

Revealing the neutrino mass through persistent homology of the cosmic web

Jiaqi Wang^{1,2*}, Willem Elbers^{2*}, Carlos S. Frenk², Shaun Cole², Xiaohu Yang^{1,3}, Ian G. McCarthy⁴ and Rien van de Weygaert⁵

¹State Key Laboratory of Dark Matter Physics, Tsung-Dao Lee Institute & School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 201210, China

²Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

³Shanghai Key Laboratory for Particle Physics and Cosmology, and Key Laboratory for Particle Physics, Astrophysics and Cosmology, Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China

⁴Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

⁵Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

Accepted 2026 May 13. Received 2026 May 13; in original form 2026 April 2

ABSTRACT

Cosmological constraints on neutrino mass offer a promising avenue for advancing our understanding of both fundamental particle physics and the evolution of cosmic large-scale structure. To overcome challenges associated with traditional probes of neutrino mass, particularly degeneracies with other parameters, we consider topological summaries of the cosmic web based on the formalism of persistent homology. We introduce persistence strips, a novel representation that segments Betti curves by topological persistence, effectively condensing 2D persistence diagrams into a set of 1D curves. Applied to the FLAMINGO suite of cosmological simulations, these topological descriptors demonstrate pronounced sensitivity to neutrino mass. By constructing an emulator spanning a 10D w_0w_a CDM + ν cosmological parameter space that includes parameters degenerate with neutrino masses in conventional approaches, assuming a volume of only $(350 \text{ Mpc})^3$, we obtain neutrino mass constraints with an uncertainty of 0.05 eV for the total matter field and 0.13 eV for the dark matter-only field, with the strongest constraints coming from void topology. Persistence strips exhibit roughly twice the constraining power of unbinned Betti curves and, through their multiscale, environment-dependent sensitivity, systematically break degeneracies between neutrino mass and other cosmological parameters. We pinpoint the precise physical origin of the signal, identifying two equally important contributions: sensitivity to the neutrino mass fraction, which is highest in underdense regions, and the impact of neutrinos on the distribution of dark matter. Our findings highlight the particular promise of applying topological statistics to weak lensing measurements, which directly probe the total matter distribution.

Key words: neutrinos – dark matter – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Massive neutrinos represent a crucial extension of the standard model of particle physics and play a vital role in the evolution of the Universe. As the only known form of hot dark matter, they exhibit key differences from the cold dark matter (CDM) that makes up most of the matter budget in the standard cosmological paradigm. While CDM clumps under gravity to form cosmic structures, neutrinos suppress the growth of structure on small scales. Their significant thermal motions prevent them from clustering effectively, thereby smoothing out the matter distribution (J. R. Bond, G. Efstathiou & J. Silk 1980; W. Hu, D. J. Eisenstein &

M. Tegmark 1998; J. Brandbyge et al. 2008). In addition, neutrinos alter the cosmic expansion history by behaving as non-relativistic matter only at late times. Consequently, cosmological observations provide powerful and complementary constraints on the total neutrino mass, $\sum m_\nu$, to those derived from laboratory experiments (e.g. K. N. Abazajian et al. 2011; I. Esteban et al. 2024).

The cosmological imprints of neutrino mass can be probed through a variety of astrophysical observables, usually by combining probes of the cosmic large-scale structure with information derived from the cosmic microwave background (CMB), where the latter provides an anchor at large scales and early times when neutrinos are still relativistic (e.g. K. N. Abazajian et al. 2015). The strongest current constraints are derived from CMB data from the *Planck* satellite (Planck Collaboration VI 2020) and the ground-based Atacama Cosmology and South Pole Telescopes

* E-mail: jiaqi.wang@durham.ac.uk, wangjia777@sjtu.edu.cn (JW); willem.h.elbers@durham.ac.uk (WE)

(E. Calabrese et al. 2025; E. Camphuis et al. 2026), combined with measurements of baryon acoustic oscillations (BAO) in the distribution of galaxies, quasars, and the Lyman- α forest from the second data release of the Dark Energy Spectroscopic Instrument (M. Abdul Karim et al. 2025; DESI Collaboration 2025). When analysed within the standard Λ CDM framework, these data provide 95 per cent upper bounds in the range $\sum m_\nu < 0.04 - 0.06$ eV (E. Camphuis et al. 2026; W. Elbers et al. 2025b). This is close to or below the lower bounds derived from terrestrial neutrino oscillation measurements, $\sum m_\nu > 0.059$ eV, giving rise to a neutrino mass tension that could point to new neutrino physics or a non-standard cosmological evolution (e.g. N. Craig et al. 2024; D. Naredo-Tuero et al. 2024; H. E. Noriega & A. Aviles 2025; O. Avsajanishvili 2025; D. Green & J. Meyers 2025; J.-Q. Jiang et al. 2025; W. Elbers et al. 2025a, b). While conventional CMB and BAO techniques currently provide the strongest and most reliable constraints on $\sum m_\nu$, the latest hints of a neutrino mass tension motivate the development of alternative approaches that could add complementary information.

Additional information can be extracted by modelling the shape of the power spectrum of density fluctuations probed by galaxy clustering (e.g. W. Hu et al. 1998; Ø. Elgarøy et al. 2002; M. Tegmark et al. 2004), weak lensing (e.g. A. R. Cooray 1999; K. Abazajian & S. Dodelson 2003), and the absorption spectra of the Lyman- α forest (e.g. R. A. C. Croft, W. Hu & R. Davé 1999; U. Seljak et al. 2005). While these probes provide valuable information, their constraining power is subject to challenges, including theoretical modelling uncertainties, degeneracies with other parameters (particularly the amplitude of matter fluctuations, σ_8 ¹, and the spectral index, n_s), and limited signal-to-noise ratios in current observations.

Recognizing these limitations, several studies have proposed that higher-order statistics could provide improved constraints on neutrino masses (C.-T. Chiang et al. 2018; C. Hahn & F. Villaescusa-Navarro 2021). For instance, C. Hahn & F. Villaescusa-Navarro (2021) demonstrated that the bispectrum can supplement information from the power spectrum and help break the degeneracy between $\sum m_\nu$ and σ_8 , despite the challenges associated with its modelling. However, the utility of such statistics may be limited in certain contexts; A. E. Bayer, A. Banerjee & U. Seljak (2022) argued that for 2D probes like the halo field and weak lensing at scales $k \leq 1h \text{ Mpc}^{-1}$, higher-order statistics contribute little information beyond the linear power spectrum. Although the 3D matter field may contain more information, obtaining high-quality observational data remains difficult.

An alternative strategy involves focusing on specific cosmic environments, such as voids (e.g. E. Massara et al. 2015; C. D. Kreisch et al. 2019; L. Thiele et al. 2024). Nevertheless, these methods face challenges: results are sensitive to void definitions that vary across studies, and analytical models of void statistics often rely on simplifying assumptions that limit their applicability to the full void population. These issues underscore the need for an effective and unambiguous summary statistic that balances conciseness with informational richness. For example, A. E. Bayer et al. (2024) showed that the halo mass function in environments split according to the emptiness can yield tight bounds on neutrino masses.

¹Defined as the standard deviation of the present-day linear matter field averaged in spheres of radius $8 h^{-1} \text{ Mpc}$.

Recent years have seen the emergence of several novel cosmological probes that target the geometry or topology of the cosmic web, including the minimum spanning tree (K. Naidoo, E. Massara & O. Lahav 2022; L. N. L. Simões et al. 2025), k -Nearest Neighbour distributions (k NN; K. Gangopadhyay, A. Banerjee & T. Abel 2025; A. Ouellette & G. Holder 2025), the clustering of cosmic web components (J. Sunseri, A. E. Bayer & J. Liu 2025), and topological data analysis (R. de Weygaert et al. 2011, TDA hereafter). Among these, TDA has shown particular promise as a powerful tool for characterizing the multiscale topological structure of the cosmic web. Through persistent homology, a method that quantifies the evolution of topological features across a range of physical scales or density thresholds, TDA provides a rigorous framework for describing the ‘shape’ of large-scale structure (G. Carlsson 2009). This approach encodes information in the form of persistence diagrams, which track the birth and death scales of topological features and contain significantly more information than the traditional Euler characteristic. Moreover, persistent homology captures structural aspects of the density field complementary to conventional N -point statistics and offers a natural definition of the signal-to-noise ratio through the scale-persistence relationship of features in the persistence diagrams (PDs hereafter; P. Pranav et al. 2017). TDA has found diverse applications in astrophysics, including the detection of non-Gaussianity (M. Biagetti, A. Cole & G. Shiu 2021), the analysis of the topology of reionization (W. Elbers & R. van de Weygaert 2019; S. K. Giri & G. Mellema 2021; W. Elbers & R. van de Weygaert 2023), constraints on redshift-space distortions (Y. Liu et al. 2024), and the extraction of cosmological information from cosmic shear (e.g. S. Heydenreich, B. Brück & J. Harnois-Déraps 2021). Within large-scale structure studies, TDA offers a powerful framework for quantifying the complex topology of the cosmic web (R. de Weygaert et al. 2011) and has been used to constrain both cosmological parameters and galaxy evolution processes (J. H. T. Yip et al. 2024). Common topological methods include the Euler characteristic (R. de Weygaert et al. 2011), the β -skeleton (F. Fang et al. 2019), Minkowski functionals (Y. Liu et al. 2020; W. Liu, A. Jiang & W. Fang 2022), critical points (J. Moon, G. Rossi & H. Yu 2023), and persistent homology (G. Wilding et al. 2021; M. Tsizh, V. Tymchyshyn & F. Vazza 2023). The last approach characterizes multiscale topology through Betti curves, which track the evolution of topological features, such as connected components, loops, and voids, across a hierarchy of filtration parameters.

Recent work by M. H. Jalali Kanafi, S. Ansarifard & S. M. S. Movahed (2024) has demonstrated the particular value of Betti curves in constraining neutrino mass. Through Fisher forecasts using simulations with varying $\sum m_\nu$ and σ_8 (with other parameters held fixed), they showed that Betti curves not only provide tight constraints on $\sum m_\nu$, but also help break its degeneracy with σ_8 . Notably, they found that the degeneracy-breaking capability differs between the total matter field (including neutrinos) and the combined field of dark matter and baryons—a phenomenon whose physical origin warrants further investigation. To extend this research, we consider a broader cosmological parameter space, including Ω_m and h , and utilize higher-resolution neutrino simulations from the FLAMINGO suite (R. Kugel et al. 2023; J. Schaye et al. 2023), which enable exploration of smaller scales than previously possible with Quijote (F. Villaescusa-Navarro et al. 2020). This approach allows us to derive more comprehensive parameter constraints and better understand the scale-dependent information content of topological summaries.

To address these challenges, we conduct a comprehensive investigation across a high-dimensional cosmological parameter space. However, modelling topological statistics presents significant difficulties, particularly in such an expansive parameter regime where analytical approaches become intractable. To overcome this limitation, we employ emulation techniques, constructing surrogate models that accurately replicate the output of full-scale cosmological simulations at a fraction of the computational cost. This approach offers dual advantages in model fitting: dramatically reduced computation times and enabling the use of sophisticated statistical methods, including gradient-based optimization and efficient sampling algorithms that would otherwise be prohibitively expensive. Emulation has demonstrated remarkable success across cosmological applications, providing an effective methodology for balancing precision and computational efficiency in data-intensive studies.

In this work, we aim to characterize the multiscale topological structure and connectivity of the cosmic web using persistent homology and Betti curves. We develop an emulator capable of predicting topological statistics across a 10D cosmological parameter space and check the cosmological parameter inference. The paper is organized as follows: Section 2 describes the simulations used in our analysis. Section 3 introduces the theoretical framework of TDA and examines the response of Betti curves to neutrino mass variations. Section 4 details the emulator design and training. Section 5 presents our main results on parameter dependence and cosmological constraints. Section 6 discusses the physical origin of the signals. Section 7 checks some of the observational effects and presents further discussion. Finally, Section 8 summarizes our findings and conclusions. Note that in this paper log is in base 10.

2 DATA

We make use of two sets of simulations in our analysis: a subset of simulations from the FLAMINGO suite (Section 2.1) and a larger set of 110 simulations spanning a cosmological parameter grid (Section 2.2). The first are used for a systematic exploration of the topological effects of massive neutrinos, while the second are used to construct our persistent homology emulator.

2.1 FLAMINGO simulations

The Virgo Consortium’s FLAMINGO suite represents a state-of-the-art set of cosmological simulations designed to advance our understanding of structure formation, galaxy cluster physics, and the role of neutrinos and baryons in cosmic evolution (R. Kugel et al. 2023; J. Schaye et al. 2023, Helly et al. in preparation). The original suite contains 12 dark matter-only simulations and 16 hydrodynamical simulations, which were later complemented with additional neutrino mass variations and extensions with decaying dark matter (W. Elbers et al. 2025c). The suite comprises large simulations in multiple volumes and resolutions, incorporating both gravitational physics and detailed prescriptions for galaxy formation. While FLAMINGO offers many baryonic physics variations, we do not use the hydrodynamical simulations in this study, where our focus is on the gravitational impact of neutrinos. We exploit the detailed implementation of massive neutrinos in the FLAMINGO simulations to investigate the constraining power of TDA on neutrino mass, as well as its degeneracies with other cosmological parameters.

Table 1. Allowed ranges of each cosmological parameter in our emulator. The upper rows show the 10 original parameters used to construct the hypercube. We also list the range of the σ_8 parameter used in this study.

Parameter	Minimum	Maximum
Ω_m	0.17	0.38
f_b	0.14	0.17
h	0.6	0.8
n_s	0.9	1
$\ln(10^{10} A_s)$	2.97	3.11
w_0	-1.19	-0.30
w_a	-2.74	1.08
M_ν [eV]	0.005	0.45
α_s	-0.03	0.03
$\Gamma_{\text{cdm}} [10^{-3} \text{Gyr}^{-1}]$	0.33	19.8
σ_8	0.48	1.05

The FLAMINGO simulations were run with the modern and highly scalable cosmological hydrodynamics code SWIFT (M. Schaller et al. 2018; M. Schaller et al. 2024). Massive neutrinos are included using the computationally efficient δf method (W. Elbers et al. 2021), which significantly reduces shot noise and allows for accurate, self-consistent evolution of neutrino perturbations. The simulations are started from higher-order Lagrangian multi-fluid initial conditions, incorporating corrections for the presence of massive neutrinos (O. Hahn, C. Rampf & C. Uhlemann 2021; W. Elbers 2022; W. Elbers et al. 2022).

For this study, we primarily rely on FLAMINGO simulations with a 1 Gpc^3 volume, consisting of 1800^3 cold dark matter particles and 1000^3 neutrino particles, for which multiple neutrino mass variations are available. We also use the larger 2.8 Gpc^3 simulation with 5040^3 cold dark matter particles and 2800^3 neutrino particles to estimate the covariance matrix of the statistical errors of our estimates. To explore the effect of neutrino mass, we use three different values: 0.06 eV (the baseline), 0.24 eV, and 0.48 eV. This range enables a detailed examination of the imprint of neutrino mass on the topological structure of the cosmic web, including its interplay with other cosmological parameters.

2.2 The emulator

In order to constrain cosmological parameters, including the total mass of neutrinos, we developed an emulator to model our topological measurements. The emulator is based on a suite of 50 cosmological N -body simulations with a side length of $L = 350 \text{ Mpc}$ and 50 additional simulations with a side length of $L = 700 \text{ Mpc}$, each containing 1260^3 cold dark matter particles and 700^3 neutrino particles, evolved under periodic boundary conditions. The simulations were run with the SWIFT code, and their treatment of neutrinos was identical to the FLAMINGO simulations described above.

The simulations are designed for the sampling of 10 cosmological parameters: $\{\Omega_m, f_b, h, n_s, A_s, w_0, w_a, M_\nu, \alpha_s, \Gamma_{\text{cdm}}\}$, where the parameter ranges are shown in Table 1. This includes the standard Λ CDM parameters: the matter density Ω_m , baryon fraction $f_b \equiv \Omega_b/\Omega_m$, Hubble parameter h , power spectrum amplitude A_s , and spectral index n_s . Beyond these parameters, the emulator incorporates extensions for dynamical dark energy with an evolving equation of state parametrized by w_0 and w_a (M. Chevallier & D. Polarski 2001; E. V. Linder 2003), non-zero neutrino masses $M_\nu \equiv \sum m_\nu$, and running of the spectral index α_s . Finally, Γ_{cdm} indicates the decay rate of cold dark matter in units

of $\text{km s}^{-1} \text{Mpc}^{-1}$, as described in W. Elbers et al. (2025c). The parameters were sampled according to a space-filling orthogonal Latin hypercube, but points were transformed in the (w_0, w_a) -plane to primarily sample the space around the degeneracy direction inferred from the DESI DR1 results (A. G. DESI Collaboration 2025). To evaluate the performance in the presence of neutrino mass, four additional simulations with $L = 350 \text{ Mpc}$ are run with varying neutrino mass while holding other parameters fixed. This simulation suite will be described in greater detail in forthcoming publications (McCarthy et al., in prep.).

