



UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXV CICLO DEL DOTTORATO DI RICERCA IN
FISICA

**INTERNAL DYNAMICS OF GALAXY
CLUSTERS FROM COSMOLOGICAL
HYDRODYNAMICAL SIMULATIONS**

Settore scientifico-disciplinare: FIS/05

DOTTORANDA:
ILARIA MARINI

COORDINATORE:
PROF. FRANCESCO LONGO

SUPERVISORE DI TESI:
PROF. STEFANO BORGANI

Co-SUPERVISORE DI TESI:
PROF. ALEXANDRO SARO

ANNO ACCADEMICO 2021/2022

“George,” said Fred, “I think we’ve outgrown full-time education.”
“Yeah, I’ve been feeling that way myself,” said George lightly.

— J. K. Rowling, *Harry Potter and the Order of the Phoenix*

ABSTRACT

Galaxy clusters are the most massive systems in the Universe. They are usually located at the nodes of the cosmic web from which they continuously accrete matter. In this work, by combining cosmological simulations and local Universe observations, we examined several properties of the different collisionless tracers of the internal dynamics of galaxy clusters – namely Dark Matter (DM), stars, and galaxies – to gain insights into the main processes operating in structure formation and evolution. We base our analysis on the DIANOGA zoom-in simulation set which is composed of 29 Lagrangian regions at different levels of resolution and under varying physical conditions (full hydrodynamical and/or N-body simulations).

Recent measurements (Biviano et al., 2013, 2016; Capasso et al., 2019) of the pseudo-entropy ($\sigma^2/\rho^{2/3}$, where σ is the velocity dispersion and ρ the density of the collisionless tracer) allowed us to study its role in the evolution of clusters as dynamical attractor (e.g. Taylor and Navarro, 2001; Dehnen and McLaughlin, 2005). Its fingerprint is the universal radial profile described by a simple power-law. We find good agreement in both normalisation and slope between observations and simulations. A significant tension is present with the galaxy member population, we discuss in detail the probable reasons behind this finding.

A large body of spectroscopic measurements (Loubser et al., 2018, 2020; Sohn et al., 2020, 2022) were able to provide a large statistical sample to study the dynamics of the Brightest Cluster Galaxy (BCG). We compare scaling relations between the BCG and cluster velocity dispersions and corresponding masses: we find in general a good agreement with observational results for the former and significant tension in the latter. We analyse the key features of the velocity dispersion profiles, as traced by stars, DM, and galaxies (Sartoris et al., 2016) and they are in excellent agreement with simulations. We also quantify the assumed impact of the IntraCluster Light (ICL) in these measurements.

Furthermore, given the existing dynamical distinction between BCG and ICL, we developed a Machine Learning (ML) method based on a supervised Random Forest to classify stars in simulated galaxy clusters in these two classes. We employ matched stellar catalogues (built from a modified version of Subfind, Dolag et al. 2010) to train and test the classifier. The input features are cluster mass, normalised particle clustercentric distance, and rest-frame velocity. The model is found to correctly identify most of the stars, while the larger errors are exhibited at the BCG outskirts, where the differences between the physical properties of the two components are less obvious. We find that our classifier provides consistent results in simulations for clusters at $z < 1$, using different numerical resolutions and implementations of the feedback. The last part of the project has focused on creating a ML framework to bridge the observational analysis

with predictions from simulations. Measuring the ICL in observations is a difficult task which is often solved by fitting functional profiles to the BCG+ICL light profile, but often providing significantly different results. We developed a method based on convolutional neural networks to identify the ICL distribution in mock images of galaxy clusters, according to the dynamical classification we routinely perform in simulations. We construct several sets of mock images based on different observables (i.e., magnitudes, line-of-sight velocity, and velocity dispersion) that can be employed as input by the network to predict the ICL distribution in such images. This project has highlighted the dependence of the ICL build-up on the numerical resolution of the simulations, a problem which requires further investigations.

CONTENTS

List of Figures	xiii
List of Tables	xxxi
Acronyms	xxxiii
1 INTRODUCTION	1
I SCIENTIFIC BACKGROUND	
2 CLUSTERS OF GALAXIES	9
2.1 Galaxy clusters in the cosmological environment	9
2.1.1 How do they form?	10
2.1.2 How do they evolve?	11
2.2 Properties of the galaxy population	12
2.2.1 Brightest cluster galaxy	13
2.2.2 The intracluster light	16
2.2.3 Cluster galaxies	18
II METHODS	
3 NUMERICAL COSMOLOGY	35
3.1 Cosmological simulations	35
3.2 N-body solvers	38
3.2.1 PM code	38
3.2.2 The Tree algorithm	39
3.2.3 Tree-PM code	41
3.3 Hydrodynamical simulations	42
3.3.1 Smoothed Particle Hydrodynamics	43
3.3.2 Mesh codes	44
3.4 The GADGET-3 code	45
3.4.1 Radiative cooling and photoionisation heating	45
3.4.2 Star formation	46
3.4.3 Chemical enrichment	47
3.4.4 AGN feedback	48
3.5 Structures identification	50
3.5.1 The Friends of Friends group finder	50

3.5.2	The Subfind algorithm	51
3.5.3	ICL-Subfind	53
3.6	The Smac visualisation tool	55
3.6.1	Assigning optical/IR luminosities to star particles . .	58
3.6.2	Photometric maps with Smac	59
4	MACHINE LEARNING	61
4.1	Basic concepts	61
4.1.1	What is a machine learning model?	61
4.1.2	Machine learning versus deep learning	65
4.2	Random Forest	67
4.2.1	Feature importance	68
4.2.2	Confusion matrix and performance scores	69
4.3	Convolutional Neural Networks	71
4.3.1	Regularisation and optimisation	73
4.3.2	Learning rate range test	75
4.3.3	The U-Net architecture	75
III RESULTS		
5	ON THE PHASE-SPACE STRUCTURE OF SIMULATED GALAXY CLUSTERS	81
5.1	Scientific context	81
5.2	Simulations	83
5.2.1	Dynamical state	83
5.3	Self-similarity of pseudo-entropy profiles	84
5.4	Results	86
5.4.1	Pseudo-entropy profiles traced by DM	86
5.4.2	Pseudo-entropy profiles traced by substructures . . .	95
5.4.3	Pseudo-entropy profiles traced by stars	103
5.4.4	Comparison with observational results	106
5.5	Conclusions	112
6	VELOCITY DISPERSION OF THE BRIGHTEST CLUSTER GALAXIES	117
6.1	Scientific context	117
6.2	Simulations	119
6.3	Observables in simulations	119
6.3.1	Dynamics of the BCGs	119

6.3.2	Dynamics of the clusters	121
6.4	Connections between BCG and cluster	122
6.4.1	Scaling relations	122
6.4.2	The $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation at $z = 0$	125
6.4.3	Effect of excluding the ICL	128
6.4.4	Redshift evolution of the $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation	129
6.5	Velocity dispersion profiles	130
6.5.1	Dynamical properties of cluster core regions	133
6.5.2	Gradients of the stellar velocity dispersion profile	135
6.6	Conclusions	138
7	IDENTIFYING INTRACLUSTER LIGHT WITH MACHINE LEARNING	141
7.1	Scientific context	141
7.2	Simulations	143
7.3	Random Forest	144
7.3.1	Data set and training phase	144
7.3.2	Input features	145
7.3.3	Classification performance	148
7.4	Results	148
7.4.1	Testing on a single simulated cluster	149
7.4.2	Testing on a simulated cluster population	152
7.4.3	Testing the robustness of the classifier	156
7.4.4	Phase-space structure	162
7.4.5	Hints for the identification of shell galaxies	163
7.5	Conclusions	165
8	DETECTING INTRACLUSTER LIGHT IN MOCK IMAGES OF GALAXY CLUSTERS	171
8.1	Scientific context	172
8.2	Simulations	173
8.2.1	Input and output data	173
8.3	Results	177
8.3.1	Optimisation	177
8.3.2	Training phase	178
8.3.3	Stellar mock images	179
8.3.4	Quantifying the accuracy	185
8.3.5	Increasing the numerical resolution	187

8.4	Conclusions and future developments	189
IV FINAL REMARKS		
9	CONCLUSIONS	195
9.1	This work	195
9.2	Future perspectives	198
A	BACKGROUND ON COSMIC STRUCTURE FORMATION	203
A.1	Statistical properties	204
A.2	Linear evolution of density perturbations	205
A.3	Gravitational instability	206
A.4	Spherical collapse model	208
A.5	Basics on dynamics of a collisionless self-gravitating system	210
A.5.1	The Vlasov equation	211
A.5.2	Jeans equation	212
A.5.3	Virial Theorem	214
	BIBLIOGRAPHY	221

LIST OF FIGURES

Figure 1.1	A slice through the Millennium XXL simulation showing the simulated cosmic web. The bright yellow regions are the high density clusters that form at the intersection of filaments. The simulation follows the non-linear growth of DM structures within a cubical region 4.1096 Gpc on a side. The particle mass is $8.456 \times 10^9 M_{\odot}$. Credits: http://galformod.mpa-garching.mpg.de/mxxlbrowser/	2
Figure 2.1	The Bullet cluster. Left panel: we show the optical image taken with the Magellan telescope. Right panel: Chandra X-ray measurements of the hot plasma in the cluster halo. The green contours denote the gravitational potential modelled with weak lensing measurements. Credits: Clowe et al. (2006).	11
Figure 2.2	ICL in the Virgo cluster core. Black areas are the masked galaxies in the cluster while the strength of the white is related to the magnitude of the stars. Credits: Mihos et al. (2005)	15
Figure 2.3	Decomposing the ICL in A1668 (black dots) based on surface brightness methods. The coral-shaded region is the ICL found. Method b (left panel) defines all light fainter than a surface brightness threshold (horizontal line at $27 g' \text{ mag arcsec}^{-2}$) as ICL. Method c (central panel) takes de Vacouleurs profile (solid black) to fit the BCG light and assigns the excess to ICL. Method d (right panel) is the double Sérsic decomposition (two black lines) where the outer profile is the ICL. Credits: Kluge et al. (2021)	15
Figure 2.4	The velocity distribution of the stars in the main halo in a simulated cluster in the DIANOGA set (Bassini et al., 2020). The shaded histogram represents the data points which are fitted by the double Maxwellian (orange line). The singular components are given by the green (BCG) and red lines (ICL).	16

- Figure 2.5 Cluster surface mass density profiles for A1689 (green triangles) and CL0024 + 1654 (red squares) reconstructed from strong lensing (HST/ACS: $r < 200$ kpc) and weak lensing measurements (Subaru: $r > 200$ kpc). Solid curves are the best-fit NFW models, with a continuously steepening density profile, for A1689 (black) and CL0024 + 1654 (grey). Credits: Umetsu and Broadhurst (2008). 19
- Figure 2.6 Study of the phase-space density profiles $Q(r) \equiv \rho/\sigma^3$ (left panels) and $Q_r(r) \equiv \rho/\sigma_r^3$ (right panels) as a function of radius, within 1σ confidence regions (shaded area) for (from top to bottom) all, passive and star forming galaxy members of the cluster MACS J1206.2 – 0847 (redshift $z = 0.44$): The mass distribution is well fitted by a NFW density profile with mass $M_{200} = (1.4 \pm 0.2) \times 10^{15} M_\odot$ and concentration $c_{200} = 6 \pm 1$. Data were mostly collected by ESO VLT, NASA Hubble Space Telescope, and NASJ Subaru Telescope. Solid lines are the best-fit results from Dehnen and McLaughlin (2005). Credits: Biviano et al. (2013). 21
- Figure 2.7 Log-log plane of the stellar mass versus stellar specific angular momentum in a galaxy sample. Left panel: data points from the Magneticum simulations at redshift $z = 0.1$; right panel: the observed sample from Fall and Romanowsky (2013). The coloured solid lines have a slope around $0.5 - 0.6$ and represent the kinematical distinction between disks and bulges. The markers label the morphological distinction between elliptical (early types) and spirals. Credits: Teklu et al. (2015) 24
- Figure 2.8 An example of two galaxies from the DIANOGA simulation set divided according to their kinematical properties: fast (top) and slow (bottom) rotator. The histograms report distributions of the star particle circularities ϵ (see text) for the two galaxies. A fast rotator has this distribution peaking at $\epsilon \sim 1$ (vertical red line), while for slow rotators the distribution peaks at $\epsilon \sim 0$ (vertical blue line). 25

Figure 2.9	An example of a disk galaxy from the DIANOGA simulation set taken from two different lines of sight. The galaxy is chosen among the fast rotators, thus we expect it to have properties similar to spiral/disk-like galaxies. The simulated galaxy is coloured in RGB using the HST's <i>gri</i> filters. Left panel: the galaxy is taken face-on. The bulge at the centre is the bluest region in the cluster. We can observe a disk-like structure at larger radii. Right panel: the galaxy is taken edge-on. We can still observe a central bulge and a thin disk surrounding it.	27
Figure 2.10	UVJ colour-colour diagram in the $[0, 4]$ redshift interval for massive galaxies in the COSMOS/UltraVISTA field. The bimodality in the galaxy population is visible up to $z = 2$. The solid lines mark the cuts applied to separate star forming (top-left) from passive (bottom-right). Credits: Muzzin et al. (2013).	28
Figure 2.11	From left to right, the panels report results from different redshift intervals: $[0.2, 0.5]$, $[0.5, 1.0]$, $[1.0, 1.5]$, and $[1.5, 2.0]$. Top panels: stellar mass functions in the Muzzin et al. (2013) study for the quiescent (in red) and star forming (in blue) populations. Overplotted in the $0.2 < z < 0.5$ bin are the stellar mass functions from Cole et al. (2001), Bell et al. (2003) and Baldry et al. (2012). In the remaining redshift bins, the dotted curve is the total stellar mass function from UltraVISTA in the $0.2 < z < 0.5$ bin. Bottom panels: quenching fraction as a function of the log stellar mass. Credits: Muzzin et al. (2013).	30
Figure 2.12	Passive fraction of the cluster galaxies in three different redshift bins. The black dots represent observational data from Wetzel et al. (2012) while the coloured curves mark the simulations. Credits: L-Galaxies (Henriques et al., 2013), Eagle (Furlong et al., 2015), Illustris and IllustrisTNG (Donnari et al., 2021), selected from a presentation's slides given in Favignana 2018 at the conference "Birth, life and fate of massive galaxies and their central beating heart".	30

Figure 3.1	A snapshot of a simulated galaxy cluster showing why they are defined as multi-component systems. Following from left to right, and top to bottom, we see represented the DM distribution, the gas density, the gas temperature, the star distribution, the gas entropy and the gas metallicity. The zoomed region has size $8 h^{-1}$ Mpc. Credits: http://www.astro.yale.edu/nagai/Research.html	37
Figure 3.2	Left: computation of the forces to a single target particle with direct summation over 100 particles included in the simulation: every line corresponds to a single computation. Middle: computation of the forces of the same particles utilising the Tree-code algorithm. Cells are shown as black squares with their centres indicated by solid squares; green lines denote the cell-particle interaction. Right: calculation of all the forces among the particles in the simulation, requiring 902 cell-particle and 306 particle-particle interactions, instead of 4950. Credits: Dehnen and Read (2011).	39
Figure 3.3	Space-filling curve Peano-Hilbert in two (lower) and three (top) dimensions. Credits: Springel (2005) . . .	40
Figure 3.4	Spatial decomposition of the force in a Tree-PM code. This has been computed by randomly placing a particle of unit mass in a periodic box, and measuring the forces obtained by the simulation code for a set of randomly placed test particles. Credits: Springel (2005)	42
Figure 3.5	Graphical representation of two FoF groups, according to different linking lengths. The shorter the linking length, the smaller the resulting group. Credit: cosmosim	51
Figure 3.6	Projected two-dimensional distribution of a simulated halo in which Subfind identifies different subhalos. The FoF halo is in the top-left panel. The main halo is on the top right; the substructures are on the bottom left. The bottom-right hosts the "fuzz", which is not gravitationally bound to any structure but resides in the same spatial location as the group. Credits: Springel et al. (2001).	52
Figure 3.7	Mass-weighted maps of the stellar components as identified by the ICL-Subfind for one of the clusters in our simulations. Left panel: BCG and ICL hosted in the main halo of the cluster. Central panel: stars associated with the ICL component. Right panel: stars bound to the BCG. The stellar component is dominated by the ICL diffuse component.	53

Figure 3.8	Observed properties in the stellar populations identified by the ICL-Subfind for the same cluster in Figure 3.7 . Left panel: main halo including BCG+ICL (black), ICL (red), and BCG (blue) stellar density profiles. Right panel: velocity histograms of the main halo (black), ICL (red), and BCG (blue). We also report the single (dashed black) and double Maxwellian (dotted black) best-fit curves for the entire data set. Red and blue dashed lines are showing the individual Maxwellian distributions which are associated with the unbound ICL ($\sigma_{\text{ICL}}^* = 1002 \text{ km s}^{-1}$) and bound BCG ($\sigma_{\text{BCG}}^* = 484 \text{ km s}^{-1}$).	54
Figure 3.9	Different mock images (produced with Smac) of a galaxy cluster. We present the gas density (top-left), the X-ray emission (top-right) and the mass-weighted temperature of the gas (bottom). The cluster is selected at $z = 0$ from the DIANOGA set (25x resolution). The image is centred on the cluster's centre and it has size R_{200} of the cluster.	56
Figure 3.10	These maps (produced with Smac) are centred on the same cluster as in the previous image. We report the stellar metallicity distribution (left-top), the stellar age (right-top) and stellar luminosity (bottom). The brightness denotes the higher metallicity, younger stellar population and more luminous respectively. We can recognise structures such as massive and small galaxies within the virial radius we used to frame the image.	60
Figure 4.1	Biological neural cell (top) and artificial neuron (bottom) schemes. The analogy between the two is not random, as the operating principle of the artificial neuron is modelled after the biological one. Both neurons receive some information to process (dendrites/inputs) in order to provide an answer (terminal/output). Credits: https://www.mriquestions.com/deep-learning-dl.html	62
Figure 4.2	A deep NN composed of N hidden layers. Neurons are the fundamental unit in the hidden linear combination (layer) present in a NN architecture. Credits: https://www.mriquestions.com/deep-learning-dl.html	63

- Figure 4.3 Comparison of the performances in Goodfellow et al. (2016). The plot shows the number of parameters set in the model versus the model accuracy when assessed with the test set. This is showing that increasing the number of parameters in layers without increasing their depth is not nearly as effective at increasing test set performance. The legend indicates the depth of the network used to make each curve (3 for blue and green, 11 for red) and whether the curve represents variation in the size of the convolutional or fully connected layers. Credits: Goodfellow et al. (2016). 65
- Figure 4.4 Examples of different selections of learning rates during the stochastic gradient descent of the loss function. The learning rate controls how much we are adjusting the weights of our network with respect to the loss gradient and how much we can earn at each timestep. Left panel: the learning rate is too small and thus the learning happens at a very small pace (care must be taken in this case as the procedure might get stuck in a local minimum). Central panel: correct learning rate. Right panel: the learning rate is too large and the algorithm is possibly not able to find the global minimum. Credits: jeremyjordan.me/nn-learning-rate/ 67
- Figure 4.5 Graphical representation of a Random Forest architecture composed of 4 decision trees. Each tree has a hierarchical structure of leaves and nodes. Provided the dataset to each tree, they independently decide to which class the given example belongs. At the end of the computation, the final class is the average class coming from the ensemble. Credits: Habib et al. (2020). 68
- Figure 4.6 Graphical representation of the feature importance histograms to determine which input features are redundant in a model. This example is taken from the binary classification problem we outline in Chapter 7. 69
- Figure 4.7 Graphical representation of a confusion matrix for a binary classification. The columns yields the true labels, while the rows have predicted labels. A normalised matrix will describe the percentages of each one of the four categories (after assembling true and predicted labels of the classes). Credits: <https://towardsdatascience.com/> 70

Figure 4.8	Structure of a generic CNN with one channel. Each image patch undergoes feature extraction (composed of several layers of convolution and max pooling) before the final layer can provide the class label (output). The process can be broadly divided in two phases: feature learning (image is processed) and classification (the output is extracted). The initial image size is downsized during the processing. Credits: https://towardsdatascience.com/	71
Figure 4.9	Graphical representation of the convolution operation with a kernel size of 3×3 across an input tensor 5×5 . The kernel is applied as an element-wise product between kernel and input at each location and summed to obtain the output value in the feature map. Multiple kernels (i.e., different matrices) work as different feature extractors. Credits: Alzubaidi et al. (2021).	72
Figure 4.10	Loss function as a function of the learning rate. In the search for the optimal learning rate, the use of this type of graph allows one to search the parameter space thoroughly. During the process, the learning rate goes from a very small value to a very large value (i.e., from 10^{-7} to 100), causing the training loss to start with a plateau, descend to some minimum value, and eventually explode. Credits: https://blog.dataiku.com	76
Figure 4.11	U-Net architecture for one channel. The flow of information is denoted by the coloured arrows in the image. Each box corresponds to a multichannel feature map whose size is reported at the bottom left of each rectangle. The size of the kernels is reported on top of each box. Credits: Ronneberger et al. (2015).	77
Figure 5.1	From the top panel: density, velocity dispersion, and pseudo-entropy profiles as traced by the DM particles in the Hydro-10x at $z = 0$. The coloured lines report the profiles traced by each cluster while the black is the median. In the middle panel, the grey and pink lines represent the velocity dispersion profiles $\sigma_v(r)$ and $\widetilde{\sigma}_v(r)$ of each cluster respectively scaled and not-scaled according to Equation 33. The black lines trace the median profiles. Similarly, in the bottom panel, the grey and pink lines represent the scaled and not-scaled profiles of the pseudo-entropy $S(r)$ according to Equation 34, the black solid line illustrates the median trend while the dashed one is the best-fit.	87

- Figure 5.2 Comparison of the pseudo-entropy profiles traced by DM particles at redshift $z = 0$ to test the convergence among the distinct simulations. In the top panel, we compare the median pseudo-entropy profiles obtained from the clusters in the Hydro-1x (in ocher) and Hydro-10x (in red) runs. In the central panel, we show the median pseudo-entropy profiles obtained from the DM-10x (in green) and DM-100x (in blue) simulations. The bottom panel reports the median pseudo-entropy profiles from the simulated clusters in the Hydro-10x (in red) and DM-10x (in green). For all the curves we plot the 68th percentile region given by the cluster samples as a shaded area. 89
- Figure 5.3 Top panel: median of the pseudo-entropy profiles traced by DM particles at different redshifts (as indicated in the legend in the bottom panel) for the clusters in the Hydro-10x simulation. Bottom panel: median of the pseudo-entropy profiles traced by DM particles at different redshifts for the clusters in the DM-10x simulation. 92
- Figure 5.4 Scaling relation of the integrated pseudo-entropy (upper panel) and the virial velocity dispersion (lower panel), both computed over all the DM particles within r_{vir} , as a function of the virial mass of the clusters in the Hydro-1x simulation at $z = 0$. The dashed grey lines are the best-fit profiles for the scaling relations. The second most massive halo in the sample is marked with a black cross to signal an outlier in the distribution of the integrated pseudo-entropy profile, as pointed out in [Section 5.4.1.4](#). . . 93
- Figure 5.5 The evolution of the density maps in logarithmic scale traced by the DM particles in the second most massive cluster reported in the top panel of [Figure 5.4](#) which appears as an outlier from the $S_{\text{vir}} - M_{\text{vir}}$ scaling relation. The region is centred at the cluster centre and spans a square of 4 times the virial radius of the main halo. The white circle marks the virial radius. The red smaller circle marks the trajectory of the second most massive halo in the region and its growth given that the radius of the circle is equal to its virial radius (as provided by Subfind). The orbiting object at redshift $z \simeq 0.13$ crosses the virial radius to merge with the main halo by redshift $z = 0$ 94

- Figure 5.6 From top to bottom panel: number density, velocity dispersion, and pseudo-entropy profiles as traced by the substructures in the Hydro-10x simulation at $z = 0$. The grey lines show the single cluster profiles, whereas the black ones reproduce the median profiles. We also report with the dashed black lines the DM particles profiles (as seen in Figure 5.1) normalised to match the number density profile at $0.5 r_{\text{vir}}$ 97
- Figure 5.7 Comparison of profiles traced by substructures in DM-10x and Hydro-10x simulated clusters. Upper and central panels: the pseudo-entropy and number density profiles traced by the stack sample of substructures at $z = 0$. Solid curves are for the median profiles, while the shaded areas encompass the 16th and 84th percentiles. In the upper panel, the dashed line shows the best-fit power-law relation for the pseudo-entropy profile of substructures in the Hydro-10x simulations (plotted in Figure 5.6). Lower panel: The baryon fraction profile of substructures in the Hydro-10x simulations. Green and dot-dashed red curves are the median profiles of the baryon fraction when including stars and gas, and only stars, respectively. The horizontal dashed line marks the cosmic baryon fraction of the simulations ($\Omega_b/\Omega_M = 0.154$). 98
- Figure 5.8 Effect of mass segregation on the distribution of substructures in the Hydro-10x set of simulations. In each panel, we show with blue (orange) curves results for subhalo populations with masses larger (smaller) than the median subhalo mass found in each cluster. Shaded areas encompass the 16th – 84th percentile of the distribution of profiles. From top to bottom panels we show results for pseudo-entropy, velocity dispersion and number density profiles. 102
- Figure 5.9 Pseudo-entropy profiles traced by the stars in the main halo in the Hydro-1x (solid red line) and the Hydro-10x (solid orange line) at $z = 0$. For both profiles we also report the respective intrinsic scatter of the simulated cluster samples (the shaded areas) and the best-fit profiles (dashed lines), using the same colour code. To help the comparison, we report the best-fit profile for the DM particles (dotted black line) in the Hydro-1x simulation. . . . 103

Figure 5.10	Pseudo-entropy profiles traced by the stars bound to the BCG (in green), ICL (in yellow) and the two combined (in red) in the Hydro-1x simulation at $z = 0$. The dashed black line is the best-fit profile traced by all the stars while the dotted black line is the best-fit profile for the DM particles which has been shifted in order to match the pseudo-entropy value of the ICL profile at $0.1 r_{\text{vir}}$. The vertical grey line helps the eye to visualise until which radius ($\sim 0.55 r_{\text{vir}}$) the power-law holds for the sample of all the stars at this resolution.	104
Figure 5.11	Phase-space density profiles as traced by DM within the Hydro-10x (in blue), DM-10x (in orange) and from the observed dataset from Capasso et al. (2019) (in red). We also report the phase-space density traced by both DM particles and baryons (stars and gas) within the Hydro-10x simulations (in green). The profiles are shown for the same 5 different redshift bins within which the analysis of observational data has been carried out. Profiles from DM particles in the Hydro-10x have been rescaled by $(1 - f_b)$ (f_b is the cosmic baryon fraction assumed in the simulations). For simulations, lines show the median profiles while shaded areas encompass the 16 th to 84 th percentiles.	106
Figure 5.12	Phase-space density profiles as traced by substructures within the Hydro-10x simulations (in blue) and galaxies in the observational analysis in Capasso et al. (2019). For each redshift interval, profiles have been normalised to match at $0.5 r_{\text{vir}}$. Lines and shaded areas have the same meaning as in Figure 5.11.	107
Figure 5.13	Left panel: pseudo-entropy profiles (DM particles) for a relaxed and a disturbed cluster in the Hydro-1x simulation at $z = 0$. The dashed lines are obtained by stacking the true profiles (in cyan) or those obtained by the Jeans equation procedure (in orange), while the solid lines are for the profiles of the specific cluster. Right panel: density maps in logarithmic scale traced by the DM particles in the same objects. The brightest colours indicate the areas at higher density. The white circles identify the $0.5 r_{\text{vir}}$ spherical regions centred on the cluster centre as computed by Subfind.	111

Figure 6.1	<p>A comparison of some of the general properties of our simulated set of clusters with the observational results by Sohn et al. (2020). Left panel: scaling relation between cluster velocity dispersion, σ_{200} and total cluster mass, M_{200}; central panel: scaling relation between BCG stellar mass, M_{BCG}^* and M_{200}; right panel: scaling relation between BCG velocity dispersion, σ_{BCG}^* and M_{BCG}^*. Legend is common to all panels: dark green and golden points indicate respectively the Hydro-1x and Hydro-10x samples using all the stars in the main halo enclosed in a bi-dimensional aperture of 50 kpc. The grey circles stand for the observational results by Sohn et al. (2020). Velocity dispersions are measured for each simulated cluster within three orthogonal projections.</p>	120
Figure 6.2	<p>The velocity bias as a function of cluster mass in the Hydro-1x (dark green points) and Hydro-10x samples (golden points). The dashed black line reports the line for no bias ($\sigma_{200, \text{DM}}/\sigma_{200, \text{gal}}=1$). Velocity dispersions are measured for each simulated cluster within three orthogonal projections.</p>	123
Figure 6.3	<p>Relationship between BCG and cluster line of sight velocity dispersion from clusters in Sohn et al. (2020) (in grey), and at $z = 0$ the Hydro-1x and Hydro-10x simulations. The coloured points are the result of selecting all stars in the main halo within the cylindrical projection. Velocity dispersions are measured for each simulated cluster within three orthogonal projections.</p>	123
Figure 6.4	<p>Stellar velocity dispersion of the BCG (blue) and ICL (red) components in the main halo of our set of simulated galaxy clusters, obtained from the double Maxwellian fit of the particle velocity distribution as a function of host cluster mass M_{200}. The distinct markers encode the two distinct simulations: simple circles for the Hydro-1x and crossed circles for the Hydro-10x. The solid lines represent the best-fit results from Dolag et al. (2010), while the dashed lines mark the best-fit from our sample.</p>	126
Figure 6.5	<p>Stellar BCG velocity dispersion versus ICL velocity dispersion in the Hydro-1x (simple circles) and Hydro-10x (crossed circles) runs. Solid line reports the best-fit curve from Dolag et al. (2010) while the dashed line is derived from our cluster set. We use a colour-coded description to mark the ICL mass fraction in each cluster.</p>	126

- Figure 6.6 A comparison of the scaling relations for the stellar population of our simulated clusters in both BCG and BCG+ICL definitions, for the Hydro-10x sample. Left panel: scaling relation between BCG stellar velocity dispersion, σ_{BCG}^* and total cluster mass, M_{200} ; central panel: scaling relation between M_{BCG}^* and M_{200} ; right panel: scaling relation between BCG velocity dispersion, σ_{BCG}^* and M_{BCG}^* . In the last panel, we also report the predictions from the virial theorem for both samples with coloured dashed lines. Legend is common to all panels: golden points indicate the BCG+ICL sample while the brown crosses stand for the BCG sample. Velocity dispersions are measured for each simulated cluster within three orthogonal projections. 129
- Figure 6.7 Relationship between line of sight BCG velocity dispersion as a function of cluster velocity dispersion, at five different redshifts. Clusters at each redshift are binned in σ_{200} . For each bin, we plot the corresponding median value of σ_{BCG}^* . The grey-shaded band marks the standard deviation of the clusters at $z = 0$, with comparable dispersions at different redshifts. At the different redshifts, $z = 2, 1, 0.5, 0.2, 0$, we have the following corresponding number of clusters: 113, 80, 84, 80, and 93. The value of σ_{BCG}^* is obtained by also including the contribution of ICL. 131
- Figure 6.8 Evolution of the scatter relation between the velocity dispersion of the BCG (normalised at $z = 0$) versus the cluster velocity dispersion (normalised at $z = 0$) from 8 clusters in the Hydro-10x run. The redshift bins ($z = 2, 1, 0.5, 0.2, 0$) are labelled with different colours. 132
- Figure 6.9 Comparison with the line of sight velocity dispersion from Sartoris et al. (2020) (points with error bars) and the projected velocity dispersion profiles from one of our simulated clusters at $z = 0$ in the Hydro-10x set. Error bars in the observational points refer to a 68 per cent confidence level. The solid lines are the mean of the three projected velocity dispersion profiles (faded lines). 134

- Figure 6.10 Comparison of the observed line of sight velocity dispersion from Sartoris et al. (2020), with results from the combined sample of simulated clusters at $z = 0$ from the Hydro-10x set. Error bars in the observational points refer to a 68 per cent confidence level. The shaded areas represent respectively the 16th and 84th percentiles given by the simulated cluster sample. 134
- Figure 6.11 Top panels: comparison of the slope of the stellar velocity dispersion η with the observational results from Newman et al. (2013) (maroon triangles) and Loubser et al. (2018) for BCG (in light turquoise circles). The legend is common to all panels: dark green and golden points indicate the Hydro-1x and Hydro-10x's BCG+ICL samples, respectively; light green and brown crosses label the BCG-only sample from the Hydro-1x and Hydro10x runs, respectively. Following the approach adopted in Loubser et al. (2018), we report the slopes η obtained by fitting the velocity dispersion profiles over the radial range from a central aperture of $r_0 = 5$ kpc and 15 kpc. We show results only for the Hydro-10x sample. Bottom panels: the same analysis as in the top panels, but computing the slope of the velocity dispersion profiles between $0.1 R_{500}$ and $0.05 R_{500}$, with the solid lines showing the best-fit results. The left and right panels show the slope η of the velocity dispersion profiles versus the central velocity dispersion σ_0 and the cluster velocity dispersion σ_{200} , respectively. 136
- Figure 7.1 Probability density functions of the input features associated with the ICL (red) and BCG (blue) in one of the clusters. Left panel: distribution of the logarithmic cluster distance over R_{200} . Right panel: stellar rest-frame velocities scaled by the virial circular velocity V_{200} 145
- Figure 7.2 Importance of the input features including all initial input features to predict ICL and BCG components. From top to bottom: the logarithm of the cluster mass M_{200} , the stellar metallicity, age, the logarithmic cluster distance over R_{200} , and the rest-frame velocities scaled by the virial circular velocity V_{200} . The black bars encode the standard deviation when sampling the importance from the trees. 147

Figure 7.3	Top panel: distribution of the input features drawn from the simulated cluster. In each panel, we plot the predicted (bars) and true (line) number counts of both ICL (red) and BCG (blue) associated with the star particles. More in detail we have (from left to right) the distributions of the logarithmic cluster-centric distance normalised for R_{200} and the stellar rest-frame velocities normalised for V_{200} . Bottom panel: percentage residuals measured between the true and predicted counts in each bin. 150
Figure 7.4	Mass-weighted maps of the stellar components (ICL in the top panels, BCG in the bottom ones) in the same cluster as in Figure 7.3 . Left panels report the results from the stellar division provided by ICL-Subfind. Right panels show the mass-weighted maps for the stellar components identified with the predicted labels. 151
Figure 7.5	Top: density profiles of the BCG (blue) and ICL (red) in the selected cluster. The dashed lines are the profiles computed with the predicted labels, while solid lines report the profiles computed with the true labels. Bottom: residuals (in per cent) between the histograms of the labels from the ICL-Subfind and the ML classifier. 152
Figure 7.6	Comparison of the velocity dispersions derived from the fit of the double Maxwellian in Equation 24 in the stellar velocity distribution labelled by the ML classifier and the ICL-Subfind. Each point represents the result from the component of a single cluster, coloured according to the ICL mass fraction estimated in each cluster by the ICL-Subfind. Furthermore, we mark the clusters from the training set with a cross. The dashed grey line is the reference line to a 1 : 1 relation. Left panels: the velocity dispersions in the ICL component. Right panels: the velocity dispersions in the BCG component. 153
Figure 7.7	Stack of the metric score profiles relative to the BCG particles classification in the entire cluster set. The profiles are reported as a function of the cluster distance normalised by R_{200} . Solid lines refer to P (red), R (blue) and FS (brown). The dashed line marks the value 0.5 on the y-axis. 155

- Figure 7.8 Scatter plots of the (weighted) mean metric scores (P, R, and FS from top to bottom panel) as a function of log cluster mass (left panels) and dynamical state (right panels) for all the clusters under study. We mark the pure testing set with coloured points, while the clusters in the training set are recorded with empty squares. The colour legend follows the ICL fraction estimated by ICL-Subfind. The y-axis is limited to exclude D7, D11 and D13 having very low metric scores. 157
- Figure 7.9 Density profiles of the stellar content in D2 taken at three different snapshots: from top to bottom the corresponding redshifts are $z \simeq 1, 0.5, 0$. For each of these plots, we show two panels. In the top panel, we report the profiles normalised at R_{200} of both ICL (red) and BCG (blue) identified by ICL-Subfind (solid line) and the ML classifier (dashed line). The lower panel shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels. 159
- Figure 7.10 Density profiles of the stellar content in D2 simulated at higher resolution. In the top panel, we report the profiles normalised at R_{200} of both ICL (red) and BCG (blue) identified by ICL-Subfind (solid line) and the ML classifier (dashed line). The lower panel shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels. Differences account for a per cent level only between the profiles. 160
- Figure 7.11 Density profiles of the stellar content in D2 simulated with no AGN feedback scheme. In the top panel, we report the profiles normalised at R_{200} of both ICL (red) and BCG (blue) identified by ICL-Subfind (solid line) and the ML classifier (dashed line). Additionally, we overplot the BCG and ICL profiles from the same cluster when including the AGN feedback model (dotted line). The lower panel shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels. 161

Figure 7.12	Histograms of the distributions of the features of D2 simulated with no AGN feedback scheme. In the top row, we report the number count distributions of (from left to right): the logarithmic clustercentric distance over R_{200} and the particle rest-frame velocity over V_{200} . The lower row shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels.	161
Figure 7.13	Density (top), velocity dispersion normalised for the velocity dispersion of the stars within R_{200} (central), and phase-space density (bottom) profiles of the star particles in the BCG (blue) and ICL (red). Results from the ICL-Subfind labels are given with solid lines, on the other hand, the dashed lines mark the profiles extracted from the ML classified labels. In each panel, we report the median profile of each method (dark solid line) and the uncertainty given by the intrinsic standard deviation with the shaded coloured band.	164
Figure 7.14	Mass-weighted map of the stellar components in cluster D7: BCG+ICL (in the left panel), ICL (central) and BCG (left) spatial distributions. We only show the results from the predicted labels, since the ICL-Subfind does not converge for this case.	165
Figure 8.1	The ICL mass fraction over the ICL and BCG in one of the clusters. Darker (brighter) colours mark lower (higher) fractions.	175
Figure 8.2	One way to provide data augmentation is to perform flips and rotations to one image, as reported in the panels.	176
Figure 8.3	Loss function for one of the input datasets as a function of the learning rate. The red dot marks the point in the curve with the steepest gradient which corresponds to the suggested value for the learning rate.	178
Figure 8.4	Magnitude and colours maps	180
Figure 8.5	Predicted (left) and true (right panel) distribution of the ICL mass fraction in a cluster belonging to the testing set. The colour bar marks the region with the highest concentration with yellow towards purple for lower concentrations.	181
Figure 8.6	Input (panel a) and output (panel b) of the model using velocity maps as input images.	182
Figure 8.7	Input (panel a) and output (panel b) of the model using velocity maps as input images.	183

Figure 8.8	Input (panel a) and output (panel b) of the model using velocity dispersion maps as input images.	186
Figure 8.9	Percentage error maps for the clusters with velocity maps as the input dataset. In each central panel, we display the PE between the predicted and true ICL mass fraction for pixel. Generalising these results, we compute the MAPE given by each line of pixels. We report these errors on the top histogram (i.e., MAPE along the vertical direction) and right histogram (MAPE along the horizontal direction) with respect to the central panel. The grey dashed lines mark the necessary scale.	187
Figure 8.10	ICL fractions for one cluster as a function of the normalised clustercentric distance. The ICL fraction is determined as $ICL/(BCG+ICL)$. Lines are colour-coded for the three different numerical resolutions (i.e., 1x, 10x, 25x). The dotted horizontal lines report the median ICL fraction from the entire sample of clusters in each resolution run.	190
Figure 8.11	Cumulative stellar mass of the ICL (red) and BCG (blue) as a function of the normalised clustercentric radius for the same cluster simulated at three different numerical resolutions (i.e., 1x, 10x, 25x). Each resolution is marked with a different line style, as reported in the legend. The vertical grey lines represent the "transition radius" for each pair of simulations: these monitor the change between the BCG-dominated area to the ICL-dominated one. As the resolution increases, the shift to the ICL-dominated region moves to smaller radii.	190
Figure A.1	Map of the CMB from the Planck telescope. Colour differences correspond to differences in temperature. Credit: ESA and Planck collaboration	204
Figure A.2	Growth factor $D_+(z)$ as a function of redshift for five different cosmological models as indicated. The growth factor is normalised to unity at the present epoch, and divided by the scale factor to emphasise the differences between the models. With increasing ω , the growth factor increases towards its value for the open model with Ω_M . Credits: Bartelmann et al. (2002)	207

- Figure A.3 Upper panel: critical threshold for collapse, δ_c as a function of Ω_M in the spherical collapse model. Results are plotted for open models with $\Omega_\Lambda = 0$ (dotted line) and flat models with $\Omega_M + \Omega_\Lambda = 1$ (dashed line). Lower panel: the virial density of collapsed objects in units of the critical density; the dotted and dashed lines are as in the upper panel. Credits: Eke et al. (1996) 210
- Figure A.4 Recap of the itinerant writing of this thesis. 220

LIST OF TABLES

Table 3.1	A summary of the main characteristics of the analysed simulations. For each set of simulations, we report: DM and gas particles mass; DM, gas, star particles and black hole (Plummer-equivalent) softening lengths at redshift $z = 0$; and the total number of clusters analysed. The simulations are named "Hydro" or "DM" whether they are hydrodynamical or DM-only respectively.	46
Table 5.1	Best-fit results for the normalisation S_0 and exponent α in the power-law expression of Equation 35 for the median pseudo-entropy profiles at $z = 0$ traced by DM particles.	88
Table 5.2	Best-fit results for the normalisation S_0 and exponent α in the power-law fitting of the median pseudo-entropy profiles at $z = 0$ traced by substructures.	99
Table 5.3	Best-fit results for the normalisation S_0 and exponent α in the power-law fitting of the median pseudo-entropy profiles at $z = 0$ traced by stars. . .	104
Table 5.4	The best-fit values obtained for the logarithmic slope in the true profile and the one recovered through the Jeans equation procedure for the relaxed and disturbed cluster shown in Figure 5.13 with the 68 per cent uncertainty. In the last row, the results were obtained when fitting the median profile from the stack sample of 29 clusters in the Hydro-1x simulation at $z = 0$. In the latter, the uncertainties reported are the standard deviation derived from the distribution of best-fit values for the sample.	112
Table 7.1	A summary of the main characteristics of the simulated clusters used in the training and testing phases at $z = 0$. We report the given cluster name, cluster mass M_{200} , the cluster radius R_{200} , the cluster orbital velocity V_{200} , the stellar mass in the main halo $M_{*,gal}$ and the dynamical state. We add an asterisk to the clusters which are part of the training set.	146

Table 7.3	Results from the best-fit procedure applied to the double Maxwellian in both the ICL-Subfind and ML algorithm case. We list the convergence report of ICL-Subfind in the second column. Then, we present the ICL and BCG velocity dispersions, the fraction of ICL, and in the last columns the algorithm performance scores (of the two classes and their weighted median), namely: the P, R, and FS. . . . 169
Table 8.1	A summary of the hyperparameters used for the Adam optimiser in the different input models. . . . 179

ACRONYMS

AGN	Active Galactic Nuclei
BAO	Baryonic Acoustic Oscillations
BCG	Brightest Cluster Galaxy
BH	Black Hole
CMB	Cosmic Microwave Background
CDM	Cold Dark Matter
CNN	Convolutional Neural Network
DL	Deep Learning
DM	Dark Matter
FLRW	Friedmann-Lamaître-Robertson-Walker
FoF	Friends of Friends
FS	F-Score
ICL	IntraCluster Light
IFS	Integral Field Spectroscopy
IMF	Initial Mass Function
ML	Machine Learning
MSE	Mean Squared Error
PE	Percentage Error
MAPE	Mean Absolute Percentage Error
MFM	Meshless Finite Mass
NN	Neural Network
P	Precision
PM	Particle Mesh
R	Recall
ReLU	Rectified Linear Unit
SAM	Semi-Analytic Model
SMBH	Supermassive Black Hole

SNe Supernovae

SPH Smoothed-Particle Hydrodynamics

SSP Single Stellar Population

INTRODUCTION

According to the standard cosmological scenario, structures in the Universe have formed as a result of the gravitational instability of tiny initial perturbations in the density field, superimposed on a homogeneous and isotropic background. These initial perturbations are thought to be the result of Gaussian quantum fluctuations in a scalar field (i.e., the Inflaton; see Guth 1998) which generated an exponential expansion of the Universe. Thus, the Universe is currently believed to have originated from an extremely hot and dense phase extremely high temperature and density (i.e., the Big Bang theory), which began to expand in a process that is considered to be still ongoing today. Perhaps the most conclusive, and certainly among the most carefully examined evidence for the Big Bang is the existence of an isotropic radiation bath known as the Cosmic Microwave Background (CMB), which was first discovered by Penzias and Wilson (1965). Besides the study of the anisotropies in the CMB (Fixsen et al., 1996; Torbet et al., 1999; Melchiorri et al., 2000; Hanany et al., 2000; Hinshaw et al., 2007; Collaboration et al., 2020), other independent cosmological probes have been gathered in recent years to sustain the Λ Cold Dark Matter (CDM) standard cosmological model. Among these, we remind the observed primordial abundance of cosmic elements (Burles and Tytler, 1998), the observations of the motion of distant objects such as galaxies (see Hubble 1929; Hubble and Humason 1931) or Supernovae (SNe) type Ia (in Riess et al. 1999), the velocity dispersion profiles of galaxies within clusters (we remind the seminal papers Zwicky, 1937b,a) and the rotational curves measured in spiral galaxies (Rubin et al., 1978, 1980). Other probes come from the large scale observables, such as galaxy clustering (i.e., Baryonic Acoustic Oscillations (BAO) more details in Cole et al. 2005; Percival et al. 2007 and redshift space distortion Kaiser 1987), and cosmic shear (Abbott et al., 2018). This Λ CDM paradigm is also favoured for its simplicity as it requires only six free parameters to fully constrain the model (matter density Ω_M , baryonic density Ω_b , dark energy density Ω_Λ , the spectral index n_s , the curvature fluctuation amplitude σ_8 at 8 Mpc h^{-1} and the Hubble parameter H), once we assume the validity of the cosmological principle (homogeneity and isotropy in the Universe) and general relativity.

