocity samples. The grey band represents the 16$^{\rm th}$ to 84$^{\rm th}$ percentile range for the bulk flow magnitude expected in ΛCDM, while the single dark blue 1σ error bar indicates the motion of the Local Group in the CMB frame (Planck Collaboration et al., 2020a). This is plotted slightly offset from $R = 0$ simply for visual clarity.

plot, posing a greater challenge to ΛCDM if true. This is also evident in Watkins et al. (2023), based on the same data.

Examining the bulk flow direction as a function of radial distance, we observe that the reconstructions of C15 and L24 display similar trends, though they do differ in longitude. The two CSiBORG reconstructions show significant disagreement in latitude. This disagreement between the two BORG reconstructions is not entirely surprising. The IC inference of CSiBORG1 used an inferior gravity model and a different galaxy bias model (Stopyra et al., 2024a), leading to an overestimation of cluster masses, which in turn should affect the local velocity field. Specifically, CSiBORG1 used a 10-step particle-mesh solver for the gravitational dynamics, whereas CSiBORG2 adopted a 20-step COLA integrator. The impact of gravity model accuracy on the BORG ICs is discussed in detail by Stopyra et al. (2024a). This is reflected in the evidence comparison, where CSiBORG1 is strongly disfavoured relative to CSiBORG2. Finally, the CosmicFlows-based reconstructions demonstrate similar directional trends, but the preferred $\boldsymbol{V}_{\rm ext}$ varies significantly between the peculiar velocity samples leading to more prominent differences in bulk flow direction across samples.

Finally we investigate consistency with the one piece of information we have about the bulk flow curve: that the velocity of the Local Group (LG) in the CMB frame is 620 ± 15 km s^{-1} towards $(\ell, b) = (271.9^\circ \pm 2.0^\circ, 29.6^\circ \pm 1.4^\circ)$, as inferred from the CMB dipole (Planck Collaboration et al., 2020a). To recover the predicted LG velocity $\boldsymbol{B}_{\rm LG}$, we evaluate the velocity fields at the origin to obtain $\boldsymbol{B}_{\rm sim}(R = 0)$ and add to it $\boldsymbol{V}_{\rm ext}$. However, this cannot be directly compared to the observed CMB–LG velocity due to the σ_v scatter describing a point-by-point mismatch between predicted and observed velocities. To account for this we calculate $\boldsymbol{B}_{\rm LG}$ as

$$\boldsymbol{B}_{\rm LG} = \boldsymbol{B}_{\rm sim}(R = 0) + \boldsymbol{V}_{\rm ext} + \boldsymbol{V}_{\rm rand}, \tag{3.29}$$

where $\boldsymbol{V}_{\rm rand}$ is a random vector with components drawn from a mean-zero Gaussian with

standard deviation σ_v in each component. The result is shown in Fig. 3.11.

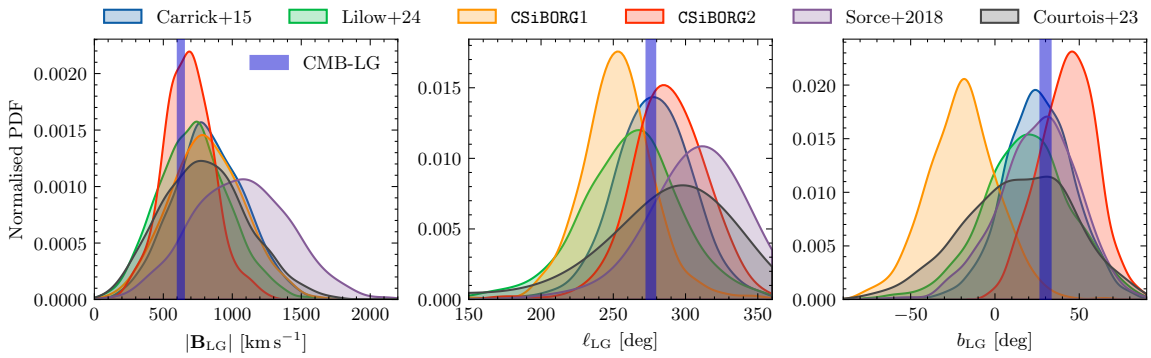


Figure 3.11: The predicted LG velocity following Eq. (3.29), where we propagate the σ_v noise. We show its magnitude, longitude, and latitude ($|\boldsymbol{B}_{\rm LG}|$, $\ell_{\rm LG}$, $b_{\rm LG}$) in the CMB frame in galactic coordinates calibrated against the CF4 TFR W1-band sample for the various reconstructions. All predicted LG velocities are consistent with the observed CMB-LG velocity (1σ region in dark blue rectangles, Planck Collaboration et al. 2020a), although CSiBORG1 shows a $\sim 2\sigma$ deviation in latitude.

We find that propagating the σ_v uncertainty significantly broadens the uncertainties on the predicted LG motion compared to when this scatter is not included, since its value is comparable to the LG–CMB speed. Additionally, the variations between the different peculiar velocity samples, each with its specific $\boldsymbol{V}_{\rm ext}$ and σ_v, are no longer significant. Therefore, in Fig. 3.11 we present the predicted LG motion for the CF4 TFR W1 sample only, which is representative of all samples' results. We find agreement with the observed CMB-LG velocity in both magnitude and direction for all models; the only minor tension is found in CSiBORG1, where the predicted latitude deviates from the CMB dipole by $\sim 2\sigma$ ($-22^\circ \pm 21^\circ$).

3.4.6 Growth rate measurement

In the reconstruction of C15, the posterior samples on $\beta^\star$, the velocity scaling parameter, are related to the parameter combination $f\sigma_{8,\rm NL}$ through Eq. (3.1). As many cosmological probes measure the amplitude of the linear field $\sigma_{8,\rm L}$, we wish to convert this to $f\sigma_{8,\rm L}$. A mapping between $\sigma_{8,\rm NL}$ and $\sigma_{8,\rm L}$ was derived by Juszkiewicz et al. (2010), but we

derive an alternative mapping using the non-linear matter power spectrum. To this end, we first obtain $\sigma_{8,\mathrm{NL}}$ from $f\sigma_{8,\mathrm{NL}}$ by using the approximation $f \approx \Omega_{\mathrm{m}}^{0.55}$ (Wang & Steinhardt, 1998). For all these conversions, we assume a flat ΛCDM with the cosmological parameters $h = 0.6766$, $\Omega_{\mathrm{m}} = 0.3111$, $\Omega_{\mathrm{b}} = 0.02242/h^2$ and $n_s = 0.9665$, except the amplitude of the primordial power spectrum A_{s} which we wish to infer. This is achieved by running a root-finding algorithm, where, for a given A_{s}, $\sigma_{8,\mathrm{NL}}$ is computed by performing the appropriate integral of the non-linear matter power spectrum, which is evaluated using the `syren-new` emulator[7] (Bartlett et al., 2024a; Sui et al., 2025). Once the corresponding value of A_{s} has been obtained, this is converted to the linear value of $\sigma_{8,\mathrm{L}}$ using the conversion given in Eq. 5 of Sui et al. (2025). For our chosen cosmological parameters, we find a correction to the mapping previously derived by Juszkiewicz et al. (2010): the non-linear matter power spectrum-based mapping from $\sigma_{8,\mathrm{NL}}$ to $\sigma_{8,\mathrm{L}}$ is approximately three per cent lower than reported by Juszkiewicz et al. (2010). The exact value of the discrepancy ranges from approximately 3–5 per cent, depending on the exact non-linear prescription used. For example, using the Mead et al. (2016) implementation of `hmcode` gives a value which is five per cent lower than the Juszkiewicz et al. formula. We choose to report values using the `syren-new` approximation of the non-linear power spectrum for this conversion due to its better agreement with N-body simulations (Bartlett et al., 2024a; Sui et al., 2025) compared to `hmcode` and `halofit` (Smith et al., 2003; Takahashi et al., 2012). We show this mapping in Fig. 3.12. Similarly, Hollinger & Hudson (2024) derived the non-linear to linear σ_8 mapping using `halofit` (Smith et al., 2003; Takahashi et al., 2012) and reported comparable disagreements with the results of Juszkiewicz et al. (2010).

In Fig. 3.13, we compare our measurements of $\beta^\star$ to various previous calibrations of C15. We find $\beta^\star = 0.459 \pm 0.013$ for the joint sample (LOSS, Foundation+CF4 TFR i-band, CF4 TFR W1-band). This measurement is under the assumption of a linear bias parameter b_1, which is inferred jointly with an uninformative prior. By contrast, assuming

[7] https://github.com/DeaglanBartlett/symbolic_pofk

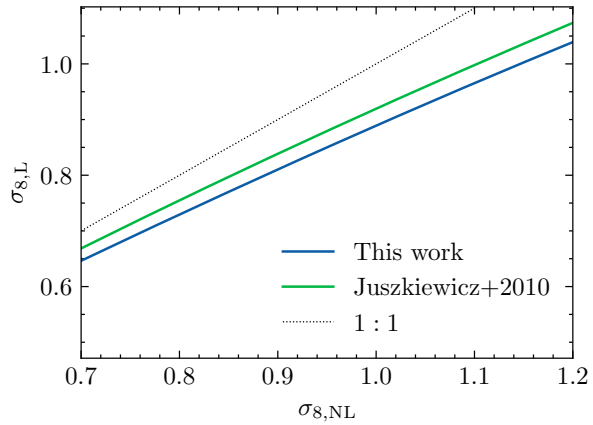


Figure 3.12: Mapping from $\sigma_{8,\mathrm{NL}}$ (computed from the non-linear matter power spectrum) to the linearly evolved $\sigma_{8,\mathrm{L}}$. For our cosmological parameters ($h = 0.6766$, $\Omega_\mathrm{m} = 0.3111$, $\Omega_\mathrm{b} = 0.02242/h^2$, $n_s = 0.9665$), $\sigma_{8,\mathrm{L}}$ is ∼3 per cent lower than derived by Juszkiewicz et al. (2010).

a power-law bias in density would systematically underpredict $\beta^\star$, since the model would inflate the probability of galaxies residing in overdense regions, where peculiar velocities are naturally higher, and thus attempt to lower $\beta^\star$. As C15 is a linear-theory reconstruction with an effective smoothing scale of 4 h^{-1} Mpc, a linear bias is more consistent with the assumptions underlying it.

The measured $\beta^\star$ is similar to both the W1 and i-band-only measurement individually. We summarise the inferred values of $\beta^\star$ in Table 3.4. The joint sample includes approximately twice as many data points, which, as expected, reduces the uncertainty in $\beta^\star$ by roughly a factor of $\sqrt{2}$. As shown in Fig. 3.13, our measurement is only marginally higher than those reported by Carrick et al. (2015) and Boruah et al. (2020), but significantly lower than the values of Said et al. (2020) and Boubel et al. (2024), suggesting inconsistency at the level of $> 4\sigma$. The difference is particularly noteworthy for Boubel et al. (2024), who also used the CF4 TFR sample, whereas Said et al. (2020) used fundamental plane data. One key difference is that Said et al. (2020), in their Eq. 23, adopt a distance prior that effectively assumes a galaxy bias of $b_1 = 1$, i.e. that the peculiar velocity sample shares the same bias as the 2M++ sample, whereas we consistently infer $b_1 > 1$ (see Table 3.4). It is noteworthy that the SFI++ sample alone favours relatively small values of $\beta^\star$, consistent with Said et al. (2020) and Boubel et al. (2024). However, because a subset of SFI++ sources is included in the CF4 compilation, we exclude it from the joint constraint. This discrepancy may indicate

a systematic effect in the SFI++ data.

Catalogue	$\beta^\star$	S_8	b_1
2MTF	0.455 ± 0.023	0.787 ± 0.046	1.208 ± 0.030
SFI++	0.369 ± 0.020	0.650 ± 0.041	1.212 ± 0.018
LOSS	0.395 ± 0.107	0.687 ± 0.178	1.295 ± 0.178
Foundation	0.294 ± 0.077	0.523 ± 0.136	1.141 ± 0.104
CF4 i	0.429 ± 0.018	0.747 ± 0.039	1.071 ± 0.016
CF4 W1	0.475 ± 0.019	0.817 ± 0.041	1.191 ± 0.011
Joint	**0.459 ± 0.013**	**0.793 ± 0.035**	-

Table 3.4: Inferred values of $\beta^\star$, S_8 and b_1 assuming the linear-theory reconstruction of Carrick et al. (2015) for the peculiar velocity samples considered in this work. The "joint" sample denotes a combined inference from LOSS, Foundation, and the CF4 i- and W1-band TFR data. We convert $\beta^\star$ to S_8 following the approach of Section 3.4.6. We report the posterior mean and standard deviation. We do not report b_1 for the joint sample, as each catalogue has a separate b_1.

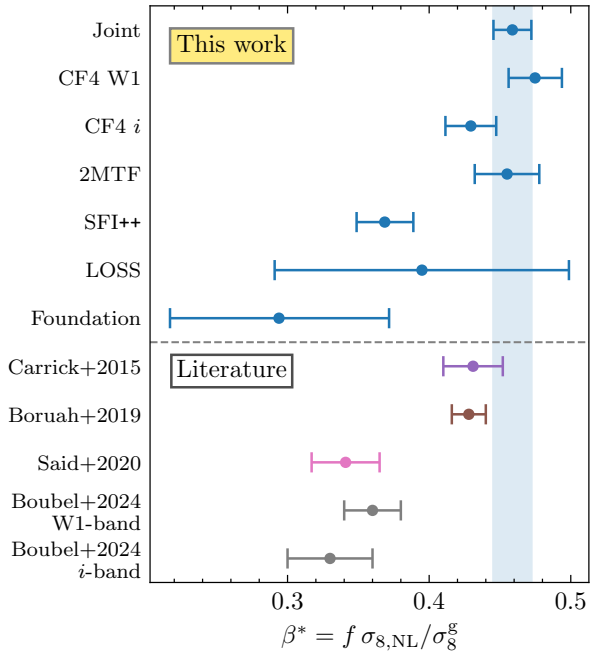


Figure 3.13: Velocity field scaling $\beta^\star$ of Carrick et al. (2015) calibrated against various peculiar velocity samples, compared to literature values (Carrick et al., 2015; Boruah et al., 2020; Said et al., 2020; Boubel et al., 2024). "Joint" denotes LOSS+Foundation+CF4 TFR W1+i. Error bars are 1σ.

Finally, we transform the posterior samples of $\beta^\star$ into $S_8 \equiv \sigma_{8,\mathrm{L}}\sqrt{\Omega_\mathrm{m}/0.3}$. Using the value of $\beta^\star$ calibrated on the joint sample, we find $S_8 = 0.793 \pm 0.035$ under the assumption that $\Omega_\mathrm{m} = 0.3111$, based on the mapping of σ_8 described above. In contrast, applying the mapping of Juszkiewicz et al. (2010) yields $S_8 = 0.820 \pm 0.036$ for the same $\beta^\star$. Our inferred S_8 is in good agreement with Boruah et al. (2020), while using larger TFR samples.

In contrast, we find a substantially higher value than Said et al. (2020), who reported $S_8 = 0.637 \pm 0.054$. Assuming Ω_m values of either 0.25 or 0.35 alters our results by less than one per cent.

In Fig. 3.14 we compare the inferred S_8 to various values from the literature. Our results are in good agreement with other peculiar velocity measurements of S_8 of Boruah et al. (2020); Nusser (2017); Huterer et al. (2017), but disagree significantly with Boubel et al. (2024); Said et al. (2020), who claimed a $> 3\sigma$ growth-rate tension with the CMB. However, our constraints agree with weak lensing results to within 1σ (Wright et al., 2025; García-García et al., 2024; Dark Energy Survey and Kilo-Degree Survey Collaboration et al., 2023; Abbott et al., 2022) and to within 1.4σ with *Planck* TT, TE, EE+lowE+lensing (Planck Collaboration et al., 2020b). Hollinger & Hudson (2024) used mock 2M++ realisations in conjunction with (smaller than CF4) mock TFR catalogues to show that the sample variance in $f\sigma_{8,\mathrm{NL}}$ is approximately five per cent. Incorporating this additional (conservative) five per cent uncertainty into the S_8 constraint would result in $S_8 = 0.793 \pm 0.053$.

3.5 Discussion

In this section, we provide a detailed summary of the results (Section 3.5.1), discuss their interpretation (Section 3.5.2), and address the caveats of our analysis (Section 3.5.3). We then compare our results with the literature (Section 3.5.4) and suggest possible future studies (Section 3.5.5).

3.5.1 Summary of results

We have evaluated several local Universe reconstructions using various peculiar velocity samples and metrics, ranging from a quantitative to more qualitative analyses. Our first metric is the Bayesian evidence comparison of the reconstructions, shown in Fig. 3.4. This comparison consistently and strongly favoured `CSiBORG2` over the other reconstructions.

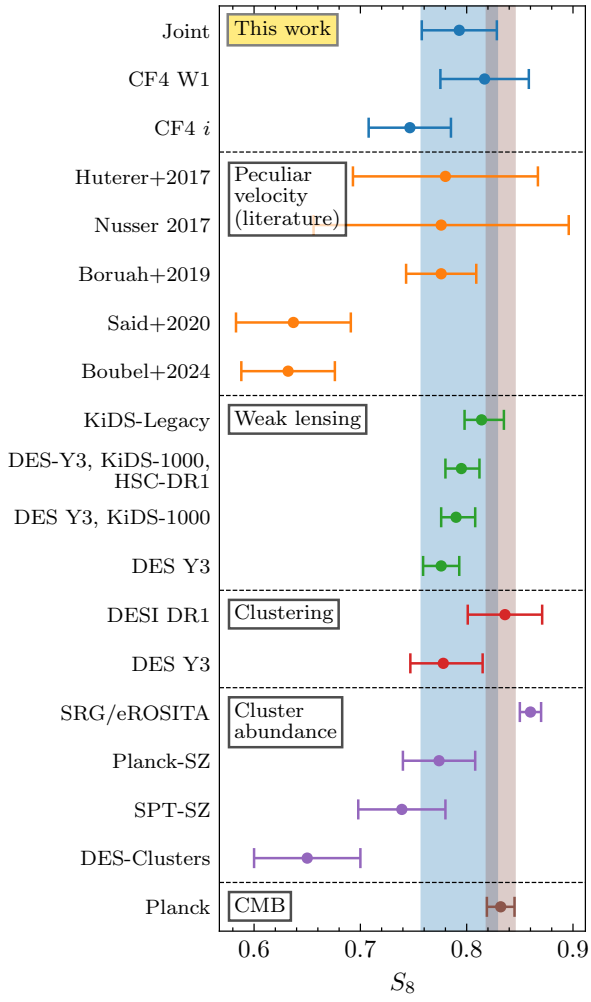


Figure 3.14: $S_8 \equiv \sigma_{8,\mathrm{L}}\sqrt{\Omega_\mathrm{m}/0.3}$ inferred with Carrick et al. (2015) calibrated against the joint peculiar velocity sample (LOSS+Foundation+CF4 TFR W1+i). Comparison values from peculiar velocities (Huterer et al., 2017; Nusser, 2017; Boruah et al., 2020; Said et al., 2020), weak lensing (Wright et al., 2025; García-García et al., 2024; Dark Energy Survey and Kilo-Degree Survey Collaboration et al., 2023; Abbott et al., 2022), clustering (Adame et al., 2025b; Porredon et al., 2022), cluster abundance (Ghirardini et al., 2024; Bocquet et al., 2019; Planck Collaboration et al., 2016; Abbott et al., 2020), and *Planck* TT,TE,EE+lowE+lensing (Planck Collaboration et al., 2020b). Errors are 1σ.

The second and third ranking reconstructions are those of L24 and C15, respectively, both of which, like CSiBORG2, are constrained using the 2M++ data. The earlier version of BORG, CSiBORG1, performs slightly worse than CSiBORG2, likely stemming from the fact that the ICs of CSiBORG2 were constrained using a more accurate, 20-step COLA gravity model, whereas the ICs of CSiBORG1 were based on a 10-step particle-mesh gravity model (for further details, see Stopyra et al. 2024a). In contrast, the CosmicFlows-based reconstructions of S18 and C23 consistently underperform when tested against both SN and TFR samples. This outcome is particularly surprising when testing these reconstructions on the samples that were used to constrain them, for which the other models still perform better, so that the comparison is not independent for S18 or C23. We comment on this further in Section 3.5.3.

Previous studies typically set σ_v to 150 $\mathrm{km\,s^{-1}}$ based on a comparison to N-body simulations (Carrick et al., 2015; Boruah et al., 2020; Said et al., 2020). However, σ_v depends not only on the reconstruction method but also on the peculiar velocity sample. For this reason, we instead treat σ_v as a free parameter of the model with a reference (scale-invariant) prior $\pi(\sigma_v) \propto 1/\sigma_v$. As shown in Fig. 3.5 and Fig. 3.6, the inferred σ_v correlates inversely with the goodness-of-fit and hence σ_v provides a good indicator of the quality of a reconstruction. However, because of the dependence of σ_v on the sample in question, it should not be compared between samples, but only between reconstructions for a given sample. For example, a galaxy catalogue with primarily field galaxies will have a lower σ_v than a catalogue with galaxies in clusters because of the greater small-scale motions (not captured by the reconstructions) in the latter.

We then examined the parameter β, which scales the predicted velocity field. For C15, we denote it as $\beta^\star$ as it is related to the growth rate of structure (see Section 3.4.6). For the other reconstructions, where velocities are consistently sourced by the total matter distribution (except L24), scaling the predicted velocity field with β should not be necessary, implying $\beta = 1$. However, even in these cases, if allowed to vary, the recovered β can be

affected by the resolution at which the field is constrained, inaccurate reconstruction or discrepancies between the cosmological parameters assumed by the reconstruction and the actual description of the Universe. For example, an overly smoothed field may exhibit suppressed velocities, necessitating $\beta > 1$. We tested this hypothesis in Fig. 3.7 by treating β as a free parameter for all reconstructions. We find that CSiBORG2 is consistent with $\beta = 1$ for most samples, except for CF4 TFR i-band samples, where it is biased high (though at less than 3σ). Interestingly, both S18 and C23 require higher values of β on this catalogue as well. On the other hand, both CSiBORG1 and L24 prefer a lower β value across all samples, suggesting that they overpredict peculiar velocities.

Regarding the external flow magnitude, it is noteworthy that L24 indicate their neural network reconstruction is sensitive to "super-survey" scales through redshift-space distortions. Our findings in Fig. 3.8 support this claim, as the external flow magnitude required by L24 is nearly consistent with zero, a distinction not seen in other reconstructions which also use the same redshift-space data. The preferred model, CSiBORG2, also favours a relatively small external flow, whereas the less favoured CSiBORG1 requires a substantially larger flow magnitude. However, similar to β, there is no clear relationship between the goodness-of-fit and the external flow magnitude.

The BORG inferences self-consistently model redshift-space distortions: using the sampled density and velocity field, forward-modelled to $z = 0$, the predicted galaxy counts are mapped into redshift space using the corresponding velocity field and directly compared to the observed redshift-space distribution. However, the finite size of the box used to constrain the ICs prevents the modelling of the largest-scale power spectrum modes. As a result, some large-scale power may be missing in the redshift-space mapping (or in the density field). In the modelling employed here, we assume an existing reconstruction and forward-model the redshifts from the real-space positions constrained by the direct-distance tracers, thereby self-consistently accounting for redshift-space distortions. To address the missing large-scale power, we include $\boldsymbol{V}_{\rm ext}$, representing the contribution to the reconstructed velocity

field sourced by external matter.

Lastly, we analysed the predicted bulk flow curves by combining the internal bulk flow of each reconstruction with the external flow. Near the origin, our position provides a constraint on the bulk flow through the LG-CMB velocity, as derived from the CMB dipole (Planck Collaboration et al., 2020a). As shown in Fig. 3.11, all reconstructions (with the possible exception of CSiBORG1) are consistent with this velocity when accounting for σ_v, which reflects the expected scatter between a given galaxy's velocity and the velocity field estimate at its position.

Regarding the radial dependence, we find that while each reconstruction tracks a somewhat distinct bulk flow curve, there is a general agreement in the bulk flow direction at radii of approximately 100 h^{-1} Mpc among C15, S18, C23, and L24. The BORG reconstructions, however, exhibit bulk flows that point in different directions, with CSiBORG2 aligning more closely with those mentioned but slightly shifted towards the Shapley Supercluster (approximately $\ell = 305°$, $b = 32°$). Since our peculiar velocity samples are generally limited to redshifts below $z = 0.05$, we do not draw comparisons with Watkins et al. (2023), who report unexpectedly large and increasing bulk flows at radii of 150 to 250 h^{-1} Mpc.

3.5.2 Interpretation of results

While we have determined the relative preferences for the local Universe reconstructions, the reasons underlying these preferences remain unclear. To address this, we extend our analysis in two ways. First, we divide the galaxy samples into subsets based on their observed redshift, inclination errors, or absolute magnitude, and repeat the evidence comparison for each subset. Second, we correlate differences in log-likelihoods between reconstructions with galaxy properties in an attempt to discover which galaxy types drive the preference of some reconstructions over others.

We begin by considering tracers with observed redshift $z_{\rm CMB} < 0.025$ to examine the inner regions of the reconstructions, compared to the usual threshold of $z_{\rm CMB} < 0.05$.

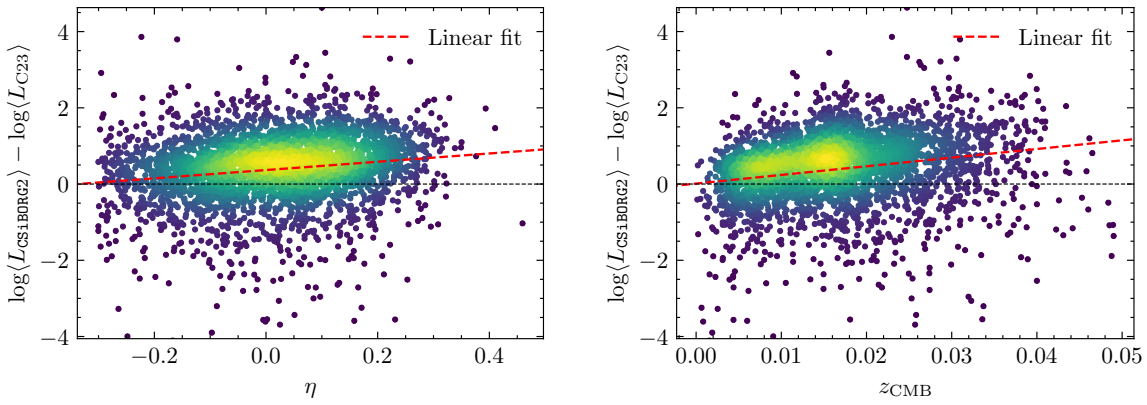
The evidence ratios for all peculiar velocity samples remain qualitatively identical to those presented in Fig. 3.4: CSiBORG2 is the most preferred reconstruction, L24 and C15 exhibit the second- and third-best goodness-of-fit, respectively, with similar performance, while the CosmicFlows reconstructions are disfavoured. Next, we focus only on the CF4 TFR i- and W1-band samples, dividing them into subsets of galaxies with either below- or above-average inclination errors, and analyse the evidence ratios in both cases. As before, the conclusions from our fiducial analysis are robust to these variations. Since inclination errors propagate into the linewidth measurements and are already accounted for, this outcome is unsurprising. It does however indicate that there are no systematic errors associated with the linewidth that scale with the observational uncertainties. Lastly, we estimate the absolute magnitude of each galaxy from its apparent magnitude and redshift-space distance (converting the observed redshift in the CMB frame into a distance modulus). The two CF4 TFR samples are then split into subsets of the 25 per cent brightest and faintest galaxies, and the evidence ratios are recomputed. Yet again, we find that our qualitative conclusions are unchanged.

Next, we repeat our fiducial analysis on the CF4 TFR sample without splitting the data. We record the log-likelihood of each galaxy at each posterior sample and compute the average likelihood $\langle L \rangle$. Subsequently, we calculate the differences in $\log\langle L \rangle$ between reconstructions on the same peculiar velocity samples and correlate these differences with galaxy properties: sky position, observed redshift, apparent magnitude, linewidth, linewidth error, inclination, and inclination error.

Among these features, significant trends emerge only for linewidth and observed redshift. The most pronounced differences occur when comparing CSiBORG2 with C23. In Fig. 3.15a, we show a scatter plot of the differences in $\log\langle L \rangle$ between CSiBORG2 and C23 as a function of linewidth in the CF4 TFR W1-band sample. We find a Spearman correlation coefficient of 0.18 (p-value of 10^{-27}), indicating that CSiBORG2 is favoured more for galaxies with larger linewidths, corresponding to more massive objects. Since these trace higher-density

regions, this result reaffirms that the improved performance of CSiBORG2 is partly due to its better non-linear modelling via BORG. Similar significant trends are observed when comparing CSiBORG2 with all other reconstructions, including CSiBORG1.

Fig. 3.15b illustrates that CSiBORG2 consistently achieves higher likelihoods for galaxies at greater distances when compared to the least effective reconstructions C23 (Spearman correlation coefficient of 0.23, p-value of 10^{-42}) or S18. This suggests that the constraints from CosmicFlows-based reconstructions degrade more rapidly with distance, despite the reconstructed volume in C23 extending up to $z = 0.08$. However, even though CSiBORG2 demonstrates better performance for brighter and more distant galaxies, we have confirmed that its preference remains significant even when considering only nearby or faint objects. The trend with observed redshift is not found for either C15 or L24 when compared to CSiBORG2. Interestingly, there is no preference between CSiBORG2 and C23 for the nearest or lowest-linewidth galaxies.