2.3 Density contrast

In order to save computational costs, we focus in this study on the density field as input data and employ the density contrast $1 + \delta(\mathbf{r}) \equiv n(\mathbf{r})/\bar{n}$ as our filtration (in Section 3.1). The density contrast for the total matter field is defined as,

$$\begin{aligned} \delta_m &\equiv \frac{\Omega_{\text{CDM}}}{\Omega_m} \delta_{\text{CDM}} + \frac{\Omega_b}{\Omega_m} \delta_b + \frac{\Omega_\nu}{\Omega_m} \delta_\nu \\ &= \frac{\Omega_{\text{CDM}} + \Omega_b}{\Omega_m} \delta_{\text{CDM}} + \frac{\Omega_\nu}{\Omega_m} \delta_\nu, \end{aligned} \quad (1)$$

where $\Omega_m = \Omega_{\text{CDM}} + \Omega_b + \Omega_\nu$. The second equality comes from the fact that we assume the baryonic field to be the same as the CDM field since we do not use hydrodynamical simulations.

We build the density contrast fields from simulation snapshots for different mass components using the Pylians (F. Villaescusa-Navarro 2018) package² on a 512^3 grid using the Cloud-In-Cell (CIC hereafter) mass-assignment scheme (S. Laux 1996). We generally consider smoothed fields in our analysis, convolving the underlying field with a Gaussian window function. As described in Section 7.1, we consider a range of different smoothing scales, with a fiducial choice of $R = 2 \text{ Mpc}$.

3 PERSISTENT HOMOLOGY AND COSMIC WEB

The conceptual foundation for understanding walls, filaments, and voids in the large-scale structure was established by the Zeldovich school through the Zeldovich approximation (Y. B. Zel'dovich 1970) and the adhesion model (S. F. Shandarin & Y. B. Zeldovich 1989). The term ‘cosmic web’ was later coined by J. R. Bond, L. Kofman & D. Pogosyan (1996) to emphasize the interconnected, hierarchical nature of this structure. Subsequent theoretical work has further elucidated the multiscale evolution of voids (V. Sahni, B. S. Sathyaprakash & S. F. Shandarin 1994; R. K. Sheth & R. van de Weygaert 2004), the broader cosmic web (J. Shen et al. 2006), filaments (J. Feldbrugge & R. van de Weygaert 2025), and walls (B. Hertzsch et al. 2026).

Topology provides a complementary perspective by characterizing this structure as a manifold, focusing on its intrinsic connectivity rather than its precise metric properties (H. Edelsbrunner & J. Harer 2010; L. Wasserman 2018; G. Carlsson & M. Vejdemo-Johansson 2021). Persistent homology enables the systematic tracking of the emergence and disappearance of structures classified as k -cycles: connected components (0-cycles), loops enclosing tunnels (1-cycles), and shells enclosing cavities (2-cycles), which correspond to halos, filaments, and voids, respectively (see J. R. Munkres 1984; A. Hatcher 2002). Throughout this study,

we calculate the topological properties of the field data using the PYTHON package GUDHI (P. Dlotko 2025; T. G. Project 2025)³. The field data are represented as a cubical complex with periodic boundary conditions, from which the topological properties can be calculated efficiently. In general, Betti curves depend both on the simulation box size and the number of grid points that are used to calculate the density contrast. Thus, we use normalized numbers in units of $[\text{Mpc}^{-3}]$ to make the model predictions.

3.1 Persistent homology

Persistent homology is an important method in TDA that quantitatively describes features in data across a range of scales. It constructs a concise summary of the shape of a dataset, robust to noise and input perturbations. The topology of a field is entirely determined by the maxima, saddles, and minima of the field, with topological changes occurring precisely when the density threshold passes through the value of a singularity according to Morse theory (J. Milnor 1963).

This process begins with a filtration, a nested sequence of structures built from any well-defined field (e.g. continuous fields or discrete point clouds) by varying the chosen parameter. Superlevel and sublevel filtrations are two examples of filtrations based on excursion sets, constructed by traversing the field in opposite directions and selecting the regions that exceed or fall below a certain threshold. In practice, by analysing the negative of the field, a superlevel filtration becomes equivalent to a sublevel filtration of the inverse field. In N dimensions, Alexander duality provides a further correspondence between the k th homology of a manifold and the $(N - k - 1)$ th cohomology of the complement. This implies that the persistent homology of a superlevel filtration can be obtained by analysing the sublevel filtration of the inverse field. However, there are subtle differences in practice (see appendix C of O. Bobrowski & P. Skraba 2020, for an illustrative example in the case of cubical complexes). In this work, we adopt a superlevel filtration, i.e. start at a high density threshold and gradually decrease the density threshold, to keep aligned with previous works (e.g. G. Wilding et al. 2021; M. H. Jalali Kanafi et al. 2024). However, since the GUDHI package computes sublevel filtrations, we rely on the correspondence discussed above.

During the filtration process, topological features appear, evolve, and eventually disappear. This yields topological quantities that vary as functions of the threshold, $\vartheta = 1 + \delta(\mathbf{r})$. To illustrate this process, Fig. 1 shows the structures identified above the specific thresholds ϑ , alongside their corresponding Betti curves. Halos, quantified by β_0 , appear first at the highest density filtration ranges. As the density filtration decreases gradually, halos are connected to filaments and tunnels, reflected in reducing β_0 and raising β_1 . As the threshold decreases further, filaments are connected to walls enclosing voids, manifested as reducing β_1 and raising β_2 .

3.2 Persistence diagrams and Betti curves

To analyse the multiscale nature and interactions of topological features in greater detail than Betti curves allow, we represent them using persistence diagrams. These diagrams feature each structure in the ‘birth’ density and ‘death’ density coordinates,

²<https://pylians3.readthedocs.io/en/master/index.html>

³<https://gudhi.inria.fr/python/latest/>

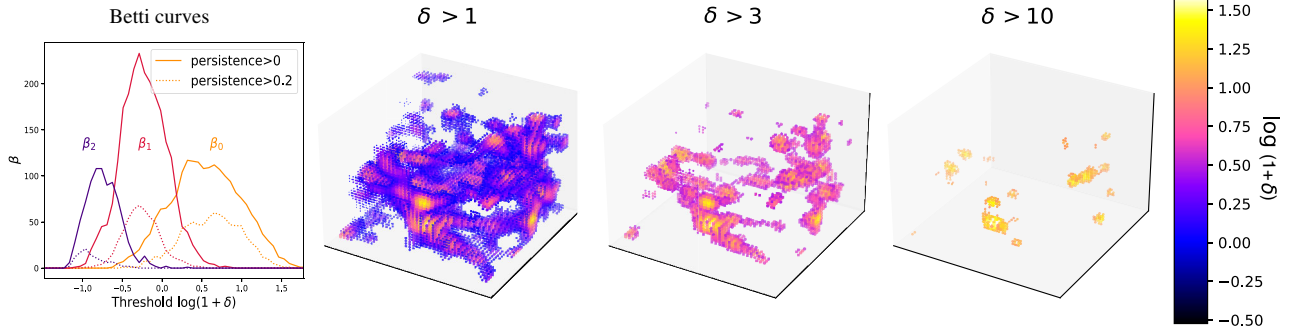


Figure 1. Topological structure of a 3D cosmic density field. (Left) Betti curves β_k as a function of density threshold, showing the number of k -dimensional topological features. From left to right: cosmic voids (β_2 , purple), filamentary tunnels (β_1 , red), and haloes (β_0 , orange). Solid and dotted lines correspond to persistence thresholds of 0 and 0.2, respectively, applied to the underlying persistence diagrams. (Right) Visualizations of the identified structures at different density thresholds. Colour in the 3D renderings represents the local matter overdensity, as indicated by the colour bar.

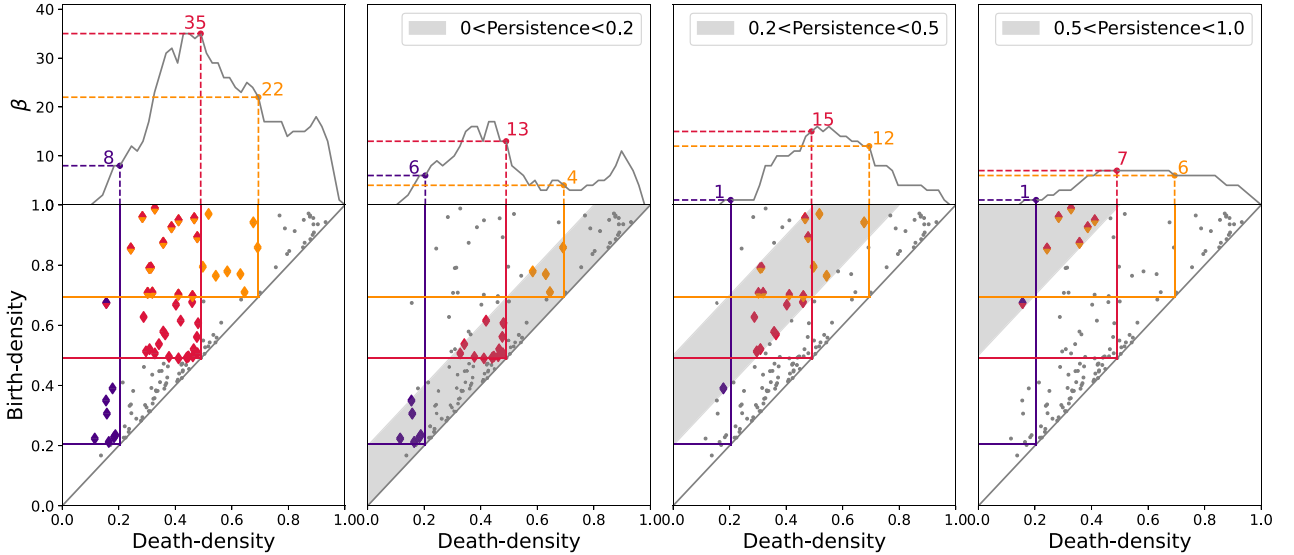


Figure 2. The relationship between Betti curves and persistence diagrams. Upper panel: Betti curves showing the count of topological features as a function of density threshold. The coloured dots on the curves mark the Betti numbers at three specific density thresholds: 0.2 (red), 0.5 (orange), and 1.0 (purple). Bottom panel: The corresponding persistence diagrams, where each point represents a topological feature's birth and death density. The coloured diamonds highlight the specific features that are active (i.e. born before and dying after) at each of the three density thresholds. Features active at multiple thresholds are shown with multiple colours. The three columns on the right demonstrate the effect of applying different persistence cuts (i.e. minimum feature lifetimes) when computing the Betti curves: $[0, 0.2]$, $[0.2, 0.5]$, and $[0.5, 1.0]$, as indicated by the shaded regions in the persistence diagrams.

corresponding to the thresholds at which the feature first appears and at which it subsequently disappears. The ‘persistence’ of a feature thus provides a quantitative measure of its significance across scales. All points in a persistence diagram lie above the diagonal line ($\vartheta_{\text{birth}} = \vartheta_{\text{death}}$), since a feature must be born before it can die. In a superlevel filtration, this means the birth density is always higher than the death density.

The persistence diagram preserves the complete birth-death pairing, thereby encoding the entire lifecycle of each topological feature. Meanwhile, Betti curves integrate out the pairing information between birth and death events, retaining only the net balance of features at each density threshold. This additional information enables persistence diagrams to distinguish between features of different lifetimes that would be indistinguishable in a

Betti curve, and to capture the nested hierarchy of features across scales.

Fig. 2 illustrates the relationship between Betti curves (top panel) and persistence diagrams (bottom panel) using simplified mock data. The Betti number $\beta_k(\vartheta)$ for topological features of dimension k is mathematically related to the persistence diagram through the expression,

$$\beta_k(\vartheta) = \sum_{i=1}^{n_d} \Theta(\vartheta_{i, \text{birth}}^k - \vartheta) \Theta(\vartheta - \vartheta_{i, \text{death}}^k), \quad (2)$$

where Θ denotes the Heaviside step function. This corresponds to counting the number of topological features in the persistence diagram that were born at a density threshold $\vartheta < \vartheta_{i, \text{birth}}$ and die at a threshold $\vartheta > \vartheta_{i, \text{death}}$, that is, points lying to the left of

and above the coordinates $(\vartheta_{i,\text{death}}, \vartheta_{i,\text{birth}})$, as shown in Fig. 2. We demonstrate this correspondence explicitly using three representative density thresholds (0.2, 0.5, and 0.7), indicated by red, yellow, and purple highlighting, respectively. Solid lines mark the counting regions, and diamond points represent features counted at each threshold. Points marked with multiple colours correspond to persistent topological features that contribute to the Betti counts across several density thresholds.

Furthermore, the perpendicular distance of a point from the diagonal in a persistence diagram quantifies the persistence (or lifetime) of the corresponding topological feature, defined as,

$$\vartheta_{i,\text{pers}}^k \equiv |\vartheta_{i,\text{birth}}^k - \vartheta_{i,\text{death}}^k|. \quad (3)$$

This measure serves a dual purpose. First, features with short persistence typically originate from noise, offering a natural criterion to distinguish significant structures from random fluctuations. Second, the distribution of feature lifetimes itself encodes valuable physical and cosmological information. In the three right subpanels of Fig. 2, the grey shaded regions indicate distinct persistence intervals: $[0, 0.2]$, $[0.2, 0.5]$, and $[0.5, 1.0]$. The Betti curves derived from the lowest persistence range (0–0.2) exhibit high fluctuations, indicative of noise-dominated features. In contrast, Betti curves constructed from features with higher persistence (e.g. 0.2–0.5 or 0.5–1.0) show smoother behaviour and reveal more robust large-scale topological structure. The morphological differences between these filtered Betti curves highlight the multiscale nature of the cosmic web and demonstrate how persistence-based filtering can isolate physically distinct components of the matter distribution.

3.3 Persistence strips

Since persistence quantifies the lifetime of topological features across filtration scales, it is natural to incorporate it as an additional dimension in the analysis of cosmic web structures. We introduce the concept of persistence strips, defined as follows:

$$\beta_{k,\pi_p}(\vartheta) = \sum_{i=1}^{n_k} \Theta(\vartheta_{i,\text{birth}}^k - \vartheta) \Theta(\vartheta - \vartheta_{i,\text{death}}^k) \Theta(\vartheta_{i,\text{pers}}^k - \pi_{\text{low},p}) \Theta(\pi_{\text{high},p} - \vartheta_{i,\text{pers}}^k), \quad (4)$$

where $\pi_{\text{low},p}$ is the p^{th} lower persistence cut and $\pi_{\text{high},p}$ the higher cut. These persistence strips can reveal additional information compared to the Betti curves in the following ways. This formulation counts the number of topological features of dimension k that are present at the density threshold ϑ and whose persistence (lifetime) is in the range $[\pi_{\text{low},p}, \pi_{\text{high},p}]$.

The persistence strips provide a summary that captures both the abundance and robustness of topological features across densities and scales, circumventing the computational cost associated with full 2D image-based approaches. Furthermore, the persistence (lifespan) of identified features serves as a natural indicator of significance: long-persisting structures generally correspond to physically meaningful signals, while short-lived features can reflect the noise level. This intrinsic property allows for straightforward and physically interpretable filtering based on signal-to-noise ratio, enhancing the robustness of topological summaries in cosmological analysis.

3.4 Topological structure dependence on neutrino mass

In this subsection, we check the persistent homology of the simulated universe using FLAMINGO simulations. We illustrate the basic features and the neutrino imprints on Betti curves and persistence strips for different mass components and different structures. We reach two preliminary conclusions. First, both the cold dark matter and all matter fields contain imprints of the neutrinos. Second, among all topological structures, the 2-cycles, which refer to shell-enclosing voids, are most sensitive to neutrino mass.

In Fig. 3, we show the Betti curves and persistence diagrams built under the superlevel filtration for different matter density fields of FLAMINGO realizations with different neutrino masses, 0.06, 0.24, and 0.48 eV, and other cosmological parameters fixed. The neutrino field has far fewer structures in total than the cold dark matter field and the total matter field. Lighter neutrinos further suppress neutrino structures, driving them closer to the mean density. In contrast, the topological structures of the cold dark matter field exhibit the opposite response, becoming more pronounced and moving further away from the mean density with lighter neutrino mass.

Fig. 4 presents the density superlevel sets and Betti curves of the neutrino field and cold dark matter field. The comparison directly reveals the key feature of the neutrino field relative to cold dark matter: it traces significantly larger scales and thus produces fewer structures overall. Moreover, voids in the neutrino field occur at markedly different density thresholds than their CDM counterparts. This behaviour directly reflects the formation and evolution of the hierarchical void population: neutrinos give rise to shallow, large-volume voids, in contrast to the intricate, multiscale cold dark matter voids (see e.g. R. K. Sheth & R. van de Weygaert 2004; E. Platen, R. van de Weygaert & B. J. T. Jones 2008).

