The dwarf stellar mass function in different environments and the lack of a generic missing dwarfs problem in ΛCDM

I. Lazar [,1]★ S. Kaviraj [,1] G. Martin [,2] A. Watkins,[1] D. Kakkad [,1] B. Bichang'a [,1] K. Kraljic [,3]
S. K. Yi,[4] Y. Dubois,[5] J. E. G. Devriendt,[6] S. Peirani[7,8] and C. Pichon[5,9]

[1]*Centre for Astrophysics Research, Department of Physics, Astronomy and Mathematics, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK*
[2]*School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[3]*Observatoire Astronomique de Strasbourg, Université de Strasbourg, CNRS, UMR 7550, F-67000 Strasbourg, France*
[4]*Department of Astronomy and Yonsei University Observatory, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea*
[5]*Institut d'Astrophysique de Paris, Sorbonne Universités, UMPC Univ Paris 06 et CNRS, UMP 7095, 98 bis bd Arago, F-75014 Paris, France*
[6]*Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[7]*ILANCE, CNRS – University of Tokyo International Research Laboratory, Kashiwa, Chiba 277-8582, Japan*
[8]*Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[9]*Dept. of Astronomy & Space Science, Kyung Hee University, Yongin-shi, Gyeonggi-do 17104, Republic of Korea*

ABSTRACT

We combine deep photometric data in the COSMOS and XMM-LSS fields with high-resolution cosmological hydrodynamical simulations to explore two key questions: (1) how does the galaxy stellar mass function, particularly in the dwarf ($M_\star < 10^{9.5}\,\mathrm{M_\odot}$) regime, vary with environment, defined as the distance from large-scale structure (LSS) traced by nodes and filaments in the cosmic web? (2) is there a generic 'missing dwarfs' problem in Lambda cold dark matter (ΛCDM) predictions when all environments – and not just satellites around Milky Way like galaxies – are considered? The depth of the observational data used here enables us to construct complete, unbiased samples of galaxies, down to $M_\star \sim 10^7\,\mathrm{M_\odot}$ and out to $z \sim 0.4$. Strong environmental differences are found for the galaxy stellar mass function when considering distance from LSS. As we move closer to LSS, the dwarf mass function becomes progressively flatter and the knee of the mass function shifts to larger stellar masses, both of which result in a higher ratio of massive to dwarf galaxies. While the stellar mass functions from the three simulations (NEWHORIZON, TNG50, and FIREBOX) considered here do not completely agree across the dwarf regime, there is no evidence of a generic missing dwarfs problem in the context of ΛCDM, akin to the results of recent work that demonstrates that there is no missing satellites problem around Galactic analogues.

Key words: galaxies: dwarf – galaxies: evolution – galaxies: formation – galaxies: luminosity function, mass function.

1 INTRODUCTION

In the hierarchical Lambda cold dark matter (ΛCDM) paradigm, dwarf ($M_\star < 10^{9.5}\,\mathrm{M_\odot}$) galaxies represent an important population of astrophysical objects. Their numerical dominance enables this population to bring about macroscopic evolutionary changes, such as making a large contribution to the re-ionization of the Universe (e.g. H. Atek et al. 2015; S. Salvadori, Á. Skúladóttir & E. Tolstoy 2015). The shallow gravitational potential wells of these systems makes them sensitive laboratories for studying both internal processes, such as baryonic feedback (S. Dong, D. N. C. Lin & S. D. Murray 2003; G. Martin et al. 2025), and external, environmentally-driven processes like tidal perturbations (e.g. A. Koch et al. 2012; H. Martel, P. Barai & W. Brito 2012; A. Fattahi et al. 2018; R. A. Jackson et al. 2021a; A. E. Watkins et al. 2023) and ram pressure stripping (e.g. M. Mori & A. Burkert 2000; L. Mayer et al. 2006; P. Steyrleithner, G. Hensler & A.

Boselli 2020; A. Boselli, M. Fossati & M. Sun 2022). Their high ratio of dark to baryonic mass also makes dwarf galaxies good testbeds for the study of dark matter (e.g. R. A. Jackson et al. 2021b; G. Battaglia & C. Nipoti 2022).

Despite their significance, the dwarf population remains much more poorly understood than massive galaxies. In the nearby Universe, dwarfs have been studied in detail largely in our local neighbourhood (and in relatively high-density environments), such as in the Local Group (e.g. E. Tolstoy, V. Hill & M. Tosi 2009), around nearby massive galaxies (e.g. M. Poulain et al. 2021; I. Trujillo et al. 2021; M. Geha et al. 2024) or in nearby galaxy clusters (e.g. J. A. L. Aguerri et al. 2005; A. Boselli et al. 2008; N. Choque-Challapa et al. 2021). Such studies have either employed deep targeted observations or have observed regions which are close enough that relatively complete samples of dwarfs can be assembled using the data available from existing surveys.

It is worth noting, however, that typical dwarfs are not bright enough to be detectable outside the local Universe in past wide

★ E-mail: i.lazar@herts.ac.uk

area surveys like the Sloan Digital Sky Survey (SDSS), which offer large footprints (that facilitate statistical studies at low redshift) but are relatively shallow. Dwarfs that do appear in such data sets and reside outside the local neighbourhood have high star formation rates (SFRs), which are needed to boost their luminosities above the detection limits of the shallow images. Thus, while large samples of dwarfs do exist in surveys like the SDSS, they represent a biased, blue star-forming subset of the dwarf population (e.g. S. Kaviraj et al. 2025). This bias may lead to erroneous conclusions, such as underestimated red/quenched fractions, fractions of active galactic nuclei (AGN) that are too low (because the strong star formation may swamp the signatures of an AGN) and a morphological mix that is likely skewed towards late-type galaxies (because late-types typically host stronger star formation).

Assembling unbiased, statistical samples of dwarfs in low-density environments has traditionally been challenging because it requires surveys that are both deep and wide. While this will become routinely possible in the new era of surveys like LSST (e.g. Ž. Ivezić et al. 2019) and *Euclid* (R. Laureijs et al. 2011; J. C. Cuillandre et al. 2025), some deep multiwavelength data sets do now exist which enable such studies, albeit over relatively small areas of the sky. The ability to access large, complete samples of dwarf galaxies outside the local Universe offers the opportunity for unbiased explorations of many properties (e.g. star formation, quenching, morphology, and the presence of AGN) that we were, until recently, limited to exploring largely in massive galaxies only. For example, in a series of recent studies, we have used the deep photometric data available in the 2 deg^2 COSMOS field (N. Scoville et al. 2007) to explore the properties of the dwarf population outside the very local Universe (B. Bichang'a et al. 2024; I. Lazar et al. 2024a, b; S. Kaviraj et al. 2025). Notwithstanding the relatively modest area of a field like COSMOS, the shape of the galaxy mass function[1] means that complete, statistical samples of dwarfs can be constructed in these parts of the sky out to moderate redshift ($z \sim 0.4$).

These papers have demonstrated interesting differences between the properties of dwarfs and massive galaxies, which often diverge from the results of past dwarf studies underpinned by shallow surveys, due to the biases described above. For example, while the well-known 'early-type' and 'late-type' morphological classes that dominate the Hubble sequence (E. P. Hubble 1936) in massive galaxies also exist in dwarfs, a progressively larger fraction of galaxies appears to be 'featureless' towards lower stellar masses in the dwarf regime (I. Lazar et al. 2024a). This featureless class is akin to galaxies which are described as 'dwarf spheroidals' near the Milky Way and in nearby clusters (e.g. H. Shapley 1938; G. Reaves 1956; S. den Bergh 1959; J. Kormendy 1985). In contrast to their massive counterparts, dwarf early-types exhibit a lower incidence of interactions, are less concentrated, show positive colour gradients and share similar rest-frame colours as dwarf late-types (I. Lazar et al. 2024b). This suggests that their formation histories are likely to be shaped by secular processes (rather than by interactions) and that they grow 'outside in' rather than 'inside out', in contrast with their massive counterparts.

Interactions in dwarfs appear to enhance star formation across the entire body of the galaxy (I. Lazar et al. 2024b), rather than in the galaxy centres, as is often the case in massive galaxies. Finally, contrary to what has been suggested by dwarf studies that are underpinned by shallow surveys like the SDSS, the AGN fractions in

dwarfs appear similar (or perhaps even higher) than those in massive galaxies (e.g. B. Bichang'a et al. 2024, see also M. Mezcua & H. Domínguez Sánchez 2024). This suggests that the role of AGN in dwarfs – which remains poorly understood – may be important in the low-mass regime (e.g. J. Silk 2017; G. Dashyan et al. 2018; S. Koudmani, N. A. Henden and D. Sijacki 2021; F. Davis et al. 2022; S. Koudmani, D. Sijacki & M. C. Smith 2022). Together with the recent literature, these studies demonstrate that significant advances in our understanding of galaxy evolution can be made by studying complete, unbiased samples of galaxies in the dwarf regime.

In this paper, we use deep, multiwavelength data to probe the galaxy stellar mass function in the nearby Universe down to $M_\star \sim 10^7 \, M_\odot$ and out to $z \sim 0.4$, which represents the parameter space within which mass-complete samples of galaxies can be constructed, given the depth of the available data (see S. Kaviraj et al. 2025 and the discussion in Section 3). We first study how the galaxy stellar mass function varies with environment, defined as the distance from large-scale structure (LSS) traced by the nodes and filaments in the cosmic web. We then compare our observed mass functions to those predicted by high-resolution cosmological hydrodynamical simulations (NEWHORIZON, TNG50, and FIREBOX), to explore whether there is a 'missing dwarfs' problem in ΛCDM predictions when galaxies across all environments are considered. This is a generalized version of the 'missing satellites' problem, which is a longstanding potential tension between the predicted and observed number densities of low mass galaxies in the standard paradigm (e.g. A. Klypin et al. 1999; B. Moore et al. 1999; M. Boylan-Kolchin, J. S. Bullock & M. Kaplinghat 2011; J. S. Bullock & M. Boylan-Kolchin 2017). The issue of missing dwarfs has so far been probed in satellite populations around the Milky Way or other local massive galaxies. However, if such a problem does exist, it will not be restricted just to the Milky Way neighbourhood but rather manifest itself as a generic missing dwarfs problem across all environments. The deep data now available enables just such an extended exploration of missing dwarfs outside the local neighbourhood, albeit restricted to larger stellar masses than has been probed in the context of the missing satellites problem.