One of main avenues to investigate the fundamental ingredients in the history of the Universe evolution is to study the matter (or light) distribution, which is assembled in the so-called *cosmic web*. As the name suggests, the cosmic web is an interconnected structure stretched between high-density (rich in stars and galaxies) and low-density regions through thin *filaments* and *voids*, as depicted in Figure 1.1 showing a slice of the simulated universe by the Millennium XXL (Kuhlen et al., 2012). At the intersections, we find galaxy clusters, which are structures in quasi-dynamical

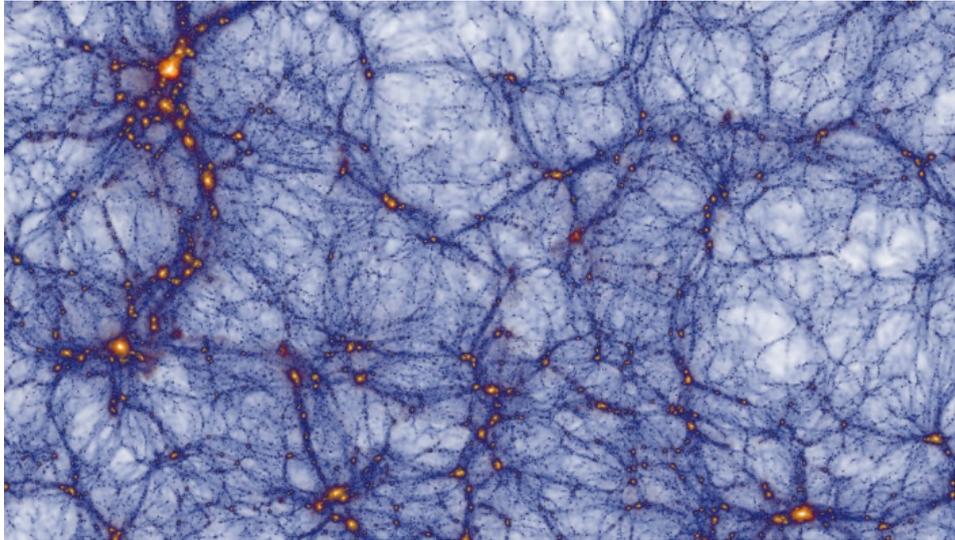


Figure 1.1: A slice through the Millennium XXL simulation showing the simulated cosmic web. The bright yellow regions are the high density clusters that form at the intersection of filaments. The simulation follows the non-linear growth of DM structures within a cubical region 4.1096 Gpc on a side. The particle mass is $8.456 \times 10^9 M_{\odot}$. Credits: <http://galformod.mpa-garching.mpg.de/mxxlbrowser/>

equilibrium with a characteristic mass of $10^{14} - 10^{15} M_{\odot}$ that typically contains several hundreds, or even thousands, of galaxies. Galaxy clusters have proved to be excellent cosmological probes, besides their importance as astrophysical systems. They are the largest virialised objects in the Universe, formed of baryonic and Dark Matter (DM), since they are associated with the collapse of the largest gravitationally bound overdensities in the initial density field of the Universe (e.g., Kravtsov and Borgani, 2012). The number and mass distributions of galaxy clusters are powerful probes to constrain the cosmological parameters (e.g., Allen et al., 2011). Along with the cluster mass estimates, member galaxies orbiting in the cluster potential provide important information on the build-up of galaxy clusters (Dolag et al., 2009). Structures form via merging of smaller systems, and the signature of the infall process is imprinted in the dynamics of these objects. Conversely, given the high-density environment in which galaxies live, evolutionary processes may also impact their evolution, for example tidal stripping (Binney and Tremaine, 1987) and dynamical friction (Chandrasekhar, 1943). Thus, the extreme environment within which cluster galaxies are embedded allows us to study the interplay of a large variety of different astrophysical processes under extreme conditions. For all these reasons, galaxy clusters represent nowadays exceptional laboratories for astrophysical and cosmological purposes.

Given such remarkable properties, it is no surprise that the quest to understand the formation and evolution of galaxy clusters has become one of the central efforts in modern astrophysics over the past several decades. Examining the internal dynamics of galaxy clusters gives us insights to

improve the current paradigm of structure formation: identifying the most important processes shaping the observed properties of clusters and their evolution is the key to this prospect. Such a complete description requires detailed modelling of the non-linear processes of collapse and the dissipative physics of baryons. Gas is heated to high thermally bremsstrahlung X-ray emitting temperatures by adiabatic compression and shocks during the collapse and settles in hydrostatic equilibrium within the cluster potential well. Once the gas is sufficiently dense, it cools, process that can feed both star formation and accretion onto Supermassive Black Holes (SMBHs) harboured by the massive cluster galaxies. The process of cooling and formation of stars and SMBHs can then result in energetic feedback due to SNe or Active Galactic Nuclei (AGN), which can inject substantial amounts of heat into the intracluster medium and spread heavy elements throughout the cluster volume. To describe such complex interplay, we can resort to Semi-Analytic Models (SAMs) and/or numerical simulations which are capable of providing us with complete information of the system at each instant, thus allowing us to trace in detail how the dynamics of cluster formation and evolution define the phase-space structure of galaxy clusters.

This thesis aims to carry out, through the detailed analysis of an extended set of cosmological hydrodynamical simulations of galaxy clusters, a study focused on the dynamical properties of the collisionless components forming galaxy clusters (i.e., DM, stars and galaxies). Recent observational measurements in the local Universe have provided us with a wealth of information on the phase-space properties of galaxy clusters and their components which can be used to assess the quality of predictions from our simulations. We particularly focus our attention on the properties of the stellar and galactic population, which can be directly observed, and on the underlying DM distribution. One of the many scopes in our analysis is to quantify the impact of modelling and numerical resolution in our results, which is why we encouraged the use of multiple simulation runs and subgrid models.

The thesis is structured as follows.

CHAPTER 2 presents a brief overview of the observational properties of galaxy clusters. In this chapter we mainly discuss the galaxy member population, the stellar component in the central galaxy and the IntraCluster Light (ICL).

CHAPTER 3 illustrates the numerical methods developed to describe the dynamics of cosmic structure formation in the non-linear regime. This chapter is dedicated to numerical cosmological simulations. More in detail, we firstly present a description of algorithms employed in the N-body and hydrodynamical simulations, going further into the details of the Tree-Particle Mesh (PM)/Smoothed-Particle Hydrodynamics (SPH) code in GADGET-3, a modified version of the GADGET-2 (Springel et al., 2005), which is the code used to carry out the simulations analysed in this thesis. Later on, we provide

a general outline of the astrophysical subgrid models introduced in GADGET-3, which include radiative cooling, chemical enrichment, star formation and AGN feedback. We conclude the chapter by presenting Smac, a map making tool to create mock images of galaxy clusters from cosmological simulations.

CHAPTER 4 outlines the Machine Learning (ML) techniques we employed to carry out this thesis work. Considered the vastness of the topic, we focused only on the general aspects common to all ML models. Furthermore, two basic ML architectures are presented, i.e. Random Forest and Convolutional Neural Network (CNN), both of which will be used in this work.

CHAPTER 5 provides a detailed description of our analysis of the phase-space structure of cosmological simulations of galaxy clusters. More specifically, we analysed four sets of simulations from the DIANOGA project, which differ in the physics included (the presence or not of baryons) and resolution. At the beginning of this chapter, we define the adopted parameters employed in the simulations (such as the cosmological parameters, mass and force resolution) with some specifications on the method of analysis. Finally, we present the results obtained on the dynamical properties of the phase-space of the different components within galaxy clusters (DM, stars and substructures) focusing especially on pseudo-entropy profiles. Universal features are commented to provide a general picture of how the cluster population evolves in a Λ CDM universe. With these results, it is possible to investigate the pseudo-entropy (for redshifts $z < 2$), and evolution of the pseudo-entropy traced by two distinct dynamical components of stars, i.e. ICL and stars in the Brightest Cluster Galaxy (BCG). In the final part of this chapter, we show a comparison with the observational results (as described in Capasso et al., 2019; Biviano et al., 2016). The results presented in this chapter have been published in Marini et al. (2021a).

CHAPTER 6 describes the results of our analysis of the velocity dispersion of the stellar component in the central galaxy and making up the ICL. These two components exhibit rather different dynamical properties and we aim at quantifying such differences in a systemic approach provided by the set of cosmological hydrodynamical simulations of galaxy clusters in the DIANOGA set. The content of this chapter has been published in Marini et al. (2021b).

CHAPTER 7 presents a ML-based classification method that we specifically developed to separate the stellar population in the main halo of simulated galaxy clusters into stars bound to the BCG (stars with low entropy) and those bound to the cluster potential (higher entropy) forming the ICL. These two dynamically distinct stellar components are the result of different mechanisms simultaneously operating in the cluster halo. Stars in the central galaxies have possibly formed

after in-situ star formation and have thus got rid of the excessive orbital energy in order to settle into the galaxy's potential. ICL is thought to be the outcome of tidal stripping of stars from the orbiting galaxies in the cluster. Therefore, being able to dynamically separate these two stellar components may help unveiling information on the history of the cluster itself and the mechanisms in act. This algorithm exploits a Random Forest to identify differences in the phase-space of star particles. The supervised learning process is based on pre-compiled stellar catalogues produced with a modified version of the halo finder Subfind (extensively described Dolag et al., 2009) which, in turn, labels stars according to an unbinding procedure. In this chapter, we critically assess the accuracy compared to the traditional method. Furthermore, we discuss the limits and benefits of using the Random Forest. Our findings are reported in Marini et al. (2022).

CHAPTER 8 portrays a preliminary study dedicated to detecting ICL in observed images of galaxy clusters. We create mock images of the stellar population in the main halos of our simulated galaxy cluster population with different observables (i.e., magnitudes, colours, line of sight velocity and velocity dispersion) and we train a U-Net (Ronneberger et al., 2015) to infer the ICL distribution from them. This is possible thanks to the precompiled catalogues of the BCG and ICL population by the ICL-Subfind which we use to obtain the true label in the ICL distribution. We mimic the possible presence of interlopers and de-centering. The best model results to be the one trained with the velocity maps as input feature, although we suggest the implementation of a few measures to possibly improve the model in future iterations. This project has highlighted the dependence of the ICL mass fraction with the numerical resolution of the simulation, as critical issue which further needs to be addressed.

CHAPTER 9 summarises the main conclusions of this thesis and presents future lines of investigation based on the results presented.

APPENDIX A presents the general theoretical background describing the evolution of the initial perturbations in the cosmic density field to the formation of structures due to gravitational instability. The ultimate stage of such collapse is a configuration of equilibrium in which structures are governed by collisionless dynamics. Thus, in the final part of this chapter, we outline the main equations governing the collisionless systems, namely the continuity equation (which describes the mass conservation), the Euler equation (the equations of motion) and the Poisson equation (describing the gravitational field).

Part I

SCIENTIFIC BACKGROUND

Clusters of galaxies are the largest gravitationally bound structures in the Universe: typical masses range from 10^{14} to few $10^{15}M_{\odot}$. Their abundance and vast spatial distribution in the sky are important cosmological probes that allow us to test the initial conditions of the Universe and constrain cosmological parameters (Allen et al., 2011). At the same time, the extreme environment in which they are embedded permits us to study the interplay of a large variety of different astrophysical processes under extreme conditions. Aiming to obtain results from both astrophysics and cosmology, along with many aspects concerning cosmic evolution, thermodynamics of the intergalactic medium and galaxy evolution, galaxy clusters represent unique laboratories.

In this chapter, we describe in general terms how galaxy clusters are expected to form according to the Λ CDM model. After discussing the physical mechanisms that cause these systems to reach equilibrium, we describe the dynamics of collisionless equilibrium (conditions under which galaxies, DM and stars in halos operate). On the other hand, a theoretical treatment of the diffuse gas – that represents a sensitive portion of the cluster mass and follows a dynamics complicated by the collisional and thermodynamical interactions – will not be part of the following discussion. The rest of the chapter will be devoted to present properties and scaling relations observed in the cluster galaxy population. The purpose of this introductory chapter is to sketch our current understanding of galaxies and galaxy clusters and their formation process, without going into any detail. A more exhaustive treatment of these topics can be found in Binney and Tremaine (1987), Kravtsov and Borgani (2012), Mo et al. (2010) and Pratt et al. (2019).

2.1 GALAXY CLUSTERS IN THE COSMOLOGICAL ENVIRONMENT

In the standard cosmological model, structure formation is addressed in terms of the growth of small perturbations in the initial cosmic density field described by a random Gaussian field, through *gravitational instability*. In this picture, structures – such as clusters of galaxies – are the last systems having formed in an expanding Universe thanks to the gravitational force acting from the primordial epochs (Mo et al., 2010). Although the scientific community has not yet reached a common agreement on how these fluctuations may have been generated, the most widely accepted hypothesis indicates them as the natural result of a primordial inflationary epoch (Guth, 1998; Linde, 2005).

2.1.1 *How do they form?*

At early times, when the perturbations are still in the so-called "linear regime" (i.e., density perturbations with $\delta\rho/\rho \ll 1$ with respect to the background density) their physical size follows the Hubble flow, expanding with the Universe. Ultimately gravity causes overdensities to reach $\delta\rho/\rho \simeq 1$. This is the turnaround limit considered to break the linear approximation whereby perturbations detach from cosmic expansion and start to collapse. The subsequent non-linear, gravitational collapse depends on the matter content of the perturbation. Halos associated with galaxy clusters are multi-component systems consisting of DM and baryons in several phases: Black Holes (BHs), stars, and gas (in different chemical compositions and temperatures). If the perturbation consists of ordinary baryonic gas, the collapse creates strong shocks that induces a raise in entropy. If radiative cooling is inefficient, the system relaxes to hydrostatic equilibrium, with its self-gravity balanced by pressure gradients. On the other hand, if the perturbation consists of collisionless matter (e.g. CDM), no shocks develop, but the system still relaxes to a quasi-equilibrium state with an almost universal structure after violent relaxation (Lynden-Bell, 1967). Non-linear, quasi-equilibrium DM objects are called DM halos. Their predicted structure has been thoroughly explored using numerical simulations, and they play a pivotal role in modern theories of galaxy formation, as we will discuss more in [Appendix A](#).

After decades of studies via N-body and hydrodynamical simulations (see Borgani and Kravtsov, 2011 for a review), models of the detailed evolution of the baryons in clusters are growing in capability to describe an increasingly large and rich volume of observations. As such, we can theorise the main steps in the formation of clusters. The gas is heated due to the shocks and adiabatic compression undergoing as the structure forms via gravitational collapse. Temperature measurements of the extended hot gas – the so-called intracluster medium – have supported the theoretical predictions of a DM halo to explain the depth of the gravitational potential in clusters. The most striking evidence was provided in the seminal paper by Clowe et al. (2006) with the detection of both the gravitational potential and X-ray radiation in the Bullet cluster. The Bullet cluster is a post-merging structure in which the collisional and collisionless components, namely gas and DM (for the most part), have different spatial centres. In [Figure 2.1](#), we show the optical counterpart (left panel) with the green contours denoting the bimodal distribution of the gravitational potential obtained from the weak lensing modelling; in the right panel, we show the Chandra X-ray data measurements tracing the hot plasma in the galaxies: dynamical friction slows down the gas which thermalises at the centre of the two DM peaks. Indeed, during the interaction, gas and plasma will undergo shocks and heating, while the collisionless DM component will pass right through. If the gas settles into a configuration of equilibrium, it cools down and, if it is sufficiently dense, it forms the stellar population in the cluster. The stellar mass accounts for only 12 ± 2 per cent of the

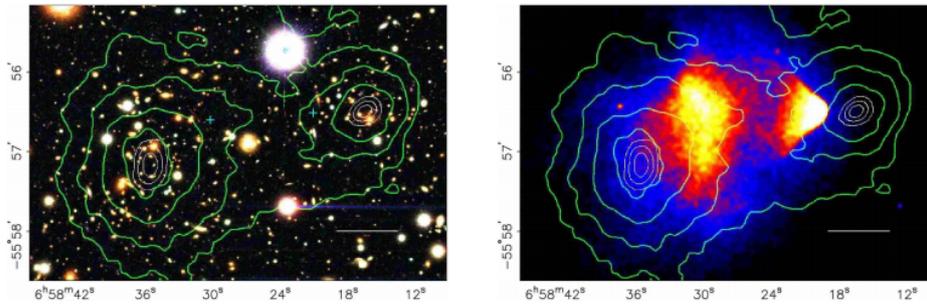


Figure 2.1: The Bullet cluster. Left panel: we show the optical image taken with the Magellan telescope. Right panel: Chandra X-ray measurements of the hot plasma in the cluster halo. The green contours denote the gravitational potential modelled with weak lensing measurements. Credits: Clowe et al. (2006).

total baryon budget in the most massive halos (Giodini et al., 2009; Chiu et al., 2018 and references therein) and it can be divided into two different components: stars bound to individual galaxies and those part of a diffuse component bound to the cluster potential – the ICL. Radiative cooling of gas is regulated by feedback from various sources, including mechanical and radiative input from SNe winds and BH jets, thermal conduction along with other plasma processes, ablation and harassment during gravitational encounters.

Hence, most clusters appear with an extended gas halo (predicted to trace the DM halo spatial distribution) which envelopes the galactic population in the potential. A small contribution to the matter budget comes from the stellar population, which generally resides in the orbiting galaxies or to a smaller extent in the cluster halo. Other elements (e.g., SMBH and dust) will be ignored at this level.

2.1.2 How do they evolve?

At the end of the gravitational collapse, matter settles into an equilibrium configuration. The collisionless component satisfies the condition of equilibrium as theorised by the Jeans equation (fully derived in Section A.5) which, for a non-rotating spherically symmetric system, takes the form:

$$M(< r) = -\frac{r\sigma_r^2}{G} \left(\frac{d \ln v}{d \ln r} + \frac{d \ln \sigma_r^2}{d \ln r} + 2\beta \right) \quad (1)$$

and relates the total mass $M(< r)$ (enclosed within r) with the orbital properties of the components in the virialised volume. The orbital anisotropy parameter $\beta = 1 - \sigma_t/(2\sigma_r)$ is defined in terms of the tangential velocity dispersion σ_t and the radial dispersion σ_r , whereby

- $\beta = 0$: perfectly isotropic orbits;
- $\beta = 1$: all orbits are radial;

- $\beta = -\infty$: perfectly tangential orbits.

Proof of the validity of the virial theorem in clusters (and groups) is often given by the dynamical properties of the member galaxies. In relaxed clusters, observations have shown that their orbits tend to be radially anisotropic (Biviano et al., 2013; Capasso et al., 2019). The term $v(r)$ denotes the number density profile of the tracers. In principle, galaxies in groups and clusters are not strictly collisionless, yet interactions between members are relatively rare and the Jeans equation should be quite accurate.

NATURE OR NURTURE? Broadly speaking, galaxies can form in a dense environment (in clusters or groups) or as isolated systems in the so-called "field". This difference can dictate the galaxy's growth and thus its properties, such as morphology and colour. In other words, one may ask to what extent are galaxies governed by initial conditions (*nature*), and to what extent are they driven by environmental effects (*nurture*). Disentangling causes and effects is a challenging task.

Galaxies in rich and dense environments have typically more evolved (elliptical) morphologies (e.g., Oemler Jr, 1974; Dressler, 1980), lower gas masses (from past interactions or lost in triggered events of star formation Mihos, 2004), redder colours (e.g., Siudek et al., 2022), more compact structures (Boselli et al., 2022) and lower star formation rates (De Lucia et al., 2019; Lotz et al., 2019) than galaxies of similar mass living in regions of the Universe with mean density. Modern theoretical models of galaxy formation describe this as the result of the combined action of many physical processes taking place through galaxy-galaxy (e.g., tidal stripping) or galaxy-environment (e.g., ram pressure and strangulation) interactions. This is why we cannot expect galaxies in clusters, groups and field to evolve altogether along a similar evolutionary path. On the other hand, this evidence can be accounted to our advantage: properties of cluster galaxies can become tracers of the evolutionary path of the cluster itself. Efficient quenching of star formation in central regions can be hinting at the presence of AGN feedback (Miller et al., 2003) while orbital properties of cluster galaxies seem to be tightly connected with the star forming properties of these systems (Lotz et al., 2019).

In this thesis, galaxies forming within the dense environment of simulated clusters will be often employed likewise as a diagnostic to investigate the host cluster features. We expect galaxies to exhibit differences between those populating the innermost region compared to the satellite population. We will see, in the next sections, that all these systems have interesting phase-space properties.

2.2 PROPERTIES OF THE GALAXY POPULATION

Regardless of the ambiguity in the definitions of "central" and "satellite" galaxies both in simulations and observations (Werner et al., 2022), it is undeniable that environmental effects have different impacts at different scales in galaxy clusters. Observations have established that the dense

galaxy cluster environment impacting the galaxy population (as opposed to the field hosting isolated galaxies) extends out to 2 – 3 virial radii, much further than previously assumed (Balogh et al., 2000; Solanes et al., 2002; Verdugo et al., 2008; Wetzel et al., 2012). In the galaxy population itself, we find examples of these differences: central galaxies (i.e., BCGs) are long recognized as different from the other cluster member galaxies, possibly due to the privileged position at the halo centre whereby galaxy interactions are the most frequent. A satellite galaxy, on the other hand, is subjected to various environmental effects, such as ram-pressure stripping, tidal stripping and strangulation that can strip its associated gas reservoir, thereby suppressing star formation. In addition, satellite galaxies are also subjected to various interactions with the host halo and other galaxies in the halo, which can lead to a morphological transformation. Note that mergers among satellites are expected to be rare. To further understand these differences, we will review some of the most interesting results for central (in Section 2.2.1) and satellite galaxies (in Section 2.2.3).

Furthermore, in recent years, observational studies in both local universe (Feldmeier et al., 2004; Mihos et al., 2005; Doherty et al., 2009; Kluge et al., 2020) and intermediate redshift (Feldmeier et al., 2004; Zibetti et al., 2005; Gonzalez et al., 2007) have demonstrated the existence of two different stellar components in galaxy clusters. A substantial fraction of stars is confined within the central galaxy while the other fraction – accounting between 10 per cent (Zibetti et al., 2005) and 35 per cent (Gonzalez et al., 2007) of the total stellar mass in clusters – is not gravitationally bound to any particular galaxy and constitutes the so-called ICL. Most of the incurring differences have arisen from the different formation channels and environments.

In the following, we will describe a few observables in these three classes of objects (i.e., BCG, ICL, and satellite galaxies) and we will give some insights into the theoretical predictions from simulations.

2.2.1 *Brightest cluster galaxy*

Brightest Cluster Galaxy (BCG) is a peculiar family of objects: being the most luminous (and typically the most massive) galaxies in the Universe, they are often located at the bottom of the gravitational potential of galaxy clusters (e.g. Lin and Mohr, 2004; Von Der Linden et al., 2007). Because of their privileged position, their properties are severely influenced, and in turn heavily affect, by the extreme environmental conditions of cluster centres, the latter being sites of interesting evolutionary phenomena (e.g. dynamical friction, mergers, galactic cannibalism). The singular properties shown by BCGs strongly support the idea of them having undergone a distinct formation history compared to the rest of the galaxies in a cluster (Dressler, 1980). Also employed as standard candles (Postman and Lauer, 1995), they include the most massive galaxies of the cluster. Furthermore, the evolution of the host halo is thought to be tightly connected to that of the BCG, since their mutual alignment is predicted to be present since

at least $z \leq 2$ (Ragone-Figueroa et al., 2020). BCGs often host the largest SMBHs in the Universe, whose presence is usually responsible for intense star formation histories due to the interplay between AGNs and stellar feedback (e.g. Chen et al., 2018). Likewise, BCGs represent the dominant population at the massive end of the galaxy luminosity function. It is precisely for these reasons that the resulting properties (e.g. high luminosity, extended stellar envelope, quenched star formation) are influenced both by their large masses and the surrounding cluster environment (Von Der Linden et al., 2007). Both factors combined make them attractive targets to benchmark models of galaxy formation.

To understand how the hierarchical growth of collapsed halos regulates the observational properties of galaxies hosted in such halos, we often rely on the correlation between the properties of the central galaxy and the host DM halo. Among these, the correlation between the BCG stellar mass and the cluster mass has been studied over the years in both observational (e.g. Bellstedt et al., 2016; Kravtsov et al., 2018; Erfanianfar et al., 2019) and theoretical works (e.g. Bahé et al., 2017; Pillepich et al., 2018; Ragone-Figueroa et al., 2018). Indeed, several theoretical models based both on semi-analytical models of galaxy formation (e.g. De Lucia and Blaizot, 2007) and hydrodynamical simulations (e.g. Ragone-Figueroa et al., 2018; Pillepich et al., 2018) suggest that the mass growth of the BCGs at late times is controlled by hierarchical accretion rather than in-situ star formation. However, the debate about the mechanism of BCG formation is not yet settled, while only recently simulations have started to provide predictions that agree with observational measurements (Ragone-Figueroa et al., 2018). An important factor in the BCG evolution is the role played by the AGN feedback which regulates the shape and the star formation rates of these massive objects. If radiative cooling were the only process active within cluster cores it would cause a rapid condensation of cold gas within the inner regions of $100 - 1000 M_{\odot} \text{ yr}^{-1}$ (e.g., White et al., 1997; Fabian et al., 2022). Such high cooling rates would imply blue colours and highly star forming BCGs, while current measurements suggest star formation rates of $\sim 1 - 10 M_{\odot} \text{ yr}^{-1}$ for massive clusters (Fraser-McKelvie et al., 2014). The absence of such high star formation rates is explained by invoking a heating feedback mechanism, now identified as the feedback by the AGN, at the centre of galaxy groups and clusters. This feedback prevents catastrophic cooling by injecting energy in the intracluster medium in a self-regulated fashion, such that cooling and heating energies balance each other within 10 per cent (see Fabian, 2012, for a review). Observationally, the presence of an AGN at the centre of clusters can be inferred from the detection of X-ray cavities around the cluster core, which in turn are commonly interpreted as the effect of energetic outflows of radio plasma from the central AGN that displaces the hot X-ray emitting gas.

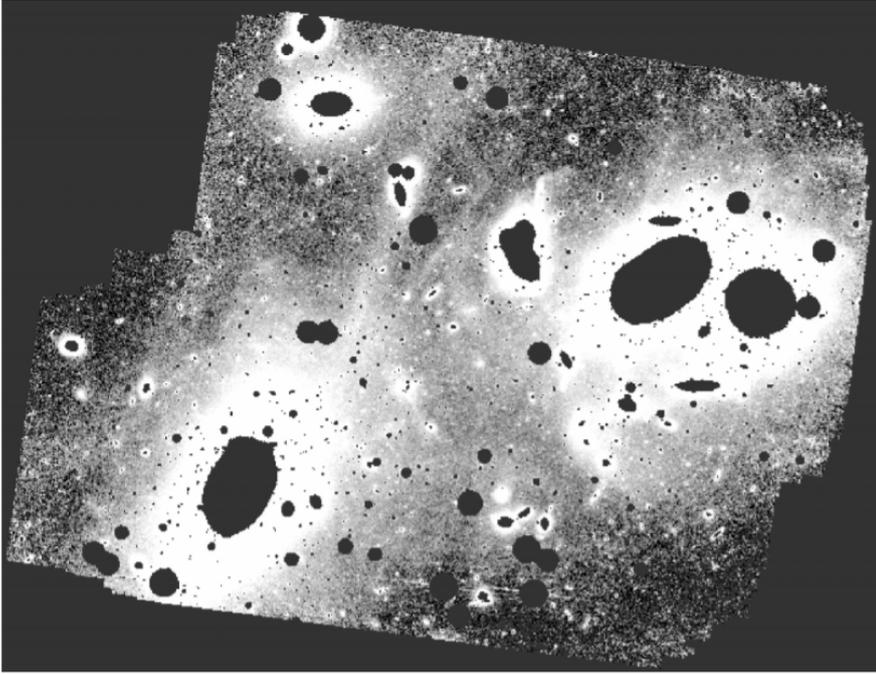


Figure 2.2: ICL in the Virgo cluster core. Black areas are the masked galaxies in the cluster while the strength of the white is related to the magnitude of the stars. Credits: Mihos et al. (2005)

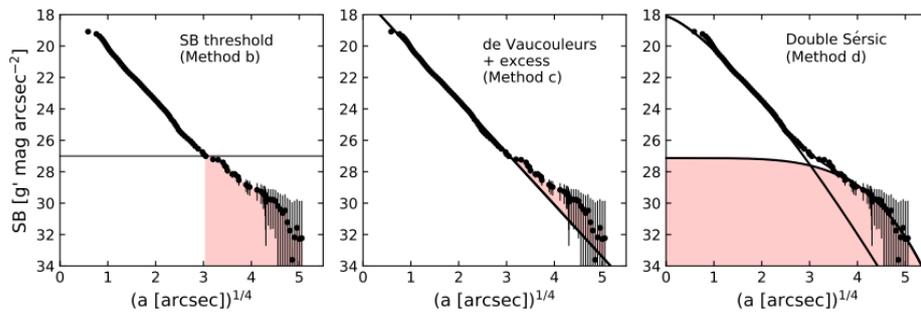


Figure 2.3: Decomposing the ICL in A1668 (black dots) based on surface brightness methods. The coral-shaded region is the ICL found. Method b (left panel) defines all light fainter than a surface brightness threshold (horizontal line at $27 g' \text{ mag arcsec}^{-2}$) as ICL. Method c (central panel) takes de Vaucouleurs profile (solid black) to fit the BCG light and assigns the excess to ICL. Method d (right panel) is the double Sérsic decomposition (two black lines) where the outer profile is the ICL. Credits: Kluge et al. (2021)

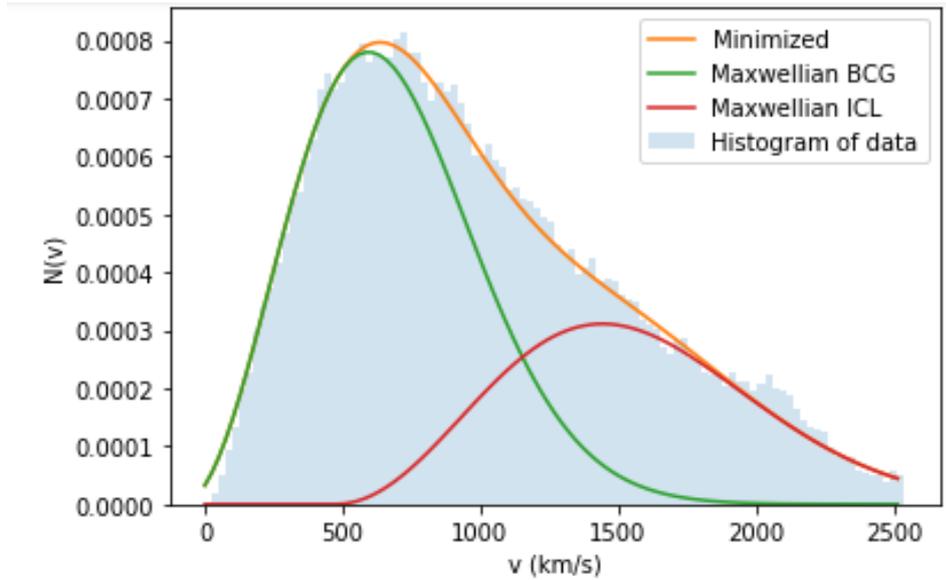


Figure 2.4: The velocity distribution of the stars in the main halo in a simulated cluster in the DIANOGA set (Bassini et al., 2020). The shaded histogram represents the data points which are fitted by the double Maxwellian (orange line). The singular components are given by the green (BCG) and red lines (ICL).

2.2.2 The intracluster light

The stellar halo bound to the cluster’s potential, the so-called ICL, has interesting properties worth discussing separately. The ICL is observed to be distributed around the central galaxy of clusters and extends to several hundred kiloparsecs away from the cluster centre (e.g. Zibetti, 2007; Murante et al., 2004). In the Virgo cluster core (Figure 2.2), the ICL is shown in contrast to the stars gravitationally bound to galaxies (masked in black in the image). The physical scales of the ICL resemble those of the DM distribution (e.g. Dubinski, 1998), so it is reasonable to expect that this component may trace the global gravitational potential of its host cluster (Montes and Trujillo, 2018, 2019; Montes, 2022). A natural question then is why the ICL is such an accurate tracer of the underlying mass distribution. The answer to this question is connected with the origin of the ICL. The stripped stars forming the ICL are not bound to any particular galaxy but to the cluster itself. Furthermore, both the ICL and the DM behave as collisionless components following the same gravitational potential. On the contrary, hot intracluster plasma is highly collisional, and consequently, its dynamics and spatial distribution could in principle differ from those of the collisionless components. As shown by Montes and Trujillo (2018), the ICL can be used to reconstruct the gravitational potential of the cluster as accurately as it is done already with the gravitational lensing.

Observationally, the ICL is usually detected using photometry in various bands albeit the identification remains difficult and uncertain. The major problems include its extremely low surface brightness (e.g. Mihos et al.,

2005; Zibetti, 2007) and contamination by foreground and background galaxies. The assumed total amount of ICL per cluster depends on how the surface brightness profile is separated from the BCG in the inner regions and extrapolated beyond the outermost measurable radius. The uncertainty of the extrapolation was examined in Kluge et al. (2020), while Kluge et al. (2021) attempted to quantify differences in the photometric separation between BCG and ICL by applying different methods reported in the literature, as shown in Figure 2.3. The surface brightness profile of A1668 is decomposed in several ways to obtain the ICL contribution. Going from left to right panel, the ICL (coral shaded region) is estimated with a threshold brightness (in method b), fitting a de Vaucouleurs in the inner region and taking the excess (method c) and fitting a double Sérsic profile (method d). It is straightforward to notice that the three approaches provide significantly different ICL fractions, demonstrating there is no unique and self-consistent definition of ICL in the literature. Recently, there has been a push for more complex algorithms aiming at separating galaxies and diffuse light in two dimensions. This has been done using 2D fitting algorithms like GALFIT to most of the galaxies in the cluster (Giallongo et al., 2014; Morishita et al., 2017) or more sophisticated wavelet-like decomposition techniques (Da Rocha and De Oliveira, 2005; Jiménez-Teja and Dupke, 2016; Ellien et al., 2021).

Simulations predict that the ICL forms relatively late ($z < 1$; e.g., Monaco et al., 2006; Murante et al., 2007; Contini et al., 2013) and it is thought to form primarily by the tidal stripping of stars from galaxies during the hierarchical accretion history of the cluster (e.g., Murante et al., 2004; Contini et al., 2013). If this scenario is correct, the colour, metallicity, spatial distribution, and surface brightness of the ICL should reflect the properties of galaxies from which the intracluster stars originated, effectively encoding the formation history of the cluster. Therefore, the characterisation of the ICL provides a direct way of determining the assembly mechanisms occurring inside galaxy clusters.

The main issue with the analysis of the diffuse component in numerical simulation is often its very identification. Rudick et al. (2006) firstly suggested building simulated surface brightness maps and detecting an ICL component with a scheme similar to that employed with observational data. This has the advantage of being somewhat easier to compare with observations, at the cost of losing the intrinsic advantage of simulations, that is, the knowledge of the dynamics of each star particle. The other possibility is to benefit from the dynamical distinction between the two stellar components, in terms of the gravitational boundedness to the cluster potential, as it has been performed in Murante et al. (2004), Dolag et al. (2010) and Marini et al. (2021b). These differences can be further exploited by using ML techniques to disentangle the two components, as is presented in Chapter 7. The only disadvantage may arise because such distinction may or may not coincide with the observational definition of the ICL, but clearly, this method has its benefits. Observations are in agreement with such scenario: integrated light spectroscopy (Dressler, 1979; Kelson

et al., 2002; Edwards et al., 2016; Gu et al., 2020) and planetary nebulae kinematics (Arnaboldi et al., 1996; Longobardi et al., 2018) of nearby BCGs show that the radial velocity dispersion increases with radius reaching the velocity dispersion of the galaxies of the cluster. All this together is suggestive that this extended component is kinematically distinct, as confirmed in the nearby Virgo cluster (Longobardi et al., 2015).

This dynamical behaviour of the ICL is clearly visible in simulations. While the velocity dispersion of the stars in the BCG has values similar to those of the stars in galaxies, the ICL presents velocity dispersion values that are much larger and comparable to those of the DM halo (e.g. Dolag et al., 2010; Longobardi et al., 2015; Mihos et al., 2016). In Marini et al. (2021b, 2022), we show that these features are a fundamental difference between the two stellar components. As an example, in Figure 2.4 we present the results obtained for the velocity distributions of the stars belonging to the main halo of one of our simulated clusters: we show that the distinction between the two populations of stars associated with the BCG and the ICL is in agreement with our data. Thus, this motivates us to proceed in considering the two components as two different dynamical populations coexisting in the same environment. The importance of this result is strictly related to our interest in studying the dynamical properties of the phase-space of galaxy clusters: new possibilities open up in our analysis as now we can also investigate the features of the phase-space of these two stellar components.

2.2.3 Cluster galaxies

If we now focus on the other galaxy members in a cluster (i.e., satellites), we observe similar formation histories and properties. Let us discuss some of them.

2.2.3.1 Internal structure of DM and galaxies

The inner structure of DM halos reveals information about the past history of halo assembly. To a first approximation, we can model a DM halo as a spherical object. In this case, the internal mass distribution is fully described by a density profile. As discussed above, different halos have different formation histories, so we may expect a significant halo-to-halo variation in density profiles. On the other hand, DM halos are highly non-linear objects, and it may be that information regarding their past histories has largely been erased by their non-linear collapse. In the latter case, density profiles may be more closely related to the violent relaxation process than to initial conditions. This section describes a few of these properties in the coarse-grained phase-space structure.

DENSITY PROFILES. According to CDM cosmological numerical simulations, the radial mass distribution of DM halos (see Merritt, 1987) and member galaxies follows a universal trend in the cluster population as a function of the radial distance at least out to the virial

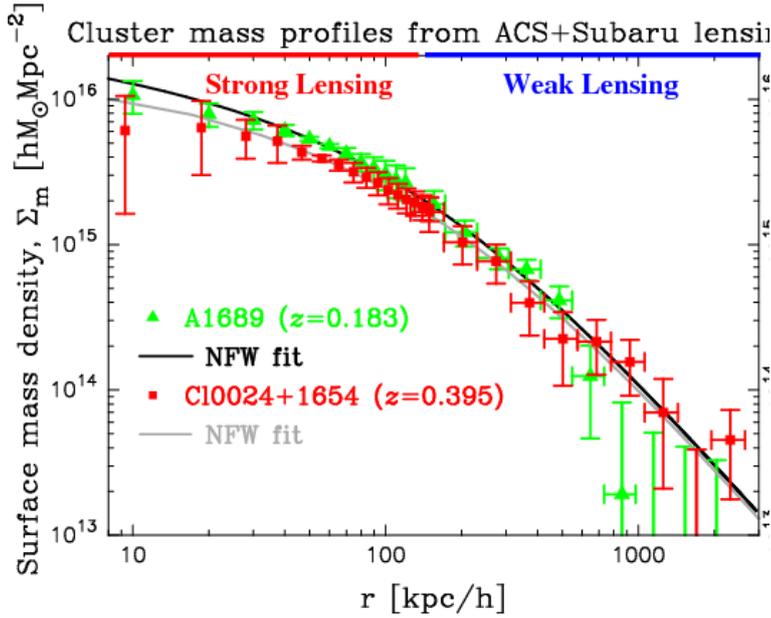


Figure 2.5: Cluster surface mass density profiles for A1689 (green triangles) and CL0024 + 1654 (red squares) reconstructed from strong lensing (HST/ACS: $r < 200$ kpc) and weak lensing measurements (Subaru: $r > 200$ kpc). Solid curves are the best-fit NFW models, with a continuously steepening density profile, for A1689 (black) and CL0024 + 1654 (grey). Credits: Umetsu and Broadhurst (2008).

radius. Despite findings in early works asserting the trend to be simple power-laws $\rho(r) \propto r^{-\gamma}$, nowadays it is generally agreed that the universal profile steepens monotonically with the radius being shallower than an isothermal profile (i.e. $\gamma < 2$, Mo et al., 2010) near the centre but steeper around the virial radius. At least in the central part, this universal behaviour can be explained by the initial fast assembly phase of halos (Huss et al., 1999; Wang and White, 2009; Lapi and Cavaliere, 2011) in which dynamical processes such as violent and collective relaxations occur repeatedly (Henriksen, 2006; Lynden-Bell, 1967); whereas the following slower accretion may be responsible for the outer slope of the density profile (Hiotelis and Del Popolo, 2006). Navarro (1996) and Navarro et al. (1997) – hereafter NFW – encouraged the use of a function described by the following expression:

$$\rho_{\text{NFW}}(r) = \frac{\rho_0}{x(1+x)^2} \quad (2)$$

whereby $x = r/r_s$ with r_s corresponds to the clustercentric distance at which the slope of the NFW profile is equal to -2 and ρ_0 is a normalisation. As shown in Figure 2.5, the profile is characterised by a composite power-law with exponent -1 in the inner region and -3 on the outskirts providing an adequate fit to the mass distribution of relaxed clusters reconstructed with different methods (Biviano and Girardi, 2003; Rines et al., 2003; Diemer and Kravtsov, 2015).

PHASE-SPACE DENSITY PROFILE. Further information on the dynamics of DM halos can also be obtained from the coarse-grained phase-space of galaxy clusters. Taylor and Navarro (2001) argued for the first time that the gravitational assembly in halo formation leads to a simple power-law radial stratification of the phase-space density given by a sort of phase-space density:

$$Q(r) = \frac{\rho(r)}{\sigma^3(r)} \propto r^{-\alpha}, \quad (3)$$

with $\rho(r)$ the density profile and $\sigma(r)$ the velocity dispersion profile. Similarly, the radial phase-space density $Q_r(r)$ is computed using the radial projected velocity dispersion profile $\sigma_r(r)$. Later on, Dehnen and McLaughlin (2005) found a similar trend in the phase-space density profiles when conditions for dynamical equilibrium of the system hold and there is the validity of a linear relationship between the orbital anisotropy parameter β and the logarithmic slope of the density profile γ . The existence of this relation has been found by Hansen and Moore (2006) to hold in a variety of halos extracted from numerical simulations,

$$\beta(r) = -0.15 - 0.19 \gamma(r).$$

Observations and theoretical predictions tell us that in the regime of dynamical equilibrium the phase-space density profile in a DM halo is well approximated by a power-law of slope $\alpha = 1.84$ while radial phase-space density profiles $Q_r(r)$ have slopes with $\alpha = 1.92$ (Dehnen and McLaughlin, 2005; Biviano et al., 2013, 2021). This is true for passive galaxies, as they have undergone violent relaxation and have reached dynamical equilibrium while star forming galaxies have marginal deviations (see Figure 2.6). Another recent study (Aguerri et al., 2017) showed that $Q_r(r)$ follows the theoretical predictions regardless of the galaxy luminosity or colour, concluding that different populations of galaxies within clusters all consistently trace the dynamics of a virialised object. Capasso et al. (2019) found the same results arguing that galaxies behave approximately as collisionless particles and are as relaxed as the DM particles in halos forming within cosmological structure formation in simulations. In conclusion, it seems that the power-law behaviour of the phase-space density may be the key driver in the formation of clusters (or more generally of DM halos).

FROM PHASE-SPACE DENSITY TO PSEUDO-ENTROPY. An equivalent way of looking at this is by focusing on a quantity closely related to phase-space density, i.e. "pseudo-entropy". In X-ray astronomy, it is customary to define the entropy of the intracluster plasma as (Voit et al., 2005):

$$S_X = \frac{T}{n_e^{2/3}} \quad (4)$$

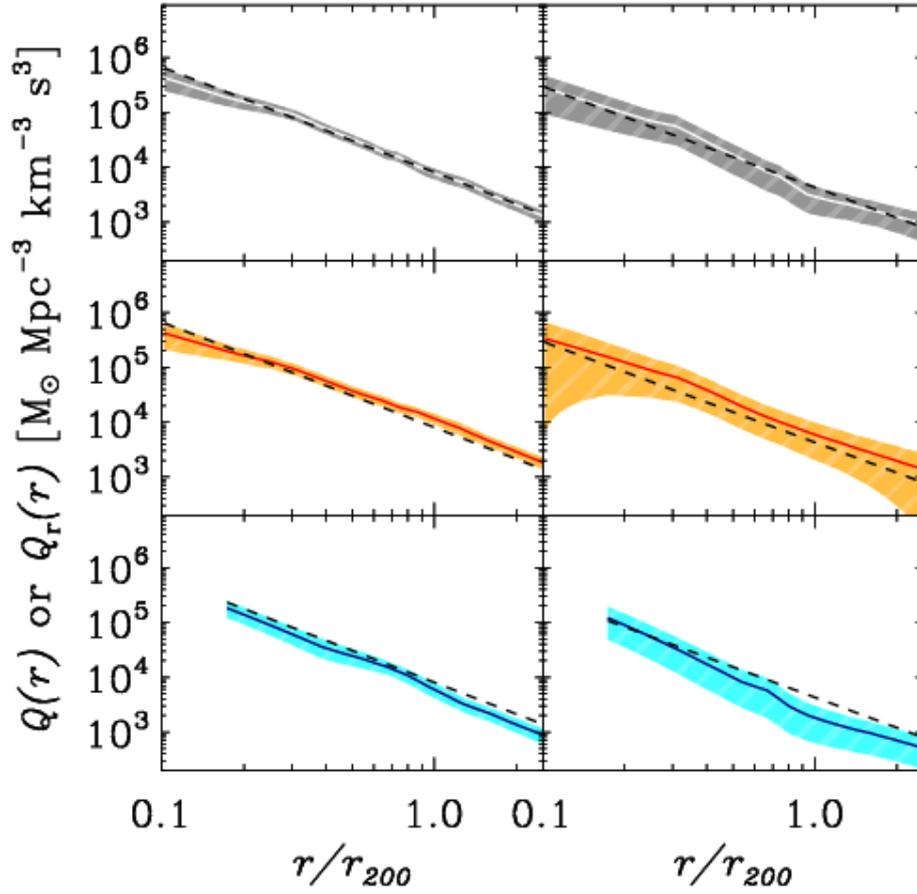


Figure 2.6: Study of the phase-space density profiles $Q(r) \equiv \rho/\sigma^3$ (left panels) and $Q_r(r) \equiv \rho/\sigma_r^3$ (right panels) as a function of radius, within 1σ confidence regions (shaded area) for (from top to bottom) all, passive and star forming galaxy members of the cluster MACS J1206.2 – 0847 (redshift $z = 0.44$): The mass distribution is well fitted by a NFW density profile with mass $M_{200} = (1.4 \pm 0.2) \times 10^{15} M_\odot$ and concentration $c_{200} = 6 \pm 1$. Data were mostly collected by ESO VLT, NASA Hubble Space Telescope, and NASJ Subaru Telescope. Solid lines are the best-fit results from Dehnen and McLaughlin (2005). Credits: Biviano et al. (2013).

where T is the temperature and n_e is the electron number density directly linked to ρ_{gas} for a fully ionised plasma. So defined, the quantity S_X describes the thermodynamic properties of the intracluster medium, i.e. the dominant collisional component of galaxy clusters. In analogy, we can define a "pseudo-entropy" for the collisionless component, identifying the temperature with the velocity dispersion of the collisionless tracers,

$$T \propto \sigma^2$$

so by using this equality, we can recast [Equation 4](#) as follows:

$$S(r) = \frac{\sigma^2(r)}{\rho^{2/3}(r)}. \quad (5)$$

We remind here that ρ is the mass or number density of the tracer. This is the analogue to the thermodynamic entropy for a collisionless system. By comparing [Equation 3](#) to [Equation 5](#), we notice that $S(r) = Q^{-2/3}(r)$. As expected, the higher the phase-space density, the lower the entropy of the system. In [Chapter 5](#), we will use $S(r)$ or $Q(r)$ interchangeably for our analysis.