The total matter field, as the combination of cold dark matter and neutrinos, exhibits Betti curves similar to those of the cold dark matter field. Counterintuitively, however, the differences between neutrino mass scenarios are slightly more pronounced in the total matter Betti curves, particularly in the low-density region where the relative neutrino density is higher and thus their contribution is more significant. This arises because the rescaling effect of neutrinos on the density field is substantially stronger than the clustering of neutrinos themselves. This will be further discussed in Section 6. Among all topological structures in the CDM and total matter fields, 2-cycles—which trace the shells enclosing voids—exhibit the greatest sensitivity to neutrino mass. This finding aligns with previous work showing that cosmic voids are more sensitive to neutrino mass than haloes (C. D. Kreisch et al. 2019; G. Zhang et al. 2020), while filaments also serve as a competitive probe (L. N. L. Simões et al. 2025; G. Rossi, H. Yu & M. Michaux 2026).

Unfortunately, we cannot directly detect or even trace the neutrino field in real observation data. Weak lensing observations probe the total matter distribution, including neutrinos, but generally in 2D projection, while galaxy clustering in redshift space is a biased tracer of the 3D cold matter distribution, which excludes neutrinos to first order. Thus, we will focus on the cold dark matter (D) field and total matter (T) field hereafter.

As shown in Fig. 3, β_1 structures (filaments) outnumber both β_0 (haloes) and β_2 (voids) by roughly a factor of two. This abundance aligns with the function of filaments in connecting haloes and surrounding voids. In our analysis, the total number and

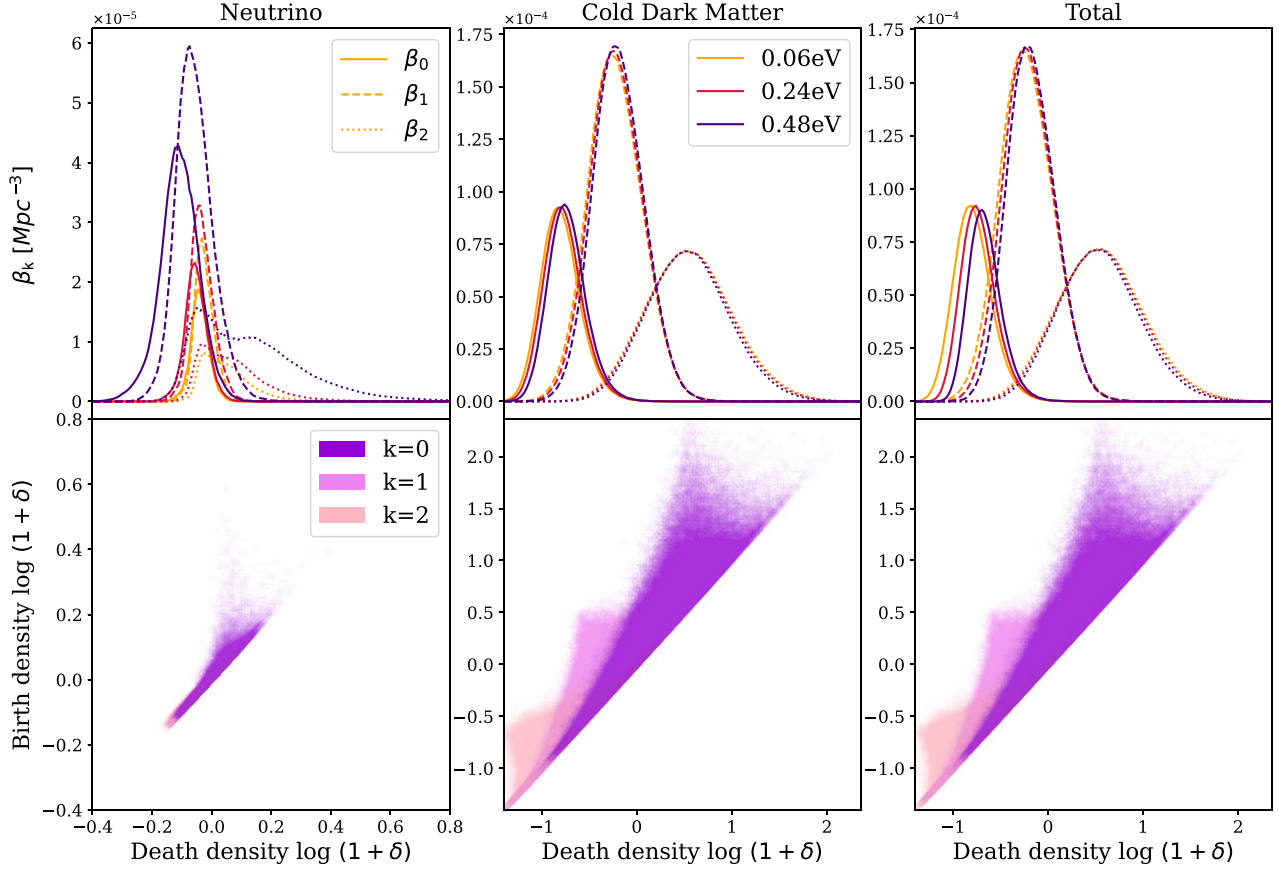


Figure 3. Topological structure of different mass components in the FLAMINGO simulations. Top panel: Betti curves β_k for the neutrino, cold dark matter, and total matter density fields. The curves show the count of connected components (β_0 , dotted lines), loops (β_1 , dashed lines), and voids (β_2 , solid lines) as a function of density threshold. Colours indicate the total neutrino mass: 0.06 eV (yellow), 0.24 eV (red), and 0.48 eV (purple). Bottom panel: Persistence diagrams for the 0.06 eV neutrino mass case, corresponding to the density fields shown in the top row. Each point represents a topological feature, with its birth and death coordinates indicating the density thresholds at which it appears and disappears. Structures in dimension $k = 0, 1$, and 2 are represented in purple, violet, and pink, respectively. This figure illustrates the imprint of neutrino mass on the multiscale topology of cosmic density fields, providing a qualitative basis for comparing the simulations.

peak height of β_2 features exceed those of β_0 . This contrasts with the findings of G. Wilding et al. (2021) and M. H. Jalali Kanafi et al. (2024), who report more persistent structures in $k = 0$ than in $k = 2$. We have verified that this is not due to the higher particle resolution, smaller smoothing scale, or differences between the simulations used in our analysis and those used in prior works. We obtain similar results when down-sampling the particles to match the resolution of these prior works, when increasing the smoothing scale to 5 Mpc, or when analysing snapshots from the Quijote simulations instead of FLAMINGO. We speculate that the difference may relate to asymmetry between superlevel and sublevel filtrations (see Section 3.1). Nevertheless, we observe that a large fraction of structures in $k = 2$ exhibit low persistence. If we adopt a persistence threshold of 0.2, the number of haloes exceeds that of voids.

In Fig. 5, we show the persistence strips of D and T fields across five persistence bins in FLAMINGO realisations with different neutrino masses. Rather than using identical persistence bins across dimensions, we manually select bin boundaries for each dimension to ensure roughly comparable Betti curve binning. This approach is motivated by Fig. 3, which shows that topological features of different dimensions exhibit distinct persistence ranges, with $k = 2$ structures displaying the largest spread. Both

the D and T fields exhibit a clear capacity to distinguish neutrino masses with filamentary and void structures (topological dimensions $k = 1, 2$) across all persistence thresholds π_p , with the T field generally providing higher discriminatory power. As persistence increases, the peak positions of β_{0,π_p} and β_{1,π_p} shift slightly toward higher densities, whereas β_{2,π_p} peaks at lower densities. This indicates that longer-lived haloes and filaments are associated with higher-density regions, while long-lived voids reside in lower-density environments.

Neutrino mass significantly influences the amplitude and shape of the topological distributions. Higher neutrino mass pushes the peak closer to the mean density across all three topological dimensions, as is the case for Betti curves. In low-persistence bins, higher neutrino mass leads to an increased number of structures. However, this trend reverses at high persistence: the amplitude decreases with increasing neutrino mass, suggesting that massive neutrinos suppress the formation of long-lived structures. A transition region between these regimes shows minimal sensitivity to neutrino mass. Notably, the relative difference in amplitude between neutrino masses is largest in the highest persistence bin.

In low-persistence bins, an increase in neutrino mass produces a tighter distribution of topological features. This ‘tightening’

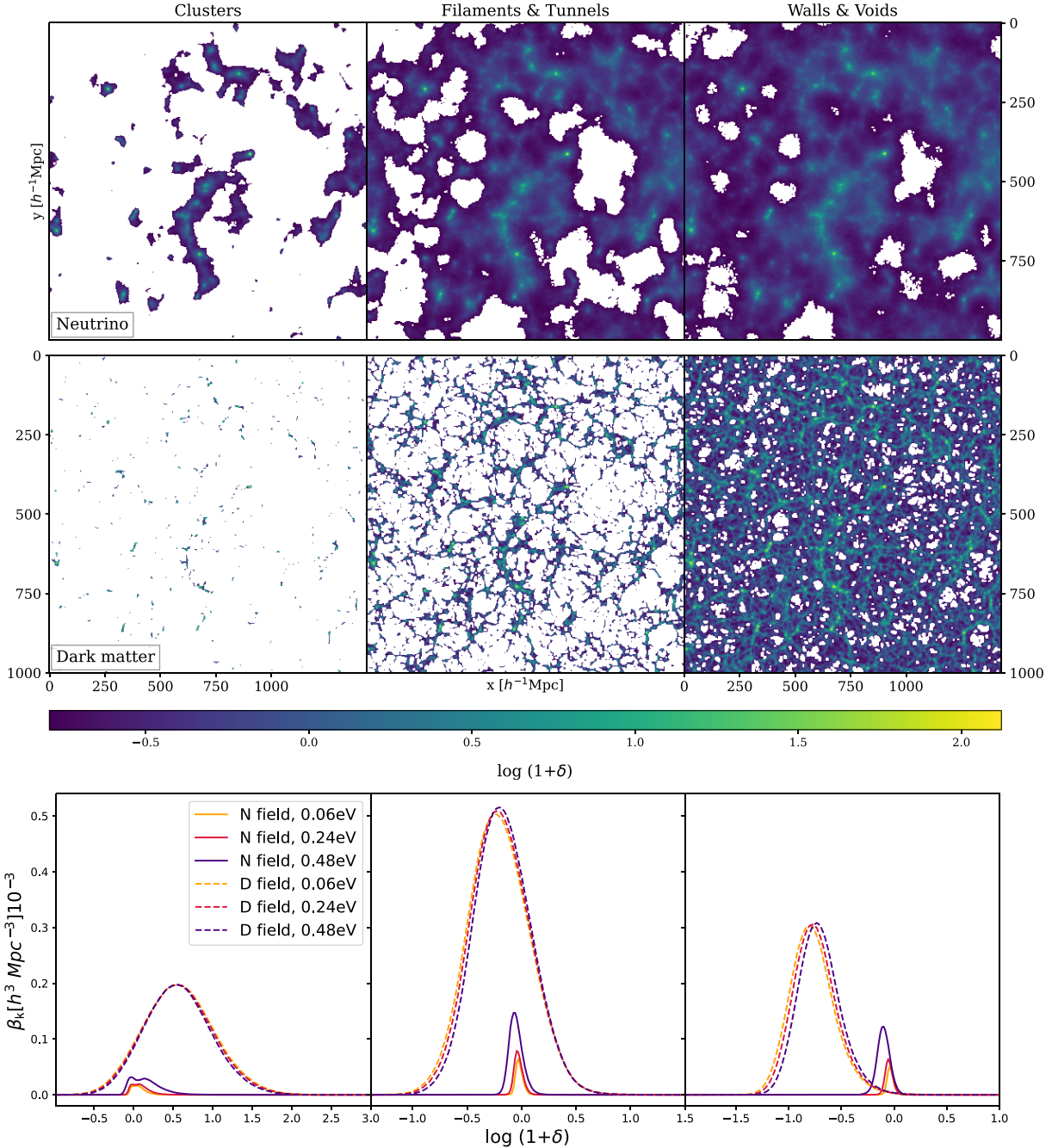


Figure 4. Density superlevel sets and Betti curves of different mass components and structures. Top panel: Superlevel sets of different mass components (neutrinos on the top and dark matter on the bottom) in a 6.8 Mpc slice around a height of 275 Mpc, in the baseline simulation. The structure is depicted as superlevel sets, adopting density thresholds equal to the peak positions of the respective Betti curves to illustrate the disjoint nodes of the cosmic web ($\vartheta = 3.53, 1.11$), its filamentary structure ($\vartheta = 0.56, 0.93$), and the walls enclosing the cosmic voids ($\vartheta = 0.16, 0.91$). Bottom panel: Betti curves as a function of density threshold for connected components (β_0 , left), loops (β_1 , middle), and voids (β_2 , right) of neutrino (solid lines) and cold dark matter (dashed lines).

pattern signifies that massive neutrinos compress the evolutionary timeline of transient structures: they cause structures to emerge at lower density thresholds and dissolve at higher densities during the filtration process. This behaviour implies a sharply peaked dissolution rate, which rises rapidly to a maximum before declining steeply. In contrast, high-persistence structures exhibit a flatter distribution across densities, reflecting a more uniform dissolution rate and highlighting the distinct evolutionary behaviour of structures with varying persistence.

4 EMULATION AND COSMOLOGICAL PARAMETER EXTRACTION

To evaluate the sensitivity of Betti curves to cosmological parameters and assess their capacity to break degeneracies among them, we construct an emulator based on Gaussian process regression across a 10D cosmological parameter space. This emulator enables efficient interpolation of Betti curves as functions of cosmological parameters and density thresholds, supporting

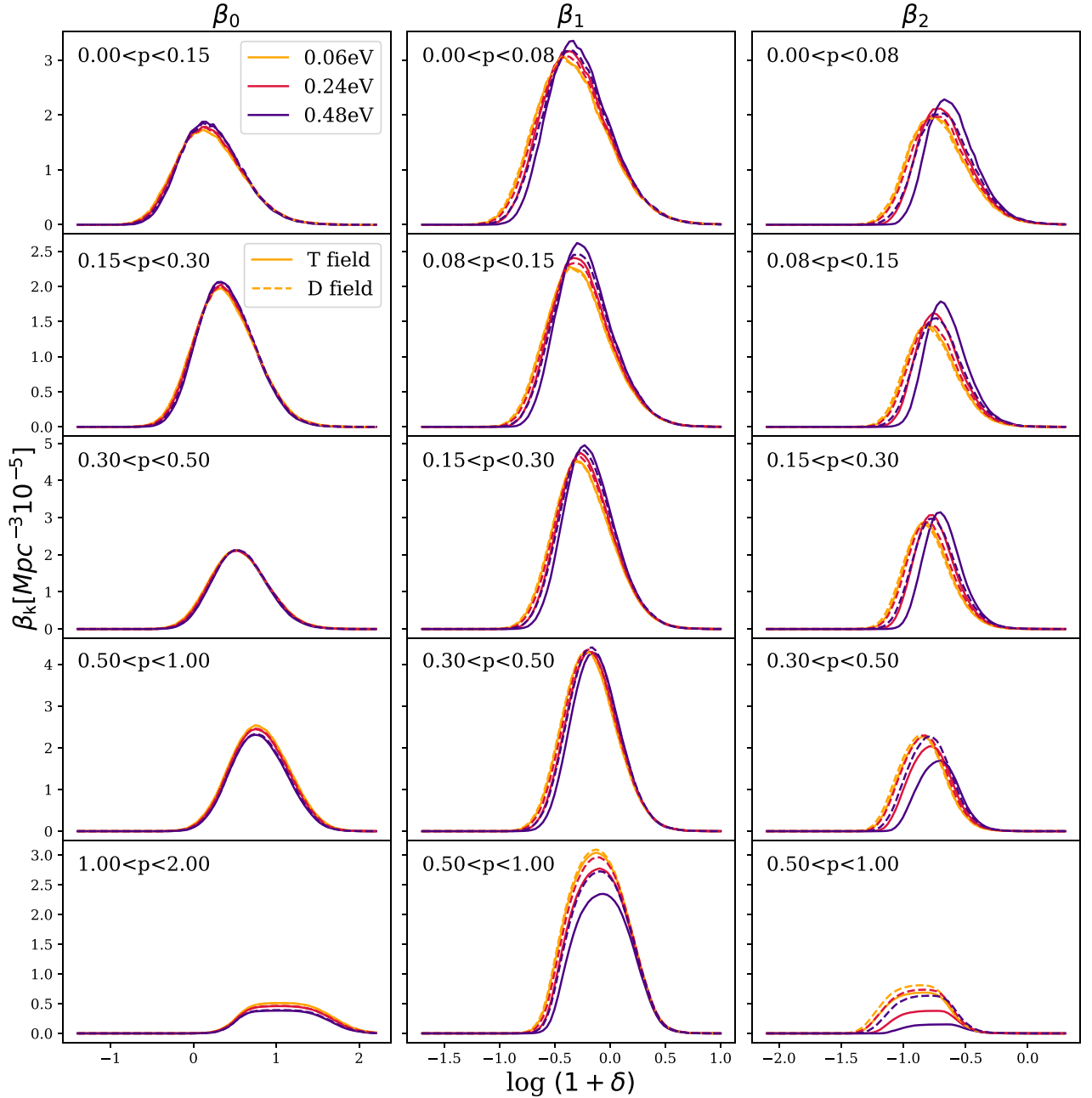


Figure 5. Persistence strips for different mass components in the FLAMINGO simulation, shown for connected components ($k = 0$), loops ($k = 1$), and voids ($k = 2$) from left to right. Each subpanel corresponds to a different persistence bin, as indicated. Solid and dashed lines represent the T field and D field, respectively. Colours indicate different neutrino mass sums: 0.06 eV (red), 0.24 eV (orange), and 0.48 eV (green).

subsequent cosmological inference through Bayesian posterior sampling and evidence estimation.