The plan for this paper is as follows. In Section 2, we describe the data sets used in this study, the algorithm used for calculating environmental parameters and the cosmological hydrodynamical simulations that we compare to the mass functions derived from the observational data. In Section 3, we calculate the lower stellar mass and upper redshift limits out to which we can construct complete galaxy samples, given the depth of the observational data. In Section 4, we describe how the uncertainties on the mass functions are calculated. In Section 5, we study how the stellar mass function, particularly in the dwarf regime, varies with distance from LSS traced by the nodes and filaments that define the cosmic web. In Section 6, we probe whether there is a generic missing dwarfs problem as described above. We summarize our findings in Section 7. Throughout this work, we assume a Hubble constant (H_0) of 70.5 km s^{-1} Mpc^{-1}.

2 DATA

In this study, we use deep multiwavelength observational survey data in the COSMOS and XMM-LSS fields, in conjunction with three high-resolution cosmological hydrodynamical simulations. In the sections below we describe each data set and the extraction of environmental parameters and briefly outline the characteristics of the simulations.

---

[1]The mass function measures the number density of galaxies as a function of stellar mass.

2.1 The COSMOS field

We use the Classic version of the COSMOS2020 catalogue (J. R. Weaver et al. 2022), which provides physical parameters, such as photometric redshifts, stellar masses and SFRs for around 1.7 million sources in the $\sim 2\,\mathrm{deg}^2$ COSMOS (N. Scoville et al. 2007) field (centred at 10h, $+02°$). The physical parameters are calculated using deep photometry in around 40 broad and medium band filters, from the UV through to the mid-infrared, from the following instruments: *GALEX* (M. A. Zamojski et al. 2007), MegaCam/CFHT (M. Sawicki et al. 2019), ACS/*HST* (A. Leauthaud et al. 2007), Hyper Suprime-Cam (H. Aihara et al. 2019), Subaru/Suprime-Cam (Y. Taniguchi et al. 2007, 2015), VIRCAM/VISTA (H. J. McCracken et al. 2012) and IRAC/*Spitzer* (M. L. N. Ashby et al. 2013; C. L. Steinhardt et al. 2014; M. L. N. Ashby et al. 2015, 2018).

The detection image used for object identification in COS-MOS2020 includes optical (i, z) images from the Ultradeep layer of the HSC Subaru Strategic Programme (HSC-SSP), which has a point-source depth of $\sim 28\,\mathrm{mag}$ (H. Aihara et al. 2019). As a comparison, this is $\sim 10\,\mathrm{mag}$ fainter than the magnitude limit of the SDSS spectroscopic main galaxy sample (e.g. S. Alam et al. 2015). The optical and infrared aperture photometry is extracted using the SExtractor (E. Bertin & S. Arnouts 1996) and IRACLEAN (B.-C. Hsieh et al. 2012) codes, respectively. Parameter estimation is then performed using the LePhare SED-fitting algorithm (S. Arnouts et al. 2002; O. Ilbert et al. 2006) run on all broad and medium band filters from UV through to IRAC Channel 2 (which has a central wavelength of $4.5\,\mu\mathrm{m}$). The deep data, combined with the large number of photometric filters, results in photometric redshift accuracies better than ~ 1 and ~ 4 per cent for bright ($i < 22.5\,\mathrm{mag}$) and faint ($25 < i < 27\,\mathrm{mag}$) galaxies, respectively. In our analysis below, we use the redshifts and stellar masses directly from the COSMOS2020 catalogue.

2.2 The XMM-LSS field

We use deep u band to near-infrared photometry, produced using SExtractor, from the CLAUDS XMM-LSS catalogue (G. Desprez et al. 2023; V. Picouet et al. 2023), which is based on data within a $4.1\,\mathrm{deg}^2$ footprint in the XMM-LSS field (centred at 02h, $-04°$) from the following instruments: MegaCam/CFHT (M. Sawicki et al. 2019), Hyper Suprime-Cam (H. Aihara et al. 2019), and VIRCAM/VISTA (H. J. McCracken et al. 2012).

The detection image used for object identification in this data set includes optical (i, z) images from the Deep layer of the HSC-SSP, which has a point-source depth of $\sim 27\,\mathrm{mag}$ (1 mag shallower than the HSC-SSP images in COSMOS). Photometric redshifts and stellar masses are calculated in an identical fashion to COSMOS2020 via LePhare, using the broad-band $ugrizyYJHK$ photometry available in this catalogue. The accuracy of the physical parameters are comparable to the ones obtained for the COSMOS2020 catalogue. The different locations and sizes of the COSMOS and XMM-LSS fields aids in reducing the uncertainties related to cosmic variance when considering the galaxy stellar mass function.

2.3 DisPerSE: measuring environmental parameters

In the sections below, we study how the stellar mass function varies with environment, defined as the distance from LSS. To perform this analysis, we measure the projected distances of our dwarf galaxies from the nodes and filaments that define the cosmic web using the DisPerSE algorithm (T. Sousbie 2011). We follow I. Lazar et al.
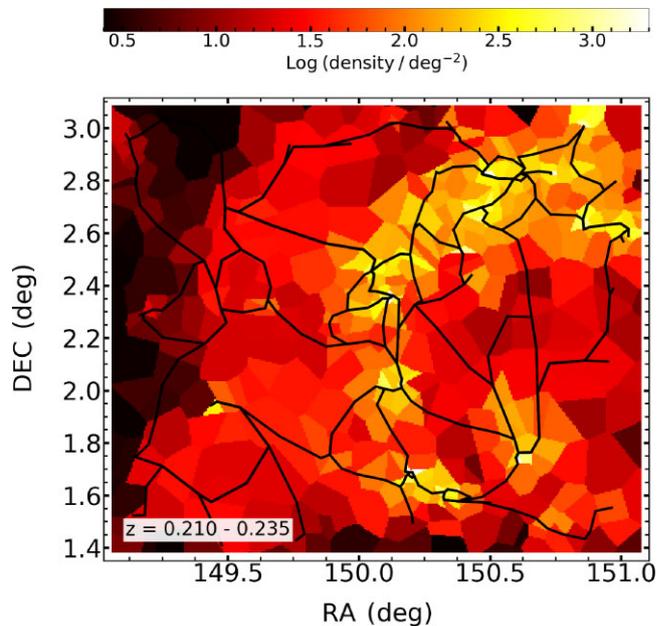

Figure 1. An example number density map in the COSMOS field at $0.210 < z < 0.235$ created using DisPerSE. The colours indicate the local density (see colour bar), while the solid black lines show the locations of the filaments.

(2023) and B. Bichang'a et al. (2024), who have performed an identical density analysis using the COSMOS2020 catalogue. DisPerSE starts by producing a density map using Delaunay tessellations, which is calculated using the spatial positions of galaxies (W. E. Schaap & R. van de Weygaert 2000). Stationary (critical) points in the density map (i.e. the maxima, minima, and saddles) correspond to local nodes, voids and the centres of filaments respectively. Segments are used to connect the local maxima with saddle points, forming a set of ridges that defines the network of filaments that describes the cosmic web. A 'persistence' parameter (N) is used to set a threshold value for defining pairs of critical points that are used to produce the final density map. Only critical pairs with Poisson probabilities above $N\sigma$ from the mean are retained. Following I. Lazar et al. (2023) and B. Bichang'a et al. (2024) we use a persistence value of 2 in our study, which removes ridges close to the noise level, where structures could be spurious.

We use massive ($M_\star > 10^{10}\,\mathrm{M}_\odot$) galaxies to construct the density maps because they exhibit the smallest redshift errors ($\delta z \sim 0.008$ at $z \sim 0.2$) and dominate their local gravitational potential wells. The accuracy of the COSMOS2020 and CLAUDS XMM-LSS redshifts enables us to employ well-defined and relatively narrow redshift slices within which we build our 2D density maps. When constructing each map, individual galaxies are weighted by the area under their redshift probability density function that is contained within the slice in question. This ensures that the uncertainties in the photometric redshifts are propagated through the construction of the maps. Once the maps have been constructed and the stationary points located, we calculate the projected distances from the nearest nodes and filaments for each galaxy in our catalogues. Finally, as noted in I. Lazar et al. (2023), the low number counts of massive galaxies at $z < 0.2$ makes it difficult to create density maps at these redshifts. Our analysis of how the mass function varies as a function of location in the cosmic web therefore focuses on the redshift range $0.2 < z < 0.4$. Fig. 1 shows an example density map from our analysis.

We complete this section by considering the types of environments that are likely to be present in our COSMOS and XMM-LSS

Table 1. Key characteristics of the cosmological hydrodynamical simulations used in this study.

	NEWHORIZON	TNG50	FIREBOX
Code	RAMSES (R. Teyssier 2002)	AREPO (V. Springel 2010)	GIZMO (P. F. Hopkins 2015)
Volume	Zoom of 20 Mpc spherical region	50 Mpc box	22 Mpc box
Spatial resolution	34 pc	100–140 pc	20 pc
Mass resolution	$m_\star = 1.3 \times 10^4\,M_\odot$	$m_\star = 8.5 \times 10^4\,M_\odot$	$m_\star = 6 \times 10^4\,M_\odot$
Environment	Maximum halo mass $\sim 10^{13}\,M_\odot$	Maximum halo mass $\sim 10^{14}\,M_\odot$	Maximum halo mass $\sim 10^{13}\,M_\odot$
Star formation	Dense, self-gravitating gas; efficiency modulated by turbulence	Schmidt law in two-phase ISM; density threshold-based (V. Springel & L. Hernquist 2003)	Cold, dense, self-gravitating gas; 100% efficiency per freefall time (P. F. Hopkins et al. 2018)
SN feedback	Mechanical feedback from SN Type II (T. Kimm & R. Cen 2014)	Direct heating and delayed kinetic winds from SNe I/II and stellar winds (V. Springel & L. Hernquist 2003)	Mechanical feedback from SNe I/II, stellar winds and radiation (P. F. Hopkins et al. 2018), no AGN feedback
ISM physics	Partially resolved multiphase ISM	Idealized two-phase model with effective equation of state	Partially resolved multiphase ISM

footprints in the redshift range $0.2 < z < 0.4$, by considering the M_{200} values (which are proxies for the virial masses) of groups identified in the literature (A. Finoguenov et al. 2007; M. R. George et al. 2011; G. Gozaliasl et al. 2014, 2019). The virial masses of groups in COSMOS lie in the range $10^{13}\,M_\odot < M_{200} < 10^{14.5}\,M_\odot$, with a median value of $10^{13.4}\,M_\odot$. The virial masses of groups in XMM-LSS lie in the range $10^{13.3}\,M_\odot < M_{200} < 10^{14.1}\,M_\odot$, with a similar median value of $10^{13.6}\,M_\odot$. For comparison, a small cluster like Fornax has a virial mass of $\sim 10^{13.9}\,M_\odot$ (M. J. Drinkwater, M. D. Gregg & M. Colless 2001), while large clusters like Virgo and Coma have virial masses of $\sim 10^{15}\,M_\odot$ (e.g. P. Fouqué et al. 2001; R. Gavazzi et al. 2009). The galaxy population considered in this study therefore resides mainly in relatively low-density environments (i.e. groups and the field) and not in rich clusters. It is worth noting that the simulations used in Section 6 bracket the types of environments in the observations and are also not focused on clusters, making the comparisons between data and theory consistent.