2.2.3.2 Dynamical properties

When speaking of the dynamical properties of galaxies, one can be referring to two possible observables: *(i)* the dynamics followed by galaxies as a statistical sample of tracers in a cluster or *(ii)* the inner dynamics of the single galaxy. In this section, we provide an example of these two cases. Indeed, virialised galaxies are a useful tracer of the dynamical evolution of clusters. Galaxies should enter clusters on fairly radial orbits. In the inner regions, the two relevant time-scales, i.e. the two-body relaxation time and the crossing time, are both shorter than the age of the Universe. Hence, here galaxies should have forgotten their initial trajectories and the inner population should have isotropic velocities. At this point, galaxies behave as collisionless tracers of the gravitational potential: they will obey the Jeans equation having the density and velocity distribution in equilibrium with the gravitational potential. In principle, we could infer the total mass profile of clusters from the measurements of the projected number density of galaxies and their line of sight velocities under the assumption of equilibrium and spherical symmetry. Nevertheless, it is also necessary to assume a fitting profile to their orbital anisotropy profile β . This generates the so-called "mass-orbit degeneracy" (see Merritt, 1987 for further details).

Conversely, the internal dynamics of galaxies may be left with traces of the environment in which the galaxy is embedded (and thus, of the cluster environment). As first suggested by Hoyle (1951), the spin of a proto-galaxy arises from the tidal field of the neighbouring structures, although the late stages of the non-linear collapse and mergers with other halos may change the motion. This has early prompted detailed studies into angular momentum growth within the framework of the gravitational instability (e.g. Peebles, 1969; Doroshkevich, 1970; White, 1984).

ORBITS. During their evolution clusters grow by mergers. In the limit of very minor mergers, clusters are relatively isolated systems accreting individual galaxies, whose orbits should be fairly radial on their first infall. Conversely, during major mergers, galaxies suffer violent relaxation that isotropises their orbits. Moreover, the angular momentum of the merging cluster will be transferred to individual galaxies, which may lead to an excess of more circular orbits. The measurement of the elongations of galaxy orbits is therefore a diagnostic to understand the formation of clusters, and how galaxy orbits vary with cluster mass, elongation, and large scale environment. Mamon et al. (2019) investigated how orbital elongations depend on the morphological type, revealing that spiral galaxies yield outer radial orbits before changing morphological type. Early type galaxies are found on less radial orbits possibly related to the longer time spent by these galaxies in the cluster. On the simulations side, Lotz et al. (2019) found that star forming satellite galaxies prefer radial orbits, while their star formation is quenched during their first passage. This is in agreement with observations, which also find that galaxies experience strong quenching during their first passage through the cluster (Mahajan et al., 2011; Wetzel et al., 2013; Jaffé et al., 2015, 2016; Oman and Hudson, 2016).

Orbiting galaxies allow to determine the *splashback radius*, a dynamical radius defining the size of the host halo (Adhikari et al., 2014; More et al., 2015, 2016). In principle, there exists a physical boundary to a halo if one takes into account the continuous stream of particles accreting the halo. We define this boundary as the distance at which the most recently accreted matter reaches the farthest point in its orbit after it falls into its cluster (i.e., the apogee). In other words, this radius separates infalling material from orbiting with a caustic, which is an infinitely sharp drop in density. In the spherical approximation, this distance divides the space into two distinct regions. The inner region has objects that have DM shells and streams that are in orbits that constantly cross each other; the exterior region has still infalling matter that has never completed a single orbit within the halo potential. For this reason, the velocity dispersion of matter at this radius is expected to average zero. Several observations have detected the splashback radius in clusters at different wavelengths (Umetsu and Diemer, 2017; Zürcher and More, 2019; Shin et al., 2019).

ANGULAR MOMENTUM Angular momentum, in addition to the mass, arguably governs most stages of galaxy formation and evolution (Fall and Efstathiou, 1980). From the primordial tidal torque exerted on galaxy-size DM halos which can lead to the formation of disk-like structures (Hoyle, 1951; Peebles, 1969) to the consequent morphology of present-day galaxies (Romanowsky and Fall, 2012; Fall and Romanowsky, 2013), the angular momentum \vec{J} , or specific angular

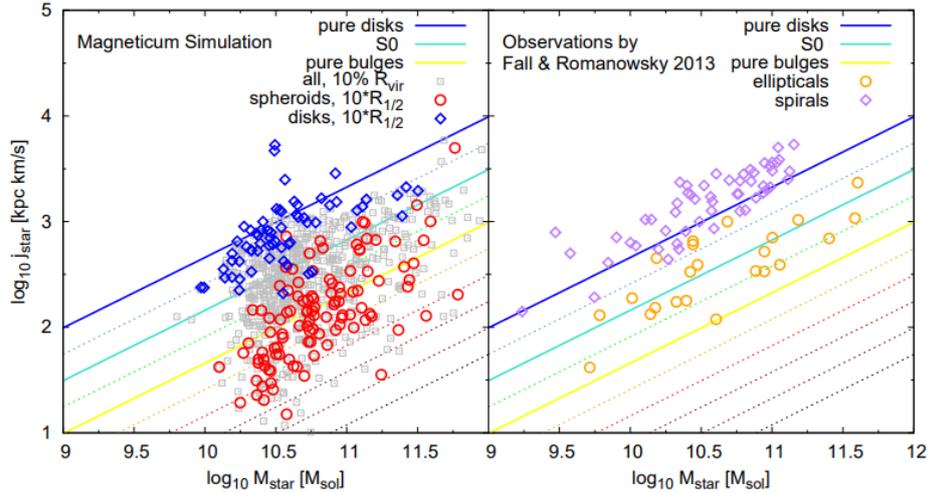
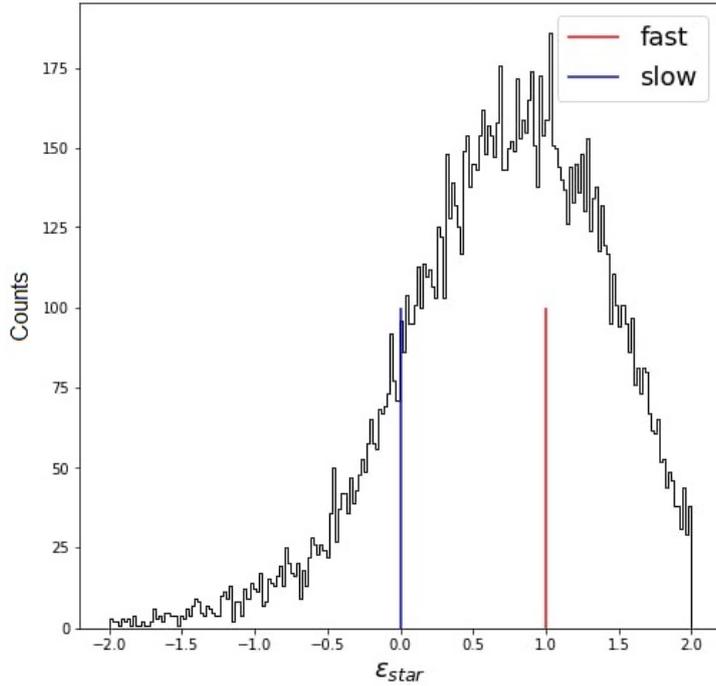


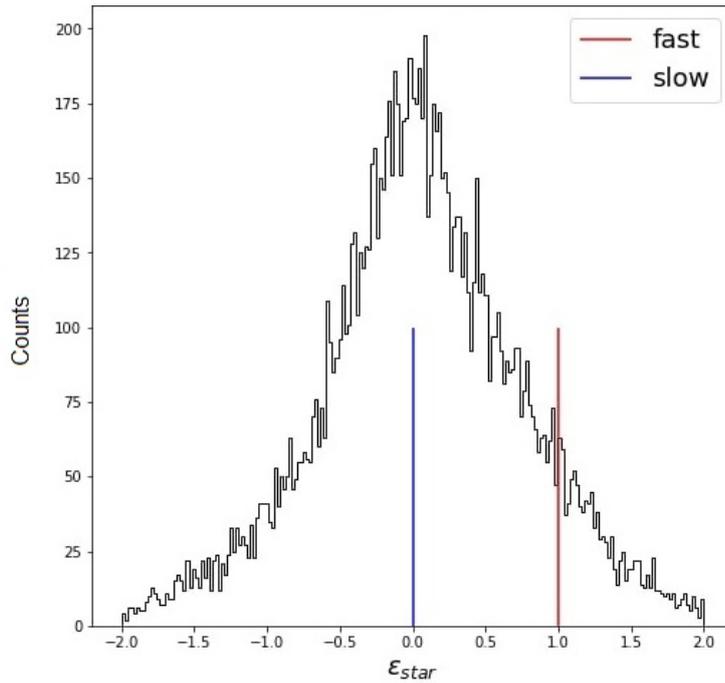
Figure 2.7: Log-log plane of the stellar mass versus stellar specific angular momentum in a galaxy sample. Left panel: data points from the Magneticum simulations at redshift $z = 0.1$; right panel: the observed sample from Fall and Romanowsky (2013). The coloured solid lines have a slope around $0.5 - 0.6$ and represent the kinematical distinction between disks and bulges. The markers label the morphological distinction between elliptical (early types) and spirals. Credits: Teklu et al. (2015)

momentum \vec{j} (the angular momentum in units of mass), represents a basic ingredient to include in galaxy models. The general picture emerging from observational studies (Romanowsky and Fall, 2012; Fall and Romanowsky, 2013) supports the claim that spiral galaxies define a tight sequence in the $j_{\star} - M_{\star}$ (i.e., total stellar angular momentum versus stellar mass) plane, whereas ellipticals (or bulges) tend to follow a similar trend but with a lower intercept, as reported on the right-hand panel of Figure 2.7. The transition among the different types is continuous. Notice that the distinction between spirals and early types can also be traduced into a kinematical decomposition of galaxies composed of a dispersion-supported component (bulge) or the rotation-supported one (disk) (Fall and Romanowsky, 2018). Such interesting results can be compared with simulated galaxies (left-hand panel in Figure 2.7) to analyse the accuracy in reproducing these small-scale features in a cosmological framework, such as the one described by the Magneticum simulations in the plot (Teklu et al., 2015).

In this regard, we tested the galaxy evolution model in our simulations by performing a similar exercise with the simulation set in our hands (see Chapter 3): we investigate whether we can also predict this morphology-kinematics connection. Following Teklu et al. (2015),



(a) Fast rotator



(b) Slow rotator

Figure 2.8: An example of two galaxies from the DIANOGA simulation set divided according to their kinematical properties: fast (top) and slow (bottom) rotator. The histograms report distributions of the star particle circularities ϵ (see text) for the two galaxies. A fast rotator has this distribution peaking at $\epsilon \sim 1$ (vertical red line), while for slow rotators the distribution peaks at $\epsilon \sim 0$ (vertical blue line).

we define the specific angular momentum \vec{j}_k of the particles in a galactic halo:

$$\vec{j}_k = \frac{\sum_{i \in N_k} m_{k,i} (\vec{r}_{k,i} \times \vec{v}_{k,i})}{\sum_i m_i}, \quad (6)$$

where N_k is the number of particles in the k^{th} specie (i.e., gas, star, and DM). In the equation, we define the mass $m_{k,i}$, the position $\vec{r}_{k,i}$, and the velocity $\vec{v}_{k,i}$ of the particle i^{th} . From the $j_* - M_*$, we identify several galaxies at the two extremes of the distribution (i.e., disk or bulge galaxies). For each of them, we compute the circularity ϵ characterised by the ratio

$$\epsilon = \frac{j_z}{j_{\text{circ}}} = \frac{j_z}{r V_{\text{circ}}}, \quad (7)$$

i.e., the specific angular momentum in the direction perpendicular to the disk j_z over the specific angular momentum of a reference circular orbit j_{circ} . This is redefined in terms of $V_{\text{circ}} = \sqrt{GM(< r)}/r$ being the circular velocity of each particle at the distance r from the galaxy centre. In a dispersion-dominated system, there is usually a broad peak in the distribution at $\epsilon \simeq 0$, while in a rotation-supported system there is usually a broad peak at $\epsilon \simeq 1$ (Teklu et al., 2015), as shown in Figure 2.8. One of the most striking examples in our simulated galaxies is shown in Figure 2.9: we selected one galaxy among the fast rotators at $z = 0$ and investigated its morphological properties in the *gri* bands with the filter bandpasses from the HST¹. In the left panel, we show the galaxy taken face-on: we can recognize the bulge at the centre and the surrounding disk at a larger radius. In the right panel, we see the galaxy observed edge-on exhibiting the thin disk. The general structure is remarkably similar to that of a disk-like galaxy and considering the blueness of the image, we can expect high star formation rates both in the bulge and disk. We point out that these high values of star formation in the centre are quite different from observed ones in the local Universe. However, high star formation rates (especially in central galaxies) which cause very massive BCGs are a well known problem in simulations (see Section 2.2.1).

2.2.3.3 Other statistical properties of the galaxy population

Galaxies are observed to have different intrinsic properties (luminosity, morphology, size, mass, nuclear activity, etc.) which, in turn, are shaped by multiple processes playing a role during the subsequent evolution. Here, we will further address the question of how galaxies are statistically distributed with respect to such intrinsic properties, providing examples from state-of-art simulations and observations.

¹ Hubble Space Telescope (HST): <https://hubblesite.org/>

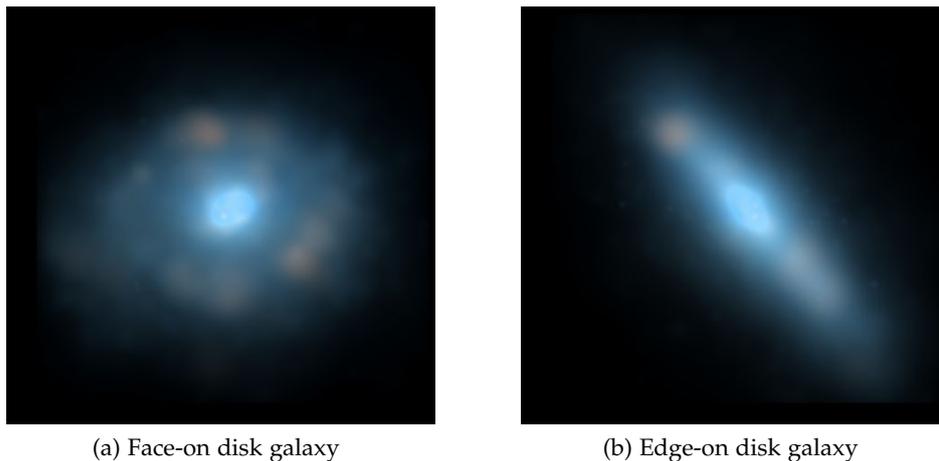


Figure 2.9: An example of a disk galaxy from the DIANOGA simulation set taken from two different lines of sight. The galaxy is chosen among the fast rotators, thus we expect it to have properties similar to spiral/disk-like galaxies. The simulated galaxy is coloured in RGB using the HST’s *gri* filters. Left panel: the galaxy is taken face-on. The bulge at the centre is the bluest region in the cluster. We can observe a disk-like structure at larger radii. Right panel: the galaxy is taken edge-on. We can still observe a central bulge and a thin disk surrounding it.

LUMINOSITY AND MASS FUNCTION. The luminosity function reflects the internal processes taking place in a cluster relative to the environmental effects on the member galaxies or the presence of active star formation. The luminosity function of galaxies $\phi(L)$ is defined as (Mo et al., 2010):

$$dn(L) = \phi(L) dL \quad (8)$$

where $dn(L)$ is the comoving number density of galaxies with luminosity L in dL , in a given band. Schechter (1976) proposed an analytical expression which reads:

$$\phi(L) dL = \phi_* \left(\frac{L}{L_*} \right)^\alpha \exp^{-\frac{L}{L_*}} \frac{dL}{L_*}. \quad (9)$$

Here L_* is a characteristic luminosity, while ϕ_* is the galaxy number density in the unit volume at luminosity L_* . The exponent α is a constant which indicates the slope of the relation at the faint end. Average values for the Schechter luminosity function are $\alpha = 0.7$ and $\phi_* = 0.02 h^3 \text{ Mpc}^{-3}$ (Longair, 2007). Notice that a large source of uncertainty is present at the very faint end, mainly because in a magnitude-limited sample these galaxies can only be observed within a very small, local volume so that the errors are dominated by cosmic variance (Mo et al., 2010). Contrary to the luminosity function, the stellar mass function (a fundamental tool used in interpreting the evolution of galaxies) is not directly observable. Thus, it is customary to estimate the stellar mass function from multi-band

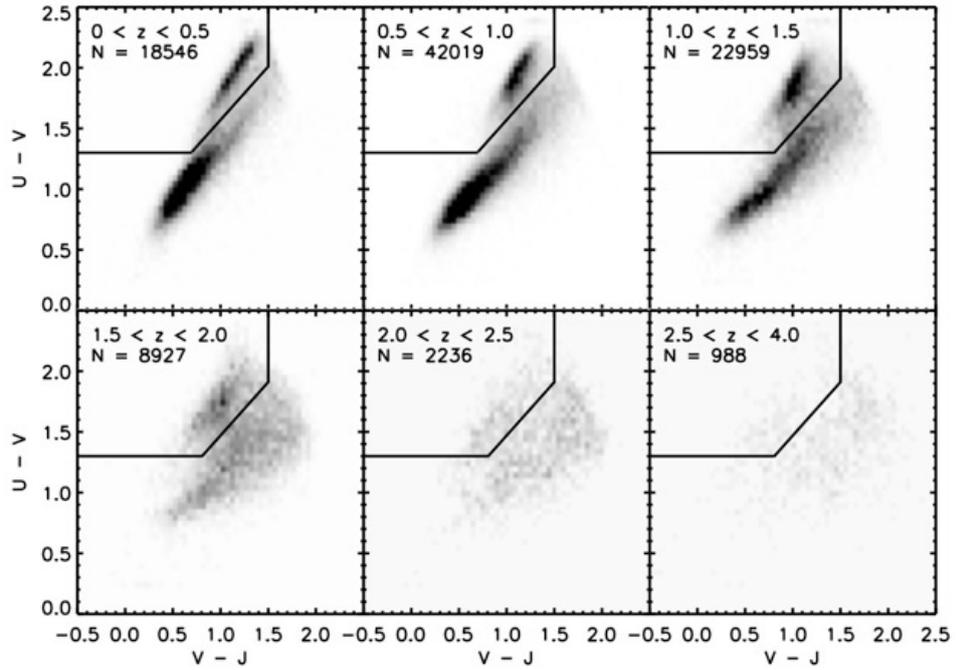


Figure 2.10: UVJ colour-colour diagram in the $[0, 4]$ redshift interval for massive galaxies in the COSMOS/UltraVISTA field. The bimodality in the galaxy population is visible up to $z = 2$. The solid lines mark the cuts applied to separate star forming (top-left) from passive (bottom-right). Credits: Muzzin et al. (2013).

photometry (and, possibly, high-quality spectroscopy) parametrising the stellar spectrum of the single galaxies and thus estimating their masses (e.g., Merluzzi et al., 2010). This modelling requires a formal description of the stellar population synthesis model. More details on the implementation of this scheme in our simulations can be found in Section 3.6.1.

COLOUR – COLOUR DIAGRAM. The colour – colour diagram (and, similarly, the colour – magnitude diagram) is another diagnostic to describe the galaxy population in clusters. In principle, colour is a measure of how efficient is the star formation rate: the bluer is the galaxy the higher the relative number of blue newborn stars. Nevertheless, other factors contribute to determining the colour of a galaxy such as stellar metallicity or dust contamination. As shown in Figure 2.10, the galaxy population at low redshift exhibits a clear bimodal distribution in the colour-colour plane, at all redshifts. Indeed, many studies use this specific combination of rest-frame colours ($U - V$ versus $V - J$ in the so-called UVJ diagram) to discriminate star forming from passive galaxies (e.g., Daddi et al., 2004; Williams et al., 2009; Arnouts et al., 2013), and on this technique we will discuss more in the next section. We point out that just recently several cosmological simulations (Valentini et al., 2017; Nelson et al.,

2018; Donnari et al., 2019) have started to model galaxy luminosity and colour consistently with observations.

STAR FORMING AND PASSIVE GALAXIES. The population of galaxies can be broadly divided into two distinct types according to the star formation rate. Star forming galaxies have blue colours, typically disk-like morphologies and a relatively high star formation rate, whereas quiescent galaxies have redder colours, more spheroidal morphologies, and very low, if not absent, star formation. Generally speaking, star forming galaxies are abundant at relatively early times and at low redshift they tend to have lower average stellar masses (Peng et al., 2010; Muzzin et al., 2013).

A central question is how galaxies transform from actively star forming systems to passive quiescent; such question goes under two lines of research: "mass quenching" (Peng et al., 2010) or "environmental quenching" depending on whether these mechanisms are mass-dependent (e.g., due to feedback from SNe, galactic winds and AGN – Di Matteo et al., 2005; Oppenheimer et al., 2010) or they are mostly caused by external conditions. In the latter, we usually encounter: stripping of cold (ram pressure, Gunn and Gott III, 1972) and hot gas (strangulation, starvation; Larson et al., 1980), galaxy harassment, and dry mergers. Addressing these issues allows us to infer what is the role played by each process in quenching galaxies and the distribution of properties in the host clusters population. One additional scenario is presented when galaxies transform before their infall in the cluster region, and thus we speak of "pre-processing" as an additional quenching mechanism. Quantitative studies of passive and star forming populations, across different redshifts, give us useful insights into environmental quenching through stellar mass functions and quenching efficiencies. The latter is the fraction of passive galaxies in a cluster with respect to the passive fraction in the field: often it is used to determine the star formation history of a cluster across time. Figure 2.11 shows the stellar mass function from COSMOS/UltraVISTA field (in the top panels) of passive and star forming (red and blue respectively); in the bottom panels, the quiescent fraction is plotted as a function of log stellar mass. From left to right, the redshift intervals are [0.2, 0.5], [0.5, 1.0], [1.0, 1.5], and [1.5, 2.0]. In the hypothesis of environmental quenching/pre-processing of the star forming blue galaxies as driving mechanism, we expect the blue galaxies to represent the young population not yet influenced by the evolutionary mechanisms occurring in cluster environments. On the other hand, red galaxies, generally older and more evolved, have a stellar mass distribution which is significantly affected at both the low and high mass end. Accordingly, quenching fractions (in the lower panels) demonstrate the existence of a quenching mechanism linked to the stellar galaxy mass.

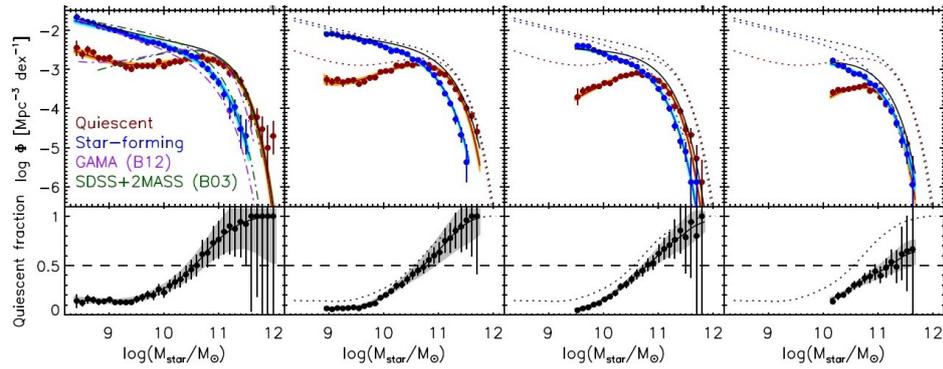


Figure 2.11: From left to right, the panels report results from different redshift intervals: $[0.2, 0.5]$, $[0.5, 1.0]$, $[1.0, 1.5]$, and $[1.5, 2.0]$. Top panels: stellar mass functions in the Muzzin et al. (2013) study for the quiescent (in red) and star forming (in blue) populations. Over-plotted in the $0.2 < z < 0.5$ bin are the stellar mass functions from Cole et al. (2001), Bell et al. (2003) and Baldry et al. (2012). In the remaining redshift bins, the dotted curve is the total stellar mass function from UltraVISTA in the $0.2 < z < 0.5$ bin. Bottom panels: quenching fraction as a function of the log stellar mass. Credits: Muzzin et al. (2013).

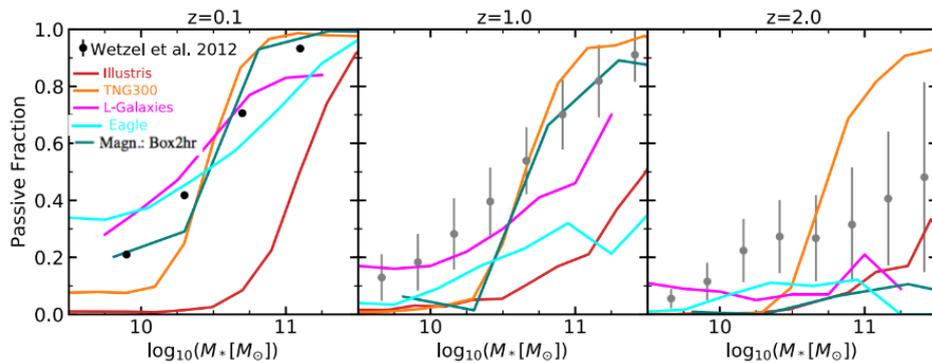


Figure 2.12: Passive fraction of the cluster galaxies in three different redshift bins. The black dots represent observational data from Wetzel et al. (2012) while the coloured curves mark the simulations. Credits: L-Galaxies (Henriques et al., 2013), Eagle (Furlong et al., 2015), Illustris and IllustrisTNG (Donnari et al., 2021), selected from a presentation's slides given in Favignana 2018 at the conference "Birth, life and fate of massive galaxies and their central beating heart".

Comparing these results with simulations is not a trivial task. In principle, to measure star formation rates of simulated galaxies we can count all gas elements and/or star particles (bound to the given galaxy according to the halo finder) which have formed in the last snapshot. However, this method often provides null star formation rates, since it is an instantaneous measurement (obtained from the particles in one snapshot). Another possibility is to account for one of the properties of Single Stellar Populations (SSPs), whereby for each star particle we have the redshift of formation: we sum the mass of star particles in a galaxy over the last Y years, where the value of Y (of the order of few hundreds Myr) depends on how continuous is the description of star formation provided by the simulation. With this approach, we can take into account a small interval of time when stars may have formed, rather than the number of those formed between the previous and present snapshot. According to these definitions, we can arbitrarily choose a criterion to determine if a galaxy is passive or star forming. In the literature, there are a handful of possibilities, we list a few of these conditions for a galaxy to be passive.

Distance from the main sequence. Galaxies in the $\text{SFR} - M_*$ \log - \log plane can be labelled as quenched if (i) the distance from the main sequence is more than 1 dex or (ii) more than 2σ , i.e. at least twice the standard deviation of the main sequence distribution. Notice that this definition can be applied in both observational and simulated data.

Specific star formation rate threshold. The specific star formation rate is the mass-weighted star formation of a galaxy and can be a proxy of quenching phases. Mostly, a $s\text{SFR} < 10^{-11} \text{ yr}^{-1}$ is a valid criterion to observe a bimodal distribution of galaxies (Donnari et al., 2019). Note that this case does not account for the evolution of the $\text{SFR} - M_*$ plane and it is valid only for a narrow range of redshifts. Similar approaches base their star formation threshold value to be a fraction of the Hubble time.

UVJ diagram. An extensively used tool in observations (and therefore, mimicked in simulations) is the colour bimodality occurring in the UVJ diagram (e.g., in Figure 2.10). The diagram shows a bimodality in the galaxy distribution observed up to $z = 2$ whereby galaxies located approximately on the top-left corner are passive, and star forming if located towards the bottom right. Assuming that dust attenuation is rightly taken into account, the connection between colours and star formation relies on the understanding that star forming galaxies (thus, with many young stars) tend to be bluer than passive galaxies which are populated by older stellar populations. In the plot, the solid lines mark the cuts.

Figure 2.12 shows the comparison among different state-of-art simulations on the predicted passive fraction of galaxies as a function of stellar mass. This comparison allows us to have insights into the quality of our predictions in simulations and how they compare with the other numerical schemes. For sake of simplicity, in this case, the fraction of quenched galaxies is determined by the specific star formation rate threshold method. Simulations vary for a number of factors, specifically Illustris², IllustrisTNG³, Eagle⁴ and Magneticum⁵ are hydrodynamical simulations, while L-Galaxies⁶ is a semi-analytic model. It is worth pointing out that different hydrodynamical simulations have different resolutions, box sizes, hydrodynamical solvers and, most important, different implementations of star formation and feedback from SNe and AGNs. A selection of observational data is indicated as black and grey symbols. The three panels mark three different redshifts in the simulations. At $z = 0$, IllustrisTNG and Magneticum are the best at reproducing the observed trend, although the low-mass end is underestimated by the former. At higher redshifts, in comparison to the other cosmological models, IllustrisTNG returns systematically higher quenched fractions. As opposed to the IllustrisTNG model, the fraction of quenched galaxies more massive than $10^{10.5} M_{\odot}$ predicted by Magneticum and EAGLE (where passive galaxies are selected via the specific star formation rate cut) at $z \sim 1$ and 2, falls below the observed ones (Muzzin et al., 2013, and others) by $\sim 10 - 15$ per cent. Similarly, the passive fractions predicted by L-Galaxies at $z = 0.86$ (where passive galaxies are identified with a $U - B$ versus M_{\star} cut) fall below constraints from COSMOS (compilation by Peng et al., 2010) by 15 – 20 per cent too, suggesting that quenching mechanisms are less effective in those models than what observations imply.

Broadly speaking these results hint at some crucial aspects of state-of-art cosmological simulations: many observables in the galaxy populations are not recovered in simulations or still need to be addressed. One of these is the passive and star forming fractions both in the field and in the group/cluster environment. Also, colours and magnitudes still remain a challenge not yet fully resolved by simulators. This highlights how using different sets of simulations to perform comparisons between predicted and observed data allows us to have information on the different implementations in simulations of the physical processes driving galaxy formation.

² Illustris: <https://www.illustris-project.org/>

³ IllustrisTNG: [tng-project.org](https://www.tng-project.org/)

⁴ Eagle: <https://icc.dur.ac.uk/Eagle/>

⁵ Magneticum: <http://www.magneticum.org/index.html>

⁶ L-Galaxies: <https://lgalaxiespublicrelease.github.io/>

Part II

METHODS

In this chapter, we introduce one of the two main tools on which this thesis work is based. We overview the methods for cosmological simulations. This section will include a treatment on both the gravity (Section 3.2) and hydrodynamics (Section 3.3) solvers, with a particular emphasis on the algorithms used in our standard code GADGET. In Section 3.4, we discuss more in detail GADGET-3, including a treatment of the subgrid models used to parametrise several astrophysical processes (i.e., radiative cooling, photoionisation heating, star formation, chemical enrichment and AGN feedback). Within the simulations, we identify the halos and substructures hierarchy through a combination of halo-finding techniques, presented in Section 3.5. In conclusion, Section 3.6 presents a map-making tool called Smac to create mock images from simulated data. Considering the sizable volume of the literature on the topic, we aim at providing only a general picture of how these techniques generally operate, while focusing our more detailed discussion on a few points key to our work. The interested reader can find a more in-depth treatment on the subject in Benz (1990), Dolag et al. (2005, 2008, 2009), Dehnen and Read (2011), Borgani and Kravtsov (2011) and Springel (2016).

3.1 COSMOLOGICAL SIMULATIONS

Modelling the evolution of structures in the Universe with a cosmological framework requires going beyond the perturbation growth with the linear theory and addressing the gravitational collapse of overdensities in the non-linear regime. Thanks to the revolutionary development of powerful calculators in the XX century, numerical simulations have made it possible to tackle the issue. This effort rewards cosmologists and astrophysicists with a realistic representation of the Universe at different scales used to make predictions on the underlying mechanisms. The gravitational force will be the prime responsible for matter accretion in structure formation, although several other physical processes need to be accounted for. In broad terms, the main processes involved can be split into three main categories: gravitational, gas-dynamical and galactic processes. In the following coloured box, a few key aspects for each category have been summarised.

GRAVITY SOLVERS: N-BODY CODES A complete description of the single collisionless component (Binney and Tremaine, 1987), such as DM particles and stars, can be given by the continuous distribution function $f(\vec{x}, \vec{v}, t)$, which describes the number density of particles in the phase-space $\vec{\omega} = (\vec{x}, \vec{v})$. This function satisfies the Boltzmann (or Vlasov) equation:

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla} f - \vec{\nabla} \phi \cdot \frac{\partial f}{\partial \vec{v}} = 0. \quad (10)$$

where ϕ relates via the Poisson equation $\vec{\nabla}^2 \phi(\vec{x}, t) = 4\pi G \rho(\vec{r}, t)$. This can be solved by discretising the initial distribution with particles and evolving the resulting N-body system with simple interparticle interactions, whose computational cost rapidly increases with the number N of particles $\mathcal{O}(N(N-1)/2)$. To reduce the number of operations, there are different approximation techniques, such as Particle Mesh codes or Tree-codes, that can attain computationally sustainable results (see Section 3.2).

HYDRODYNAMICS The collisional component in cosmological simulations (i.e., gas) is rather complex to treat. Through numerical methods, there can be two approaches:

Eulerian techniques We record the evolution of the flow properties in space (i.e., field description of the system). Thus, field properties (e.g., the density $\rho(\vec{x}, t)$ and the velocity $\vec{u}(\vec{x}, t)$) at a specified point depend on location and time.

Lagrangian techniques The fluid is discretised by a distribution of particles. Field variables are attached to particles and their evolution is followed along particles' trajectories. Lagrangian methods, thanks to their intrinsic resolution-adaptive nature, are well suited to follow the gravitational growth of structures and automatically increase the resolution in the central regions of galactic halos, which are the areas of primary interest in cosmology.

Both methods are widely used in the field and the choice of which one to use is more driven by the advantages of each numerical scheme towards solving that numerical setup with the highest possible accuracy (i.e., convergence). Seek Section 3.3 for more details on our implementation.

GALACTIC PROCESSES A complete and coherent description of cosmic structures requires the coupling of baryon hydrodynamical equations with the astrophysical processes shaping our observables. Nevertheless, modelling galaxy formation in cluster simulations is extremely challenging because it requires sampling scales of a few tens of comoving megaparsec, while star formation and relevant gas-dynamical processes in galaxies occur below the parsec scales. One way to approach the issue is to implement phenomenological "sub-resolution" models describing what occurs at the scales not resolved by simulations, but whose effects propagate to explicitly resolved scales (see Section 3.4).

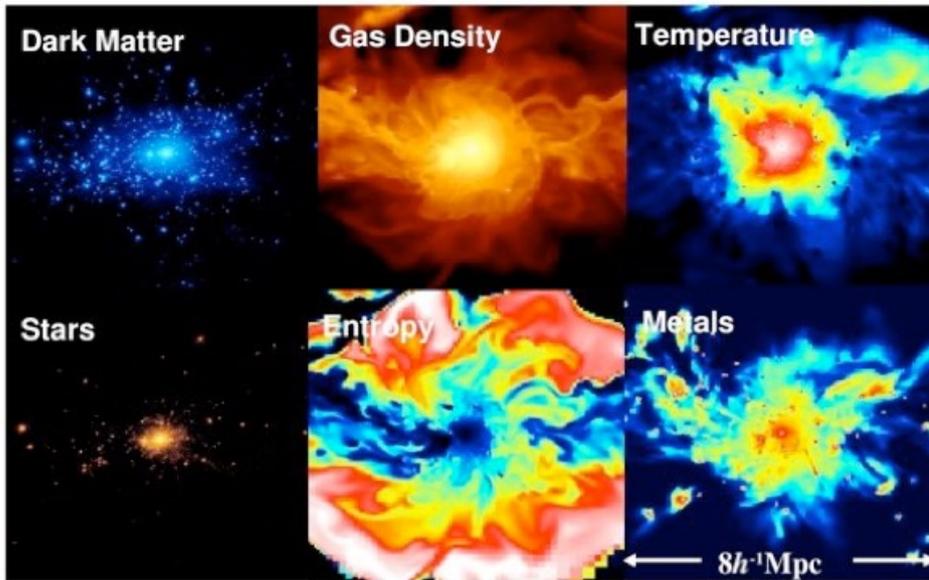


Figure 3.1: A snapshot of a simulated galaxy cluster showing why they are defined as multi-component systems. Following from left to right, and top to bottom, we see represented the DM distribution, the gas density, the gas temperature, the star distribution, the gas entropy and the gas metallicity. The zoomed region has size $8 h^{-1} \text{ Mpc}$. Credits: <http://www.astro.yale.edu/nagai/Research.html>.

Figure 3.1 shows an example of what a galaxy cluster might look like in a simulation when taking into account all its components. From left to right (going top to bottom) there are represented the DM distribution, the gas density, the gas temperature, the star distribution, the gas entropy and the gas metals. Generally, the gas distribution is smoother than the stellar and DM ones, a difference determined by the kind of interaction these components have (i.e., collisional and collisionless). Stars tend to be concentrated in small regions as a consequence of the radiative collapse from which they have formed which usually originates in restricted volumes. Such multi-components studies allow us to address a variety of physical effects altogether.

Having determined the framework necessary to resolve the given set of equations, in principle we are ready to carry out our simulations. Although for cosmological simulations we should switch to general relativistic equations of motion, we can take advantage of a very useful approximation. On large scales, the Universe is nearly isotropic and well described by a smooth Friedmann-Lemaître-Robertson-Walker (FLRW) metric. On the other hand, on small scales, the Universe must tend towards a locally inertial frame in which Newton's laws are valid. This can be used to prove that the cosmological expansion has essentially no effect on the local dynamics, even on galaxy cluster scales. Thus, we may think of the Universe as being filled with a self-gravitating fluid that locally obeys Newton's law of gravitation, while expanding as an FLRW metric on large

scales. The initial conditions set in the simulation will resemble those of the primordial Universe, determined by the precise CMB measurements of the cosmological parameters (for a recent review Collaboration et al., 2020), which also shed light on the nature of the cosmological perturbations that have formed the observed structures.

3.2 N-BODY SOLVERS

3.2.1 PM code

A large class of N-body codes is based on computing the large-scale gravitational field using Fourier techniques (Hockney and Eastwood, 1988). In this approach, one divides the space into a grid of density values and computes the force by solving on this grid the Poisson equation in comoving coordinates,

$$\vec{\nabla}^2 \phi(\vec{x}, t) = 4\pi G a^2(t) [\rho(\vec{x}, t) - \bar{\rho}(t)]. \quad (11)$$

The cubic box of the simulation (assumed to have volume L^3) is divided up into a grid of M^3 meshes with constant length $l = L/M$. Each one of them is identified by a mesh point (with coordinates $\vec{X}_q = l\vec{q}$) and a value of the mass m according to the particle distribution within the grid. For example, if at some time t the particle positions are \vec{r}_i , ($i=1\dots N$), we can select a window function W_q , assumed to be normalised such as $\int W(\vec{r}) d\vec{r} = 1$, so that the density field at the position \vec{q} of a grid point can be computed by interpolation according to:

$$\rho(\vec{q}) = \frac{m}{L^3} \sum_{i=1}^N W(\vec{r}_i - \vec{X}_q). \quad (12)$$

Rephrasing the problem in terms of Fourier space makes it a little easier to implement. The transformed Poisson equation $-k^2 \phi_{\vec{k}} = 4\pi G \rho_{\vec{k}}$ shows a simple relation between the gravitational potential and the distribution of the particles. If the computational box has a size much larger than the scales of interest, the mass distribution can be assumed to be periodic on the opposite surfaces of the box, and the Fourier transform of $\rho(\vec{q})$ can be obtained speedily using the fast Fourier transform (FFT) (e.g. Press et al. (1992)). The gravitational force at each grid point can then be obtained by $\vec{F}_{\vec{k}} = -i\phi_{\vec{k}}\vec{k}$.

It is clear that in the PM algorithm the total computational cost for force evaluation scales linearly with the number of meshes M ; a large number of meshes must be used to achieve high accuracy. The PM algorithm's strength relies on the saving of computational costs, although its resolution is sensitively limited by the grid spacing. The P^3M algorithm attempts to remedy this by calculating the potential through a direct sum for particles that are found at scales below a few mesh sizes and through the PM method for the others. Another technique to increase resolution in grid-based methods is based on adaptive mesh refinement, which builds in the

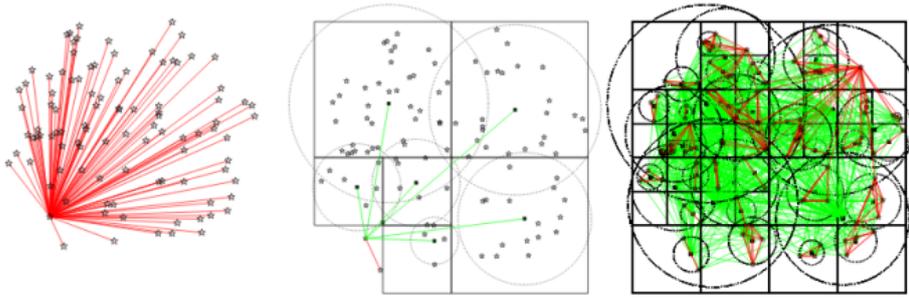


Figure 3.2: Left: computation of the forces to a single target particle with direct summation over 100 particles included in the simulation: every line corresponds to a single computation. Middle: computation of the forces of the same particles utilising the Tree-code algorithm. Cells are shown as black squares with their centres indicated by solid squares; green lines denote the cell-particle interaction. Right: calculation of all the forces among the particles in the simulation, requiring 902 cell-particle and 306 particle-particle interactions, instead of 4950. Credits: Dehnen and Read (2011)

high-density region a finer grid to capture more details while a coarser grid is used in lower-density regions.

3.2.2 The Tree algorithm

The Tree-code (Barnes and Hut, 1986) is constructed to maintain a coarse-grained representation of the system at large scales while reaching higher resolution on small scales with respect to the PM scheme. This method makes use of a hierarchical spatial tree to define localised groups of particles by grouping them systematically according to their distances to the particle for which the force is being calculated. Instead of requiring $N - 1$ partial force evaluations per particle, as needed in a direct summation approach, the gravitational force on a single particle can be computed with substantially fewer operations, because distant groups are treated as “macro” particles in the sum. Thus, the force from each group is replaced by its multipole expansion and the total sum usually reduces to $N \log(N)$ for homogeneous particle distributions, whereas the scaling for strongly inhomogeneous distributions, as present in evolved cosmological structures, can be less efficient (Dolag et al., 2008).

Following the prescription given by Barnes and Hut, 1986, the algorithm is divided into two steps: (i) creating the tree and (ii) walking the tree to calculate inter-particle distances. In the first step, the space is partitioned into small cubic cells which are recursively divided into smaller regions (in three dimensions, the initial cube is divided into eight sub-cubes) until either one or zero particles are in a cell. In both cases, and at each stage of the recursive process, the centre of mass of the particle distribution in

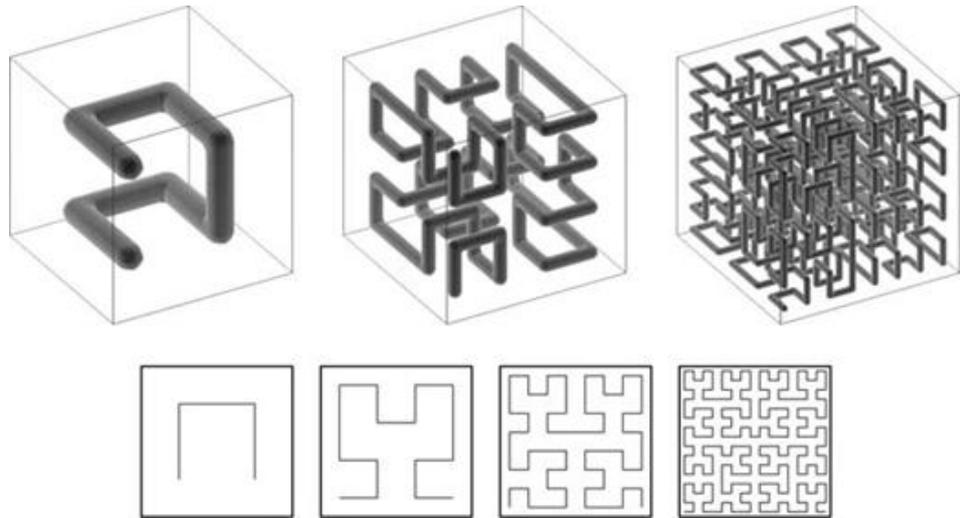


Figure 3.3: Space-filling curve Peano-Hilbert in two (lower) and three (top) dimensions. Credits: Springel (2005)

each cell is calculated. As this recursive process of subdividing the space continues, the results of every stage can be stored in a tree structure: each node will contain the centre of mass coordinates and total mass associated with the array of particles in the cell.