(a) The difference in $\log\langle L\rangle$ as a function of linewidth η. The Spearman correlation coefficient is 0.18 (p-value of 10^{-27}).

(b) The difference in $\log\langle L\rangle$ as a function of observed redshift z_{CMB}. The Spearman correlation coefficient is 0.23 (p-value of 10^{-42}).

Figure 3.15: Differences in the logarithms of the average likelihoods of individual galaxies $\log\langle L\rangle$, averaged over the posterior samples, between CSiBORG2 and the reconstruction by Courtois et al. (2023). The plots illustrate these differences as functions of linewidth η (left) and observed redshift z_{CMB} (right). Positive values indicate preference for CSiBORG2 and we include a line of best fit. This indicates that CSiBORG2 is preferentially better than the field of Courtois et al. (2023) for higher linewidth (mass) galaxies and at higher redshifts. No significant trends are found with other galaxy properties.

3.5.3 Caveats

One caveat in our comparison is that the goodness-of-fit depends on the resolution of the reconstruction. Higher resolution reconstructions would be expected to be preferred simply because they can capture smaller scale motions (before they become noise-dominated). To test this, we smoothed all fields to the same, coarse resolution shown in Fig. 3.4, which did not change our main conclusions: CosmicFlows-based reconstructions remain disfavoured, though L24 becomes preferred over CSiBORG2. However, reaching higher resolutions is an advantage of a reconstruction method as, in principle, the goal is to reconstruct the local Universe down to non-linear scales, reducing the required extra random scatter σ_v. The preference of the 2M++ derived reconstructions over CosmicFlows-derived ones suggests either that the former dataset is more informative, or that there are limitations of the methods applied to the CosmicFlows data. On the one hand, this result is surprising as the peculiar velocities probe directly the matter distribution of the Universe, whereas working with galaxy fields requires a careful bias treatment. On the other hand, while the 2M++ galaxy redshift-space positions are in practice nearly only shot-noise dominated, the CosmicFlows peculiar velocity data is noisy and sparse: out of the $38,000$ objects in the grouped CF4 catalogue used by C23 only $18,000$ are within $z_{\rm obs} < 0.05$.

We aimed for a broad comparison of the reconstructions available in the literature, without separating the effect of the reconstruction technique and the input data. For instance, S18 is based on a grouped version of the CF2 catalogue, which contains approximately only 5,000 galaxies. In contrast, C23, another reconstruction based on peculiar velocities, uses the larger CF4 catalogue, yet performs similarly to S18 according to our metrics. The input data also set how independent each test is. The reconstructions constrained by galaxy redshift-space positions (CSiBORG1, CSiBORG2, and C15 from 2M++, and L24 from 2MRS) use information broadly independent of the TFR and SN peculiar velocities tested here, with the incidental galaxy overlap carrying negligible shared information, so their double count-

ing is insignificant. The CosmicFlows-based reconstructions instead use peculiar velocities that overlap our test data: C23 draws on CF4, of which our CF4 TFR sample is a subset, so for these the comparison is not a fully independent check. We demonstrated that the non-linear BORG-based reconstructions outperform the other methods. However, this improvement does come at a significantly higher computational cost to obtain the ICs. Once the BORG-based reconstructions are produced, though, this computational cost is irrelevant for any follow-up analysis. In this respect it is worth noting that reconstruction of the velocity field is not a main aim of the BORG programme, but rather it comes "for free." Nevertheless, it is worth emphasising that the linear method of C15, while orders of magnitude less computationally expensive than BORG, still performs better than both S18 and C23.

We do not evaluate the quality of the reconstructions in specific regions but instead use all-sky samples. The two SN samples are an exception: both are small and share a similar, predominantly northern footprint, limiting their leverage on the bulk flow and preventing them from independently testing sky-dependent systematics. Crucially, the broader TFR set includes all-sky or near-all-sky samples, with the CF4 W1-band catalogue in particular covering the southern sky absent from the SNe; the CF4 SDSS i-band sample is limited to the SDSS footprint, but all the TFR samples nevertheless yield a qualitatively consistent evidence ranking (Fig. 3.4), so our conclusions are not driven by the SN footprint. The TFR samples mainly include disk (field) galaxies, leading to a focus on lower-density regions rather than galaxy clusters. Evaluating velocities in clusters, which could be done with a fundamental plane sample, is beyond our present scope. However, we discovered in Section 3.5.2 that the preference for BORG is even greater for brighter galaxies. Since these trace on average higher-density regions, it is likely that the fundamental plane would provide even greater evidence for BORG. An example of related work is that by Sorce et al. (2024), who examined velocities along the LOS to nearby clusters.

A potential systematic effect relevant for all large-scale structure analyses is the Doppler modulation induced by the observer's motion with respect to the cosmic rest frame (e.g. Ellis

& Baldwin 1984; Kaiser 1987; Nusser & Davis 2011; Bahr-Kalus et al. 2021). This effect introduces a dipolar anisotropy in the observed source distribution, and can bias both the inferred density and peculiar velocity fields. In the literature, related systematics are sometimes discussed under the heading of the "Kaiser-rocket effect," which more commonly refers to the dipolar modulation of the two-point correlation function resulting from the observer's motion (Kaiser, 1987). While the Kaiser-rocket effect in its original formulation is a concern primarily for correlation function analyses, its underlying physical origin (Doppler boosting) can more generally impact forward-modelling frameworks such as ours if not properly accounted for.

This arises because the observer's motion induces an apparent overdensity of sources in redshift space along the direction of motion. For example, BORG mitigates this effect by boosting the observed redshifts into the CMB frame, thereby adopting the CMB frame as the cosmic rest frame. Moreover, since e.g. both C15 and BORG self-consistently compute the peculiar velocities, in practice this effect is not a problem. In this work, we also use CMB-boosted redshifts and additionally introduce $\boldsymbol{V}_{\rm ext}$ as a free parameter to capture any residual effects of the observer's motion (or of the bulk motion of the sample relative to the CMB). A relevant test of this effect was performed by Nusser et al. (2014), who demonstrated using simulations that, without imposing the cosmic rest frame a priori, the motion of the LG can only be recovered to within 150–200 $\mathrm{km\,s^{-1}}$ in amplitude and 10° in direction (assuming linear theory). This is of similar order of magnitude as the behaviour shown in Fig. 3.11, where we propagate the σ_v term (representing the expected redshift scatter between observations and predictions from direct distance tracers and the assumed peculiar velocity), which is typically of order 300 $\mathrm{km\,s^{-1}}$, to the predicted LG velocity.

3.5.4 Comparison with literature

Peculiar velocity corrections are essential for any local Universe cosmology to model deviations from the Hubble flow. For instance, in the main Pantheon+ SN analysis (Brout

et al., 2022), Carr et al. (2022) define "Hubble diagram redshifts," $z_{\rm HD}$, which are derived by converting the observed heliocentric SN (host) redshifts to the CMB frame and applying corrections for peculiar motion based on C15.

The selection of C15 was previously motivated by Peterson et al. (2022), who found it to be the most effective in reducing scatter in the SN Hubble diagram. Although the only common reconstruction between Peterson et al. (2022) and this work is C15, their analysis also included a CosmicFlows3-based reconstruction (Tully et al., 2016), which was found to increase the scatter rather than reduce it. This conclusion aligns qualitatively with our findings, where we identify the reconstruction by C15 as preferred over the CosmicFlows-based reconstructions. However, we also demonstrate that the latest BORG reconstruction has the potential to reduce scatter in the Hubble diagram even further. Moreover, unlike C15, BORG provides a measure of reconstruction uncertainty, allowing for propagation of this uncertainty into the resulting Hubble diagram.

A related study is that of Boruah et al. (2021), who compared C15 to kernel-smoothed velocity field estimates derived from either fundamental plane or TFR samples by calculating a Bayesian model evidence. They show that, because of the noise in the velocity estimates going into the kernel-smoothed velocity estimates, the latter are disfavoured when compared to independent peculiar velocity samples.

3.5.5 Future work

An evident extension of our work is the inclusion of additional peculiar velocity tracers. Two promising candidates are larger SN samples and fundamental plane galaxy samples. The primary advantage of SN samples lies in their smaller distance errors, approximately 5 per cent, compared to the 20 per cent TFR errors. However, SN samples are generally smaller in size.

One major SN dataset is Pantheon+ (Brout et al., 2022), which combines several low-redshift SN surveys. We excluded this sample from our analysis because it would require

careful systematic treatment of the constituent surveys, relying on the covariance matrix provided by Brout et al. (2022). However, this matrix already includes peculiar velocity corrections based on C15, and disentangling these corrections is non-trivial. Simply subtracting their contributions results in a singular, non-invertible covariance matrix (that parts of the Pantheon+ covariance matrix are not positive semi-definite was already noted by Lane et al. 2025). Another promising SN dataset is the Zwicky Transient Facility (ZTF) SN Ia DR2, a volume-limited, unbiased sample of 1000 SNe below $z_{\rm CMB} \sim 0.06$ (Rigault et al., 2025; Amenouche et al., 2025). However, this dataset is not yet publicly available. Two relevant fundamental plane samples are those from the 6dF Galaxy Survey (Campbell et al., 2014) and SDSS (Howlett et al., 2022). Both have been utilised by Said et al. (2020) to calibrate the field of C15, albeit their approach relied on forward modelling the galaxy angular size rather than galaxy distance as we do here. Nevertheless, the consistent evidence comparison across all SN and TFR samples used in this work indicates that our results are robust to choice of dataset, and hence similar conclusions would likely be reached using Pantheon+ or ZTF DR2.

3.6 Conclusion

We introduce a novel framework for jointly calibrating distance indicator relations and velocity field reconstructions using peculiar velocity data in the local Universe. Unlike previous studies, we carefully account for uncertainties in distance-indicating properties within a Bayesian hierarchical model. We use reconstructions summarised in Table 3.1 that include linear theory (Carrick et al., 2015; Courtois et al., 2023), machine learning (Lilow et al., 2024), and resimulations of density field constraints derived from Bayesian forward modelling of either redshift-space galaxy positions (BORG; Jasche & Lavaux 2019; Stopyra et al. 2024a) or peculiar velocities (Sorce & Tempel, 2018). We apply this framework to TFR and SN samples summarised in Table 3.2, with future plans to extend it to fundamental

plane samples. We compute the Bayesian model evidence for the joint distance and flow calibration described in Section 3.3.2, enabling us to compare how well different local Universe reconstructions explain the available peculiar velocity data.

We show that while by our metrics the older BORG reconstruction did not outperform the linear reconstruction of Carrick et al. (2015), the more recent and improved BORG reconstruction based on the ICs of Stopyra et al. (2024a) is significantly preferred. This suggests a strong future potential for non-linear modelling within the BORG framework. We also evaluated the neural-network-based reconstruction of Lilow et al. (2024), finding that it performs marginally better than Carrick et al. (2015). Finally, the CosmicFlows-based reconstructions by Sorce & Tempel (2018) and Courtois et al. (2023) were disfavoured.

We further established that σ_v, the scatter between predicted and observed redshifts, serves as a reliable indicator of the ability of a reconstruction to explain the data. Moreover, β, the velocity field scaling parameter, is generally consistent with unity for the preferred CSiBORG2 reconstruction (as expected), except in the CF4 TFR i-band sample, where it exceeds unity. This suggests that the velocities need to be scaled up to match observations, which is interestingly also true of the two CosmicFlows-based reconstructions.

We also studied the predicted LG motion in the CMB frame. However, due to the σ_v scatter, we found that all reconstructions are broadly consistent with the CMB-derived LG velocity (Planck Collaboration et al., 2020a). We also compared the predicted bulk flow curves as a function of radius and found similar qualitative trends, though the quantitative details vary.

Finally, we calibrated the $\beta^\star$ velocity scaling of the linear theory reconstruction of Carrick et al. (2015) using our peculiar velocity samples, enabling us to place a stringent constraint on the parameter $S_8 \equiv \sigma_{8,\mathrm{L}}\sqrt{\Omega_\mathrm{m}/0.3}$. For the joint sample, we determined $\beta^\star = 0.459 \pm 0.013$, which translates to $S_8 = 0.793 \pm 0.035$. This estimate of S_8 agrees well with both weak lensing results and *Planck*. It also matches some peculiar velocity measurements, but disagrees with others (Boubel et al., 2024; Said et al., 2020).

In conclusion, this work provides a comprehensive comparison of various velocity field reconstructions available in the literature, identifying the latest BORG reconstruction based on ICs from Stopyra et al. (2024a) as the most strongly preferred. Our framework and results will be valuable not only for evaluating local Universe reconstructions but also for cosmological analyses that rely on peculiar velocity corrections, such as SNe or the dark siren method for gravitational waves as well as growth of structure measurements from peculiar velocities. This is highly relevant in the context of the Hubble tension and, more broadly, for any inference requiring physical distances to galaxies.

3.A SN flow model

In Section 3.3.2, we derived the model for joint density and velocity field calibration, alongside the calibration of the TFR. We now extend this framework for SN calibration, aiming to infer the Tripp parameters $\mathcal{A}$ and $\mathcal{B}$, as well as the absolute magnitude of Type Ia SNe, $M_{\rm SN}$. For this purpose, we consider the observed apparent magnitude $m_{\rm obs}$, the "observed" SALT2 light curve parameters: stretch $x_{1,\rm obs}$ and colour correction $c_{\rm obs}$, along with their corresponding "true" values $x_{1,\rm true}$ and $c_{\rm true}$.

The first difference compared to the TFR model is that we assume all sources to have the same absolute magnitude $M_{\rm SN}$ and adopt the distance–magnitude relation

$$m_{\rm standard} = \mu(r) + M_{\rm SN}, \tag{3.30}$$

where $m_{\rm standard}$ is the standardised apparent magnitude. Following the Tripp parametrisation in Eq. (3.4), the predicted SN apparent magnitude is

$$m_{\rm pred} = \mu(r) + M_{\rm SN} - \mathcal{A}\, x_{1,\rm true} + \mathcal{B}\, c_{\rm true}, \tag{3.31}$$

where we explicitly use the true stretch and colour corrections. In this case there are three

latent parameters per source: r, $x_{1,\mathrm{true}}$, and c_{true}. We sample r from the same prior as in Eq. (3.12), and place a (correlated) Gaussian prior on $x_{1,\mathrm{true}}$ and c_{true},

$$\pi(x_{1,\mathrm{true}},\, c_{\mathrm{true}} \mid \boldsymbol{\theta}) = \mathcal{N}\big(\{x_{1,\mathrm{true}},\, c_{\mathrm{true}}\};\, \boldsymbol{\mu}_{\mathrm{SN}},\, \mathbf{C}_{\mathrm{SN}}\big)\,, \tag{3.32}$$

where $\boldsymbol{\mu}_{\mathrm{SN}}$ is the hyperprior mean of $x_{1,\mathrm{true}}$ and c_{true}, inferred with uniform priors. The 2×2 covariance matrix $\mathbf{C}_{\mathrm{SN}}$ has standard deviations σ_i, inferred with the reference (scale-invariant) prior $\pi(\sigma_i) \propto 1/\sigma_i$, and a correlation coefficient assigned a uniform prior in the range $[-1,\, 1]$.

Unlike the TFR case, where we numerically marginalise over both r and η_{true} at each MCMC step, here we marginalise only over r and explicitly sample $x_{1,\mathrm{true}}$ and c_{true} with HMC to avoid evaluating a three-dimensional numerical integral at every step.[8] This results in a high-dimensional posterior, which renders evidence calculation infeasible. For evidence estimation, we therefore approximate the posterior by fixing $x_{1,\mathrm{true}}$ and c_{true} to their observed values, as described in Section 3.3.3. This approach is implicitly followed by Boruah et al. (2020) and other earlier works. The rationale for this method is that the observational uncertainties are generally subdominant to the intrinsic scatter. Here, we directly compare the two methods to verify that the approximate method is reliable. Fig. 3.16 shows the posteriors from the two methods for the Foundation SN sample and the C15 reconstruction. We find that the only parameter significantly affected by use of the approximate method is $\mathcal{B}$. The remaining parameters remain unaffected.

Due to the high dimensionality of the posterior, it is not straightforward to directly assess how this approximation impacts the evidence ratios. We instead compute the Bayesian information criterion (BIC), defined as

$$\mathrm{BIC} = D\, \ln\, N_{\mathrm{gal}} - 2\, \ln\, \hat{\mathcal{L}}, \tag{3.33}$$

[8] Since $x_{1,\mathrm{true}}$ and c_{true} enter linearly through the Tripp formula, they can in fact be analytically marginalised.

where D is the dimensionality of the posterior, $N_{\rm gal}$ is the number of galaxies, and $\hat{\mathcal{L}}$ is the maximum-likelihood value. The BIC can serve as a model selection criterion analogous to Bayesian evidence, where lower BIC values indicate a preferred model. We note that the BIC is a reliable model selection indicator only when the number of observations is significantly larger than the dimensionality of the model. Nonetheless, we compute the BIC values for the full posterior and find good qualitative agreement with our main results presented in Fig. 3.4, which were obtained using the approximated posterior. The ranking of the models remains unchanged, and the magnitude of the differences between them is approximately preserved. Yet another way to test this approximation is to assume a Gaussian posterior and use the Laplace approximation. However, we do not pursue this additional test here, since for the more informative TFR we explicitly marginalise over the per-galaxy latent variables instead, which involves no approximation.

The advantage of this framework is that corrections to the Tripp formula can be incorporated directly by modifying Eq. (3.31) and introducing additional terms that may depend, for example, on host-galaxy mass or environment. In the present analysis we assume independence between SNe by omitting any covariance matrix, since the Tripp calibration is inferred jointly, and we neglect correlations between the SALT2 light curve parameters. Additional covariance can, however, be included in the likelihood of either the observed redshift or apparent magnitude. While this would render the one-dimensional marginalisation over radial distance computationally prohibitive, the distances could instead be sampled explicitly. Incorporating more general selection effects in this context is non-trivial and may require a simulation-based approach (Boyd et al., 2024, 2026).

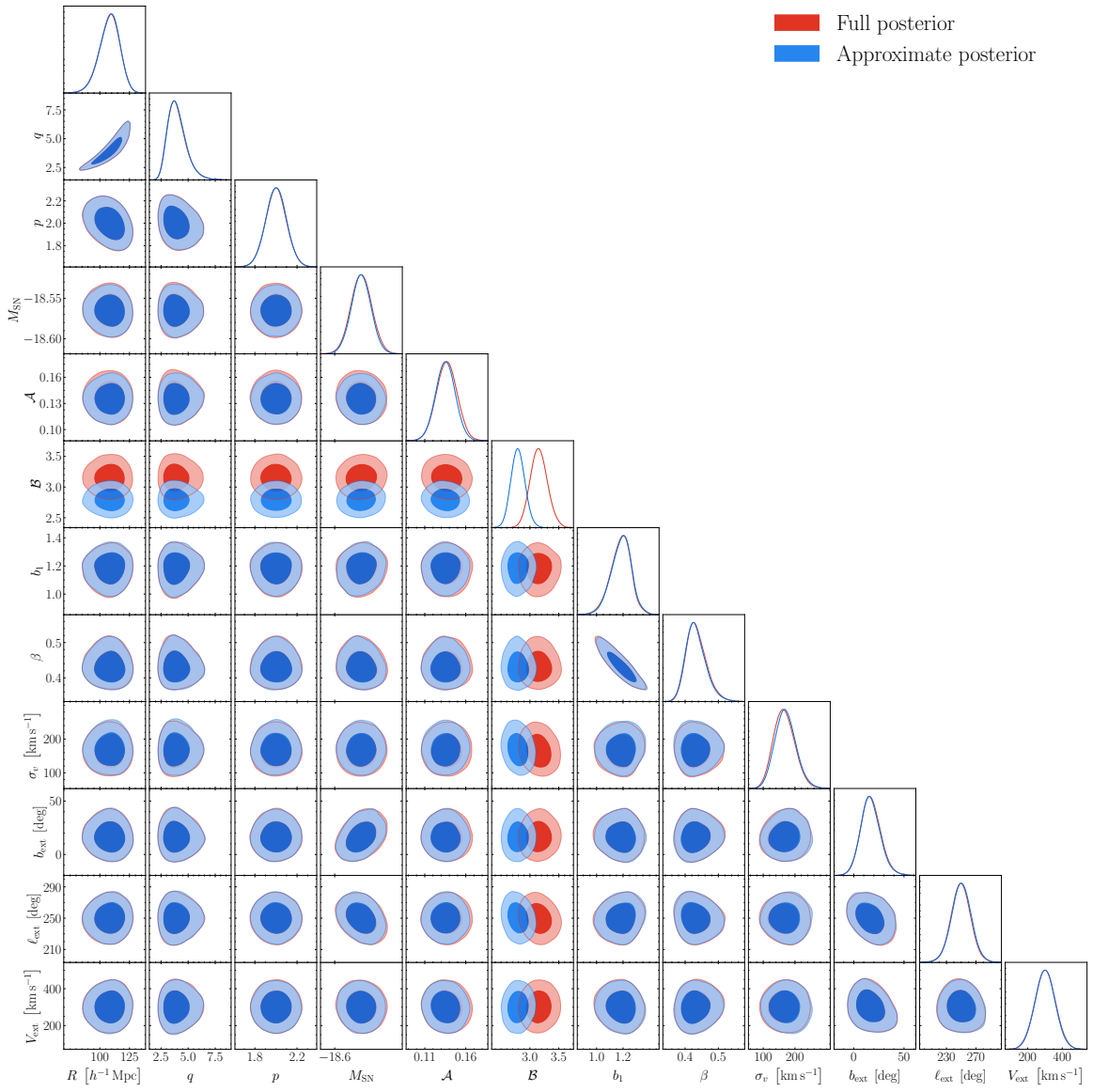


Figure 3.16: Example comparison of the inferred posteriors on the model parameters using the full SN likelihood from Eq. (3.23) and Section 3.A (red) versus the approximate likelihood used for evidence calculation (Section 3.3.3; blue). The underlying reconstruction is Carrick et al. (2015), and the peculiar velocity sample is the Foundation SN sample. The posteriors are nearly identical except for $\mathcal{B}$. Contours represent the 1σ and 2σ credible regions.

4

Measurement of the Hubble constant from Cepheids alone

The posterior predictive check of the previous chapter established the BORG velocity field as the most accurate reconstruction of the local Universe to date. This chapter applies this velocity field to infer H_0 from a supernova-free, two-rung distance ladder built from Cepheids and geometric anchors alone.

Based on Stiskalek et al. (2026e), published in Monthly Notices of the Royal Astronomical Society.

4.1 Introduction

The Hubble tension, a 5.8σ discrepancy between cosmic microwave background (CMB) and distance-ladder determinations of H_0 (Riess et al., 2022a; Breuval et al., 2024), has withstood extensive cross-checks on both sides and remains without a compelling theoretical resolution. A battery of systematic investigations on the SH0ES distance ladder has not identified any single effect that accounts for the discrepancy, consolidating the view that the tension is real.

A key approach for assessing the robustness of the SH0ES inference, and that of the local distance ladder more generally, is to drop or swap out several of the star types used. In par-

ticular, the supernovae (SNe), which in SH0ES extend the Hubble diagram from $z = 0.0233$ to $z = 0.15$, are sometimes considered a weak link: they require standardisation based on light curve and environmental properties, complex dust corrections, absolute-magnitude calibration from lower-rung distance data and elaborate simulation-based selection modelling (e.g. Wojtak & Hjorth 2022, 2024; Gall et al. 2024; Hoyt et al. 2026; Efstathiou 2025; Wojtak & Hjorth 2025). This raises the question of whether H_0 may be inferred (albeit to lower precision) without them. This was in fact the original method of Hubble (1929), and was also employed by Willick & Batra (2001), who measured H_0 from geometry- and Cepheid-based distances—just as we do here. A discrepancy could then imply systematics in the full SH0ES analysis; as this is the only H_0 analysis that exhibits a statistically significant discrepancy with *Planck*, this has the potential to pull the rug out from under the Hubble tension itself.

Kenworthy et al. (2022) (hereafter K22) pursued this approach earlier. They selected 35 galaxies with both Cepheid and SN measurements from the SH0ES project. In SH0ES the Cepheids are used merely to calibrate the SN absolute magnitude (i.e. the redshifts of the hosts are unused); here, by contrast, the SNe are discarded and H_0 is inferred from the Cepheid distances and host galaxy redshifts. The Cepheid Period-Luminosity Relation (CPLR; Leavitt & Pickering 1912b) is calibrated from the first, geometric rung of the distance ladder, namely parallax using *Gaia* within the Milky Way (MW) (Riess et al., 2018, 2021), detached eclipsing binaries within the Large Magellanic Cloud (LMC) (Pietrzyński et al., 2019) and the water maser in NGC 4258 (Reid et al., 2019). At these distances, peculiar velocities are a significant perturbation to the cosmological redshifts and must be modelled. K22 developed and deployed models for these peculiar velocities, as well as possible sample selection effects, to find an overall result $H_0 = 72.9^{+2.4}_{-2.2}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This is in agreement with the SH0ES constraint and in mild, 2.3σ tension with *Planck*, leading the authors to conclude that the SNe in SH0ES are not appreciably altering the best-fit H_0, and therefore are not responsible for the Hubble tension.

The present study is a reanalysis of that data and a reassessment of that conclusion. We construct a Bayesian forward model of the distance ladder for inferring H_0 (and a slew of nuisance parameters) from the Cepheid properties, geometric anchor data and host galaxy redshifts. Our hierarchical treatment is similar to that of Feeney et al. (2018), who built a Bayesian hierarchical model of the then-current local distance ladder of Riess et al. (2016), spanning the geometric anchors, the Cepheids, and the Hubble-flow SNe. They replaced the outlier rejection of Riess et al. (2016) with non-Gaussian likelihoods that model the SN (and, where available, Cepheid) outliers directly, inferring $H_0 = 73.15 \pm 1.78\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and reaffirming the tension with *Planck*. Crucially, however, they did not model sample selection and assumed a uniform-in-distance-modulus prior on the host distances, which is non-physical. We adopt the same hierarchical philosophy but truncate the ladder to its first two rungs, discarding the Hubble-flow SNe. We present a principled Bayesian framework to account for sample selection and demonstrate its impact on the inferred H_0. In addition, we employ state-of-the-art peculiar velocity modelling with the `Manticore-Local` reconstruction. In comparison with K22, we find a lower central value of H_0, and a much reduced error bar. To understand the differences, we highlight several crucial improvements over their analysis.

The structure of this chapter is as follows. Section 4.2 describes the SH0ES data that we use, which is identical to that of K22. Section 4.3 presents our method: the inference framework (Section 4.3.1), peculiar velocity models (Section 4.3.2), and selection function modelling (Section 4.3.3). Section 4.4 details our results: we set the stage with a distance-only analysis not involving redshifts (Section 4.4.1), before moving on to the H_0 inference and its robustness to peculiar velocity modelling and selection effects (Section 4.4.2). Section 4.5.1 discusses the importance of selection function modelling; Section 4.5.2 provides a detailed comparison with the method of K22; Section 4.5.3 discusses the broader implications of our results and suggests further work; and Section 4.6 concludes. Appendix 4.D lists the properties of the Cepheid host galaxies used in our analysis; Appendix 4.A details

the mock tests that we develop to demonstrate that our model is unbiased; Appendix 4.B describes the `Manticore-Local` and Carrick et al. (2015) peculiar velocity reconstructions; and Appendix 4.C summarises the calculation of the Λ-cold dark matter (ΛCDM) peculiar velocity covariance.

4.2 Data

We use data from the SH0ES programme (Riess et al., 2022a), which aims to measure the Hubble constant H_0 using the local distance ladder. The catalogue combines observations of Cepheids in the MW and nearby galaxies, including hosts with both Cepheids and Type Ia SNe.

The CPLR is anchored by MW parallaxes (Riess et al., 2018, 2021), detached eclipsing binaries in the LMC (Pietrzyński et al., 2019), and megamasers in NGC 4258 (Reid et al., 2019). Reid et al. (2019) measure a distance modulus to NGC 4258 of 29.398 ± 0.032 mag, while Pietrzyński et al. (2019) determine a distance modulus to the LMC of 18.477 ± 0.026 mag; both are adopted by SH0ES as geometric anchors for the Cepheid calibration.[1] SH0ES also incorporates 55 Cepheids in M31 observed with the Hubble Space Telescope (HST) using the same three-filter system, as presented by Li et al. (2021). This sample provides precise measurements that constrain the slope of the period–luminosity relation.

Beyond the MW, LMC, M31, and NGC 4258, the SH0ES sample includes 35 galaxies hosting both Cepheids and Type Ia SNe. An additional 254 Type Ia SNe, observed in galaxies without Cepheids, extend the measurement into the Hubble flow, with their calibration tied to the Cepheid-SN hosts. However, in this work we do not use any SH0ES SN data, but instead focus on a two-rung distance ladder involving only geometric anchors and Cepheids. This has the crucial advantage of affording a SN-independent determination of H_0, helping to check for potential systematics in the SN analysis. We now provide a brief overview of

[1] Recently, the Small Magellanic Cloud (SMC) has also been used as a geometric anchor by Breuval et al. (2024).

the geometric anchor and Cepheid distance ladder in SH0ES.