2.4 Cosmological hydrodynamical simulations

We briefly describe the cosmological hydrodynamical simulations that are compared to the observational data in Section 6 below. Table 1 summarizes the properties of each simulation, while the sections below provide a brief outline of their key characteristics.

2.4.1 NEWHORIZON

NEWHORIZON (Y. Dubois et al. 2021) is a high-resolution zoom-in of a 20 Mpc diameter spherical volume of the parent Horizon-AGN simulation (Y. Dubois et al. 2014; S. Kaviraj et al. 2017), which uses the adaptive mesh refinement code RAMSES (R. Teyssier 2002). NEWHORIZON achieves a maximum spatial refinement of 34 pc and a stellar mass resolution of $1.3 \times 10^4\,M_\odot$ (Y. Dubois et al. 2021). It is not calibrated to match specific observational data, except for AGN feedback, whose efficiency is calibrated using the local $M_{BH}-M_*$ relation. Star formation is based on dense, self-gravitating gas, with an efficiency regulated by local turbulence. Feedback is implemented via mechanical supernova feedback from Type II supernovae (T. Kimm & R. Cen 2014) and the interstellar medium (ISM) is partially resolved, allowing for multiphase structure to emerge.

2.4.2 TNG50

TNG50 (D. Nelson et al. 2019a; A. Pillepich et al. 2019) is part of the IllustrisTNG suite (A. Pillepich et al. 2018; V. Springel et al. 2018; D. Nelson et al. 2018; J. P. Naiman 2018; F. Marinacci et al. 2018; D. Nelson et al. 2019b) and simulates a 50 Mpc cosmological box using the moving mesh code AREPO (V. Springel 2010). It includes

magneto-hydrodynamics and is explicitly calibrated to match a range of low-redshift observables, such as the stellar mass function and galaxy sizes in the intermediate and high-mass regime. Star formation follows a Kennicutt–Schmidt relation (R. C. Kennicutt 1998) within an idealized two-phase ISM governed by an effective equation of state (V. Springel & L. Hernquist 2003). Feedback from AGN, Type I and Type II supernovae, and stellar winds is implemented through a combination of thermal and kinetic modes. Stellar feedback is injected as kinetic winds and initially decoupled from the dense ISM. AGN feedback includes both thermal energy injection in the high-accretion regime and decoupled kinetic winds in the low-accretion regime.

2.4.3 FIREBOX

FIREBOX (R. Feldmann et al. 2023) simulates a 22 Mpc cosmological box based on the FIRE-2 model, which is run using the mesh-free finite mass code GIZMO (P. F. Hopkins 2015) and provides high spatial (20 pc) and mass ($m_\star = 6 \times 10^4\,M_\odot$) resolution. The code is designed to minimize the number of tunable parameters, eschewing empirical calibration in favour of a physics-based approach that aims to model well-understood physical processes at sufficient resolution. Star formation occurs in explicitly cold, dense, self-gravitating gas with an assumed 100 per cent efficiency per local freefall time (P. F. Hopkins et al. 2018). Feedback includes mechanical energy and momentum from both Type I and Type II supernovae, stellar winds and radiation, but omits AGN feedback. The ISM is partially resolved, allowing for multiphase gas dynamics.

3 COMPLETENESS

We begin by quantifying the mass and redshift ranges within which mass-complete samples of galaxies can be constructed, building on our previous work. Using complete samples is desirable because it avoids the need to apply completeness corrections which carry associated uncertainties and biases (see e.g. A. K. Weigel, K. Schawinski & C. Bruderer 2016). In S. Kaviraj et al. (2025), we have used a purely old simple stellar population (SSP), that forms at $z \sim 2$, to estimate the redshifts at which galaxy populations are likely to be complete as a function of stellar mass. An SSP is defined as one where all stars form simultaneously. Assuming that this hypothetical stellar population is a faintest limiting case, it follows that, if this population is detectable in a given survey, then all galaxies will also be detectable and the galaxy population will therefore be mass-complete.

For example, S. Kaviraj et al. (2025) estimate that galaxies down to $10^8\,M_\odot$ should be complete out to at least $z \sim 0.4$ at the depth of the HSC data in the COSMOS field (see their Fig. 4). Note that
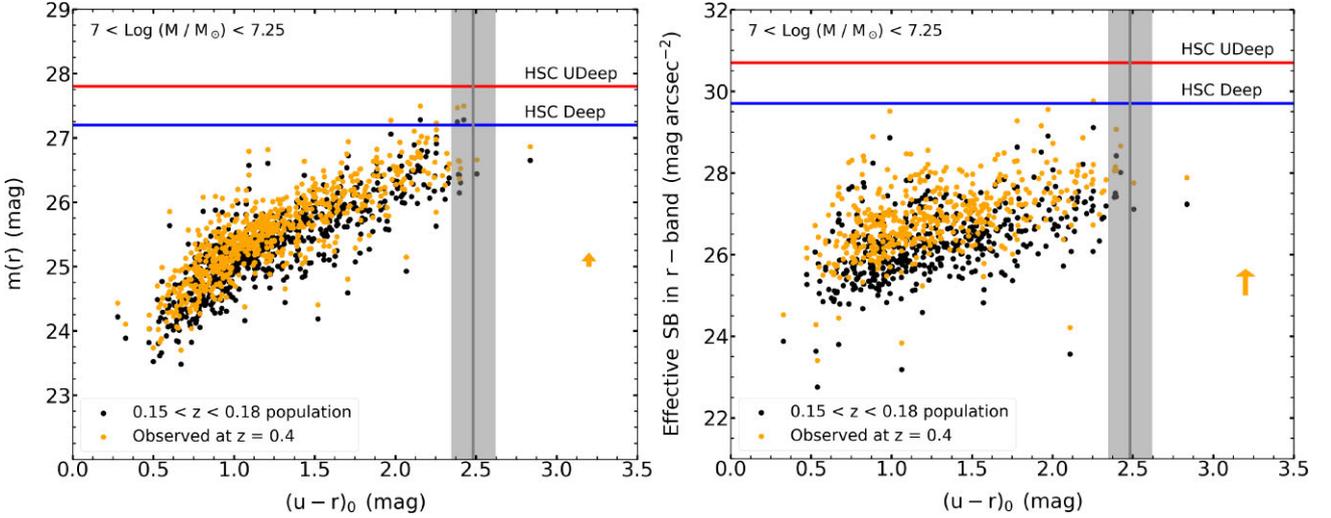
Figure 2. Total magnitudes (left) and effective surface brightnesses (right) versus rest-frame $(u − r)$ colours of COSMOS2020 galaxies in the stellar mass and redshift ranges of $10^7 \, M_\odot < M_\star < 10^{7.25} \, M_\odot$ and $0.15 < z < 0.18$, respectively. The black points indicate the COSMOS2020 galaxies. The grey shaded region indicates the location of a purely old SSP that forms at $z = 2$ (assuming half solar metallicity). The upper and lower limits of the grey shaded region indicate the rest-frame $(u − r)$ colour of this SSP with and without a reddening of $E(B − V) = 0.1$, respectively. The solid grey vertical line indicates the mid-point of the grey shaded region. The red and blue horizontal lines indicate the detection limits of the Deep and Ultradeep layers of the HSC-SSP imaging.

J. R. Weaver et al. (2022) come to an identical conclusion using a different technique. However, S. Kaviraj et al. (2025) also note that the completeness redshifts thus derived are pessimistic because the vast majority of galaxies at any stellar mass (but particularly in the dwarf regime) do not consist of purely old stellar populations (see their Fig. 5). Here we refine the completeness redshifts presented in Fig. 4 in S. Kaviraj et al. (2025) using observed dwarf galaxies from the COSMOS2020 catalogue.

In Fig. 2, we plot the total magnitudes and effective surface brightnesses versus the rest-frame $(u − r)$ colours of galaxies in the stellar mass and redshift ranges of $10^7 \, M_\odot < M_\star \, 10^{7.25} \, M_\odot$ and $0.15 < z < 0.18$, respectively. We select these galaxies because S. Kaviraj et al. (2025) demonstrate that dwarfs with $M_\star \sim 10^7 \, M_\odot$ are complete out to at least $z \sim 0.18$. The black points indicate the COSMOS2020 galaxies. The grey shaded region indicates the location of a purely old SSP that forms at $z = 2$ (assuming half solar metallicity). The lower and upper limits of the grey shaded region indicate the rest-frame $(u − r)$ colour of this SSP without any reddening and with a reddening of $E(B − V) = 0.1$, respectively. The solid grey vertical line indicates the mid-point of the grey shaded region. The red and blue horizontal lines indicate the detection limits of the Deep and Ultradeep layers of the HSC-SSP imaging.

As already noted by S. Kaviraj et al. (2025), less than 1 per cent of dwarfs are actually consistent with a purely old SSP (i.e. reside inside the grey shaded region). It is also apparent that the vast majority of galaxies in this mass range are brighter than the detection limit of both the HSC Deep and Ultradeep imaging. The orange points show how our COSMOS2020 galaxy population moves in this parameter space if they are redshifted to $z = 0.4$. The shift, indicated by the orange arrows, includes both distance dimming and a k-correction. Here we apply the k-correction from our purely old SSP that forms at $z = 2$, which produces a larger k-correction than stellar populations with younger ages. In other words, the k-correction, and therefore the shift in the galaxies applied here, is likely to be maximal.

We find that, while S. Kaviraj et al. (2025) show that the galaxy population at $M_\star \sim 10^7 \, M_\odot$ is complete to *at least* $z \sim 0.18$, in reality, since hardly any galaxies are consistent with purely old

stellar populations, galaxies with $M_\star \sim 10^7 \, M_\odot$ are likely to remain complete out to around $z \sim 0.4$ in both the Deep and Ultradeep layers of the HSC imaging. In these mass and redshift ranges complete samples of galaxies can be constructed, removing the need to apply completeness corrections.

4 MEASUREMENT OF MASS FUNCTION UNCERTAINTIES

The uncertainty associated with the galaxy stellar mass function (σ_Φ) is estimated by adding, in quadrature, the Poisson noise (σ_P), the uncertainties on the stellar masses that are calculated via SED fitting (σ_M) and cosmic variance (σ_{CV}), such that $\sigma_\Phi = \sqrt{\sigma_P^2 + \sigma_M^2 + \sigma_{CV}^2}$. We briefly describe each contribution to the uncertainty below.