4.1 Emulator construction and training

Gaussian process (GP hereafter) provides a non-parametric, probabilistic framework for approximating smooth functions from finite training data. A GP defines a distribution of functions, characterized by a mean function $\mu(\theta)$ and a covariance kernel $k(\theta, \theta')$, which models the similarity between the input points θ and θ' . Given a training set of cosmological simulations with parameter sets $\{\theta\}_j$, where j refers to the training samples index, the

GP predicts the function value of persistence strips $\beta_{d,\pi_p}(\vartheta; \{\theta\}_j)$ at a new point $\{\theta\}_*$ as a Gaussian distribution,

$$\beta_{d,\pi_p}(\vartheta; \{\theta\}_*) \sim \mathcal{GP}(\mu(\theta_*), k(\theta, \theta_*) + \sigma_n^2 I), \quad (5)$$

where σ_n^2 accounts for observational and modelling noise. For a more detailed discussion of GP and its features, see C. E. Rasmussen & C. K. I. Williams (2006).

We model each persistence strip across death density thresholds ϑ using an independent GP. To incorporate the density dimension, we treat ϑ as an additional input parameter to the 10 cosmological parameters, allowing the emulator to smoothly interpolate topological measurements across both cosmological

and density spaces. The full emulator comprises 15 independent 11D GPs (one for each of the three topology dimensions and five persistence cuts), each yielding predictions with quantified uncertainties.

The choice of representation for the target observable significantly influences emulation accuracy. For instance, the CosmicEmu (E. Lawrence et al. 2010) represents the matter power spectrum as $k^{3/2}P(k)$, the Euclid Emulator (Euclid Collaboration 2019) and the CSST Emulator (Z. Chen et al. 2025) improve performance by dividing out a reference spectrum. In this work, we introduce a cumulative representation of the persistence strips to enhance smoothness (especially for the lower density slope) and improve emulation efficacy. The cumulative statistic is defined as:

$$\text{CUM}_{k,\pi_p}(x) = \log_{10} \left(\sum_{n=x}^{\infty} \beta_{k,\pi_p}(\vartheta_n) + C \right), \quad (6)$$

where $C = 500$ is a constant added to avoid logarithmic numerical instabilities. This transformation significantly improves the smoothness and differentiability of the target function, leading to more robust GP interpolation. A similar cumulative approach has also been adopted in the CSST Cosmological Emulator for halo mass function emulation, where it is shown that the cumulative form is independent of the binning scheme and yields improved emulation performance (Z. Chen & Y. Yu 2025).

For the Gaussian kernel choice, we adopt a composite kernel $k(x, x') = k_{\text{constant}} + k_{\text{Matern32}}(x, x')$ by combining a constant kernel and the Matérn32 kernel,

$$k_{\text{Matern32}}(r^2) = \sigma_f^2 \left(1 + \sqrt{3r^2} \right) \exp \left(-\sqrt{3r^2} \right), \quad (7)$$

where $r = (x - x')/l$, l is the characteristic length-scale governing the smoothness of the function, σ_f^2 represents the signal variance, and r denotes the scaled Euclidean distance between input points. The constant kernel captures global, non-zero mean trends in the function, while the Matérn32 kernel models smooth, but not infinitely differentiable, local variations. The hyperparameters of the kernel, including σ_f and l , are optimized by maximizing the log marginal likelihood,

$$\ln p(\mathbf{y} | X) = -\frac{1}{2} \mathbf{y}^\top K^{-1} \mathbf{y} - \frac{1}{2} \ln |K| - \frac{n}{2} \ln 2\pi, \quad (8)$$

where $K = k(x, x') + \sigma_n^2 I$ is the covariance matrix of the observed data, $K(X, X)$ is the kernel matrix evaluated at the training points, σ_n^2 is the noise variance, and \mathbf{y} is the vector of observed function values.

In this work, we employ the PYTHON code GEORGE3 developed by S. Ambikasaran et al. (2015)⁴ to develop our emulator. The emulator is constructed based on the 50 simulations as outlined in Section 2.2, where 40 simulations are used as the primary training sample, with the remaining 20 per cent of simulations held out as the testing set to validate performance. To ensure robustness against specific random subdivisions, this train-test partitioning process was repeated multiple times with different random seeds. The cosmological parameter space explored in this work is defined by the 10D Latin hypercube design detailed in the top panel of Table 1. In addition to this baseline parametrization, we systematically tested alternative parametrizations by substituting $\ln A_s$ with σ_8 . A comparative analysis of emulator performance under the different parametrization is provided in Appendix B.

⁴<http://george.readthedocs.io/en/latest/>

We find σ_8 yields much higher emulation accuracy to $\ln A_s$, indicating its strong correspondence with topological features of the cosmic web.

4.2 Emulator accuracy

Performance was assessed using an iterative holdout validation: in each iteration, a random 20 per cent of the simulations were withheld as a test set, and the emulator was retrained on the remaining 80 per cent. This process was repeated over 10 iterations to ensure statistical robustness. The accuracy of the emulator was quantified using the relative residual of the cumulative statistic:

$$\chi_{\text{CUM}_{k,\pi_p}(x)} = \frac{\text{CUM}_{k,\pi_p}(x) - \langle \text{CUM}_{k,\pi_p}(x) \rangle_j}{\langle \text{CUM}_{k,\pi_p}(x) \rangle_j}, \quad (9)$$

where j represent the repeating iterations and x denotes the density threshold used for cumulating. Our analysis focuses on the accuracy of the cumulative persistence strips, as this is the format directly utilized in our cosmological parameter inference. We have examined the residuals of the non-cumulative (original) format, which, as expected, exhibit numerical instabilities, particularly at the edges of the curves.

Fig. 6 displays the corresponding residuals, with individual realisations shown as coloured lines. Dark and light shaded regions indicate the 68 and 95 per cent confidence intervals, respectively, across all test samples. The emulator achieves high accuracy, with residuals ≤ 1 per cent generally. For the structures in the highest persistence bin or with $k = 2$, the error is relatively higher, reflecting the greater dynamic range and structural complexity in those regimes. These results demonstrate the emulator's reliability across a wide range of topological scales and cosmological models.

4.3 Bayesian posterior and evidence estimation

For cosmological parameter inference, we employ the Nautilus sampler (J. U. Lange 2023)⁵, a Bayesian nested sampling algorithm that utilizes neural network-based importance sampling to significantly reduce the number of required likelihood evaluations. This approach efficiently estimates both posterior distributions and Bayesian evidence while incorporating all likelihood evaluations during the sampling process. We adopt flat priors on all cosmological parameters, with ranges identical to those used in the Latin hypercube design of our training simulations. For this analysis, we find that 6000 live points provide sufficient sampling accuracy and stability.

The covariance matrix used in the likelihood evaluation combines contributions from both cosmic variance and emulator prediction errors. Cosmic variance is estimated via bootstrap resampling of the FLAMINGO 2.8, Gpc simulation box. The simulation, configured with the same resolution and smoothing scales as the emulator, is divided into 512 sub-boxes of 350, Mpc³ each, from which the covariance is robustly estimated. The emulator uncertainty is incorporated as a diagonal matrix, where each diagonal element corresponds to the Gaussian process prediction error in the respective density bin.

To incorporate limited cosmological information from the CMB, we add the CMB prior term to the total likelihood function

⁵<https://nautilus-sampler.readthedocs.io/en/stable/>

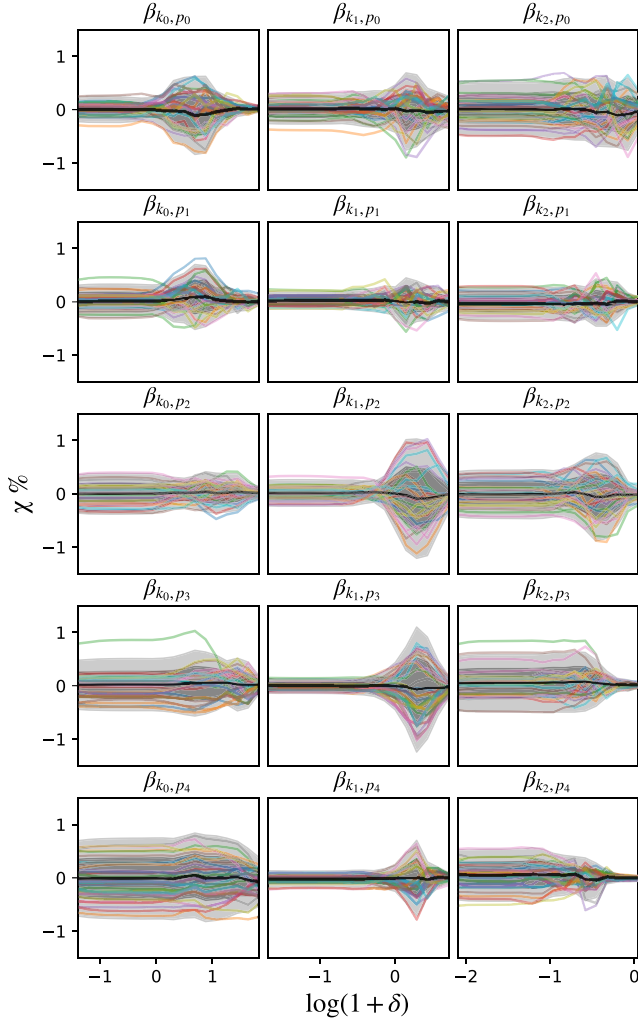


Figure 6. The emulator accuracy of selected testing samples. Different panels show results for different dimensions and different persistence ranges as indicated on top of each panel. Each coloured line represents one testing sample, and the black solid line shows the average error between the prediction and the truth. The dark and light grey shaded regions show the 68 and 95 per cent confidence intervals, respectively.

following M. Abdul Karim et al. (2025),

$$\chi_{\text{CMB}}^2 = \begin{pmatrix} \theta_{*,100} - \theta_{*,100}^{\text{true}} \\ \omega_{\text{cdm}} - \omega_{\text{cdm}}^{\text{true}} \\ \omega_b - \omega_b^{\text{true}} \end{pmatrix}^T C^{-1} \begin{pmatrix} \theta_{*,100} - \theta_{*,100}^{\text{true}} \\ \omega_{\text{cdm}} - \omega_{\text{cdm}}^{\text{true}} \\ \omega_b - \omega_b^{\text{true}} \end{pmatrix}, \quad (10)$$

where $\omega_b \equiv \Omega_b h^2$, $\omega_{\text{cdm}} \equiv \Omega_{\text{cdm}} h^2$, and $\theta_{*,100}$ is 100 times the angular acoustic scale of the CMB power spectrum, defined as $\theta_* = r_*/D_M(z_*)$, where r_* is the comoving sound horizon at recombination and $D_M(z_*)$ is the comoving angular diameter distance to the last scattering surface at redshift z_* (P. Lemos & A. Lewis 2023). The true value of these parameters is calculated from the cosmology of the simulation to be estimated. The covariance matrix C of these three parameters is derived conservatively from *Planck* and ACT data, following W. Elbers et al. (2025b). This term ensures consistency between our inferred parameters and early-universe constraints, thereby enhancing the robustness of the cosmological inference.

5 RESULTS

In this section, we present an analysis of the cosmological dependence captured by our emulator and discuss the resulting cosmological constraints derived from these topological summaries. We examine how this dependence varies across different cosmic environments and investigate the cosmological inference for key parameters from different probes and different fields.

5.1 Cosmological parameter dependence

Among the ten cosmological parameters considered, we find that the neutrino mass sum M_ν , the matter density Ω_m , the amplitude of matter fluctuations σ_8 , and the Hubble constant h exert the most pronounced influence on the topology of the cosmic web, as quantified by the Betti curves. As illustrated in Fig. 7, each parameter affects the Betti curves in a distinct way, with other parameters fixed at the centre of the hypercube. Persistence strips shown in each row correspond to different persistence ranges, as defined in Fig. 5.

The topological statistics of clusters ($k = 0$), displayed in the first column of each panel in Fig. 7, exhibit a comparatively modest response to variations in cosmological parameters. The most substantial influences arise from Ω_m and σ_8 , followed by a weaker dependence on h ; the response to M_ν is nearly negligible (at fixed Ω_m and σ_8). The matter density Ω_m systematically enhances the abundance of cluster structures across all persistence levels, with a magnitude that scales with the persistence value. In contrast, the amplitude of matter fluctuations σ_8 exerts a more complex influence: it suppresses the number of low-persistence clusters while enhancing the population of high-persistence ones. This redistribution consequently shifts the peak of the low-persistence distribution to lower density thresholds. The Hubble parameter h exhibits the converse trend to σ_8 ; an increase in h shifts structures from high to low persistence bins and moves the peak of the low-persistence distribution toward higher densities. In the case of neutrino mass, the $k = 0$ response is too small to be statistically significant.

Filamentary structures ($k = 1$), as illustrated in the second column of each panel in Fig. 7, generally encode richer cosmological information than $k = 0$. Among the four cosmological parameters studied, σ_8 exhibits the most pronounced influence, followed by Ω_m , with h and M_ν showing comparatively weaker but still discernible effects. An increase in neutrino mass M_ν suppresses the abundance of high-persistence filaments while enhancing low-persistence ones, accompanied by a systematic peak shift toward higher density thresholds. In contrast, a higher matter density Ω_m primarily boosts the population of high-persistence filaments and shifts the distribution peak toward lower densities. The amplitude of matter fluctuations σ_8 displays a trend consistent with that observed in the cluster environment, promoting the transition of structures from low to high persistence bins. Similarly, a larger Hubble parameter h results in a sharper filament distribution across all persistence levels and an increase in the number of high-persistence filaments.