The jth Cepheid in the ith host has apparent magnitude

$$m_{W,ij} = \mu_i + M_W + b_W \log\left(P_{ij}/\hat{P}\right) + Z_W [\mathrm{O/H}]_{ij}, \tag{4.1}$$

where μ_i is the host distance modulus, M_W is the fiducial absolute magnitude of a Cepheid with a period of 10 days and solar metallicity (zero-point), b_W is the period–luminosity slope, P_{ij} is the Cepheid period, $\hat{P} = 10$ days, Z_W is the coefficient of the metallicity correction, and $[\mathrm{O/H}]_{ij}$ is the metallicity at the Cepheid's galactocentric position.[2]

To account for systematic uncertainty in the metallicity, SH0ES propagates the difference between the mean of nine calibrations of strong-line abundance measurements and the scale of Pettini & Pagel (2004), which matches well with direct extragalactic stellar abundances (Bresolin et al., 2016).

This dependence is encoded in the covariance matrix of the apparent magnitudes, which depends on Z_W, whose fiducial value in Riess et al. (2022a) is determined from the joint optimisation of the geometric anchor, Cepheid, and SN components of the SH0ES distance ladder. We adopt the same approach and neglect the dependence of the covariance on Z_W during inference, assuming the same covariance matrix with a fiducial Z_W as Riess et al. (2022a).

An additional contribution to the Cepheid magnitude covariance arises from uncertainties in the sky background estimation due to crowding. This is quantified by injecting artificial stars near each Cepheid at expected magnitudes from trial period–luminosity fits. The discrepancy between input and recovered magnitudes informs a correction to the photometry and yields a background-induced covariance term for Cepheids in the same host

[2]The subscript W denotes the reddening-free Wesenheit magnitude $m_W \equiv m_H - R(m_V - m_I)$; as discussed in Section 1.5.1, this cancellation is exact only for the adopted reddening law, while intrinsic Cepheid colour variation and spatial variation in the extinction law can still contribute residual scatter or population bias. It is more important for the SNe in these dusty calibration galaxies: Wojtak & Hjorth (2024) argue that adopting an extinction law consistent with the Cepheid reddening curve lowers H_0 by 1.7–2.9 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Our SN-independent two-rung ladder avoids this SN extinction systematic entirely.

galaxy. The resulting uncertainty, which reflects correlated photometric biases among Cepheids in the same host, contributes an error floor of 0.03 to 0.06 mag per host. We denote the final Cepheid magnitude covariance matrix as $\mathbf{\Sigma}_{\rm Ceph}$, which combines the contributions from the systematic uncertainties in the metallicity scale, the background-induced photometric biases, and the per-Cepheid uncorrelated covariance terms, including each Cepheid's individual photometric measurement error and the intrinsic astrophysical scatter about the period–luminosity relation set by the finite width of the instability strip. For more details, see Section 2.1 of Riess et al. (2022a).

The SH0ES analysis incorporates an external calibration of the fiducial Cepheid absolute magnitude using MW Cepheids with trigonometric parallax distances. Two samples are used: eight Cepheids with high-precision HST spatial-scan parallaxes (Riess et al., 2018) and 75 Cepheids with *Gaia* EDR3 parallaxes (Riess et al., 2021), both with fluxes measured on the same HST photometric system and with direct spectroscopic metallicities. These samples provide independent constraints on M_W, denoted $M_W^{\rm HST} = -5.804 \pm 0.082$ and $M_W^{\rm Gaia} = -5.903 \pm 0.025$, derived from HST and *Gaia* parallaxes, respectively. In addition to the MW calibration, SH0ES uses Cepheid observations in the LMC to further constrain the luminosity scale. A systematic offset $\Delta_{\rm ZP}$ with uncertainty $\sigma_{\rm grnd} = 0.10$ mag is introduced to account for differences between ground- and space-based photometry in the LMC calibration. The anchor constraints are summarised in Table 4.1.

Unlike the SH0ES analysis, which does not use the redshifts of Cepheid host galaxies but instead infers H_0 from more distant SNe (Riess et al., 2022a), we infer H_0 directly from the redshifts of the Cepheid hosts, without relying on SNe. Specifically, we use the observed redshifts of the 37 galaxies that host both Cepheids and Type Ia SNe, excluding the geometric anchors (LMC, M31, and NGC 4258). These redshifts, converted to the CMB frame, are taken from the Pantheon+ sample, which includes these galaxies as part of its SN compilation (Brout et al., 2022). Similarly to K22, we exclude the two most distant host galaxies out of the 37, as they were targeted with a different HST programme and are

therefore subject to a distinct selection function. Our final sample therefore comprises 35 galaxies, whose coordinates are also listed in Appendix 4.D.

All Cepheid host galaxies are confined to within approximately 40 Mpc (or redshift less than 0.011), indicating the presence of some selection effects in constructing the sample. This selection does not arise from incompleteness in SN detection at these distances. Rather, the selection reflects the fact that only a subset of detected SNe with $m_{\rm SN} \lesssim 14$ mag or $cz \lesssim 3300\,{\rm km\,s^{-1}}$ were selected for HST Cepheid follow-up observations; this selection is consistent with simple random subsampling in either SN magnitude or host redshift below an effective limiting value. We test this consistency by comparing the SN magnitudes and redshifts of the Cepheid host galaxies to the full Pantheon+sample. It is not clear whether this selection is in the SN apparent magnitude, Cepheid apparent magnitude, the host galaxy redshift or their combination, given that Cepheid observations were assembled over many years from independent proposals without a unified selection strategy (K22). We discuss our approach to modelling the selection function, which differs from that of K22, in Section 4.3.3. To model selection in SN apparent magnitude, we use the observed apparent magnitudes of the SNe in the 35 host galaxies, adopting the bias-corrected apparent magnitudes from the SH0ES sample together with the corresponding covariance matrix $\mathbf{\Sigma}_{\rm SN}$.

In Fig. 4.1 we show the distribution of the SN apparent magnitudes for the 35 host galaxies and the distribution of their host observed redshifts. The left panel compares the apparent magnitudes of the SH0ES SNe to those of the full Pantheon+ sample. We assume that Pantheon+ is effectively complete at such small redshifts ($z < 0.011$) down to $m_{\rm SN} = 14$ mag.[3] The distribution of the 35 Cepheid hosts' SN magnitudes, which are below $m_{\rm SN} = 14$ mag, is consistent with Pantheon+ ($p = 0.94$ from a Kolmogorov–Smirnov test),

[3]Type Ia SNe are luminous standard candles ($M_B \approx -19.3$), so a fiducial event at $z = 0.011$ ($cz \approx 3300\,{\rm km\,s^{-1}}$) peaks near $m_{\rm SN} \approx 14$ mag, comfortably brighter than the detection threshold of low-redshift searches such as LOSS/KAIT (Ganeshalingam et al., 2013). This detectability argument does not by itself prove Pantheon+ spectroscopic completeness; it motivates an effective-completeness assumption that we test internally through the low-redshift magnitude and redshift distributions in Fig. 4.1. Along these lines, Tsaprazi et al. (2022), for example, state that the ZTF SN Ia sample (which we do not use) is volume-complete out to $z < 0.036$, well beyond our cut.

which implies that it is plausible that the Cepheid host sample has been selected with a hard limiting estimated magnitude of $m_{\rm SN} = 14$ mag. The same trend can be seen in the right-hand side panel, where we compare the redshift distributions between the Pantheon+ and Cepheid host samples (mutually consistent with $p = 0.60$).

Although the real selection criteria may be complicated and not universal, further support for modelling the selection on the basis of apparent SN magnitude comes from the distributions of magnitude and redshift. Had the sample been selected on either the basis of magnitudes or redshifts but not both, a tail would be visible in the other distribution. However, as shown in Fig. 4.1, the subset of Pantheon+ with SN magnitudes below 14 leads to a sharp truncation in the observed redshift distribution at $3300\,{\rm km\,s^{-1}}$, whereas the subset of Pantheon+ hosts with redshift below $3300\,{\rm km\,s^{-1}}$ shows a small tail in the SN apparent magnitude above 14. We therefore conclude on the basis of current evidence that SN selection is the most likely model. It was also a primary model considered by K22. Both the magnitude and redshift distributions show a small deficit relative to Pantheon+ near the limiting values. However, these deficiencies are not statistically significant if accounting for the expected Poisson uncertainties. Nevertheless, we account for non-trivial selection near the edge by modelling a smooth detection probability function rather than assuming a sharp selection threshold.

Constraint	*Value*	*Reference*
HST Milky Way Cepheid zero-point	-5.804 ± 0.082 mag	Riess et al. (2018)
Gaia Milky Way Cepheid zero-point	-5.903 ± 0.025 mag	Riess et al. (2021)
LMC distance modulus	18.477 ± 0.026 mag	Pietrzyński et al. (2019)
NGC 4258 distance modulus	29.398 ± 0.032 mag	Reid et al. (2019)

Table 4.1: External observational constraints entering the Bayesian hierarchical model. The HST and *Gaia* MW calibrations constrain the Cepheid absolute magnitude zero-point, while the geometric anchors constrain the distances to the anchor galaxies LMC and NGC 4258. We use these measurements as they have been reported by Riess et al. (2022a).

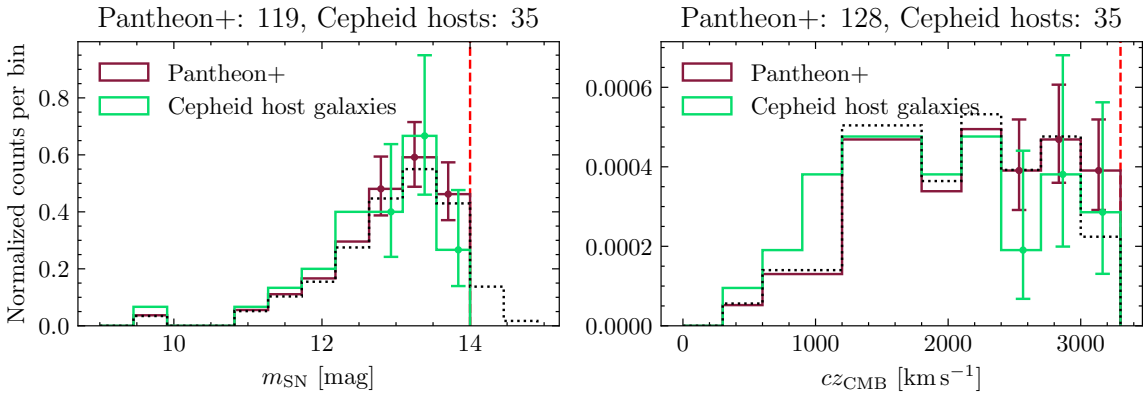


Figure 4.1: Distribution of SN apparent magnitudes for the 35 Cepheid host galaxies and their observed host redshifts either below a limiting SN magnitude of 14 mag or host observed redshift of 3300 km s^{-1} (left and right panels, respectively). Following a Kolmogorov-Smirnov test, the two samples are mutually consistent with $p = 0.94$ and $p = 0.60$. Under the assumption that the Pantheon+ sample is effectively complete within ~40 Mpc, this is consistent with simple random subsampling of Cepheid hosts below a limiting SN apparent magnitude of 14 and a host observed redshift of 3300 km s^{-1}. The error bars are 1σ Poisson counting errors. In addition, the dotted black lines show the magnitude distribution of Pantheon+ hosts with observed redshifts below 3300 km s^{-1} in the left panel, and the redshift distribution of hosts with magnitudes below 14 in the right panel. Within Pantheon+, a truncation in magnitude produces a sharply truncated redshift distribution, whereas a truncation in redshift yields a small tail toward higher magnitudes. We therefore regard SN magnitude selection as marginally more likely and adopt it as our fiducial scenario.

4.3 Methodology

4.3.1 Bayesian forward modelling framework

We now describe our Bayesian hierarchical model. Our approach differs from that of K22 in that, rather than using pre-computed distances from the fiducial SH0ES analysis (Riess et al., 2022a) to predict the observed redshifts of the Cepheid host galaxies in a two-step inference, we forward-model the Cepheid observables directly in a self-contained single-step inference. The data consist of Cepheid magnitudes, periods, and metallicities, and the estimated redshifts of their host galaxies. We also include as observational inputs the MW calibration of the Cepheid absolute magnitude, and the geometric distance moduli of the LMC and NGC 4258. Under the assumption of the SN magnitude selection of the host galaxy sample, we also include these magnitudes as inputs to our model, as detailed in Section 4.3.3.4. The model parameters for this two-rung distance ladder are the Cepheid absolute magnitude M_W, the period-luminosity slope b_W, the metallicity dependence Z_W, and the Hubble constant H_0 as well as the standardized SN magnitude M_B in the case that we model SN selection. Each host galaxy is assigned a latent distance modulus μ_i, providing another 35 parameters to infer. Additional free parameters relate to the peculiar velocity field, for which we consider several models as described in Section 4.3.2. In Section 4.3.3, we describe our approach to modelling selection effects in either SN magnitude, Cepheid magnitude or host galaxy redshift. We illustrate the model with a directed acyclic graph in Fig. 4.2 and summarise the prior distributions of the model's free parameters in Table 4.2.

We sample M_W, b_W, and Z_W from broad uniform priors. Similarly, H_0 is sampled uniformly from 10 to 100 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The MW Cepheid calibration is imposed through two Gaussian likelihood terms, corresponding to the HST and *Gaia* geometric calibration

of the fiducial absolute magnitude:

$$\mathcal{L}(M_W^{\rm HST} \mid M_W) = \mathcal{N}(M_W^{\rm HST}; M_W, \sigma^2_{\rm HST}), \tag{4.2}$$

$$\mathcal{L}(M_W^{\rm Gaia} \mid M_W) = \mathcal{N}(M_W^{\rm Gaia}; M_W, \sigma^2_{\rm Gaia}), \tag{4.3}$$

where $\sigma_{\rm HST}$ and $\sigma_{\rm Gaia}$ are the uncertainties associated with the HST and *Gaia* calibrations, respectively. We impose priors on the distance moduli μ_i, which describe the distances to the Cepheid host galaxies, and are related to luminosity distance $D_{{\rm L},i}$ of the $i^{\rm th}$ host as

$$\mu_i = 5 \log \frac{D_{{\rm L},i}}{\rm Mpc} + 25, \tag{4.4}$$

which is related to comoving distance r as $D_L = (1 + z_{\rm cosmo})r$, where $z_{\rm cosmo}$ is the cosmological redshift. For distances, we choose a prior that is uniform in volume: along a given line-of-sight (LOS), the probability of finding a galaxy increases as the square of the distance due to the volume element. Sampling the distance moduli μ_i implies the prior

$$p(\mu_i) = p(r_i) \left| \frac{{\rm d}r}{{\rm d}\mu} \right|_{\mu_i}, \tag{4.5}$$

where r_i is the physical distance, $p(r_i) \propto r_i^2$ and $|{\rm d}r/{\rm d}\mu|$ is the Jacobian for the transformation from distance to distance modulus. (We later introduce further modification of it, the "inhomogeneous Malmquist bias", to account for the tendency of galaxies to reside in dense environments.) This prior is important, as we discuss further in Section 4.4.1. For the LMC and NGC 4258, we treat the geometric distance calibrations as additional observational constraints with Gaussian likelihoods:

$$\mathcal{L}(\tilde{\mu}_k \mid \mu_k) = \mathcal{N}(\tilde{\mu}_k; \mu_k, \sigma_k^2), \tag{4.6}$$

where k represents either LMC or NGC 4258, μ_k is the sampled distance modulus, $\tilde{\mu}_k$

is the reported measurement from either Pietrzyński et al. (2019) or Reid et al. (2019), respectively, and σ_k is the corresponding measurement uncertainty. The Cepheid apparent magnitudes are predicted according to Eq. (4.1), leading to the likelihood

$$\mathcal{L}(\boldsymbol{m}_W \mid \boldsymbol{m}_W^{\mathrm{pred}}) = \mathcal{N}(\boldsymbol{m}_W; \boldsymbol{m}_W^{\mathrm{pred}}, \boldsymbol{\Sigma}_{\mathrm{Cep}}), \tag{4.7}$$

where the bold font denotes a vector over all 3130 Cepheids.

Up to this point, we have used only the geometric anchors and the CPLR relation to constrain the physical distances to the Cepheid host galaxies. Constraining H_0 requires also including redshift information. The predicted host redshifts decompose as in Eq. (1.38),

$$1 + z_{\mathrm{CMB}}^{\mathrm{pred}} = (1 + z_{\mathrm{cosmo}})(1 + z_{\mathrm{pec}}), \tag{4.8}$$

where z_{cosmo} is computed from the sampled distance modulus assuming flat ΛCDM with $\Omega_{\mathrm{m}} = 0.3$ (the dependence on Ω_{m} is negligible at these redshifts) and $z_{\mathrm{pec}} = V_{\mathrm{pec}}/c$. The redshift likelihood, assuming Gaussian covariance $\boldsymbol{\Sigma}_{\mathrm{cz}}$, is

$$\mathcal{L}(z_{\mathrm{CMB}} \mid z_{\mathrm{CMB}}^{\mathrm{pred}}) = \mathcal{N}(cz_{\mathrm{CMB}}; cz_{\mathrm{CMB}}^{\mathrm{pred}}, \boldsymbol{\Sigma}_{cz}). \tag{4.9}$$

There are additional likelihood and prior terms to consider: the likelihood of the observed sky position and the prior on the true sky position. In conventional astronomy, sky positions are measured sufficiently precisely (unlike in, for example, gravitational-wave astronomy) that the sky-position likelihood is effectively a delta function. This still leaves a prior term, now evaluated at the observed sky position, that must be included in the model. If the model assumes isotropy, this term reduces to a constant $1/4\pi$ and can be dropped. However, if the model is anisotropic, e.g. because the local source number density enters through the inhomogeneous Malmquist bias, the term becomes sky-dependent and must be retained. In practice, one must then work with the prior on the full three-dimensional position rather

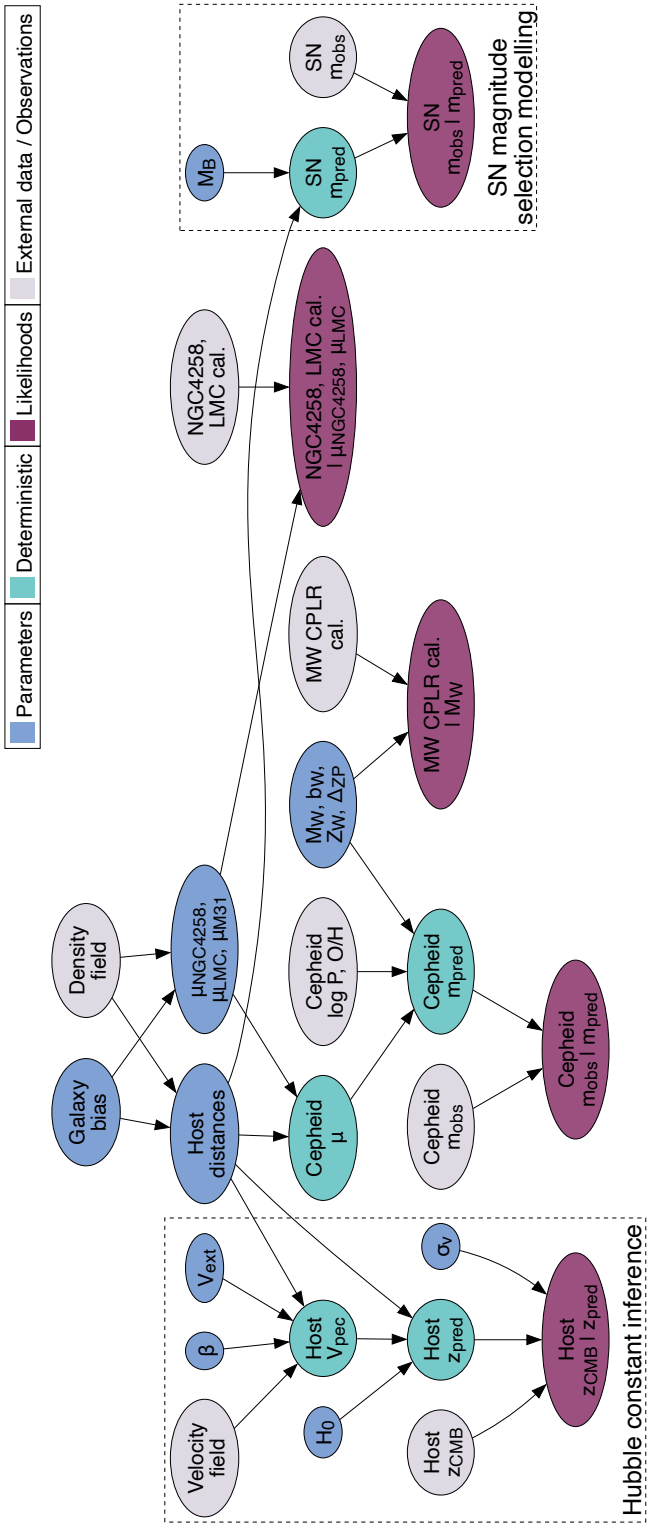


Figure 4.2: Directed acyclic graph of the probabilistic model. The left dashed box delineates the H_0 constraint; excluding it yields the distance-only analysis of Section 4.4.1. The right dashed box marks the SN magnitude selection model, in which M_B is sampled and SN apparent magnitudes are forward-modelled; this step is omitted under redshift selection. See Section 4.3.3 for details.

than simply on the distance. It so happens that the normalisation of this prior, computed over all of space, cancels with the terms arising from the selection function.

We sample the posterior using the NUTS variant of Hamiltonian Monte Carlo (HMC) (Section 1.3.4), running twelve chains of 6000 samples with 1000 burn-in and requiring $\hat{R} - 1 \leq 0.01$.

4.3.2 Peculiar velocity modelling

Since the Cepheid host galaxies are nearby ($cz < 3000\,\mathrm{km\,s^{-1}}$ for the great majority), peculiar velocities are a significant contributor to their redshifts and must be modelled reliably. We adopt a hierarchy of six models, following the same progression as in Chapter 3 (see also Section 1.5.3), with increasing physical realism. In all models σ_v denotes the residual small-scale velocity dispersion, sampled from a reference prior $\pi(\sigma_v) \propto 1/\sigma_v$.

1. **Uncorrelated dispersion.** No coherent flows ($z_{\mathrm{pec}} = 0$ on average); diagonal covariance $\sigma_v^2 + \sigma_{cz}^2$, where σ_{cz} is the redshift measurement uncertainty.

2. **External flow.** As above, plus a constant bulk velocity $\boldsymbol{V}_{\mathrm{ext}}$ with $V_{\mathrm{pec},i} = \boldsymbol{V}_{\mathrm{ext}} \cdot \hat{\boldsymbol{r}}_i$, sampled uniformly in magnitude and direction.

3. **ΛCDM covariance.** Marginalises over all ΛCDM velocity field realisations via the theoretical covariance $\boldsymbol{\Sigma}_{\Lambda\mathrm{CDM}}$ (Section 4.C), with total covariance $\boldsymbol{\Sigma}_{cz} = \boldsymbol{\Sigma}_{\Lambda\mathrm{CDM}} + \mathbf{I}(\sigma_v^2 + \sigma_{cz}^2)$. We show the resulting correlation coefficients in Fig. 4.3.

4. **Scaled ΛCDM covariance.** Introduces a scaling parameter A multiplying $\boldsymbol{\Sigma}_{\Lambda\mathrm{CDM}}$ (Huterer et al., 2015), with a truncated Gaussian prior ($\mu = 1$, $\sigma = 0.5$, $A \geq 0$). The limits $A = 1$ and $A = 0$ recover models 3 and 1, respectively.

5. **C15 reconstruction.** The field-level velocity and density are evaluated at each host's sampled position (Section 4.B.2), with $\beta \equiv f/b_1$ as a free parameter, an additional $\boldsymbol{V}_{\mathrm{ext}}$, and diagonal residual covariance. The local density is used to model

inhomogeneous Malmquist bias. We investigate sub-resolution residual covariance in Section 4.5.2.3.

6. **`Manticore-Local` (fiducial).** Our most realistic model uses the `BORG`-based `Manticore-Local` reconstruction (McAlpine et al., 2025), shown in Chapter 3 to be the most accurate velocity field available. Unlike C15, `Manticore-Local` provides 30 independent posterior samples over which we marginalise, quantifying reconstruction uncertainty (Section 4.B.1). We work in real-space, avoiding triple-valued regions. As in model 5, we evaluate the LOS density and velocity, include $\boldsymbol{V}_{\rm ext}$, and model residuals with constant σ_v (a density-dependent variant, Eq. (4.33), makes negligible difference). This model provides our headline H_0 posterior.

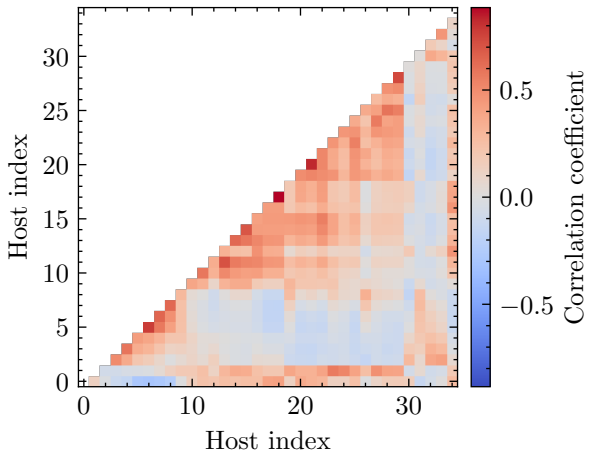


Figure 4.3: Expected peculiar velocity correlation coefficients from the ΛCDM covariance matrix (Section 4.C) for the 35 Cepheid host galaxies. A large fraction are strongly correlated, reducing the effective sample size and motivating the use of a field-level reconstruction.

4.3.3 Selection function modelling

4.3.3.1 General framework

The selection-corrected posterior for n observed sources drawn from a population characterised by $\boldsymbol{\Lambda}$ is derived in Section 1.5.4 (following Kelly et al. 2008). The result, after

Parameter	*Description*	*Prior*
H_0	Hubble constant	$\pi(H_0) = \mathcal{U}(10, 100)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$
Cepheid period–luminosity relation		
M_W	Absolute magnitude zero-point	$\pi(M_W) = \mathcal{U}(-7, -5)$
b_W	period–luminosity slope	$\pi(b_W) = \mathcal{U}(-6, 0)$
Z_W	Metallicity correction	$\pi(Z_W) = \mathcal{U}(-2, 2)$
Δ_{ZP}	Zero-point offset between ground- and space-based LMC photometry	$\pi(\Delta_{\mathrm{ZP}}) = \mathcal{N}(0, 0.1)$
Physical distances		
μ_i	Distance moduli of 35 Cepheid host galaxies, LMC, M31 and NGC 4258	$\pi(\mu_i) = \pi(r_i)\left\lvert\frac{\mathrm{d}r}{\mathrm{d}\mu}\right\rvert_{\mu_i}$ with $\pi(r_i) \propto r_i^2$
α_{low}, α_{high}, $\ln \rho_t$	`Manticore-Local` galaxy bias parameters	$\pi(\alpha_{\mathrm{low}}) = \mathcal{N}(1, 1)$, $\pi(\alpha_{\mathrm{high}}) = \mathcal{N}(0.5, 1)$ (both truncated below at zero) and $\pi(\ln \rho_t) = \mathcal{N}(0, 2)$
Peculiar velocity modelling		
σ_v	Small-scale velocity dispersion	$\pi(\sigma_v) \propto 1/\sigma_v$
$\sigma_{v,\mathrm{low}}$, $\sigma_{v,\mathrm{high}}$, $\ln \rho_{\sigma_v}$, k_{σ_v}	Density-dependent dispersion parameters of Eq. (4.33) (optional)	$\pi(\sigma_{v,\mathrm{low}}) = \pi(\sigma_{v,\mathrm{high}}) = \mathrm{Maxwell}(\mathrm{scale} = 200\,\mathrm{km\,s^{-1}})$, $\pi(\ln \rho_{\sigma_v}) = \mathcal{N}(1, 5)$, $\pi(k_{\sigma_v}) = \mathcal{N}(1, 1)$ truncated at $k_{\sigma_v} \geq 0$
$\boldsymbol{V}_{\mathrm{ext}}$	Constant (external) flow vector	$\pi(\lvert\boldsymbol{V}_{\mathrm{ext}}\rvert) = \mathcal{U}(0, 1000)\,\mathrm{km\,s^{-1}}$ and uniform in sky direction
A	Scaling of the ΛCDM velocity covariance matrix	$\pi(A) = \mathcal{N}(1, 0.5)$ truncated at $A \geq 0$
β	Carrick et al. (2015) velocity field scaling parameter	$\pi(\beta) = \mathcal{N}(0.43, 0.02)$
Selection function modelling		
M_B	Standardized SN absolute magnitude	$\pi(M_B) = \mathcal{U}(-22, -18)$

Table 4.2: Free parameters of the Bayesian hierarchical model described in Section 4.3, with their corresponding priors. The inclusion of specific peculiar velocity parameters depends on the chosen velocity modelling approach; see Section 4.3.2. $\mathcal{U}(a, b)$ denotes a uniform prior over the interval $[a, b]$, while $\mathcal{N}(\mu, \sigma^2)$ denotes a normal prior with mean μ and variance σ^2. The width of the Δ_{ZP} prior follows Riess et al. (2022a).

marginalising over the unknown total population size N under a log-uniform prior, is

$$\mathcal{P}(\boldsymbol{\Lambda} \mid \boldsymbol{d}_{\rm obs}) \propto \pi(\boldsymbol{\Lambda})\,[p(S=1 \mid \boldsymbol{\Lambda})]^{-n} \prod_{i \in \mathcal{A}_{\rm obs}} \mathcal{L}(\boldsymbol{d}_i \mid \boldsymbol{\Lambda},\, \hat{u}_i), \tag{4.10}$$

where $p(S=1 \mid \boldsymbol{\Lambda})$ is the detection probability averaged over all source parameters and sky positions. For a smooth selection threshold at $d_{\rm lim}$ with width $\sigma_{\rm sel}$, the detection probability at a given sky position reduces to

$$p(S=1 \mid \boldsymbol{\Lambda}) = \iint \mathrm{d}\hat{u}\,\mathrm{d}\boldsymbol{\theta}\,\Phi\left(\frac{d_{\rm lim} - d_{\rm pred}}{\sqrt{\sigma_{\rm sel}^2 + \sigma_d^2}}\right)\pi(\boldsymbol{\theta} \mid \boldsymbol{\Lambda},\, \hat{u})\,\pi(\hat{u}), \tag{4.11}$$

where $d_{\rm pred}$ is the predicted observable (e.g. apparent magnitude or redshift), σ_d is its measurement uncertainty, $\boldsymbol{\theta}$ denotes the source parameters (e.g. distance), and Φ is the standard normal cumulative distribution function (CDF). Marginalising the smooth selection over the Gaussian measurement uncertainty convolves the selection CDF with a Gaussian density, which again yields a CDF,

$$\int \Phi\left(\frac{d_{\rm lim} - x}{\sigma_{\rm sel}}\right)\mathcal{N}(x;\, d_{\rm pred},\, \sigma_d^2)\,\mathrm{d}x = \Phi\left(\frac{d_{\rm lim} - d_{\rm pred}}{\sqrt{\sigma_{\rm sel}^2 + \sigma_d^2}}\right), \tag{4.12}$$

which is why $\sigma_{\rm sel}^2$ and σ_d^2 enter in quadrature. We now apply this framework separately to SN magnitude-selected and host-galaxy redshift-selected samples, and verify on mock data that the H_0 inference is unbiased (Section 4.A).