Once the tree is constructed, we must “walk the tree” to carry out the force calculations. Keeping in mind that the primary objective is to maximise the efficiency in calculating inter-particle distances, we note that if an array of particles is “far enough” away from an individual particle, the array can be treated as a single particle with a composite mass, located at the centre of mass of the array. The numerical condition to satisfy is that a node is viewed under an angle smaller than the critical opening angle θ_c , i.e., $\theta \leq \theta_c$. Otherwise, the tree is walked through until the condition is met. In other words, starting at the root node, a decision is made as to whether or not the multipole expansion of the node provides an accurate force. If the answer is “yes”, the multipole force is used and the walk along this branch of the tree can be terminated; if the answer is negative, the node is opened, thus its daughter nodes are considered in turn. The procedure continues until all groups are broken down as far as they need to be. The strength of this technique lies in the use of low-order multipole expansion to compute the gravitational potential of distant groups of particles, reducing the time consumption since one doesn’t necessarily descend through the tree.

Figure 3.2 illustrates the Tree-code work and compares it graphically to the direct summation approach, by applying both to the task of computing the force for one of 100 particles (left and middle panels). The Tree-code requires much fewer calculations — it is straightforward to show that the number of force computations per particle scales like the depth of the

tree, i.e. $\log N$, such that the cost of computing all N forces is $\mathcal{O}(N \log N)$ (Barnes and Hut, 1986; Dubinski, 1996).

Much of the memory overhead does not only come from building the tree, but also from the communication among CPUs (or tasks, if the code is parallelised). A popular choice to solve the issue is to order the particles on a so-called Peano-Hilbert curve which is the geometric analogue to a Tree-code realisation. The idea is hereby to order particles on the Peano-Hilbert structure while building the tree and send the particles to the various CPUs following the constructed Peano-Hilbert curve. This ensures that particles close in space are likely to be on the same CPU which reduces the communication cost between processors in a very efficient manner. We show an idealised structure of the Peano-Hilbert curve in [Figure 3.3](#).

3.2.3 Tree-PM code

Improved implementation of the above techniques is the Tree-PM first developed by Hernquist and Katz (1989) and Xu (1995). The algorithm combines the PM method on large scales with a Tree-code to handle particle-particle interactions at smaller separations. PM methods offer the fastest schemes for computing the gravitational force, but for scales below one or two mesh cells, the force is heavily suppressed. As a result, this technique seems not to be well suited for simulations requiring high resolution. After the global PM forces are calculated in the box, spatially distinct regions above a given density contrast are located; the Tree-code calculates the gravitational interactions inside these denser regions at higher spatial and temporal resolution. It is worth noticing that for each tree this portion of the code is self-contained; that is, once the particle data and potential mesh for a given tree have been collected together no further information is required to evolve that tree forward. This makes Tree-PM well suited for parallel processing on distributed memory systems.

In Springel (2005), this is performed by splitting the potential in Fourier space into a long-range and a short-range component according to:

$$\phi_{\vec{k}} = \phi_{\vec{k}}^{\text{long}} + \phi_{\vec{k}}^{\text{short}}.$$

In [Figure 3.4](#) we illustrate the partition between the two forces. Contributions from the long-range are

$$\phi_{\vec{k}}^{\text{long}} = \phi_{\vec{k}} \exp(-k^2 r_s^2),$$

where r_s is a spatial scale of the force split. Analogously, short-range can be solved in real space by noting that for $r_s \ll L$ the short-range part of the real space solution is given by

$$\phi_{\vec{k}}^{\text{short}}(\vec{x}) = -G \sum_i \frac{m_i}{r_i} \operatorname{erfc} \left(\frac{r_i}{2r_s} \right),$$

where $r_i = |\vec{x} - \vec{r}_i|$ is the distance of any of the particle i to the point \vec{x} . If periodic boundary conditions are enabled, we take the minimum of the

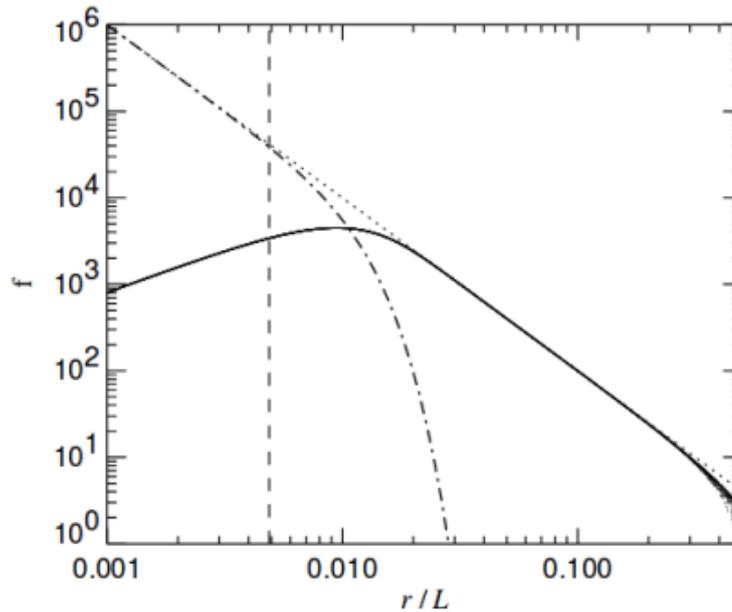


Figure 3.4: Spatial decomposition of the force in a Tree-PM code. This has been computed by randomly placing a particle of unit mass in a periodic box, and measuring the forces obtained by the simulation code for a set of randomly placed test particles. Credits: Springel (2005)

difference in r_i . The error function establishes a natural cut-off for $r > r_s$. The Tree-PM ensures an enhancement in performance and accuracy. We point out that such a scheme is implemented in GADGET-3, which is the code run for the simulations employed for this thesis.

3.3 HYDRODYNAMICAL SIMULATIONS

Over the past two decades, two dominant approaches have been used to model the collisional component in numerical simulations: the SAM, where baryonic physics is modelled at the scale of an entire galaxy and applied in post-processing on top of DM simulations, and hydrodynamical simulations, where the evolution of the gaseous component is treated resolving the fluid dynamics equations. The latter approach enables the complex interaction of the different baryonic components (gas, stars, and BHs) to be treated at a much smaller scale, ideally yielding self-consistent and powerfully predictive calculations. In the GADGET-3 code (see Springel 2005 for a description of a previous version of this code) used for this analysis, a SPH method is applied, so we will briefly illustrate this hydrodynamical solver in the next section. We also mention mesh codes, as a valid alternative to solve hydrodynamic schemes in cosmological simulations.

3.3.1 Smoothed Particle Hydrodynamics

Smoothed-Particle Hydrodynamics (SPH), which was independently developed by Lucy (1977) and Gingold and Monaghan (1977), is a Lagrangian (particle-based) technique that follows the motion of individual elements in a fluid (see Springel 2010, 2016 for a review). The SPH is particularly well suited to be used in association with a Tree/N-body code since their underlying principles are similar. Both techniques are fully Lagrangian and as such their resolution is intrinsically adaptive. The resulting code naturally allows following the evolution of a large number of particles within a Lagrangian framework (Hernquist and Katz, 1989).

Let us consider a non-relativistic ideal fluid with density ρ , moving at a velocity \vec{v} under the influence of a gravitational field with potential ϕ . Its time evolution can be described by the *hydrodynamical equations* in the Lagrangian formulation as in the following:

$$\frac{d\rho}{dt} + \rho \vec{\nabla} \cdot \vec{v} = 0 \quad (13)$$

$$\frac{d\vec{v}}{dt} + \frac{\vec{\nabla} P}{\rho} + \vec{\nabla} \phi = 0 \quad (14)$$

$$\frac{du}{dt} + \frac{P}{\rho} \vec{\nabla} \cdot \vec{v} = 0 \quad (15)$$

which represent the continuity equation, the conservation of motion and energy respectively. Notice that, as it is, the system is not closed and cannot be solved: we have five unknowns ($\rho, \vec{v}, P, \phi, u$) and only three equations. We include the Poisson equation, connecting the density ρ with the gravitational potential ϕ via $\nabla^2 \phi = 4\pi G \rho$, and the equation of state $P = P(\rho)$ generally assumed to be a polytropic $P = A\rho^\Gamma$ with $\Gamma = 5/3$ to resolve it. Once the set of equations is determined, we can solve it numerically following the SPH approach.

In the SPH, to avoid statistical fluctuations due to particle-discretisation of a fluid, the distribution is smoothed through the convolution of a smoothing function W (also called smoothing kernel, smoothing kernel function or simply kernel in the dedicated literature). A generic function $f(\vec{x})$ over a volume Ω can be written in terms of the smoothing kernel $W(\vec{x} - \vec{x}', h)$:

$$\overline{f(\vec{x})} = \int_{\Omega} W(\vec{x} - \vec{x}', h) f(\vec{x}') d\Omega_{\vec{x}'} \quad (16)$$

where h defines the spatial scale of the influencing area of a particle. The volume integral of the kernel W is normalised to 1 and it goes to the Dirac δ function for $h \rightarrow 0$. The kernel is generally assumed to be only a function of the modulus of the separation vector of $|\vec{x} - \vec{x}'|$, twice differentiable, non-negative, and monotonically decreasing. These conditions ensure that the error relative to the smoothed field is of the order of $\sim h^2$ (e.g., Price, 2012).

This prescription can be used to sample any field A of the fluid, so that the field value and its spatial derivative at an arbitrary position \vec{x} can be approximated respectively by:

$$A(\vec{x}) = \sum_j W(\vec{x} - \vec{x}_j, h) \frac{A_j}{\rho_j} m_j; \quad (17a)$$

$$\vec{\nabla} A(\vec{x}) = \sum_j \vec{\nabla} W(\vec{x} - \vec{x}_j, h) \frac{A_j}{\rho_j} m_j. \quad (17b)$$

The equation provides us with an efficient way to numerically solve the evolution of fluids which all just depend on the first spatial derivatives of different fields (velocity, density and internal energy of SPH particles) that can be integrated forward in time as in the N-body simulations.

It is worth noticing that the choice of the functional form of the kernel has a significant impact on the final result. A Gaussian kernel will take into account the contributions of all particles, weighting them on their distance to the position \vec{x} , but it largely increases the computational costs since neighbours out to arbitrarily large distances need to be included. It is preferable to choose a Kernel with compact support so that only neighbours out to a finite distance need to be included in the computation of the smoothed quantities.

3.3.2 Mesh codes

Aside from the Lagrangian methods, mesh codes deserve a mention in the description of the cosmological simulation methods. Mesh based finite volume codes follow the prescriptions for the Eulerian scheme whereby the computational domain is divided on a fixed grid with thermodynamic properties given at each cell. The evolution is given by the computation of the fluxes between neighbouring cells, known as the Riemann problem which consists of the resolution of a conservation law with initial conditions given by two piece-wise constant states that meet at a plane (Springel, 2016). The first (and most important) implication of solving a Riemann problem is that these schemes automatically include the dissipative terms describing discontinuities. However, since the domain is discretised in cells, discontinuities are also introduced where the exact flux is continuous, resulting in spurious dissipation. This spurious effect can be reduced using different reconstruction methods for the states at the interfaces, instead of a constant approximation. The second implication is that energy and angular momentum are not conserved with this scheme. Thirdly, mesh codes also lack natural adaptivity in highly dense environments, much like those simulated in cosmological boxes, unless specific adaptive grids are introduced.

Further advancements in the field have introduced a mixture between the two methods (i.e., a Lagrangian meshless scheme) which guarantees the conservation properties of SPH while allowing the rise of the artificial dissipation, like in grid methods. (Hopkins, 2015) introduced a method

based on kernel discretisation of the volume to a higher-order gradient estimator and a Riemann solver between the particles inside that volume, the Meshless Finite Mass (MFM) method. This allows the particles to move with the flow; energy, mass and momentum are conserved and there is no need for artificial diffusion terms.

3.4 THE GADGET-3 CODE

GADGET-3 (GALaxies with Dark matter and Gas intERacT) code is a massively parallel Tree-PM/SPH code, capable of following a collisionless fluid with N-body methods, and an ideal gas employing SPH (for more details, we refer to the review by Springel 2005). It was used as the founding architecture of the first cosmological simulation for 10^{10} DM-only particles, in addition to very large hydrodynamic simulations (reaching 250 million particles, Springel 2005).

The code is built so that the collisionless dynamics can be followed with either a Tree code or with a hybrid Tree-PM algorithm. The formulation of the SPH manifestly conserves energy and entropy. It also includes a prescription for artificial viscosity that can capture shocks, while decaying away from such regions. Adaptive time-steps guarantee improvements in computational costs since larger temporal intervals will be used in low-density regions with almost no effects on the accuracy. In addition, GADGET-3 allows the user to choose over a large set of parameters including cosmological parameters, subgrid models and particle masses. We report a summary of the simulation parameters in Table 3.1. Additional physics described by GADGET-3 includes radiative cooling of the gas, heating by an ultraviolet background, and treatment of star formation and feedback processes associated with SNe explosions and AGN. To present a lookout for the features included in the code, the next sections will shortly describe such methods.

3.4.1 Radiative cooling and photoionisation heating

The primary cooling processes relevant for structure formation are two-body radiative processes, in which gas loses energy through radiation as a result of two-body interactions. For our purposes, processes involving three bodies can be ignored, since in cosmological simulations the involved gas densities are too low. According to the temperature range, different processes are dominant.

For temperatures above 10^6 K primordial gas (composed of hydrogen and helium only) is almost entirely ionised, and above a few 10^7 K the enriched gas (heavier elements are present) is completely ionised. The only significant radiative cooling at these high temperatures is bremsstrahlung due to the acceleration of electrons as they encounter atomic nuclei.

At lower temperatures, several other processes contribute. The phenomenon is caused by atomic processes such as recombination, collisional ionisation and excitation. Each interaction removes from the gas a certain

Table 3.1: A summary of the main characteristics of the analysed simulations. For each set of simulations, we report: DM and gas particles mass; DM, gas, star particles and black hole (Plummer-equivalent) softening lengths at redshift $z = 0$; and the total number of clusters analysed. The simulations are named "Hydro" or "DM" whether they are hydrodynamical or DM-only respectively.

	Hydro-1x	Hydro-10x	Hydro-25x	DM-10x	DM-100x
$M_{\text{DM}} [10^8 h^{-1} M_{\odot}]$	8.3	0.83	0.34	1.00	0.10
$M_{\text{gas}} [10^8 h^{-1} M_{\odot}]$	1.56	0.16	0.06	-	-
$\epsilon_{\text{DM}} [\text{kpc h}^{-1}]$	3.75	1.4	1.0	1.4	0.6
$\epsilon_{\text{gas}} [\text{kpc h}^{-1}]$	3.75	1.4	1.0	-	-
$\epsilon_{\text{star}} [\text{kpc h}^{-1}]$	1.0	0.35	0.25	-	-
$\epsilon_{\text{BH}} [\text{kpc h}^{-1}]$	1.0	0.35	0.25	-	-
N Lagrangian Regions	29	11	10	29	12

amount of internal energy equal to the element ionisation or excitation potential, thus causing the cooling of the gas. In the code, the cooling process is modelled following Katz et al. (1996): radiative cooling is implemented as a process which acts on an optically thin gas in ionisation equilibrium with a specified time-dependent background of ultraviolet radiation. On the other hand, the background radiation is expected from an early population of quasars (Madau et al., 1999) which reionised the Universe at $z \simeq 6$. To compute the radiative cooling rates, one needs to know the abundances of the different ionic species of the gas, since in the code it is a metallicity-dependent process. For reasons of computational costs, these rates are often implemented via a lookup table which is updated every time the ionising background changes.

3.4.2 Star formation

Radiative cooling of the gas may determine star formation in the right environmental and physical conditions. The star formation model implemented in GADGET-3 follows the prescription described in Yepes et al. (1997) and Springel (2005). The authors argued the instrumental advantage of using a multiphase component to describe the different states of the gas in GADGET-3. In this model, each SPH particle is described by a cold and a hot phase with densities ρ_c and ρ_h , in pressure equilibrium.

Star formation occurs when cold clouds are converted into stars after a characteristic time scale t_* at a rate expressed by Schmidt’s law (Schmidt, 1959). A small fraction of cooling gas that turns into stars β_{SNe} returns sufficiently massive stars ($M > 10^8 M_\odot$) to evolve as SNe, thus the change of star density ρ_* in time will be regulated by:

$$\frac{d\rho_*}{dt} = (1 - \beta_{\text{SNe}}) \frac{d\rho_c}{dt}. \quad (18)$$

Different Initial Mass Function (IMF) will change the expected value for β_{SNe} . In our simulations, we always use a Chabrier IMF (Chabrier, 2003).

SNe contribute with a certain amount of hot gas enriched with metals – see Section 3.4.3 – and the injection of energy in the system. Assuming that cold clouds dissipate by thermal conduction inside the hot bubbles hosting the SNe, we have cloud evaporation whose efficiency depends on the density of the gas. In this scheme, a phenomenological model of the galactic winds is introduced. This feedback process partly solves the overcooling problem by expelling gas from the centres of low-mass halos (Nath and Trentham, 1997; Aguirre et al., 2001; Madau et al., 2001). Since winds can reheat and transport collapsed material from the centre of a galaxy back to its extended DM halo and even beyond, they can help reduce the overall cosmic star formation rate to a level consistent with observational constraints. In simulations, winds are associated with gas particles stochastically selected among the multiphase particles with a probability proportional to the local star formation rate.

3.4.3 Chemical enrichment

In GADGET-3, the star formation prescription is coupled to a description of chemical evolution. For a detailed description, we leave for the interested reader a few references (Tornatore et al., 2007a; Matteucci, 2003; Borgani et al., 2008).

In short, each star particle produces chemical elements (in our simulations, He, C, Ca, O, N, Ne, Mg, S, Si, Fe, Na, Al, and Ar in addition to keeping track of H) via three different channels: SNe Ia, II, and AGB stars. The chemical stellar evolution is determined by the lifetime function, the IMF, and the yields for SNe, intermediate, and low mass stars. Binary systems – which explodes as SNe Ia – in the single-degenerate scenario Greggio and Renzini (1983) and Matteucci and Greggio (1986) are also accounted for. This represents an important source of energy feedback that is caused by the explosion of white dwarfs in the process of accretion from the companion in binary stellar systems. Furthermore, it is generally agreed that stars having masses above $8M_\odot$ at the end of their lifetime undergo an electron capture process, leading to a core-collapse. A third channel to eject heavy elements in the interstellar medium is the mass loss of low and intermediate stars by stellar winds.

To increase the precision of a metal enrichment model that allows for separate treatments of type I and II SNe, it is necessary to take into account

the lifetimes of different stellar populations, as it has been independently implemented by Tornatore et al. (2007b) and Scannapieco et al. (2005). The adopted lifetime function $\tau(m)$ aims to distinguish the different energy supplies from stars of different masses. Following in our simulations the prescription given by Padovani and Matteucci (1993), we assumed the following lifetime function:

$$\tau(M) = \begin{cases} 10^{[(1.34 - \sqrt{1.79 - 0.22(7.76 - \log(M))})/0.11] - 9}, & \text{if } M \leq 6.6M_{\odot} \\ 1.2M^{-1.85} + 0.003, & \text{otherwise} \end{cases} \quad (19)$$

which is correct for intermediate stellar masses. However, for very high ($> 60M_{\odot}$) or low ($< 1.3M_{\odot}$) masses, it is advisable to use different lifetime functions, as has been pointed out by the same authors. Different lifetimes will produce different evolutions of both absolute and relative abundances (e.g. Talbot, Arnett et al., 1973; Güsten and Mezger, 1982).

The stellar yields, which specify the number of metal species that are released during the evolution of a single stellar population, are taken from metallicity-dependent stellar yields. We take Thielemann et al. (2003) for SNe Ia, yields from Woosley and Weaver (1995) combined with Romano et al. (2010) for SNe II, and Karakas (2010) for AGB stars.

3.4.4 AGN feedback

Energy feedback released by gas accretion onto a SMBH is performed by following the scheme originally introduced by Springel (2005), (see also Di Matteo et al. 2005, 2008).

Primordial BHs ($M_{\text{BH}} = 10^5 h^{-1} M_{\odot}$) are seeded in DM halos whose masses are larger than a threshold mass value depending on the resolution of the DM particles (e.g., $M_{\text{th}} = 5 \times 10^{10} h^{-1} M_{\odot}$ at our base resolution). They are represented by collisionless sink particles of initially small mass, which can evolve during the simulation through two channels: via merging with other BHs during a close encounter or via gas accretion. Bondi (1952), Bondi and Hoyle (1944) and Hoyle and Lyttleton (1941) describe gas accretion rate for a BH of mass M_{BH} and velocity v_{BH} embedded in an environment of density ρ as,

$$M_{\text{B},\alpha} = \alpha \frac{4\pi\rho G^2 M_{\text{BH}}^2}{(c_s^2 + v_{\text{BH}}^2)^{2/3}}. \quad (20)$$

Here, α is a proportionality constant that is not determined by the steady-state equations, though it is constrained within certain limits. The boosting factor α is needed as the Bondi formula has been derived under quite simplistic assumptions, such as spherical symmetry and unperturbed and steady initial conditions. In a more realistic scenario, the accretion happens more chaotically, with the overall effect of enhancing the accretion by a factor of 10 – 100.

However, for any compact object, there is a physical limitation to the BH accretion rate due to the Eddington limit. This arises because both

the attractive gravitational force acting on an electron-ion pair and the repulsive force due to radiation pressure decreases inversely with the square of the distance. When the luminosity exceeds the Eddington limit, which is given by

$$L_{\text{Edd}} = \epsilon_{\text{r}} c^2 \dot{M}_{\text{Edd}}$$

the gas will be blown away by the radiation. Thus, we find the Eddington accretion limit to be:

$$\dot{M}_{\text{Edd}} = L_{\text{Edd}} / (\epsilon_{\text{r}} c^2). \quad (21)$$

Therefore, the effective mass accretion rate is described by:

$$\dot{M}_{\text{BH}} = \min(\dot{M}_{\text{B}}, \dot{M}_{\text{Edd}}). \quad (22)$$

The parameter ϵ_{r} is the radiative efficiency and indicates the radiated energy in units of the rest-mass energy of the accreted gas. Typical values for a radiately efficient accretion onto a Schwarzschild BH are $\epsilon_{\text{r}} \simeq 0.1$ (Shakura and Sunyaev, 1973). A small fraction ϵ_{f} of the released energy thermally couples with the surrounding gas, in such a way, that we can determine the feedback energy rate provided, which reads,

$$\dot{E}_{\text{feed}} = \epsilon_{\text{r}} \epsilon_{\text{f}} c^2 \dot{M}_{\text{BH}}. \quad (23)$$

Di Matteo et al. (2005) has suggested $\epsilon_{\text{f}} \simeq 0.05$ because the authors were able to reproduce the observed $M_{\text{BH}} - \sigma$ bulge velocity dispersion versus BH mass relation (Magorrian et al., 1998). In simulations, gas swallowed by the BH is implemented in a stochastic way, by assigning to each neighbour gas particle a probability of contributing to the accretion proportional to the SPH kernel weight computed at the particle position. Assuming that each SPH particle contributes to the accretion with 1/3 of its mass, instead of being completely swallowed, a larger number of particles contributes to the accretion, allowing us to follow it continuously.

Besides this standard implementation of BHs in cosmological simulations, a few more prescriptions are added to improve the modelling. By following the procedure in Shakura and Sunyaev (1973), BHs can be classified in a "quasar mode" or a "radio mode" according to the intensity of the accretion rate. High accretion values lead to very luminous gas, thus small ϵ_{f} in Equation 22, which is what characterises quasars at high redshifts. Instead, when $\dot{M}_{\text{BH}} / \dot{M}_{\text{Edd}} < 10^{-2}$, the BHs are assumed to switch to "radio mode" in which most of the energy is released as kinetic and thermalises with the gas, leaving a low accretion rate. This is true for BHs that are at lower redshifts.

Although Omma et al. (2004) and Brighenti and Mathews (2006) have pointed out that gas circulation generated by jets provides an important contribution to the stabilisation of cooling flow, the model of BH feedback implemented in our simulations provides energy only in thermal form to the gas particles surrounding the supermassive BHs. The scales resolved in our simulations are not in the regime where kinetic feedback can be explicitly resolved, thus justifying the adoption of purely thermal feedback.

3.5 STRUCTURES IDENTIFICATION

Once a cosmological simulation is completed, we still need to identify the gravitationally bound structures that have formed. As we have seen in the previous sections, cosmic growth is sensitive to several parameters that define the background expansion, as well as the nature and composition of DM and dark energy, or the presence of neutrinos (Lesgourgues and Pastor, 2006). Thus, any analysis that we want to perform requires an accurate and efficient way to estimate the distribution of gravitationally bound and virialised structures, to be associated with observable structures such as galaxies and clusters of galaxies. For our needs, we will restrict to the description of two techniques, the same used in our simulations, namely Friends of Friends and Subfind (Springel, 2005), leaving the more interested reader to other reviews, e.g., Knebe et al. (2013).

3.5.1 *The Friends of Friends group finder*

Friends of Friends (FoF) is a technique that identifies bound groups in the simulated box according to a simple geometrical method. The prescription is very simple as the algorithm selects particles separated by a distance that is less than a threshold length (the only free parameter of the algorithm), which is called linking length b : e.g. values of order $b \sim 0.2$ (in units of the mean inter-particle separation) are used to identify virialised structures. In Figure 3.5, we report an example of two FoF groups (coloured areas) formed by associating different particles with a given linking length. Notice that choosing different linking lengths may turn into different group classifications.

The first application of this algorithm to identify groups of galaxies in redshift surveys was made by Geller and Huchra (1983). Widely used is also the method described in Springel (2005), the same scheme employed in this work, which creates FoF groups by hierarchical merging, and that was specifically developed for the analysis of large cosmological simulations, like those analysed in this thesis. This was originally implemented for parallelisation on large distributed computer architectures, as it allows a very large concurrency with a simple decomposition of the problem onto spatially disjoint domains. In each iteration, the code performs a neighbour query on a selected set of points and merges the proto-clusters hosting these particles by updating the labels of all constituent points of these two. Iterations are repeated until no additional merging is possible. There are few intrinsic limitations of the FoF algorithm, as it tends to link independent structures across feeble particle bridges occasionally, or it is not capable of detecting substructures inside larger virialised objects (Springel et al., 2001).

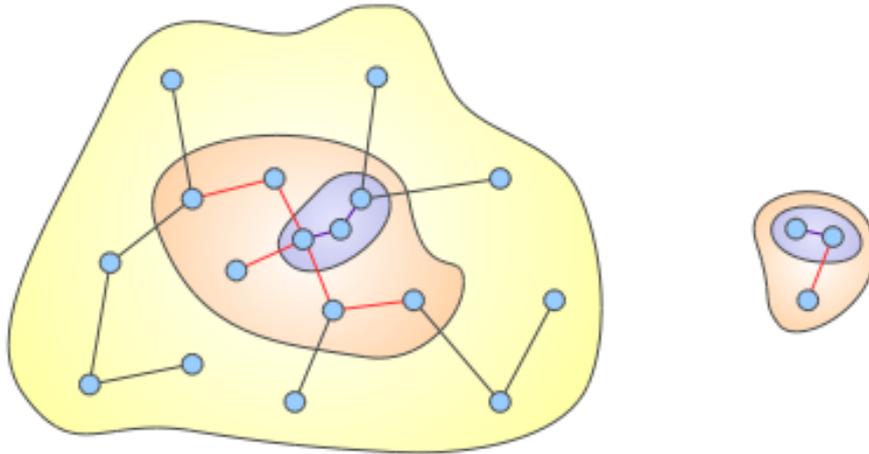


Figure 3.5: Graphical representation of two FoF groups, according to different linking lengths. The shorter the linking length, the smaller the resulting group. Credit: [cosmosim](#).

3.5.2 The Subfind algorithm

Originally implemented for DM-only simulations, (Springel et al., 2001), and then extended to hydrodynamical simulations (Dolag et al., 2009), Subfind is a refined step in the structure identification procedure. The parent group (i.e., cluster) is initially selected from the FoF output run on the DM particles only. Each gas and star particle is then associated with its nearest DM particle, i.e. if this DM particle belongs to a FoF group, then the corresponding baryonic particle is also associated with that group. From there, Subfind searches for substructure candidates as those regions enclosed by isodensity contours that traverse a saddle point. In other words, Subfind descends along the density gradient given by the particles to find the local maxima and minima of the potential. Each three-dimensional area that is encompassed by local minima (saddle points) is a potential candidate. This procedure creates a list of subhalo candidates that outlines the first determination of the substructures, as seen in Figure 3.6. In addition, Subfind implements an unbinding procedure, to include particle members of a substructure only those that are gravitationally bound to that substructure. This amounts to eliminating those particles whose internal energy is positive (unbounded particles): if more than a certain minimum number of particles (50) survives the unbinding, the substructure is identified as a genuine subhalo. The centre of each subhalo is identified with the minimum gravitational potential occurring among the member particles.

Similarly, for the cluster as a whole, we determine a suitable centre as the position of the particle that has the minimum gravitational potential. Around this point, we calculate the virial radius and mass with the spherical-overdensity approach, using the overdensity predicted by the generalised spherical top-hat collapse model (e.g. Eke et al., 1996). We

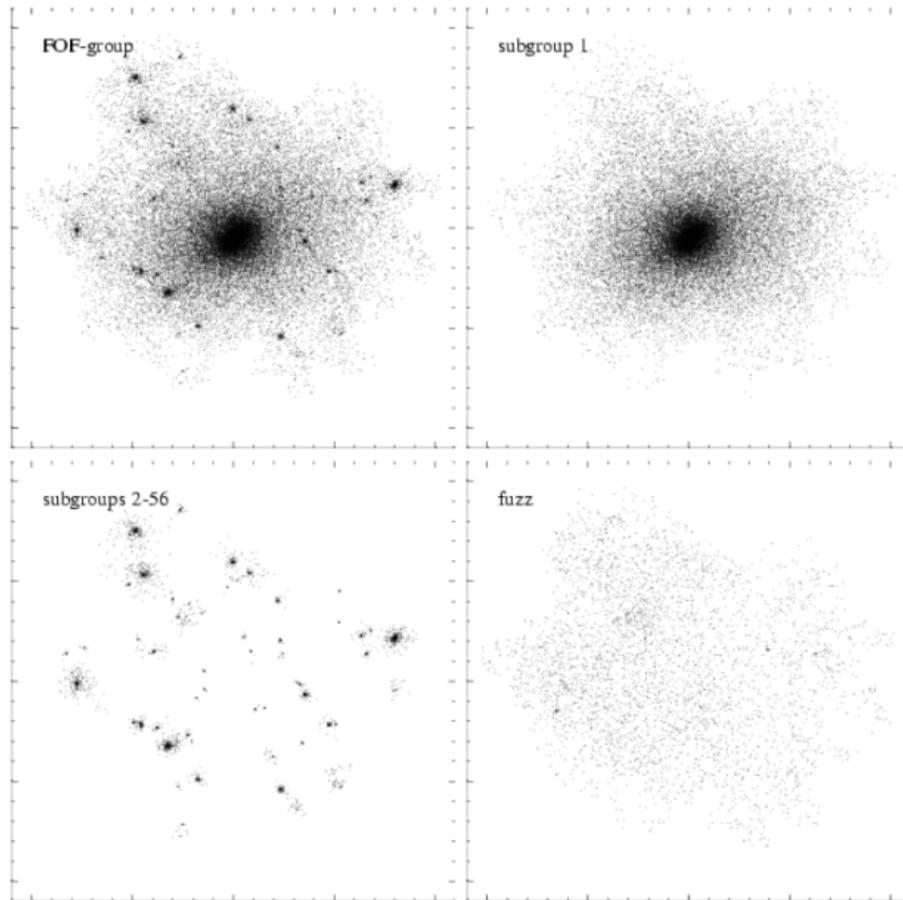


Figure 3.6: Projected two-dimensional distribution of a simulated halo in which Subfind identifies different subhalos. The FoF halo is in the top-left panel. The main halo is on the top right; the substructures are on the bottom left. The bottom-right hosts the "fuzz", which is not gravitationally bound to any structure but resides in the same spatial location as the group. Credits: Springel et al. (2001).

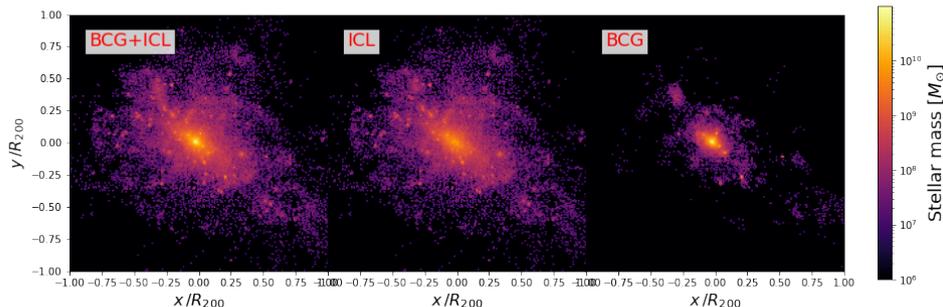


Figure 3.7: Mass-weighted maps of the stellar components as identified by the ICL-Subfind for one of the clusters in our simulations. Left panel: BCG and ICL hosted in the main halo of the cluster. Central panel: stars associated with the ICL component. Right panel: stars bound to the BCG. The stellar component is dominated by the ICL diffuse component.

define as subhalos of a cluster all substructures identified within its virial radius. Note that this can sometimes include subhalos that belong to a FoF group different from that of the main halo. Likewise, not all of the subhalos in the cluster’s FoF group are necessarily inside its virial radius.

3.5.3 ICL-Subfind

Besides the identification of bound structures, a Subfind also includes a feature which separates stars bound to the BCG from the ICL. This division is performed once all star particles of the central group are isolated, by firstly subtracting the stars associated with the member galaxies of the group with the standard version of Subfind (presented in the previous section). The details of this technique are presented in Dolag et al. (2010), here we will only provide a brief description.

The algorithm identifies the single star particles in the main halo of clusters as either bound to the BCG or ICL, solely applying a dynamical criterion. The underlying assumption is that the two velocity distributions of stars belonging to the ICL and the BCG can be each fitted by a Maxwellian shape so that the overall velocity distribution of stars not assigned to substructures is described by a double Maxwellian of the following form:

$$N(v) = k_1 v^2 \exp\left(-\frac{v^2}{2\sigma_1^2}\right) + k_2 v^2 \exp\left(-\frac{v^2}{2\sigma_2^2}\right). \quad (24)$$

The diffuse ICL is associated with the Maxwellian yielding the largest velocity dispersion, in contrast, the BCG, having colder dynamics, populates the distribution at lower dispersion.

To assign each star particle to either one of the two dynamical components, the algorithm follows an unbinding procedure, by iteratively computing the gravitational potential given by all particles within a sphere

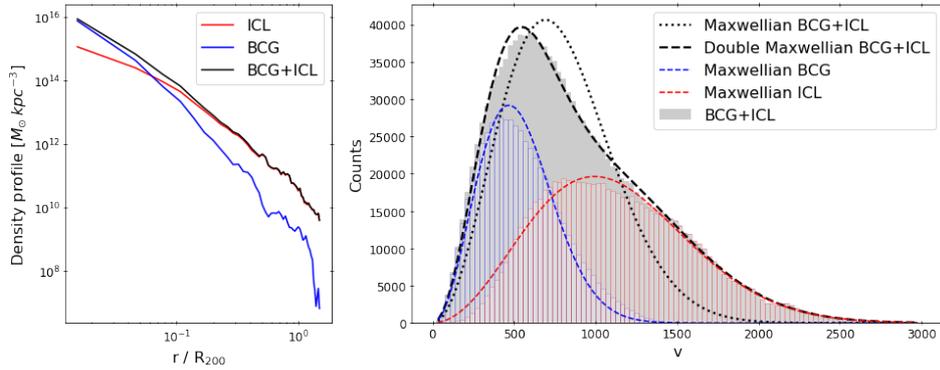


Figure 3.8: Observed properties in the stellar populations identified by the ICL-Subfind for the same cluster in Figure 3.7. Left panel: main halo including BCG+ICL (black), ICL (red), and BCG (blue) stellar density profiles. Right panel: velocity histograms of the main halo (black), ICL (red), and BCG (blue). We also report the single (dashed black) and double Maxwellian (dotted black) best-fit curves for the entire data set. Red and blue dashed lines are showing the individual Maxwellian distributions which are associated with the unbound ICL ($\sigma_{\text{ICL}}^* = 1002 \text{ km s}^{-1}$) and bound BCG ($\sigma_{\text{BCG}}^* = 484 \text{ km s}^{-1}$).

whose radius is initially equal to a fraction of the virial radius. In this framework, we compare each particle's kinetic energy with the potential energy (at the particle position) given by this spherical mass distribution. If the particle's kinetic energy is higher, then the particle is defined as "unbound" (and "bound" otherwise). Performing this operation on all star particles identifies two stellar populations which are then separately fitted with a single Maxwellian. If the best-fit parameters of the double Maxwellian match those obtained from the single initial Maxwellians, then the procedure is completed, otherwise, the radius of the sphere is changed and the computation is repeated. Note that the radius is usually adjusted to match the value of the BCG components, but provided that the algorithm does not converge, then a second attempt is made with the ICL component. The radius is varied to decrease (increase) the spherical mass distribution according to whether the BCG velocity dispersion is too high (low) compared to the expected result from the initial double Maxwellian fit. The iterative procedure stops when the ratio of the expected velocity dispersion over the fitted one differs by less than a given value and thus, one obtains the label "ICL" or "BCG" for each star particle in the main halo. Nevertheless, the algorithm may not necessarily converge if the number of iterations exceeds a limiting value provided by the user. We will see that this is the case in some cases.

As a proof of concept, Figure 3.7 illustrates the outcome of the labelling in the total stellar population of the main halo (left panel) from the stellar map weighted on the particle masses of one of the clusters in our simulation. By applying the ICL-Subfind algorithm, we can separately study the characteristics of the ICL (central panel) and BCG population (right panel). Similarly, Figure 3.8 shows the resulting density profiles (left panel) and

velocity histograms (right panel) of the same cluster. Radius and velocity are normalised to R_{200} and V_{200} ¹. The colour-coded legend is common to both panels: stars in the main halo in black, ICL in red, and BCG in blue. The density profiles show that the BCG mostly resides in the central regions, while the ICL extends to larger distances, in fact dominating the stellar component in the outskirts. In the right panel, we plot the histograms of the main halo star particles' velocity distribution (in black) and the single BCG (blue) and ICL (red). We observe that a single Maxwellian (best-fit $\sigma^* = 867 \text{ km s}^{-1}$), represented by the dotted black line, does not provide a good fit to the particle distribution. On the other hand, a double Maxwellian fitting (black dashed line) provides a much better agreement. The single Maxwellian associated with both the BCG and ICL is also reported with the dashed lines. The diffuse ICL is associated with the Maxwellian with the larger velocity dispersion ($\sigma_{\text{ICL}}^* = 1002 \text{ km s}^{-1}$), in contrast, the BCG, having colder dynamics, populates the distribution at lower dispersion ($\sigma_{\text{BCG}}^* = 484 \text{ km s}^{-1}$). Furthermore, in Dolag et al. (2010) it was tested that a triple Maxwellian does not improve the results in most cases.

3.6 THE SMAC VISUALISATION TOOL

Realistic graphical representations at the particle level of astrophysical properties extracted from an SPH cosmological simulation require special care. The use of a smoothing kernel in this prescription, not only defines the volume in which a particle has influence, but also translates the distribution of the particles to a continuous field in space. This allows us to construct smoothed images of simulated astrophysical processes (such as gas shocks, stellar luminosities, and densities) from our simulated universe. We use Smac which is a map-making utility for idealised observations. Originally, it was designed to extract full-sky maps from constrained numerical simulations to infer the statistical properties of the thermal and kinetic Sunyaev-Zel'dovich effect produced by the local supercluster (Dolag et al., 2005). Over the years, additional features were added (e.g., density maps, pressure maps, temperature maps and X-ray surface brightness maps) to overcome the continuous need for accurate visual representations of baryonic physics included in the simulations. One of its main qualities is its large versatility which allows the user to independently choose among a wide set of parameters to create the preferred image. These include, but are not limited to, the size of the image, pixels and the radius along the line of sight, the observed property, the format of the image and the centring.

Regardless of the observable to plot, its main rendering algorithm is based on standard SPH kernel projection techniques whose fundamental building unit is the smoothing kernel function. According to the so-called

¹ We can define R_Δ as the radius encompassing a mean overdensity equal to Δ times the critical density of the universe $\rho_c(z) = 3H^2(z)/8\pi G$. Equivalently, we define the circular velocity within $V_{200} = \sqrt{GM_{200}/R_{200}}$

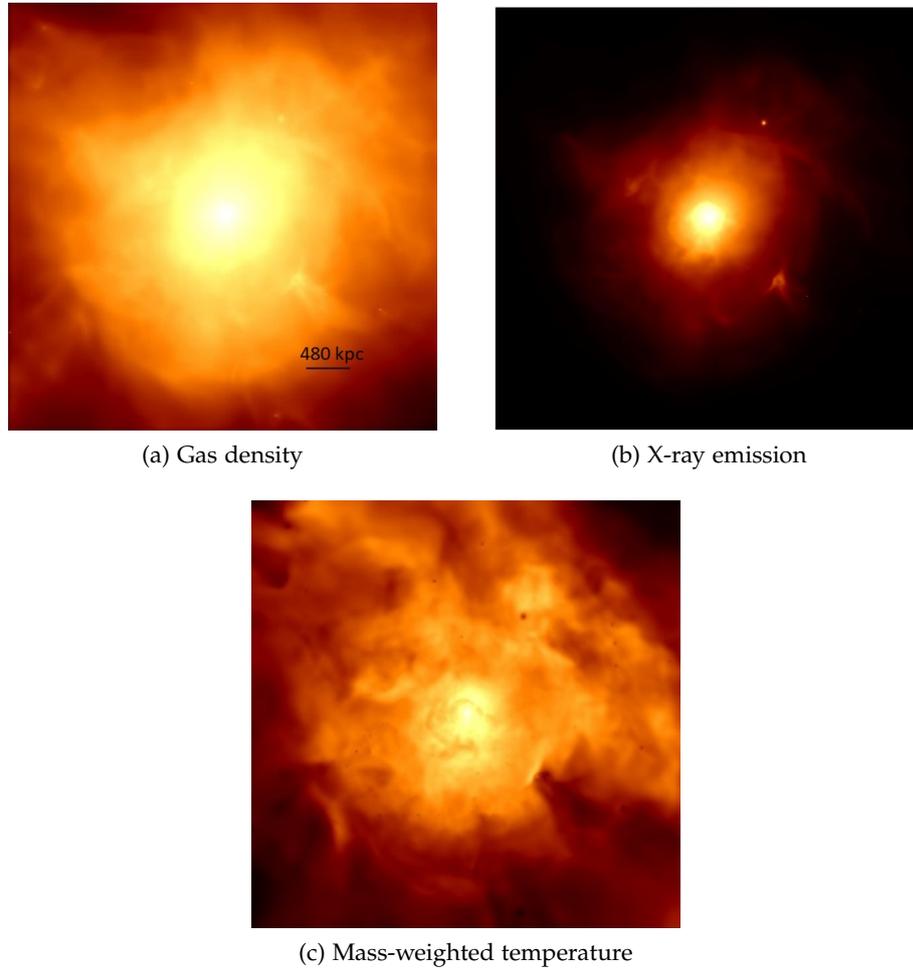


Figure 3.9: Different mock images (produced with Smac) of a galaxy cluster. We present the gas density (top-left), the X-ray emission (top-right) and the mass-weighted temperature of the gas (bottom). The cluster is selected at $z = 0$ from the DIANOGA set (25x resolution). The image is centred on the cluster's centre and it has size R_{200} of the cluster.

gather approximation, all particles whose projection either overlaps or completely falls in a given pixel are taken into account. This choice largely facilitates the overall computation. In fact, in general, provided that we aim to render a generic observable X contributed by the j^{th} SPH particle, we would define the line of sight integral as:

$$X_{\text{int}} = \int \sum_j \frac{m_j}{\rho_j} X_j W[d_j(r)/h_j] dr \quad (25)$$

where m_j and ρ_j are the mass and density of the particle, and $d_j(r)$ is the projected distance with respect to the position r along the line of sight. In principle, the sum has to be carried out over all particles, but in the case of a compact kernel where $W(r/h)$ becomes zero at large r/h (such as the one in Monaghan and Lattanzio, 1985), it can be restricted to those particles having a distance from the line of sight smaller than twice their smoothing length. With the gather approximation, the last equation becomes

$$X_{\text{pix}} = \sum_j \frac{m_j}{\rho_j} X_j N_j \frac{A_{\text{pix}}}{A_j} W_{\text{int}}\left(\frac{\tilde{d}_j}{h_j}\right), \quad (26)$$

where W_{int} is the integrated kernel, that depends on the projected distance \tilde{d}_j . The fraction A_{pix}/A_j represents the effective area of influence of the particle j^{th} in the pixel. In other words, the geometry of the configuration between the pixel and the particles determines the contribution of the particle to the given pixel. To first approximation, the area A_j associated with each SPH particle corresponds to the 2/3-rd power of the corresponding volume, i.e. $A_j = (m_j/\rho_j)$. The term N_j is a normalisation defined such that

$$\sum_{\text{pix}} N_j A_{\text{pix}} W_{\text{int}}\left(\frac{\tilde{d}_j}{h_j}\right) = A_j \quad (27)$$

to conserve the quantity X when distributed over more than one pixel.

We point out that in cases when the particle's influence sphere is contained in only one pixel, W_{int} is fixed to unity. On the other hand, if the particle only partially overlaps with the pixel, the associated area is approximated by:

$$\hat{A}_{\text{pix}} = \sqrt{A_{\text{pix}}} \left(\tilde{d}_j + 0.5\sqrt{A_{\text{pix}}} + h_j \right). \quad (28)$$

To compute the contribution to such a pixel, the weight of the integrated kernel W_{int} is then corrected by a factor $\hat{A}_{\text{pix}}/A_{\text{pix}}$. Figure 3.9 shows the output of three different maps computed with Smac on the same cluster from the DIANOGA set at $z = 0$. The images are centred on the cluster centre and they all have physical size $2 R_{200} 512 \times 512$ pixels). The maps report the gas density (top-left), the X-ray emission (top-right) and the mass-weighted temperature (bottom). Brighter colours indicate higher values of the observable. Notice that gas density, X-ray emission and temperature are observables connected to the physical state of the intracluster medium.