The topological statistics of cosmic voids ($k = 2$), presented in the right-most column of each panel in Fig. 7, exhibit the most complex and informative behaviour among all topological components, showing strong dependence on all four cosmological parameters. The influence of neutrino mass is particularly striking, characterized by a combined ‘shifting’ and ‘narrowing’ pattern across persistence bins. A higher M_ν increases the abundance of

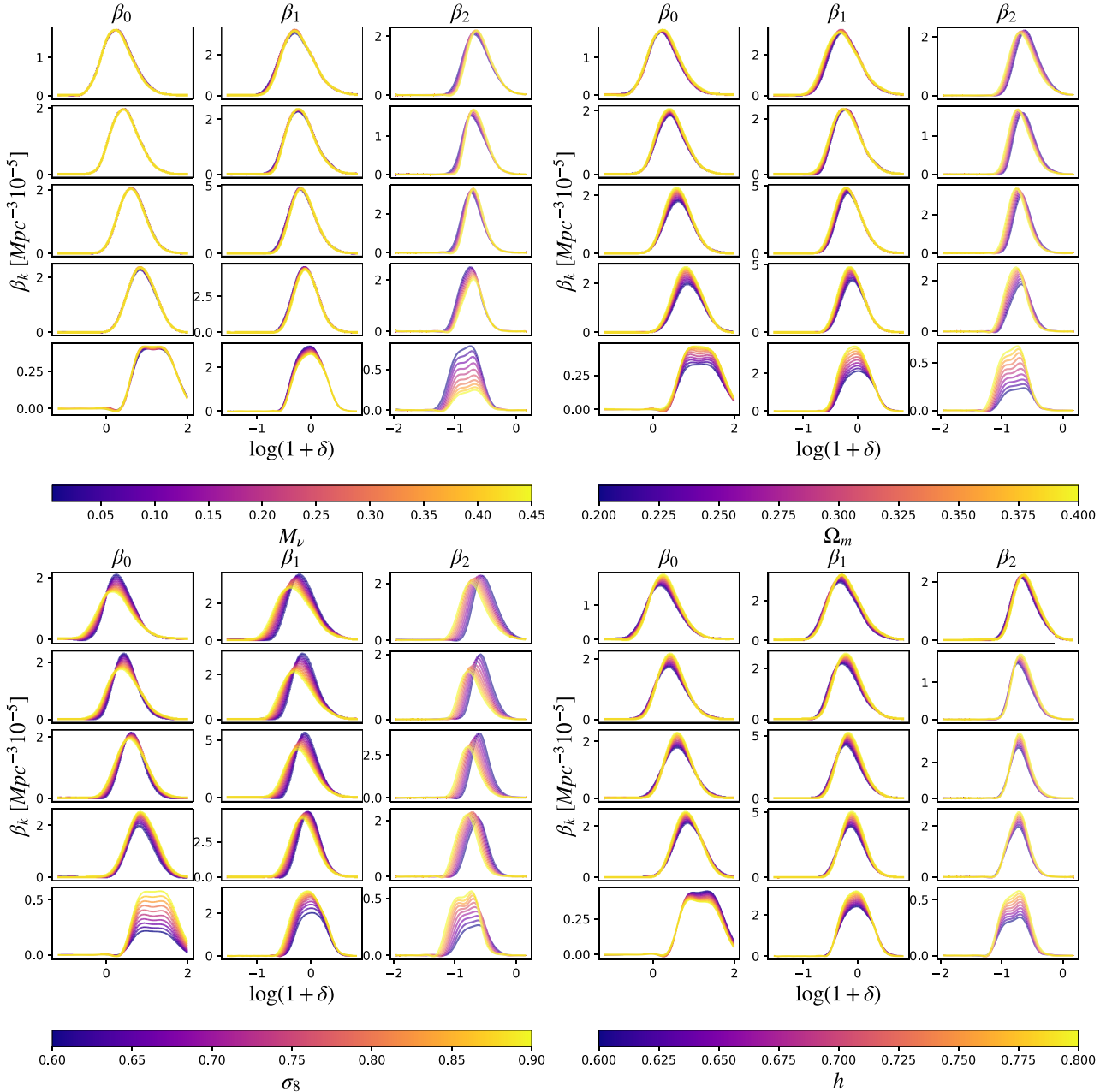


Figure 7. Dependence of persistence strips on key cosmological parameters, including the neutrino mass M_ν (upper left), the matter density Ω_m (upper right), the amplitude of matter fluctuations σ_8 (bottom left), and the Hubble constant h (bottom right). All other parameters are fixed at the centre of the hypercube. In each panel, the five rows correspond to the five persistence ranges as in Fig. 5, illustrating how the sensitivity to each parameter varies with topological persistence. Colours in each panel represent the varying value of the parameter of interest, as indicated by the colour bar.

low-persistence voids and shifts their distribution toward higher density thresholds, while simultaneously suppressing the number of high-persistence voids. This can be understood as a topological manifestation of neutrino free-streaming: the diffusive effect of heavier neutrinos elevates the density floor in deep potential wells, effectively fragmenting large, persistent voids into smaller, shallower ones. Similarly, the matter density Ω_m exerts a strong influence, significantly enhancing the number of high-persistence voids while leaving low-persistence counts largely unaffected. This is accompanied by a systematic shift of the entire distribution toward lower density thresholds. The impact of σ_8

on void statistics follows a similar pattern to that observed in filaments but with slightly larger variance. The Hubble parameter h also shows a trend analogous to its effect in the filamentary environment.

Several cosmological parameters, including the baryon fraction f_b , the scalar spectral index n_s , the running of the spectral index α_s , exhibit only a minor influence on the Betti curves. Similarly, parameters associated with decaying dark matter (Γ_{ddm}) and standard dark energy evolution (w_0, w_a) show negligible impact across all topological features. A comprehensive presentation of the effects of these parameters is provided in Appendix C. Due to

Table 2. Constraints on key cosmological parameters from topological persistence analysis. Median values with 68 per cent credible intervals for Ω_m , h , σ_8 , and M_ν are presented for a range of analyses, including different topological dimensions (k), persistence bins (τ_p), density fields, and data combinations (with/without CMB priors), all with 2 Mpc smoothing. The true parameter values are shown at the top of the table.

Probe	CMB priors	Ω_m 0.233	h 0.655	σ_8 0.607	M_ν/eV 0.236
$\beta_{k_0, p_{01234}}$	✓	$0.229^{+0.029}_{-0.039}$	$0.657^{+0.063}_{-0.036}$	$0.607^{+0.069}_{-0.032}$	$0.145^{+0.174}_{-0.092}$
$\beta_{k_1, p_{01234}}$	✓	$0.220^{+0.037}_{-0.033}$	$0.671^{+0.054}_{-0.046}$	$0.618^{+0.058}_{-0.036}$	$0.195^{+0.141}_{-0.105}$
$\beta_{k_2, p_{01234}}$	✓	$0.221^{+0.033}_{-0.026}$	$0.670^{+0.042}_{-0.041}$	$0.615^{+0.058}_{-0.035}$	$0.238^{+0.057}_{-0.061}$
$\beta_{k_0, p_{123}}$	✓	$0.229^{+0.031}_{-0.041}$	$0.658^{+0.065}_{-0.038}$	$0.607^{+0.095}_{-0.034}$	$0.146^{+0.177}_{-0.100}$
$\beta_{k_1, p_{123}}$	✓	$0.212^{+0.038}_{-0.025}$	$0.683^{+0.044}_{-0.050}$	$0.625^{+0.070}_{-0.039}$	$0.183^{+0.131}_{-0.093}$
$\beta_{k_2, p_{123}}$	✓	$0.218^{+0.039}_{-0.031}$	$0.675^{+0.052}_{-0.048}$	$0.613^{+0.106}_{-0.041}$	$0.228^{+0.100}_{-0.097}$
$\beta_{k_{012}, p_{123}}$	✓	$0.226^{+0.033}_{-0.038}$	$0.663^{+0.060}_{-0.041}$	$0.612^{+0.053}_{-0.029}$	$0.234^{+0.059}_{-0.072}$
$\beta_{k_{012}, p_{01234}}$	✓	$0.216^{+0.036}_{-0.027}$	$0.678^{+0.044}_{-0.045}$	$0.620^{+0.035}_{-0.032}$	$0.233^{+0.049}_{-0.050}$
$\beta_{k_{012}, p_{123}}$		$0.227^{+0.046}_{-0.037}$	$0.682^{+0.080}_{-0.068}$	$0.616^{+0.042}_{-0.037}$	$0.252^{+0.072}_{-0.063}$
$\beta_{k_{012}, p_{01234}}$		$0.230^{+0.045}_{-0.033}$	$0.684^{+0.072}_{-0.073}$	$0.617^{+0.035}_{-0.039}$	$0.254^{+0.057}_{-0.046}$

the insensitivity of most of the parameters, we will focus on the four main parameters in the following analysis.

It is essential to leverage the distinct responses of topological components to break degeneracies among cosmological parameters. For instance, filamentary structures (β_1) respond in opposite directions to changes in M_ν and σ_8 , suggesting a potential degeneracy between neutrino mass and matter clustering amplitude. The degeneracy can be effectively broken by leveraging the complementary parameter sensitivities of different topological environments, namely, by jointly exploiting the clustered (cluster, $k = 0$) and underdense (void, $k = 2$) structures. This multiscale approach enhances the overall constraining power and reduces parameter uncertainties.

5.2 Cosmological parameters inference

In this subsection, we first analyse the four parameters of primary interest (Ω_m , h , σ_8 , M_ν), then compare the constraints on all parameters obtained with and without CMB priors. It should be noted that no observational noise has been added in this section. Consequently, the accuracy of the recovered cosmological parameters is limited by both the emulator and the construction of the covariance matrix, which is based on a volume of $(350 \text{ Mpc})^3$. We randomly select one of the testing samples not used in the emulator training as the example. The inferred inference of the four primary cosmological parameters are listed in Table 2.

5.2.1 Four most influential parameters

We first focus on the posteriors of Ω_m , h , σ_8 , and M_ν based on the topological measurement from different dimensional probes. The constraints separately from each dimension k and jointly are presented in Fig. 8, respectively. For clarity, we only show the $k = 0$ and joint analysis cases in 2D subpanels, with true values in grey and the median value of the joint analysis in orange. We include the results from $k = 1$ and $k = 2$ for comparison in the 1D marginalized distributions.

Consistent with the dependence analysis in Section 5.1, all persistence strips individually provide different levels of constraining power on these four parameters. In general, $k = 2$ provides the tightest constraints within all dimensions, and

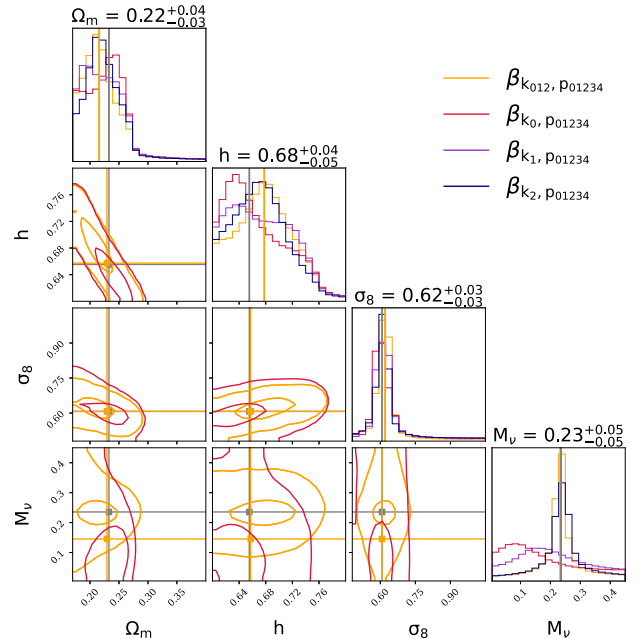


Figure 8. Cosmological constraints from persistence analysis. Confidence contours for Ω_m , h , σ_8 , and M_ν using intermediate persistence strips across dimensions $k = 0, 1, 2$ individually and jointly with CMB priors. For clarity, we only show the $k = 0$ (red contour) and joint analysis (orange contour) cases in 2D subpanels, with true values in grey and the median value of the joint analysis in orange. We also include the results from $k = 1$ (purple) and $k = 2$ (navy) in the 1D subpanels. Labels above each column indicate the median and 68 per cent credible intervals for $\beta_{k_{012}, p_{01234}}$.

combining persistence strips in all three dimensions yields tighter constraints.

For Ω_m and σ_8 , the constraining power is comparably good across all 3D topological components, showing no significant differences beyond the 1σ level. Combining 3D topological measurements tightens the constraints, indicating that the information on the matter density is encoded across multiple topological dimensions and is complementary to each other. For h , the constraints are slightly looser, as predicted in Section 5.1. The degeneracy between Ω_m and h persists across all probes and proves

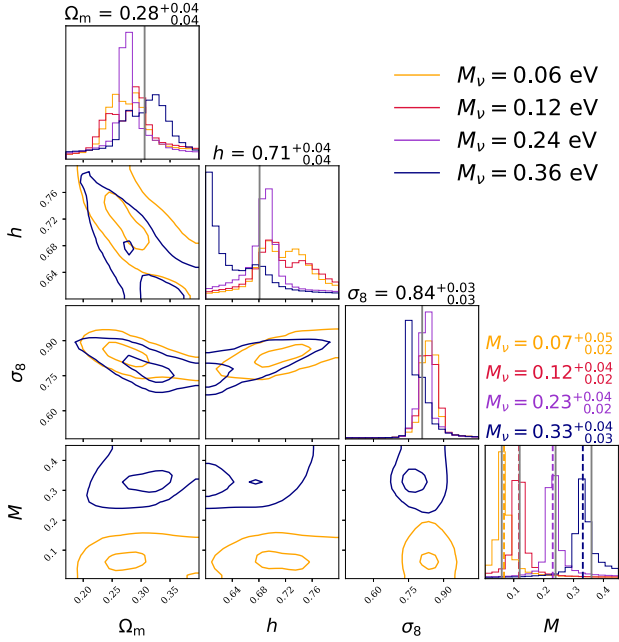


Figure 9. Similar to Fig. 8, but showing the cosmological inference dependence on the true neutrino mass. For clarity, we only show the smallest and largest neutrino mass cases in the 2D subpanels (in orange and blue, respectively). We include the results from all four cases in the 1D subpanels, with true values in grey and the median value of each analysis in the corresponding colour. Labels above each column show the median and 68 per cent credible interval: for the left three columns, this is the 0.06 eV case; for the rightmost column, it includes all four cases, with the colour the same as the corresponding lines.

difficult to break using topological information alone. This can be understood from Fig. 7, which shows that their dependence patterns are nearly identical. These results partially contrast with previous studies: A. Ouellette & G. Holder (2025) reported that Ω_m and σ_8 are constrained primarily by β_0 and β_1 , with little contribution from β_2 , while J. Calles et al. (2025) found clusters and voids to be most informative for Ω_m . Our analysis indicates that all three environments contribute significantly to constraining both Ω_m and σ_8 .

The constraints on the neutrino mass M_ν are dominated by voids ($k = 2$). Haloes ($k = 0$) contribute negligible information, while filaments ($k = 1$) offer only modest improvements. Voids give 5 times tighter constraints on M_ν than haloes (as shown in Table 2). Combining the 2D probes will tighten the constraints further to an uncertainty of 0.05 eV. A notable degeneracy between σ_8 and M_ν is evident when using haloes only ($k = 0$), while the degeneracy is effectively broken upon the inclusion of the void-based measurement, underscoring the value of multi-dimensional topological probes in breaking the parameter degeneracies and enhancing the robustness of cosmological inference.

In Fig. 9, we assess the dependence of constraining power on the true neutrino mass using four additional simulations in which only the neutrino mass is varied. The constraining power remains generally stable, with slightly looser constraints near the edges of the hypercube, which is expected. Other parameters exhibit consistent performance, confirming the robustness of our results. Given this validation and to maintain brevity, we omit results from other test samples or random seeds used in emulator

training. We note that here the constraint on neutrino mass is slightly tighter than in the case shown above. This may be because the other parameters lie closer to the centre of the hypercube, where emulator performance is typically better.

We identify three primary factors that likely contribute to the discrepancies between our constraints and those of M. H. Jalali Kanafi et al. (2024), who found an uncertainty of 0.015 eV. First, a direct comparison requires caution due to differing covariance estimations: our analysis employs 512 subboxes of $(350 \text{ Mpc})^3$, whereas theirs utilizes 5000 simulations with a $1 \text{ Gpc}^3 h^{-3}$ box size, leading to substantially different volumes and sampling variance. Second, the non-Gaussianity of the posterior distribution is a plausible explanation for the further discrepancy with the Fisher forecasts. Finally, our analysis is generalized to a wider cosmological parameter space that also varies parameters like the dark energy equation of state, dark matter lifetime, and running of the spectral index.

We present the constraints derived using only the three intermediate persistence bins ($p = 1, 2, 3$) in Table 2. This conservative selection excludes the lowest and highest bins ($p = 0$ and 4), which exhibited less stable performance in our emulator and, in the case of the lowest bin, greater susceptibility to shot noise. Even with this restricted data set, we obtain a robust neutrino mass constraint with an uncertainty of 0.06-0.07 eV. This result confirms that the primary cosmological signal does not reside in the lowest-persistence bin, reinforcing the practical applicability of our method to future observational data where such noise-sensitive regimes may be excluded.

5.2.2 Other parameters

We present constraints on all the cosmological parameters from our topological statistics, comparing inferences with and without CMB priors in Fig. 10. Topology alone yields constraints on most parameters, though with varying levels of precision; however, the inclusion of CMB priors significantly improves their tightness. The most pronounced improvement is seen for the baryon fraction, f_b , which exhibits a substantial gain in both accuracy and precision. This is expected, as the CMB priors directly provide tight constraints on the physical baryon density ω_b and the angular acoustic scale θ_* .

The CMB priors introduce a tight degeneracy between the dark energy equation of state parameters, w_0 and w_a , but the posteriors remain significantly biased from the true values. This bias likely stems from residual degeneracies with other poorly constrained parameters, such as n_s , which is itself biased. The spectral index n_s is slightly improved by the CMB prior but remains outside the 1σ confidence level. The running of the spectral index α_s shows no significant improvement with the inclusion of CMB priors. The decaying dark matter parameter Γ_{dcmm} remains entirely unconstrained in both cases.

For the four key parameters, the CMB priors yield more nuanced effects: The constraints on Ω_m , σ_8 , and M_ν show marginal improvements. The constraints on h improve significantly from a bimodal distribution to a clearly defined Gaussian peak.