4.3.3.2 Supernova magnitude selection

For SN apparent magnitude selection we have that

$$p(S=1 \mid M_B) = \iint \mathrm{d}\hat{u}\mathrm{d}r\,\Phi\left(\frac{m_{\rm SN}^{\rm lim} - m_{\rm SN}^{\rm pred}}{\sqrt{\tilde{\sigma}_{\rm SN}^2 + \sigma_{\rm SN}^2}}\right)\pi(r \mid \hat{u})\pi(\hat{u}), \tag{4.13}$$

where $m_{\rm SN}^{\rm lim}$ is the SN apparent magnitude threshold,

$$m_{\rm SN}^{\rm pred} = \mu(r) + M_B \tag{4.14}$$

is the predicted apparent magnitude of a SN at distance r with absolute magnitude M_B, $\tilde{\sigma}_{\rm SN}$ is the selection smoothness term, and $\sigma_{\rm SN}$ is the uncertainty in SN apparent magnitude. We set $\tilde{\sigma}_{\rm SN}$ to be 0.15 mag and $\sigma_{\rm SN}$ to the average square root of the diagonal of the SN covariance (approximately 0.13 mag), though the inference is not very sensitive to either choice.

Without considering the inhomogeneous Malmquist bias, the prior on distance is given by $\pi(r) \propto r^2$ and is independent of $\hat{u}$. At small physical distances, the distance modulus is related to distance (expressed in Mpc) as $\mu \approx 5\log(r/1\ \text{Mpc}) + 25$, and the detection probability simplifies to

$$p(S = 1 \mid M_B) \propto 10^{-3M_B/5}. \tag{4.15}$$

This result is independent of $\sigma_{\rm SN}$. In the expression above we wrote the dependence as $10^{-3M_B/5}$. More precisely, the relevant factor is $10^{3(M_B - m_{\rm SN}^{\rm lim})/5}$, which depends on the difference between M_B and the limiting magnitude $m_{\rm SN}^{\rm lim}$. Since the posterior dependence on $m_{\rm SN}^{\rm lim}$ is separable, it can be marginalised out and absorbed into the overall proportionality constant together with $\sigma_{\rm SN}$. We can also consider the case when the selection is a smooth function of the magnitude. Going back to Eq. (4.11), we have that

$$p(S = 1 \mid M_B) = \int \mathrm{d}r\, \mathrm{d}m_{\rm SN}^{\rm pred}\, p(S{=}1 \mid m_{\rm SN}^{\rm pred}) \mathcal{L}(m_{\rm SN}^{\rm pred} \mid M_B,\, r)\, \pi(r). \tag{4.16}$$

Through a similar change of variables we can pull out the dependence on M_B outside of the integral, so that $p(S = 1 \mid M_B) \propto 10^{-3M_B/5}$ and the shape of the selection function only changes the normalisation constant such that the posterior is independent of it.

However, because the inhomogeneous Malmquist bias makes the distance prior depen-

dent on $\hat{u}$, the integral over $\hat{u}$ no longer vanishes and must be evaluated explicitly. We model the galaxy bias using the LOS density from the reconstruction and marginalise by drawing 1000 random LOS, computing the selection probability along each, and averaging the results. We verify that this number of random LOS is sufficient to ensure independence of the results on the sampling.

When accounting for the selection in SN apparent magnitude, the inference is no longer independent of the SNe and depends explicitly on M_B. The intrinsically brighter the SNe, the more of them pass the selection threshold and are observed at greater distances. This behaviour holds regardless of SNe being observed at much greater distances than those considered here; only the host galaxies included in the sample of 35 matter. Moreover, their absolute magnitudes must be modelled self-consistently with the inferred distances to the Cepheid host galaxies and with the SN apparent magnitudes used in the selection. We treat M_B as a free parameter and sample it explicitly, and therefore we also forward-model the apparent magnitudes of SNe in Cepheid host galaxies with the likelihood term to constrain it,

$$\mathcal{L}(\boldsymbol{m}_{\rm SN} \mid \boldsymbol{m}_{\rm SN}^{\rm pred}) = \mathcal{N}(\boldsymbol{m}_{\rm SN}; \boldsymbol{m}_{\rm SN}^{\rm pred}, \boldsymbol{\Sigma}_{\rm SN}), \tag{4.17}$$

where $\boldsymbol{m}_{\rm SN}$ are the bias-corrected SN apparent magnitudes from the SH0ES sample, $\boldsymbol{m}_{\rm SN}^{\rm pred}$ are the predicted SN apparent magnitudes, and $\boldsymbol{\Sigma}_{\rm SN}$ is the covariance matrix of the SN apparent magnitudes. Some of the 35 Cepheid host galaxies contain multiple SNe; however, in such cases we select only the brightest SN per host. In deriving Eq. (4.13), we assumed the SN observations are independent, which is not valid due to correlations introduced by the standardisation procedure (i.e. non-zero off-diagonal elements in the covariance matrix). In contrast, the corresponding likelihood in Eq. (4.17) uses the covariance matrix. We verify that neither the downsampling of SNe to one per galaxy nor setting all the off-diagonal elements of the SN magnitude covariance matrix to zero affects the H_0 posterior distribution.

4.3.3.3 Redshift-limited selection

For a redshift-limited survey,

$$p(S = 1 \mid H_0) = \iint \mathrm{d}\hat{u}\, \mathrm{d}r\, \Phi\left(\frac{cz^{\mathrm{lim}} - cz^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_v^2 + \sigma_v^2}}\right) \pi(r \mid \hat{u})\, \pi(\hat{u}), \tag{4.18}$$

where z^{pred} is the predicted redshift of a galaxy at distance r and sky direction $\hat{u}$, and $\tilde{\sigma}_v$ is the redshift truncation width, assumed to be 300 km s^{-1}, though the inference is not sensitive to this choice. For illustration, if peculiar velocities and inhomogeneous Malmquist bias are neglected (so that the integrand is independent of $\hat{u}$) and σ_v is fixed, then at low redshift, where $cz_{\mathrm{CMB}} \approx H_0 r$, the detection probability for a redshift-selected survey simplifies to

$$p(S = 1 \mid H_0) \propto H_0^{-3}, \tag{4.19}$$

per host (and these probabilities multiply, yielding an effective H_0^{+105} dependence, which can shift the peak significantly), with the redshift cut-off absorbed into the proportionality constant. In practice, as before, the inhomogeneous Malmquist bias and peculiar velocities introduce a dependence on $\hat{u}$, requiring the selection probability to be marginalised over all plausible lines of sight. We follow the same procedure as previously, but now also use the LOS peculiar velocities from the reconstruction.

As with the SN magnitude selection, the detection probability in Eq. (4.18) assumes mutually independent source redshifts, corresponding to a diagonal covariance matrix with entries of σ_v^2. This assumption holds for `Manticore-Local` where we take each realisation (without covariance) independently and combine them self-consistently, and also for C15 which provides a single realization. We assume any residual correlations to be negligible. However, it fails when adopting the ΛCDM peculiar velocity covariance matrix. As demonstrated in Fig. 4.3, the covariance matrix for the 35 host galaxies exhibits significant off-diagonal correlations.

To "correct" for the loss of information due to correlated samples, we estimate the effective number of samples: given a generic covariance matrix $\mathbf{C}$ of dimension $N \times N$ we can calculate the effective rank as the exponential of the Shannon entropy of its normalised eigenvalues,

$$\mathrm{N}_{\mathrm{eff}}(\mathbf{C}) = \exp\left[-\sum_{i=1}^{N} p_i \ln p_i\right], \tag{4.20}$$

where $p_i = \lambda_i / \sum_{j=1}^{N} \lambda_j$, and λ_i are the eigenvalues of $\mathbf{C}$ (Roy & Vetterli, 2007). $\mathrm{N}_{\mathrm{eff}}(\mathbf{C})$ provides a continuous measure of the matrix's effective dimensionality, with $\mathrm{N}_{\mathrm{eff}}(\mathbf{C}) = N$ for a diagonal covariance matrix. This produces $N_{\mathrm{eff}} = 21.0$ for the ΛCDM peculiar velocity covariance matrix for the 35 host galaxies. Accordingly, we down-weight the contribution of Eq. (4.18) in the model likelihood by a factor of $N_{\mathrm{eff}}/N \approx 0.57$ for that peculiar velocity model. However, we stress that this correction is only approximate. Moreover, no such correction is required for either the SN or Cepheid magnitude selection; the resulting constraints on H_0 are unchanged when their respective covariances are diagonalised in the likelihood.

4.3.3.4 Joint supernova magnitude- and redshift-limited selection

Although not a fiducial model, we also consider the case of selection on *both* SN host magnitude and galaxy redshift. In this case, the selection probability is given by

$$p(S = 1 \mid H_0, M_B) = \iint \mathrm{d}\hat{u}\, \mathrm{d}r\, \Phi\left(\frac{m_{\mathrm{SN}}^{\mathrm{lim}} - m_{\mathrm{SN}}^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_{\mathrm{SN}}^2 + \sigma_{\mathrm{SN}}^2}}\right) \Phi\left(\frac{cz^{\mathrm{lim}} - cz^{\mathrm{pred}}}{\sqrt{\tilde{\sigma}_v^2 + \sigma_v^2}}\right) \pi(r \mid \hat{u})\, \pi(\hat{u}), \tag{4.21}$$

which is computed analogously to the previous cases, except that we have a product of the CDFs. For this selection, when modelling the peculiar velocities with the ΛCDM covariance (but crucially not with either `Manticore-Local` or C15), we use the same correction for the number of effective samples as in Section 4.3.3.3.

4.4 Results

As a preliminary to our H_0 inference, in Section 4.4.1 we infer the host galaxy distances without using their redshifts. This enables us to quantify the effect of priors and selection modelling on the inferred distances. Then, in Section 4.4.2 we add the redshift information to constrain H_0 with the full model (i.e. including the part in the dashed box of Fig. 4.2). In Table 4.4 we provide the inferred Cepheid host distances and the peculiar velocities from both C15 and `Manticore-Local`.

4.4.1 Host distances

We now consider only the Cepheid host distances, calibrated using the CPLR and two geometric anchors (LMC and NGC 4258). We examine four modelling scenarios:

1. a uniform prior in distance modulus without selection modelling (a baseline reference);
2. a uniform prior in volume without selection modelling;
3. a uniform prior in volume with modelling of SN magnitude selection;
4. a uniform prior in volume with modelling of Cepheid magnitude selection.

The uniform-in-volume prior, reflecting the fact that the Universe is three-dimensional, yields unbiased distance-ladder inference. In contrast, the uniform-in-distance-modulus prior has no physical motivation and is implicitly adopted in K22 and the SH0ES analysis (Riess et al., 2022a) where maximum-likelihood parameters are estimated by minimising a joint χ^2 statistic across the distance ladder including the host galaxy distance moduli as free parameters. The uniform-in-volume prior follows $p(r_i) \propto r_i^2$, while the uniform-in-distance-modulus prior corresponds to $p(r_i) \propto 1/r_i$. Since the uniform-in-distance-modulus prior biases the inferred distances low relative to a uniform-in-volume

prior, for fixed redshifts it biases H_0 high. This is discussed in more detail and generality in Desmond et al. (2026). Fig. 4.4 shows the inferred distance moduli to the LMC and NGC 4258 along with the Cepheid zero-point M_W. We compare our results to the SH0ES analysis (Riess et al., 2022a). The no-selection, uniform-in-distance-modulus prior (red contours) produces results in excellent agreement with the SH0ES analysis, as expected. In contrast, the no-selection, uniform-in-volume prior (blue contours) yields larger inferred distances and a brighter Cepheid absolute magnitude zero-point. However, interestingly, incorporating either the SN or Cepheid selection modelling with a uniform-in-volume prior shifts the inferred distances to the LMC and NGC 4258 and Cepheid zero-point back to values identical with those obtained under a uniform-in-distance-modulus prior without selection modelling (see Fig. 4.4).

We verify that in any case, the inferred Cepheid-zero-point M_W is not in any tension with the MW calibration, which was used as a constraint in the model. Similarly, the inferred distance to the LMC and NGC 4258 remain consistent with the imposed external calibration. Moreover, Fig. 4.4 also shows that, in the absence of selection modelling, the choice of prior shifts the distance to NGC 4258 but leaves the distance to the LMC unchanged. This suggests that the MW Cepheid zero-point calibration, with precise parallax distances, is unaffected by the prior choice.

In Fig. 4.5 we extend the comparison of inferred distances to all 35 Cepheid host galaxies. In the absence of selection modelling, we find that the uniform-in-volume prior yields, on average, a distance modulus larger by 0.045 ± 0.15 mag compared to the uniform-in-distance-modulus prior. Thus, using a uniform-in-distance-modulus prior instead of uniform-in-volume potentially biases the inferred value of H_0 high by $1.79^{+1.00}_{-0.42}$ per cent. On the other hand, a uniform-in-volume prior with SN selection modelled yields distance moduli that are not systematically offset from those obtained with a uniform-in-distance-modulus prior without selection modelling; the mean difference is -0.0049 ± 0.13 mag. Finally, the uniform-in-volume prior with Cepheid magnitude selection yields distances that

are systematically offset from those inferred with a uniform-in-distance-modulus prior by 0.022 ± 0.15 mag, approximately half the shift observed when selection effects are neglected. From this point onward, we adopt the physically motivated uniform-in-volume prior. We note that in Section 4.4.2, when inferring H_0 using either `Manticore-Local` or C15, we apply this prior to host galaxy distances and account for inhomogeneous Malmquist bias due to source density fluctuations induced by underlying density fluctuations.

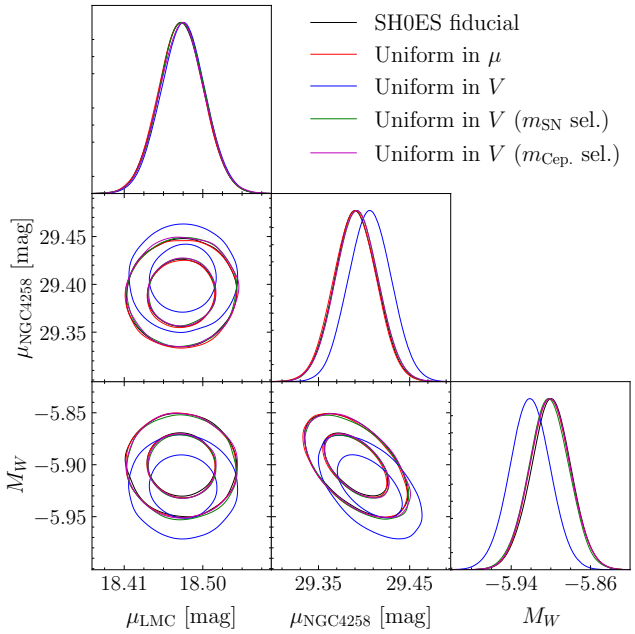


Figure 4.4: Corner plot of the inferred distance moduli to the LMC, NGC 4258, and the CPLR zero-point calibration, from analyses that do not incorporate redshift information. We compare four scenarios: a uniform-in-volume prior on host galaxy distances without selection modelling (blue), a uniform-in-distance-modulus prior without selection modelling (red), a uniform-in-volume prior with SN magnitude selection modelled (green), and a uniform-in-volume prior with Cepheid magnitude selection (violet). The uniform-in-distance-modulus prior agrees with the SH0ES analysis (Riess et al., 2022a), which implicitly assumes this prior (black). The uniform-in-volume posteriors with either SN or Cepheid magnitude selection are in excellent agreement with the SH0ES calibration. Contours denote 1σ and 2σ confidence intervals.

4.4.2 Hubble constant inference

To set the stage for our inference of H_0, and visualize the data, we present in Fig. 4.6 the Hubble diagram showing the relation between the observed redshifts of the 35 host galaxies converted to the CMB frame and their inferred distance moduli under the uniform-in-volume prior and the SN magnitude selection modelling from the previous subsection.

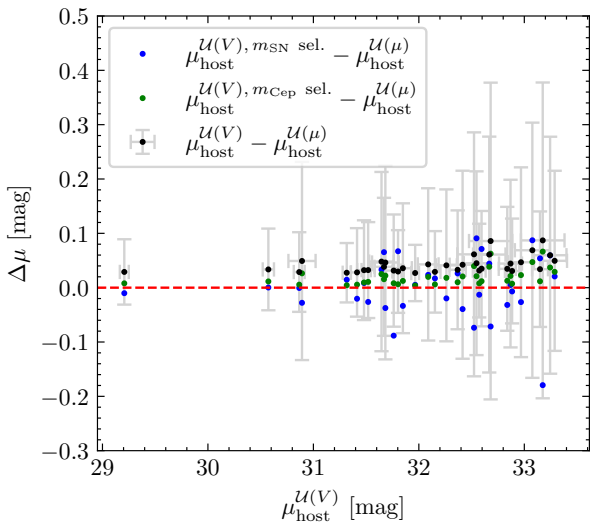


Figure 4.5: Comparison of distances to the 35 Cepheid host galaxies between a Cepheid-only distance inference using a uniform-in-volume prior without any selection (black), with SN magnitude selection (blue), or with Cepheid magnitude selection (green), relative to a uniform-in-distance-modulus prior without selection modelling. The first yields an average shift of 0.045 mag (H_0 biased high by $1.79^{+1.00}_{-0.42}$ per cent), the second yields no systematic offset, and the third yields 0.022 mag. Error bars are 1σ.

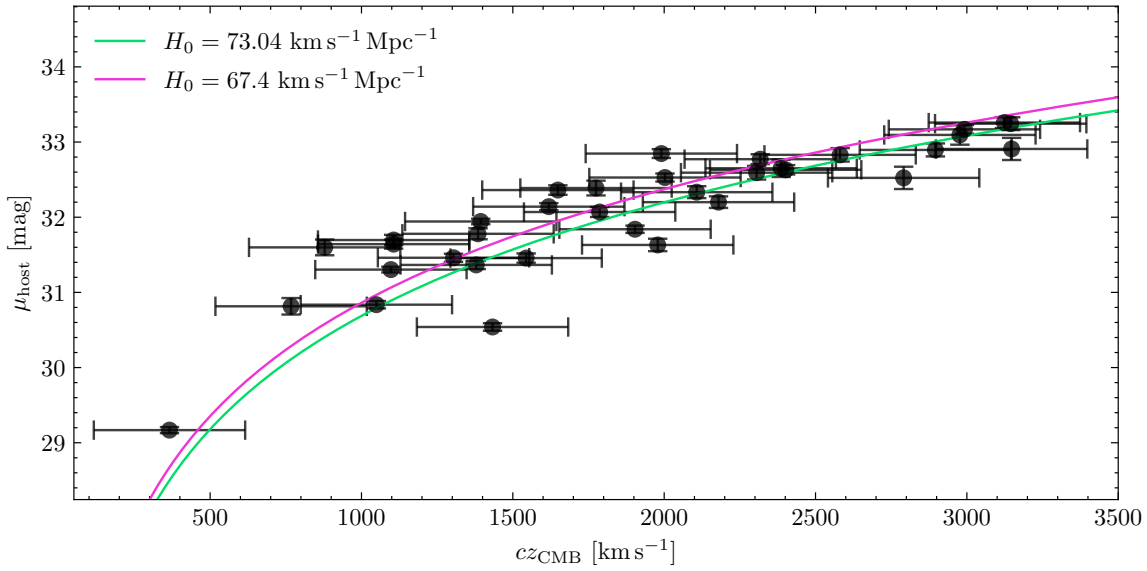


Figure 4.6: The Hubble diagram for the 35 Cepheid host galaxies, showing the relation between CMB-frame redshifts $z_{\rm CMB}$ and the inferred distance moduli from our analysis assuming a uniform-in-volume prior and a selection in SN apparent magnitude. The red curve indicates the predicted distance modulus–redshift relation for the SH0ES best-fit value of $H_0 = 73.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Riess et al., 2022a), while the blue curve shows the relation for the *Planck* value $H_0 = 67.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration et al., 2020b). For plotting purposes, we assume a redshift uncertainty corresponding to a velocity dispersion of $250\,\mathrm{km\,s^{-1}}$ though the exact value would depend on peculiar velocity modelling. The error bars represent 1σ uncertainties.

The red and blue lines indicate the predictions for the best-fit SH0ES and *Planck* values of H_0, respectively.

We now use the full forward model (see Fig. 4.2) to infer H_0 as well as the distances by folding in the Cepheid host galaxy redshift information. In Table 4.3, we tabulate the inferred values of H_0 for various peculiar velocity and selection function models considered in Section 4.3, all assuming a uniform-in-volume prior on the host galaxy distances. Figure 4.8 compares the inferred H_0 for the various selection models while modelling peculiar velocities with `Manticore-Local`. We present a stacked H_0 comparison in Fig. 4.10.

4.4.2.1 No selection

As the baseline case we completely omit selection modelling and assume a uniform-in-volume distance prior, although we note that this cannot be a realistic model as the sample is confined within approximately 40 Mpc. Assuming the underlying population of sources is indeed distributed uniformly in volume, this implies selection based on a distance-dependent observable which should be either redshift or apparent magnitude. We find $H_0 = 68.9 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ using the C15 reconstruction and $H_0 = 71.1 \pm 1.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with `Manticore-Local`. The values are lower than with any selection modelling, as described below.

4.4.2.2 Supernova magnitude selection

We now consider in more detail the case where the sample is assumed to be selected by SN apparent magnitude solely (described in Section 4.3.3.2). As outlined in Section 4.3.2, our simplest models assume uncorrelated peculiar velocities, either setting them to zero on average, or modelling them with a constant velocity vector $\boldsymbol{V}_{\mathrm{ext}}$ that is sampled jointly. The redshift scatter is described by a free parameter σ_v that is also sampled. In these cases, we find $H_0 = 68.8 \pm 2.0$ and $69.3 \pm 1.7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, respectively. However, these results should be interpreted with caution, as they rely on highly simplified treatments of peculiar

velocities. An alternative and more conservative approach adopts the ΛCDM peculiar velocity covariance matrix (described in Section 4.C). This effectively marginalises over all plausible peculiar velocity realisations consistent with the ΛCDM power spectrum. It yields $H_0 = 70.1 \pm 3.0$, or $70.1 \pm 3.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when including a global scaling factor A for the covariance matrix, with it constrained to 1.2 ± 0.3.

We now account for peculiar velocities using local Universe reconstructions. First, we use the C15 reconstruction, with $\boldsymbol{V}_{\mathrm{ext}}$ to capture large-scale flows sourced by structure outside the reconstructed volume. Smaller-scale motions are treated as uncorrelated and modelled with a velocity dispersion σ_v. As shown in Fig. 4.8, this model yields $H_0 = 72.1 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. When using C15, we find $\sigma_v = 215 \pm 31\,\mathrm{km\,s^{-1}}$, $|\boldsymbol{V}_{\mathrm{ext}}| = 199 \pm 82\,\mathrm{km\,s^{-1}}$ and $\beta = 0.42 \pm 0.02$ (sampled with a Gaussian prior with mean 0.43 and standard deviation 0.02, following C15). To account for SN magnitude selection, we infer and marginalise over the standardised SN absolute magnitude M_B. We verify that, when no selection is modelled and only the SN magnitude likelihood of Eq. (4.17) is included, we obtain $H_0 = 69.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, in good agreement with the no-selection inference without the SNe, which yields $H_0 = 68.9 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Thus the modification to H_0 is driven by the selection modelling, not by the additional SN likelihood.

Second, we use the `Manticore-Local` reconstruction, finding $H_0 = 71.7 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In this setup, the velocity field scaling parameter β is fixed to unity (since `Manticore-Local` is a dark matter resimulation assuming a fixed cosmology) and we find that $\sigma_v = 145 \pm 30\,\mathrm{km\,s^{-1}}$ and $|\boldsymbol{V}_{\mathrm{ext}}| = 223 \pm 64\,\mathrm{km\,s^{-1}}$. In Fig. 4.9, we show the posterior distribution of all model parameters besides the 35 host galaxy distances. In particular, the posterior on M_B agrees well with the reported value of $M_B = -19.253$ in Fig. 14 of Riess et al. (2022a). Letting σ_v follow the density-dependent form of Eq. (4.33) yields $H_0 = 71.6 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, almost unchanged from the fiducial value. We also consider the case where β is allowed to vary, imposing a Gaussian prior with mean 1 and standard deviation 0.5, loosely corresponding to varying the assumed value of σ_8. This produces

$H_0 = 71.5 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with $\beta = 0.81 \pm 0.17$, which is in good agreement with the $\beta = 1$ inference and β itself is consistent with unity at $\sim 1\sigma$. Using `Manticore-Local` yields the smallest uncertainty on H_0 due to its most precise (and accurate; Stiskalek et al. 2026c; McAlpine et al. 2025) determination of the peculiar velocity field. The smaller inferred σ_v than C15 also indicates that it more successfully captures local flows.

4.4.2.3 Host galaxy redshift selection

We now consider the case in which the sample of host galaxies is assumed to be selected solely based on observed redshift, as described in Section 4.3.3.3. Our two simplest models, both using a diagonal covariance matrix with a free parameter σ_v, assume either zero-mean peculiar velocities or model them with a constant velocity vector $\boldsymbol{V}_{\rm ext}$. These yield $H_0 = 73.9 \pm 2.7$ and $73.2 \pm 2.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, respectively. In contrast, assuming zero-mean peculiar velocities but including the ΛCDM peculiar velocity covariance matrix, we find $H_0 = 76.1 \pm 3.1$, or $78.5 \pm 4.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when allowing for a global scaling factor A, which is found to be 1.4 ± 0.3. However, these results require an approximate correction to the selection function to account for the effective number of independent samples (see Section 4.3.3.3), so the results obtained using the ΛCDM covariance should be interpreted with caution. No such correction is necessary when peculiar velocities are explicitly modelled and the covariance is diagonal.

Then, we again consider our two local Universe reconstruction models. Using the C15 reconstruction, we find $H_0 = 74.4 \pm 1.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, with $\beta = 0.42 \pm 0.02$, which is in good agreement with SH0ES. In contrast, using the `Manticore-Local` reconstruction yields $H_0 = 73.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Fig. 4.9 presents the posterior distribution assuming `Manticore-Local` and redshift-limited selection, showing that the inferred $\boldsymbol{V}_{\rm ext}$ and σ_v are consistent with the SN selection case. Applying the density-dependent σ_v model of Eq. (4.33) gives $H_0 = 72.9 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. If instead of fixing β to unity for `Manticore-Local` we sample it with a Gaussian prior of mean of unity and standard

deviation 0.5, we obtain consistent values of $H_0 = 72.7 \pm 1.4$ km s^{-1} Mpc^{-1} and $\beta = 0.87 \pm 0.18$.

4.4.2.4 Joint magnitude and redshift selection

For completeness, we briefly consider joint selection scenarios. If the entire host sample were selected on *both* SN apparent magnitude *and* observed redshift (see Section 4.3.3.4), then Eq. (4.13) is modified so that the integrand includes the selection in both the magnitude and redshift limits, Eq. (4.21). Such a joint selection yields H_0 values intermediate between those from SN magnitude–only and host redshift–only selections. Using C15 gives $H_0 = 73.3 \pm 1.7$ km s^{-1} Mpc^{-1} with $\beta = 0.42 \pm 0.02$, while `Manticore-Local` yields $H_0 = 72.3 \pm 1.4$ km s^{-1} Mpc^{-1}. Adopting the density-dependent σ_v of Eq. (4.33) yields $H_0 = 72.1 \pm 1.3$ km s^{-1} Mpc^{-1}. Finally, allowing β to vary for `Manticore-Local` with a Gaussian prior of $\mathcal{N}(1, 0.5)$, we find $H_0 = 72.0 \pm 1.4$ km s^{-1} Mpc^{-1} and $\beta = 0.86 \pm 0.18$, indicating no significant effect on H_0 from fixing β to unity, i.e. the assumed cosmology of `Manticore-Local`.