4.1 Poisson noise

The Poisson noise measures the uncertainty in the number counts (for low number counts the Poisson noise can make a significant contribution to the overall uncertainty in the mass function). We define the Poisson uncertainty as $\sigma_P = \sqrt{N}/V$ where N represents the number count in a given mass bin and V is the maximum volume within which objects in that mass bin can be detected, given the redshift range and mass completeness of the sample.

4.2 Uncertainty on the stellar mass estimated via SED fitting

We calculate σ_M via a Monte Carlo approach, in which 1000 independent realisations of the galaxy stellar mass function are constructed, by drawing a random stellar mass value from the stellar mass likelihood distribution of each object in each realization. We assume a Gaussian likelihood distribution with a standard deviation that is equal to the stellar mass uncertainty measured by the SED fitting (in log stellar mass space). We estimate σ_M by calculating

the standard deviation of Φ in each stellar mass bin across all realizations.[2]

4.3 Cosmic variance

Cosmic variance refers to the fact that the statistical properties of galaxies (e.g. the distribution of stellar masses) are affected by LSS. This can lead to field-to-field differences in different parts of the sky, which are larger than the Poisson noise. Larger sky areas lead to smaller cosmic variance and vice-versa. We estimate the uncertainties due to cosmic variance using the PYTHON version of the Cosmic Variance Cookbook (C. K. Jespersen et al. 2025), originally developed in IDL by B. P. Moster et al. (2011).

5 THE DWARF STELLAR MASS FUNCTION IN DIFFERENT ENVIRONMENTS

It has been shown in a variety of studies, using both simulations and observations, that galaxy properties such as SFR, colour and stellar mass, at least in the massive-galaxy regime, change as a function of distance from LSS traced by the nodes and filaments of the cosmic web. For example, galaxies that are more massive and more quiescent typically reside closer to nodes and filaments, while environments further away from LSS typically host lower mass galaxies with higher levels of star formation. This dependence on the distance from LSS has been demonstrated both in observational (e.g. C. Laigle et al. 2018) and theoretical (e.g. K. Kraljic et al. 2018; F. Hasan et al. 2023) work, at least in the low and intermediate redshift Universe ($z < 2$). Indeed, galaxy properties may change more noticeably with varying distance to LSS than with just the numerical value of local density (e.g. C. Laigle et al. 2018). Recent empirical work using the SDSS (S. Zarattini & J. A. L. Aguerri 2025) suggests that galaxy SFRs begin to show statistical differences, compared to those in the field, for galaxies located within 5 Mpc of filaments. Galaxy colours show similar changes for galaxies located within ∼1 Mpc of filaments. This suggests that distance from LSS, on scales of around 5 Mpc or less, plays an important role in influencing the physical properties of galaxies, at least in the massive-galaxy regime.

In light of these results we use the distance from LSS as our metric for environment. We note first that, while we explore the COSMOS data set down to stellar masses of $10^7 \, M_\odot$, we restrict the XMM-LSS data set to stellar masses above $10^8 \, M_\odot$. This is because the mass function in XMM-LSS shows an unexpected downturn at $M_\star < 10^8 \, M_\odot$, which is not shared by its COSMOS counterpart (see the discussion in Appendix D and Fig. D1). We speculate that this inconsistency between XMM-LSS and COSMOS may be driven by the availability of medium-band filters in the SED fitting process, since the SED fitting in COSMOS benefits from the presence of medium-band filters while the XMM-LSS data set does not. Indeed, when we use the same filter set as that employed for the SED fitting in XMM-LSS (i.e. broad-band $ugrizyJHK$ only) on COSMOS data, we obtain a similar downturn in the COSMOS mass function (see Fig. D1). This suggests that the lack of medium-band filters

---

[2] It is worth noting here that the systematic statistical bias induced by more low mass galaxies scattering to high masses than vice versa, due to the intrinsic shape of the mass function (Eddington bias, see e.g. D. Obreschkow et al. 2018), is negligible in this work. Appendix E demonstrates this by showing that the mean galaxy stellar mass functions obtained from the perturbed stellar masses are very similar (within ∼0.01 dex) of the unperturbed mass functions in both the COSMOS and XMM-LSS fields.
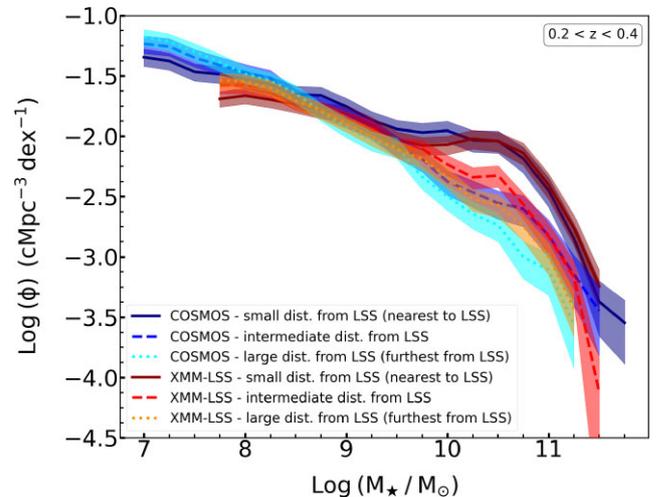

Figure 3. The galaxy stellar mass function for the COSMOS and XMM-LSS fields in different environments. Galaxies at small distances from LSS (i.e. nearest to LSS) are defined as those that reside in the lower tercile (0th–33rd percentile values) of both the distances to nodes and the distances to filaments. Galaxies at intermediate and large distances from LSS are defined as those that reside in the middle (33rd–66th percentile values) and upper (66th–100th percentile values) terciles of both the distances to nodes and the distances to filaments, respectively.

may lead to a poorer sampling of the SED which appears to affect very low mass ($M_\star < 10^8 \, M_\odot$) galaxies. Note that this downturn has also been noted in previous work that has only used broad-band filters for measuring SED fitted stellar masses in the XMM-LSS and COSMOS fields (e.g. N. J. Adams et al. 2021). Thus, as a precaution we consider the XMM-LSS data set only down to $M_\star = 10^8 \, M_\odot$ where the XMM-LSS mass function shows good consistency with its COSMOS counterpart, regardless of the filters used in the SED fitting. This issue is discussed further in Appendix D.

5.1 Variation of the mass function with distance from LSS

We proceed by using the outputs from DISPERSE to explore the relationship between the galaxy stellar mass function and the proximity of galaxies to LSS. We first split our galaxies into three terciles in terms of their projected distances from the nearest node and three terciles in terms of their projected distances from the nearest filament. We then define three galaxy populations that reside at 'small', 'intermediate' and 'large' distances from LSS as follows. Galaxies at small distances from LSS (i.e. nearest to LSS) are defined as those that reside in the lowest tercile (0th–33rd percentile values) of both the distances to nodes and the distances to filaments. Similarly, galaxies at intermediate and large distances from LSS are defined as those that reside in the middle (33rd–66th percentile values) and upper (66th–100th percentile values) terciles of both the distances to nodes and the distances to filaments, respectively. In Fig. 3, we show the galaxy stellar mass function for these three populations.

Two clear differences arise when studying how the mass function behaves with respect to distance from LSS, regardless of the field (i.e. COSMOS or XMM-LSS) being considered. First, the knee of the mass function becomes more pronounced in regions that are closer to LSS. The galaxy number density around the knee of the mass function is ∼0.4 dex larger in galaxies that are in the lower tercile in their distance to both nodes and filaments (i.e. nearest to LSS) compared to those that are in the upper tercile (i.e. furthest from LSS). In other

words, the number density of massive systems increases as we move closer to LSS. Second, the mass function in the dwarf galaxy regime at $M_\star < 10^{8.5}$ M$_\odot$ becomes flatter – i.e. there are fewer dwarfs per unit volume – as we move closer to LSS. The net effect is that the ratio of massive to dwarf galaxies increases with greater proximity to LSS. For completeness, we note that these trends also exist (although with varying significance) when considering the stellar mass function for different terciles in the distances to filaments and nodes *separately* (see discussion in Appendix C and Fig. C1). Fig. C1 demonstrates that the environmental trends are driven principally by the distances of galaxies from filaments, rather than their distances from nodes.

The trends described above are consistent with the predictions of the ΛCDM paradigm. The galaxy stellar mass distribution is predicted to depend on the surrounding large-scale environment due to biased clustering (N. Kaiser 1984), with the non-linear collapse of over-dense regions and their location within the cosmic web determining where and how halos form (e.g. J. R. Bond, L. Kofman & D. Pogosyan 1996). The galaxy bias – particularly in terms of the location of galaxies with respect to filaments (M. Musso et al. 2018) – is related to the distribution of dark matter, as a consequence of the large-scale tides that influence the assembly history of halos. K. Kraljic et al. (2019) have used a cosmological hydrodynamical simulation to demonstrate that the overall impact of this process is two-fold. The stellar mass of galaxies increases with decreasing distance from the cosmic web (in the transverse and longitudinal directions with respect to the filament-type saddle point) while the number density of massive galaxies is higher closer to filaments compared to the number density of lower-mass galaxies (see also H. Song et al. 2021), in good agreement with our empirical results.

We continue by performing a quantitative analysis of the differences that are observed in the mass function in different environments. Galaxy mass functions have often been parametrized using a Schechter function (P. Schechter 1976). An extension of this procedure is to use a double Schechter function, which provides a better representation of the observed upturn in galaxy number densities when low mass galaxies are included, at least at low and intermediate redshift (e.g. I. K. Baldry et al. 2012; J. R. Weaver et al. 2022). The double Schechter function is expressed as follows:

$$\Phi \, \mathrm{d} \log M = \ln(10) \, \mathrm{e}^{-10^{\log M - \log M^\star}}$$
$$\times [\Phi_1^\star (10^{\log M - \log M^\star})^{\alpha_1 + 1} \qquad (1)$$
$$+ \; \Phi_2^\star (10^{\log M - \log M^\star})^{\alpha_2 + 1}] \, \mathrm{d} \log M.$$

Here, M^* represents a characteristic mass scale (which corresponds to the knee of the mass function) beyond which there is an exponential decline in galaxy number densities at high stellar masses. α_1 and α_2 are the slopes of the two Schechter functions which have normalizations $\Phi_1^\star$ and $\Phi_2^\star$, respectively. The slope of the fitted function at stellar masses much lower than the knee of the mass function is controlled by the more negative α value, while the slope nearer the knee is controlled by the more positive α value.