5.3 Persistence strips over Betti curves

To maximize the information extracted from persistence diagrams and to facilitate the emulation of topological features, we

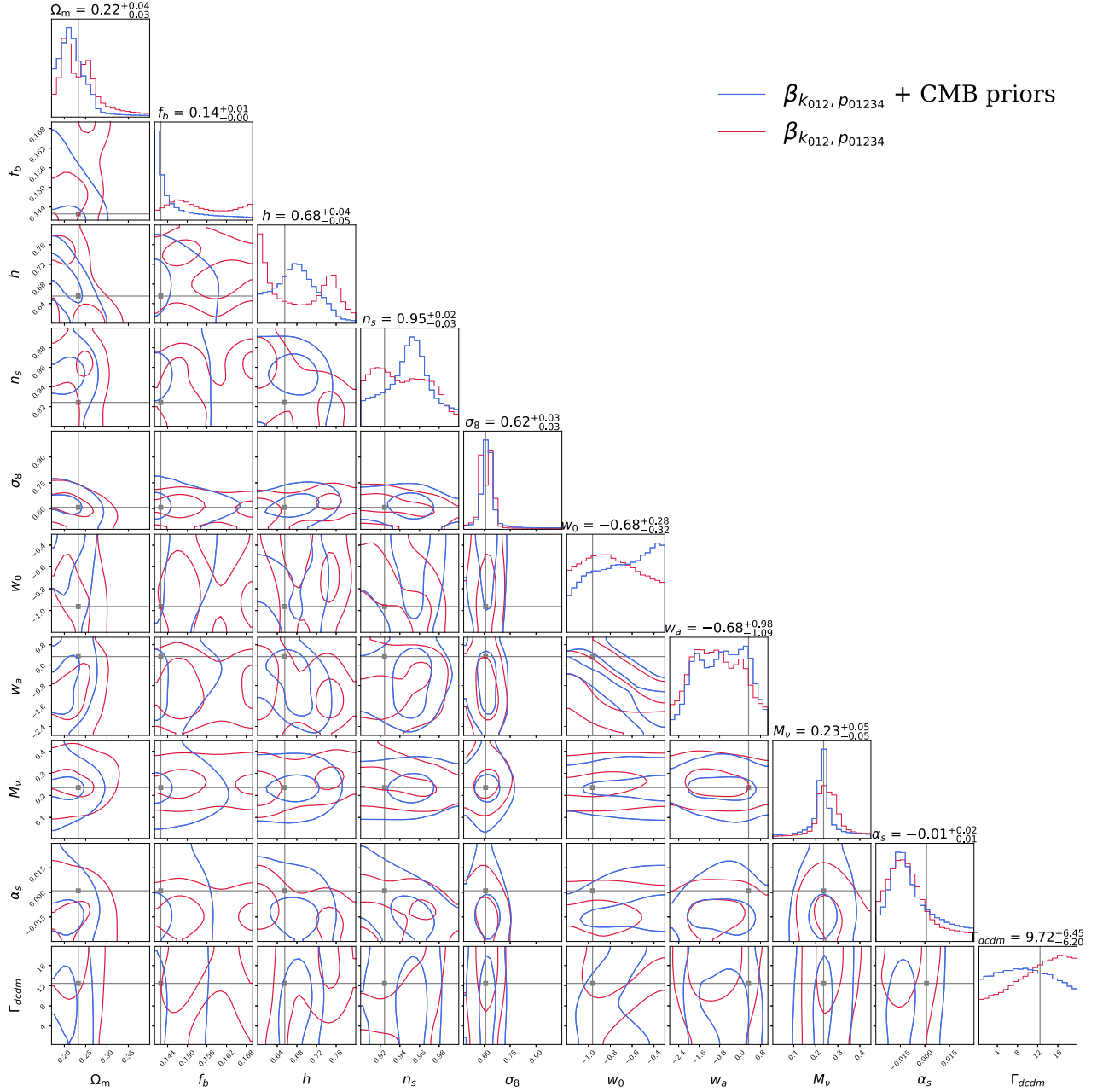


Figure 10. The contour plot of the constraints on all ten parameters in the emulator, using persistence strips with (blue contours) and without the CMB priors (red contours). The grey lines indicate the true value of each parameter. Labels in each column show the 1σ constraints with CMB priors.

have introduced the persistence strips framework. This approach transforms the complex topological data into a set of 1D curves binned by persistence, which are more amenable to interpretation and emulation than the full diagrams.

To quantify the advantage of this persistence-based binning, we construct a separate emulator using the standard Betti curves from the same simulation suite. A comparison of the cosmological constraints from both methods is presented in Fig. 11. While both the standard Betti curves and persistence strips successfully break the key degeneracy between M_ν and σ_8 , the persistence strips achieve about two times tighter constraints on M_ν (see Table 3). The marginally weaker constraints on Ω_m and σ_8 from

the persistence strips as shown in Fig. 11 occurred in only a single instance among our ten test samples. Generally, the two probes show similar performance. In contrast, the improvement in the constraint for the neutrino mass M_ν was systematic and robust across the entire test set.

This advantage is expected to be even more pronounced in real observational data. In such cases, density fluctuations and observational noise can significantly bias low-persistence structures. The inherent ability of persistence strips to separate the robust, high-persistence signal from the noisy, low-persistence features will therefore be critical for robust cosmological inference.

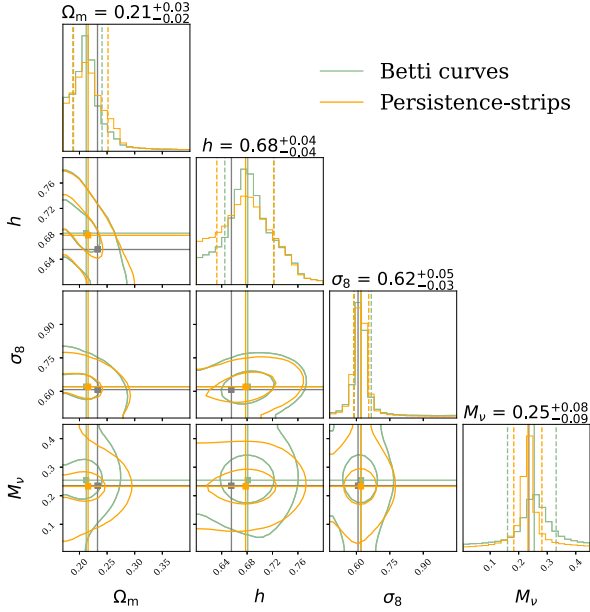


Figure 11. Similar to Fig. 8, but comparing the cosmological inference from persistence strips (orange contours) with that from normal Betti curves (green contours), all with CMB priors. The grey lines indicate the true value of each parameter. Labels on top of each column indicate the 1σ confidence level of the normal Betti curves.

5.4 Inference from dark matter-only field

Notably, when applying the emulator to the dark matter-only (D) field, we observe significantly weaker constraints on $\sum m_\nu$, across all persistence strips compared to the distinct signal identified in the total matter (T) field, as shown in Fig. 12. The constraint obtained from the D field is 0.13 eV, which aligns more closely with the Fisher forecast of 0.12 eV reported in M. H. Jalali Kanafi et al. (2024). In contrast, constraints on other parameters, such as σ_8 and Ω_m , do not show significant differences between the D and T fields.

We hypothesize that this discrepancy arises from a stronger degeneracy between neutrino mass and other cosmological parameters in the D field, which hinders the ability of topological statistics to cleanly isolate the neutrino signature. This interpretation is consistent with the findings of M. H. Jalali Kanafi et al. (2024). We further explore the physical origin of this degeneracy in Section 6.

6 PHYSICAL ORIGINS OF NEUTRINO MASS SENSITIVITY

A key objective of this work is to understand the physical mechanism that breaks this cosmological degeneracy with the ‘all-matter’ field T. In this section, we discuss and check the possible physical reasons for the sensitivity of the T-field’s persistence strips to neutrino mass. We conclude that the overall neutrino mass fraction and the distribution of the cold dark matter are equally important, while the distribution of the neutrinos is less important.

6.1 Neutrino fraction or field?

As can be seen in Fig. 3, the contribution of the distribution of neutrinos is subdominant compared to that of the suppression of the D field. To isolate the effect of the distribution of neutrinos,

we construct a ‘Fake all-matter’ field. This field is identical to the ‘True all-matter’ field except that we replace the true neutrino field with its mean value, i.e. $\delta_\nu = 0$. The field F only depends on neutrino masses in two ways: through the global average neutrino density, as a fraction of the total matter density, and through the indirect impact on the distribution of dark matter. This setup allows us to compare four distinct density fields to precisely determine the source of the constraining power. The cold dark matter (D) and neutrino (N) fields are simply

$$1 + \delta_D = 1 + \delta_{\text{cdm}}, \quad (11)$$

$$1 + \delta_N = 1 + \delta_\nu. \quad (12)$$

The true total matter field (T) is their weighted average:

$$\begin{aligned} 1 + \delta_T &= 1 + \frac{\Omega_{\text{cdm}}}{\Omega_m} \delta_{\text{cdm}} + \frac{\Omega_\nu}{\Omega_m} \delta_\nu \\ &= (1 + \delta_{\text{cdm}}) - \frac{\Omega_\nu}{\Omega_m} (\delta_{\text{cdm}} - \delta_\nu), \end{aligned} \quad (13)$$

whereas the fake total matter field (F) is defined as

$$\begin{aligned} 1 + \delta_F &= 1 + \frac{\Omega_{\text{cdm}}}{\Omega_m} \delta_{\text{cdm}} \\ &= (1 + \delta_{\text{cdm}}) - \frac{\Omega_\nu}{\Omega_m} \delta_{\text{cdm}}. \end{aligned} \quad (14)$$

We construct an emulator based on field F following the same procedure as in Section 4. The parameter dependence and cosmological inference on the four main parameters Ω_m , h , σ_8 and M_ν are very close to those for Field T (See Fig. 13 and Table 3). We show the cosmological constraints from field T and field F in Fig. 13. The figure demonstrates that the constraining power remains similar, indicating that the information does not derive from the distribution of neutrinos but rather from the neutrino fraction in the total matter density and the topology of the cold dark matter. In other words, these results show that the neutrino fraction rather than the neutrino field contributes to revealing the neutrino mass and breaking parameter degeneracies. The fluctuations in the neutrino density field itself provide only a minor contribution to the topological imprint of neutrino mass.

6.2 Mathematical explanation

This behaviour can be explained mathematically by considering the definition of the fields in equation (14). Under a simplified assumption that the cold dark matter density contrast, δ_{cdm} , remains approximately constant for different neutrino masses, the variation in Field F is primarily driven by the neutrino fraction, Ω_ν/Ω_m . A higher neutrino mass causes matter density fluctuations to be suppressed, driving structures closer to the mean density ($\delta = 0$); overdense regions become less dense, and underdense regions become less empty. This differential shift contributes significantly to the observed ‘tightening’ pattern in the persistence distributions (for curves expanding across the point $\vartheta_{\text{death}} = 1$, in this case, β_0 and β_1) and the shift in peak density (for curves with $\vartheta_{\text{death}} < 1$).

This nonlinear transformation of the density field also accounts for the observed suppression of high-persistence voids and enhancement of low-persistence ones. For example, a higher neutrino mass causes a topological feature crossing the $\delta = 0$ threshold to experience a decrease in $\vartheta_{\text{birth}} > 0$ and an increase in $\vartheta_{\text{death}} < 0$, pushing it toward lower persistence rather than simply rescaling its density thresholds. Thus, the enhanced sensitivity of

Table 3. Similar to Table 2, but showing results from different fields and probes, including the topological analysis on the F, T, and D fields, as well as the 1-pts analysis of the T field and the joint analysis of 1-pts with topological analysis on field D.

Probe	CMB priors	Field used	Ω_m 0.233	h 0.655	σ_8 0.607	M_ν/eV 0.236
$\beta_{k_{012}, P_{123}}$	✓	F	$0.214^{+0.042}_{-0.028}$	$0.680^{+0.047}_{-0.053}$	$0.619^{+0.045}_{-0.040}$	$0.236^{+0.062}_{-0.072}$
$\beta_{k_{012}, P_{01234}}$	✓	F	$0.216^{+0.042}_{-0.029}$	$0.677^{+0.048}_{-0.054}$	$0.618^{+0.035}_{-0.037}$	$0.237^{+0.051}_{-0.050}$
$\beta_{k_{012}, P_{123}}$	✓	D	$0.224^{+0.030}_{-0.034}$	$0.664^{+0.056}_{-0.036}$	$0.616^{+0.051}_{-0.029}$	$0.152^{+0.176}_{-0.094}$
$\beta_{k_{012}, P_{01234}}$	✓	D	$0.220^{+0.031}_{-0.029}$	$0.671^{+0.048}_{-0.039}$	$0.619^{+0.037}_{-0.031}$	$0.228^{+0.132}_{-0.137}$
$\beta_{k_{012}}$	✓	T	$0.212^{+0.029}_{-0.024}$	$0.681^{+0.041}_{-0.036}$	$0.621^{+0.046}_{-0.030}$	$0.254^{+0.077}_{-0.094}$
1 pts	✓	T	$0.244^{+0.025}_{-0.055}$	$0.640^{+0.080}_{-0.032}$	$0.580^{+0.151}_{-0.046}$	$0.215^{+0.137}_{-0.133}$
1 pts + $\beta_{k_{012}, P_{01234}}$	✓	T/D	$0.251^{+0.016}_{-0.059}$	$0.632^{+0.085}_{-0.022}$	$0.590^{+0.063}_{-0.015}$	$0.238^{+0.073}_{-0.084}$

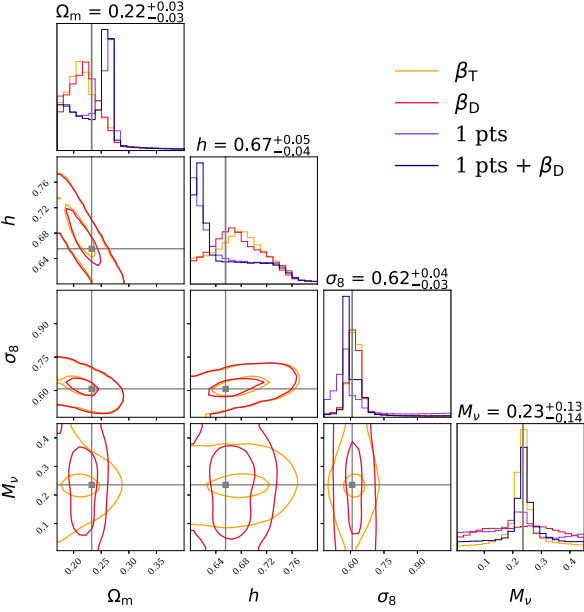


Figure 12. Similar to Fig. 8, but comparing the results from T and D fields. For clarity, we only show the confidence contours from the T and D fields’ topology (with CMB priors, in orange and red, respectively) in 2D subpanels. We also include the results from 1-pts and the joint analysis of 1-pts with field D in the 1D subpanels (purple and navy, respectively), with CMB priors. The grey lines indicate the true value of each parameter. Labels above each column indicate the median and 68 per cent credible intervals for β_D .

void topology (β_2) to neutrino mass, compared to haloes (β_0) and filaments (β_1), arises largely from the relatively higher fraction of neutrinos.

In a more realistic scenario, δ_{cdm} is also modulated by the neutrino mass, due to the suppression of the growth of CDM perturbations. Heavier neutrinos lead to a more clustred neutrino field and, in contrast, a less clustred cold dark matter field. This is the origin of the degeneracy between M_ν and σ_8 within the field D. The question remains how much information is derived from these two effects: the change in the neutrino fraction and the suppression of the D field.

6.3 Two effects with equal importance

Based on these results, we considered filtrations as a function of the volume fraction instead of the overdensity, following the rescaling method proposed by D. H. Weinberg (1988), thereby removing information from the one-point distribution

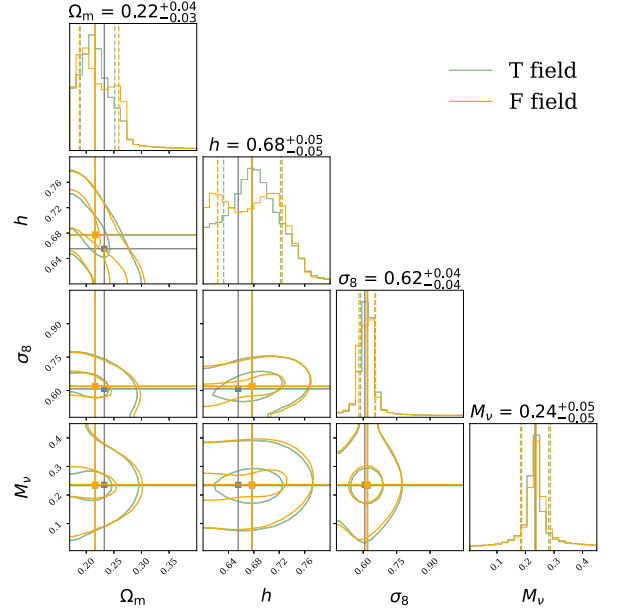


Figure 13. Similar to Fig. 8, but comparing the cosmological inference from field T (green contours) with that of field F (orange contours), all with CMB priors. The grey lines indicate the true value of each parameter.

function (1pts hereafter) of the density field, $P(\delta)$. Under this transformation, the Betti curves for the T field become nearly indistinguishable from those of the D field. This indicates that the enhanced constraining power of the total matter field originates not from a fundamentally different topology, but from the way that neutrinos modify the underlying 1pts of the T field. This is consistent with the results of J. Liu & M. S. Madhavacheril (2019), showing that 1pts are capable of constraining the neutrino mass. The case is more complex for persistence strips. Their dependence on structure lifetimes (persistence) cannot be straightforwardly rescaled by this method, suggesting that they encode information beyond the 1pts. This likely explains why persistence strips achieve approximately twice the constraining power on $\sum m_\nu$ compared to conventional summary statistics.