Alternatively, the sample could be interpreted as a concatenation of two subsamples: one selected purely on SN magnitude and the other purely on host redshift. Fig. 4.7 shows the mean inferred H_0 as a function of the number of galaxies selected by SN magnitude, varying this number from zero to 35. The left bound corresponds to purely redshift-based selection and the right bound to purely SN-magnitude-based selection. We find an approximately linear scaling between these two limiting cases, with an average H_0 uncertainty very close to $\sim$1.4 km s^{-1} Mpc^{-1} in all cases.

4.5 Discussion

In this section we discuss the impact of selection function modelling on the inferred H_0 (Section 4.5.1) and then compare our results with those of K22 (Section 4.5.2).

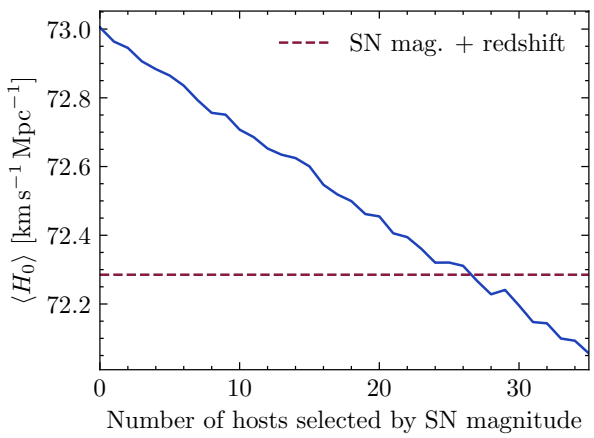


Figure 4.7: Mean inferred H_0 using `Manticore-Local` as a function of the number of hosts (out of 35) assumed to be selected by SN magnitude, with the remainder selected by redshift. The left bound is purely redshift-based selection; the right is purely SN magnitude-based. The dashed line shows the result assuming the entire sample is selected on both.

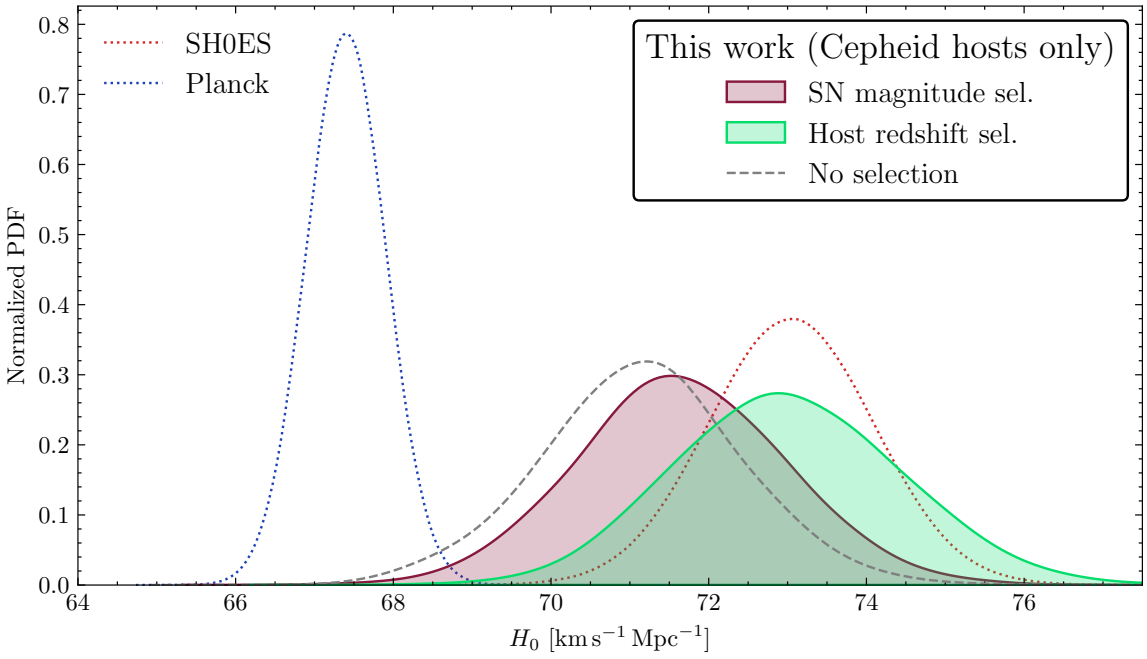


Figure 4.8: Posterior distributions of H_0 for the two main selection scenarios—SN magnitude selection (purple) and host galaxy redshift selection (green)—when accounting for peculiar velocities using `Manticore-Local` (McAlpine et al., 2025), compared to *Planck* ($67.4 \pm 0.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; Planck Collaboration et al. 2020b) and SH0ES ($73.0 \pm 1.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; Riess et al. 2022a). For completeness, we also show the H_0 posterior without modelling any selection effects, which leads to a biased inference. These results agree well with SH0ES, though they peak slightly lower except under redshift selection. Model variations are listed in Table 4.3, and Fig. 4.10 compares our results using both `Manticore-Local` and Carrick et al. (2015) with those from the literature.

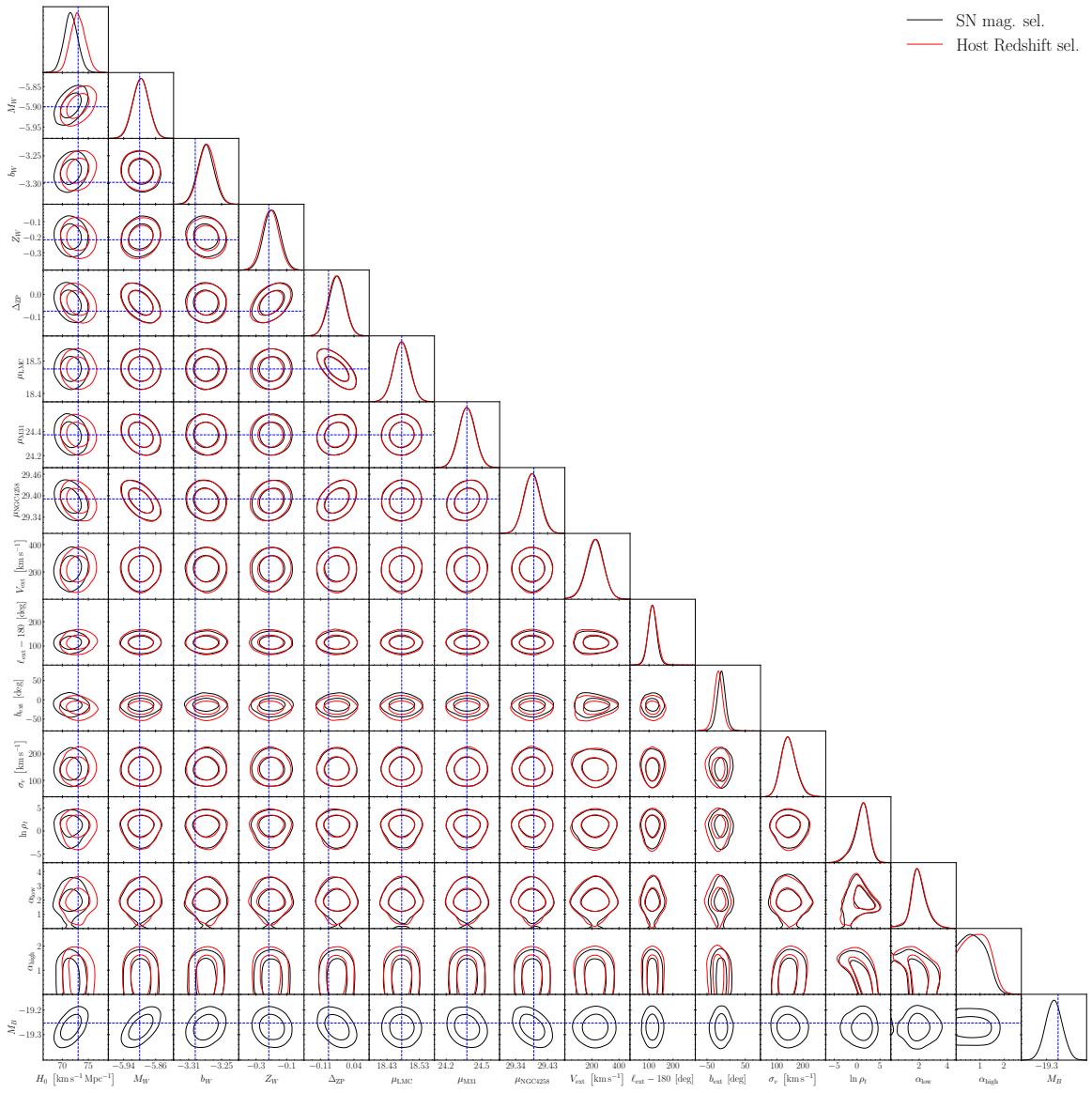


Figure 4.9: Constraints on all model parameters, except the 35 Cepheid host distances, for our fiducial peculiar velocity model based on the `Manticore-Local` reconstruction (McAlpine et al., 2025), assuming either SN magnitude (black) or host redshift selection (red). The blue lines show the best-fit parameter values from the SH0ES analysis (Riess et al., 2022a). The contours represent 1 and 2σ confidence intervals.

Peculiar velocity model	Selection model		
	None	SN mag.	Redshift
No pec. vel., σ_v	67.3 ± 1.9	68.8 ± 2.0	73.9 ± 2.7
Constant inferred flow $\boldsymbol{V}_{\rm ext}$, σ_v	67.9 ± 1.7	69.3 ± 1.7	73.2 ± 2.3
ΛCDM covariance, σ_v^2	68.3 ± 3.1	70.1 ± 3.0	76.1 ± 3.1
ΛCDM covariance, scaling A, σ_v^2	68.4 ± 3.3	70.1 ± 3.4	78.5 ± 4.0
Carrick et al. (2015), $\boldsymbol{V}_{\rm ext}$, σ_v	68.9 ± 1.6	72.1 ± 1.6	74.4 ± 1.8
`Manticore-Local`, $\boldsymbol{V}_{\rm ext}$, σ_v	71.1 ± 1.2	$\mathbf{71.7 \pm 1.3}$	73.0 ± 1.4
`Manticore-Local`, $\boldsymbol{V}_{\rm ext}$, $\sigma_v(\delta)$	71.1 ± 1.2	71.6 ± 1.3	72.9 ± 1.4
`Manticore-Local`, $\boldsymbol{V}_{\rm ext}$, σ_v, β	70.9 ± 1.2	71.5 ± 1.3	72.7 ± 1.4
External comparison			
CMB (Planck Collaboration et al., 2020b)	67.4 ± 0.5		
SH0ES (Riess et al., 2022a)	73.0 ± 1.0		
SH0ES (with the SMC, Breuval et al. 2024)	73.2 ± 0.9		
SH0ES Cepheid-only (Kenworthy et al. 2022)	72.9 ± 2.3		

Table 4.3: Inferred values of the Hubble constant H_0 in units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for different combinations of peculiar velocity models (rows) and selection models (columns). All results are based on the forward modelling approach described in Section 4.3.1. See Section 4.4.2 for details. We highlight in bold peculiar velocity modelling with `Manticore-Local` (McAlpine et al., 2025), which we consider our fiducial model. The lower block contains literature constraints for external comparison. All uncertainties are 1σ statistical except for that of Kenworthy et al., which combines statistical and systematic uncertainty.

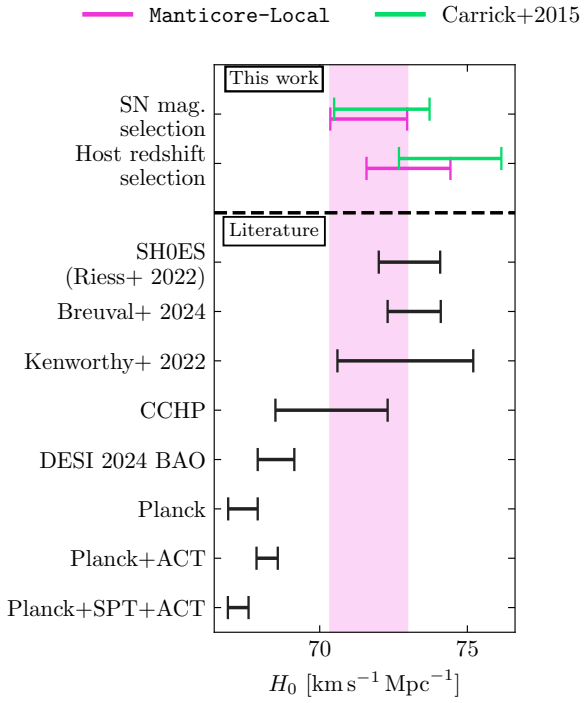


Figure 4.10: Inferred H_0 using `Manticore-Local` (McAlpine et al., 2025) and Carrick et al. (2015) for the fiducial SN magnitude selection, alongside the result under redshift selection, compared to literature values (Riess et al., 2022a; Breuval et al., 2024; Kenworthy et al., 2022; Freedman et al., 2025; Adame et al., 2025a; Planck Collaboration et al., 2020b; Louis et al., 2025; Camphuis et al., 2026). The pink band shows our fiducial result; all bands are 1σ.

4.5.1 Importance of selection function modelling

Typically (or at least ideally), the selection strategy for a survey is well-defined and imposed, for example, on apparent magnitude (i.e. a flux limit), reflecting the detector's sensitivity. However, as the sample of Cepheid host galaxies has been assembled over many years as follow-up to nearby SN hosts from various proposals, it is unclear which selection is most applicable (K22). Nonetheless, given that all galaxies in the sample are relatively nearby, some form of selection must be present. Therefore, we consider two main selection scenarios: a SN apparent magnitude limit and a host galaxy redshift limit. K22 considered these same two cases, arguing for a SN-based selection on the grounds that Cepheid hosts are typically targeted for follow-up after a SN is observed, with a preference for brighter (nearer) SNe within which (once selected) Cepheids are always observed.

Given the distributions of SN magnitudes and host redshifts shown in Fig. 4.1, we consider selection based solely on SN magnitude to be the more plausible scenario. Compared to the Pantheon+ sample, such a selection can reproduce the Cepheid host redshift distribution, whereas selection on redshift alone fails to reproduce the SN magnitude distribution. We briefly explore mixed selection scenarios in Section 4.3.3.4, finding that they simply interpolate between the two cases. The effect of Cepheid selection may however also need to be considered. This would require modelling not only the Cepheid population (e.g. period and metallicity distributions) but also Cepheids' spatial distribution within host galaxies, which affects detectability, and propagating the resulting non-trivial covariance matrix. We leave this for future work.

If the sample was selected based on the SN magnitude, neglecting this selection would bias H_0 low, as the selection preferentially retains nearby sources. If the forward model does not account for this, it may place sources at greater distances which, at fixed redshift, leads to a lower and biased estimate of H_0. Because the SN and Cepheid apparent magnitudes (which are directly used within the forward model) are correlated via the host galaxy

distance, selection on SNe magnitude alters the distribution of "true" source distances retained in the sample, thus necessitating explicit modelling as described in Section 4.3.3.2 that (weakly) couples the inference to the SNe. Consequently, the SN magnitude selection analysis is no longer independent of SNe; however, we verify that including SN apparent magnitudes in the forward model alone (without modelling their selection) has no effect on the inferred H_0. On the other hand, if selection is assumed to be imposed on either the Cepheid apparent magnitude or host galaxy redshift, then there is no need to account for SN apparent magnitudes, and the analysis remains independent of SN data.

The standard SH0ES analysis (Riess et al., 2022a) models the selection and standardisation biases of the Hubble-flow SNe through bias corrections to their apparent magnitudes, and applies Cepheid sample cuts and photometric corrections, but does not include an explicit Cepheid population-selection normalisation in the likelihood. We adopt similar logic under SN-magnitude selection, but forward-model it within a single hierarchical inference with propagated uncertainty, and compare several selection hypotheses as an explicit systematic. Crucially, by modelling the local peculiar velocity field we infer a precise H_0 from the Cepheid host redshifts alone, independent of the SNe (up to the selection modelling of the Cepheid hosts), thereby providing a test of whether the SNe are a source of systematics in the distance ladder.

4.5.2 Comparison to Kenworthy et al. 2022

While using the exact same 35 host galaxies, our approach has three principal differences to that of K22:

1. We forward model the Cepheid apparent magnitudes, rather than adopt a-priori calibrated distance estimates.

2. We employ a principled treatment of selection effects, avoiding arbitrary modifications to the prior on object distances (unless considering an effective selection model).

3. We use a state-of-the-art local Universe reconstruction `Manticore-Local`, rather than relying solely on linear-theory reconstructions. This marginalises over the full posterior rather than assuming a Gaussian covariance.

We now compare the methods in detail, showing that neither the redshift- nor distance-limited selection of K22 is well justified. A distance-based selection cannot be considered in the first place, since selection must act on the level of an observable quantity, which distance is not.

4.5.2.1 Redshift selection

For redshift selection, K22 model the case in which only the host galaxy redshift was effectively used in the selection procedure. It appears that they mean this to imply a prior on the comoving distance χ of the Cepheid hosts that models homogeneous and inhomogeneous Malmquist bias in the usual way (their Eq. 15), without any further selection modifications. However, they then write "Tests on simulated data found that the use of this prior tended to bias the recovered distances and H_0 high", and proceed to modify the prior to:

$$\pi(\chi) = \max(\chi)^\alpha \prod_i \frac{\chi_i^2}{\int_0^{\max(\chi)} \chi^2 \mathrm{d}\chi}, \tag{4.22}$$

where χ is expressed in h^{-1} Mpc and α is set to 60. This entails that:

1. The individual χ_i^2 terms are normalised at each step to the current maximum inferred host distance, effectively truncating the χ^2 prior at $\max(\chi)$.

2. The prior is not a valid probability density function, as the $\max(\chi)^\alpha$ factor breaks the normalisation. Moreover, it is not merely an improper prior on distance but it explicitly depends on h, which is a separate (and the crucial) parameter.

To quantify the bias on H_0 that this produces, we consider the relation between priors on χ_i and r_i, where r_i is physical distance in Mpc such that $\chi_i = r_i h$. If $\pi(\chi_i) = 3\chi_i^2/\chi_{\max}^3$,

which corresponds to a prior on χ_i that goes as χ_i^2 but truncated at $\chi_{\rm max}$, the corresponding prior on r_i is

$$\pi(r_i) = \pi(\chi_i) \left| \frac{\mathrm{d}\chi}{\mathrm{d}r} \right|_{r_i} = \frac{3r_i^2}{r_{\rm max}^3}, \tag{4.23}$$

where $r_{\rm max} = \chi_{\rm max}/h$. Therefore, provided that this term is normalised as it is in K22's Eq. 17, it does not introduce spurious factors of h. However, the factor of $\max(\boldsymbol{\chi})^\alpha$ introduces a problem by modifying the "prior" on individual distances to

$$\pi(\chi_i) = \chi_{\rm max}^{\alpha/n} \frac{3\chi_i^2}{\chi_{\rm max}^3}, \tag{4.24}$$

where n is the number of Cepheid host galaxies. This distribution is no longer normalized and alters the prior on r in a strange way. Since $\chi_{\rm max} = r_{\rm max}h$, this introduces a spurious factor of $(r_{\rm max}h)^{\alpha/n}$ per host galaxy. Thus, the $\max(\boldsymbol{\chi})^\alpha$ term implicitly imposes a prior on H_0 proportional to H_0^α, which K22 set to H_0^{60}. The simulations from which they "derive" this prior include a selection in the observed redshift of $z < 0.011$. We have shown that without modelling peculiar velocities or the inhomogeneous Malmquist bias, the correct method for accounting for a redshift selection introduces a factor of H_0^3 per host galaxy, as discussed in Section 4.3.3.3. Thus, for a uniform-in-volume prior and a redshift-selected sample of 35 host galaxies, the correct treatment of the selection introduces a factor of H_0^{105}, rather than H_0^{60}.

Adopting C15 as the peculiar velocity model, K22 found $H_0 = 76.4^{+2.6}_{-2.4}$ or $74.1^{+2.4}_{-2.1}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, depending on whether the redshift scatter σ_v is fixed or inferred. In contrast, with the same field and inferring σ_v we find $H_0 = 74.4 \pm 1.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Although in principle our selection treatment effectively introduces a factor of H_0^{105} rather than H_0^{60}, direct comparison is complicated by our explicit treatment of both peculiar velocities and inhomogeneous Malmquist. While K22 introduced an ad hoc prior modification in an attempt to minimise bias in redshift-limited mocks (which is ultimately incorrect), we present a principled method for modelling redshift selection, as well as properly accounting for the effects of

peculiar velocities and inhomogeneous Malmquist bias on it. We demonstrate that our method is unbiased in Section 4.A.

4.5.2.2 Distance selection

The second case considered by K22 is a distance-limited selection, which is presumably meant to model a selection on the observable SN magnitudes. In this model, they set the prior on the host galaxy distances to

$$\pi(\chi_i|d_T) = \frac{f(\chi_i, d_T)}{\int_0^{80} f(x, d_T)\mathrm{d}x}, \tag{4.25}$$

where d_T is the distance cutoff in h^{-1} Mpc, the dummy variable x is also in h^{-1} Mpc and

$$f(\chi, d_T) = \chi^2\, \Phi\left(\frac{\mu(\chi) - \mu(d_T)}{0.15}\right). \tag{4.26}$$

(The inhomogeneous Malmquist term is also included but we omit it for brevity here.) Φ is the CDF of a standard Gaussian distribution. However, since the CDF has the limits

$$\Phi(x) \to 0 \quad \text{as } x \to -\infty; \quad \Phi(x) \to 1 \quad \text{as } x \to +\infty, \tag{4.27}$$

this implies that χ values *below* d_T are cut off. We assume this is a typo and instead interpret Φ as the survival function, defined as

$$\mathrm{SF}(x) = 1 - \Phi(x), \tag{4.28}$$

with which we can reproduce the reported results. Using the CDF would force d_T to be less than the distance to the nearest host, since the model cannot assign host galaxies to distances where the prior is forced to zero by the CDF.

To verify this, we implement the model of K22 and their distance selection (described

in Section 4.5.2). Unlike our analysis, this involves predicting some a priori-calibrated distance estimates, rather than deriving them simultaneously with the other model parameters by forward-modelling the Cepheid magnitudes. For this, we use the Cepheid-derived distances from our redshift-independent analysis in Section 4.4.1 with a uniform-in-distance-modulus prior, matching the data and assumptions they used. When working in h^{-1} Mpc as they do and marginalising over peculiar velocities with the ΛCDM covariance matrix, we find $H_0 = 71.6 \pm 3.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, consistent with their reported value of $71.6 \pm 4.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ as shown in the second row of their Table 2. (Their uncertainty is larger because they over-estimate the ΛCDM peculiar velocity covariance; see the last paragraph of Section 4.C.) Conversely, if we use the model of K22 in terms of predicting "observed" distances derived from a prior analysis, but work consistently in Mpc without applying any "selection" modification to the prior, we again find a lower value of $H_0 = 68.4 \pm 3.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, consistent with our no-selection results.

However, the approach of K22 does not provide a principled method for modelling selection in apparent magnitude. In Section 4.3.3.2, we derive the correct treatment: in a Bayesian hierarchical framework, modifying the prior of latent variables is not justified; instead, the appropriate approach is to apply a correction for the unobserved sources, which is independent of the properties of the detected sources.

Although K22 do not consider the case of no selection, it is interesting to see how it would play out when measuring distances in h^{-1} Mpc units as they do. This would correspond to $\pi(\chi_i) \propto \chi_i^2$, which, if treated as an improper prior without normalisation (as would appear from their Eq. 15), would produce a factor of h^3 per galaxy when mapping it to the prior on r_i: a factor of h^2 derives from χ^2, and h from the Jacobian of the transformation from χ to r. This yields a prior scaling as H_0^{105}, which, as we have shown, is effectively the result of redshift selection. We therefore see that if one works in redshift as the independent variable one does not need to explicitly model redshift selection, analogously to how one does not need to explicitly model selection effects for a volume-limited sample when working in

physical distances.

While most of the issues we find with K22 are specific to that analysis, we note that the issue of distance priors is pervasive to distance-ladder studies. In general, adopting a uniform-in-distance-modulus or uniform-in-distance prior (r^0 or r^{-1}) puts galaxies at smaller distances relative to the physically motivated uniform-in-volume prior (r^2), biasing H_0 high. This may be contributing to a community-wide overestimation of H_0 (Desmond et al., 2026), although it is complicated by selection effects.

4.5.2.3 Modelling the peculiar velocity covariance

We effectively considered three discrete velocity models: `Manticore-Local`, the C15 field (both with additional variance), and the ΛCDM covariance without a mean field. K22 add an additional covariance to C15 describing contributions to the velocity field at small enough scales not captured by C15 yet large enough to couple galaxies, i.e. not simply a diagonal σ_v contribution. They find that the inclusion of this covariance term yields a relatively large shift in H_0, pushing it higher. To derive this additional covariance, K22 use the reconstruction of Lilow & Nusser (2021), which is a linear reconstruction based on the 2MRS galaxy density field (Huchra et al., 2012) within a $200\,h^{-1}$ Mpc radius. We do not adopt this approach: a covariance of unmodelled peculiar velocities cannot be reliably constructed from two discrete fields. However, it can be estimated from `Manticore-Local`, which provides self-consistent posterior samples of the large-scale structure, and accounted for directly by marginalising over field realisations, as we do. Moreover, this approach ensures a self-consistent treatment of the peculiar velocity correlations between host galaxies.

We also investigate directly whether the covariance between galaxies on scales smaller than those resolved by C15 is significant: we consider the contribution to the ΛCDM peculiar velocity covariance matrix from scales not captured by C15. The effective smoothing applied to the galaxy density field by C15 is $4\,h^{-1}$ Mpc, which corresponds to a wavenumber $k \approx 1.6\,h\,\mathrm{Mpc}^{-1}$. (K22 apply an additional Gaussian smoothing of $3\,h^{-1}$ Mpc.) To be

conservative we adopt a cutoff of $k > 0.5\,h\,\mathrm{Mpc}^{-1}$ (corresponding to a physical scale of $12.6\,h^{-1}\,\mathrm{Mpc}$) and set the power spectrum to zero below this limit when computing the ΛCDM peculiar velocity covariance matrix. Figure 4.11 shows the resulting correlation matrix for all 35 Cepheid host galaxies. Compared to Fig. 4.3 (computed using the full power spectrum), the correlations are significantly reduced, since it is the large-scale flows that are primarily responsible for coupling the velocities. Of the 595 possible pairs of galaxies, only 7 have an absolute expected ΛCDM peculiar velocity correlation exceeding 0.1, and only 57 exceed 0.02. In comparison, for the full covariance matrix these numbers were 350 and 534, respectively.

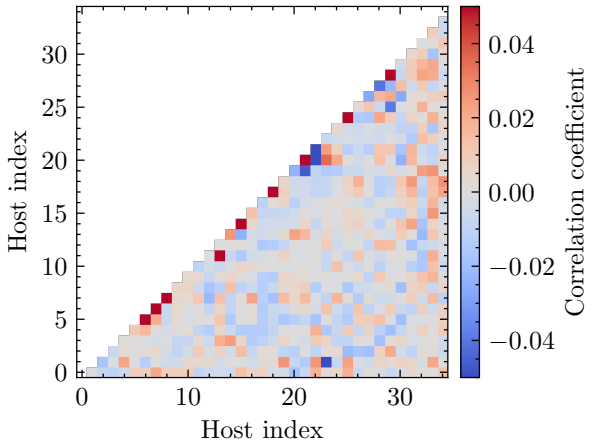


Figure 4.11: Correlation matrix of the ΛCDM peculiar velocity covariance including only modes with $k > 0.5\,h\,\mathrm{Mpc}^{-1}$, conservatively modelling scales not reconstructed by Carrick et al. (2015). Excluding large-scale contributions reduces the correlations significantly compared to Fig. 4.3, confirming that a diagonal covariance suffices for sub-reconstruction scales.

To quantify this effect, we infer H_0 using the same model as in Section 4.4.2, but compare two cases (each assuming SN magnitude selection). In the fiducial analysis, when peculiar velocities are accounted for using the C15 reconstruction we assume the covariance matrix to be diagonal with variance σ_v^2. In the second case, we additionally include the ΛCDM covariance matrix, which accounts for contributions from all scales with $k > 0.5\,h\,\mathrm{Mpc}^{-1}$. This fiducial analysis yields $H_0 = 72.1 \pm 1.6$, while the latter yields $72.1 \pm 1.6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This demonstrates that the off-diagonal contributions are negligible for scales not captured by C15, justifying our fiducial model where only a diagonal contribution (σ_v) is added to that. A similar conclusion holds for the other selection models. This disagrees with the finding of K22, which we attribute to them modelling the peculiar velocity covariance from

the two distinct fields of C15 and Lilow & Nusser (2021).