3.6.1 Assigning optical/IR luminosities to star particles

State-of-art cosmological simulations – including a treatment of star formation, chemical evolution, and stellar and AGN feedback schemes, such as our simulation set – can be used to predict the spectro-photometric properties of star particles and simulated galaxies. The basic ingredient is the stellar evolution model which can be employed to predict the evolutionary track of stars in the theoretical Hertzsprung-Russell diagram. The evolution of a star is almost completely determined by its initial mass and chemical composition. Thus, for a given initial mass and metallicity, a stellar evolution model should yield all the properties of a star particle (i.e., metallicity, mass, age) at any time after its birth. To compute these quantities, we use GALAXEV (Bruzual and Charlot, 2003). GALAXEV is a library of stellar spectral evolution that spans a substantial range of ages and metallicities in stars. Provided that each star particle in a simulation is representative of a SSPs, with its mass, metallicity and redshift of formation, we can assign a luminosity in given bands, for the used initial mass function in the corresponding simulation. For this purpose, we pre-compiled an interpolation grid of luminosity values from GALAXEV by assuming a Chabrier IMF. Consistent with the stellar evolution model implemented in the simulation code, GALAXEV keeps track of stellar mass loss, thus the interpolated value needs to be multiplied by the initial mass of the star particle. In other words, the luminosity $L_{*,v}$ of a star, treated as an SSP of mass M_* at the time t and in a given photometric band v is

$$L_{*,v}(t) = \frac{M_*(t)}{1 M_\odot} L_v(1 M_\odot). \quad (29)$$

It is worth noticing that there is not a unique definition of luminosity in the literature, for this reason, we included different ones used in photometry. Rest-frame and observer-frame luminosity differ by integrating the filter response in different spectral intervals due to the Universe expansion ($\propto \lambda(1+z)$). This requires the use of multiple grids to interpolate (thus, not just for stellar age and metallicity) as a function of redshift, so we compute the grid for a log-spaced distribution of redshifts z in the interval $[0, 25]$ and interpolate between the two files closest in time. Standard values of metallicities in the interpolation are in the range $0.005 - 2.6 Z_\odot^2$, therefore we decided to set to zero as a default answer in cases when the metallicity was above or below this interval. The final set of filters is composed of 97 filters which include the Large Synoptic Survey Telescope, James Webb Space Telescope, Hubble Space Telescope, Euclid, Herschel, and Spitzer (Rodrigo et al., 2012; Rodrigo and Solano, 2020).

The code that computes luminosities is implemented with an OpenMP scheme, allowing one to compute in parallel luminosity maps for a given number of filters. This ensures a quick way of computing several maps at once for the same object, which is a useful trick to mimic photometric maps of a telescope scanning different bands at the same time.

² The solar metallicity is set to $Z_\odot = 0.02$.

3.6.2 *Photometric maps with Smac*

The scheme to assign luminosities outlined in the previous section can be coupled to a map-making tool such as Smac to produce synthetic images of the photometric properties of simulated galaxies. We provide an example of a photometric map in [Figure 3.10](#). The images report the stellar metallicity (top-left), the stellar age (top-right) and the stellar luminosity in the Buser's V band (bottom) of the same cluster depicted in [Figure 3.9](#). These quantities have been chosen here for their connection in calculating the luminosity. The brightness marks the higher metallicity, the younger population and the higher luminosity respectively. Notice that we can recognise a dim stellar halo enveloping the central galaxy, and smaller substructures all around. In [Chapter 8](#), we will mention the use of this tool to create synthetic maps of the BCG+ICL and ICL-only distributions. Such maps can be used to train a Deep Learning (DL) model to recognise the contribution from ICL in observational images of the central regions of galaxy clusters.

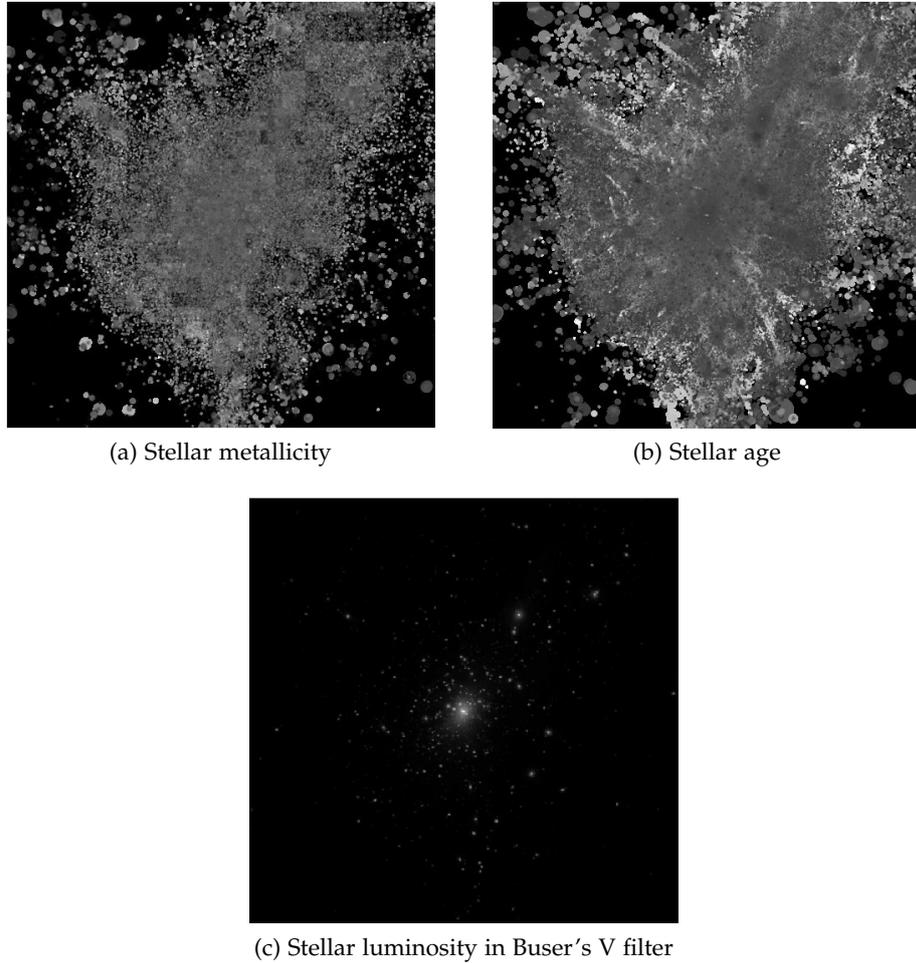


Figure 3.10: These maps (produced with Smac) are centred on the same cluster as in the previous image. We report the stellar metallicity distribution (left-top), the stellar age (right-top) and stellar luminosity (bottom). The brightness denotes the higher metallicity, younger stellar population and more luminous respectively. We can recognise structures such as massive and small galaxies within the virial radius we used to frame the image.

Since its first appearance, Machine Learning (ML) has proved its power in many scientific applications (medicine, biology, applied science, and economics, just to name a few). Rather than hard-coding knowledge in artificial intelligence, these types of algorithms can acquire knowledge by extracting patterns from raw data. Each piece of information included in the representation of the system is known as a feature. From a given set of features, the algorithm is trained to discover a hidden mapping in the data from representation to output. Therefore, the effectiveness of a ML algorithm heavily relies on the integrity of the input-data representation: poor representations will not provide the network with a suitable training set, thus causing bad performance.

The scope of this chapter is to introduce a few key concepts from the vast field of ML to provide the reader with the theoretical basics to understand the research work in [Chapter 7](#) and [Chapter 8](#). In [Section 4.1](#), we introduce some basic concepts and terminology used and we explain the main differences between ML and DL models. In [Section 4.2](#), we describe the Random Forest, a simple but effective ML architecture to resolve classification problems which is employed in [Chapter 7](#). [Section 4.3](#) presents the DL architectures employed in [Chapter 8](#) to solve an image recognition problem. We warn the reader that we do not aim to provide a complete review of the topic, but rather mention those aspects interesting for our research. We suggest to approach Bishop and Nasrabadi (2006), Goodfellow et al. (2016), Stevens et al. (2020) and Alzubaidi et al. (2021) for more complete descriptions.

4.1 BASIC CONCEPTS

4.1.1 *What is a machine learning model?*

In broad terms, ML is a discipline which attains to resolve pattern recognition problems in structured data. It is easy to imagine how such skill can be of major impact in science as hidden relations between data are intertwined with the underlying physical processes. Specifically, in ML the parameters of an adaptive model are tuned to perform this task during a phase called *learning process*. This phase can be thought of as the search for a mapping between a given input dataset (i.e., *features* \vec{x}) and an output function (i.e., *prediction* \vec{y}), where a model is said to be "trained" to recognise a mapping.

The learning process ML models undergo can be broadly categorised into two major categories: unsupervised and supervised. Supervised learning deals with labelled data. The algorithm is trained to recognise a specific

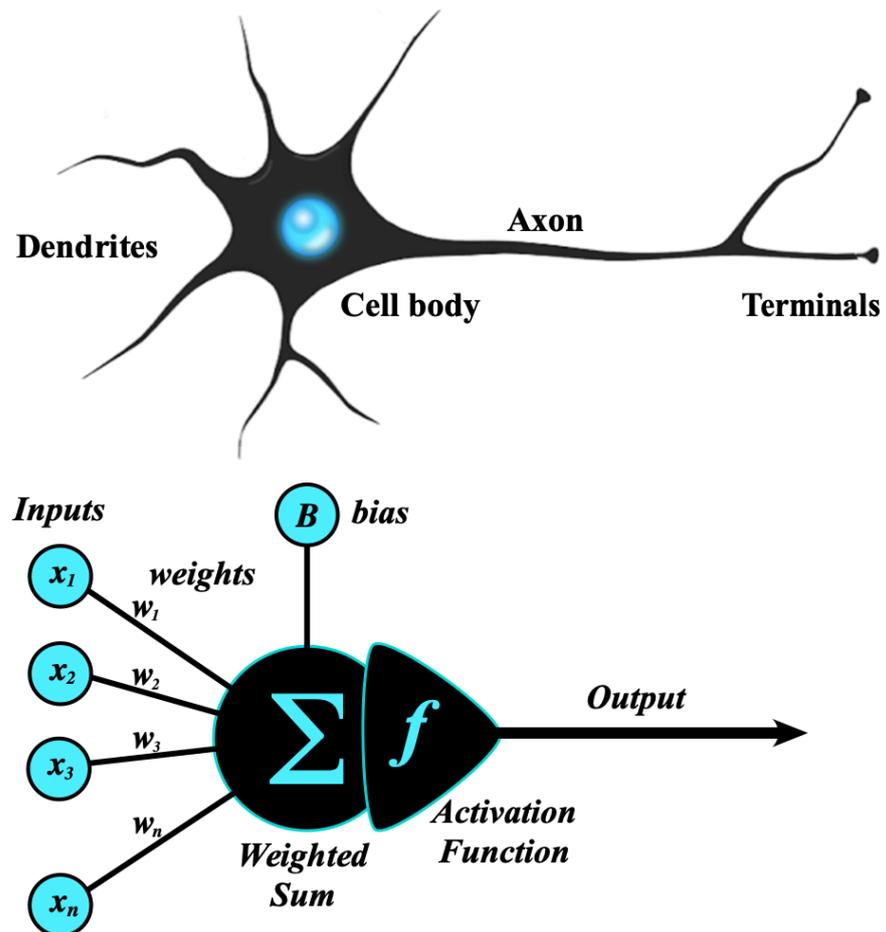


Figure 4.1: Biological neural cell (top) and artificial neuron (bottom) schemes. The analogy between the two is not random, as the operating principle of the artificial neuron is modelled after the biological one. Both neurons receive some information to process (dendrites/inputs) in order to provide an answer (terminal/output). Credits: <https://www.mriquestions.com/deep-learning-dl.html>

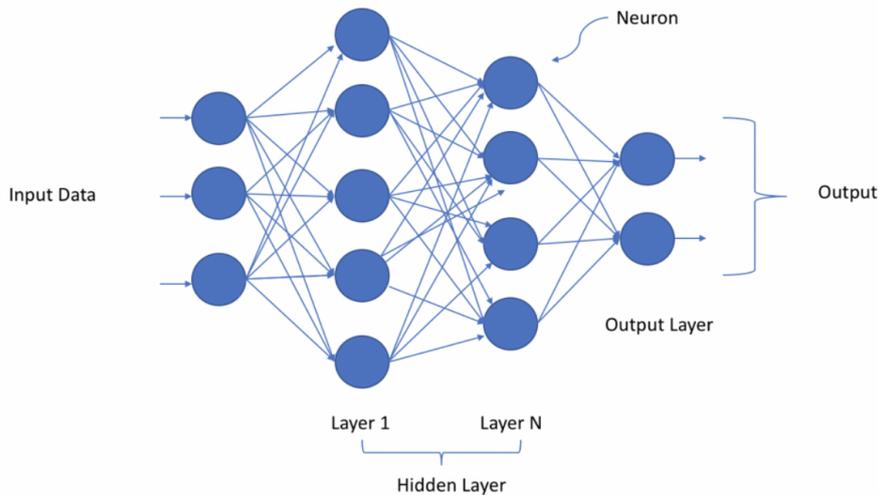


Figure 4.2: A deep NN composed of N hidden layers. Neurons are the fundamental unit in the hidden linear combination (layer) present in a NN architecture. Credits: <https://www.mriquestions.com/deep-learning-dl.html>

label (i.e., target or class) initially provided by the user. A simple example to help visualise is to imagine an algorithm trained to recognise cats among images of other animals. The model is given the information cat/no-cat for each image, and thus it learns to detect features that can be associated with a cat. On the other hand, in unsupervised learning, the network experiences an unlabelled dataset whereby it learns useful properties of the structure of this dataset, by determining the underlying probability distribution or some interesting property of the dataset itself. Taking a similar example to the previous one, we can imagine performing the same task without labelling the images. After having processed a certain number of images, we could expect a well trained model to be able to distinguish images with cats from other animals just by detecting differences in their morphology, colour, size etc.

In this work, we focus only on supervised learning as a viable way to efficiently and quickly classify large volumes of data coming from our simulations. More specifically, in [Chapter 7](#), we will use a ML model to classify stars in the main halo of simulated galaxy clusters between the two dynamical classes introduced in [Section 2.2](#): BCG and ICL. To achieve this, we train a Random Forest whose architecture is described in [Section 4.2](#). In [Chapter 8](#), we will present a more complex DL model trained to recognise the ICL contribution in mock images of galaxy clusters. These predictions can be made by a complex Neural Network (NN) such as a CNN, described in [Section 4.3](#).

Before we continue, let us briefly discuss what we generically mean by "training a ML model" and what this implies. This point will also allow us to introduce some of the terminology used in the rest of the chapter. In the ML context, training a model refers to determining a good set of

parameters $\vec{\theta}$ to solve a given task. These parameters are not to be confused with the *hyperparameters* of a model. The latter are those explicitly defined by the user to improve the learning of the model, and their values are set before starting the learning process of the model. For example, the number of input features and the depth of the architecture are hyperparameters. Notice that this initial choice will affect the performance of the model. Ideally, at the end of the training, models should be evaluated on data that were not used to build or fine-tune the model itself, so that they provide an unbiased sense of their efficiency and robustness. When dealing with a large amount of data, a set of samples can be set aside to evaluate the final model (Kuhn, Johnson et al., 2013). Usually, before the training phase, the entire dataset is split into three portions: *training, validation and test set*. The test set is the dataset which gives the final evaluation of the model performance: it benchmarks the metric scores of the model and therefore, it should not be "seen" by the network at any time during the training phase. In the training phase, the validation set supports the training set by providing a way of measuring the performance while tuning the parameters. The crucial point is that a test set, by the standard definition in the literature, provides an unbiased estimate of the generalisation error (assuming that the test set is representative of the population) and thus, it is a good measure to choose among two or more networks. The ability to categorise correctly new examples that differ from those used for the training is known as *generalisation*. In practical applications, variability in the data allows for only a tiny fraction of all possible input vectors to be "seen" by the model during the training phase, and so generalisation is a central goal in pattern recognition. In this regard, *overfitting* describes the condition whereby a model has been trained too well on the training data, and it is unable to generalise new data. In other words, the model is not able to make accurate predictions on unseen data. Notice that, a model is said to not have generalised also in the opposite case, i.e., *underfitting*. Underfitting happens when a model has not been trained enough on the data. This condition makes the model just as useless and it is not capable of making accurate predictions, even with the training data.

Most of the ML models require preliminary human intervention to preprocess the given input data, a step we need to follow to transform or encode data so that it may be easily parsed by the machine. Most of these interventions act to clean the data, include missing information and, data reduction (e.g., normalisation). Avoiding this step can also lead to a substantial increase in the learning process time.

Although we find a vast literature on the different models theorised and currently employed in ML, many of the key ideas that underpin them are common to all such problems and they have been summarised here above. As we shall see next, to use a specific architecture one needs to delve further into the subject to clarify a few aspects.

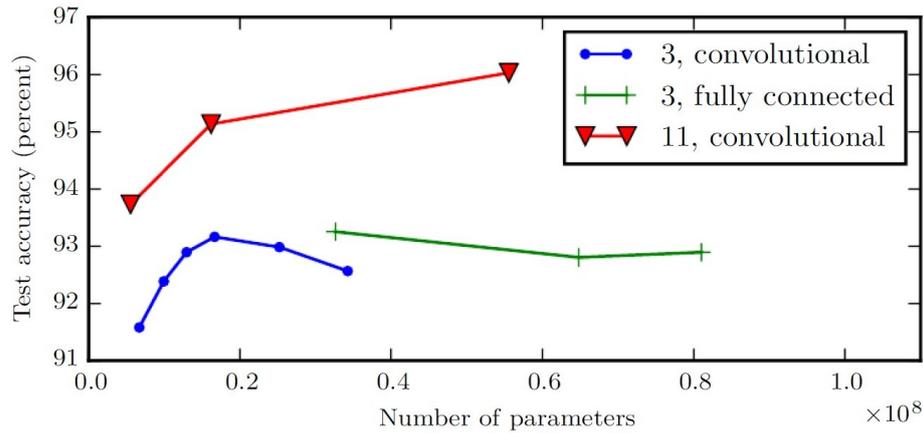


Figure 4.3: Comparison of the performances in Goodfellow et al. (2016). The plot shows the number of parameters set in the model versus the model accuracy when assessed with the test set. This is showing that increasing the number of parameters in layers without increasing their depth is not nearly as effective at increasing test set performance. The legend indicates the depth of the network used to make each curve (3 for blue and green, 11 for red) and whether the curve represents variation in the size of the convolutional or fully connected layers. Credits: Goodfellow et al. (2016).

4.1.2 Machine learning versus deep learning

Having described the general key points present in all ML models, we would like to draw attention to the differences between ML and DL in processing high-level abstract features from raw data. Conceptually, DL techniques *are* ML algorithms, thus they might be better thought of as a special field within the larger ML topic. The basic building block of any DL model is the *artificial neuron*, which is loosely modelled after the biological neural cell (Figure 4.1). A biological nerve cell (top panel) receives input stimuli from neighbouring nerves through its dendrites. If the sum of these stimuli is sufficient to create membrane depolarisation in the neuron's cell body, an electrical output signal will be transmitted down the axon to its terminals. Analogously, the artificial neuron receives a set of weighted inputs ($w_1x_1 + w_2x_2 + \dots + w_nx_n$) plus a constant bias (B). This weighted sum is then fed into an activation function that produces an output for the node. Both types of neuron process the input information thanks to a prime learning phase which trains the neuron in recognising the input data. Notice that the combination of artificial neurons in different layers often amplifies the capacity of a model to perform a task (much like what happens with brain cells), for this reason in computer science we often speak of NN.

We can distinguish DL from other ML models as they are usually characterised by (i) complex NN architectures, (ii) a low need of human intervention and (iii) a large data requirement. First and foremost, DL applications implement a layered structure of functions. We can think of

each one as an application of a different mathematical function providing a new representation of the input, thus allowing the computer to build complex notions out of simpler ones. To simplify this concept, let us take a look at [Figure 4.2](#). A deep NN is composed of an initial input layer of neurons (the first one to receive the input data), a hidden layer of arbitrary depth N (i.e., effectively corresponding to N layers of neurons) and an output layer. Each neuron (or node) is connected with trainable parameters (i.e., weights in the simple neuron scheme from [Figure 4.1](#)) in order to have different units acting in parallel. Each node will process part of the information through a function and then it will pass the information to the next one, therefore each layer is trained to recognise a specific feature in the data and to pass this information to the others. Thus, the combination of multiple neurons, such as the one represented in the figure, can break down a complex set of intertwined data thanks to the operating connections among neurons. In [Figure 4.3](#), we show an example of performance score improvement in cases when one uses a deep model, rather than a shallow one with more parameters. Thus, deep models generally outperform shallow ones. Secondly, the human intervention in DL is significantly low compared to general ML models. Specifically, one usually has to manually choose features to feed the NN. A simple example will be provided by the model trained in [Chapter 7](#) which classifies stars according to their properties in the phase-space. We show that prior to the training phase, we perform a feature importance recognition search to assess which features help the most our model to classify correctly (more on the feature importance in [Section 4.2.1](#)). On the other hand, in DL this task is completely automatised by the model. Thirdly, DL requires a large dataset to properly function. This is mainly due to the larger complexity of the model which needs more information to eliminate potential fluctuations.

In the DL context, training a model refers to determining a good set of weights and biases (i.e., the parameters $\vec{\theta}$) to solve a given task. Effectively, these values cannot be calculated using analytical methods. Instead, they must be discovered via an empirical optimisation procedure whereby the DL model minimises a *loss function* $J(\vec{\theta})$ with respect to its predictions. Therefore, the goal of training is to find a set of parameters that have low loss values across all examples, by following the *gradient descent* of the loss function. The *stochastic gradient descent* is the most diffused technique to search these minima and it replaces the actual gradient of the loss (calculated from the entire data set) with an estimate thereof taken from random points in the data. Especially in high-dimensional optimisation problems, this reduces the very high computational burden, achieving faster iterations in trade for a lower convergence rate. It is common to divide the sample in smaller subsets (i.e., *mini-batches*) and update the model parameters for each mini-batch seen by the model. This technique seeks to find a balance between the robustness of stochastic gradient descent and the efficiency of batch gradient descent (where the data is fed in a single iteration). Therefore, (mini-batch) stochastic gradient descent

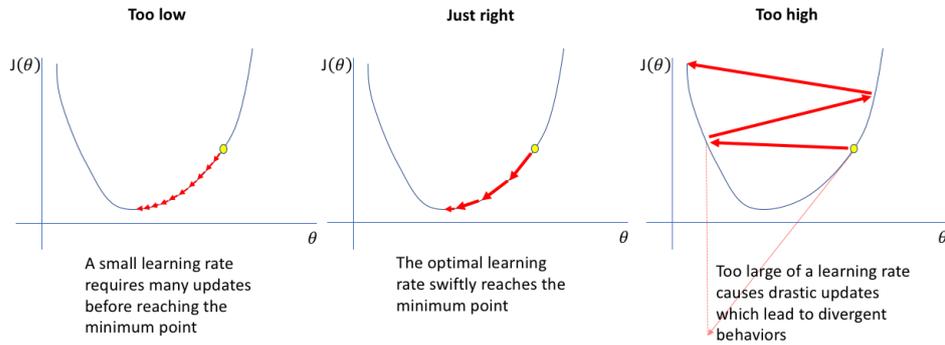


Figure 4.4: Examples of different selections of learning rates during the stochastic gradient descent of the loss function. The learning rate controls how much we are adjusting the weights of our network with respect to the loss gradient and how much we can earn at each timestep. Left panel: the learning rate is too small and thus the learning happens at a very small pace (care must be taken in this case as the procedure might get stuck in a local minimum). Central panel: correct learning rate. Right panel: the learning rate is too large and the algorithm is possibly not able to find the global minimum. Credits: [jeremyjordan.me/nn-learning-rate/](https://twitter.com/jeremyjordan)

requires the configuration of an additional hyperparameter to account for the size of the mini-batch during the learning phase. The amount of change to the model during each step of this search process, or the step size, is called the *learning rate* α and provides perhaps the most important hyperparameter to tune. The learning rate controls how much we are adjusting the weights of our network with respect to the loss gradient and how much we can earn at each timestep. We provide a few didactic examples in Figure 4.4: too small (left panel) or too large (right panel) learning rates can trouble the search for the global minimum in the loss function. We propose a technique to find the optimal one in Section 4.3.2. Finally, we define the *epoch* as one complete pass of the training set to the network during the learning process. The number of epochs determines how many times the model will see the same data for the parameter update: care must be taken when selecting the right number of epochs, as too large or too small values may cause overfitting or underfitting (see Section 4.3.1).

In conclusion, DL is a particular kind of ML that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones.

4.2 RANDOM FOREST

The Random Forest algorithm (Breiman, 2001) is a tree-based classification method that learns how to classify objects into different classes. Its

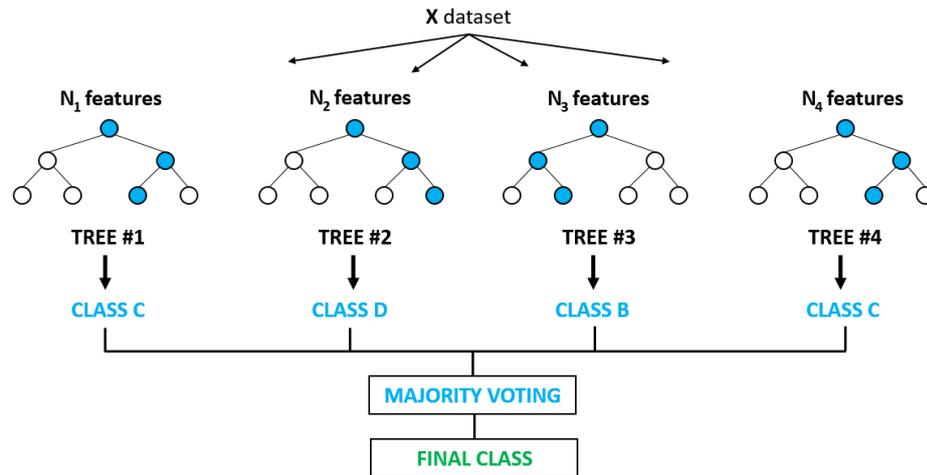


Figure 4.5: Graphical representation of a Random Forest architecture composed of 4 decision trees. Each tree has a hierarchical structure of leaves and nodes. Provided the dataset to each tree, they independently decide to which class the given example belongs. At the end of the computation, the final class is the average class coming from the ensemble. Credits: Habib et al. (2020).

founding components are the decision trees (Quinlan, 1993) that singly operate to make predictions on the single particle based on the associated input features. A Random Forest is thus a simple extension of the single operating decision tree, but it generally improves the performance of the classifier. Indeed, the advantage of constructing an ensemble of classifiers, where multiple trees fit random subsamples of the data, is that overfitting and instabilities in the data distributions may be mitigated by the averaged results of several trees. In Figure 4.5, we provide a graphical representation of the structure for a generic Random Forest. Lastly, one of the main advantages of using Random Forest is that it is a model easy to interpret.

4.2.1 Feature importance

We can interpret the output from a Random Forest in different ways. Feature importance is one of them, whereby the relevance of each feature in the classification process is investigated. Feature importances (Pedregosa et al., 2011) can be computed as the mean and standard deviation of the decrease in impurities summed for each tree in the forest. In the case of a classification problem, impurity is often defined as Gini impurity, which is a measurement of the likelihood of incorrect classification of a new instance if it were randomly classified according to the distribution of class labels in the data set. In other terms, this importance figure provides an understanding of how the ML classifier evaluates the input features relative to each other. For each feature, we collect the average decreases in impurity. The average over all trees in the forest is the measure of

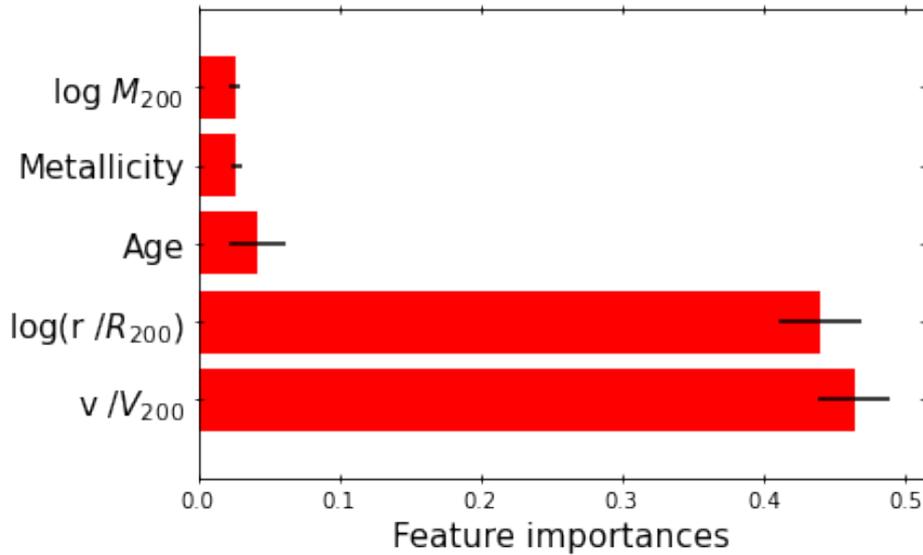


Figure 4.6: Graphical representation of the feature importance histograms to determine which input features are redundant in a model. This example is taken from the binary classification problem we outline in [Chapter 7](#).

the feature importance, the higher the importance the better. Most times these results are expressed in terms of histograms, thus allowing for an immediate visual interpretation of the most significant features during training. We anticipate one of the plots from [Chapter 7](#) to help the reader with a practical example. The binary classification problem we will face initially deals with five possible input features (i.e., $\log M_{200}$, metallicity, age, $\log(r/R_{200})$ and v/V_{200}) to train the model. Without going into the detail of what these features physically represent, we see in [Figure 4.6](#) that after training the Random Forest with this initial dataset, the features with the highest feature importance are $\log(r/R_{200})$ and v/V_{200} . In other words, the network mostly relies on these two features to discriminate one label with the other and make the final prediction. In fact, if we aimed at dealing with the simplest model to perform this classification problem, we would remove the other three input features decreasing the redundancy. On the other hand, there might be cases when one can consider retaining some of these, and we will see in [Chapter 7](#) why it will be the case in our problem.

4.2.2 Confusion matrix and performance scores

The confusion matrix is another useful tool to interpret results for general classification problems. A confusion matrix expresses how many of a classifier's predictions were correct compared to the total. In binary classification, a general confusion matrix has two rows and two columns (see [Figure 4.7](#)): columns represent the true labels and the rows represent predicted labels. Values on the diagonal represent the number (or per

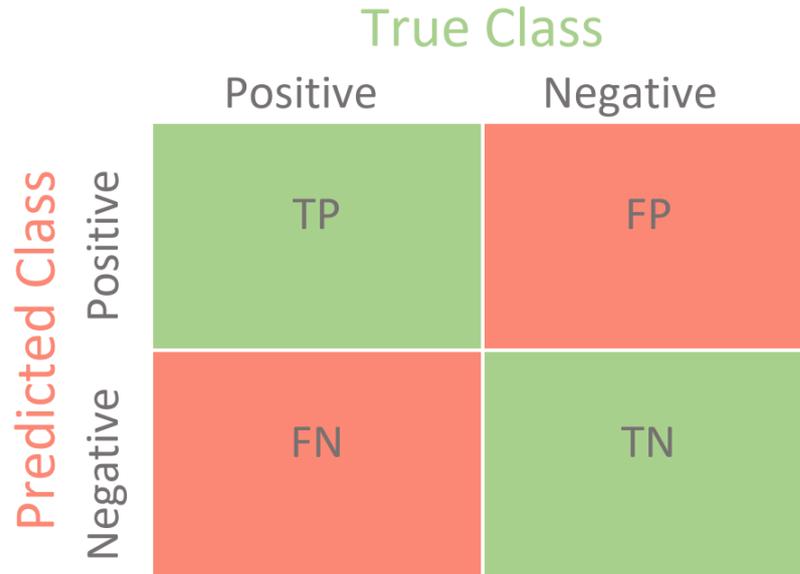


Figure 4.7: Graphical representation of a confusion matrix for a binary classification. The columns yields the true labels, while the rows have predicted labels. A normalised matrix will describe the percentages of each one of the four categories (after assembling true and predicted labels of the classes). Credits: <https://towardsdatascience.com/>

cent, in a normalised confusion matrix) of times when the predicted label matches the true label. Values in the other cells represent instances where the classifier mislabeled an observation. This is a convenient way to spot areas where the model may need a little extra training.

From a confusion matrix, one can easily obtain summary statistics to evaluate the model performance. In our case, we deal with imbalanced label problems, in other words, one class is dominant with respect to the other. Standard metrics tend to become unreliable and fail in such cases. For this reason, we use Recall (R), Precision (P), and F-Score (FS) to evaluate our model. These are defined as follows:

$$R = \frac{TP}{TP + FN}; \quad P = \frac{TP}{TP + FP}; \quad FS = 2 \frac{P \times R}{P + R}. \quad (30)$$

In the above expressions, TP, FP, and FN are the numbers drawn from the predicted labels of true positives, false positives, and false negatives, respectively. In other words, recall expresses the rate at which the model correctly predicts the class of an object; precision measures the fraction of correctly classified objects over the total number of objects labelled with that class; finally, F-Score can be interpreted as the weighted average of the precision and recall, where it reaches its best value at 1 and worst at 0.

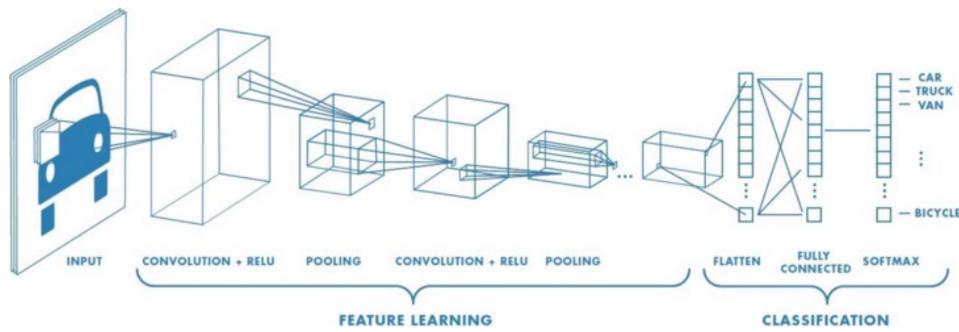


Figure 4.8: Structure of a generic CNN with one channel. Each image patch undergoes feature extraction (composed of several layers of convolution and max pooling) before the final layer can provide the class label (output). The process can be broadly divided in two phases: feature learning (image is processed) and classification (the output is extracted). The initial image size is downsized during the processing. Credits: <https://towardsdatascience.com/>

4.3 CONVOLUTIONAL NEURAL NETWORKS

CNN, a class of artificial NN that has become dominant in various computer vision tasks, is attracting interest across a variety of domains. CNNs are specialised in processing data with grid-like topology (e.g., a homogeneous time series or an image, on which we will focus our discussion) and its name indicates that the network applies a mathematical operation, i.e. a convolution, between the layers. In reality, convolution is only one of the many components in a CNN. The general architecture of a CNN with its principal building blocks (i.e., *convolution*, *pooling*, and *fully connected layers*) is shown in Figure 4.8. The input data is an image on which several operations (or *kernels*) are applied. The first phase (characterised by an alternation of convolution and pooling operations) performs feature extraction, whereas the second (a fully connected layer) maps the extracted features into the final output, for example, a classification. Let us describe in more detail each operation.

Convolution layer. This layer is composed of a collection of *convolutional filters* (i.e., kernels) performing features extraction. In this context, convolution is defined by a specialised type of linear operation (element-wise product and total sum) between the input image and the tensor representing the kernel, as presented in Figure 4.9. The output tensor is a feature map; therefore, applying several kernels on the same image produces an arbitrary number of different feature maps, each one extracting a specific feature of the input data. The key feature in this procedure is weight sharing: kernels are shared across the whole image, allowing for feature invariance in translation. Kernels are the only parameters automatically learnt during the training process in this layer; however, the size and number of the

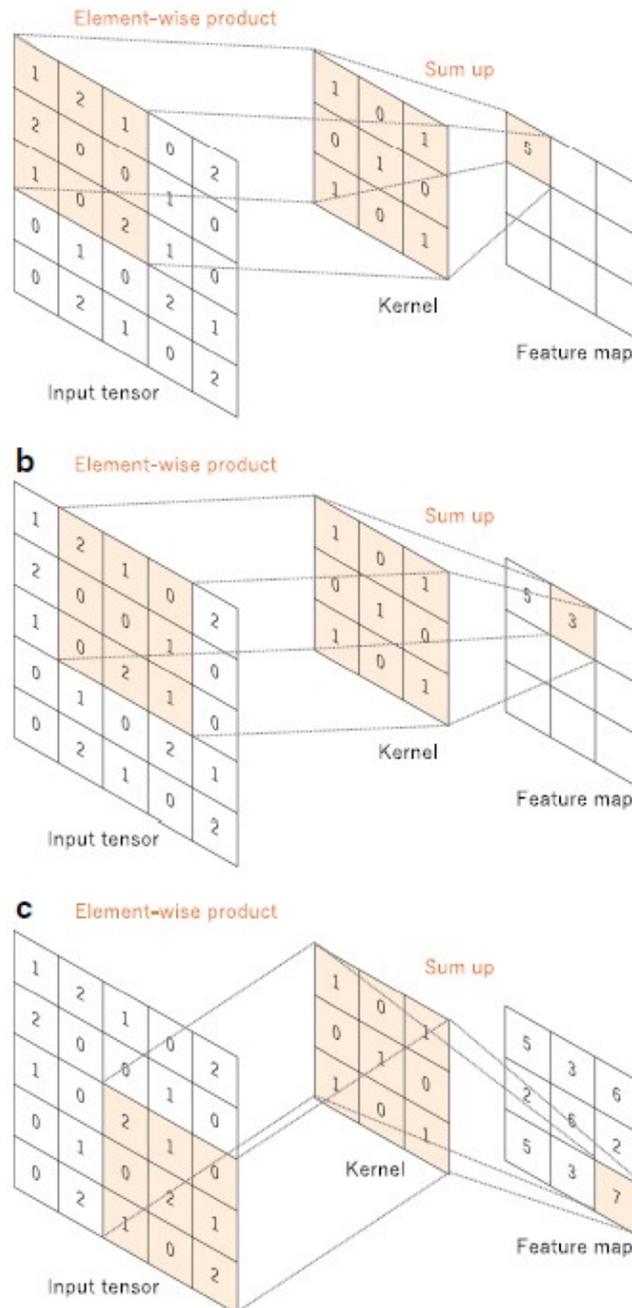


Figure 4.9: Graphical representation of the convolution operation with a kernel size of 3×3 across an input tensor 5×5 . The kernel is applied as an element-wise product between kernel and input at each location and summed to obtain the output value in the feature map. Multiple kernels (i.e., different matrices) work as different feature extractors. Credits: Alzubaidi et al. (2021).

kernel are hyperparameters set by the user at the beginning of the process.

Pooling layer. The downsampling operation is performed by the pooling layer according to the type of function set by the user. This step reduces the dimension of the feature maps in order to introduce a translation invariance to small shifts and distortions, thus effectively decreasing the number of learnable parameters in the next steps and preventing overfitting by increasing the generalisation. The basic idea is to sample only the dominant response, e.g. by taking the maximum (max pooling) or average from a given feature map.

Activation function. After every layer, the output is passed to a non-linear activation function which decides whether or not to fire a neuron with reference to a particular input. Common examples of these functions are the sigmoid and hyperbolic functions, while the most widely used is the Rectified Linear Unit (ReLU), which simply computes: $f(x) = \max(0, x)$. This non-linear performance of the activation layers means that the mapping of input to output will be non-linear; moreover, these layers give CNN the ability to learn extra-complicated things. The activation function must also have the ability to differentiate, which is an extremely significant feature, as it allows error backpropagation¹ to be used to train the network.

Fully connected layer. The last layer in a network is usually the fully connected layer which, unlike pooling and convolution, is a global operation. The input is a flattened one-dimensional array from the previous layer that is fed to a fully connected layer (i.e., every node is connected by a learnable weight). The output is usually passed through another activation function, producing outputs to target class probabilities or normalised.

These components, assembled in different ways, represent the core block of every CNN. In the training phase, the hyperparameters are updated according to the propagation rule chosen by the user.

4.3.1 Regularisation and optimisation

Implementing a DL model requires to undergo a process of testing and improving the model performance, in order to be able to generalise the

¹ Back-propagation and forward-propagation are two ways in which the network can learn to update weights during training. The method calculates the gradient of the error function with respect to the NN's weights. The "backwards" stems from the fact that the calculation of the gradient proceeds backwards through the network, with the gradient of the final layer of weights being calculated first and the gradient of the first layer of weights being calculated last. Partial computations of the gradient from one layer are reused in the computation of the gradient for the previous layer. This backwards flow of the error information allows for efficient computation of the gradient at each layer versus the naive approach of calculating the gradient of each layer separately. In forward-propagation, the computation of the derivative begins from the first layer.

assigned task, while also preventing overfitting. This task goes under *regularisation* which focuses on reducing the test and generalisation error without affecting the initial training error. An easy way to visualise the issue is to imagine fitting several points (drawn from a distribution with a given small scatter) with a polynomial function passing through each of these points. It is likely that, if we were to add another point sampled from the same distribution, the polynomial would not pass through this new point, under-representing any new data point. In other words, the polynomial (i.e., our model) would not be able to generalise the distribution. Although overfitting can be controlled by increasing the size of the training dataset, we can adopt a few precautions, such as weight decay and early stopping.

Weight decay. One explicitly adds a term to the optimisation problem.

This term could be priors, penalties, or constraints and it imposes a cost on the optimisation function to make the optimal solution unique. In simple terms, this penalty can be expressed as a change in the gradient computation $\vec{\nabla} \leftarrow \vec{\nabla} + \lambda \vec{\theta}$, where λ is the weight and $\vec{\theta}$ are the model parameters.

Early stopping. The training phase is stopped when performance on a validation dataset starts to degrade. In other words, we encourage the training phase only when we know that the model is learning through the data without overfitting, as it would happen if the validation error were to increase.

Aside from regularisation, in principle, we can improve our modelling with *optimisation*. Optimisers are algorithms or methods used to change the attributes of a NN (e.g., weights) in order to help the model to converge as fast as possible. One useful class of optimiser is the momentum optimiser.

Momentum optimiser. The idea is to accelerate the gradient descent along the loss function, by reducing the variance in every other insignificant direction. This is attained by the introduction of a momentum vector \vec{m} which stores the changes in previous gradients, since at each timestep \vec{m} is updated with $\vec{m} \leftarrow \beta \vec{m} + \alpha \vec{\nabla} J(\vec{\theta})$. In the previous expression, α is the learning rate, $J(\vec{\theta})$ is the loss function, and β is the momentum. The algorithm helps accelerate the stochastic gradient descent in the relevant direction and dampens oscillations. Parameters are updated by subtracting the momentum vector, i.e. $\vec{\theta} \leftarrow \vec{\theta} - \vec{m}$.

Both optimisation and regulations are included in Adam (Adaptive Moment Estimation, see Kingma and Ba 2014) which we use in our analysis. Adam self-adjusts the stepsize combining two gradient descent methods (i.e., Adaptive Gradient Algorithm, or AGA, and Root Mean Square Propagation, or RMSP). Instead of adapting learning rates based on the average first moment as in RMSP, Adam makes use of the average of both the first and second moments of the loss function. Therefore, the parameter vector of the model is updated with exponential decay for the first and

second moment estimates according to two momentum parameters (i.e., β_1 and β_2). The update rule for the gradient follows the weight decay technique. The ϵ parameter scales the gradient. Tuning the various hyper-parameters mentioned above requires some in-depth studying which we discuss in the next section.

4.3.2 *Learning rate range test*

Finding an optimal learning rate is said to improve performances and allows to reach faster convergence during training. To this scope, firstly introduced by Smith (2017), the Cyclical Learning Rate lets the learning rate cyclically vary between reasonable boundaries, as opposed to monotonically decreasing its value. In practice, the method is applied on a single epoch of training iteration, where the learning rate is increased at every mini-batch of data. Such an increase can be selected between a linear or exponential one, according to the user preference. With each mini-batch, we record the loss function in the learning rate curve as a function of α . The expected trend is what is shown in Figure 4.10: a plateau at the beginning, a descent at intermediate values and eventually explode. Therefore, a recommended choice for optimal learning rates is to pick the value with the sharpest downward slope rather than just the one with the lowest loss (i.e., minimum negative gradient). This suggestion accounts for two precautions.

- (i) By not choosing the lowest value in the loss function, we are taking the safe choice which allows a steep change in the gradient, while simultaneously excluding dramatic jumps.
- (ii) If using an optimiser with momentum then the loss is a moving average of the current mini-batch loss and the past losses. This algorithm calculates the exponential moving average of gradients and square gradients. Consequently, when α starts being too large, the loss plot will not explode immediately as the average will be pulled down by the past losses.

Note that the size of the mini-batch can affect the location of the optimal learning rate: a common DL practice is to take larger learning rates for larger mini-batch sizes since providing more data requires an increased need in updating weights. Furthermore, different weight initialisation can return different non-overlapping suggested ranges.