To isolate the constraining power arising purely from changes to $P(\delta)$, we build an emulator of the one-point probability distribution function following the same methodology and present the results in Fig. 12. Both the 1pts and D-field topology individually yield weak constraints on neutrino mass. However, only by combining these two probes can we closely match the constraining power achieved with the T field topology alone. This confirms

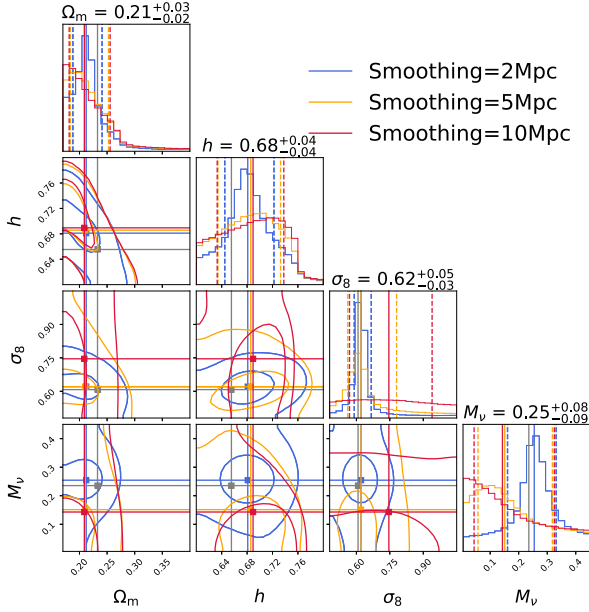


Figure 14. Similar to Fig. 8, but comparing the cosmological inference from Betti curves obtained from fields with different smoothing scales, 2 Mpc (blue), 5 Mpc (orange), and 10 Mpc (red), all with CMB priors. The grey lines indicate the true value of each parameter.

that persistent homology captures information beyond simple field rescaling, and these two effects are equally important.

7 OBSERVATIONAL CONSIDERATIONS

To assess the robustness of our findings under more realistic observational conditions, we incorporate the effect of redshift-space distortions (RSD) and test against different smoothing scales.

7.1 Smoothing scale effect

M. H. Jalali Kanafi et al. (2024) demonstrated that increasing the smoothing scale degrades constraints on neutrino mass, reporting a 68 per cent confidence interval for M_ν that widens from 0.0152 eV with a $5h^{-1}$ Mpc kernel to 0.0720 eV with a $10h^{-1}$ Mpc kernel. In this work, the δf method combined with the higher simulation resolution enables us to probe significantly smaller scales, down to 2 Mpc. To systematically investigate the smoothing scale dependence of our constraints, we construct two additional emulators for comparison with our 2 Mpc baseline, employing smoothing scales of 5 and 10 Mpc while keeping all other procedures identical. We note that the meaningful persistence range changes substantially with the smoothing scale, making it difficult to quantify this difference and introducing challenges for a fair comparison of the persistence strips across different configurations. Although we present the complete results for both Betti curves and persistence strips in Table B1, the following comparative analysis focuses primarily on the Betti curves to enable a more straightforward interpretation.

The resulting constraints from all three smoothing scales are compared in Fig. 14. We find that different smoothing scales generally yield consistent cosmological predictions for Ω_m , h , and σ_8 , but with degraded precision. Specifically, the constraining power on σ_8 almost disappears when increasing to 10 Mpc smoothing. The constraints on M_ν are largely biased to smaller values and

lose constraining power with larger smoothing scales, indicating that a large smoothing may erase the neutrino imprints. We have verified that the Bayesian inference is well-converged and consistent across several testing samples, ruling out sampling issues as the cause.

This degradation in cosmological constraining power with a larger smoothing scale can be attributed to its effect on cosmic voids. As the smoothing scale increases, it erodes the topological signature of voids. Given that our primary constraining power originates from high-persistence topological features, this erosion of void signatures provides a natural explanation for the significantly weaker cosmological constraints we obtain.

7.2 Redshift space distortion effect

To assess the robustness of our method under more realistic observational conditions, we incorporate redshift-space distortions (RSD) into the dark matter particles in the simulation and repeat the cosmological inference procedure. The comoving position of each dark matter particle in redshift space, \mathbf{s} , is calculated from its real-space position, \mathbf{x} , and peculiar velocity, \mathbf{v} , using the standard transformation

$$\mathbf{s} = \mathbf{x} + \frac{\mathbf{v} \cdot \hat{\mathbf{n}}}{aH(a)} \hat{\mathbf{n}}, \quad (15)$$

where $\hat{\mathbf{n}}$ is the unit vector along the line-of-sight direction, a is the scale factor, and $H(a)$ is the Hubble parameter at that epoch. We implement RSD using the plane-parallel approximation, which is well-justified for our simulation geometry with the observer effectively at infinity. The line-of-sight direction $\hat{\mathbf{n}}$ is aligned with one principal axis (z -axis) of the simulation box. For each dark matter particle, we use the line-of-sight component of its peculiar velocity and apply the corresponding displacement, scaled by the Hubble expansion factor $aH(a)$, accounting for periodic boundary conditions. The resulting redshift-space dark matter density field is then constructed from the displaced particle positions using the same CIC mass assignment scheme as applied to the real-space density field.

The resulting constraints, presented in Fig. 15, demonstrate that our topological approach is robust to RSD effects: the constraints on the four most prominent parameters generally remain unchanged. The introduction of RSD shifts the constraints along existing degeneracy directions towards lower Ω_m and higher h , while offering essentially the same constraint on these parameters. Crucially, the constraint on the neutrino mass M_ν remains robust, and the Betti curves continue to effectively break the degeneracy between M_ν and other cosmological parameters. These results confirm that topological statistics remain a powerful tool for constraining neutrino mass even in redshift space, providing strong support for the application of this method to real observations.

7.3 Other considerations

When addressing realistic observational effects, it is essential to recognize that galaxy surveys probe the discrete distribution of galaxies, not the continuous dark matter field. Persistent homology of discrete tracers is based on simplicial complexes (e.g. R. Bermejo et al. 2024) and may introduce several interrelated effects that are critical for interpreting galaxy survey data.

We should note a series of tracer-dependent effects. At the most basic level, we observe that more massive haloes are rarer, so

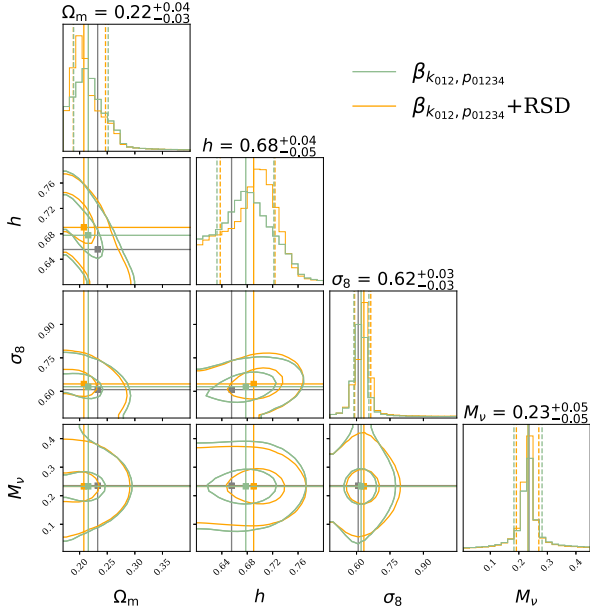


Figure 15. Similar to Fig. 8, but comparing the cosmological constraints in real and redshift space for the joint analysis with persistence strips combined with CMB priors. Green and orange contours show the 68 and 95 per cent confidence regions for real space and redshift space analyses, respectively. The true parameter values are indicated by grey lines.

only sufficiently large topological features (e.g. giant voids) are well delineated, while smaller-scale filaments and walls become poorly resolved. Second is the clustering bias effect, whereby massive haloes cluster more strongly and preferentially occupy high-density regions, leading to voids that are systematically larger and more regular than those in the full matter distribution. Most subtly, R. Bermejo et al. (2024) identify a distinct topological bias effect, wherein even after controlling for mass function and clustering, different halo populations trace systematically different topological features. In particular, more massive haloes consistently probe larger-scale structures.

Another viable technical approach involves conducting topological analyses directly on weak lensing convergence maps. However, this method faces a primary challenge of information loss, since the convergence map essentially projects the line-of-sight structure on to a 2D plane, which discards spatial information along the redshift dimension (although the limited 3D information provided by tomography offers a partial resolution). Nevertheless, works like J. Prat et al. (2025) have demonstrated the feasibility and potential of this approach. Its application to constraining neutrino mass remains an open question, which we will explore further in future work.

8 SUMMARY

This work establishes the cosmological constraining power of TDA, specifically through persistent homology, as a superior alternative to traditional two-point statistics for probing neutrino mass and breaking parameter degeneracies. Utilising the FLAMINGO simulation suite, we developed a robust emulator spanning a 10D $w_0 w_a \text{CDM} + \nu$ cosmological parameter space. Trained on only 50 simulations, it accurately predicts topological summaries (Betti curves and persistence diagrams) and enables Bayesian inference.

Our key findings show that Betti curves subdivided by persistence (the ‘persistence strips’) exhibit enhanced sensitivity to the sum of neutrino mass, M_ν . Persistence strips strike an optimal balance between information content and emulation feasibility, delivering roughly twice the constraining power of unbinned Betti curves. The strongest constraints come from 2D topological structures (voids), which exhibit a characteristic response to neutrino diffusion. Using the emulator trained on only 50 simulations, each with a volume of $(350 \text{ Mpc})^3$, we obtain a 68 per cent uncertainty of 0.05 eV, demonstrating the superior sensitivity of persistent homology.

The total matter field yields substantially tighter constraints (0.05 eV) than the dark matter-only field (0.13 eV), underscoring the importance of probes sensitive to the total matter distribution (e.g. weak lensing). We performed a detailed investigation of the source of this enhancement. By analysing a ‘fake’ total matter field, in which the neutrino density perturbations have been replaced with their mean value, we established that variations in the global neutrino mass fraction, rather than the direct imprint of the neutrino perturbations, dominate the signal. Furthermore, by constructing emulators of multiple probes, we demonstrated that the sensitivity of the total matter topology can be decomposed into two effects of comparable importance: the indirect effect of neutrino mass on the growth of cold dark matter perturbations and the imprint of the neutrino mass fraction on the 1-point distribution of the total matter field.

The method proves robust to redshift-space distortions, converges with 50 simulations, and effectively breaks degeneracies among M_ν and σ_8 , paving the way for its application to upcoming observational data from DESI and other wide-field surveys. Looking forward, we leave the inspection of hydrodynamical effects and the application of our formalism to realistic galaxy mocks and weak lensing maps, incorporating the characteristics of surveys like Euclid and DESI, for future work.

ACKNOWLEDGEMENTS

The authors acknowledge useful discussions with Carlton Baugh and Hailiang Du on the construction of the GP emulator. The authors also acknowledge support from the China Scholarship Council, the Science and Technology Facilities Council (grant numbers ST/X001075/1 and ST/Y002733/1), the National Key R&D Program of China (2023YFA1607800, 2023YFA1607804), and ‘the Fundamental Research Funds for the Central Universities’, 111 project No. B20019. RvdW acknowledges funding from EU Horizon Europe (EXCOSM, grant number 101159513). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. Part of the computations in this paper were also run on the Gravity Supercomputer at the Department of Astronomy, Shanghai Jiao Tong University.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- Abazajian K., Dodelson S., 2003, *Phys. Rev. Lett.*, 91, 041301
- Abazajian K. N. et al., 2011, *Astropart. Phys.*, 35, 177
- Abazajian K. N. et al., 2015, *Astropart. Phys.*, 63, 66
- Abdul Karim M. et al., 2025, *Phys. Rev. D*, 112, 083515
- Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O’Neil M., 2015, *IEEE Trans. Pattern Anal. Mach. Intell.*, 38, 252
- Avsajanishvili O., 2025, *Astronomische Nachrichten*, 347, id.e70084
- Bayer A. E., Banerjee A., Seljak U., 2022, *Phys. Rev. D*, 105, 123510
- Bayer A. E., Liu J., Kreisch C. D., Pisani A., 2024, *Phys. Rev. D*, 110, L061305
- Bermejo R., Wilding G., van de Weygaert R., Jones B. J. T., Vegter G., Efstathiou K., 2024, *MNRAS*, 529, 4325
- Biagetti M., Cole A., Shiu G., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 061
- Bobrowski O., Skraba P., 2020, *Phys. Rev. E*, 101, 032304
- Bond J. R., Efstathiou G., Silk J., 1980, *Phys. Rev. Lett.*, 45, 1980
- Bond J. R., Kofman L., Pogosyan D., 1996, *Nature*, 380, 603
- Brandbyge J., Hannestad S., Haugbølle T., Thomsen B., 2008, *J. Cosmol. Astropart. Phys.*, 2008, 020
- Calabrese E. et al., 2025, *J. Cosmol. Astropart. Phys.*, 2025, 063
- Calles J., Yip J. H. T., Contardo G., Noreña J., Rouhiainen A., Shiu G., 2025, *JCAP*, 2025, id.064
- Camphuis E. et al., 2026, *Phys. Rev. D*, 113, 083504
- Carlsson G., 2009, *Bull. Am. Math. Soc.*, 46, 255
- Carlsson G., Vejdemo-Johansson M., 2021, *Topological Data Analysis with Applications*. Cambridge University Press, <https://books.google.nl/books?id=NxFvzg'EACAAJ>
- Chen Z., Yu Y., 2025, *Sci. China Phys. Mech. Astron.*, 68, 109513
- Chen Z., Yu Y., Han J., Jing Y., 2025, *Sci. China Phys. Mech. Astron.*, 68, 289512
- Chevallier M., Polarski D., 2001, *Int. J. Mod. Phys. D*, 10, 213
- Chiang C.-T., Hu W., Li Y., LoVerde M., 2018, *Phys. Rev. D*, 97, 123526
- Cooray A. R., 1999, *A&A*, 348, 31
- Craig N., Green D., Meyers J., Rajendran S., 2024, *J. High Energy Phys.*, 2024, 97
- Croft R. A. C., Hu W., Davé R., 1999, *Phys. Rev. Lett.*, 83, 1092
- DESI Collaboration, 2025, *Phys. Rev. D*, 112, 083514
- DESI Collaboration, 2025, *J. Cosmol. Astropart. Phys.*, 2025, 021
- Dlotko P., 2025, in *GUDHI User and Reference Manual*, 3.11.0 edn. GUDHI Editorial Board, https://gudhi.inria.fr/doc/3.11.0/group_ubicual_complex.html
- Edelsbrunner H., Harer J., 2010, *Computational Topology—An Introduction*. American Mathematical Society, Providence, RI
- Elbers W., 2022, *J. Cosmol. Astropart. Phys.*, 2022, 058
- Elbers W., van de Weygaert R., 2019, *MNRAS*, 486, 1523
- Elbers W., van de Weygaert R., 2023, *MNRAS*, 520, 2709
- Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2021, *MNRAS*, 507, 2614
- Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2022, *MNRAS*, 516, 3821
- Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2025a, *Phys. Rev. D*, 111, 063534
- Elbers W. et al., 2025b, *Phys. Rev. D*, 112, 083513
- Elbers W. et al., 2025c, *MNRAS*, 537, 2160
- Elgarøy Ø. et al., 2002, *Phys. Rev. Lett.*, 89, 061301
- Esteban I., Gonzalez-Garcia M. C., Maltoni M., Martinez-Soler I., Pinheiro J. P., Schwetz T., 2024, *J. High Energy Phys.*, 2024, 216
- Euclid Collaboration, 2019, *MNRAS*, 484, 5509
- Fang F., Forero-Romero J., Rossi G., Li X.-D., Feng L.-L., 2019, *MNRAS*, 485, 5276
- Feldbrugge J., van de Weygaert R., 2025, *MNRAS*, 539, 873
- Gangopadhyay K., Banerjee A., Abel T., 2025, *MNRAS*, 543, 3409
- Giri S. K., Mellema G., 2021, *MNRAS*, 505, 1863
- Green D., Meyers J., 2025, *Phys. Rev. D*, 111, 083507
- Hahn C., Villaescusa-Navarro F., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 029
- Hahn O., Rampf C., Uhlemann C., 2021, *MNRAS*, 503, 426
- Hatcher A., 2002, *Algebraic Topology*. Cambridge Univ. Press, Cambridge
- Hertzsch B., Feldbrugge J., Rodriguez M., van de Weygaert R., 2026, *J. Cosmol. Astropart. Phys.*, 2026, 037
- Heydenreich S., Brück B., Harnois-Déraps J., 2021, *A&A*, 648, A74
- Hu W., Eisenstein D. J., Tegmark M., 1998, *Phys. Rev. Lett.*, 80, 5255
- Jalali Kanafi M. H., Ansarifard S., Movahed S. M. S., 2024, *MNRAS*, 535, 657
- Jiang J.-Q. et al., 2025, *J. Cosmol. Astropart. Phys.*, 2025, 153
- Kreisch C. D., Pisani A., Carbone C., Liu J., Hawken A. J., Massara E., Spergel D. N., Wandelt B. D., 2019, *MNRAS*, 488, 4413
- Kugel R. et al., 2023, *MNRAS*, 526, 6103
- Lange J. U., 2023, *MNRAS*, 525, 3181
- Laux S., 1996, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 15, 1266
- Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, *Astrophysics Source Code Library*, record ascl:1010.030
- Lemos P., Lewis A., 2023, *Phys. Rev. D*, 107, 103505
- Linder E. V., 2003, *Phys. Rev. Lett.*, 90, 091301
- Liu J., Madhavacheril M. S., 2019, *Phys. Rev. D*, 99, 083508
- Liu Y., Yu Y., Yu H.-R., Zhang P., 2020, *Phys. Rev. D*, 101, 063515
- Liu W., Jiang A., Fang W., 2022, *J. Cosmol. Astropart. Phys.*, 2022, 045
- Liu Y., Yu Y., Zhang P., Yu H.-R., 2024, *ApJS*, 273, 33
- Massara E., Villaescusa-Navarro F., Viel M., Sutter P. M., 2015, *J. Cosmol. Astropart. Phys.*, 2015, 018
- Milnor J., 1963, *Morse Theory*. Based on Lecture Notes by M. Spivak and R. Wells. *Annals of Mathematics Studies*, No. 51, Princeton University Press, Princeton, N.J.
- Moon J., Rossi G., Yu H., 2023, *ApJS*, 264, 26
- Munkres J. R., 1984, *Elements of Algebraic Topology*. Addison-Wesley Menlo Park
- Naidoo K., Massara E., Lahav O., 2022, *MNRAS*, 513, 3596
- Naredo-Tuero D., Escudero M., Enrique Fernandez-MartinezMarcano X., Poulin V., Enrique Fernandez-Martinez, 2024, *Phys. Rev. D*, 110, 123537
- Noriega H. E., Aviles A., 2025, *Phys. Rev. D*, 111, L061307
- Ouellette A., Holder G., 2025, preprint ([arXiv:2502.09709](https://arxiv.org/abs/2502.09709))
- Planck Collaboration VI, 2020, *A&A*, 641, A6
- Platen E., van de Weygaert R., Jones B. J. T., 2008, *MNRAS*, 387, 128
- Pranav P., Edelsbrunner H., van de Weygaert R., Vegter G., Kerber M., Jones B. J. T., Wintraecken M., 2017, *MNRAS*, 465, 4281
- Prat J. et al., 2025, *MNRAS*
- Project T. G., 2025, *GUDHI User and Reference Manual*, 3.11.0 edn. GUDHI Editorial Board, <https://gudhi.inria.fr/doc/3.11.0/>
- Rasmussen C. E., Williams C. K. I., 2006, *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA
- Rossi G., Yu H., Michaux M., 2026, preprint ([arXiv:2604.09148](https://arxiv.org/abs/2604.09148))
- Sahni V., Sathyaprakah B. S., Shandarin S. F., 1994, *ApJ*, 431, 20
- Schaller M., Gonnet P., Chalk A. B., Draper P. W., 2018, *Astrophysics Source Code Library*, ascl:1805.020
- Schaller M. et al., 2024, *MNRAS*, 530, 2378
- Schaye J. et al., 2023, *MNRAS*, 526, 4978
- Seljak U. et al., 2005, *Phys. Rev. D*, 71, 103515
- Shandarin S. F., Zeldovich Y. B., 1989, *Rev. Mod. Phys.*, 61, 185
- Shen J., Abel T., Mo H. J., Sheth R. K., 2006, *ApJ*, 645, 783
- Sheth R. K., van de Weygaert R., 2004, *MNRAS*, 350, 517
- Simões L. N. L., Naidoo K., Joachimi B., Elbers W., Frenk C. S., 2025, preprint ([arXiv:2512.16517](https://arxiv.org/abs/2512.16517))
- Sunseri J., Bayer A. E., Liu J., 2025, *Phys. Rev. D*, 112, 063516
- Tegmark M. et al., 2004, *Phys. Rev. D*, 69, 103501
- Thiele L., Massara E., Pisani A., Hahn C., Spergel D. N., Ho S., Wandelt B., 2024, *ApJ*, 969, 89
- Tsizh M., Tymchyshyn V., Vazza F., 2023, *MNRAS*, 522, 2697
- van de Weygaert R. et al., 2011, in *Gavrilova M. L., Tan C. J. K., Mostafavi M. A. eds, Lecture Notes in Computer Science*. Vol. 6970. Springer, Berlin, Heidelberg, p. 60
- Villaescusa-Navarro F., 2018, *Astrophysics Source Code Library*, record ascl:1811.008