Moreover, we use `Manticore-Local` which has a spatial resolution of 0.7 h^{-1} Mpc—hence self-consistently modelling scales smaller than those resolved by C15—and provides an ensemble of plausible realisations of the local peculiar velocity field describing the reduced cosmic variance given the *Bayesian Origin Reconstruction from Galaxies* (BORG) constraints. However, in all cases we find that `Manticore-Local` and C15 agree to about 1σ. In Stiskalek et al. (2026c); McAlpine et al. (2025) we benchmarked `Manticore-Local`, C15 and other reconstructions against direct distance data, finding `Manticore-Local` to be the most accurate model of the local Universe to date, with the C15 reconstruction also performing well. Compared to C15, the primary advantages of `Manticore-Local` are its more accurate gravity and galaxy bias models which allow the modelling of the density and velocity field at a higher resolution by going beyond linear theory, and its independent posterior samples of the large-scale structure, over which we marginalise.

The higher fidelity of the `Manticore-Local` reconstruction is particularly important for host galaxies located within or near galaxy clusters. In Stiskalek et al. (2026c), we showed that BORG-based models provide significantly improved modelling of galaxies in over-dense regions compared to linear modelling. This is particularly relevant here because at such low redshifts peculiar velocities are the dominant source of uncertainty, and the local dynamics is dominated by infall to nearby clusters, most notably Virgo, Fornax, and Leo. Indeed at least seven of the 35 host galaxies lie within the infall region of these clusters (this number is approximate as we do not carefully test for cluster membership). In contrast, C15 provides only a single mean field estimate.

4.5.3 Ramifications of the results and future work

We have shown that a Cepheid-only distance ladder yields H_0 values consistent with SH0ES (Riess et al., 2022a), though using the state-of-the-art `Manticore-Local` reconstruction to model galaxy bias and account for peculiar velocities gives somewhat lower

values unless the sample is assumed to be redshift-selected. We adopt `Manticore-Local` as our fiducial peculiar velocity model and compare the inferred H_0 with results obtained using the C15 reconstruction. We find good agreement between the two, though `Manticore-Local` yields slightly lower values of H_0 (see Fig. 4.10 or Table 4.3 for a summary of all H_0 variations). We picked the `Manticore-Local` reconstruction because of its validated peculiar velocity field and improved performance relative to other available reconstructions (Stiskalek et al., 2026c; McAlpine et al., 2025). Additionally, `Manticore-Local` accurately predicts both the masses and positions of all nearby clusters across all validation metrics (see McAlpine et al. 2025). We do not include any SNe in our analysis (except for modelling a SN selection), instead focusing on a detailed investigation of the two lower rungs of the distance ladder. Including the Type Ia SNe hosted by these galaxies as additional distance indicators would not significantly tighten the host distances. Each host distance is already pinned by its Cepheids to ~ 0.06 mag, whereas a single SN Ia carries a distance scatter roughly twice as large (Riess et al., 2022a; Brout et al., 2022). Moreover, the SN absolute magnitude M_B that would calibrate such distances is itself set by the same Cepheid host distances, so the SNe add little genuinely independent information. They would become useful for this specific purpose only if one supplied an external calibration of M_B or reintroduced a Hubble-flow SN rung, neither of which belongs to the two-rung test considered here.

Using `Manticore-Local` to account for peculiar velocities (and their covariance), the uncertainty on the inferred H_0 with only 35 Cepheid host galaxies ranges between 1.2 to 1.4 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$, which is less than two per cent. Relative to the C15 model, `Manticore-Local` yields a 20 per cent reduction in the uncertainty on H_0. The relatively small volume probed by the second-rung objects underscores the need for increasingly faithful representations of the local Universe such as those enabled by the `BORG` programme (Jasche & Wandelt, 2013; Jasche et al., 2015; Lavaux & Jasche, 2016; Leclercq et al., 2017; Lavaux et al., 2019; Jasche & Lavaux, 2019; Porqueres et al., 2019; Stopyra et al.,

2024a; McAlpine et al., 2025) to suppress both the statistical and systematic uncertainty due to peculiar velocities. However, assuming these (and other) systematic uncertainties are already under control, further crucial evidence concerning the Hubble tension could be obtained by incorporating more second-rung data.

Even restricting to Cepheids, the host galaxies we used represent only a small subset of available Cepheid measurements. We have used the SH0ES sample because of its high-quality (e.g. Riess et al., 2022b, 2023; Breuval et al., 2023; Bhardwaj et al., 2023; Riess et al., 2024; Breuval et al., 2024) and self-consistent (Nájera & Desmond, 2025) measurements, including well-characterised covariances. However, in principle more Cepheid hosts could be incorporated. Tully et al. (2023) compile Cepheid distances for 76 galaxies (of which 40 are from SH0ES), so relaxing the SH0ES SN requirement could roughly double the 35 hosts used here and lower the statistical uncertainty on H_0. However, these additional hosts derive from uncoordinated programmes with distinct photometric systems, metallicity scales, and crowding treatments, so that, unlike the homogeneous SH0ES set we adopt, their cross-sample systematic covariance is poorly characterised, or absent altogether as in Tully et al. (2023), and their selection functions are unknown, compounding the selection modelling that already dominates our systematic budget. Beyond that, several alternative second-rung indicators exist, including the Tip of the Red Giant Branch (TRGB; Freedman et al. 2019; Li et al. 2024), Type II SNe (Vogl et al., 2025), J-region asymptotic giant branch (JAGB; Freedman & Madore 2020), Mira variables (Huang et al., 2024), Surface Brightness Fluctuations (SBF; Cantiello & Blakeslee 2023; Jensen et al. 2025) and potentially other less explored methods (Nájera & Desmond, 2025). Blakeslee et al. (2021) used an SBF calibration anchored to both Cepheid and TRGB distances to infer $H_0 = 73.3 \pm 0.7\,(\mathrm{stat}) \pm 2.4\,(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This measurement has recently been updated with HST calibration by Jensen et al. (2025), who found $H_0 = 73.8 \pm 0.7\,(\mathrm{stat}) \pm 2.3\,(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The dominant systematic error arises primarily from the distance calibration, suggesting the value of a principled joint calibration approach similar to the one we undertake here. A

future analysis should revisit this kind of approach to model jointly the Cepheid, TRGB, and SBF (and perhaps other) distances using the framework outlined here.

It is likely that with additional distance data and ever-improving peculiar velocity modelling (such as with `Manticore-Local`), second-rung inferences could become even more precise, with the key added benefit of being independent of SNe and associated potential systematics. This is even before considering future data: Anand et al. (2025) argue that only 24 new Cepheid hosts observed by the proposed Habitable Worlds Observatory (HWO) would be sufficient to reach a one per cent determination of H_0 using the SH0ES error model. As of now, with `Manticore-Local` we achieve a determination of H_0 to better than two per cent (up to the caveat of uncertain selection).

A two-rung ladder is particularly useful because of its complete elimination of SNe, beyond their use to select nearby hosts. Indeed, there have been several studies suggesting possible miscalibration of SNe (e.g. Seifert et al. 2025; Efstathiou 2025; Wojtak & Hjorth 2022; Hoyt et al. 2026; Son et al. 2025). As an example, a recent exploration by Wojtak & Hjorth (2025) modelled the SH0ES SN sample as comprising two populations with distinct intrinsic and extrinsic properties—primarily differences in mean absolute magnitude and extinction coefficients—finding a preference for lower H_0 values and reducing the Hubble tension by 30 to 50 per cent. Indeed, an alternative programme to SH0ES—the Chicago-Carnegie Hubble Program (CCHP; Freedman et al. 2019)—infers $H_0 = 69.8 \pm 0.8$ (stat) $\pm$ 1.7 (sys) km s^{-1} Mpc^{-1} using the TRGB. This was recently refined using James Webb Space Telescope data of TRGB, Cepheids and the J-region asymptotic giant branch, finding very similar results (Freedman et al., 2025).

4.6 Conclusion

The persistence of the Hubble tension demands that we multiply cross-check the cosmic distance ladder to vet the local H_0 measurement fully. To this end, we have considered

dropping the SNe from the SH0ES pipeline to constrain H_0 from the Cepheid distances (with geometric anchors) and host galaxy redshifts alone. To minimise the risk of introducing new systematics, we keep our data as close as possible to the SH0ES pipeline (Riess et al., 2022a). This follows in part the analysis of K22, who find $H_0 = 72.9^{+2.4}_{-2.2}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, where the uncertainty includes both statistics and systematics. We find several issues with this analysis, mostly to do with selection modelling. We introduce a principled framework for modelling selection effects and also employ state-of-the-art peculiar velocity modelling with `Manticore-Local` (McAlpine et al., 2025), including marginalisation over plausible realisations of the local Universe. We summarise the inferred H_0 in Fig. 4.10. Our main findings are as follows:

1. Our fiducial model, which accounts for peculiar velocities using `Manticore-Local`, yields $71.7 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when the Cepheid host sample is assumed to be selected on the basis of SN apparent magnitudes, which we consider the more plausible scenario. If however the sample were selected on observed redshift, we would infer $H_0 = 73.0 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. These two cases bracket the range of plausible H_0 values, with mixed selection scenarios lying in between. Our headline result is in 3.3σ tension with combined CMB data from ACT, SPT, and Planck (Camphuis et al., 2026).

2. While Stiskalek et al. (2026c); McAlpine et al. (2025) have demonstrated that `Manticore-Local` provides the most accurate velocity field among all currently in the literature, we explore others to investigate possible systematic uncertainties. Using the C15 reconstruction shifts the inferred H_0 marginally higher, as well as increasing the uncertainty on H_0 by twenty per cent (producing $H_0 = 72.1 \pm 1.6$ and $74.4 \pm 1.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for SN magnitude and redshift selection, respectively). Any H_0 inference with second-rung objects only is limited by the peculiar velocity modelling, highlighting the need for a state-of-the-art reconstruction like `Manticore-Local`.

3. We identify several problems in the K22 analysis of the same data: *i)* they adopt an unphysical prior on the galaxy distances used as input to their analysis (r^{-1} rather than the uniform-in-volume r^2; see also Desmond et al. 2026), *ii)* they model selection effects as acting on the prior of the latent true distances, rather than modelling the unobserved sources, and *iii)* by representing distances in h^{-1} Mpc rather than Mpc units (i.e. actually redshifts), they introduce spurious high powers of H_0 in the prior.

We conclude that the SH0ES Cepheid sample alone yields a Hubble constant consistent with the full three-rung SH0ES analysis and with only slightly larger uncertainty, though with a $\sim 1\sigma$ lower central value in the fiducial magnitude-selection scenario. This conclusion relies on careful modelling of both the selection function and peculiar velocities using the state-of-the-art `Manticore-Local` model. The ongoing challenge of understanding and resolving the Hubble tension motivates further work on both the second rung of the distance ladder, independent of SNe, and on the modelling of SNe themselves (e.g. `BayeSN`; Mandel et al. 2022; Grayling et al. 2024).

More constructively, our analysis demonstrates that H_0 may be inferred with high precision (less than two per cent statistical uncertainty) from first-plus-second-rung data alone, extending only to $z \approx 0.01$. This suggests that bringing further such data to bear may readily afford a Hubble constant inference even more precise than SH0ES and crucially avoiding the third rung (e.g. SNe) altogether. We stress that while the second-rung is independent of SNe (except when modelling the host selection), it is susceptible to significant systematics arising from selection effects or peculiar velocity modelling if these are not properly accounted for. Thus, this endeavour requires principled methods for accounting for selection (as we have implemented here) as well as accurate density and velocity field models of the local Universe, which are growing in quality through Bayesian forward modelling of the galaxy number density field (`BORG`). Finally, we remark that the sensitivity to selection effects underscores the need for observing programmes to adopt well-defined selection procedures that render the required cosmological inference procedure unambiguous.

4.A Mock data bias tests

We assess the impact of selection on either SN apparent magnitude or host galaxy redshift using mock catalogues. We generate mock catalogues designed to match the properties of the observed data and demonstrate that the inferred value of H_0 remains unbiased.

The SN magnitude-selected sample is generated as follows. For simplicity, rather than modelling the full CPLR, we assume all Cepheid stars have identical absolute magnitudes. We draw source galaxy distances from a uniform-in-volume (r^2) distribution. For each trial, we compute the SN apparent magnitude, given an assumed $M_B = -19.25$, adding a scatter of $\sigma_{\rm SN} = 0.15$ dex. If the resulting apparent magnitude is below the selection threshold, we retain the sampled distance; otherwise, we repeat the process. For each accepted host galaxy, we generate five Cepheids (the precise number is unimportant), each at the same distance, with absolute magnitude $\tilde{M}_W = -18$ and a scatter in observed magnitude of $\sigma_{\rm Ceph} = 0.1$ dex. The observed redshift of the host galaxy is then computed, assuming a given $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and including a velocity dispersion of $\sigma_v = 250\,{\rm km\,s^{-1}}$ to mimic peculiar velocities. This procedure is repeated until a sample of 35 host galaxies is obtained. Finally, to mimic the inclusion of the LMC and NGC 4258 as calibrators, we add two nearby sources. For each, we assign a true distance modulus and sample an "observed" distance modulus with a scatter of 0.25 dex, similar to the values reported in the literature. The host galaxy redshift-limited sample is generated analogously, except that SN apparent magnitudes are not modelled; instead, the rejection process is based on whether the host galaxy redshift falls below a specified threshold. When sources are drawn from an r^2 distribution, a maximum distance must be imposed for practical reasons. We adopt an upper limit of 150 Mpc, which ensures that all sources passing either the magnitude or redshift selection criteria lie well below this threshold (typically within 50 Mpc). We also test an alternative procedure in which we generate a fixed total number of galaxies and then retain those that pass the cut, rather than resampling until a fixed number pass the cut. In

both cases our procedure is unbiased.

In Fig. 4.12, we present bias tests in H_0 for the SN magnitude- and redshift-limited selection. In each case, we use identical injected parameter values but draw the mock catalogue using different random seeds. This procedure is repeated 500 times. The bias is defined as

$$\mathrm{Bias}(H_0) = \frac{\langle H_0 \rangle - H_0^{\mathrm{true}}}{\sigma_{H_0}}, \tag{4.29}$$

where $\langle H_0 \rangle$ and σ_{H_0} denote the posterior mean and standard deviation of the inferred H_0, and H_0^{true} is the true (injected) value. If the inferred H_0 is unbiased (and its posterior is Gaussian), then the distribution of $\mathrm{Bias}(H_0)$ across the mock realisations should follow a standard normal distribution. Fig. 4.12 demonstrates that, when selection effects are properly modelled, the recovered H_0 is unbiased on average in all cases. In contrast, neglecting selection effects introduces a significant bias in the inferred H_0 for each scenario.

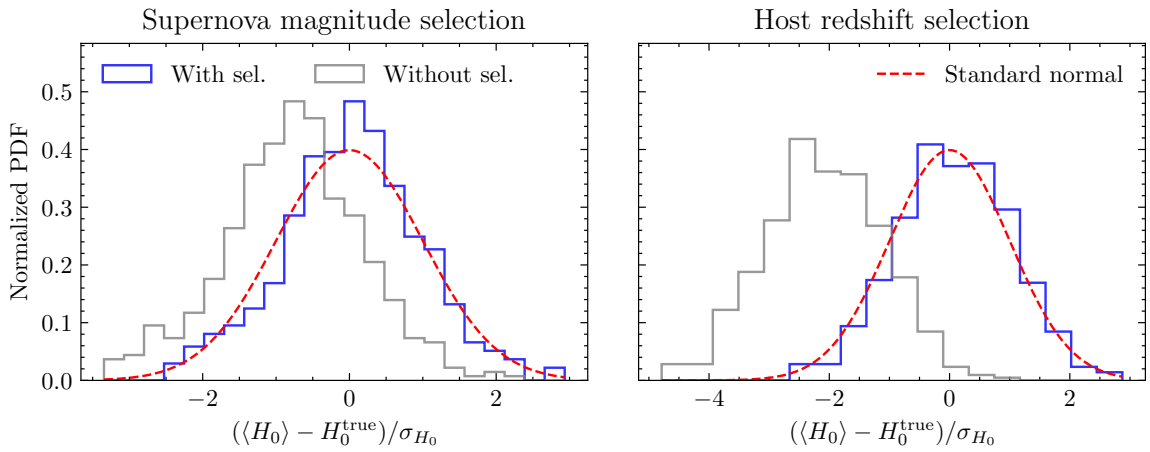


Figure 4.12: Bias test of H_0 for "Cepheid"-only mock inferences, as described in Section 4.A, under host selection by either SN magnitude or host redshift. When selection effects are not modelled (grey), the inferred H_0 is significantly biased low. By contrast, when selection is properly modelled (blue), the inference is unbiased in both cases, as indicated by excellent agreement with the standard normal distribution (red dashed).

4.B Local Universe reconstructions

As discussed in Section 4.3, some of our peculiar velocity models use the density and velocity field reconstruction of either `Manticore-Local` or C15. This enables us to both model the inhomogeneous Malmquist bias as a function of the local density, as well as the contribution of the LOS peculiar velocities to observed redshifts. Both reconstructions are based on the 2M++ galaxy catalogue (Lavaux & Hudson, 2011), which contains 69 160 galaxies and combines 2MASS photometry (Skrutskie et al., 2006) with redshifts from 2MRS (Huchra et al., 2012), 6dF (Jones et al., 2009), and Sloan Digital Sky Survey (SDSS) DR7 (Abazajian et al., 2009). Apparent magnitudes are corrected for Galactic extinction, k-corrections, evolution, and surface brightness dimming. The 2M++ sample is magnitude-limited to $K < 11.5$ in the 2MRS region and $K < 12.5$ in the 6dF and SDSS regions.

The density and velocity field of either `Manticore-Local` or C15 is reconstructed in real space. Since the real-space distances to the host galaxies are not known a priori, we consider all plausible velocities along the LOS to each source. The predicted peculiar velocity of the ith host is given by

$$V_{\text{pec},i}^{\text{C15}} = (\beta \boldsymbol{V}(\boldsymbol{r}_i) + \boldsymbol{V}_{\text{ext}}) \cdot \hat{\boldsymbol{r}}_i, \tag{4.30}$$

where $\boldsymbol{V}(\boldsymbol{r}_i)$ is the velocity field evaluated at candidate galaxy position $\boldsymbol{r_i}$. β is set to unity for `Manticore-Local` and treated as a free parameter when using the reconstruction of C15 (where it is related to σ_8). The associated density field can be used to model galaxy bias $n_g(r)$, i.e. the so-called inhomogeneous Malmquist bias, which is the prior preference for galaxies to be located in regions of higher matter density. This leads to the following prior probability for the physical distance (in Mpc) to a given galaxy

$$\pi(r \mid \boldsymbol{\theta}) = \frac{r^2 n_g(r, \boldsymbol{\theta})}{\int_0^R r'^2 n_g(r', \boldsymbol{\theta}) \mathrm{d}r'}, \tag{4.31}$$

where the integral is evaluated along the LOS to the galaxy.[4] Without $n_g(r)$ the prior could not be normalised but this would not matter because the normalisation would be independent of the model parameters and hence the prior could be left improper. However, because of the varying LOS densities per source, the normalisation in modelling the galaxy bias is no longer source-independent. It may also depend on inferred parameters and, therefore, must be computed explicitly for each source. We discuss the galaxy bias choices for `Manticore-Local` and C15 below.

Ideally, the integral in Eq. (4.31) should extend to $R \to \infty$, but the reconstructed density field is available only within approximately 200 Mpc. We therefore assume $n_g(r) = 1$ outside the reconstruction, so the integrand reduces smoothly to r^2. To avoid sharp discontinuities, we extrapolate the overdensity field $\delta(r)$ beyond the reconstructed volume with an exponential decay towards zero with a $5\,h^{-1}$ Mpc scale length. We have checked that the results are not sensitive to this choice. The normalisation integral is then computed numerically up to R. Contributions beyond R are independent of the model parameters and can be neglected; the posterior is therefore independent of R provided it is > 200 Mpc. We note that using an improper prior proportional to χ^2, where χ denotes comoving distance in h^{-1} Mpc, leads to biased inference because the prior explicitly depends on h. This is effectively putting a prior on redshift, where it should be on physical distance.

4.B.1 The `Manticore-Local` reconstruction

`Manticore` is the latest reconstruction based on the `BORG` algorithm (Section 1.4.5; Jasche & Lavaux 2019; Stopyra et al. 2024a; McAlpine et al. 2025), constrained by the 2M++ galaxy catalogue. To increase the resolution beyond the native `BORG` grid, the posterior resimulation suite `Manticore-Local` (McAlpine et al., 2025) comprises 30 N-body simulations with the

[4] As noted above, this line-of-sight normalisation should be replaced with one over the full volume, in which case it cancels exactly with the prior term in the selection contribution. We verified that the difference has a negligible impact on the results and will correct the published paper accordingly. The following discussion is retained for completeness.

SWIFT code (Schaller et al., 2024), each based on independent posterior initial conditions inferred on a 256^3 grid (spatial resolution 2.65 h^{-1} Mpc). The simulations use 1024^3 dark matter particles (no baryonic physics), yielding a mean inter-particle separation of 0.7 h^{-1} Mpc and a particle mass of $2 \times 10^{10} h^{-1}\,\mathrm{M}_\odot$. Each box has a size of 681.1 h^{-1} Mpc; the constrained volume extends to $\sim$155 h^{-1} Mpc, set by the completeness of the 2M++ survey. From the $z = 0$ particle snapshots, we construct density and velocity fields using the smooth particle hydrodynamics technique (Monaghan, 1992; Colombi et al., 2007). These fields are rendered on a 1024^3 grid with a voxel size of 0.7 h^{-1} Mpc. In Fig. 4.13 we compare the LOS velocities for a sample of Cepheid host galaxies between the `Manticore-Local` and C15 reconstructions, finding very good agreement.

`Manticore-Local` assumes cosmological parameters from the Dark Energy Survey Year 3 (DES Y3; Abbott et al. 2022) "3×2pt + All Ext." ΛCDM cosmology: $h = 0.681$, $\Omega_\mathrm{m} = 0.306$, $\Omega_\Lambda = 0.694$, $\Omega_\mathrm{b} = 0.0486$, $A_\mathrm{s} = 2.099 \times 10^{-9}$, $n_\mathrm{s} = 0.967$, and $\sigma_8 = 0.807$. Thus, unlike C15, which is derived independently of the growth factor and consequently of σ_8, these are assumed in `Manticore-Local`. However, by introducing a free parameter β (see Section 4.B.2) to scale the predicted peculiar velocities, we can (approximately) assess the impact of varying σ_8—we find it to have no impact on the inferred H_0.

Moreover, while `Manticore-Local` assumes $H_0 = 68.1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, we assume our inference to be insensitive to using it while varying H_0. This is because the deviations we probe from the fiducial H_0 are small and because the hosts typically reside in relatively low-density regions (see Tsaprazi & Heavens 2025 and Table 2 of Tsaprazi et al. 2022). Moreover, when querying `Manticore-Local` for either density or velocity, we consistently convert sampled distances from Mpc to h^{-1} Mpc using the sampled value of H_0. This removes the first-order effect of varying H_0.

To model the galaxy bias $n_g(r)$ we adopt the double power-law of Section 3.3.2,

$$n_g(r, \boldsymbol{\theta}) = \left(\frac{\rho}{\rho_t}\right)^{\alpha_\mathrm{low}} \left(1 + \frac{\rho}{\rho_t}\right)^{\alpha_\mathrm{high} - \alpha_\mathrm{low}}, \tag{4.32}$$

where $\rho \equiv \rho(r)$ is the density at distance r along the LOS to the host, $\alpha_{\rm low}$ and $\alpha_{\rm high}$ are the two slopes, and ρ_t is the transition density, all of which we infer jointly with the other parameters. We place a truncated Gaussian prior on $\alpha_{\rm low}$, centred at 1 with unit standard deviation and restricted to positive values, and on $\alpha_{\rm high}$, centred at 0.5 with unit standard deviation and likewise truncated at zero. For $\ln \rho_t$ we adopt a zero-centred Gaussian prior with standard deviation of two.

We also consider a model in which the small-scale velocity dispersion depends on local density to capture higher unresolved dispersion in clusters,

$$\sigma_v(\delta) = \sigma_{v,\rm low} + \frac{\sigma_{v,\rm high} - \sigma_{v,\rm low}}{1 + \left(\frac{1+\delta}{\rho_{\sigma_v}}\right)^{-k_{\sigma_v}}}, \tag{4.33}$$

where δ is the matter overdensity, $\sigma_{v,\rm low}$ and $\sigma_{v,\rm high}$ set the asymptotic dispersions in low and high density regions, ρ_{σ_v} is the transition density, and k_{σ_v} controls the steepness of the sigmoid. We adopt the following priors: $\sigma_{v,\rm low}$ and $\sigma_{v,\rm high}$ each follow a Maxwell distribution with scale $200\,\mathrm{km\,s^{-1}}$ (mean of approximately $320\,\mathrm{km\,s^{-1}}$), $\ln \rho_{\sigma_v}$ follows a normal distribution with mean of unity and standard deviation of five, and k_{σ_v} follows a truncated normal with mean of unity, standard deviation of unity, and lower bound at zero. Our fiducial analysis retains the simpler assumption of a constant σ_v, but Table 4.3 shows that this density-dependent model yields almost identical results. As in Chapter 3, we marginalise over the 30 `Manticore-Local` posterior samples by averaging the likelihood (Stiskalek et al., 2026c).

4.B.2 The Carrick et al. 2015 reconstruction

The C15 reconstruction (described in Chapter 3) derives the velocity field from the 2M++ luminosity-weighted density field using linear theory, scaled by a free parameter

$$\beta \equiv \frac{f\sigma_{8,\rm NL}}{\sigma_{8,g}}, \tag{4.34}$$

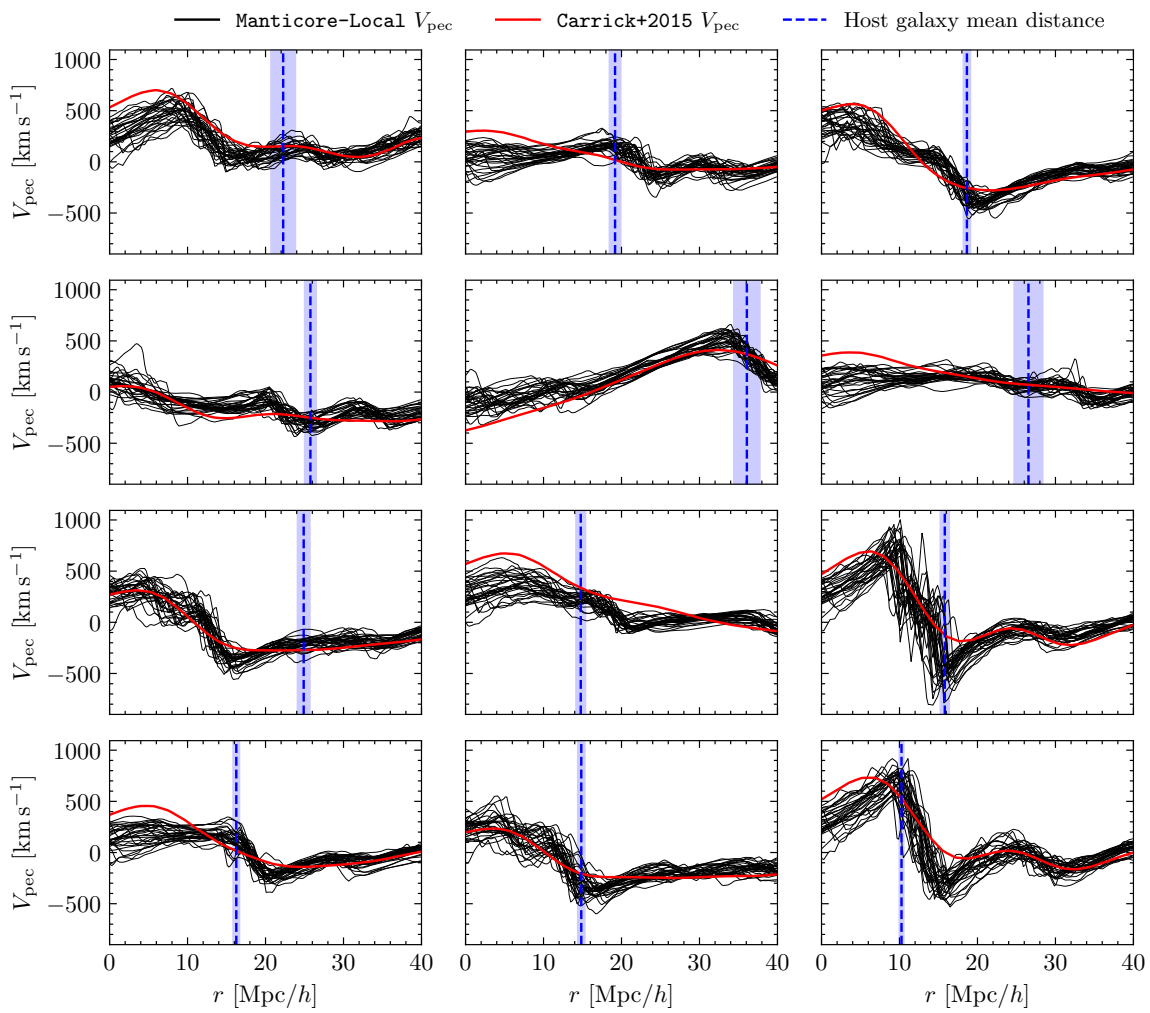


Figure 4.13: Comparison of radial peculiar velocities, $V_{\rm pec}$, between `Manticore-Local` (black) and Carrick et al. (2015) (red) along the LOS to 12 randomly selected Cepheid host galaxies, whose mean inferred distances are shown by the blue vertical lines with shaded bands indicating 1σ uncertainty. Each panel shows 30 `Manticore-Local` realisations, representing independent draws from the BORG posterior, while Carrick et al. (2015) provides only a mean field estimate. The two reconstructions are generally in good agreement. In several cases, the host galaxies reside in clusters, as indicated by the characteristic infall signature in $V_{\rm pec}$.

where $f \approx \Omega_{\rm m}^{0.55}$, $\sigma_{8,g} = 0.99 \pm 0.04$ (Carrick et al., 2015) is the galaxy-field fluctuation amplitude, and $\sigma_{8,\rm NL}$ that of the non-linear matter field. Since the reconstruction is linear, we adopt a linear bias model

$$n_g(r) = 1 + b_1 \delta(r), \tag{4.35}$$

where b_1 is the linear bias parameter. To maintain consistency with the earlier definition of β, we set $b_1 \equiv f/\beta$.