We fit this double Schechter function to our mass functions in different environments. In Fig. 4, we present the values of M^* and α_1 (the more negative of the α values, which represents the slope of the mass function in the dwarf regime) for the three galaxy populations defined above. Depending on the field in question, the values of M^* and α_1 become more positive by 0.3–0.4 and 0.1–0.2 dex, respectively, in galaxies that are in the lower tercile in their distances to both nodes and filaments (i.e. nearest to LSS), compared to those that are in the upper tercile (i.e. furthest from LSS). In other words, the knee of the mass function moves to larger stellar masses and the slope in the dwarf regime becomes flatter as we move closer
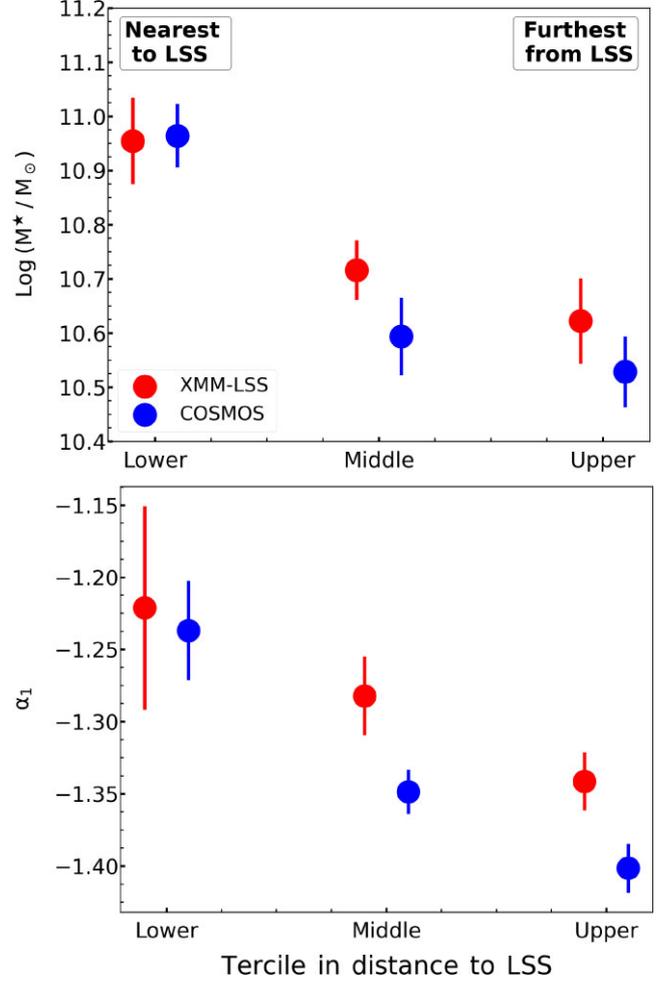

Figure 4. Best-fitting parameter values and their uncertainties for $M^\star$ and α_1 from the double Schechter fits to the mass functions in the COSMOS and XMM-LSS fields in different environments. Galaxies at small distances from LSS (i.e. nearest to LSS) are defined as those that reside in the lower tercile (0th–33$^{\mathrm{rd}}$ percentile values) of both the distances to nodes and the distances to filaments. Galaxies at intermediate and large distances from LSS are defined as those that reside in the middle (33rd–66th percentile values) and upper (66th–100th percentile values) terciles of both the distances to nodes and the distances to filaments, respectively.

to LSS (as is evident from visual inspection of Fig. 3). The best-fitting parameters of our Schechter fits are presented in Table B1.

5.2 Comparison with the observational literature

We complete this section by comparing our results to the recent observational literature. It is worth noting first that complete consensus about the role of environment on the shape of the mass function, especially in the dwarf regime, does not yet exist. This is likely to be due – at least partly – to the heterogeneity of methods that are employed to define environment in the literature and the effects of cosmic variance. Nevertheless, our results are in good agreement with some recent studies that also use deep photometric data to extract mass functions (typically via SED fitting) and probe its evolution as a function of environment.

For example, N. Malavasi et al. (2017) find that, in the UltraVISTA-COSMOS field in our redshift range of interest, denser environments exhibit a larger number density of massive galaxies and
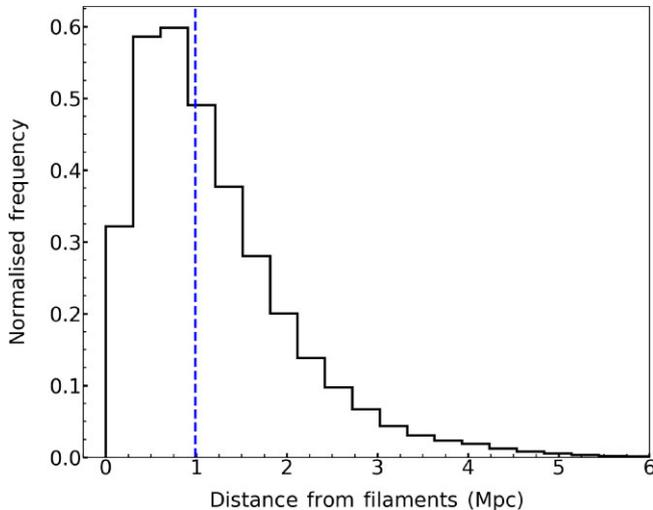
Figure 5. The distribution of projected distances from the nearest filaments for galaxies in the COSMOS and XMM-LSS fields. The median distance is shown using the dashed blue vertical line.

a smaller number density of dwarfs. A very similar result is found by J. Etherington et al. (2017), using a statistical sample of ~3.2 million galaxies observed as part of the Dark Energy Survey (DES). They find that the number density of dwarfs in denser environments is lower than that in more sparse environments by ~0.4 dex and that massive galaxies in denser environments outnumber their counterparts in sparse environments by a similar amount. The results of both these studies are very similar to the results of this study, shown in Fig. 3. Similarly, A. Mortlock et al. (2015) use deep *HST* observations in the GOODS and UDS fields to show that more massive galaxies exist per unit volume in denser environments at our redshifts of interest (but interestingly do not report significant differences in the dwarf regime in different environments).

All three studies described above define environment by calculating galaxy number counts within cylinders or conical frusta, where the projected radii of these shapes is around 1 Mpc or less, indicating that they trace local environment on relatively small spatial scales. As a comparison we show, in Fig. 5, the distribution of distances from the nearest filaments for the galaxy population studied here. The median value of this distribution, shown using the vertical blue line, lies close to 1 Mpc while the distances reach a maximum value of ~6 Mpc. The spatial scales probed here are therefore similar to those in the studies described above, which explains the consistency between our results.

Together with the literature, our results suggest that distance from LSS has a strong influence on galaxy stellar mass, on physical scales of around ~1 Mpc or less. Interestingly, both N. Malavasi et al. (2017) and S. Zarattini & J. A. L. Aguerri (2025), as discussed above, suggest that these effects start to diminish once physical scales larger than around 4–5 Mpc are probed.

6 IS THERE A MISSING DWARFS PROBLEM IN THE CONTEXT OF THE ΛCDM PARADIGM?

As noted in the introduction, the missing satellites problem is a longstanding potential tension in the ΛCDM paradigm between the predicted and observed number densities of low mass galaxies. Specifically, it has been claimed that observational data may show a deficit of dwarfs compared to the predictions of simulations, which

makes the observed stellar mass function of low-mass galaxies in our local neighbourhood much flatter than its theoretical counterpart (e.g. A. Klypin et al. 1999; B. Moore et al. 1999). However, detailed explorations of this issue are largely restricted to dwarfs around either the Milky Way or other local massive galaxies (e.g. A. Smercina et al. 2018; M. Tanaka et al. 2018; P. Bennet et al. 2020; M. Nashimoto et al. 2022; Y. Mao et al. 2024; E. Kado-Fong et al. 2025 and references therein).

A potential solution to the missing satellites problem involves invoking warm or fuzzy dark matter, which suppresses the formation of low-mass halos and therefore the number density of dwarf galaxies (e.g. W. Hu, R. Barkana & A. Gruzinov 2000; E. Polisensky & M. Ricotti 2011; M. R. Lovell et al. 2012; D. J. E. Marsh 2016). Alternatively, the inclusion of baryonic physics can also reduce the number density of dwarf galaxies within the framework of the ΛCDM paradigm, without altering the properties of dark matter, largely through the reduction of galaxy SFRs via star formation feedback (e.g. S. Geen, A. Slyz & J. Devriendt 2013; R. Hazenfratz et al. 2024) and AGN (e.g. G. Dashyan et al. 2018; S. Kaviraj, G. Martin & J. Silk 2019; S. Koudmani et al. 2021; F. Davis et al. 2022) and ionising UV radiation (e.g. J. C. Forbes et al. 2016; A. Emerick, G. L. Bryan & M.-M. Mac Low 2019). Recent work using high-resolution hydrodynamical simulations suggests that, when such realistic baryonic physics is implemented, the number densities of low-mass galaxies around Milky Way like galaxies in ΛCDM-based simulations can be consistent with their observed counterparts (e.g. T. Sawala et al. 2016; I. M. E. Santos-Santos et al. 2022; M. Jung et al. 2024; S. Jeon et al. 2025).

It is worth noting that dwarfs in close proximity to the Milky Way are, by construction, satellites. This is why past explorations of this issue have been crafted in terms of missing *satellite* galaxies. However, if such a problem does indeed exist, it cannot be restricted to just the vicinity of massive galaxies like the Milky Way. In other words, a discrepancy between the predicted and observed number densities should manifest itself as a missing *dwarfs* problem in the context of the ΛCDM paradigm. The deep data now available provides an unprecedented opportunity to explore this issue in the general dwarf population across all environments selected from a deep-wide survey, rather than in only the relatively high-density neighbourhoods of very nearby massive Milky Way like galaxies. Note, however, that in our case, this exercise is restricted to $M_\star > 10^7 \, M_\odot$ and $z < 0.4$, where we can construct complete galaxy samples.

We complete our study by exploring whether a comparison between the observed and theoretical stellar mass functions suggests the presence of a generic missing dwarfs problem in the context of ΛCDM. In Fig. 6, we compare the observed mass functions in the COSMOS and XMM-LSS fields, and an ensemble of empirical results drawn from the literature, to the predicted mass functions from three cosmological hydrodynamical simulations (NEWHORIZON, TNG50, and FIREBOX). For this exercise, we study the total mass function, constructed from all galaxies in both the observations and the simulations in our mass and redshift ranges of interest. There is excellent agreement between the observed mass functions calculated in this study and their empirical counterparts from the literature.