- Villaescusa-Navarro F. et al., 2020, *ApJS*, 250, 2
 Wasserman L., 2018, *Ann. Rev. Stat. Appl.*, 5, 501
 Weinberg D. H., 1988, *PASP*, 100, 1373
 Wilding G., Nevenzeel K., van de Weygaert R., Vegter G., Pranav P., Jones B. J. T., Efstathiou K., Feldbrugge J., 2021, *MNRAS*, 507, 2968
 Yip J. H. T., Biagetti M., Cole A., Viswanathan K., Shiu G., 2024, *J. Cosmol. Astropart. Phys.*, 2024, 034
 Zel'dovich Y. B., 1970, *A&A*, 5, 84
 Zhang G., Li Z., Liu J., Spergel D. N., Kreisch C. D., Pisani A., Wandelt B. D., 2020, *Phys. Rev. D*, 102, 083537

APPENDIX A: THE EMULATOR CONVERGENCE TEST

We assess the convergence of our emulator by comparing constraints derived from ensembles of 50 and 100 simulations. As shown in Fig. A1, the resulting cosmological constraints from the two ensembles are nearly identical. This indicates that the emulator built on 50 simulations has sufficiently converged, with a negligible gain in precision from doubling the size of the training set.

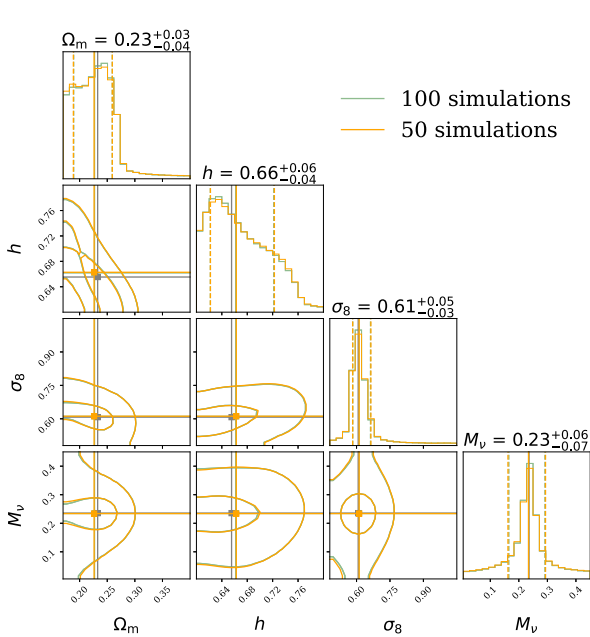


Figure A1. Similar to Fig. 8, but comparing the results for different training sets. The constraints are derived from ensembles of 50 and 100 simulations, all with CMB priors. The grey lines indicate the true value of each parameter.

APPENDIX B: EMULATOR WITH DIFFERENT COSMOLOGICAL PARAMETERS

An intriguing finding emerges from the comparison of fluctuation amplitude parameters. While the amplitude of the primordial power spectrum, A_s , and the late-time fluctuation amplitude, σ_8 , are traditionally degenerate and equally effective indicators for the power spectrum, this equivalence does not hold for topological statistics. We find that Betti curves exhibit a strong, well-defined response to σ_8 , but only a weak and poorly constrained response to A_s (see Fig. B1 and Table B1).

Consequently, this work utilizes σ_8 as the primary parameter for characterizing matter fluctuations. Since our simulation hypercube was designed with a flat prior on $\ln A_s$, the derived values of σ_8 follow a distribution that is approximately Gaussian. To mitigate potential biases from sparsely sampled regions of the parameter space, we restrict our testing samples to the range $\sigma_8 \in [0.6, 0.9]$.

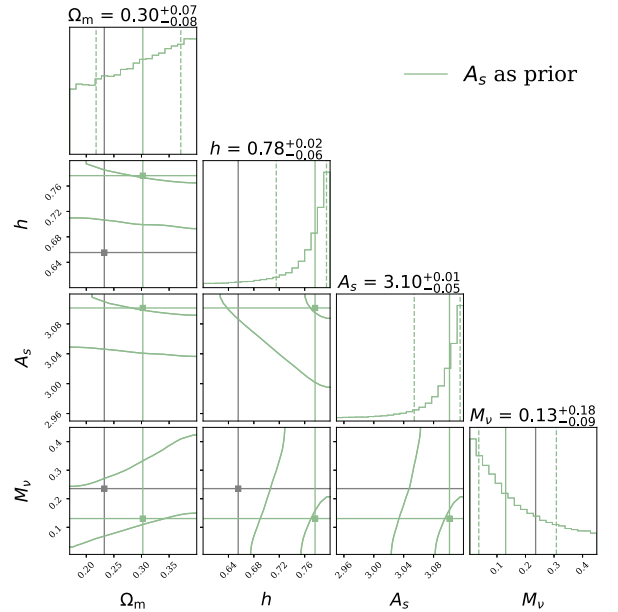


Figure B1. Similar to Fig. 8, but with A_s as prior. The grey lines indicate the true value of each parameter.

Table B1. Similar to Table 2, but with a different smoothing scale or density fluctuation parameter.

Probe	Smoothing	CMB scale (Mpc)	Field priors	Parameters used	Ω_m	h	σ_8	M_V
					0.233	0.655	0.607	0.236
$\beta_{k_{012}, P_{123}}$	5	✓	T	σ_8	$0.251^{+0.014}_{-0.041}$	$0.630^{+0.061}_{-0.016}$	$0.588^{+0.009}_{-0.005}$	$0.248^{+0.116}_{-0.136}$
$\beta_{k_{012}, P_{123}}$	10	✓	T	σ_8	$0.212^{+0.015}_{-0.030}$	$0.685^{+0.052}_{-0.022}$	$0.639^{+0.033}_{-0.024}$	$0.397^{+0.028}_{-0.272}$
$\beta_{k_{012}}$	2	✓	T	σ_8	$0.212^{+0.029}_{-0.024}$	$0.681^{+0.041}_{-0.036}$	$0.621^{+0.046}_{-0.030}$	$0.254^{+0.077}_{-0.094}$
$\beta_{k_{012}}$	5	✓	T	σ_8	$0.210^{+0.042}_{-0.027}$	$0.686^{+0.047}_{-0.052}$	$0.621^{+0.159}_{-0.060}$	$0.151^{+0.167}_{-0.094}$
$\beta_{k_{012}}$	10	✓	T	σ_8	$0.209^{+0.047}_{-0.027}$	$0.689^{+0.048}_{-0.057}$	$0.746^{+0.194}_{-0.177}$	$0.143^{+0.182}_{-0.099}$
$\beta_{k_{012}, P_{01234}}$	2		T	σ_8	$0.227^{+0.046}_{-0.037}$	$0.682^{+0.080}_{-0.068}$	$0.616^{+0.042}_{-0.037}$	$0.252^{+0.072}_{-0.063}$
$\beta_{k_{012}, P_{01234}}$	2		T	A_s	$0.302^{+0.069}_{-0.085}$	$0.776^{+0.018}_{-0.061}$	$3.101^{+0.014}_{-0.047}$	$0.131^{+0.178}_{-0.094}$

APPENDIX C: BETTI CURVE DEPENDENCE ON OTHER COSMOLOGICAL PARAMETERS

The cosmological dependence of six less prominent parameters: the baryon fraction f_b , the scalar spectral index n_s , its running α_s , the decaying dark matter rate Γ_{dcdm} , and the dark energy equation of state parameters (w_0 , w_a) are shown in Fig. C1. A measurable dependence is observed primarily for α_s and n_s (most

notably in high-persistence regimes), followed by a weaker dependence on w_0 and w_a . This indicates that the topology of the cosmic web is subtly influenced by both the initial conditions of the power spectrum and the dynamics of dark energy. Conversely, the topological statistics exhibit minimal sensitivity to the baryon fraction f_b and cannot meaningfully constrain the decaying dark matter parameter Γ_{dcdm} .

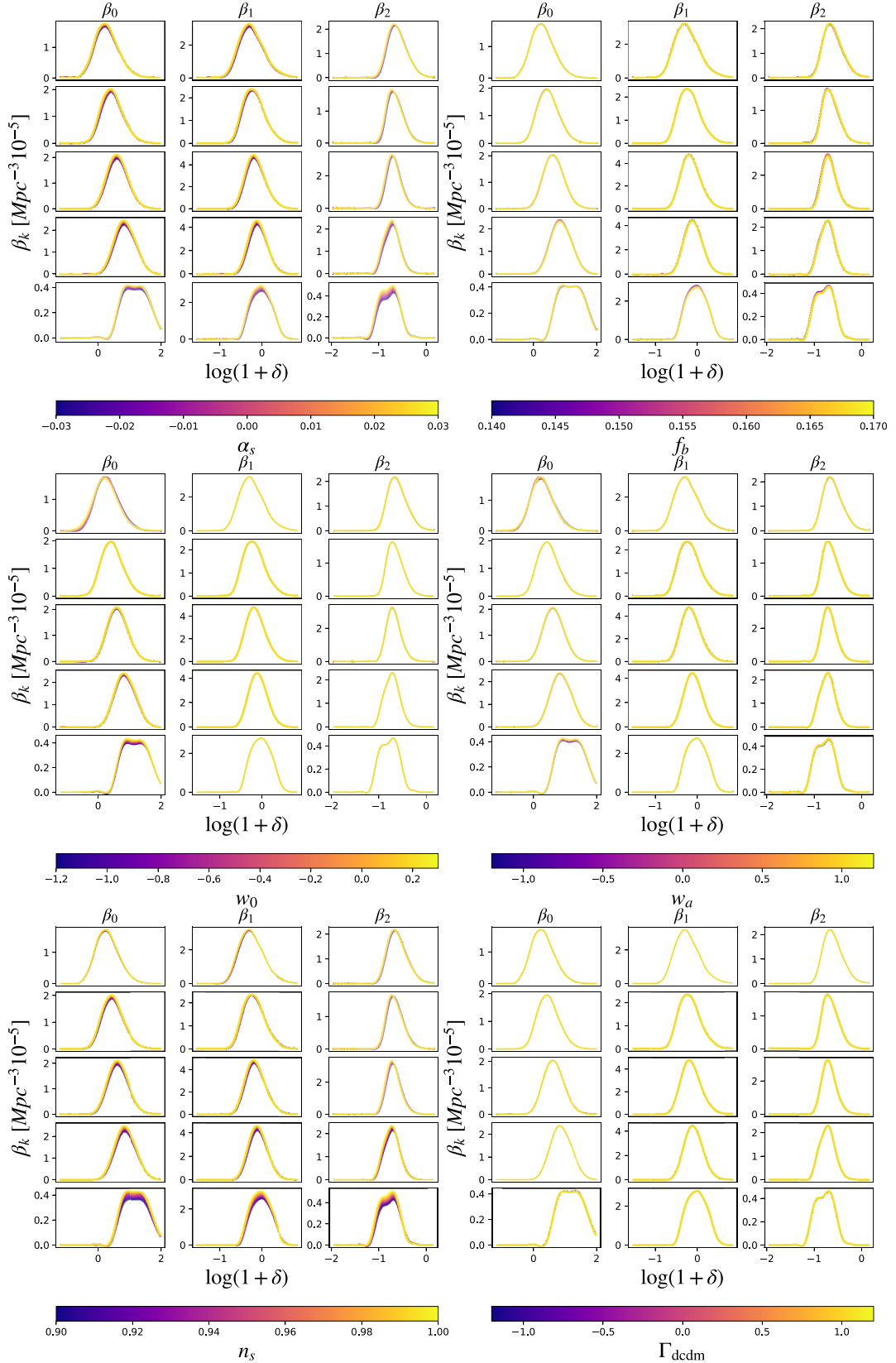


Figure C1. Similar to Fig. 7, but showing the dependence on 6 less-prominent parameters.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

© The Author(s) 2026.

Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.