4.C Theory peculiar velocity covariance

The peculiar velocity covariance matrix quantifies correlations in LOS velocities induced by large-scale structure (Kaiser, 1988; Hui & Greene, 2006; Davis et al., 2011). For galaxies indexed by i and j, it is defined as

$$\xi_{ij} = \left\langle (\boldsymbol{v}_i \cdot \hat{\boldsymbol{r}}_i)(\boldsymbol{v}_j \cdot \hat{\boldsymbol{r}}_j) \right\rangle, \tag{4.36}$$

and is computed from the ΛCDM matter power spectrum as

$$\xi_{ij} = \frac{\mathrm{d}D_i}{d\tau}\frac{\mathrm{d}D_j}{d\tau}\int \frac{\mathrm{d}k}{2\pi^2} P(k) F(kr_i,\, kr_j,\, \hat{\boldsymbol{r}}_i \cdot \hat{\boldsymbol{r}}_j), \tag{4.37}$$

where D_i is the linear growth factor at the redshift of the $i^{\rm th}$ galaxy, and τ is the conformal time with $d\tau \equiv dt/a$. F is defined by

$$F(u, v, \cos\theta) = \sum_{\ell=0}^{\infty} (2\ell + 1)\, j_\ell'(u)\, j_\ell'(v)\, P_\ell(\cos\theta) \tag{4.38}$$

where $j_\ell'(x)$ is the derivative of the spherical Bessel function of order ℓ with respect to its argument and $P_\ell(x)$ is the Legendre polynomial of order ℓ. We evaluate Eq. (4.37) using the non-linear matter power spectrum $P(k)$ from `CAMB` (Lewis & Challinor, 2011) over the interval $k \in [10^{-4}, 20]$ h/Mpc, assuming the fiducial *Planck* cosmology (Planck

Collaboration et al., 2020b). We verify that changing the background cosmology, including variations in the assumed value of H_0 when computing $P(k)$ (e.g. switching from the *Planck* to the SH0ES value), does not affect the inferred value of H_0. We truncate the sum in Eq. (4.38) at $\ell = 2000$, which is sufficient for convergence.

For the special case of considering the same object so that $\cos\theta = 1$ and $r_i = r_j = r$, we have that $F(kr, kr, 1) = 1/3$, yielding the diagonal elements

$$\xi_{ii} = \left(\frac{\mathrm{d}D_i}{\mathrm{d}\tau}\right)^2 \int \frac{\mathrm{d}k}{6\pi^2} P(k). \tag{4.39}$$

We find a value of approximately $255\,\mathrm{km\,s^{-1}}$ for our choice of the non-linear matter power spectrum $P(k)$. Adopting a linear $P(k)$ instead yields diagonal elements of approximately $200\,\mathrm{km\,s^{-1}}$. This also contrasts with K22, who incorrectly claim that the diagonal velocity dispersion is approximately $380\,\mathrm{km\,s^{-1}}$ when computed using a non-linear matter power spectrum, leading to an overestimation of their uncertainty. We have managed to reproduce this number by using inconsistent units for $P(k)$ and k.

4.D Cepheid host properties

In Table 4.4 we list the Cepheid host galaxy Galactic coordinates, CMB-frame redshifts, distance moduli, and peculiar velocities as inferred from both `Manticore-Local` and C15. For both inferences, we assume SN magnitude selection. Comparing their peculiar velocities, we find the two models to be generally consistent. The `Manticore-Local` peculiar velocities assume $\beta = 1$, include the sampled $\boldsymbol{V}_{\mathrm{ext}}$, and are computed by stacking the velocities from each Markov chain Monte Carlo (MCMC) step and across the 30 `Manticore-Local` realisations. In contrast, the peculiar velocities from C15 are computed from the mean reconstructed velocity field while jointly sampling β and $\boldsymbol{V}_{\mathrm{ext}}$; their quoted uncertainties therefore represent only the posterior width in those parameters, not of the large-scale structure reconstruction uncertainty. As a comparison, the `Manticore-Local` model yields

a residual velocity scatter of $\sigma_v = 148.2 \pm 31.7\,\mathrm{km\,s^{-1}}$, whereas the C15 model gives $\sigma_v = 221.1 \pm 33.0\,\mathrm{km\,s^{-1}}$.

Name	ℓ [deg]	b [deg]	cz_{CMB} [$\mathrm{km\,s^{-1}}$]	μ_{host} [mag]	$V_{\mathrm{pec}}^{\texttt{Manticore-Local}}$ [$\mathrm{km\,s^{-1}}$]	$V_{\mathrm{pec}}^{\mathrm{C15}}$ [$\mathrm{km\,s^{-1}}$]
M101	102	60	366	$29.17^{+0.04}_{-0.04}$	-81^{+99}_{-84}	-14^{+73}_{-66}
M1337	303	53	2896	$32.91^{+0.07}_{-0.07}$	155^{+73}_{-73}	105^{+59}_{-55}
N0691	141	-39	2581	$32.76^{+0.07}_{-0.08}$	147^{+82}_{-72}	191^{+64}_{-67}
N1015	172	-54	2401	$32.55^{+0.06}_{-0.06}$	29^{+76}_{-66}	-2^{+55}_{-57}
N1309	202	-53	2003	$32.51^{+0.05}_{-0.05}$	-180^{+99}_{-105}	-155^{+59}_{-56}
N1365	238	-55	1379	$31.36^{+0.05}_{-0.05}$	-17^{+150}_{-108}	-46^{+65}_{-62}
N1448	252	-51	1097	$31.28^{+0.04}_{-0.04}$	-46^{+132}_{-129}	-50^{+68}_{-67}
N1559	274	-41	1304	$31.46^{+0.05}_{-0.05}$	-93^{+108}_{-94}	-76^{+75}_{-78}
N2442	281	-22	1544	$31.46^{+0.06}_{-0.06}$	166^{+80}_{-88}	140^{+80}_{-82}
N2525	232	11	1787	$32.00^{+0.07}_{-0.07}$	25^{+81}_{-79}	-65^{+74}_{-78}
N2608	195	34	2386	$32.70^{+0.06}_{-0.06}$	-191^{+78}_{-78}	-298^{+79}_{-81}
N3021	192	51	1775	$32.37^{+0.10}_{-0.09}$	-355^{+102}_{-91}	-394^{+78}_{-75}
N3147	136	39	2977	$33.18^{+0.07}_{-0.07}$	-294^{+94}_{-89}	-319^{+89}_{-78}
N3254	200	59	1649	$32.34^{+0.06}_{-0.06}$	-430^{+83}_{-87}	-405^{+71}_{-70}
N3370	225	60	1619	$32.13^{+0.05}_{-0.05}$	-353^{+105}_{-97}	-329^{+64}_{-67}
N3447	228	61	1394	$31.93^{+0.04}_{-0.03}$	-205^{+98}_{-104}	-296^{+64}_{-68}
N3583	158	62	2317	$32.80^{+0.06}_{-0.06}$	-313^{+68}_{-74}	-403^{+75}_{-69}
N3972	139	60	1106	$31.69^{+0.06}_{-0.07}$	-465^{+120}_{-99}	-364^{+79}_{-71}
N3982	139	60	1106	$31.65^{+0.06}_{-0.06}$	-453^{+129}_{-107}	-359^{+78}_{-70}
N4038	287	42	1979	$31.66^{+0.10}_{-0.09}$	340^{+75}_{-86}	350^{+66}_{-59}
N4424	284	71	767	$30.83^{+0.12}_{-0.10}$	452^{+340}_{-298}	401^{+82}_{-84}
N4536	293	65	1049	$30.84^{+0.05}_{-0.05}$	646^{+145}_{-243}	467^{+63}_{-64}
N4639	294	76	1385	$31.79^{+0.08}_{-0.10}$	-414^{+199}_{-125}	-182^{+57}_{-58}
N4680	301	51	2791	$32.73^{+0.11}_{-0.12}$	207^{+77}_{-79}	169^{+58}_{-54}
N5468	335	53	2992	$33.10^{+0.06}_{-0.05}$	28^{+75}_{-68}	-30^{+57}_{-52}
N5584	345	55	1904	$31.81^{+0.05}_{-0.05}$	142^{+99}_{-94}	22^{+56}_{-55}
N5643	321	15	1433	$30.52^{+0.05}_{-0.05}$	351^{+91}_{-105}	451^{+78}_{-79}
N5728	337	38	3148	$33.02^{+0.10}_{-0.07}$	150^{+73}_{-74}	121^{+69}_{-63}
N5861	349	39	2179	$32.28^{+0.08}_{-0.09}$	175^{+84}_{-97}	51^{+70}_{-63}
N5917	355	40	2108	$32.36^{+0.07}_{-0.07}$	1^{+95}_{-100}	-55^{+70}_{-63}
N7250	94	-14	878	$31.52^{+0.10}_{-0.10}$	-375^{+73}_{-70}	-398^{+69}_{-74}
N7329	321	-46	3124	$33.26^{+0.07}_{-0.07}$	79^{+84}_{-89}	11^{+79}_{-90}
N7541	83	-51	2305	$32.60^{+0.07}_{-0.07}$	-18^{+80}_{-94}	-70^{+58}_{-57}
N7678	99	-37	3145	$33.23^{+0.07}_{-0.07}$	75^{+84}_{-74}	88^{+57}_{-64}
U9391	101	53	1991	$32.79^{+0.06}_{-0.06}$	-446^{+81}_{-82}	-385^{+76}_{-71}

Table 4.4: Summary of the 35 Cepheid host galaxies used in our analysis. Columns list galaxy name, Galactic longitude, latitude, CMB-frame redshift, host distance modulus inferred in the `Manticore-Local` analysis, and the predicted line-of-sight peculiar velocities from `Manticore-Local` and C15. Uncertainties are 1σ. All values assume SN magnitude selection.

5

Conclusions

Throughout this thesis, I have worked with field-level inference through the *Bayesian Origin Reconstruction from Galaxies* (BORG) framework and developed novel Bayesian forward-modelling approaches to peculiar velocity inference and the distance ladder. The principles of forward modelling are straightforward, yet the approach remains uncommon. One apparent barrier is the large number of latent variables that a forward model typically introduces, which can be expensive to sample explicitly. In practice, many of these can be marginalised analytically, and those that cannot, often reduce to cheap one-dimensional numerical integrals, such as marginalising over the distance to a source. Computational costs are not the obstacle either: even Python, with just-in-time compilation via JAX or numba, handles such integrals trivially, and C and Fortran, the languages in which most modern astrophysics codes are already built, have long offered the same capability. The more likely explanation is convenience and inertia. The field has grown accustomed to χ^2 methods that are nominally motivated by frequentist statistics, yet are typically not valid frequentist procedures: they rarely involve mock realisations, and the quantity minimised is not, in general, guaranteed to follow a χ^2 distribution. Over the past decades, as Bayesian methods gained popularity in astronomy, the community recognised that the χ^2 statistic resembles a negative Gaussian log-likelihood and began grafting Bayesian language onto an otherwise unchanged procedure. The important point is simple: every Bayesian analysis rests on a generative model of the observations, i.e. a forward model. Constructing such a model

comes naturally when building simulations, and the same procedure should be applied when analysing data. Simulation-based inference (Cranmer et al., 2020), or implicit likelihood inference (Alsing et al., 2019), has embraced this perspective, but the underlying logic is identical for explicit likelihood inference. Both share a generative model, but the latter additionally requires a noise model that admits a closed-form likelihood for comparing observations to model predictions.

In this thesis and in the other papers I have written over the course of my DPhil, I have adopted the forward-modelling approach. The way I approached it has itself evolved over time, as I learnt more about topics such as selection effects. This work is part of a broader theme: a robust and precise inference of the initial conditions of the local Universe. As argued in Chapter 1, uncovering these will tighten cosmological constraints and enable follow-up analyses ranging from tests of galaxy formation and the nature of dark matter (DM) to measurements of the Hubble constant from alternative data. With modelling frameworks such as `BORG`, or any future successors, the local Universe becomes a laboratory for precision cosmology, and the approach can be extended to higher redshift to infer the initial conditions of the observable Universe (Andrews et al., 2026). This target is simple to state but difficult in practice: a self-consistent forward model of the observable Universe, from initial conditions through the primary cosmic microwave background (CMB) anisotropies, large-scale structure and its secondary imprints on the CMB, and weak lensing, down to the population of individual galaxies. Today, the `BORG` inference of `Manticore-Local` is the highest-dimensional robust Bayesian inference ever performed, and it is being expanded to even larger scales within the *Aquila* Consortium. Throughout my DPhil, I have been fortunate to contribute to this effort.

The three chapters build on one another: the first two establish reconstruction precision at the halo level and in the large-scale velocity field, and the third exploits them for a high-precision H_0 measurement from Cepheids alone. Chapter 2 quantifies the precision of halo properties across the `BORG` posterior. Cluster-mass haloes are consistently reconstructed

in position, mass, and bulk velocity, while secondary properties remain unconstrained in the BORG generation used there. Improving them requires advances to both the gravity solver and the galaxy bias model. Chapter 3 compares six local-Universe velocity field reconstructions against peculiar velocity data and establishes the first metric for such a comparison, demonstrating that the latest BORG models are strongly preferred. Chapter 4 constructs a two-rung distance ladder bypassing supernovae (SNe) entirely, leading to a 3σ tension with the CMB. Here, the BORG modelling was crucial to suppress cosmic variance and demonstrate that with a high-fidelity local Universe model, even a small sample of distance tracers is sufficient for precise cosmological measurements.

Three themes run through the thesis: field-level inference of initial conditions, Bayesian forward modelling, and a rigorous treatment of selection effects. Field-level reconstructions of the local Universe are increasingly mature, improving on a year-by-year basis, and ripe to augment follow-up analyses. Bayesian forward modelling treats every problem from the perspective of a generative model of the data, and the latter two threads are closely linked: selection is often much simpler to encode in a generative model than to account for after the fact, and mismodelling selection effects can produce large biases in mock analyses, as was often my experience. A further recurring point is the importance of marginalising over all latent variables within a Bayesian hierarchical framework, rather than fixing them to fiducial values or treating them as perfectly known. For example, mapping observed redshifts deterministically to distances neglects both peculiar velocities and redshift measurement uncertainty, and can lead to biased inferences unless the prior and selection function are carefully modelled.

Several concrete steps follow, which I intend to pursue in the coming years. The first is to push the BORG inference framework to GPUs and speed up the development cycle. At present, running a single `Manticore-Local` chain takes several weeks on a few hundred CPUs and costs close to a million CPU hours. This is achievable and does not pose a fundamental difficulty, but it implies a slow development process; porting the BORG code to

(multi-)GPUs and emulating the gravity model are necessary steps that I have already begun. With this in hand, my scientific goal is to extend the BORG inference beyond the current redshift-space data to jointly model peculiar velocities and constrain initial conditions, using the modelling framework developed throughout my DPhil. This will inform the inferred galaxy bias (though the modelling of it will also need to be improved), and after that I intend to move towards a field-level cosmological inference with BORG, rather than assuming a fixed cosmological model. A longer-term goal is to couple BORG to the CMB. I took a first step in Stiskalek & Desmond (2026), benchmarking BORG posterior fields against *Planck* thermal Sunyaev–Zel'dovich maps and using them to calibrate the cluster mass–observable relation. The next stage is to forward-model the thermal Sunyaev–Zel'dovich signal from BORG posterior fields to the map level, using baryonification to paint baryonic gas profiles onto the dark matter structures (e.g. Schneider & Teyssier, 2015; Aricò et al., 2020). Ultimately, this would yield a forward model that evolves primordial fluctuations through gravitational dynamics, models galaxy formation and the baryonic gas distribution, and predicts both the observed galaxy field with its peculiar velocities and the secondary imprints of large-scale structure on the CMB. Such a model would constrain cosmology and the initial conditions simultaneously.

The second step, which I have already started, is to extend the hierarchical forward-modelling framework to the third rung of the distance ladder, introducing a new distance ladder inference that marginalises over all latent variables and their systematics. These and other advances will be driven by the rapidly expanding observational programme. Spectroscopic surveys, SNe samples, and peculiar velocity catalogues are all growing by orders of magnitude. This presents both an opportunity and a challenge: the sample sizes will deliver high precision, but ensuring accuracy will demand ever more rigorous modelling.

Delivering that accuracy requires considering the limitations of current methods. Several limitations of the BORG framework have surfaced throughout this thesis, and it is worth

mentioning the most important ones here. The first is the dependence on an assumed cosmological model and a fixed set of cosmological parameters. This is a practical rather than fundamental limitation, and for the H_0 inference presented in this thesis, the assumed values in the reconstruction have been shown to have a minimal effect (Tsaprazi & Heavens, 2025). Relaxing this assumption, as discussed above, is a natural next step once field-level cosmological inference becomes feasible.

Potentially more consequential is the modelling of galaxy bias. In the latest BORG generation, the expected galaxy count in a voxel is a truncated double power law of the local DM density, an approximation that does not hold in practice. Bartlett et al. (2024b) showed that the relation between local DM density and voxel halo count is a non-local function of density, even on scales much larger than 10 h^{-1} Mpc. There are several ways to address this, and it is the limitation I am most keen to tackle. Small convolutional neural networks that pull information from neighbouring voxels can predict the expected galaxy count of a central voxel, and this has been demonstrated on mock data within BORG (Ding et al., 2024). A more ambitious route would be to run higher-resolution simulations within the BORG inference and apply a differentiable halo finder to identify haloes directly in the forward model (Horowitz & Bayer, 2026), reducing the bias relation to a simpler halo–galaxy mapping. The difficulty is that FoF haloes are discrete particle linkages, so propagating gradients through the halo finder yields unstable gradients that would drive any Hamiltonian Monte Carlo (HMC) sampler towards a vanishing step size (and similarly for other halo definitions). One way around this is to remove the requirement for gradient-based sampling altogether. Doeser et al. (2025) proposed exactly this: although currently limited to optimisation rather than full posterior sampling, the approach allows coupling a non-differentiable forward model such as GADGET4 to a BORG-like inference, converging within $\mathcal{O}(10)$ optimiser steps and, in principle, enabling initial-condition inference with a state-of-the-art hydrodynamical simulation in the loop. One could also replace the explicit likelihood formulation entirely and use diffusion models to learn the mapping between initial conditions and the observed

galaxy field (e.g. Legin et al., 2024). This has been shown to be feasible in simulations, though current diffusion models recover the mean field well but struggle to capture the full posterior distribution.

Turning to the distance-ladder work in the second half of this thesis, I would argue that the dominant limitation is the treatment of selection effects, and the difficulty is twofold. For catalogues compiled over many years, such as `CosmicFlows-4`, the selection function is not well characterised, precluding rigorous selection modelling and any precision cosmological inference. Conversely, even when the selection is known in principle, it can be complex—varying on a night-by-night basis with telescope conditions—and therefore difficult to encode in a forward model. If the selection cannot be modelled sufficiently well, then any inference with such data will be subject to significant systematic uncertainty. This is particularly acute for SNe: the standardisation must be modelled, and the selection function for H_0 samples is complex because they are composed of heterogeneous surveys, collected over decades with different instruments. Bayesian forward models of SN light curves (Mandel et al., 2022) or simulation-based inference approaches (Karchev et al., 2026) will be crucial to this effort, as will scaling any such treatment to the large number of SNe expected from future surveys. More broadly, this calls for greater cooperation between observers and modellers: the selection function must be well characterised at the observing stage so that it can be incorporated into inference, rather than reconstructed post hoc without a clearly defined algorithmic procedure. While there are clear procedures for modelling selection explicitly within a Bayesian inference, as pursued here, or through simulation-based approaches on mock data, neither can help if the telescope selection itself is insufficiently documented.

These considerations bear directly on the Hubble tension. It is now the best-established tension in modern cosmology, and no convincing theoretical resolution is in sight. Regardless, the possibility of systematics in the local distance ladder, whether in the data or in the modelling, must be taken seriously. Beyond the astrophysical systematics commonly

discussed, the statistical treatment itself can be a source of bias: approximations in the likelihood, neglected covariances, unmodelled selection effects, and the fixing of latent variables to point estimates can each shift inferred parameters by comparable amounts. In Chapter 4, I constructed a two-rung distance ladder using Cepheid host galaxies to infer H_0, bypassing SNe. Selection is again immediate: the unknown selection function of the Cepheid host galaxies, targeted because they hosted an SN, contributes a $\sim 1\sigma$ systematic uncertainty. The selection is complex, involving a dual process—first the discovery of the SN and then the follow-up for Cepheid photometry—that is nearly impossible to model precisely for transient catalogues collected over decades. The results of Chapter 4 are mildly lower than the SH0ES measurement, hinting at a small reduction of the tension.[1]

If one takes a sceptical view of SN-related systematics, the next step is to continue inferring H_0 without SNe, using as rigorous a modelling framework as possible while admitting that the selection function will sometimes be imprecisely known. Two possibilities: the Tip of the Red Giant Branch (TRGB) sample of Anand et al. (2022), with nearly 500 sources all within $\sim$10 Mpc and so highly sensitive to peculiar velocity modelling; and the megamaser sample of Pesce et al. (2020b), which yields H_0 in agreement with SH0ES but with an unmodelled selection function, likely offset by its uniform-in-distance prior. Upcoming data will improve second-rung precision (Anand et al., 2025), but even with present samples, I expect these H_0 measurements to remain in tension with the CMB, albeit with a smaller offset from *Planck* than SH0ES. A rigorous forward model of the distance ladder, including the third rung, remains an open question.

As cosmology enters a data-rich regime, ever more precise constraints on cosmological parameters will follow, along with many spurious tensions. Gaining confidence in such tensions will require modern, rigorous statistical frameworks, at whose heart lies the forward modelling approach.

[1] The results in Chapter 4 are published, though I later found a small mismodelling; the correction shifts H_0 lower by less than 1σ.

Bibliography

Aaronson M., Huchra J., Mould J., Schechter P. L., Tully R. B., 1982, ApJ, 258, 64

Abazajian K. N., et al., 2009, ApJS, 182, 543

Abbott B. P., et al., 2017, Nature, 551, 85

Abbott B. P., et al., 2018, Living Reviews in Relativity, 21, 3

Abbott T. M. C., et al., 2020, Phys. Rev. D, 102, 023509

Abbott B. P., et al., 2021, ApJ, 909, 218

Abbott T. M. C., et al., 2022, Phys. Rev. D, 105, 023520

Abbott R., et al., 2023, ApJ, 949, 76

Abdalla E., et al., 2022, Journal of High Energy Astrophysics, 34, 49

Abdul Karim M., et al., 2025, Phys. Rev. D, 112, 083515

Adame A. G., et al., 2025a, J. Cosmology Astropart. Phys., 2025, 021

Adame A. G., et al., 2025b, J. Cosmology Astropart. Phys., 2025, 028

Adams C., Blake C., 2017, MNRAS, 471, 839

Adams C., Blake C., 2020, MNRAS, 494, 3275

Akeson R., et al., 2019, arXiv e-prints, p. arXiv:1902.05569

Alsing J., Charnock T., Feeney S., Wandelt B., 2019, MNRAS, 488, 4440

Amenouche M., et al., 2025, A&A, 694, A3

Amon A., et al., 2022, Phys. Rev. D, 105, 023514

Anand G. S., Tully R. B., Rizzi L., Riess A. G., Yuan W., 2022, ApJ, 932, 15

Anand G., Durbin M., Beaton R., Jensen J., Riess A., 2025, arXiv e-prints, p. arXiv:2507.02056

Andrews A., Loureiro A., Jasche J., McAlpine S., Lavaux G., Leclercq F., 2026, arXiv e-prints, p. arXiv:2602.02363

Angulo R. E., Hahn O., 2022, Living Reviews in Computational Astrophysics, 8, 1

Aricò G., Angulo R. E., Hernández-Monteagudo C., Contreras S., Zennaro M., Pellejero-Ibañez M., Rosas-Guevara Y., 2020, MNRAS, 495, 4800

Astropy Collaboration et al., 2013, A&A, 558, A33

Astropy Collaboration et al., 2018, AJ, 156, 123

Astropy Collaboration et al., 2022, ApJ, 935, 167

Bahr-Kalus B., Bertacca D., Verde L., Heavens A., 2021, J. Cosmology Astropart. Phys., 2021, 027

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barnes J., Hut P., 1986, Nature, 324, 446

Bartlett D. J., Desmond H., 2023, The Open Journal of Astrophysics, 6, 42

Bartlett D. J., Desmond H., Ferreira P. G., 2021, Phys. Rev. D, 103, 023523

Bartlett D. J., Kostić A., Desmond H., Jasche J., Lavaux G., 2022, Phys. Rev. D, 106, 103526

Bartlett D. J., Wandelt B. D., Zennaro M., Ferreira P. G., Desmond H., 2024a, A&A, 686, A150
Bartlett D. J., Ho M., Wandelt B. D., 2024b, ApJL, 977, L44
Bernardo J. M., 1979, Journal of the Royal Statistical Society. Series B (Methodological), 41, 113
Bertschinger E., 1987, ApJL, 323, L103
Bertschinger E., 1998, ARA&A, 36, 599
Bertschinger E., Jain B., 1994, ApJ, 431, 486
Bhardwaj A., et al., 2023, ApJL, 955, L13
Birrer S., et al., 2020, A&A, 643, A165
Blake C., et al., 2011, MNRAS, 415, 2876
Blakeslee J. P., Jensen J. B., Ma C.-P., Milne P. A., Greene J. E., 2021, ApJ, 911, 65
Blas D., Lesgourgues J., Tram T., 2011, J. Cosmology Astropart. Phys., 2011, 034
Bocquet S., et al., 2019, ApJ, 878, 55
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Boruah S. S., Hudson M. J., Lavaux G., 2020, MNRAS, 498, 2703
Boruah S. S., Hudson M. J., Lavaux G., 2021, MNRAS, 507, 2697
Boruah S. S., Lavaux G., Hudson M. J., 2022, MNRAS, 517, 4529
Boubel P., Colless M., Said K., Staveley-Smith L., 2024, MNRAS, 531, 84
Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, A&A, 296, 575
Boyd B. M., Grayling M., Thorp S., Mandel K. S., 2024, arXiv e-prints, p. arXiv:2407.15923
Boyd B. M., et al., 2026, MNRAS, 550, stag1046
Braatz J. A., Reid M. J., Humphreys E. M. L., Henkel C., Condon J. J., Lo K. Y., 2010, ApJ, 718, 657
Bradbury J., et al., 2021, JAX: Autograd and XLA, Astrophysics Source Code Library, record ascl:2111.002 (ascl:2111.002)
Bresolin F., Kudritzki R.-P., Urbaneja M. A., Gieren W., Ho I.-T., Pietrzyński G., 2016, ApJ, 830, 64
Breuval L., et al., 2023, ApJ, 951, 118
Breuval L., et al., 2024, ApJ, 973, 30
Brout D., et al., 2022, ApJ, 938, 110
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Bryan G. L., et al., 2014, ApJS, 211, 19
Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343
Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, ApJ, 555, 240
Burns C. R., et al., 2018, ApJ, 869, 56
Butsky I., et al., 2016, MNRAS, 462, 663
Cadiou C., Pontzen A., Peiris H. V., 2021, MNRAS, 502, 5480
Calabrese E., Slosar A., Melchiorri A., Smoot G. F., Zahn O., 2008, Phys. Rev. D, 77, 123531
Campbell L. A., et al., 2014, MNRAS, 443, 1231
Camphuis E., et al., 2026, Phys. Rev. D, 113, 083504

Cantiello M., Blakeslee J. P., 2023, arXiv e-prints, p. arXiv:2307.03116

Carr A., Davis T. M., Scolnic D., Said K., Brout D., Peterson E. R., Kessler R., 2022, PASA, 39, e046

Carreres B., et al., 2025, A&A, 694, A8

Carrick J., Turnbull S. J., Lavaux G., Hudson M. J., 2015, MNRAS, 450, 317

Carron J., Neyrinck M. C., 2012, ApJ, 750, 28

Cataldi P., Pedrosa S. E., Tissera P. B., Artale M. C., 2021, MNRAS, 501, 5679

Chevallier M., Polarski D., 2001, International Journal of Modern Physics D, 10, 213

Colless M., et al., 2001, MNRAS, 328, 1039

Colombi S., Chodorowski M. J., Teyssier R., 2007, MNRAS, 375, 348

Courtois H. M., Dupuy A., Guinet D., Baulieu G., Ruppin F., Brenas P., 2023, A&A, 670, L15

Cranmer K., Brehmer J., Louppe G., 2020, Proceedings of the National Academy of Science, 117, 30055

Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369

DES Collaboration et al., 2024, ApJL, 973, L14

DESI Collaboration et al., 2024, AJ, 168, 58

Da Costa G. S., Armandroff T. E., 1990, AJ, 100, 162

Dálya G., et al., 2022, MNRAS, 514, 1403

Dark Energy Survey and Kilo-Degree Survey Collaboration et al., 2023, The Open Journal of Astrophysics, 6, 36

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Davis T. M., et al., 2011, ApJ, 741, 67

Dehnen W., Read J. I., 2011, European Physical Journal Plus, 126, 55

Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1

Desmond H., 2023, MNRAS, 526, 3342

Desmond H., Ferreira P. G., 2020, Phys. Rev. D, 102, 104060

Desmond H., Mao Y.-Y., Wechsler R. H., Crain R. A., Schaye J., 2017, MNRAS, 471, L11

Desmond H., Hutt M. L., Devriendt J., Slyz A., 2022, MNRAS, 511, L45

Desmond H., Stiskalek R., Nájera J. A., Banik I., 2026, MNRAS,

Di Valentino E., et al., 2021, Classical and Quantum Gravity, 38, 153001

Di Valentino E., et al., 2025, Physics of the Dark Universe, 49, 101965

Dickey J. M., 1971, The Annals of Mathematical Statistics, 42, 204

Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108

Ding S., Lavaux G., Jasche J., 2024, A&A, 690, A236

Djorgovski S., Davis M., 1987, ApJ, 313, 59

Dodelson S., Schmidt F., 2020, Modern Cosmology, doi:10.1016/C2017-0-01943-2.