We note that the mass functions from the three simulations do not completely agree with each other. While there is good agreement at the low mass end of the galaxy population considered in this study ($M_\star < 10^8 \, M_\odot$), the disagreement between the different theoretical mass functions can be as large as ~0.5 dex at some points within the dwarf regime. The survivability and evolution of galaxies at
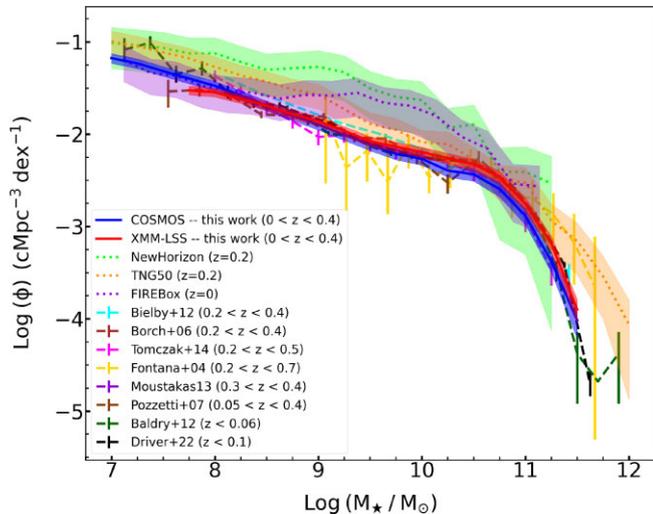
Figure 6. The observed galaxy stellar mass function in the nearby Universe in the COSMOS and XMM-LSS fields and an ensemble of observations from the literature, compared to theoretical mass functions from three high resolution cosmological hydrodynamical simulations (NEWHORIZON, TNG50, and FIREBOX). The mass functions from the literature are taken from S. P. Driver et al. (2022), I. K. Baldry et al. (2012), R. Bielby et al. (2012), A. Borch et al. (2006), A. R. Tomczak et al. (2014), A. Fontana et al. (2004), J. Moustakas et al. (2013), and L. Pozzetti et al. (2007).

dwarf masses can depend on simulation resolution, particularly in cluster environments (G. Martin et al. 2024). But, given the low-density environments probed by these simulations and the similar mass/spatial resolutions and cosmology adopted by each model, these discrepancies are more likely to be driven by differences in the implementation of baryonic physics, such as baryonic feedback and ISM gas physics (e.g. G. Martin et al. 2025).

It is also worth noting that the recent literature suggests that stellar masses derived from SED fitting systematically underestimates the true stellar masses of galaxies (e.g. C. Laigle et al. 2019; M. A. C. de los Reyes et al. 2025). In particular, C. Laigle et al. (2019) demonstrate that performing SED fitting using the same filters as those used in this study underestimates the stellar mass by $\sim$0.12 dex at $M_\star \sim 10^9\,M_\odot$, with the discrepancy increasing with decreasing stellar mass. Extrapolation of the results of their study suggests that the underestimate could be $\sim$0.25 dex at $M_\star \sim 10^7\,M_\odot$, with an average value of $\sim$0.2 dex in the dwarf regime. While Fig. 6 already demonstrates reasonable consistency between theory and data, raising the observed mass function by $\sim$0.2 dex improves the consistency between the observed and theoretical mass functions even further. To illustrate this point we show, in Fig. A1, a version of Fig. 6 where the observed mass functions in the dwarf regime have been shifted upwards by 0.2 dex.

Given these two points, we do not find evidence that there is a missing dwarfs problem in the observational data, in line with recent work that suggests that the missing satellites problem is also solved when baryonic physics is employed in a realistic manner in simulations. Nevertheless, further work is desirable to fully explore this issue in low density environments, using both higher resolution simulations in larger volumes, in order to probe the dwarf regime down to lower stellar masses, and data from forthcoming surveys like LSST (Ž. Ivezić et al. 2019) which offer large footprints that can alleviate the effect of cosmic variance.

7 SUMMARY

We have combined deep photometric data in the COSMOS and XMM-LSS fields with high resolution cosmological hydrodynamical simulations to explore two key open questions in our understanding of dwarf-galaxy evolution: (1) how does the stellar mass function, particularly in the dwarf regime, vary with environment, defined here as the distance from LSS? (2) is there a generic 'missing dwarfs' problem in the context of the ΛCDM paradigm, when all environments – and not just the regions around Milky Way like galaxies – are considered? The unprecedented depth of the observational survey data now available in COSMOS and XMM-LSS facilitates the construction of mass-complete galaxy populations, down to $M_\star \sim 10^7\,M_\odot$ and out to $z \sim 0.4$. This enables us to tackle these questions in an unbiased manner, without the need for completeness corrections. Our main conclusions are as follows:

(i) Distance from LSS has a strong influence on the shape of the galaxy stellar mass function at both the dwarf and massive-galaxy ends. As we move closer to LSS, the number densities of massive and dwarf galaxies increase and decrease respectively, resulting in a higher ratio of massive to dwarf galaxies.

(ii) Depending on the field in question (i.e. COSMOS or XMM-LSS), the values of M^* and α_1 become more positive by 0.3–0.4 and 0.1–0.2 dex, respectively, in galaxies that are nearest to LSS (i.e. in the lower tercile in their distances to both nodes and filaments) compared to those that are furthest from LSS (i.e. in the upper tercile in their distances to both nodes and filaments). In other words, the knee of the mass function moves to larger stellar masses and the slope in the dwarf regime becomes flatter as we move closer to LSS.

(iii) While there is reasonable agreement between the predicted stellar mass functions of the TNG50, FIREBOX, and NewHorizon simulations in the stellar mass range $M_\star < 10^8\,M_\odot$, the disagreement between the different theoretical mass functions can be as large as $\sim$0.5 dex at some points within the dwarf regime (e.g. just below the knee). These discrepancies are likely to be driven by differences in the implementation of baryonic physics in the different simulations.

(iv) A comparison between the observed and theoretical stellar mass functions suggests that there is no generic missing dwarfs problem in ΛCDM, when all environments – and not just the regions around Milky Way like galaxies – are considered. This is a broader, generalized version of recent results which have shown that there is no missing satellites problem in ΛCDM when satellite populations around massive galaxies like the Milky Way are considered.

A0070402192 by Genci, and as a 'Grand Challenge' project granted by GENCI on the AMD Rome extension of the Joliot Curie supercomputer at TGCC. The simulation was also supported by KISTI under KSC-2017-G2-0003, KSC-2020-CRE-0055, and KSC-2020-CRE-0280 runs. The large data transfer was supported by KREONET, which is managed and operated by KISTI.

The Hyper Suprime-Cam (HSC) collaboration includes the astronomical communities of Japan and Taiwan, and Princeton University. The HSC instrumentation and software were developed by the National Astronomical Observatory of Japan (NAOJ), the Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), the University of Tokyo, the High Energy Accelerator Research Organization (KEK), the Academia Sinica Institute for Astronomy and Astrophysics in Taiwan (ASIAA), and Princeton University. Funding was contributed by the FIRST program from the Japanese Cabinet Office, the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Japan Society for the Promotion of Science (JSPS), Japan Science and Technology Agency (JST), the Toray Science Foundation, NAOJ, Kavli IPMU, KEK, ASIAA, and Princeton University. This paper makes use of software developed for Vera C. Rubin Observatory. We thank the Rubin Observatory for making their code available as free software at http://pipelines.lsst.io/.

This paper is based on data collected at the Subaru Telescope and retrieved from the HSC data archive system, which is operated by the Subaru Telescope and Astronomy Data Center (ADC) at NAOJ. Data analysis was in part carried out with the cooperation of Center for Computational Astrophysics (CfCA), NAOJ. We are honored and grateful for the opportunity of observing the Universe from Maunakea, which has the cultural, historical and natural significance in Hawaii. This paper used data that is based on observations collected at the European Southern Observatory under ESO programme ID 179.A-2005 and on data products produced by CALET and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium.

DATA AVAILABILITY

The data used in this study are taken from J. R. Weaver et al. (2022), G. Desprez et al. (2023), and V. Picouet et al. (2023). The density maps were created using the DISPERSE algorithm which is described in T. Sousbie (2011).

REFERENCES

Adams N. J., Bowler R. A. A., Jarvis M. J., Häußler B., Lagos C. D. P., 2021, MNRAS, 506, 4933
Aguerri J. A. L., Iglesias-Páramo J., Vílchez J. M., Muñoz-Tuñón C., Sánchez-Janssen R., 2005, AJ, 130, 475
Aihara H. et al., 2019, PASJ, 71, 114
Alam S. et al., 2015, ApJS, 219, 12
Arnouts S. et al., 2002, MNRAS, 329, 355
Ashby M. L. N. et al., 2013, ApJ, 769, 80
Ashby M. L. N. et al., 2015, ApJS, 218, 33
Ashby M. L. N. et al., 2018, ApJS, 237, 39
Atek H. et al., 2015, ApJ, 814, 69
Baldry I. K. et al., 2012, MNRAS, 421, 621
Battaglia G., Nipoti C., 2022, Nat. Astron., 6, 659
Bennet P., Sand D. J., Crnojević D., Spekkens K., Karunakaran A., Zaritsky D., Mutlu-Pakdil B., 2020, ApJ, 893, L9
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bichang'a B., Kaviraj S., Lazar I., Jackson R. A., Das S., Smith D. J. B., Watkins A. E., Martin G., 2024, MNRAS, 532, 613