4.3.3 *The U-Net architecture*

In many visual tasks, the desired output from a DL model is a class label assigned to each pixel, rather than a single class label for the image. Many examples can be drawn in biomedical images, where different portions of the image can reveal different health-related problems. In such class of problems, the U-Net model (Ronneberger et al., 2015) has

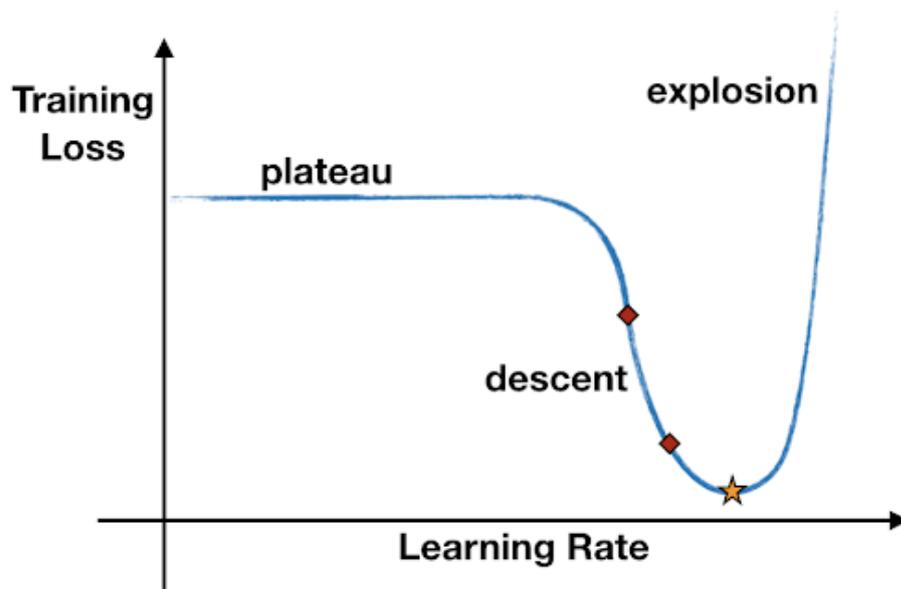


Figure 4.10: Loss function as a function of the learning rate. In the search for the optimal learning rate, the use of this type of graph allows one to search the parameter space thoroughly. During the process, the learning rate goes from a very small value to a very large value (i.e., from 10^{-7} to 100), causing the training loss to start with a plateau, descend to some minimum value, and eventually explode. Credits: <https://blog.dataiku.com>.

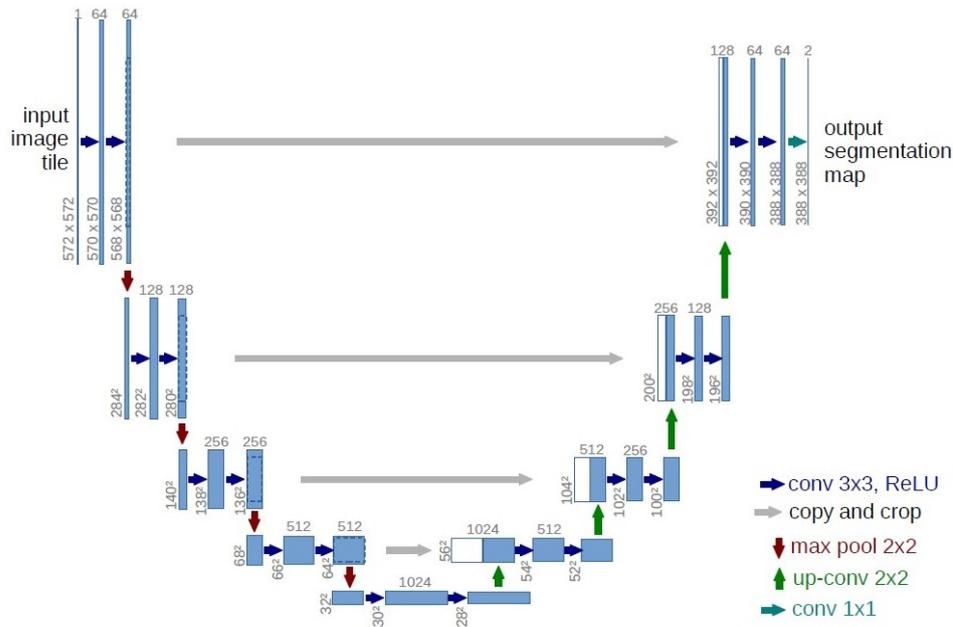


Figure 4.11: U-Net architecture for one channel. The flow of information is denoted by the coloured arrows in the image. Each box corresponds to a multichannel feature map whose size is reported at the bottom left of each rectangle. The size of the kernels is reported on top of each box. Credits: Ronneberger et al. (2015).

proved to provide precise results. For our scopes, it is adapted to resolve pixel-to-pixel regression/classification problems in astronomical images. We employ this model in [Chapter 8](#) to predict the ICL mass fraction in mock images of galaxy clusters, basing our information on the dynamical properties of the stellar population in clusters. More in detail, we aim at predicting such fraction in each pixel to recover complete ICL distribution maps. The U-Net architecture (see [Figure 4.11](#)) consists of a contracting path to capture context and a symmetric expanding path that enables precise localisation. This double action increases the resolution of the output: high-resolution features from the contracting path are combined with the upsampled output allowing better localisation of the features.

The contracting path is built following the typical architecture of a CNN: a theory of 3×3 convolutions (with ReLU activation functions) and 2×2 max pooling operations to downsample. After each max pool, the feature channels are doubled (as it is shown at the top of the rectangular boxes in [Figure 4.11](#)). Every step in the expansive path consists of an upsampling of the feature map followed by a 2×2 convolution that halves the number of feature channels and two 3×3 convolutions, each followed by a ReLU. In the final layer, a 1×1 convolution is used to map each 64-component feature vector to the desired number of classes. Notice that the input channels may be more than 1, thus the final layer combines the different channels in one or more outputs, as required by the classification. In total the network has 23 convolutional layers.

Part III

RESULTS

ON THE PHASE-SPACE STRUCTURE OF SIMULATED GALAXY CLUSTERS

In this chapter, we investigate the phase-space structure of collisionless tracers in galaxy clusters, such as DM particles, stars, and galaxies by comparing simulation predictions to the observational results from nearby clusters. We aim to answer the following scientific questions.

What is a possible dynamical driver of structure formation? Can we competitively describe the phase-space properties of simulated galaxy clusters compared to the observed ones? Can we infer significant differences among the phase-space structure of the different collisionless components reacting to the cluster's gravitational potential?

The content of this chapter largely reflects the content presented in a paper published in Monthly Notices of the Royal Astronomical Society.

- **Marini I**, Saro A, Borgani S, Murante G, Rasia E, Dolag K, Lin W, Napolitano N R, Ragagnin A, Tornatore L, Wang Y "On the phase-space structure of galaxy clusters from cosmological simulations". Monthly Notices of the Royal Astronomical Society, 2021, 500.3: 3462-3480.

5.1 SCIENTIFIC CONTEXT

While several cosmological simulations demonstrated that the equilibrium configuration of DM halos is characterised by a quasi-universal density profile at least out to the virial radius (NFW density profile, Navarro, 1996; Navarro et al., 1997), more recent pieces of evidence (e.g., Taylor and Navarro, 2001; Dehnen and McLaughlin, 2005) have shown that rather than to its density profile, cluster and DM halo formation and evolution might be more deeply connected to another, possibly more fundamental quantity: the "pseudo-entropy" profile $S(r)$. This quantity is defined in terms of the velocity dispersion profile $\sigma_v(r)$ and the density profile $\rho(r)$ of the tracer:

$$S(r) = \frac{\sigma_v^2(r)}{\rho^{2/3}(r)}. \quad (31)$$

The "phase-space density" $Q(r) = S^{-3/2}(r)$ is equivalently discussed in the literature. Empirically, it has been shown that $S(r)$ (or analogously Q) closely follows a power-law in radius in simulated galaxy-size (Taylor and Navarro, 2001) and cluster-size halos (Rasia et al., 2004; Ascasibar et al., 2004), a result that has been confirmed by observations (e.g. Biviano et al., 2013, 2016; Capasso et al., 2019). Taylor and Navarro (2001) found

from N-body simulations $Q \propto r^{-1.82}$ (corresponding to $S(r) \propto r^\alpha$ with $\alpha = 1.21$). Similarly, Rasia et al. (2004) derived $Q \propto r^{-1.85}$, i.e. $\alpha \simeq 1.23$, in agreement with the analysis of Dehnen and McLaughlin (2005). Moreover, the universality of halo density profiles can be recovered starting from the power-law behaviour of the phase-space density profile and the Jeans equation, under the assumption of an isotropic, spherically symmetric equilibrium mass distribution (Dehnen and McLaughlin, 2005). This result motivates the study of pseudo-entropy as a quantity intrinsically connected to the process of halo formation. However, the underlying physical reason leading to the power-law dependence of pseudo-entropy is still unclear.

The power-law behaviour of pseudo-entropy profiles in DM halos was also independently derived by Faltenbacher et al. (2007), starting from the analogy with the entropy of the intracluster gas $S_X(r)$ which is generally defined as $S_X \propto T_g \rho_g^{-2/3}$, where T_g is the gas temperature and ρ_g is the gas density. Spherical gas accretion models predict gas entropy to scale with the clustercentric distance as $S_X \propto r^{1.1}$ (Tozzi and Norman, 2001; Voit et al., 2003). Indeed outside the central region, mostly affected by non-gravitational processes (e.g. Borgani and Kravtsov, 2011), the slope obtained in non-radiative hydrodynamical simulations agrees with the observed values (e.g. Voit et al. 2005) and gas and DM entropy profiles follow one another very closely.

Stars in the main halo, which have been proven to be composed of two different dynamical populations, were further investigated. Several studies of both observational data and simulations have shown the existence of two different stellar components in the main halo of galaxy clusters. A substantial fraction of these stars is confined within the BCG. Dubinski (1998) investigated the origin of the ICL showing that close encounters and halo merging naturally produce a massive central galaxy with surface brightness and velocity dispersion profiles resembling those of the BCGs. The other fraction is not gravitationally bound to any particular galaxy and constitutes the so-called ICL. The distribution of the ICL involves physical scales comparable to those over which the DM component is distributed (e.g. Dubinski, 1998), so it is reasonable to expect that this component traces the global gravitational potential of its hosting cluster (Montes and Trujillo, 2018). Simulations predict that the ICL forms at relatively late times ($z < 1$; e.g., Contini et al., 2013; Monaco et al., 2006; Murante et al., 2007) and it is thought to arise primarily from the tidal stripping of stars from infalling groups and satellite galaxies during the hierarchical accretion of the cluster (e.g. Murante et al., 2004). We test the differences in their dynamical properties in relation to the pseudo-entropy profiles traced singularly by the two to disentangle the distinct contributions to the pseudo-entropy profile of all the stars. Indeed, in the hypothesis of a strong correlation between distinct formation mechanisms and phase-space structure, we expect to detect a corresponding difference in the pseudo-entropy profiles.

In recent years, some observational analyses have deepened the study on the pseudo-entropy as traced by the hosted galaxies in galaxy clusters

(Biviano et al., 2013; Annunziatella et al., 2016; Capasso et al., 2019). These results have shown the existence of the power-law feature, also for this tracer. We plan to provide a computational counterpart in that regard: the dynamics of real galaxies is expected to be traced by the substructures within the simulated halos.

5.2 SIMULATIONS

We analyse a set of DM-only simulations and a set of hydrodynamical simulations (Hydro). For each set, we carried out simulations at two different levels of resolution. At the base resolution (1x hereafter), for the Hydro set we adopt a DM mass particle of $8.3 \times 10^8 h^{-1} M_{\odot}$ and an initial mass of the gas particle of $3.3 \times 10^8 h^{-1} M_{\odot}$. At intermediate (10x hereafter) resolution, we have both DM-only and hydrodynamical simulations, while at high resolution (100x hereafter) we include a set of DM-only simulations. The characteristics of the four sets of simulations are summarised in Table 3.1.

The set of simulated clusters, named DIANOGA (Bassini et al., 2020, and references therein), were extracted from a parent N-body box of size $1 h^{-1}$ Gpc and resimulated adopting the zoom-in technique as implemented by Tormen and Bertschinger (1996). The adopted cosmology is a Λ CDM model with $\Omega_M = 0.24$, $\Omega_b = 0.037$ for the total matter and baryon density parameters, $n_s = 0.96$ for the slope of the primordial power spectrum, $\sigma_8 = 0.8$ for the normalisation of the power spectrum, $h_0 = 0.72$ for the Hubble parameter in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. At the base resolution, each simulation describes the evolution of Lagrangian regions centered on the 24 most massive clusters in the initial simulated box, all having mass $M_{200} \geq 8 \times 10^{14} h^{-1} M_{\odot}$ and 5 isolated smaller systems with M_{200} in the range $(1 - 4) \times 10^{14} h^{-1} M_{\odot}$. For the other resolutions, the number of Lagrangian regions varies for each set, we report the exact numbers in the last row of Table 3.1. Additionally, we point out that the set of Hydro-1x simulations is the one originally presented in Ragone-Figueroa et al. (2018), while the set of Hydro-10x has been presented in Bassini et al. (2020).

5.2.1 Dynamical state

Estimates of the dynamical state of single clusters became important when investigating the impact of the internal equilibrium of these systems in relation to their pseudo-entropy profile. We classify only the clusters in the Hydro-1x sample. This is performed following the prescription described in Biffi et al. (2016) to which we refer for further details.

The method is based on two properties: the centre shift (identified as the distance between the position of the minimum of the gravitational potential \mathbf{x}_{\min} and the centre of mass \mathbf{x}_{cm}) and the fraction of mass in

substructures f_{sub} . A halo is classified as relaxed if both the following conditions are satisfied:

$$\begin{cases} \delta r = \|\mathbf{x}_{\text{min}} - \mathbf{x}_{\text{cm}}\|/r_{200} < 0.07 \\ f_{\text{sub}} = \frac{M_{\text{TOT,sub}}}{M_{\text{TOT}}} < 0.1 \end{cases} \quad (32)$$

where M_{TOT} is the total mass and $M_{\text{TOT,sub}}$ is the total mass in substructures. If neither is satisfied, then the cluster is classified as disturbed, while it is tagged as partially disturbed if only one of the above two criteria is not satisfied. After applying this classification to the 29 main clusters of the Hydro-1X set at redshift $z = 0$, we find 6 relaxed, 8 disturbed systems, and 15 intermediate cases.

5.3 SELF-SIMILARITY OF PSEUDO-ENTROPY PROFILES

Ludlow et al. (2010) found that the pseudo-entropy profiles may not hold the power-law behaviour when approaching the virial radius due to the proximity of the last accreted shell. We aim to broaden the analysis of the universality of the pseudo-entropy profiles traced by different collisionless components in the simulated clusters.

Our analysis is extended out to redshift $z = 2$, which is the epoch when massive clusters assemble. All the particle positions and velocities are computed in the rest-frame of the cluster centre, which is identified as the particle within the central FoF group or the main halo having the minimum value of the gravitational potential. Cluster radii are defined in units of the virial radius r_{vir} , to better capture the universal behaviour of the self-similar scaling of the simulated clusters. The virial radius of a halo at redshift z is defined as the radius encompassing a mean halo density of $\Delta_{\text{vir}}(r)\rho_c(z)$, where $\rho_c(z)$ is the critical cosmic density at redshift z and Δ_{vir} is the redshift-dependent virial overdensity predicted by spherical-collapse for a given cosmological model (e.g. Bryan and Norman, 1998; Eke et al., 1996)¹. We provide here below a short description of the scaling associated with the velocity dispersion profiles, and consequently with the pseudo-entropy profiles, while the full derivation is provided in the Appendix.

In the self-similar model, particles within a sphere of radius r_{vir} at redshift z have a measured velocity dispersion profile $\widetilde{\sigma}_v(r, z)$ that scales as a function of the virial radius r_{vir} (i.e. halo mass) and redshift

$$\sigma_v(r, z) = \frac{\widetilde{\sigma}_v(r, z)}{r_{\text{vir}}} \left[\frac{\Delta_{\text{vir}}(z)}{\Delta_{\text{vir}}(0)} E^2(z) \right]^{-1/2}, \quad (33)$$

¹ Similarly, we can define r_Δ as the radius encompassing a mean overdensity equal to Δ times the critical density of the universe at that redshift $\rho_c(z)$.

where $E(z) = [\Omega_M(1+z)^3 + \Omega_\Lambda]^{1/2}$ provides the redshift dependence of the Hubble parameter: $H(z) = E(z)H_0$. From this relation, we derive the scaling on the measured pseudo-entropy $\tilde{S}(r, z)$:

$$S(r, z) = \frac{\tilde{S}(r, z)}{r_{\text{vir}}^2 E^{2/3}(z) \left(\frac{\Delta_{\text{vir}}(z)}{\Delta_{\text{vir}}(0)}\right)^{1/3}}. \quad (34)$$

We show in [Figure 5.1](#) how the rescaling proposed in [Equation 33](#) and [34](#) effectively captures the universal behavior of the individual pseudo-entropy profiles. More in detail, we illustrate the radial profiles of density, velocity dispersion, and pseudo-entropy, as traced by DM particles in the main halo given by Subfind within $2 r_{\text{vir}}$ for the entire cluster sample of the Hydro-10x simulation at $z = 0$. Individual cluster density profiles are shown as thin grey lines, while the solid black line represents the median profile. We note that no further scaling needs to be applied to the density profiles (upper panel) once the radial distance is expressed in terms of the virial radius. On the other hand, simply rescaling radii in units of the virial radius is not sufficient to properly capture the universality of the other two phase-space quantities (velocity dispersion $\tilde{\sigma}_v$ and pseudo-entropy \tilde{S} profiles), shown in pink thin lines in the central and bottom panels. For velocity dispersion and pseudo-entropy profiles, universality is recovered once such quantities are expressed in terms of σ_v and S , or in other words, are rescaled according to [Equation 33](#) and [34](#) (thin grey lines).

[Figure 5.1](#) confirms the remarkable power-law shape of the pseudo-entropy profiles (dashed black line) of the DM component, which is stable from the innermost resolved radius out to nearly the virial radius, in simulations including hydrodynamics and baryonic physics. As discussed in [Ludlow et al. \(2010\)](#), the outer region is most likely associated with the transition from the inner, relaxed parts, to the dynamically more active outer parts, where infalling material has not yet had time to undergo phase-mixing and relaxation. Such an upturn is present also in the self-similar solution of [Bertschinger \(1985\)](#) and it might be a general feature of the outer pseudo-entropy profiles of DM halos. On the other hand, the density profile corresponding to the Bertschinger solution (a power-law with constant slope) differs significantly from the density profiles of DM halos as shown in the top panel of [Figure 5.1](#), which are better described by NFW profiles, whereby the logarithmic slope smoothly changes from -1 in the central regions to -3 in the outer regions. Similarly, in the central panel, we show velocity dispersion profiles, which also clearly show departures from a scale-free behaviour. As the main driver of the mechanisms involved in halo formation (phase-mixing and violent relaxation) is gravity (which has a scale-free behaviour), it is reasonable to expect that closely associated phase-space density quantities retain a scale-invariant behaviour. A simple power-law suggests, therefore, the possibility of interpreting the pseudo-entropy as a key quantity in structure formation, or in other words a dynamical attractor more fundamental than either the velocity dispersion or density profile which individually do not have a power-law trend

(surprisingly given the fact that it is derived from the combination of the density and the velocity dispersion profiles which, we stress, are not power-laws if taken singularly). In the following sections, we present our results on the simulated clusters in more detail, focusing on the different tracers of the phase-space, namely DM particles, substructures, stars belonging to the BCG and the diffuse stellar component surrounding the BCG.

5.4 RESULTS

5.4.1 Pseudo-entropy profiles traced by DM

Having motivated why the pseudo-entropy profile traced by DM particles is considered a fundamental diagnostic for the description of halo formation, we now investigate its behaviour for simulations with different resolutions and including the description of different physical processes as well as studying its evolution. This will allow us to assess the robustness of its shape against both numerical resolution and physical processes that add to gravitational instability.

5.4.1.1 The effect of resolution

As shown in [Figure 5.1](#), DM particles in simulated clusters distribute in phase-space in such a way as to predict a tight power-law shape of the pseudo-entropy profile. [Figure 5.2](#) presents the pseudo-entropy profiles traced by the DM particles in the stack of clusters common to the four sets of simulations, all at $z = 0$. The upper and the central panels compare each a pair of simulation sets including the same physics, but with different resolutions: fully hydrodynamics and DM-only simulations, respectively. The bottom panel compares instead DM-only and Hydro simulations at the same resolution. Despite resolving structures with different sensitivity, due to the different mass resolutions and the presence of baryons, the phase-space robustly describes a similar power-law profile in all cases. We report the median profiles (solid lines) and the 68 per cent dispersion (shaded areas) given by the cluster samples. The profiles are all in excellent agreement with each other, with small differences only in the innermost regions, where the effects of both resolution and baryonic processes become more relevant.

We then quantify the dependence of the pseudo-entropy profiles, traced by DM particles, on the resolution by computing the normalisation S_0 and the logarithmic slope α for the DM-only and Hydro runs. We assume a power-law as:

$$S(r) = S_0 \left(\frac{r}{0.5 r_{\text{vir}}} \right)^\alpha. \quad (35)$$

The fitting procedure is carried out with *emcee* (Foreman-Mackey et al., 2013), a Python implementation of the affine-invariant ensemble sampler for Markov Chain Monte Carlo. We use the median profile and the associated error to fit the profiles. Results are reported for the four sets of

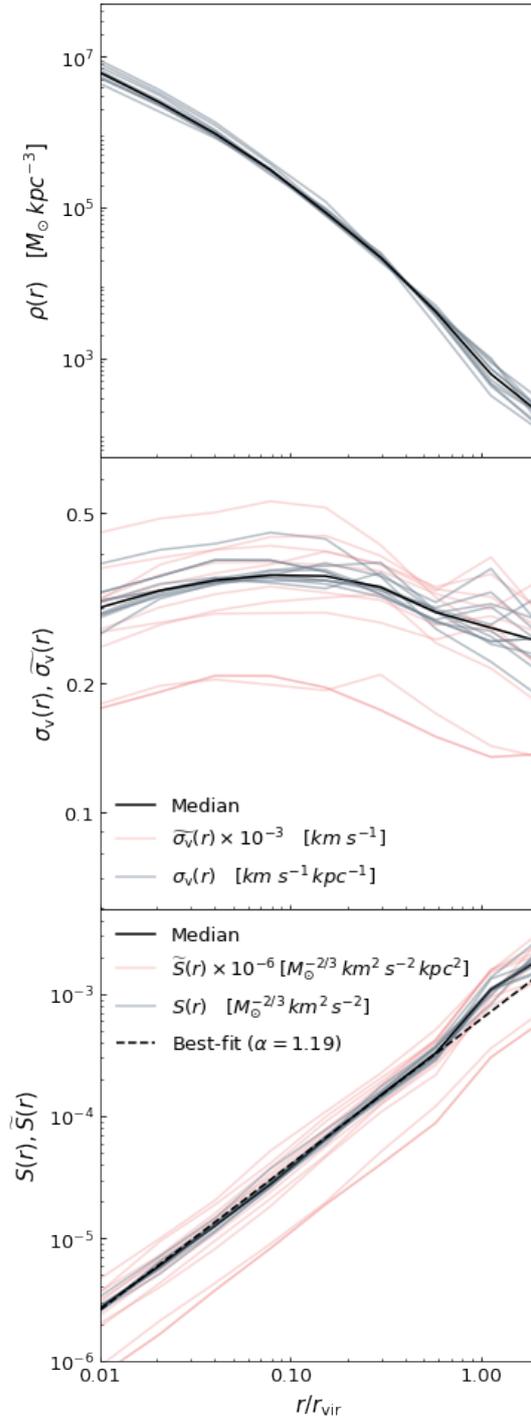


Figure 5.1: From the top panel: density, velocity dispersion, and pseudo-entropy profiles as traced by the DM particles in the Hydro-10x at $z = 0$. The coloured lines report the profiles traced by each cluster while the black is the median. In the middle panel, the grey and pink lines represent the velocity dispersion profiles $\sigma_v(r)$ and $\tilde{\sigma}_v(r)$ of each cluster respectively scaled and not-scaled according to Equation 33. The black lines trace the median profiles. Similarly, in the bottom panel, the grey and pink lines represent the scaled and not-scaled profiles of the pseudo-entropy $S(r)$ according to Equation 34, the black solid line illustrates the median trend while the dashed one is the best-fit.

Table 5.1: Best-fit results for the normalisation S_0 and exponent α in the power-law expression of Equation 35 for the median pseudo-entropy profiles at $z = 0$ traced by DM particles.

	S_0	α
Hydro-1x	$(2.79^{+0.04}_{-0.04}) \times 10^{-4}$	$1.20^{+0.01}_{-0.01}$
Hydro-10x	$(2.72^{+0.05}_{-0.05}) \times 10^{-4}$	$1.19^{+0.01}_{-0.01}$
DM-10x	$(2.32^{+0.02}_{-0.02}) \times 10^{-4}$	$1.23^{+0.01}_{-0.01}$
DM-100x	$(2.39^{+0.06}_{-0.06}) \times 10^{-4}$	$1.25^{+0.01}_{-0.01}$

simulations in Table 5.1. We find that results on the normalisation and slope for both the DM-only and the Hydro runs are consistent within 2σ for the different resolutions, thus guaranteeing the convergence of our results against the resolution.

5.4.1.2 The impact of baryons

The bottom panel of Figure 5.2 shows the profiles for DM-only (in green) and Hydro (in red) simulations at the same resolution. Pseudo-entropy profiles derived from the DM particles in the Hydro runs have been rescaled by $(1 - \Omega_b/\Omega_M)^{2/3}$ to properly compare them to their DM-only counterparts. After this correction, we observe the two profiles to be almost indistinguishable over the entire radial range, besides the core regions, where the impact of baryons mostly influences the distribution of the DM particles. A new fit of the Hydro simulations, considering the aforementioned correction factor, gives values consistent with the DM-only results: for the Hydro-1x $S_0 = 2.48^{+0.04}_{-0.04}$ and the Hydro-10x $S_0 = 2.42^{+0.04}_{-0.04}$. We note that the logarithmic slope α of the Hydro run profiles is slightly shallower ($\simeq 3\sigma$) than its DM-only counterpart. Thus, the emerging picture of the pseudo-entropy profiles traced by the DM particles in the simulations agrees with the general result of a power-law with a fixed slope, a result that is supported against the numerical resolution, indicating that this is a key quantity in the description of the gravity-driven collapse of non-linear structures.

5.4.1.3 Evolution in redshift

Figure 5.3 describes the redshift evolution of the pseudo-entropy profiles. The top panel shows the median profiles traced by DM particles in the Hydro-10x runs for 6 different redshifts within $0 \leq z \leq 1.6$. The bottom panel illustrates the median profiles when traced by DM particles in the DM-10x. To not overload the plot we do not show the associated 68th percentile regions (which we verify do not change significantly as a function of redshift). In both cases, the close similarity of the profiles highlights the self-similar scaling of the redshift evolution of pseudo-

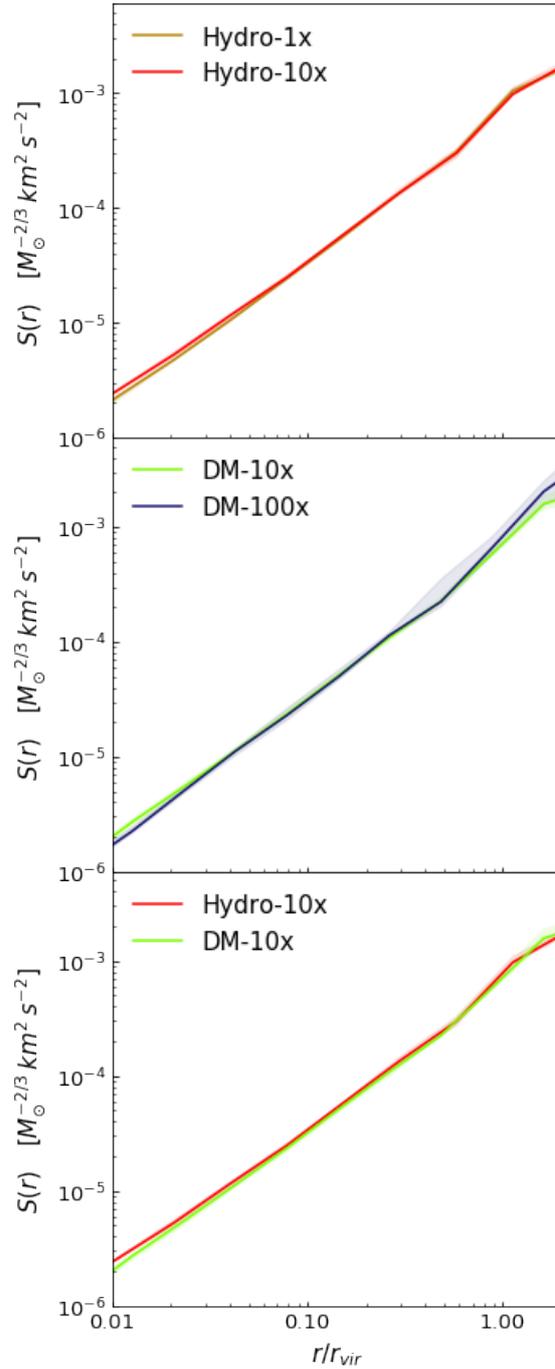


Figure 5.2: Comparison of the pseudo-entropy profiles traced by DM particles at redshift $z = 0$ to test the convergence among the distinct simulations. In the top panel, we compare the median pseudo-entropy profiles obtained from the clusters in the Hydro-1x (in orange) and Hydro-10x (in red) runs. In the central panel, we show the median pseudo-entropy profiles obtained from the DM-10x (in green) and DM-100x (in blue) simulations. The bottom panel reports the median pseudo-entropy profiles from the simulated clusters in the Hydro-10x (in red) and DM-10x (in green). For all the curves we plot the 68th percentile region given by the cluster samples as a shaded area.

entropy profiles, as discussed in [Section 5.3](#). In particular, we note that the profiles are well described by a power-law behaviour, within the considered redshift range.

The hydrodynamical run presents some tension in the innermost regions against the pure self-similar behaviour in the DM-only case. These small deviations emerge for clusters at early times, because they tend to present higher entropy profiles with respect to their lower-redshift descendant. This reflects in a resulting systematic trend of decreasing in the slope α as a function of redshift.

5.4.1.4 The $M_{\text{vir}} - S_{\text{vir}}$ relation

Simple scaling relations between basic cluster properties, such as the total virial mass M_{vir} and the dispersion velocity within the virial radius $\sigma_{\text{v, vir}}$ (Bryan and Norman, 1998; Borgani et al., 1999; Evrard et al., 2008; Munari et al., 2013; Saro et al., 2013), are naturally predicted by the self-similar model (Kaiser, 1986, 1992). From the analysis of an extended set of N-body simulations of galaxy clusters, Evrard et al. (2008) found that massive DM halos closely adhere to the relation

$$\sigma_{\text{v, vir}}(M_{\text{vir}}, z) = \sigma_{\text{v, 15}} \left(\frac{E(z)M_{\text{vir}}}{10^{15}h^{-1}M_{\odot}} \right)^{\gamma} \quad (36)$$

with a remarkably modest scatter $\sigma_{\sigma_{\text{v, vir}}|M_{\text{vir}}} \simeq 0.04$, where $\sigma_{\text{v, 15}} = 1082.9 \pm 4.0 \text{ km s}^{-1}$ is the normalisation at mass $10^{15} h^{-1} M_{\odot}$ and $\gamma = 0.3361 \pm 0.0026$ is the logarithmic slope, found to be within the virial expectation $\gamma = 1/3$ considering the associated uncertainty. Also, we assumed

$$E(z) = [\Omega_{\Lambda, 0} + (1 + \Omega_{\text{M}, 0} - \Omega_{\Lambda, 0})(1 + z)^2 + \Omega_{\text{M}, 0}(1 + z)^3]^{1/2} \quad (37)$$

where $\Omega_{\text{X}, 0}$ are the corresponding energy (mass) content at the present time. The tight scatter in this relation makes $\sigma_{\text{v, vir}}$ a rather accurate mass proxy. The best-fit scaling for our $M_{\text{vir}} - \sigma_{\text{v, vir}}$ relation is quite close to the virial expectation: we obtain $\gamma = 0.347 \pm 0.013$ and an intrinsic logarithmic scatter of $\sigma_{\sigma_{\text{v, vir}}|M_{\text{vir}}} = 0.048 \pm 0.007$, corresponding to a fractional uncertainty in mass at fixed observable of $\sigma_{M_{\text{vir}}|\sigma_{\text{v, vir}}} = 0.132 \pm 0.003$.

Given the strong similarity of pseudo-entropy profiles, we can argue whether pseudo-entropy computed within the virial radius, could also provide an accurate, low-scatter halo mass proxy. To this purpose, we define S_{vir} to be the integrated pseudo-entropy enclosed within the virial radius r_{vir} :

$$S_{\text{vir}} = 4\pi \int_0^{r_{\text{vir}}} \tilde{S}(r) r^2 dr. \quad (38)$$

We assume the following scaling of the integrated pseudo-entropy with virial mass:

$$S_{\text{vir}}(M_{\text{vir}}, z) = S_{15} \left(\frac{E(z)M_{\text{vir}}}{10^{15}h^{-1}M_{\odot}} \right)^{\gamma'}. \quad (39)$$

We study DM particles in clusters from the Hydro-1x simulation at $z = 0$ and compare the results obtained computing S_{vir} and the 3D velocity dispersion $\sigma_{v,\text{vir}}$ as reported in the top and bottom panel of [Figure 5.4](#) respectively. For the $M_{\text{vir}} - S_{\text{vir}}$ relation, we find the slope $\gamma' = 1.74 \pm 0.05$ with an intrinsic logarithmic scatter of $\sigma_{S_{\text{vir}}|M_{\text{vir}}} = 0.20 \pm 0.03$, therefore corresponding to a fractional uncertainty in mass at fixed observable S_{vir} equal to $\sigma_{M_{\text{vir}}|S_{\text{vir}}} = 0.11 \pm 0.02$. This is a strong indication that the integrated pseudo-entropy might be a better mass proxy than the velocity dispersion because its scatter against halo mass is even smaller than that of $\sigma_{v,\text{vir}}$. This would be of particular relevance because, to estimate both quantities, one needs the same information on the cluster dynamics. However, to fully verify this claim, one requires a larger sample of clusters over a reasonably wider mass range to enhance the statistics. Moreover, observationally speaking, the integrated pseudo-entropy is recovered from the combination of two quantities which both carry their uncertainties, thus one might expect to have an increase in the internal distribution due to the associated observational scatter.

5.4.1.5 *Outliers*

Self-similarity can be broken down if the scale-free evolution of a halo is distressed. In this event, one expects the disturbed cluster to not follow the scaling relation, but rather to be an outlier in the overall distribution. To this end, we note the presence of an obvious outlier in the $M_{\text{vir}} - S_{\text{vir}}$ relation shown in [Figure 5.4](#) represented by the second most massive cluster which we signal in the plot with a black cross. To understand the nature of the outlier, we illustrate in [Figure 5.5](#) the recent evolution of its density maps. The four panels follow the density maps traced by DM particles within a region centred on the cluster centre with size $4 r_{\text{vir}}$, at four redshifts. Brighter colours indicate higher densities. The white circle marks the virial radius. The red smaller circle indicates a second halo which is falling into the cluster potential and reaching the central regions at around $z = 0.13$, as displayed in the left bottom panel. The size of the circle is equal to the virial radius of this second halo, as provided by Subfind. At this late redshift, the second halo has crossed the volume enclosed by the virial radius of the main halo and it is being incorporated. This merging process is completed by $z = 0$. The mass ratio of 1 : 5 is fairly large, thus the recent merging represents an event that strongly impacts the dynamical equilibrium of the main halo. Indeed, the sudden change in the internal dynamics is reflected in the pseudo-entropy, which significantly increases at late times, while the system has not yet had the time to virialise and settle into a new equilibrium. The same tension is not registered as significantly in the velocity dispersion distribution. It seems plausible that the recent major merger may have affected the pseudo-entropy in a much stronger way than it has on the velocity dispersion. If this is the case, we expect in the near future that the system will virialise and reduce the scatter with the scaling relation. Therefore, this reasoning advocates that entropy (or pseudo-entropy) has the potential of being a good estimator for

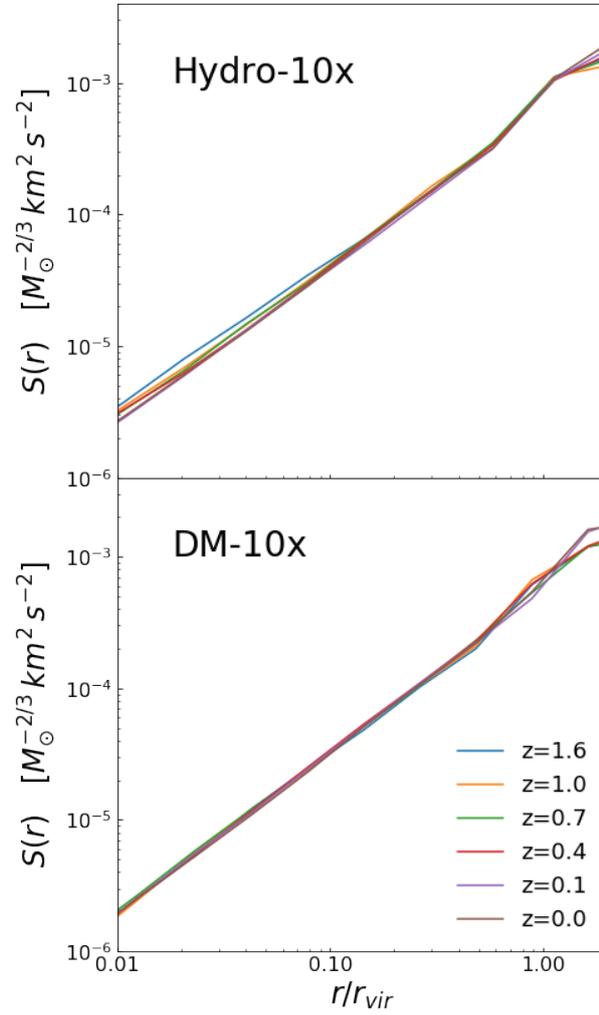


Figure 5.3: Top panel: median of the pseudo-entropy profiles traced by DM particles at different redshifts (as indicated in the legend in the bottom panel) for the clusters in the Hydro-10x simulation. Bottom panel: median of the pseudo-entropy profiles traced by DM particles at different redshifts for the clusters in the DM-10x simulation.

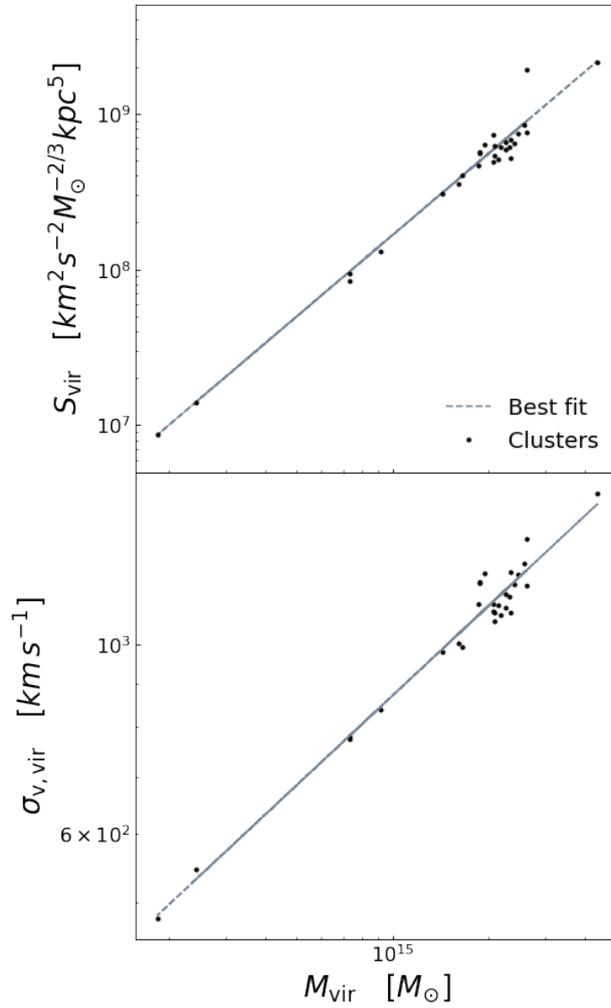


Figure 5.4: Scaling relation of the integrated pseudo-entropy (upper panel) and the virial velocity dispersion (lower panel), both computed over all the DM particles within r_{vir} , as a function of the virial mass of the clusters in the Hydro-1x simulation at $z = 0$. The dashed grey lines are the best-fit profiles for the scaling relations. The second most massive halo in the sample is marked with a black cross to signal an outlier in the distribution of the integrated pseudo-entropy profile, as pointed out in [Section 5.4.1.4](#).

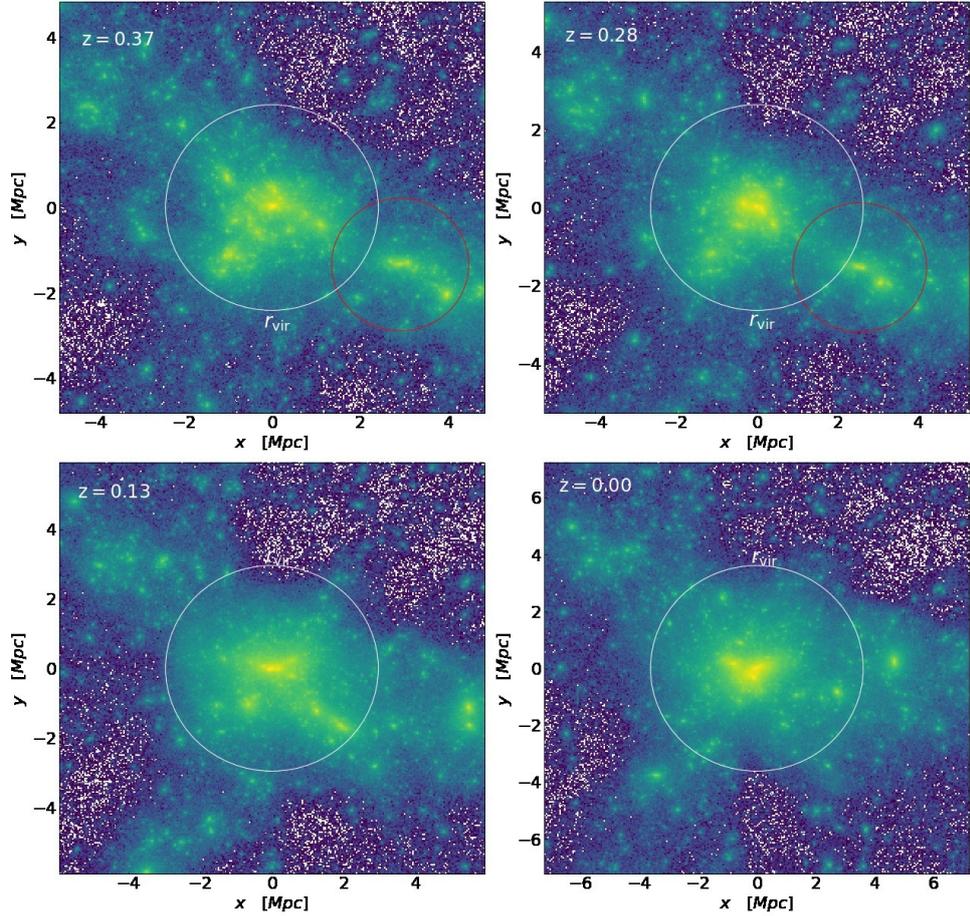


Figure 5.5: The evolution of the density maps in logarithmic scale traced by the DM particles in the second most massive cluster reported in the top panel of Figure 5.4 which appears as an outlier from the $S_{\text{vir}} - M_{\text{vir}}$ scaling relation. The region is centred at the cluster centre and spans a square of 4 times the virial radius of the main halo. The white circle marks the virial radius. The red smaller circle marks the trajectory of the second most massive halo in the region and its growth given that the radius of the circle is equal to its virial radius (as provided by Subfind). The orbiting object at redshift $z \simeq 0.13$ crosses the virial radius to merge with the main halo by redshift $z = 0$.

detecting recent major mergers. After removing this outlier, the logarithmic scatter is further reduced to $\sigma_{S_{\text{vir}}|M_{\text{vir}}} = 0.12 \pm 0.02$, with $\alpha = 1.72 \pm 0.03$, which in turn corresponds to a lower fractional uncertainty in mass at fixed observable S_{vir} equal to $\sigma_{M_{\text{vir}}|S_{\text{vir}}} = 0.067 \pm 0.003$.

Inferring S_{vir} from observations relies on integrating the pseudo-entropy profile within the virial radius, whose knowledge is equivalent to that of the virial mass. Therefore, the use of S_{vir} as a mass proxy may be plagued by a circularity in the argument. The issue can be addressed via an iterative procedure, which is similar in spirit to that described by Kravtsov et al. (2006) for estimating cluster masses from the Y_X mass proxy. This quantity is defined as the product of gas mass and core-excised ICM temperature, both estimated within R_{500} from X-ray observations. In fact, the procedure allows estimating the mass M_{vir} when one does not know a priori r_{vir} . Our approach would require relying on the velocity dispersion $\sigma_{v,\text{vir}}$ in place of the X-ray temperature, to make a first rough estimate of the virial radius through a $M_{\text{vir}} - \sigma_{v,\text{vir}}$ relation. By relying then on a pre-calibrated $M_{\text{vir}} - \sigma_{v,\text{vir}}$ relation (e.g. from high-quality observations of a selected cluster sample and/or from simulations), one can then compute r_{vir} . The procedure can be iterated until convergence. While exploiting the potential of $\sigma_{v,\text{vir}}$ as a mass proxy goes beyond the scope of this paper, we plan to address this issue in a future analysis.

In conclusion, we have shown that pseudo-entropy is not only a faithful tracer of the phase-space structure of a halo but also a potentially useful proxy of its total mass (in the absence of recent major mergers), thus making it an interesting tool for both dynamical studies of galaxy clusters and their cosmological application.

5.4.2 Pseudo-entropy profiles traced by substructures

Having established a remarkable regularity in the pseudo-entropy structure of the DM halo component, we now move to the analysis of the same quantity as traced by substructures. As discussed in Section 5.1, observational studies (e.g., Biviano et al., 2013, 2016; Capasso et al., 2019) demonstrated the power-law relation of the pseudo-entropy profile traced by galaxies in clusters. In our simulated clusters, *bona fide* galaxies correspond to gravitationally bound substructures, which we identify through the Subfind algorithm (see Springel et al., 2001). We estimate the phase-space halo structure, as traced by such substructures, from their number density profiles $N(r)$ and velocity dispersion profiles $\sigma_v(r)$,

$$S(r) = \frac{\sigma_v^2(r)}{N^{2/3}(r)}. \quad (40)$$

This case differs from the previous one with DM particles since the density employed is not the mass density, but rather the substructure number density within each cluster and the velocity dispersion profiles are derived from the statistical distribution of velocities of the substructures. For this analysis, we used the full hydrodynamical set (Hydro-1x and Hydro-10x)

although we show in [Figure 5.6](#) only the profiles of the 10 clusters in the Hydro-10x. The plot illustrates the resulting number density (upper panel), velocity dispersion (central panel), and pseudo-entropy (lower panel) profiles of the single clusters along with the associated median value (solid black). In each panel, we show with the dashed curve the corresponding median profile obtained for the DM particles (as seen in [Figure 5.1](#)). The density (and correspondingly, the pseudo-entropy) is normalised to match the substructures number density profiles at $0.5 r_{\text{vir}}$: at these large radii the two distributions are very close to each other and with this normalisation, one can better appreciate the differences in the central region.