Doeser L., Ata M., Jasche J., 2025, MNRAS, 542, 1403

Dolag K., Sorce J. G., Pilipenko S., Hernández-Martínez E., Valentini M., Gottlöber S., Aghanim N., Khabibullin I., 2023, A&A, 677, A169

Douglass K., et al., 2025, arXiv e-prints, p. arXiv:2512.03227

Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, ApJ, 313, 42

Duane S., Kennedy A. D., Pendleton B. J., Roweth D., 1987, Physics Letters B, 195, 216

Dupuy A., Courtois H. M., Kubik B., 2019, MNRAS, 486, 440

Dvali G., Gabadadze G., Porrati M., 2000, Physics Letters B, 485, 208

Efstathiou G., 2025, MNRAS, 538, 875

Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, ApJS, 57, 241

Efstathiou G., Kaiser N., Saunders W., Lawrence A., Rowan-Robinson M., Ellis R. S., Frenk C. S., 1990, MNRAS, 247, 10P

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Eisenstein D. J., Hut P., 1998, ApJ, 498, 137

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Ellis G. F. R., Baldwin J. E., 1984, MNRAS, 206, 377

Euclid Collaboration et al., 2025, A&A, 697, A1

Faber S. M., Jackson R. E., 1976, ApJ, 204, 668

Feeney S. M., Mortlock D. J., Dalmasso N., 2018, MNRAS, 476, 3861

Feix M., Branchini E., Nusser A., 2017, MNRAS, 468, 1420

Feldman H. A., Watkins R., Hudson M. J., 2010, MNRAS, 407, 2328

Foley R. J., et al., 2018, MNRAS, 475, 193

Freedman W. L., 2021, ApJ, 919, 16

Freedman W. L., Madore B. F., 2020, ApJ, 899, 67

Freedman W. L., et al., 2019, ApJ, 882, 34

Freedman W. L., Madore B. F., Hoyt T. J., Jang I. S., Lee A. J., Owens K. A., 2025, ApJ, 985, 203

Gaia Collaboration et al., 2023, A&A, 674, A1

Gall C., Izzo L., Wojtak R., Hjorth J., 2024, arXiv e-prints, p. arXiv:2411.05642

Gallagher C., Yasin T., Stiskalek R., Desmond H., Jarvis M. J., 2026, MNRAS, 546, stag108

Ganeshaiah Veena P., Lilow R., Nusser A., 2023, MNRAS, 522, 5291

Ganeshalingam M., Li W., Filippenko A. V., 2013, MNRAS, 433, 2240

Ganon G., Hoffman Y., 1993, ApJL, 415, L5

Gao F., et al., 2016, ApJ, 817, 128

García-García C., Zennaro M., Aricò G., Alonso D., Angulo R. E., 2024, J. Cosmology Astropart. Phys., 2024, 024

Gelman A., Rubin D. B., 1992, Statistical Science, 7, 457

Gelman A., Yao Y., 2021, Journal of Physics G Nuclear Physics, 48, 014002

Ghirardini V., et al., 2024, A&A, 689, A298

Ginolin M., et al., 2025, A&A, 695, A140

Gottloeber S., Hoffman Y., Yepes G., 2010, arXiv e-prints, p. arXiv:1005.2687

Gray R., et al., 2023, J. Cosmology Astropart. Phys., 2023, 023

Grayling M., Thorp S., Mandel K. S., Dhawan S., Uzsoy A. S. M., Boyd B. M., Hayes E. E., Ward S. M., 2024, MNRAS, 531, 953

Graziani R., Courtois H. M., Lavaux G., Hoffman Y., Tully R. B., Copin Y., Pomarède D., 2019, MNRAS, 488, 5438

Griffen B. F., Ji A. P., Dooley G. A., Gómez F. A., Vogelsberger M., O'Shea B. W., Frebel A., 2016, ApJ, 818, 10

Gunes Baydin A., Pearlmutter B. A., Andreyevich Radul A., Siskind J. M., 2015, arXiv e-prints, p. arXiv:1502.05767

Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1

Guth A. H., Pi S.-Y., 1982, Phys. Rev. Lett., 49, 1110

Guy J., et al., 2007, A&A, 466, 11

Guzzo L., et al., 2008, Nature, 451, 541

H0DN Collaboration et al., 2026, A&A, 708, A166

Hastings W. K., 1970, Biometrika, 57, 97

Hawking S. W., 1982, Physics Letters B, 115, 295

Hellwing W. A., Bilicki M., Libeskind N. I., 2018, Phys. Rev. D, 97, 103519

Hergt L. T., Handley W. J., Hobson M. P., Lasenby A. N., 2021, Phys. Rev. D, 103, 123511

Herrnstein J. R., et al., 1999, Nature, 400, 539

Heß S., Kitaura F.-S., Gottlöber S., 2013, MNRAS, 435, 2065

Heymans C., et al., 2021, A&A, 646, A140

Hinshaw G., et al., 2009, ApJS, 180, 225

Hoffman M. D., Gelman A., 2011, arXiv e-prints, p. arXiv:1111.4246

Hoffman Y., Ribak E., 1991, ApJL, 380, L5

Hoffman Y., Courtois H. M., Tully R. B., 2015, MNRAS, 449, 4494

Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, ApJ, 624, 54

Hollinger A. M., Hudson M. J., 2024, MNRAS, 531, 788

Hollinger A. M., Hudson M. J., 2025, J. Cosmology Astropart. Phys., 2025, 021

Hong T., et al., 2019, MNRAS, 487, 2061

Hopkins P. F., 2015, MNRAS, 450, 53

Horowitz B., Bayer A. E., 2026, MNRAS,

Howlett C., Said K., Lucey J. R., Colless M., Qin F., Lai Y., Tully R. B., Davis T. M., 2022, MNRAS, 515, 953

Hoyt T. J., Jang I. S., Freedman W. L., Madore B. F., Owens K. A., Lee A. J., 2026, ApJ, 1002, 11

Huang C. D., et al., 2024, ApJ, 963, 83

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huchra J. P., et al., 2012, ApJS, 199, 26

Hudson R. S., Ostro S. J., 1994, Science, 263, 940

Hui L., Greene P. B., 2006, Phys. Rev. D, 73, 123526

Humphreys E. M. L., Reid M. J., Moran J. M., Greenhill L. J., Argon A. L., 2013, ApJ, 775, 13

Huterer D., Shafer D. L., Schmidt F., 2015, J. Cosmology Astropart. Phys., 2015, 033

Huterer D., Shafer D. L., Scolnic D. M., Schmidt F., 2017, J. Cosmology Astropart. Phys., 2017,

015

Hutt M. L., Desmond H., Devriendt J., Slyz A., 2022, MNRAS, 516, 3592

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jackson J. C., 1972, MNRAS, 156, 1P

Jasche J., Lavaux G., 2019, A&A, 625, A64

Jasche J., Wandelt B. D., 2013, MNRAS, 432, 894

Jasche J., Leclercq F., Wandelt B. D., 2015, J. Cosmology Astropart. Phys., 2015, 036

Jaynes E. T., 2003, Probability Theory: The Logic of Science. Cambridge University Press, doi:10.1017/CBO9780511790423

Jeffreys H., 1939, Theory of Probability

Jensen J. B., et al., 2021, ApJS, 255, 21

Jensen J. B., Blakeslee J. P., Cantiello M., Cowles M., Anand G. S., Tully R. B., Kourkchi E., Raimondo G., 2025, ApJ, 987, 87

Jones D. H., et al., 2009, MNRAS, 399, 683

Jones D. O., et al., 2019, ApJ, 881, 19

Juszkiewicz R., Feldman H. A., Fry J. N., Jaffe A. H., 2010, J. Cosmology Astropart. Phys., 2010, 021

Kaiser N., 1984, ApJL, 284, L9

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., 1988, MNRAS, 231, 149

Karchev K., Trotta R., Jiménez R., 2026, Nature Astronomy,

Keller B. W., Wadsley J. W., 2017, ApJL, 835, L17

Kelly B. C., Fan X., Vestergaard M., 2008, ApJ, 682, 874

Kenworthy W. D., et al., 2022, ApJ, 935, 83

Kessler R., Scolnic D., 2017, ApJ, 836, 56

Kitaura F.-S., 2013, MNRAS, 429, L84

Klypin A., Hoffman Y., Kravtsov A. V., Gottlöber S., 2003, ApJ, 596, 19

Knebe A., et al., 2013, MNRAS, 435, 1618

Koribalski B. S., et al., 2020, Ap&SS, 365, 118

Kostić A., Bartlett D. J., Desmond H., 2026, Phys. Rev. D, 113, 063539

Kourkchi E., Tully R. B., Neill J. D., Seibert M., Courtois H. M., Dupuy A., 2019, ApJ, 884, 82

Kourkchi E., Tully R. B., Anand G. S., Courtois H. M., Dupuy A., Neill J. D., Rizzi L., Seibert M., 2020a, ApJ, 896, 3

Kourkchi E., et al., 2020b, ApJ, 902, 145

Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73

Kravtsov A. V., Klypin A., Hoffman Y., 2002, ApJ, 571, 563

Kuo C. Y., Braatz J. A., Reid M. J., Lo K. Y., Condon J. J., Impellizzeri C. M. V., Henkel C., 2013, ApJ, 767, 155

Lane Z. G., Seifert A., Ridden-Harper R., Wiltshire D. L., 2025, MNRAS, 536, 1752

Lavaux G., 2016, MNRAS, 457, 172

Lavaux G., Hudson M. J., 2011, MNRAS, 416, 2840

Lavaux G., Jasche J., 2016, MNRAS, 455, 3169

Lavaux G., Jasche J., Leclercq F., 2019, arXiv e-prints, p. arXiv:1909.06396

Leavitt H. S., Pickering E. C., 1912a, Harvard College Observatory Circular, 173, 1

Leavitt H. S., Pickering E. C., 1912b, Harvard College Observatory Circular, 173, 1

Leclercq F., Jasche J., Lavaux G., Wandelt B., Percival W., 2017, J. Cosmology Astropart. Phys., 2017, 049

Lee A. J., Freedman W. L., Madore B. F., Jang I. S., Owens K. A., Hoyt T. J., 2025, ApJ, 985, 182

Legin R., Ho M., Lemos P., Perreault-Levasseur L., Ho S., Hezaveh Y., Wandelt B., 2024, MNRAS, 527, L173

Lewis A., Challinor A., 2011, CAMB: Code for Anisotropies in the Microwave Background, Astrophysics Source Code Library, record ascl:1102.026 (ascl:1102.026)

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Li W. D., et al., 2000, in Holt S. S., Zhang W. W., eds, American Institute of Physics Conference Series Vol. 522, Cosmic Explosions: Tenth AstroPhysics Conference. AIP, pp 103–106 (arXiv:astro-ph/9912336), doi:10.1063/1.1291702

Li S., Riess A. G., Busch M. P., Casertano S., Macri L. M., Yuan W., 2021, ApJ, 920, 84

Li S., et al., 2024, ApJ, 976, 177

Lilow R., Nusser A., 2021, MNRAS, 507, 1557

Lilow R., Ganeshaiah Veena P., Nusser A., 2024, A&A, 689, A226

Linder E. V., 2003, Phys. Rev. Lett., 90, 091301

Linder E. V., 2005, Phys. Rev. D, 72, 043529

Linder E. V., Cahn R. N., 2007, Astroparticle Physics, 28, 481

Louis T., et al., 2025, J. Cosmology Astropart. Phys., 2025, 062

Ludlow A. D., et al., 2017, Phys. Rev. Lett., 118, 161103

Lukić Z., Reed D., Habib S., Heitmann K., 2009, ApJ, 692, 217

Lynden-Bell D., 1967, MNRAS, 136, 101

Lynden-Bell D., Faber S. M., Burstein D., Davies R. L., Dressler A., Terlevich R. J., Wegner G., 1988, ApJ, 326, 19

Ma C.-P., Bertschinger E., 1995, ApJ, 455, 7

Madhavacheril M. S., et al., 2024, ApJ, 962, 113

Madore B. F., 1982, ApJ, 253, 575

Malandrino R., Lavaux G., Wandelt B. D., McAlpine S., Jasche J., 2026, A&A, 705, A160

Malmquist K. G., 1922, Meddelanden fran Lunds Astronomiska Observatorium Serie I, 100, 1

Mandel K. S., Scolnic D. M., Shariff H., Foley R. J., Kirshner R. P., 2017, ApJ, 842, 93

Mandel K. S., Thorp S., Narayan G., Friedman A. S., Avelino A., 2022, MNRAS, 510, 3939

March M. C., Trotta R., Berkes P., Starkman G. D., Vaudrevange P. M., 2011, MNRAS, 418, 2308

March M. C., Karpenka N. V., Feroz F., Hobson M. P., 2014, MNRAS, 437, 3298

March M. C., Wolf R. C., Sako m., D'Andrea C., Brout D., 2018, arXiv e-prints, p. arXiv:1804.02474

Masters K. L., Springob C. M., Haynes M. P., Giovanelli R., 2006, ApJ, 653, 861

Masters K. L., Springob C. M., Huchra J. P., 2008, AJ, 135, 1738

McAlpine S., 2026, The Open Journal of Astrophysics, 9, 59142

McAlpine S., et al., 2022, MNRAS, 512, 5823

McAlpine S., Jasche J., Ata M., Lavaux G., Stiskalek R., Frenk C. S., Jenkins A., 2025, MNRAS, 540, 716

McEwen J. D., Wallis C. G. R., Price M. A., Spurio Mancini A., 2021, arXiv e-prints, p. arXiv:2111.12720

McGaugh S. S., Schombert J. M., Bothun G. D., de Blok W. J. G., 2000, ApJL, 533, L99

McGaugh S. S., Lelli F., Schombert J. M., 2016, Phys. Rev. Lett., 117, 201101

Mead A. J., Heymans C., Lombriser L., Peacock J. A., Steele O. I., Winther H. A., 2016, MNRAS, 459, 1468

Meszaros P., 1974, A&A, 37, 225

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087

Milgrom M., 1983, ApJ, 270, 365

Mitchell P. D., et al., 2018, MNRAS, 474, 492

Miyoshi M., Moran J., Herrnstein J., Greenhill L., Nakai N., Diamond P., Inoue M., 1995, Nature, 373, 127

Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

Monaghan J. J., 1992, ARA&A, 30, 543

Mukhanov V. F., 1985, Soviet Journal of Experimental and Theoretical Physics Letters, 41, 493

Mukhanov V. F., Chibisov G. V., 1981, Soviet Journal of Experimental and Theoretical Physics Letters, 33, 532

Nájera J. A., Desmond H., 2025, MNRAS, 541, 671

Nakamura T. T., Suto Y., 1997, Progress of Theoretical Physics, 97, 49

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., Benítez-Llambay A., Fattahi A., Frenk C. S., Ludlow A. D., Oman K. A., Schaller M., Theuns T., 2017, MNRAS, 471, 1841

Neal R., 2011, in , Handbook of Markov Chain Monte Carlo. pp 113–162, doi:10.1201/b10905

Nesseris S., García-Bellido J., 2013, J. Cosmology Astropart. Phys., 2013, 036

Nusser A., 2016, MNRAS, 455, 178

Nusser A., 2017, MNRAS, 470, 445

Nusser A., Davis M., 2011, ApJ, 736, 93

Nusser A., Davis M., Branchini E., 2014, ApJ, 788, 157

Oayda O. T., Lewis G. F., 2026, MNRAS, 546, stag248

Ong D. D. Y., Yallup D., Handley W., 2025, arXiv e-prints, p. arXiv:2511.10631

Ong D. D. Y., Yallup D., Handley W., 2026, arXiv e-prints, p. arXiv:2603.05472

Onions J., et al., 2012, MNRAS, 423, 1200

Paszke A., et al., 2019, arXiv e-prints, p. arXiv:1912.01703

Peebles P. J. E., 1969, ApJ, 155, 393

Peebles P. J. E., 1980, The large-scale structure of the universe

Peery S., Watkins R., Feldman H. A., 2018, MNRAS, 481, 1368

Perlmutter S., et al., 1999, ApJ, 517, 565

Pesce D. W., et al., 2020a, ApJ, 890, 118

Pesce D. W., et al., 2020b, ApJL, 891, L1

Peterson E. R., et al., 2022, ApJ, 938, 112

Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59

Pfeifer S., Valade A., Gottlöber S., Hoffman Y., Libeskind N. I., Hellwing W. A., 2023, MNRAS, 523, 5985

Phan D., Pradhan N., Jankowiak M., 2019, arXiv e-prints, p. arXiv:1912.11554

Phillips M. M., 1993, ApJL, 413, L105

Pietrzyński G., et al., 2019, Nature, 567, 200

Pike R. W., Hudson M. J., 2005, ApJ, 635, 11

Planck Collaboration et al., 2014, A&A, 571, A1

Planck Collaboration et al., 2016, A&A, 594, A24

Planck Collaboration et al., 2020a, A&A, 641, A1

Planck Collaboration et al., 2020b, A&A, 641, A6

Planck Collaboration et al., 2020c, A&A, 641, A9

Polanska A., Price M. A., Piras D., Spurio Mancini A., McEwen J. D., 2025, The Open Journal of Astrophysics, 8, 156

Pontzen A., Slosar A., Roth N., Peiris H. V., 2016, Phys. Rev. D, 93, 103519

Popovic B., et al., 2026, MNRAS, 548, stag632

Porqueres N., Kodi Ramanah D., Jasche J., Lavaux G., 2019, A&A, 624, A115

Porqueres N., Heavens A., Mortlock D., Lavaux G., 2021, MNRAS, 502, 3035

Porredon A., et al., 2022, Phys. Rev. D, 106, 103530

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14

Press W. H., Schechter P., 1974, ApJ, 187, 425

Prideaux-Ghee J., Leclercq F., Lavaux G., Heavens A., Jasche J., 2023, MNRAS, 518, 4191

Reid M. J., Braatz J. A., Condon J. J., Greenhill L. J., Henkel C., Lo K. Y., 2009, ApJ, 695, 287

Reid M. J., Braatz J. A., Condon J. J., Lo K. Y., Kuo C. Y., Impellizzeri C. M. V., Henkel C., 2013, ApJ, 767, 154

Reid M. J., Pesce D. W., Riess A. G., 2019, ApJL, 886, L27

Rey M. P., Pontzen A., Saintonge A., 2019, MNRAS, 485, 1906

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2011, ApJ, 730, 119

Riess A. G., et al., 2016, ApJ, 826, 56

Riess A. G., et al., 2018, ApJ, 861, 126

Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021, ApJL, 908, L6

Riess A. G., et al., 2022a, ApJL, 934, L7
Riess A. G., et al., 2022b, ApJ, 938, 36
Riess A. G., et al., 2023, ApJL, 956, L18
Riess A. G., et al., 2024, ApJL, 962, L17
Rigault M., et al., 2025, A&A, 694, A1
Rimes C. D., Hamilton A. J. S., 2005, MNRAS, 360, L82
Ripepi V., et al., 2026, A&A, 708, A216
Romaniello M., et al., 2022, A&A, 658, A29
Roy O., Vetterli M., 2007, in 2007 15th European Signal Processing Conference. IEEE, p. 124
Rubin D., et al., 2015, ApJ, 813, 137
Rubin D., et al., 2025, ApJ, 986, 231
Said K., Colless M., Magoulas C., Lucey J. R., Hudson M. J., 2020, MNRAS, 497, 1275
Said K., et al., 2025, MNRAS, 539, 3627
Sawala T., McAlpine S., Jasche J., Lavaux G., Jenkins A., Johansson P. H., Frenk C. S., 2022, MNRAS, 509, 1432
Schaller M., et al., 2024, MNRAS, 530, 2378
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Schmidt F., 2025, J. Cosmology Astropart. Phys., 2025, 056
Schneider A., Teyssier R., 2015, J. Cosmology Astropart. Phys., 2015, 049
Schöneberg N., 2024, J. Cosmology Astropart. Phys., 2024, 006
Schöneberg N., Abellán G. F., Sánchez A. P., Witte S. J., Poulin V., Lesgourgues J., 2022, Phys. Rep., 984, 1
Scolnic D., et al., 2022, ApJ, 938, 113
Scrimgeour M. I., et al., 2012, MNRAS, 425, 116
Scrimgeour M. I., et al., 2016, MNRAS, 455, 386
Secrest N. J., von Hausegger S., Rameez M., Mohayaee R., Sarkar S., Colin J., 2021, ApJL, 908, L51
Secrest N. J., von Hausegger S., Rameez M., Mohayaee R., Sarkar S., 2022, ApJL, 937, L31
Secrest N., von Hausegger S., Rameez M., Mohayaee R., Sarkar S., 2025, Reviews of Modern Physics, 97, 041001
Seifert A., Lane Z. G., Galoppo M., Ridden-Harper R., Wiltshire D. L., 2025, MNRAS, 537, L55
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Skilling J., 2006, Bayesian Analysis, 1, 833
Skrutskie M. F., et al., 2006, AJ, 131, 1163
Smith R. E., et al., 2003, MNRAS, 341, 1311
Son J., Lee Y.-W., Chung C., Park S., Cho H., 2025, MNRAS, 544, 975
Sorce J. G., 2020, MNRAS, 495, 4463
Sorce J. G., Tempel E., 2018, MNRAS, 476, 4362
Sorce J. G., Courtois H. M., Gottlöber S., Hoffman Y., Tully R. B., 2014, MNRAS, 437, 3586
Sorce J. G., et al., 2016a, MNRAS, 455, 2078

Sorce J. G., Gottlöber S., Hoffman Y., Yepes G., 2016b, MNRAS, 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
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Springel V., et al., 2008, MNRAS, 391, 1685
Springel V., Pakmor R., Zier O., Reinecke M., 2021, MNRAS, 506, 2871
Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, ApJS, 172, 599
Springob C. M., et al., 2014, MNRAS, 445, 2677
Srinivasan R., Crisostomi M., Trotta R., Barausse E., Breschi M., 2024, Phys. Rev. D, 110, 123007
Starobinsky A. A., 1982, Physics Letters B, 117, 175
Steinmetz M., Navarro J. F., 1999, ApJ, 513, 555
Stiskalek R., 2025, arXiv e-prints, p. arXiv:2509.20235
Stiskalek R., Desmond H., 2023, MNRAS, 525, 6130
Stiskalek R., Desmond H., 2026, The Open Journal of Astrophysics, 9, 62811
Stiskalek R., Desmond H., Devriendt J., Slyz A., 2024, MNRAS, 534, 3120
Stiskalek R., Desmond H., McAlpine S., Lavaux G., Jasche J., Hudson M. J., 2026a, arXiv e-prints, p. arXiv:2601.08524
Stiskalek R., Riess A., Desmond H., Lavaux G., Scolnic D., 2026b, arXiv e-prints, p. arXiv:2603.09880
Stiskalek R., Desmond H., Devriendt J., Slyz A., Lavaux G., Hudson M. J., Bartlett D. J., Courtois H. M., 2026c, MNRAS, 545, staf1960
Stiskalek R., Desmond H., Lavaux G., 2026d, MNRAS, 546, staf2048
Stiskalek R., Desmond H., Tsaprazi E., Heavens A., Lavaux G., McAlpine S., Jasche J., 2026e, MNRAS, 546, staf2260
Stopyra S., Peiris H. V., Pontzen A., Jasche J., Natarajan P., 2021, MNRAS, 507, 5425
Stopyra S., Peiris H. V., Pontzen A., Jasche J., Lavaux G., 2024a, MNRAS, 527, 1244
Stopyra S., Peiris H. V., Pontzen A., Jasche J., Lavaux G., 2024b, MNRAS, 531, 2213
Strauss M. A., Willick J. A., 1995, Phys. Rep., 261, 271
Sui C., Bartlett D. J., Pandey S., Desmond H., Ferreira P. G., Wandelt B. D., 2025, A&A, 698, A1
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, J. Cosmology Astropart. Phys., 2013, 036
Teyssier R., 2002, A&A, 385, 337
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tonry J., Schneider D. P., 1988, AJ, 96, 807
Tripp R., 1998, A&A, 331, 815
Tristram M., et al., 2024, A&A, 682, A37
Trotta R., 2008, Contemporary Physics, 49, 71

Tsaprazi E., Heavens A. F., 2025, MNRAS, 539, 1448
Tsaprazi E., et al., 2022, MNRAS, 510, 366
Tully R. B., Fisher J. R., 1977, A&A, 54, 661
Tully R. B., et al., 2013, AJ, 146, 86
Tully R. B., Courtois H. M., Sorce J. G., 2016, AJ, 152, 50
Tully R. B., et al., 2023, ApJ, 944, 94
Valade A., Hoffman Y., Libeskind N. I., Graziani R., 2022, MNRAS, 513, 5148
Valade A., Libeskind N. I., Hoffman Y., Pfeifer S., 2023, MNRAS, 519, 2981
Valade A., Libeskind N., Pomarède D., Stiskalek R., Hoffman Y., Gottlöber S., Tully R. B., 2026, arXiv e-prints, p. arXiv:2602.03699
Verdinelli I., Wasserman L., 1995, Journal of the American Statistical Association, 90, 614
Villaescusa-Navarro F., et al., 2020, ApJS, 250, 2
Virtanen P., et al., 2020, Nature Medicine, 17, 261
Vogl C., et al., 2025, A&A, 702, A41
Wagenveld J. D., Klöckner H.-R., Schwarz D. J., 2023, A&A, 675, A72
Wang L., Steinhardt P. J., 1998, ApJ, 508, 483
Wang H., Mo H. J., Yang X., Jing Y. P., Lin W. P., 2014, ApJ, 794, 94
Wang H., et al., 2016, ApJ, 831, 164
Wang H., et al., 2018, ApJ, 852, 31
Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881
Watkins R., Feldman H. A., 2015, MNRAS, 447, 132
Watkins R., Feldman H. A., Hudson M. J., 2009, MNRAS, 392, 743
Watkins R., et al., 2023, MNRAS, 524, 1885
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Weinberg S., 1989, Reviews of Modern Physics, 61, 1
Weinberg S., 2008, Cosmology
Westover M., 2007, PhD thesis, Harvard University, Massachusetts
White S. D. M., 1984, ApJ, 286, 38
Whitford A. M., Howlett C., Davis T. M., 2023, MNRAS, 526, 3051
Willick J. A., Batra P., 2001, ApJ, 548, 564
Wojtak R., Hjorth J., 2022, MNRAS, 515, 2790
Wojtak R., Hjorth J., 2024, MNRAS, 533, 2319
Wojtak R., Hjorth J., 2025, A&A, 702, A176
Wong K. C., et al., 2020, MNRAS, 498, 1420
Wright A. H., et al., 2025, A&A, 703, A158
Xu X., Yang X., Xu H., Zhang Y., 2024, MNRAS, 527, 7013
Yahil A., Strauss M. A., Davis M., Huchra J. P., 1991, ApJ, 372, 380
Yang X., et al., 2018, ApJ, 860, 30
Yasin T., Stiskalek R., Desmond H., von Hausegger S., Ferreira P. G., 2026, arXiv e-prints, p.

arXiv:2602.06007

York D. G., et al., 2000, AJ, 120, 1579

Zaldarriaga M., Seljak U., 1997, Phys. Rev. D, 55, 1830

Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, ApJ, 449, 446

Zaroubi S., Hoffman Y., Dekel A., 1999, ApJ, 520, 413

Zel'dovich Y. B., 1970, A&A, 5, 84

Zhang Y., Yang X., Guo H., 2022, MNRAS, 517, 3579

de Jaeger T., Galbany L., Riess A. G., Stahl B. E., Shappee B. J., Filippenko A. V., Zheng W., 2022, MNRAS, 514, 4620

de Jong R. S., et al., 2019, The Messenger, 175, 3

de Lapparent V., Geller M. J., Huchra J. P., 1986, ApJL, 302, L1

van den Bosch F. C., Jiang F., 2016, MNRAS, 458, 2870

van den Bosch F. C., Ogiya G., 2018, MNRAS, 475, 4066

von Hausegger S., Secrest N., Desmond H., Rameez M., Mohayaee R., Sarkar S., 2026, MNRAS, 546, stag201