Bielby R. et al., 2012, A&A, 545, A23
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Borch A. et al., 2006, A&A, 453, 869
Boselli A., Boissier S., Cortese L., Gavazzi G., 2008, ApJ, 674, 742
Boselli A., Fossati M., Sun M., 2022, A&AR, 30, 3
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40
Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343
Choque-Challapa N., Aguerri J. A. L., Mancera Piña P. E., Peletier R., Venhola A., Verheijen M., 2021, MNRAS, 507, 6045
Cuillandre J. C. et al., 2025, A&A, 697, A6
Dashyan G., Silk J., Mamon G. A., Dubois Y., Hartwig T., 2018, MNRAS, 473, 5698
Davis F. et al., 2022, MNRAS, 511, 4109
de los Reyes M. A. C. et al., 2025, ApJ, 989, 91
Desprez G. et al., 2023, A&A, 670, A82
Dong S., Lin D. N. C., Murray S. D., 2003, ApJ, 596, 930
Drinkwater M. J., Gregg M. D., Colless M., 2001, ApJ, 548, L139
Driver S. P. et al., 2022, MNRAS, 513, 439
Dubois Y. et al., 2014, MNRAS, 444, 1453
Dubois Y. et al., 2021, A&A, 651, A109
Emerick A., Bryan G. L., Mac Low M.-M., 2019, MNRAS, 482, 1304
Etherington J. et al., 2017, MNRAS, 466, 228
Fattahi A., Navarro J. F., Frenk C. S., Oman K. A., Sawala T., Schaller M., 2018, MNRAS, 476, 3816
Feldmann R. et al., 2023, MNRAS, 522, 3831
Finoguenov A. et al., 2007, ApJS, 172, 182
Fontana A. et al., 2004, A&A, 424, 23
Forbes J. C., Krumholz M. R., Goldbaum N. J., Dekel A., 2016, Nature, 535, 523
Fouqué P., Solanes J. M., Sanchis T., Balkowski C., 2001, A&A, 375, 770
Gavazzi R., Adami C., Durret F., Cuillandre J. C., Ilbert O., Mazure A., Pelló R., Ulmer M. P., 2009, A&A, 498, L33
Geen S., Slyz A., Devriendt J., 2013, MNRAS, 429, 633
Geha M. et al., 2024, ApJ, 976, 118
George M. R. et al., 2011, ApJ, 742, 125
Gozaliasl G. et al., 2014, A&A, 566, A140
Gozaliasl G. et al., 2019, MNRAS, 483, 3545
Hasan F. et al., 2023, ApJ, 950, 114
Hazenfratz R., Barai P., Lanfranchi G. A., Caproni A., 2024, ApJ, 969, 65
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hopkins P. F., 2015, MNRAS, 450, 53
Hsieh B.-C., Wang W.-H., Hsieh C.-C., Lin L., Yan H., Lim J., Ho P. T. P., 2012, ApJS, 203, 23
Hu W., Barkana R., Gruzinov A., 2000, Phys. Rev. Lett., 85, 1158
Hubble E. P., 1936, Realm of the Nebulae. Yale Univ. Press, New Haven
Ilbert O. et al., 2006, A&A, 457, 841
Ivezić Ž. et al., 2019, ApJ, 873, 111
Jackson R. A. et al., 2021a, MNRAS, 502, 1785
Jackson R. A. et al., 2021b, MNRAS, 502, 4262
Jeon S. et al., 2025, ApJ, 988, 136
Jespersen C. K., Steinhardt C. L., Somerville R. S., Lovell C. C., 2025, ApJ, 982, 23
Jung M. et al., 2024, ApJ, 964, 123
Kado-Fong E. et al. ., 2025, preprint (arXiv:2509.20444)
Kaiser N., 1984, ApJ, 284, L9
Kaviraj S. et al., 2017, MNRAS, 467, 4739
Kaviraj S., Lazar I., Watkins A. E., Laigle C., Martin G., Jackson R. A., 2025, MNRAS, 538, 153
Kaviraj S., Martin G., Silk J., 2019, MNRAS, 489, L12
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Kimm T., Cen R., 2014, ApJ, 788, 121
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Koch A., Burkert A., Rich R. M., Collins M. L. M., Black C. S., Hilker M., Benson A. J., 2012, ApJ, 755, L13
Kormendy J., 1985, ApJ, 295, 73
Koudmani S., Henden N. A., Sijacki D., 2021, MNRAS, 503, 3568
Koudmani S., Sijacki D., Smith M. C., 2022, MNRAS, 516, 2112
Kraljic K. et al., 2018, MNRAS, 474, 547

Kraljic K. et al., 2019, MNRAS, 483, 3227
Laigle C. et al., 2018, MNRAS, 474, 5437
Laigle C. et al., 2019, MNRAS, 486, 5104
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lazar I., Kaviraj S., Martin G., Laigle C., Watkins A., Jackson R. A., 2023, MNRAS, 520, 2109
Lazar I., Kaviraj S., Watkins A. E., Martin G., Bichang'a B., Jackson R. A., 2024a, MNRAS, 529, 499
Lazar I., Kaviraj S., Watkins A. E., Martin G., Bichang'a B., Jackson R. A., 2024b, MNRAS, 533, 3771
Leauthaud A. et al., 2007, ApJS, 172, 219
Lovell M. R. et al., 2012, MNRAS, 420, 2318
Malavasi N., Pozzetti L., Cucciati O., Bardelli S., Ilbert O., Cimatti A., 2017, MNRAS, 470, 1274
Mao Y. et al., 2024, ApJ, 976, 117
Marinacci F. et al., 2018, MNRAS, 480, 5113
Marsh D. J. E., 2016, preprint (arXiv:1605.05973)
Martel H., Barai P., Brito W., 2012, ApJ, 757, 48
Martin G. et al., 2025, MNRAS, 541, 1831
Martin G., Pearce F. R., Hatch N. A., Contreras-Santos A., Knebe A., Cui W., 2024, MNRAS, 535, 2375
Mayer L., Mastropietro C., Wadsley J., Stadel J., Moore B., 2006, MNRAS, 369, 1021
McCracken H. J. et al., 2012, A&A, 544, A156
Mezcua M., Domínguez Sánchez H., 2024, MNRAS, 528, 5252
Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19
Mori M., Burkert A., 2000, ApJ, 538, 559
Mortlock A. et al., 2015, MNRAS, 447, 2
Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, ApJ, 731, 113
Moustakas J. et al., 2013, ApJ, 767, 50
Musso M., Cadiou C., Pichon C., Codis S., Kraljic K., Dubois Y., 2018, MNRAS, 476, 4877
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Nashimoto M., Tanaka M., Chiba M., Hayashi K., Komiyama Y., Okamoto T., 2022, ApJ, 936, 38
Nelson D. et al. 2018, MNRAS, 475, 624
Nelson D. et al., 2019a, MNRAS, 490, 3234
Nelson D. et al., 2019b, Computational Astrophysics and Cosmology, 6, 2
Obreschkow D., Murray S. G., Robotham A. S. G., Westmeier T., 2018, MNRAS, 474, 5500
Picouet V. et al., 2023, A&A, 675, A164
Pillepich A. et al., 2019, MNRAS, 490, 3196
Pillepich A. et al., 2018, MNRAS, 475, 648
Polisensky E., Ricotti M., 2011, Phys. Rev. D, 83, 043506
Poulain M. et al., 2021, MNRAS, 506, 5494
Pozzetti L. et al., 2007, A&A, 474, 443
Reaves G., 1956, AJ, 61, 69
Salvadori S., Skúladóttir Á., Tolstoy E., 2015, MNRAS, 454, 1320
Santos-Santos I. M. E., Sales L. V., Fattahi A., Navarro J. F., 2022, MNRAS, 515, 3685
Sawala T. et al., 2016, MNRAS, 457, 1931
Sawicki M. et al., 2019, MNRAS, 489, 5202
Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
Schechter P., 1976, ApJ, 203, 297
Scoville N. et al., 2007, ApJS, 172, 1
Shapley H., 1938, Harv. Coll. Obs. Bull., 908, 1
Silk J., 2017, ApJ, 839, L13
Smercina A., Bell E. F., Price P. A., D'Souza R., Slater C. T., Bailin J., Monachesi A., Nidever D., 2018, ApJ, 863, 152
Song H. et al., 2021, MNRAS, 501, 4635
Sousbie T., 2011, MNRAS, 414, 350
Springel V., 2010, MNRAS, 401, 791
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V. et al., 2018, MNRAS, 475, 676
Steinhardt C. L. et al., 2014, ApJ, 791, L25
Steyrleithner P., Hensler G., Boselli A., 2020, MNRAS, 494, 1114

Tanaka M., Chiba M., Hayashi K., Komiyama Y., Okamoto T., Cooper A. P., Okamoto S., Spitler L., 2018, ApJ, 865, 125
Taniguchi Y. et al., 2007, ApJS, 172, 9
Taniguchi Y. et al., 2015, PASJ, 67, 104
Teyssier R., 2002, A&A, 385, 337
Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371
Tomczak A. R. et al., 2014, ApJ, 783, 85
Trujillo I. et al., 2021, A&A, 654, A40
van den Bergh S., 1959, Publ. David Dunlap Obs., 2, 147
Watkins A. E., Salo H., Kaviraj S., Collins C. A., Knapen J. H., Venhola A., Román J., 2023, MNRAS, 521, 2012
Weaver J. R. et al., 2022, ApJS, 258, 11
Weigel A. K., Schawinski K., Bruderer C., 2016, MNRAS, 459, 2150
Zamojski M. A. et al., 2007, ApJS, 172, 468
Zarattini S., Aguerri J. A. L., 2025, A&A, 698, A196

APPENDIX A: MASS FUNCTION COMPARISON TAKING INTO ACCOUNT UNDERESTIMATES OF SED-FITTED STELLAR MASSES

As noted in Section 6, the stellar masses derived from SED fitting underestimate the true stellar masses of galaxies, with the average underestimate in the dwarf regime likely to be ∼0.2 dex. In Fig. A1, we show a version of Fig. 6 where the original observed mass functions (shown in grey) have been shifted by 0.2 dex in the dwarf regime (shown in red and blue). While Fig. 6 already indicates consistency between data and theory, Fig. A1 illustrates the point that raising the observed mass function by ∼0.2 dex in the dwarf regime would result in good agreement between the observations and the ensemble of theoretical simulations explored in this study.
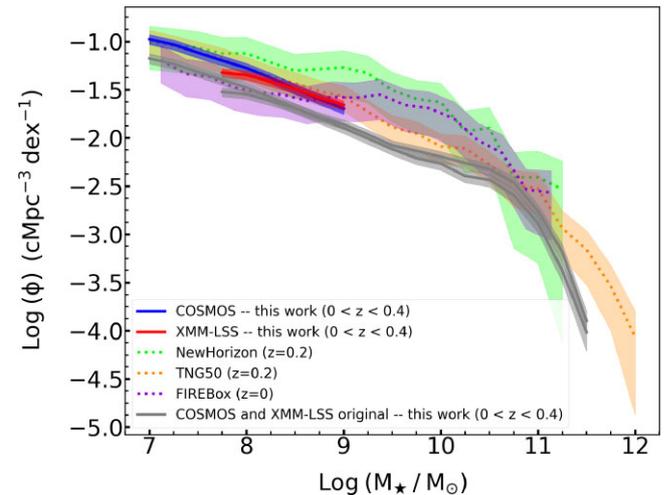

Figure A1. The observed galaxy stellar mass function in the nearby Universe in the COSMOS and XMM-LSS fields compared to theoretical mass functions from three high resolution cosmological hydrodynamical simulations (NEWHORIZON, TNG50, and FIREBOX). The original observed mass functions in the COSMOS and XMM-LSS fields, from Fig. 6, are shown in grey. The observed mass functions in red and blue show the original mass functions, in the dwarf regime only, shifted upwards by 0.2 dex, the approximate offset induced by the fact that SED-fitted stellar masses are typically underestimated in this regime (see text in Section 6 and Appendix A for more details).