Note that the same universal rescaling with mass and redshift discussed in [Section 5.3](#) has been applied to the quantities shown in [Figure 5.6](#). Although we do not report the not-scaled $\tilde{\sigma}_v$ and \tilde{S} as we did in [Figure 5.1](#), the internal scatter within profiles is significantly reduced after applying [Equation 33](#) and [Equation 34](#). The fact that both the number density and velocity dispersion profiles of substructures present the same universal scaling as the mass-density and velocity dispersion profiles traced by DM particles confirms that the self-similarity of the gravity-driven internal dynamics of clusters is preserved when traced by substructures.

As for the density profiles, we note that substructures trace profiles that are shallower than the NFW profile traced by DM particles. This result confirms previous findings (eg. Saro et al., 2006; Bosch and Ogiya, 2018; Green and Bosch, 2019) which pointed out that tidal removal of mass from merging substructures makes them more fragile in the central cluster regions, thus causing the corresponding number density profiles to flatten with respect to that traced by DM. While the velocity dispersion profiles traced by DM particles and substructures look more similar to their density profiles, we still see that substructures are characterised by a generally higher velocity dispersion, an effect that is more pronounced in central regions. This velocity bias, that has been also pointed out in previous studies (e.g., Munari et al., 2013; Diemand et al., 2004; Faltenbacher et al., 2005; Faltenbacher et al., 2007; Faltenbacher and Diemand, 2006; Lau et al., 2009; Armitage et al., 2018), is due to the effect of tidal stripping which is more effective for substructures moving with lower orbital speed. As a result, these structures tend to merge into the main halo. This effect turns into a selective removal of lower-velocity substructures, thereby increasing the velocity dispersion of substructures. The resulting profiles of pseudo-entropy are thus shallower than those of DM particles, an effect that is mainly driven by the change in the density profiles.

The best-fitting parameters describing the power-law shape of the pseudo-entropy profiles are reported in [Table 5.2](#) for the Hydro-1x and Hydro-10x sets of simulated clusters. Confirming the visual impression from [Figure 5.6](#), the slope $\alpha \simeq 0.9$ is shallower than the one of $S(r)$ traced by DM particles. Furthermore, these profiles are robust against the resolution, both in shape and in normalisation.

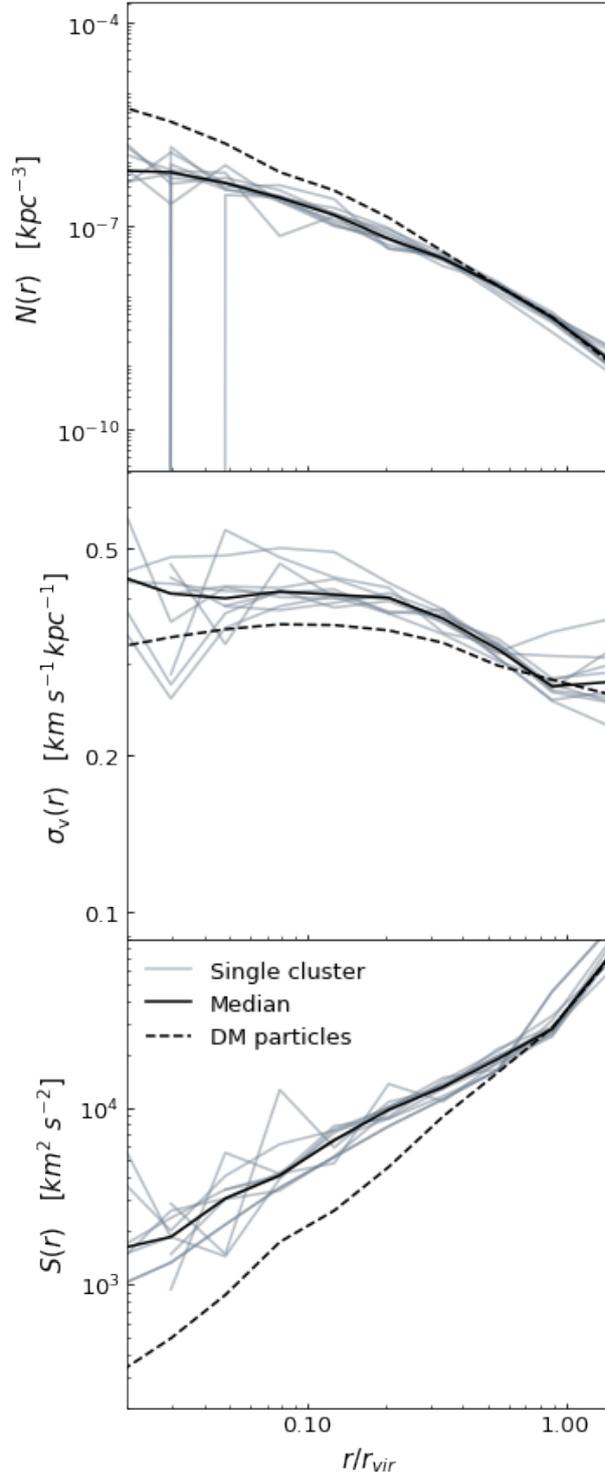


Figure 5.6: From top to bottom panel: number density, velocity dispersion, and pseudo-entropy profiles as traced by the substructures in the Hydro-10x simulation at $z = 0$. The grey lines show the single cluster profiles, whereas the black ones reproduce the median profiles. We also report with the dashed black lines the DM particles profiles (as seen in [Figure 5.1](#)) normalised to match the number density profile at $0.5 r_{\text{vir}}$.

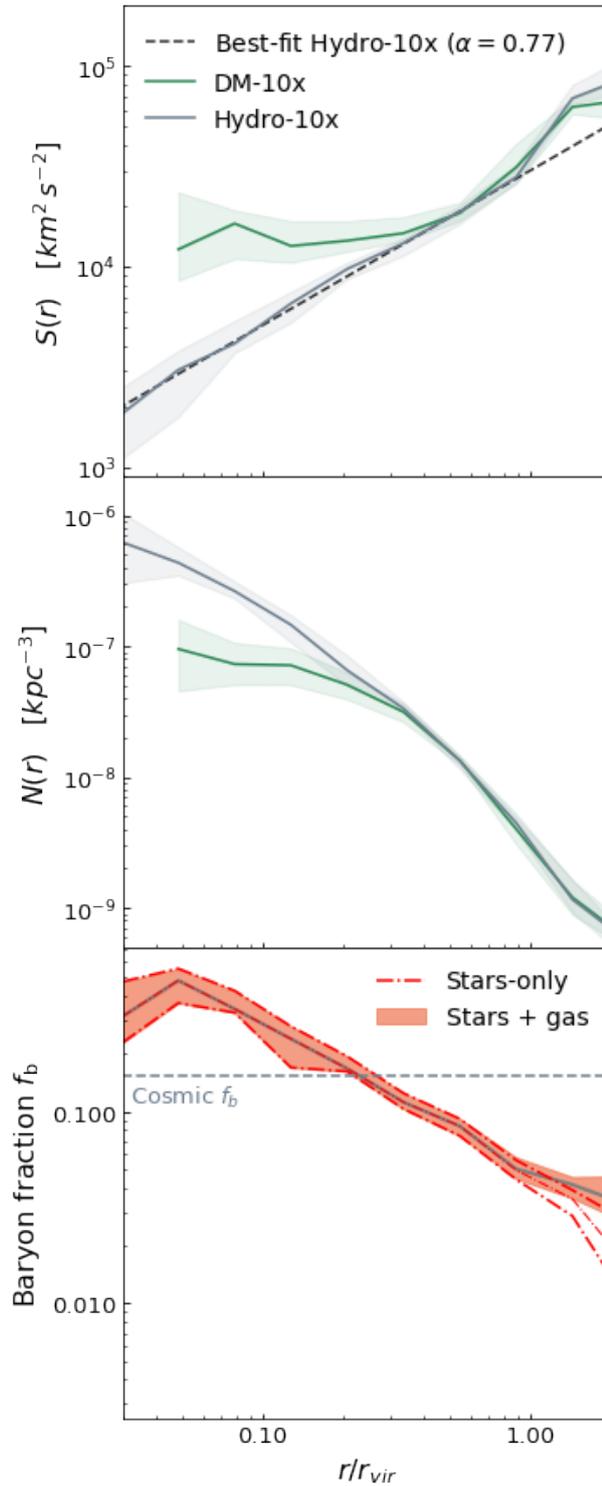


Figure 5.7: Comparison of profiles traced by substructures in DM-10x and Hydro-10x simulated clusters. Upper and central panels: the pseudo-entropy and number density profiles traced by the stack sample of substructures at $z = 0$. Solid curves are for the median profiles, while the shaded areas encompass the 16th and 84th percentiles. In the upper panel, the dashed line shows the best-fit power-law relation for the pseudo-entropy profile of substructures in the Hydro-10x simulations (plotted in Figure 5.6). Lower panel: The baryon fraction profile of substructures in the Hydro-10x simulations. Green and dot-dashed red curves are the median profiles of the baryon fraction when including stars and gas, and only stars, respectively. The horizontal dashed line marks the cosmic baryon fraction of the simulations ($\Omega_b/\Omega_M = 0.154$).

Table 5.2: Best-fit results for the normalisation S_0 and exponent α in the power-law fitting of the median pseudo-entropy profiles at $z = 0$ traced by substructures.

	S_0	α
Hydro-1x	$(4.44^{+0.16}_{-0.16}) \times 10^4$	$0.89^{+0.06}_{-0.06}$
Hydro-10x	$(4.34^{+0.15}_{-0.15}) \times 10^4$	$0.86^{+0.03}_{-0.03}$

5.4.2.1 *The impact of baryons on substructures*

The results discussed so far and presented in Figure 5.6 refer to substructures identified in radiative hydrodynamical simulations. As such, they contain not only DM, but also gas and, most importantly, star particles. The latter, being originated from the dissipative collapse of gas undergoing radiative cooling, are expected to have colder dynamics and, therefore, lower pseudo-entropy than the DM component of the subhalos (Dolag et al., 2009). As such, star particles are also expected to be more gravitationally bound and then more resilient against tidal disruption than the DM component. To elaborate more on this point, we compare in Figure 5.7 the profiles traced by substructures in the DM-10x and Hydro-10x sets of simulations. Having the same resolution, this comparison allows us to determine the effect of dissipative gas dynamics on the (nearly) dissipationless dynamics traced by substructures. The top panel of Figure 5.7 shows that while the pseudo-entropy profiles of substructures in these two simulation sets agree in the outer cluster regions ($r > 0.5 r_{\text{vir}}$), they significantly differ at radii $r < 0.5 r_{\text{vir}}$. In particular, in DM-only simulations substructures are characterised by a plateau of pseudo-entropy in the cluster core, with no substructure found within $0.05 r_{\text{vir}}$. On the other hand, pseudo-entropy profiles traced by substructures in the Hydro-10x simulations are consistent with a power-law behaviour over the whole $0.05 \lesssim r_{\text{vir}} \lesssim 1$. We further investigate the origin of this difference and find that the velocity dispersion profiles of subhalos are consistent between the DM-only and Hydro runs. On the other hand, the number density profiles of substructures (central panel of Figure 5.7) highlight the relative deficit of subhalos in the cluster core for the DM-only case. This is expected, since the presence of baryons (and in particular their stellar content which dominates the central region of subhalos) has the effect of making galaxies more gravitationally bound, therefore making them more resistant against disruption caused by the strong central tidal fields (Dolag et al., 2009).

The deficiency of subhalos in the DM-only simulations compared to the hydrodynamical simulation could be also attributed to artificial disruption. Several authors (e.g. Muldrew et al., 2011) argue that the halo finder may be incorrectly identifying subhalos in DM-only; while others suggest that DM simulations suffer from significant overmerging due to numerical artefacts and could be avoided by following certain criteria (Bosch and

Ogiya, 2018). Nonetheless, others assessed results compatible with our findings claiming the differences in the radial distribution to be the result of tidal stripping (e.g. Weinberg et al., 2008).

To further reinforce this hypothesis, we computed the baryonic fraction within substructures (f_b) as a function of clustercentric distance in the Hydro-10x runs. The baryonic fraction of each substructure is defined as the ratio of the baryonic mass (which is the sum of the stellar mass M_* , the gas mass M_{gas} , and the black hole mass M_{BH}) over the total mass contained in the substructure (that includes also the DM component):

$$f_b = \frac{M_* + M_{\text{gas}} + M_{\text{BH}}}{M_{\text{TOT}}}. \quad (41)$$

The bottom panel of Figure 5.7 describes the radial distribution of the baryon fraction within substructures. In this panel, the grey line displays the median baryon fraction in substructures, while the orange shaded area is its 68 per cent dispersion within the set of simulated clusters sample. Tidal forces in the cluster centre strip more easily the outer region of substructures, which is dominated by the DM component. As a result, subhalos in the central cluster region – where tidal forces are stronger – are characterised by a higher baryon fraction, which even exceeds the cosmic baryon fraction assumed in our simulations for $r \lesssim 0.2 r_{\text{vir}}$. Previous studies (e.g. Armitage et al., 2018) had already shown this effect, highlighting that galaxies selected by their stellar mass, rather than by their total mass, have a significantly lower scatter in dynamical scaling relations. Substructures in the outskirts of clusters, that have yet not felt significant effects of tidal forces, have on average baryon fractions that decrease with radial distance. In these regions, the baryon fraction within substructures falls well below the cosmic value assumed in the simulation ($\Omega_b/\Omega_M = 0.154$). This is because the gas, that surrounds the subhalos, is ram-pressure stripped by the cluster hot atmosphere, thus leaving behind only the DM component and a minor fraction of cold star forming gas. This hypothesis is also justified by the median distribution of the baryon fraction in the cluster when accounting only for M_* in Equation 41, shown with the red curves in the bottom panel of Figure 5.7: when approaching r_{vir} , the baryon budget of substructures is entirely dominated by stars, while a significant contribution of gas is detected when approaching $\simeq 2 r_{\text{vir}}$. Therefore, substructures retain gas particles only when they are at radial distances of $r > r_{\text{vir}}$, just before starting orbiting closer to the centre of the cluster and being completely deprived (Lotz et al., 2019; Annunziatella et al., 2016). Regardless of our motivated discussion, we believe that our results are not definite, and they would certainly require some more investigation to undoubtedly resolve the controversy.

5.4.2.2 Mass segregation in substructures

Tidal stripping is only one of the two main mechanisms responsible for the bias between DM and galaxies. As pointed out in the previous sections, we expect this phenomenon to shape the number density profile

of substructures, by selectively removing lower-velocity substructures, thereby increasing their velocity dispersion and, consequently, causing a density profile shallower than that traced by DM particles. As a further test of the robustness of our results, we investigate whether the mass of the substructures introduced biases due to selection effects on the construction of the phase-space of clusters. These effects may originate from the impact of dynamical friction, which depends on the infalling mass of the orbiting substructures (Chandrasekhar, 1943); while mass-selection biases might also be associated with the early disruption of low-mass subhalos which are more easily stripped by the strong gravitational tidal fields in the cluster central regions. To test for these effects, we split the subhalo population of each cluster within the Hydro-10x sample at their median mass $M_{0.5}$. We then compute the spatial and velocity radial distribution, and the pseudo-entropy profiles for these two equally populated subsamples of subhalos. Figure 5.8 displays the median phase-space properties (and associated 68 per cent standard deviation) of the 10 clusters in the Hydro-10x divided into the high-mass group (in blue) and the low-mass group (in orange). No significant evidence of mass segregation is found in pseudo-entropy, velocity dispersion, and number density profiles, at least within the statistics allowed by our simulations. This result provides evidence that no significant mass segregation effects given by dynamical friction impact our results.

In conclusion, the general picture emerging from the analysis of our simulations on how substructures evolve within a galaxy cluster can be summarised as follows.

- Subhalos in the outer regions are deprived of most of their diffuse gas component presumably through the ram-pressure stripping, which takes place already at distances beyond the virial radius;
- During the infall within the cluster potential, substructures are stripped of their DM component, which is less resistant to tidal forces, while still preserving part of the baryonic component in the form of stars; this causes an increase of the baryonic fraction within substructures at small clustercentric radii;
- In the central regions, we find substructures with high baryon fractions (of about 25–45 per cent) for the most part due to the presence of star particles that are gravitationally more bound, thus more resilient to tidal disruption;
- These effects seem to be independent of the mass of the subhalos populating our sample of clusters because we report no evidence of mass segregation

As a general consequence, substructures in DM-only simulations, which do not include a stellar component, are more easily destroyed and do not survive for long times in the central regions of galaxy clusters.

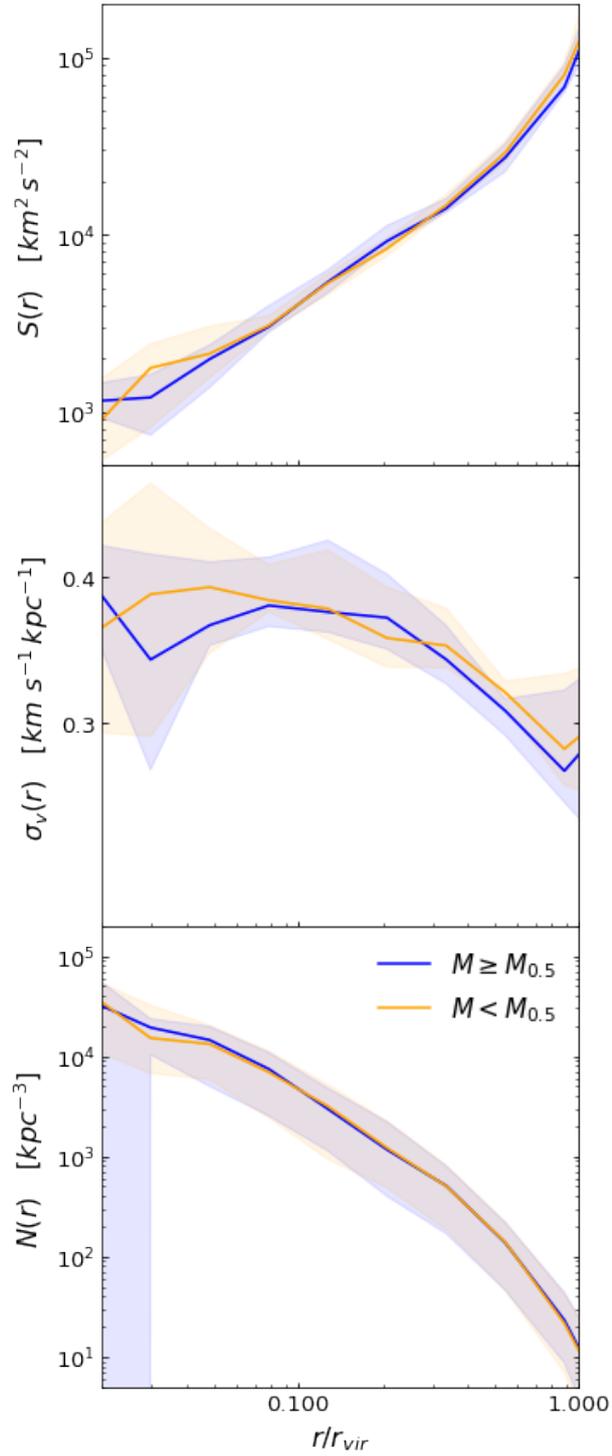


Figure 5.8: Effect of mass segregation on the distribution of substructures in the Hydro-10x set of simulations. In each panel, we show with blue (orange) curves results for subhalo populations with masses larger (smaller) than the median subhalo mass found in each cluster. Shaded areas encompass the 16th–84th percentile of the distribution of profiles. From top to bottom panels we show results for pseudo-entropy, velocity dispersion and number density profiles.

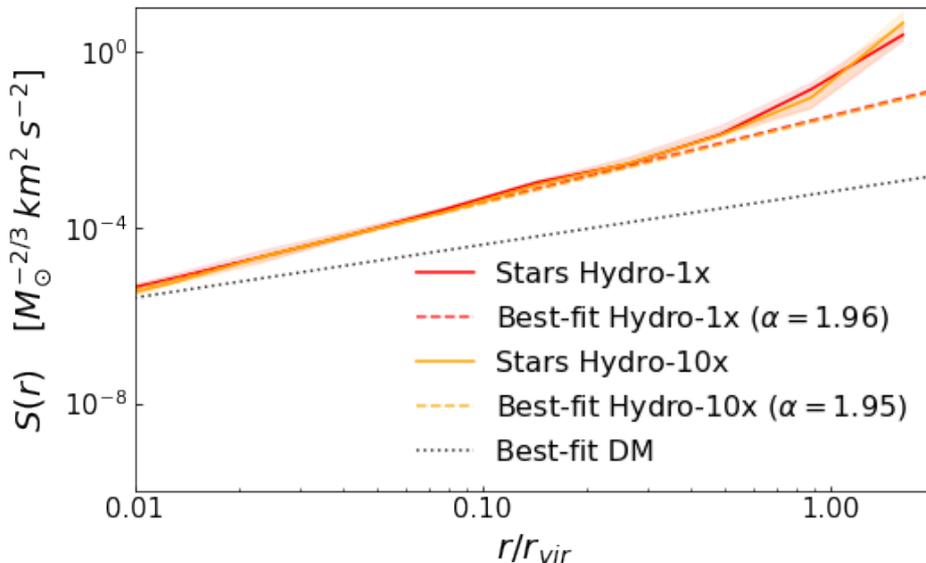


Figure 5.9: Pseudo-entropy profiles traced by the stars in the main halo in the Hydro-1x (solid red line) and the Hydro-10x (solid orange line) at $z = 0$. For both profiles we also report the respective intrinsic scatter of the simulated cluster samples (the shaded areas) and the best-fit profiles (dashed lines), using the same colour code. To help the comparison, we report the best-fit profile for the DM particles (dotted black line) in the Hydro-1x simulation.

5.4.3 Pseudo-entropy profiles traced by stars

Stars represent a collisionless component in hydrodynamical simulations, as the above discussed DM particle component. However, since they are generated from the dissipative collapse of gas particles, their phase-space structure is expected to be different from that of DM particles, whose collapse is completely non-dissipative. Therefore, we investigate the pseudo-entropy profile as traced by the stellar component as a tool to understand the different physical processes involved in its formation. Figure 5.9 illustrates the comparison between the median pseudo-entropy profiles traced by the stars belonging to the main halos of the clusters in the Hydro-1x (in red) and in the Hydro-10x (in orange) simulations. This means that here we are including only the stars belonging to the main halo while excluding all the star particles bound to substructures. For both profiles we also report the respective intrinsic scatter of the simulated cluster samples (the shaded areas) and the best-fit profiles (dashed lines). For reference, the dotted black line represents the best-fit curve traced by the DM particles in the Hydro-1x simulation. The stellar component recovers the power-law feature in the pseudo-entropy profiles only for $r \lesssim 0.55 r_{\text{vir}}$ in both the Hydro-1x and the Hydro-10x runs, which corresponds to the clustercentric distance within which most of the stellar mass of the main halo is contained (more than 90 per cent in all the clusters). In Table 5.3 we show the best-fit parameters for both simulations: we find that the two simulations

Table 5.3: Best-fit results for the normalisation S_0 and exponent α in the power-law fitting of the median pseudo-entropy profiles at $z = 0$ traced by stars.

	S_0	α
Hydro-1x	$(8.70^{+0.37}_{-0.36}) \times 10^{-3}$	$1.96^{+0.02}_{-0.02}$
Hydro-10x	$(8.39^{+0.21}_{-0.22}) \times 10^{-3}$	$1.95^{+0.02}_{-0.02}$

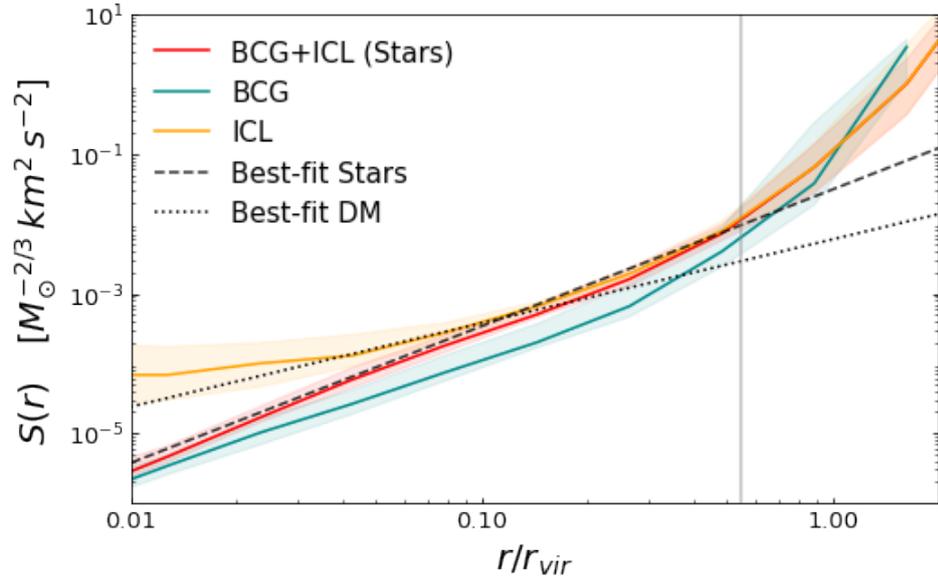


Figure 5.10: Pseudo-entropy profiles traced by the stars bound to the BCG (in green), ICL (in yellow) and the two combined (in red) in the Hydro-1x simulation at $z = 0$. The dashed black line is the best-fit profile traced by all the stars while the dotted black line is the best-fit profile for the DM particles which has been shifted in order to match the pseudo-entropy value of the ICL profile at $0.1 r_{\text{vir}}$. The vertical grey line helps the eye to visualise until which radius ($\sim 0.55 r_{\text{vir}}$) the power-law holds for the sample of all the stars at this resolution.

are consistent within 1σ for the different resolutions, thus guaranteeing the convergence of our results against the resolution.

5.4.3.1 ICL and BCG

In the hypothesis of tracing the real structure of the phase-space, one can further investigate whether the distinct dynamics and structure of the ICL and BCG composing the stars in the main halo of galaxy clusters emerge in tracing the pseudo-entropy profile. Therefore, we make use of the modified version of Subfind based on the definition of binding energy, as described in Springel et al. (2001), to split the stars in ICL and BCG and we recover their single pseudo-entropy profiles.

The decomposition assigns on average 65 per cent of the total stellar mass in the main halo to the ICL, while the rest is concentrated in the cent-

ral BCG. This definition of ICL is not fully comparable with the majority of definitions often applied in simulations and observations (Rudick et al., 2011). For instance, several authors have identified the ICL from observational data as the stellar component with luminosity below a limiting surface brightness (e.g. Zibetti et al., 2005), while other authors have modelled idealised galaxy profiles and subtracted them from the total stellar luminosity, taking the excess as ICL (e.g. Gonzalez et al., 2007). Recent attempts to separate the diffuse component from the BCG in simulations were oriented by excluding a given central aperture (e.g. Pillepich et al., 2018; DeMaio et al., 2020) and orbiting substructures. Whereas these approaches are closer to the ones followed in the observational analysis, they implicitly assume spherical symmetry that could result in a considerable simplification of the problem. Accordingly, this may lead to possible contamination from the two dynamical components which in turn can lead to biases in an analysis that aims at identifying differences in the phase-space structure of the two components. A detailed investigation of the different approaches used to separate the ICL and the BCG components is beyond the purpose of this paper, and we refer the reader to Dolag et al. (2010) for a more comprehensive description of the method adopted in this work.

Figure 5.10 shows the median pseudo-entropy profiles in the Hydro-1x run: namely the sample of all the stars (red line), the BCG (in green), the ICL (in yellow) and best-fit power-law relation for stars (reported with the dashed black line) and DM particles (the dotted black line normalised to match the ICL profile at $0.1 r_{\text{vir}}$). According to Figure 5.10, the pseudo-entropy profile traced by the stars bound to the BCG follows a power-law shape within $r \lesssim 0.3 r_{\text{vir}}$ with a slope shallower than in the BCG+ICL case. On the contrary, the ICL significantly deviates from a power-law shape at all radii. However, their composition clearly follows a power-law marked by the dashed curve, at least out to the grey vertical line at $0.55 r_{\text{vir}}$. Seemingly, despite having distinct dynamics and formation histories, ICL and BCG combine in the phase-space to form a power-law over a fairly large radial range, much like for DM particles, albeit with a different slope. In the innermost region (for $r < 0.04 r_{\text{vir}}$) the pseudo-entropy traced by stars resembles the behaviour followed by the BCG. The dissipative collapse that generates these stars in the core strongly affects their dynamics: on average they are characterised by "colder" dynamics, with relatively lower velocity dispersion, that reduces the entropy and impacts the pseudo-entropy profile. However, moving away from the cluster core, the contribution of the high-velocity dispersion of the "thermalised" population of the ICL becomes more dominant. For distances $r > 0.05 r_{\text{vir}}$, the ICL component prevails in the total profile.

Having been stripped from merging galaxies by tidal interactions and having undergone phase mixing during the hierarchical halo assembly, the diffuse stellar component is often found to be closely mapping the distribution of the DM particles (e.g. Montes and Trujillo, 2018). However, Alonso Asensio et al. (2020), Contini and Gu (2020) and Sampaio-Santos et al. (2021) have demonstrated that, even if this seems plausible, their radial

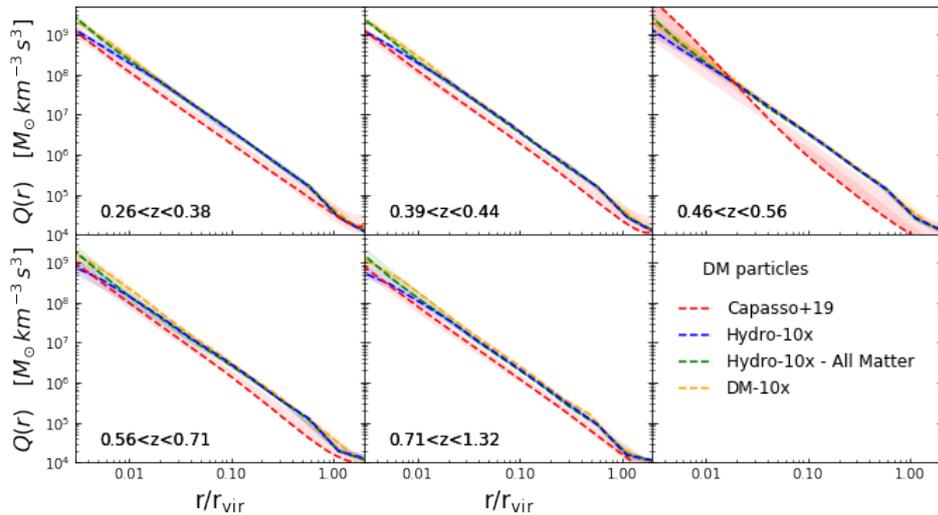


Figure 5.11: Phase-space density profiles as traced by DM within the Hydro-10x (in blue), DM-10x (in orange) and from the observed dataset from Capasso et al. (2019) (in red). We also report the phase-space density traced by both DM particles and baryons (stars and gas) within the Hydro-10x simulations (in green). The profiles are shown for the same 5 different redshift bins within which the analysis of observational data has been carried out. Profiles from DM particles in the Hydro-10x have been rescaled by $(1 - f_b)$ (f_b is the cosmic baryon fraction assumed in the simulations). For simulations, lines show the median profiles while shaded areas encompass the 16th to 84th percentiles.

profiles differ substantially. In this regard, we expect to observe these differences emerge also in their phase-space structure. By plotting the best-fit profile of the pseudo-entropy traced by the DM particles, we provide support to this claim: not only in our simulation we find a significant tension in the pseudo-entropy traced by DM particles and ICL, but these differences are significant in their density and velocity dispersion profiles too (which we do not show).

In conclusion, the remarkable picture arising is the universal power-law behaviour given by the composition of two dynamically distinct stellar components, which have different formation histories and are characterised by different, non-power-law, pseudo-entropy profiles.

5.4.4 Comparison with observational results

Comparing our findings with observational data is fundamental to understanding the capability of our simulations to correctly describe the dynamical processes leading to the formation and evolution of galaxy clusters and, ultimately, their predictive power. We compare our results from the DM-10x and Hydro-10x simulations with those in Capasso et al. (2019). Capasso et al. (2019) carried out an analysis of the phase-space density of clusters selected using the Sunyaev-Zel'dovich (SZ) effect in

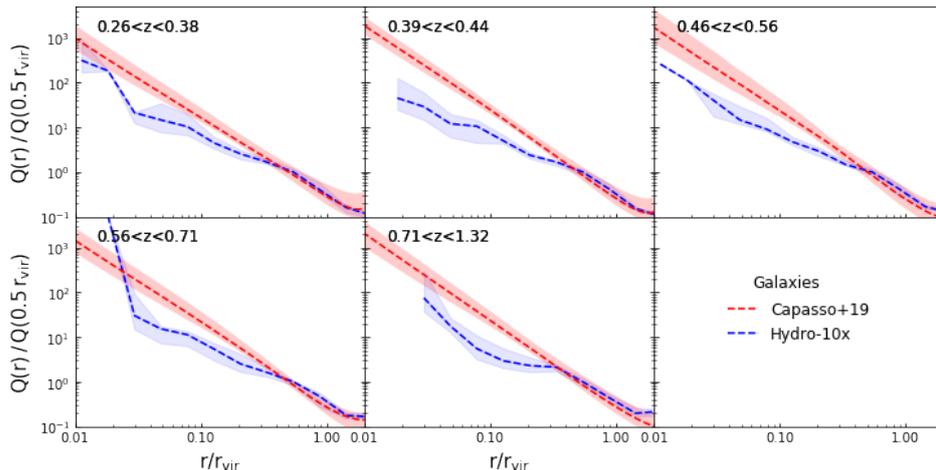


Figure 5.12: Phase-space density profiles as traced by substructures within the Hydro-10x simulations (in blue) and galaxies in the observational analysis in Capasso et al. (2019). For each redshift interval, profiles have been normalised to match at $0.5 r_{\text{vir}}$. Lines and shaded areas have the same meaning as in Figure 5.11.

the 2500 deg^2 South Pole Telescope (SPT)-SZ survey in the redshift range $0.2 < z < 1.3$. The reconstruction of the phase-space of these objects is performed with MAMPOSSt (Mamon et al., 2013) which, adopting parametric expressions for the mass and velocity anisotropy profiles, solves the Jeans equation in spherical symmetry and recovers the 3D velocity dispersion and tracers distribution. The central 50 kpc region is excluded from their analysis because it is identified as the cluster region affected by the presence of the BCGs.

5.4.4.1 Q from phase-space density

From the recovered mass density and velocity dispersion profiles, Capasso et al. (2019) derived the profile of phase-space density, $Q(r)$, by combining results from clusters within different redshift intervals. The results of this observational analysis are shown with the red dashed lines in Figure 5.11. The 5 different redshifts at which simulation results are shown have been chosen by selecting a snapshot at a redshift within the interval reported in each panel.

In the same figure, we also report with the orange and blue curves the median profiles traced by the DM particles of the DM-10x and Hydro-10x simulated clusters, respectively. For the latter, the profiles have been rescaled by a factor $(1 - f_b)^{-1}$ to take into account the baryon fraction. Finally, the green curves report the median profiles of the total mass density in the Hydro-10x simulations, i.e. adding to the DM also the contribution from the baryonic components in gas and stars.

Quite remarkably, simulations and observations produce profiles with a similar slope for all the redshift ranges. The only exception is represented by the redshift interval $0.46 < z < 0.56$, for which the observed phase-

space density profile is significantly different not only from the simulated ones but also from the observed ones at the other redshifts. Most likely this is due to the reconstructed velocity dispersion profile in this redshift bin, which is quite irregular in comparison with that recovered in the other redshift intervals. Despite such good agreement in slope, we note a slight systematic offset in normalisation between observed and simulated profiles that amounts to about 20 – 30 per cent. While the origin of this difference is not clear, it is worth reminding that the $Q(r)$ profiles from simulations have not been obtained by reproducing as close as possible the observational procedure based on the application of the Jeans equations and the deprojection of observed profiles of the number density of tracers and line of sight velocity dispersion. Nevertheless, we regard it as quite relevant that a power-law shape of $Q(r)$ is consistently produced by observational data and simulations when using matter density in the definition of phase-space density. Finally, we note that profiles traced by DM particles tend to slightly flatten in the innermost regions, $r \lesssim 3 \times 10^{-2} r_{\text{vir}}$, even if in this region the observational results are extrapolated. On the other hand, including in the analysis of the Hydro-10x simulations also the contribution of the baryonic (mostly stellar) components preserves a power-law profile extending to the innermost regions resolved in the simulations.

5.4.4.2 Q from galaxy number density

As a further term of comparison, we also use the results by Capasso et al. (2019) on the profiles of $Q(r)$ obtained by using the number density of tracers (galaxies), instead of the mass density profiles. Their results are shown with the red curves in Figure 5.12, while the blue curves show the $Q(r)$ profiles obtained from simulations using substructures as tracers of halo phase-space. Owing to the difficulty of properly normalising the number density profiles of the simulated substructures and observed galaxies, we arbitrarily fixed the normalisation in such a way that simulated and observed profiles match at $0.5 r_{\text{vir}}$. Quite remarkably, pseudo-entropy profiles from simulations are shallower than in observations if substructures/galaxies are used to trace the phase-space structure of clusters at $r \lesssim 0.4 r_{\text{vir}}$, while the two profiles recover the same slope at larger radii (Meneghetti et al., 2020, accepted). This result is in line with previous findings from simulations, indicating that the observed phase space traced by cluster galaxies is not accurately described by substructures identified in simulated clusters (Weinmann et al., 2012; Nierenberg et al., 2016; Hirschmann et al., 2016).

Thus, while the slopes of phase-space density profiles traced by the total mass density are quite consistent in observational data and simulations, the latter tend to produce lower values of phase-space density traced by substructures than in observational data. Simulations at higher resolution are needed to understand whether this is due to resolution effects, which could make substructures in simulations exceedingly fragile against the action of the tidal field in central cluster regions.

5.4.4.3 Mass reconstruction from Jeans equation

We discuss more in detail the limits built into the method used to derive the observed pseudo-entropy profiles. We see that these may introduce potential biases and underestimate the impact of mergers. The profiles of phase-space density from observational studies have been obtained by using the Jeans equation for a spherical system to recover the mass density profiles of galaxy clusters (Wolf et al., 2010; Mamon et al., 2013). However, a possible lack of dynamical equilibrium or departure from spherical symmetry could introduce biases in the recovery of such mass profiles. In addition, uncertainties in the correct modelling of the orbit anisotropy profile are also expected to affect a correct mass density reconstruction (Merritt, 1987). Therefore, one may wonder whether our comparison between observed and simulated profiles of phase-space density is affected by the assumptions underlying the application of the Jeans equation. To address this issue, we decided to reconstruct pseudo-entropy profiles in simulated clusters using the Jeans equation and compare them with the intrinsic profiles.

For a spherically symmetric system in equilibrium, the Jeans equation in spherical coordinates can be cast as

$$\frac{d(\nu\sigma_{v,r}^2)}{dr} + \frac{\nu}{r} [2\sigma_{v,r}^2 - (\sigma_{v,\theta}^2 + \sigma_{v,\phi}^2)] = -\nu \frac{d\phi}{dr}, \quad (42)$$

where ν is the number density profile of the tracer galaxy population, ϕ is the gravitational potential and $\sigma_{v,i}$ are the components of the velocity dispersion along the three spherical coordinates r, θ, ϕ . After integrating this equation for the radial component of the velocity dispersion profile, one obtains

$$\sigma_{v,r}^2(r) = \frac{1}{\nu(r)} \int_r^\infty \exp \left[2 \int_r^s \beta(t) \frac{dt}{t} \right] \nu(s) \frac{GM(s)}{s^2} ds \quad (43)$$

where G is the gravitational constant, $M(r)$ is the mass enclosed within r and $\beta \equiv 1 - \frac{\sigma_{v,\theta}^2 + \sigma_{v,\phi}^2}{2\sigma_{v,r}^2}$ is the velocity anisotropy profile. From this equation, the main idea is to provide a mass modelling technique which performs a Maximum Likelihood fit of the tracers' distribution $\nu(r)$, assuming parametric shapes for the gravitational potential ϕ (or equivalently $M(r)$, the mass profile), and the velocity anisotropy profile β . In observational analyses, this computation also involves a deprojection method to pass from the observed velocity dispersion and tracer number density profiles to their 3D counterparts. In the following, we do not address the issue of deprojection that we leave to future analysis, while we directly start from 3D information provided by simulations, to focus on the assumptions entered in the Jeans equation. Since the tracer population $\nu(r)$ does not necessarily follow the mass distribution $M(r)$, we must treat the two separately in the best-fit evaluation. Owing to the accurate fit provided to the density profiles produced by simulations, the NFW profile (Navarro, 1996; Navarro et al., 1997) defined by

$$\rho(r) = \frac{\rho_0}{x(1+x)^2} \quad (44)$$

is the functional form assumed for both the mass density and tracers number density profiles. In the above equation, $x = r/r_s$, where the scale radius r_s is by definition the radius at which the logarithmic slope is -2 . In this way, the profile is determined by r_s and by the normalisation ρ_0 , which are two parameters to be fitted independently for the mass profile $M(r)$ and the tracers number density profile $\nu(r)$.

As for the velocity anisotropy profile $\beta(r)$, we assume its expression to be given by the Tiret model (Tiret et al., 2007), which proved to provide a good description for cosmological simulations of cluster-mass halos (Mamon et al., 2013, 2010):

$$\beta(r) = \beta_0 + \frac{r}{r + r_s} \theta_B, \quad (45)$$

with r_s the scale radius of the NFW profile of the tracer distribution, β_0 a normalisation and θ_B the asymptotic value of the orbit anisotropy.

For each simulated cluster, we perform the best-fit of the tracer distribution $\nu(r)$ and its anisotropy velocity $\beta(r)$ profiles from simulation data. In order to recover the mass profiles in a way similar to what is done with observational data, we apply a maximum likelihood method to recover the normalisation ρ_0^M and scale radius r_s^M of the mass profile, so that the radial velocity dispersion profile inferred from Equation 43 matches the true one measured in simulations. For each cluster, the total velocity dispersion profiles (normalised by $\sqrt{3}$) can be computed for a given model of velocity anisotropy according to $\sigma_v(r) = \sigma_{v,r}(r) \sqrt{3 - 2\beta(r)}$. Having reconstructed the mass density profile from the Jeans equation and the profiles of total velocity dispersion, we finally derive the reconstructed profiles of pseudo-entropy, to be compared with the true profiles.

Furthermore, in order to verify the accuracy of this procedure based on the Jeans equation to recover pseudo-entropy profiles as a function of the dynamical state of a cluster, we divided our set of simulated halos into relaxed and disturbed. To carry out this classification, we followed the prescription in Biffi et al. (2016) and described herein Section 5.2.

In Figure 5.13 we compare pseudo-entropy profiles recovered from the Jeans equation procedure with the true intrinsic profiles, for both relaxed and disturbed systems from the Hydro-1x simulation. In the right panels, we show their density maps in logarithmic scale traced by the DM particles in two clusters within the $0.5 r_{\text{vir}}$ spheres (in white). In each of the two panels on the left, we compare recovered and intrinsic profiles of $S(r)$ for both the selected cluster (solid lines), whose density map is reported on the right and the entire cluster sample (dashed lines). Upper panels display the profiles for relaxed clusters, while lower panels are for disturbed systems. The solid lines are specific to the single cluster, being in cyan the true profile and in orange the one produced by the Jeans equation procedure while the dotted lines refer to the median values of all 29 clusters in the simulation set. In the upper panel, we see that true and Jeans equation profiles overlap at almost all radii, showing that the Jeans equation procedure has correctly reproduced the true profile, which is also in line with the median result. In the lower panel, we show the same

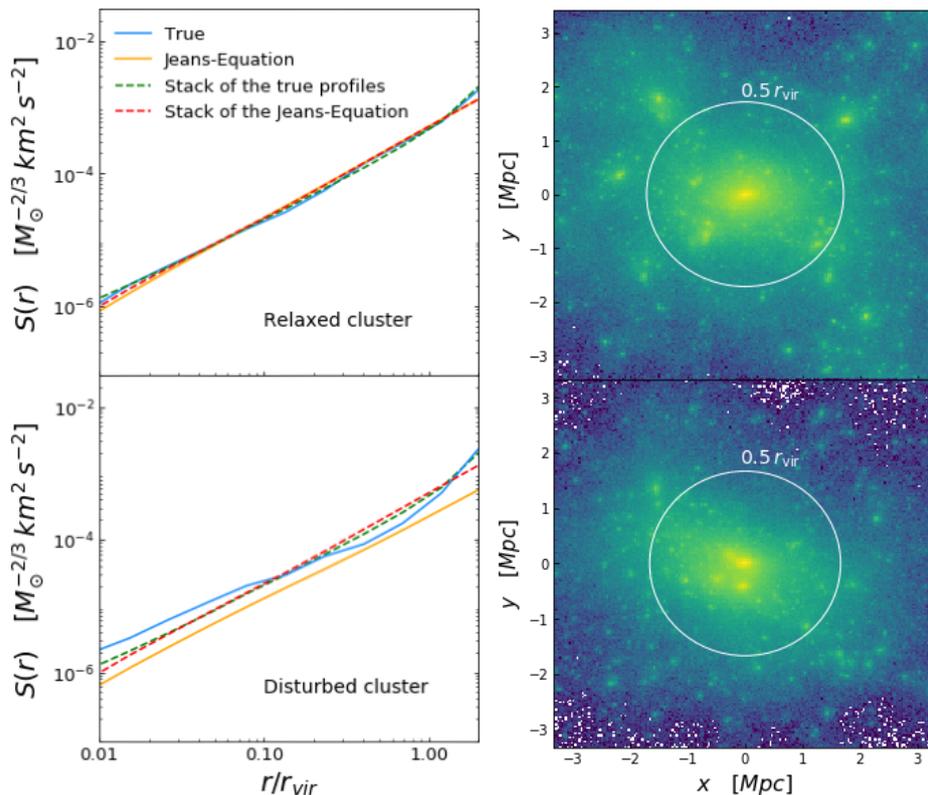


Figure 5.13: Left panel: pseudo-entropy profiles (DM particles) for a relaxed and a disturbed cluster in the Hydro-1x simulation at $z = 0$. The dashed lines are obtained by stacking the true profiles (in cyan) or those obtained by the Jeans equation procedure (in orange), while the solid lines are for the profiles of the specific cluster. Right panel: density maps in logarithmic scale traced by the DM particles in the same objects. The brightest colours indicate the areas at higher density. The white circles identify the $0.5 r_{vir}$ spherical regions centred on the cluster centre as computed by Subfind.

profiles but when obtained for a cluster labelled as "disturbed", since a major merger is occurring as shown by its density map. In this cluster, the true profile consistently deviates from the median and the Jeans equation profiles. The latter, in turn, suffers a shift in normalisation but not in the slope with respect to the median profiles. To quantitatively describe the differences between recovered and true pseudo-entropy profiles, we follow the same procedure described in Section 5.4.1 to fit a power-law profile, which depends on two parameters, and concentrate our interest on the accuracy in recovering the slope α . In the first two lines of Table 5.4, we report the true and the Jeans equation recovered slopes for the two relaxed and disturbed clusters shown in Figure 5.13. Consistently with the results shown in this figure, the slope is accurately recovered for the relaxed cluster, while a significant difference is found for the disturbed object. This is in line with the expectation that the procedure based on the Jeans equation, which implicitly assumes dynamical equilibrium, can introduce biases when applied to unrelaxed systems. The results from

Table 5.4: The best-fit values obtained for the logarithmic slope in the true profile and the one recovered through the Jeans equation procedure for the relaxed and disturbed cluster shown in Figure 5.13 with the 68 per cent uncertainty. In the last row, the results were obtained when fitting the median profile from the stack sample of 29 clusters in the Hydro-1x simulation at $z = 0$. In the latter, the uncertainties reported are the standard deviation derived from the distribution of best-fit values for the sample.