Table B1. Best-fitting parameters and their associated uncertainties from double Schechter fits to the mass functions in the COSMOS and XMM-LSS fields in different environments. Galaxies in the lower tercile in the distance from LSS are defined as those that reside in the lowest tercile (0th–33rd percentile values) of both the distances to nodes and the distances to filaments. Similarly, galaxies in the middle and upper terciles in the distance from LSS are defined as those that reside in the middle (33rd–66th percentile values) and upper (66th–100th percentile values) terciles of both the distances to nodes and the distances to filaments respectively. Galaxies in the lower tercile in the distance from LSS are the ones nearest to LSS, while galaxies in the upper tercile in the distance from LSS are furthest away from LSS.

Field	Environment	$\mathrm{Log}_{10} M^{\star}$ ($M_{\odot}$)	α_1	$\Phi_1 \times 10^{-3}$ (Mpc^{-3} dex^{-1})	α_2	$\Phi_2 \times 10^{-3}$ (Mpc^{-3} dex^{-1})
COSMOS	Lower tercile in dist. from LSS	10.96 ± 0.06	-1.24 ± 0.03	2.37 ± 0.72	-0.68 ± 0.31	2.39 ± 0.83
	Middle tercile in dist. from LSS	10.59 ± 0.07	-1.34 ± 0.02	1.68 ± 0.22	1.01 ± 0.35	1.26 ± 0.25
	Upper tercile in dist. from LSS	10.53 ± 0.07	-1.40 ± 0.02	1.32 ± 0.18	1.41 ± 0.40	0.41 ± 0.13
XMM-LSS	Lower tercile in dist. from LSS	10.95 ± 0.08	-1.22 ± 0.07	2.01 ± 1.01	-0.56 ± 0.36	3.51 ± 1.24
	Middle tercile in dist. from LSS	10.71 ± 0.06	-1.28 ± 0.03	1.93 ± 0.31	0.20 ± 0.39	1.54 ± 0.34
	Upper tercile in dist. from LSS	10.62 ± 0.08	-1.34 ± 0.02	1.55 ± 0.22	0.66 ± 0.37	0.91 ± 0.18

APPENDIX B: DOUBLE SCHECHTER FITS AND ASSOCIATED PARAMETERS

Table B1 presents the parameters derived from the double Schechter fits to the stellar mass functions discussed in Section 5.1.

APPENDIX C: VARIATION OF THE MASS FUNCTION WITH DISTANCE FROM NODES AND FILAMENTS

Fig. C1 shows the observed stellar mass function for different terciles in the distances to filaments (top panel) and nodes (bottom panel) separately. In other words, these panels show versions of Fig. 3 where the distances to nodes and filaments are considered separately. Fig. C1 demonstrates that the differences between galaxies in the different distance terciles are much more significant when considering the distances to filaments than the distances to nodes. Thus, the overall trends seen in Fig. 3 are driven predominantly by the distances of galaxies from filaments rather than their distances to nodes.
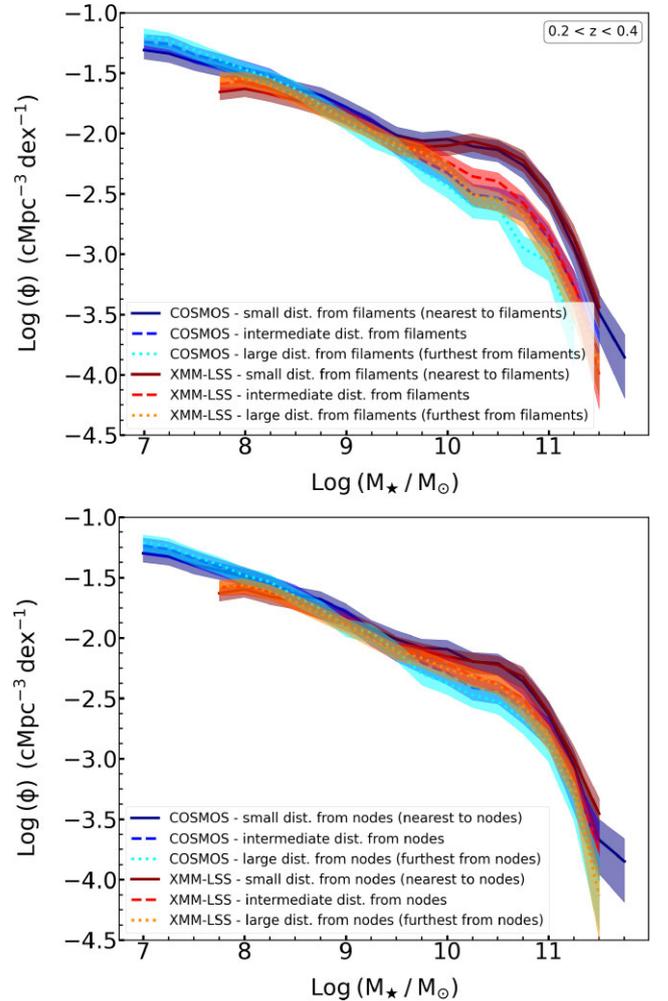

Figure C1. The galaxy stellar mass function for different terciles in the distances to filaments (top) and nodes (bottom) separately. These panels show versions of Fig. 3 where the distances to nodes and filaments are considered separately.
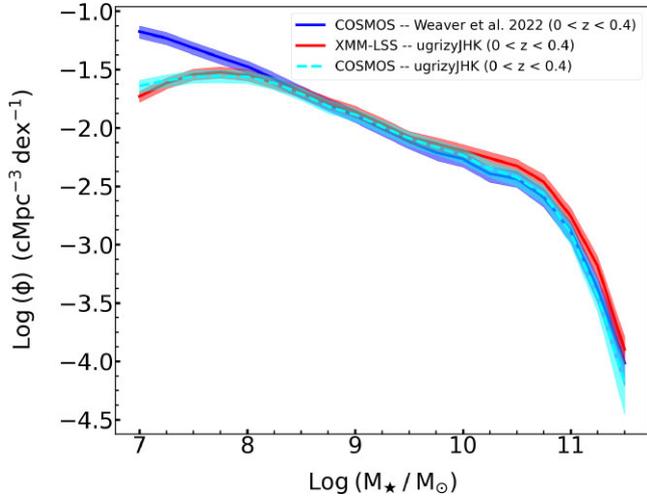
Figure D1. The observed galaxy stellar mass functions in the COSMOS and XMM-LSS fields for $z < 0.4$ derived using different filter combinations. The blue curve uses stellar masses in COSMOS that are derived using broad and medium band photometry from the COSMOS2020 catalogue. The red curve uses stellar masses in XMM-LSS that are derived using only broad-band $ugrizyJHK$ filters. The dashed cyan curve uses stellar masses in COSMOS that are derived using only broad-band $ugrizyJHK$ filters like in XMM-LSS.

APPENDIX D: SED-FITTED STELLAR MASSES WITH AND WITHOUT MEDIUM-BAND FILTERS

In this section, we further explain our reasoning for exploring the COSMOS data set down to stellar masses of $10^7\,M_\odot$ but restricting the XMM-LSS data set to stellar masses above $10^8\,M_\odot$. The mass function in COSMOS, derived using both broad-band and medium band filters, is shown using the blue curve. The mass function in XMM-LSS, derived using the nine available broad-band filters spanning the u-band to rest-frame near-infrared ($ugrizyJHK$), is shown using the red curve. The XMM-LSS mass function shows good consistency with the COSMOS mass function down to $M_\star \sim 10^8\,M_\odot$, after which it shows a downturn which is not shared by its COSMOS counterpart.

The dashed cyan curve shows the mass function in COSMOS produced by an identical SED fitting procedure to that used to create the red (XMM-LSS) curve, i.e. using $ugrizyJHK$ only with the medium band filters omitted. Interestingly, this version of the COSMOS mass function also shows a downturn which closely follows that in XMM-LSS (which, recall, does not have any medium band filters). Thus, we speculate that the inconsistency, at $M_\star < 10^8\,M_\odot$, between the blue curve (which employs SED fitting on broad and medium band filters) and its red and cyan counterparts (which employ SED fitting on broad-band filters only) could be

driven by the unavailability of medium band filters. This leads to a poorer sampling of the SED that seems to affect very low mass galaxies ($M_\star < 10^8\,M_\odot$). Due to this, as a precaution, we consider the XMM-LSS data set only down to $M_\star = 10^8\,M_\odot$ where it shows good consistency with the COSMOS mass function that is derived using both broad and medium band filters.

APPENDIX E: EDDINGTON BIAS EFFECTS ON THE GALAXY STELLAR MASS FUNCTION

In this section, we show that the systematic Eddington bias has a negligible influence on the shape of the galaxy stellar mass functions derived in this work. In particular, the COSMOS (blue curve) and XMM-LSS (red curve) stellar mass functions are very similar to the mean of the galaxy stellar mass functions derived using the perturbed stellar masses as described in Section 4.2. As shown in Fig. E1, they are typically within ~0.01 dex until $M_\star \sim 10^{11}\,M_\odot$, with a maximum deviation of ~0.08 dex at the very massive end of the XMM-LSS mass function ($M_\star \sim 10^{11.5}\,M_\odot$).
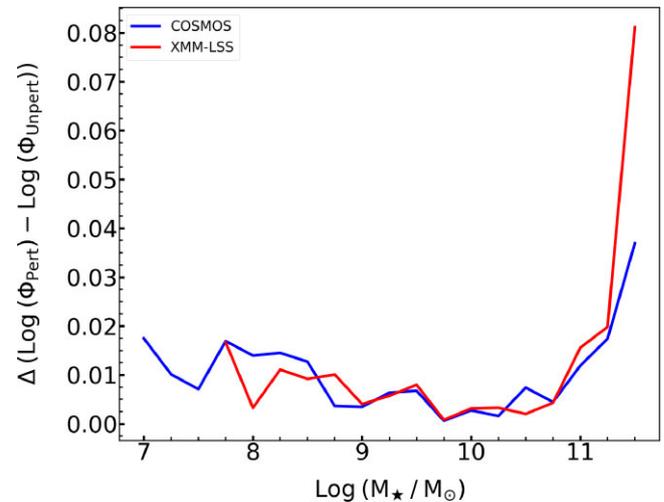

Figure E1. The blue and red curves show the difference between the unperturbed galaxy stellar mass functions in the COSMOS and XMM-LSS fields, respectively, for $z < 0.4$ (same as in Fig. 6) and the mean of the galaxy stellar mass function derived using the perturbed stellar masses via the Monte Carlo approach described in Section 4.2. The perturbed and unperturbed mass functions are denoted using Φ_{Pert} and Φ_{Unpert}, respectively.

This paper has been typeset from a TEX/LATEX file prepared by the author.