α	True	Jeans equation
Relaxed	$1.203^{+0.002}_{-0.001}$	$1.204^{+0.002}_{-0.001}$
Disturbed	$1.050^{+0.002}_{-0.001}$	$1.211^{+0.003}_{-0.002}$
Stack	1.222 ± 0.002	1.238 ± 0.003

the combination of 29 simulated clusters are reported in the last row of Table 5.4. On average, the Jeans equation sample seems to slightly overestimate the true value of the slope, although the difference between true and recovered slopes are relatively small once averaging over a sample of clusters with a representative mix of dynamical state.

In conclusion, the assumption of dynamical equilibrium imposed by the Jeans equation is shown to potentially introduce a significant bias in the reconstruction of the phase-space structure of unrelaxed clusters. It is worth reminding that the analysis presented here assumes the knowledge of the 3D distribution of tracers and the velocity dispersion profiles. A full analysis aimed at including observational effects on the measurements of phase-space density of galaxy clusters would require a proper account for projection effects, which we defer to future analysis.

5.5 CONCLUSIONS

In this chapter, we presented an extensive analysis of the phase-space structure of simulated galaxy clusters. In particular, we studied the pseudo-entropy profiles $S(r)$, or equivalently the phase-space density $Q(r)$, traced by different collisionless components: DM particles, substructures and star particles. The analysis is based on the DIANOGA set of cosmological simulations of galaxy clusters (Bassini et al., 2020), that have been carried out with the GADGET-3 code at different resolutions and including different physics: pure N-body and hydrodynamical simulations with radiative cooling, star formation and stellar feedback models implemented following Springel and Hernquist (2003), metal enrichment and chemical evolution following the formulation described in Tornatore et al. (2007b); AGN feedback as described by Ragone-Figueroa et al. (2013). Our analysis aimed at investigating the mechanisms involved in building the phase-space structure of galaxy clusters, and in comparing predictions of simulations to observational data. We note that our analysis provides for the first time

an analysis of the pseudo-entropy profiles traced by substructures, which should correspond to galaxies, and stars within the main halo of galaxy clusters.

The main results can be summarised as follows:

- Pseudo-entropy profiles from simulations, as traced by all the three collisionless components (substructures, DM, and star particles), are always close to power-laws for radii smaller than the virial radius. Furthermore, these profiles scale self-similarly with mass and redshift, at least from $z = 2$ which is the largest redshift we analysed. Substructures present a profile shallower than that outlined by DM particles as an effect of the change in the density profile. Stars, on the other hand, show a steeper pseudo-entropy profile which is a consequence of the dissipative collapse of gas by radiative cooling leading to star formation.
- Stars in the main halo (i.e. not belonging to substructures) have been separated into two dynamically distinct components: those bound to the BCGs and those belonging to the ICL. These two components are characterised by different pseudo-entropy profiles, e.g. BCG stars, that dominate in central regions, have a steeper slope for $S(r)$, thus turning into a lower level of pseudo-entropy at small radii. Quite remarkably, while the only BCG $S(r)$ profile is a power-law out to $0.3 r_{\text{vir}}$, they combine to provide an accurate power-law for the $S(r)$ profiles of the total stellar component of the main halo, extending at least out to $0.55 r_{\text{vir}}$. We verified the BCG to be responsible for the shape of the total profile in the core regions ($r \leq 0.04 r_{\text{vir}}$): stars here originate from the dissipative collapse of gas, lowering their velocity dispersion and reducing the entropy. For $r > 0.55 r_{\text{vir}}$ the total profile follows the profile traced by the ICL, formed by the stripping of the stellar matter from satellite galaxies.
- The pseudo-entropy of DM particles integrated within the virial radius provides an accurate proxy for the total mass of galaxy clusters, with an intrinsic scatter at fixed mass of $\sigma_{M_{\text{vir}}|S_{\text{vir}}} = 0.067 \pm 0.003$. This is even smaller, by about a factor of two, than that associated to the velocity dispersion $\sigma_{v,\text{vir}}$ ($\sigma_{M_{\text{vir}}|\sigma_{v,\text{vir}}} = 0.132 \pm 0.003$), which is considered the tightest proxy of cluster mass. The predicted scaling is $S_{\text{vir}} = S_{\text{DM}}[h(z)M_{\text{vir}}/(10^{15}M_{\odot})]^{\gamma'}$ where $h(z)$ is the Hubble parameter in units of 100 km s^{-1} and $\gamma' = 1.74 \pm 0.05$.
- Several factors contribute to affecting the phase-space structure of clusters, resulting in relative deviations in the power-law feature. More in detail, we found the presence of baryons to cause modest differences in the pseudo-entropy profiles traced by DM particles between the hydrodynamical and DM-only runs, but most significantly it intervened in the phase-space distribution of substructures within clusters (leading to a flattening of the pseudo-entropy profile in the central regions of clusters evolved in DM-only simulations).

Furthermore, as discussed in [Section 5.4.4.3](#), the dynamical state of the single cluster can also impact the phase-space structure of these objects. This was shown to introduce a non-trivial bias in the analysis of observational data when estimates of the mass profiles were made through the resolution of the Jeans equation. The resulting pseudo-entropy profile in disturbed objects appeared to deviate in both normalisation and slope from the true profile.

- The comparisons with the observed phase-space density profiles described in [Capasso et al. \(2019\)](#) offered the opportunity to assess the capability of simulations to predict the phase-space structure of galaxy clusters, and which evolutionary processes are responsible for it. The phase-space structure traced by DM particles is generally in good agreement with observed clusters out to the highest redshift, $z \simeq 1.3$, at which these studies have been carried out so far. On the other hand, the profiles traced by the real galaxies (in [Figure 5.12](#)) are found to be significantly steeper than those constructed by the substructures in simulated clusters. This result is specifically evident in the central regions at all redshifts and it establishes the existing limits in cosmological simulations in reproducing the phase-space traced by satellite galaxies in clusters, a well known problem in the literature.

One of the general conclusions of our analysis is that pseudo-entropy profiles provide an important characterisation of the phase-space structure of cluster-size halos: despite being defined from the combination of density and velocity dispersion profiles, each having a non-power-law shape, $S(r)$ as traced by DM particles has a shape which is accurately described by a power-law over a fairly wide range of scales and redshift, with a normalisation that scales self-similarly. Quite interestingly, also the stellar halo component develops a power-law shape of the pseudo-entropy profiles, which extends over a fairly large radial range, although only the BCG $S(r)$ tends to form a power-law, albeit over a narrower radial range. Different shapes for $S(r)$ of DM and stars are understood in terms of the different nature of the gravitational collapse determining their respective evolution, non-dissipative for the former and dissipative for the latter. These results lend support to the idea that pseudo-entropy is a fundamental quantity, possibly more fundamental than density profiles, to characterise the non-linear evolution of a collisionless self-gravitating fluid, leading to the formation of galaxy clusters. This is also reinforced by the tiny scatter that pseudo-entropy has in the scaling relation against the total halo mass, thus possibly promoting it also to the role of precise mass proxy for cosmological applications of galaxy clusters.

Quite interestingly, our comparison with observational data shows that a good agreement is attained only when using total density to trace pseudo-entropy. In this respect, substructures in simulations appear to trace a pseudo-entropy level in central regions, which is higher than that traced by galaxies in observational data. This is led by tidal disruption, which causes

substructures with relatively lower orbital velocity to become fragile in central regions. As a result, only a relatively small number of substructures, with relatively high orbital velocities, survive, thus causing an excess of the pseudo-entropy with respect to what is observed. A detailed analysis, also based on higher resolution simulations, will be required to assess whether this disagreement is merely due to the numerical limitations of our simulations, or it is rather indicating a more fundamental lack of understanding of the processes determining the evolution of substructures inside massive cosmological halos.

VELOCITY DISPERSION OF THE BRIGHTEST CLUSTER GALAXIES

In this chapter, we analyse the dynamics traced by stars belonging to the BCGs and the surrounding diffuse component, forming the ICL, and compare it to the dynamics traced by DM and galaxies identified in the simulations. We address the question of whether simulations can correctly describe the dynamics of BCGs and their surrounding stellar envelope, as determined by the past star formation and assembly histories of the most massive galaxies of the Universe. Thus, we aim to answer the following scientific questions.

How do the different stellar components react to the gravitational potential of the cluster? Can we detect the contribution of the ICL in the outer regions of the stellar profile traced by the BCG? Can we infer significant differences between the two components? Can we correlate the tracers of the phase-space structure to the dynamical and astrophysical processes determining the assembly of the BCG and the cluster as a whole? Does this relation provide constraints on structure formation models? Do state-of-the-art simulations reproduce the observed trends?

The content of this chapter largely reflects what is presented in two papers published in Monthly Notices of the Royal Astronomical Society.

- **Marini I**, Borgani S, Saro A, Granato GL, Ragone-Figueroa C, Sartoris B, Dolag K, Murante G, Ragagnin A, Tornatore L, Wang Y "Velocity dispersion of brightest cluster galaxies in cosmological simulations." Monthly Notices of the Royal Astronomical Society, 2021, 507.4: 5780-5795.
- **Marini I**, Borgani S, Saro A, Murante G, Granato GL, Ragone-Figueroa C, Taffoni G "Machine learning to identify ICL and BCG in simulated galaxy clusters." Monthly Notices of the Royal Astronomical Society, 2022, 514.2: 3082-3096.

6.1 SCIENTIFIC CONTEXT

Recent studies have concentrated on the observational study of stellar velocity dispersion in BCGs, as a way to obtain information on the dynamical history of their assembly (e.g. Remus et al., 2017; Sohn et al., 2020; Bose and Loeb, 2021; Sohn et al., 2022). Sohn et al. (2020) investigated the relationship between the stellar velocity dispersion of the BCG and the cluster velocity dispersion for a large spectroscopic sample. The observed correlation is tight and velocity dispersion measurements are less affected

by systematics which instead are inevitably introduced in estimates of the stellar mass of the BCG. More recently, Sohn et al. (2022) confirmed that the ratio between the BCG velocity dispersion σ_{BCG}^* and the overall velocity dispersion of clusters σ_{200} is a steady decreasing function of the cluster velocity dispersion, in contrast with predictions from cosmological hydrodynamical simulations by Dolag et al. (2010) and Remus et al. (2017). In any case, constraining this relation will open up more possibilities to investigate BCG properties and correlation with the host cluster.

Although promising, analyses of the BCG velocity dispersions pose several issues that need to be taken into account. Part of the challenge lies in the different approaches followed to estimate the BCG properties in observations and simulations. These include, but are not limited to, the choice of aperture and the different selection criteria of the galaxy in cluster regions (e.g. most luminous or most massive galaxy). To complicate the comparison, the evolution of the BCG is tightly connected to the build-up of the ICL, whose properties are rather unique. The ICL is predicted to form at relatively late times (primarily from tidal stripping and galaxy mergers) and to be gravitationally bound to the cluster as a whole rather than to a single galaxy (e.g. Murante et al., 2007; Contini et al., 2014, 2018; Montes and Trujillo, 2018, 2019). The spatial extent of this component, which overlaps the BCG, can reach up to hundreds of kpc, similarly to the DM in galaxy clusters (Pillepich et al., 2018; Montes and Trujillo, 2019). However, determining the amount and extent of the ICL component from observations is quite challenging, as it requires both wide and deep observations, capable of capturing spatially extended low-surface brightness regions (in MACS J1206.2-0847, Presotto et al., 2014). The comparison of the observational analyses is further complicated by the use of different apertures to define the BCG stellar mass, which can lead to the inclusion of a more or less important fraction of the ICL component. A consistent choice among different studies on including ICL stars may alleviate the tension observed in the estimated evolution of the BCG stellar mass with redshift (Zhang et al., 2016; Ragone-Figueroa et al., 2018).

Moreover, BCGs also offer a valuable route for studying the innercore of their host halos, where the DM distribution is a key prediction in cosmological models (see the seminal papers by Navarro et al. 1996, 1997; more recently, He et al. 2020). However, standard methods to determine accurate mass estimates of cluster halos (X-ray combined with Sunyaev-Zeldovich effect, and weak lensing measurements) often fail to probe the innermost regions of clusters. X-ray measurements of the intracluster medium are sensitive to the presence of cool fronts and temperature fluctuations near the cluster centre (Arabadjis et al., 2004), additionally, they are affected by systematic uncertainties due to deviations from hydrostatic equilibrium, albeit especially in the outer regions (Rasia et al., 2006). On the other hand, in weak lensing surveys, clusters are imaged deeply to obtain a resolved lensing signal but usually only cover limited areas around each cluster. This issue is present especially at low redshifts, where this regime requires imaging of greatly high power of resolution (e.g. Joffre et al., 2000).

On the other hand, the stellar kinematics of BCGs provides us with information on the total mass at small radii, given that the stellar velocity dispersion profile directly relates to the total gravitational potential well (and thus, the mass distribution) in which stars are moving. Indeed, Sartoris et al. (2020) performed a dynamical analysis in the core region of Abell S1063, via the MUSE integral field spectroscopy. The authors were able to reconstruct the inner logarithmic slope of the DM density profile, inferring the latter from solving the spherical Jeans equation with the stellar velocity dispersion and the projected phase-space distribution of the other cluster galaxies, to find a value consistent with the cold DM model predictions. These results were in contrast with previous observational measurements, based on a combination of strong lensing and BCG kinematics (i.e. Newman et al., 2013), thus stimulating a discussion on the possible reasons behind this discrepancy (He et al., 2020). In general, we expect stellar velocity dispersion profiles of BCGs to echo gravitational mechanisms operating at the cluster centres. However, as the size of the spectroscopic samples has been increasing, we have come across a large variety in the slopes of the velocity dispersion in the centres of BCGs (Loubser et al., 2018, 2020). Whether this diversity of slopes is due to the differences in the mass profiles of the galaxies (Barnes et al., 2007) or their different evolutionary paths, this is an open question whose answer requires a better understanding of its dynamical origin.

6.2 SIMULATIONS

The set of simulated galaxy clusters is the hydrodynamical run from the DIANOGA set at the two different levels of resolution available to test for numerical convergence. The initial gas and DM particle masses are reported in Table 3.1.

Our final sample is selected in Hydro-10x from among each Lagrangian region, as the main halo of the FoF groups with total DM mass (defined as the sum of all DM particles bound to the group) larger than $5 \times 10^9 h^{-1} M_{\odot}$, this comprises 57 halos. At the base resolution, we only select the main halo from the largest FoF groups, for a total of 29 halos.

6.3 OBSERVABLES IN SIMULATIONS

Here below we briefly describe the different ways we used to characterise BCG and cluster properties in our simulations. Given that our analysis mainly focuses on mass and velocity dispersion, we will provide their definitions in this Section.

6.3.1 Dynamics of the BCGs

Distinguishing the BCG from the stellar diffuse envelope within clusters is notoriously non-trivial. Although BCGs are usually the brightest (and

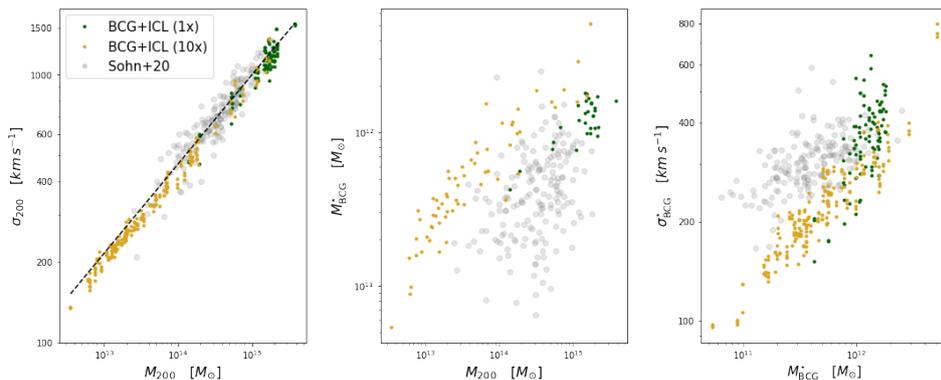


Figure 6.1: A comparison of some of the general properties of our simulated set of clusters with the observational results by Sohn et al. (2020). Left panel: scaling relation between cluster velocity dispersion, σ_{200} and total cluster mass, M_{200} ; central panel: scaling relation between BCG stellar mass, M_{BCG}^* and M_{200} ; right panel: scaling relation between BCG velocity dispersion, σ_{BCG}^* and M_{BCG}^* . Legend is common to all panels: dark green and golden points indicate respectively the Hydro-1x and Hydro-10x samples using all the stars in the main halo enclosed in a bi-dimensional aperture of 50 kpc. The grey circles stand for the observational results by Sohn et al. (2020). Velocity dispersions are measured for each simulated cluster within three orthogonal projections.

most massive) galaxies in their cluster, positioned in a privileged position at the bottom of the gravitational potential, many factors contribute to complicating the analysis, especially in observations. In some cases, the central galaxy may not coincide with the brightest galaxy in the cluster, or its position may be shifted with respect to the geometrical centre of the cluster halo (e.g. De Propris et al., 2021). This adds up to the difficulties of defining physical boundaries to the galaxy, given that its stellar population is wrapped in a diffuse stellar component, the so-called ICL, which is gravitationally bound to the cluster potential. A complete discussion on the methods to discern the two is beyond the scope of this paper, we refer the interested reader to Rudick et al. (2011) for a comprehensive discussion.

In the analysis of observed clusters, all the properties of the BCG are measured within a circular aperture centred on the brightest galaxy member in a given band (e.g Sohn et al., 2020) or on the galaxy closest to the X-ray peak (Loubser et al., 2018). The choice of the aperture radius is somewhat arbitrary in the literature, still masses measured within a fixed physical radius allow us to neatly compare predictions from simulations with observations. We choose to identify all BCG observables as the properties yielded by the star particles in a cylinder long R_{200} centred on the BCG and within a 50 kpc physical radius (we note that by reducing the aperture to 30 kpc we do not expect major differences in the BCG properties out to $z \simeq 1.5$; Ragone-Figueroa et al., 2018). We select only stars in the main halo, excluding those which Subfind identifies as bound

to other substructures, and we measure the physical properties within this circular area. We point out that with this method we include not only the dynamical information on the BCG but also the contribution from the ICL. In the following, we will refer to this definition as "BCG+ICL".

We define the line of sight stellar velocity dispersion σ_{BCG}^* as the r.m.s. (i.e., root mean square) of the velocity distribution of the star particles (identified as BCG+ICL or BCG-only) within the cylindrical projection. We tested several estimators, but using the r.m.s. estimator, we obtained stellar velocity dispersions that closely agree with the observational data. We note that using the biweight estimator (which, on the other hand, is used to compute the total cluster velocity dispersion) does not guarantee the same level of agreement. Additionally, velocity dispersions are measured within the three orthogonal projections. All the observational datasets employed in this analysis comprise velocity dispersion measurements recovered from the galaxy spectra with the Gauss-Hermite series (Cappellari and Emsellem, 2004). The BCG stellar mass M_{BCG}^* is computed as the median of the sum of the selected star particles within the same bi-dimensional aperture.

6.3.2 Dynamics of the clusters

Observed cluster properties are usually retrieved from the dynamics of galaxy members. In our simulated clusters, *bona fide* galaxies correspond to gravitationally bound substructures, which we identify through the Subfind algorithm (Springel et al., 2001). However, given that substructures in simulations sometimes fail in reproducing the phase-space structure of the real galaxy population in clusters (e.g. Hirschmann et al., 2016; Marini et al., 2021a), it is advisable to also employ DM particles to characterise the cluster dynamical properties. These particles have been selected from the main halo of the central group, excluding the contribution from those bound to substructures. In this regard, the dynamics followed by DM particles should be a robust tracer of the cluster potential, unlike substructures that are dependent on the halo finder, and possibly also suffering from the effect of dynamical friction. In Section 6.4 we further comment on the impact of choosing DM particles with respect to galaxies.

Following the observational approach, we select DM particles within a cylinder long $2R_{200}$ whose base of radius R_{200} is centred on the cluster centre, corresponding to the position of the particle with the minimum value of the gravitational potential. We determine the DM particles kinematics (i.e. the cluster velocity dispersion σ_{200}) with the biweight estimator. Given that galaxies are most likely affected by fore-/back-ground interlopers in the observational analysis, observers often choose to use the biweight because it underweights the tails of the velocity distribution which are expected to be populated by outliers (Beers et al., 1990). Analogously, we define the cluster mass M_{200} as the total particle mass within the same cylinder.

6.4 CONNECTIONS BETWEEN BCG AND CLUSTER

In the following, we will present several properties of our set of simulated clusters, and compare them with recent observational results by Sohn et al. (2020). These authors analysed the HeCS-omnibus cluster sample which includes 227 objects ($M_{200} = 2.5 - 18.4 \times 10^{14} M_{\odot}$) observed from a combination of both photometric and spectroscopic surveys, over the redshift range $0.02 \leq z \leq 0.3$. Cluster masses M_{200} are measured via caustic method (Diaferio and Geller, 1997; Serra and Diaferio, 2013), while stellar masses are estimated with synthetic spectral distribution (SED) model fitting (M_{BCG}^* ranges $6.46 - 251.19 \times 10^{10} M_{\odot}$). The authors investigated several properties connecting the BCG with the hosting DM halo, in particular, by measuring the correlation between the central stellar velocity dispersion of the BCG with the cluster velocity dispersion, finding it to be remarkably tight. One of the main questions left open by Sohn et al. (2020) is whether state-of-art numerical simulations are capable of reproducing similar results, a question which we will address in this section. In the first part, we will focus on examining the differences between the observational dataset and the results from our simulations in the approach closest to the observational procedure (namely, the BCG+ICL sample). In the second part, we will investigate the effects of excluding the ICL from the stellar component.

6.4.1 Scaling relations

Figure 6.1 illustrates (from left to right panel) σ_{200} versus M_{200} , M_{BCG}^* versus M_{200} and σ_{BCG}^* versus M_{BCG}^* from the cluster sample in our simulations at $z = 0$ compared to the observed population provided by Sohn et al. (2020). In this regard, we decided to employ the simulations at $z = 0$, given that in any case, our results do not significantly change at late times. Legend is common to all panels: dark green and golden points indicate respectively the Hydro-1x and Hydro-10x samples using all the stars in the main halo enclosed in a bi-dimensional aperture of 50 kpc (BCG+ICL sample). We remind that the BCG+ICL sample is the sample selected with an approach closest to the observational one and shall be regarded as the right sample for comparison. The gray circles stand for the BCG measurements studied by Sohn et al. (2020). More in detail, in the left panel, we examine the relationship between the cluster line of sight velocity dispersion σ_{200} and cluster mass M_{200} , to understand the capability of our simulations to recover this well constrained correlation (e.g. Evrard et al., 2008; Saro et al., 2013). Indeed, once calibrated, the cluster velocity dispersion is an observable extremely sensitive to the cluster mass. We find that DM particles trace a relation fully consistent with both the theoretical expectation, a power-law with slope $1/3$ (black dashed line), and the observational data. Given this tight correlation, in the following, we will often refer to "high-mass" or "low-mass" cluster samples as those with high or low-velocity dispersion values respectively.

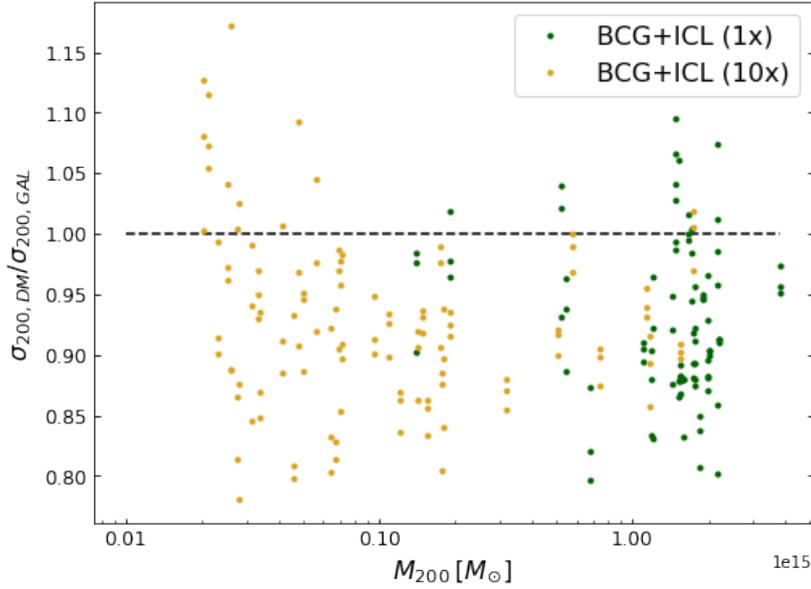


Figure 6.2: The velocity bias as a function of cluster mass in the Hydro-1x (dark green points) and Hydro-10x samples (golden points). The dashed black line reports the line for no bias ($\sigma_{200, DM} / \sigma_{200, gal} = 1$). Velocity dispersions are measured for each simulated cluster within three orthogonal projections.

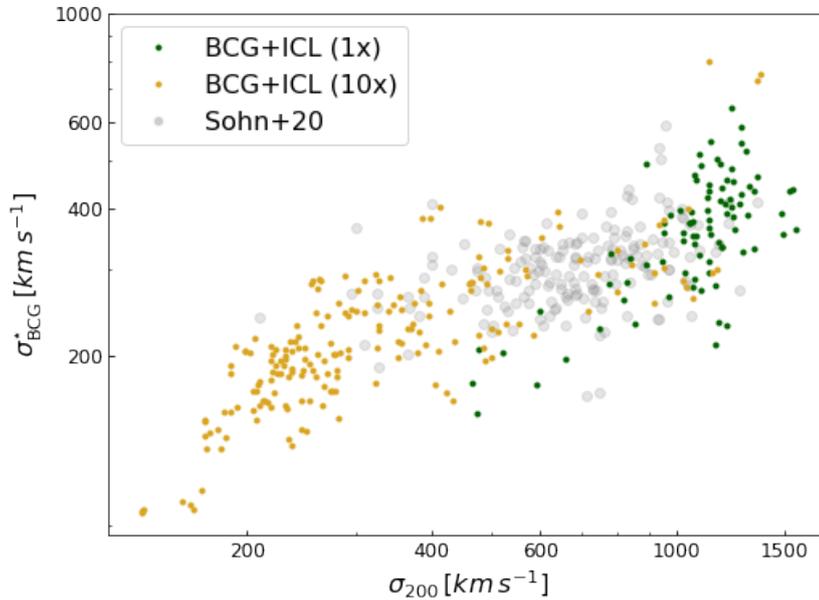


Figure 6.3: Relationship between BCG and cluster line of sight velocity dispersion from clusters in Sohn et al. (2020) (in grey), and at $z = 0$ the Hydro-1x and Hydro-10x simulations. The coloured points are the result of selecting all stars in the main halo within the cylindrical projection. Velocity dispersions are measured for each simulated cluster within three orthogonal projections.

As for the central and right panels, we verify whether the BCG properties in simulations have realistic values compared to the observational data. Among the observational properties of BCGs, that have been studied over the years, the scaling relation between the BCG stellar mass M_{BCG}^* and the host halo mass M_{200} holds a special place given its direct connection with the hierarchical growth of structure. We expect both masses to depend on a combined action of different physical processes which include (but are not limited to) halo assembly history, gas accretion, AGN feedback, and galaxy mergers. In the central panel, we examine this relation which is generally known to be difficult to reproduce in simulations, as BCG stellar masses are sensitive to the details of the feedback processes. We note that Bassini et al. (2020) and Ragone-Figueroa et al. (2018) previously discussed at length these same results on both resolutions, finding that the Hydro-10x set tends to overestimate the stellar mass of the BCGs with respect to the observational data. On the contrary, the Hydro-1x clusters have BCG stellar masses consistent with the observed ones. As discussed by Bassini et al. (2020), rather than a resolution effect, the difference in the BCG stellar masses between the Hydro-10x and Hydro-1x DIANOGA clusters lies in the different implementation of AGN feedback. Furthermore, in Ragone-Figueroa et al. (2018) the M_{200} versus M_{BCG}^* relation is compared also to other simulation results, presented by other groups (Bahé et al., 2017; Hahn et al., 2017; Pillepich et al., 2018), yielding comparable results. We note that this discrepancy is larger at the lower masses.

On the other hand, the BCG stellar mass in the Hydro-1x is compatible with the measurements by Sohn et al. (2020). The right panel shows a tension between the $M_{\text{BCG}}^* - \sigma_{\text{BCG}}^*$ scaling relation for observed and simulated clusters, at least for the Hydro-10x set. Among other factors, an overestimate of the BCG stellar mass contributes to a shift towards the right for the scaling relation predicted by our simulations. In the same plot, we note the presence of three isolated points derived from a cluster in a Hydro-10x simulation with $\sigma_{\text{BCG}}^* > 600 \text{ km s}^{-1}$. These points correspond to the three orthogonal projections of the line of sight velocity dispersion of the same cluster. To understand the origin of this outlier, we analyse its recent evolution. These extreme features can be traced back to a sequence of major mergers at late times ($z \leq 0.5$). By studying the velocity distribution of the star particles associated with the main halo we note that this structure did not yet reach dynamical equilibrium by $z = 0$, with the distribution showing two separate peaks: a central one, corresponding to the most massive subhalo, and a smaller one, corresponding to the merging structure. The reason why this is particularly evident in the Hydro-10x case, and not in the Hydro-1x, can be twofold. Firstly, it may be due to how Subfind assigns particles to one halo or the other at two resolutions. Secondly, as the resolution changes the timing of the merging also slightly changes. We point out that this feature had already emerged in Marini et al. (2021a), where we computed the scaling relation with the integrated pseudo-entropy and compared it to the velocity dispersion.

6.4.2 The $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation at $z = 0$

Stellar masses of the BCGs and the total cluster masses are not directly observable quantities, instead, they are inferred from other observable quantities. For instance, scaling relations applied on different observables, such as luminosity (e.g. Zhang et al., 2011), temperature in the X-ray (Vikhlinin et al., 2006) and velocity dispersion (Sohn et al., 2020) allow to obtain a measure of cluster masses. As for stellar masses, they are often derived by comparing synthetic spectral energy distribution models for single galaxies or through spectroscopy. For this reason, it is useful to also provide a comparison between the stellar velocity dispersion of the BCGs and the global cluster velocity dispersion, which are both quantities directly measured from spectroscopic observations.

In Figure 6.3, we show the line of sight velocity dispersion of the BCGs as a function of the cluster line of sight velocity dispersion. Our simulations span the full range of velocity dispersion probed by observations, extending to even smaller σ_{200} ($< 200 \text{ km s}^{-1}$) or, equivalently smaller masses. We expect these low- σ_{200} galaxies to be isolated early-type galaxies. On the other hand, in the mass range probed by observations ($\sigma_{200} > 200 \text{ km s}^{-1}$), at least for the case of the Hydro-10x, simulated structures show a rather good agreement with the observed stellar dynamics with a steepening of the relation on the very high-mass end.

Quite interestingly, the agreement that we find in Figure 6.3 between simulations and observations is not in line with the conclusions reached by Sohn et al. (2020) and Sohn et al. (2022), who pointed out a significant tension with the results from the hydrodynamical simulations by Dolag et al. (2010). In the latter (and more recently, including the galaxy group mass scale, in Remus et al. 2017), the stellar BCG velocity dispersions are derived from the Maxwellian fits of the velocity distribution of BCG and ICL separately. The authors find a self-similar scaling between both BCG and ICL velocity dispersions with cluster mass which suggests the existence of self-similarity between the two components over the entire mass range probed. However, such self-similar scaling is not retrieved in the observed cluster set. In the following, we will extensively discuss the main differences between our analysis of simulated galaxy clusters and that presented in Dolag et al. (2010) and Remus et al. (2017) to understand where these differences specifically lie, in the effort to clarify the origin of the previously reported disagreement between observations and simulations. Since the comparison in Sohn et al. (2020) is performed between observed velocity dispersions (thus, including ICL) and three-dimensional BCGs-only velocity dispersions, as a first step we test whether accounting for projection effects and the inclusion of the ICL contribution affect the results, but these effects turn out to be negligible in our analysis. Therefore, we test whether our cluster sets differ in the virial scaling relations. For this purpose we show the stellar velocity dispersion of BCG

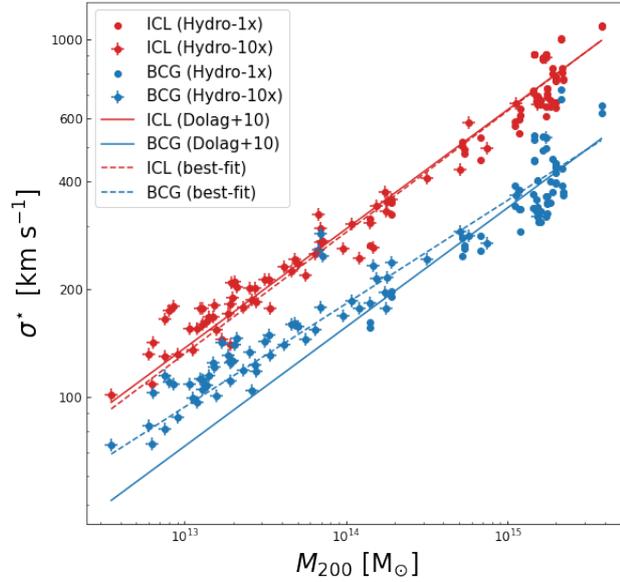


Figure 6.4: Stellar velocity dispersion of the BCG (blue) and ICL (red) components in the main halo of our set of simulated galaxy clusters, obtained from the double Maxwellian fit of the particle velocity distribution as a function of host cluster mass M_{200} . The distinct markers encode the two distinct simulations: simple circles for the Hydro-1x and crossed circles for the Hydro-10x. The solid lines represent the best-fit results from Dolag et al. (2010), while the dashed lines mark the best-fit from our sample.

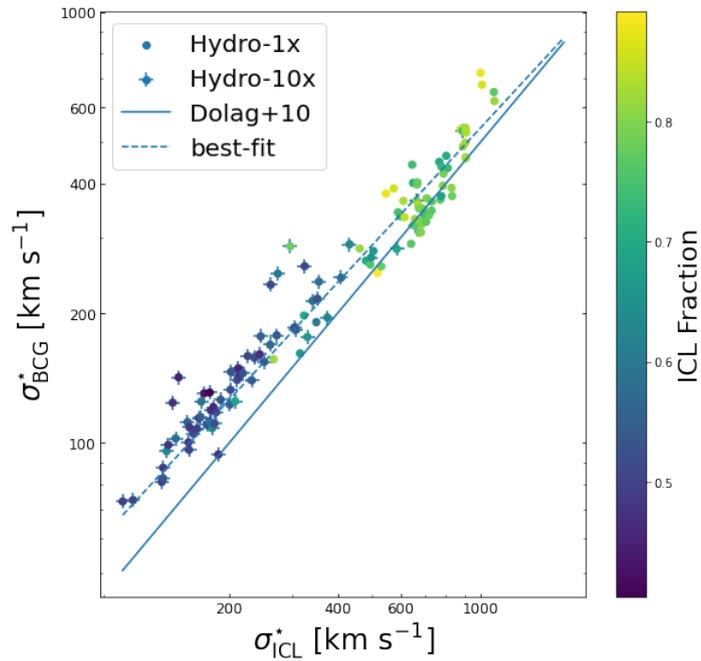


Figure 6.5: Stellar BCG velocity dispersion versus ICL velocity dispersion in the Hydro-1x (simple circles) and Hydro-10x (crossed circles) runs. Solid line reports the best-fit curve from Dolag et al. (2010) while the dashed line is derived from our cluster set. We use a colour-coded description to mark the ICL mass fraction in each cluster.

(in blue) and ICL (red) as a function of the cluster mass M_{200} in Figure 6.4. The scaling between velocity dispersion and cluster mass is modelled as:

$$M = A \left(\frac{\sqrt{3} \sigma}{10^3 \text{km s}^{-1}} \right)^\alpha \times 10^{14} h^{-1} M_\odot \quad (46)$$

where $\alpha = 3$ is the value expected from virial equilibrium and it is kept fixed in the analysis by Dolag et al. (2010). In our analysis, we decided to treat α as a fitting parameter, given that a simple comparison by eye between the virial expectation and our dataset, unlike that from Dolag et al. (2010), showed a consistent discrepancy in the slope. We distinguish between results from the Hydro-1x (simple circles) and Hydro-10x (crossed circles) simulations to clarify the dependence on resolution. Additionally, we report the best-fit relations as provided by Dolag et al. (2010) (dashed lines) and as obtained from the analysis of our set of simulated clusters (solid lines). The best-fit curves from the two distinct fits in the scaling $\sigma_{\text{ICL}}^* - M_{200}$ are comparable to each other, and our fitting value is very close to the virial expectation, with $1/\alpha \simeq 0.339$. The difference in the normalisation is minimal and still within the internal scatter. As for the $\sigma_{\text{BCG}}^* - M_{200}$ relation, our BCG set has a significantly different behaviour with respect to the best-fit line from Dolag et al. (2010), especially when expanding the dynamical range down to the mass range of galaxy groups. It is precisely this difference that explains the discrepancy with respect to Dolag et al. (2010) and Remus et al. (2017). A higher BCG velocity dispersion at fixed cluster mass (or global velocity dispersion) in the mass range of galaxy groups brings our measurements in agreement with observational results. This difference is most relevant in the low-mass end, $\sim 5 \times 10^{13} M_\odot$. The slope of our best-fit relation is shallower than the virial value ($1/\alpha \simeq 0.289$) and indicates substantial differences in the mechanisms operating in the BCG accretion with respect to those in the ICL, which traces more the overall virial dynamics of the host halo. These two effects are consistent with the different dynamical evolution of two stellar components. The ICL is associated with the stellar envelope bound to the cluster potential, and thus we expect this component to dynamically co-evolve with the DM halo, following the self-similar scaling described by the virial theorem. On the other hand, the BCG is expected to form from the dissipative collapse of the gas cooling and fueling star formation within the proto-BCG building blocks, which later assemble into the BCG dry mergers.

Furthermore, we analysed the correlation $\sigma_{\text{BCG}}^* - \sigma_{\text{ICL}}^*$. In their analysis, Dolag et al. (2010) found these two velocity dispersions to be proportional to each other, according to

$$\sigma_{\text{BCG}}^* = 0.5 \sigma_{\text{ICL}}^*. \quad (47)$$

We expect this relation not to hold for our set of simulated clusters and groups, given that our σ_{BCG}^* distribution is different in the low-mass range. Figure 6.5 presents our datapoints (circles mark the clusters in the Hydro-1x run, while circled crosses are for the Hydro-10x) with the corresponding

best-fit curve (dashed line) for a power-law with varying normalisation and slope and the best-fit in Dolag et al. (2010) (solid line). In the light of previous results, it is not surprising that the best-fit slope for our set of simulated clusters is different than 1, as implied by Equation 47, and the power-law is in fact shallower. Interestingly, we colour-coded the data points with the ICL fraction over the total stellar mass associated with the central galaxy, given that Remus et al. (2017) found that outliers in their cluster distribution were due to clusters with below 5 per cent BCG mass fractions. We find a colour gradient along the best-fit curve: data points laying on this line have larger ICL mass fractions as we move towards higher σ_{ICL}^* .

In conclusion, we discussed the origin of the different conclusions reached by Sohn et al. (2020, 2022) and by our analysis on the comparison between observed and simulated results for the BCG velocity dispersion. Since we expect the dynamics of the BCG to be different from what was observed in Dolag et al. (2010), the measurements of the stellar velocity dispersions within the innermost regions will also be different, as we have seen in Figure 6.3. This still holds even if the stellar mass of the BCG in our simulations is overestimated when compared to observational results. Hence, while it is clear that the modelling of BCG stellar mass accretion is extremely sensitive to several different factors, yet the "cold" stellar dynamics of the stars in the central regions is well reproduced at $z = 0$ in our "Hydro-10x" simulations.

6.4.3 Effect of excluding the ICL

As a final test, we investigate the effects of excluding the contribution of the ICL to the stellar velocity dispersion of the BCG. This test is based on the BCG sample which, contrary to the BCG+ICL sample, includes only the stars bound to the BCG in a cylinder long R_{200} , centred on the BCG, and within a 50 kpc aperture. We stress that this test has the only purpose of displaying the impact of removing the ICL with a dynamically motivated procedure, rather than for a comparison with observational results, which cannot separate the effect of the diffuse stellar component to the velocity dispersion measured in projection. In Figure 6.6, we show the comparison of this dataset with the BCG+ICL sample from the Hydro-10x simulation only, since we checked that results do not change with the lower resolution settings. Legend is common to all panels: brown crosses mark the BCG sample, while golden points label the BCG+ICL sample. From left to right we plot the $\sigma_{\text{BCG}}^* - \sigma_{200}$, the $M_{\text{BCG}}^* - M_{200}$ and $\sigma_{\text{BCG}}^* - \sigma_{200}$ relations respectively.

The plot in the left panel is complementary to Figure 6.3 and it shows that the velocity dispersions of the "BCG" and "BCG+ICL" cases are different albeit not significantly (within 50 kpc from the centre). Similarly, the central panel examines the amount of ICL mass that is still accounted for in the 50 kpc bi-dimensional aperture centred on the central galaxy. Unsurprisingly stellar BCG masses in the BCG-only sample are lower than

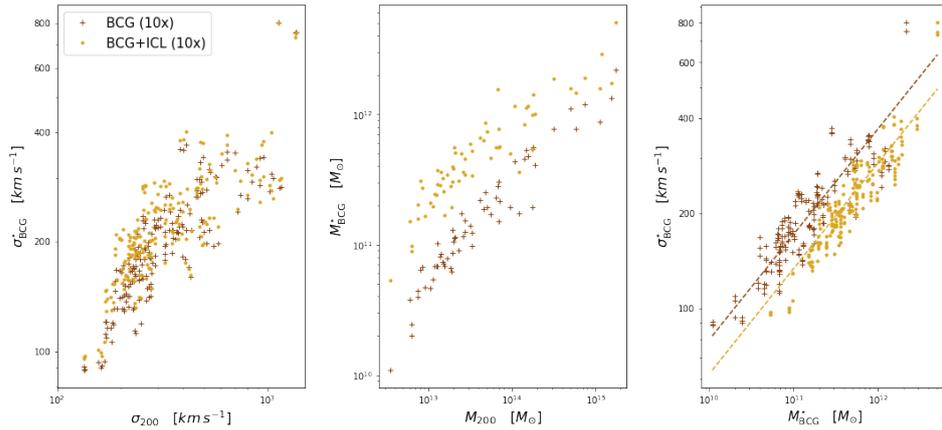


Figure 6.6: A comparison of the scaling relations for the stellar population of our simulated clusters in both BCG and BCG+ICL definitions, for the Hydro-10x sample. Left panel: scaling relation between BCG stellar velocity dispersion, σ_{BCG}^* and total cluster mass, M_{200} ; central panel: scaling relation between M_{BCG}^* and M_{200} ; right panel: scaling relation between BCG velocity dispersion, σ_{BCG}^* and M_{BCG}^* . In the last panel, we also report the predictions from the virial theorem for both samples with coloured dashed lines. Legend is common to all panels: golden points indicate the BCG+ICL sample while the brown crosses stand for the BCG sample. Velocity dispersions are measured for each simulated cluster within three orthogonal projections.

the BCG+ICL set, given that the former does not take into account the mass contribution from the stellar component assumed to be part of the ICL envelope. This mass difference accounts for a factor ~ 3 . As for the right panel, the stellar velocity dispersion as a function of the stellar mass, we note that the reduction of the stellar mass is only partially compensated by the reduced velocity dispersion, once the ICL contribution to such two quantities is removed. As a consequence, the scaling relation between velocity dispersion and stellar mass has a higher normalisation for the "BCG" case. It is striking how the predictions from the virial theorem are still conserved in both cases (a power-law with slope $1/3$ as reported by the dashed coloured lines).

In conclusion, we find that excluding the ICL component from the 50 kpc aperture on the central galaxy implies lower velocity dispersions, compatible with a cut in dynamically hot stellar population which makes up the ICL, and lower stellar masses.

6.4.4 Redshift evolution of the $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation

Having established that the observed $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation is correctly predicted by our simulations, we investigate its redshift evolution, similar to what is done for the evolution of the stellar mass of BCGs. We select five different redshifts ($z = 2, 1, 0.5, 0.2, 0$) from the Hydro-1x and Hydro-10x simulations for a total of 113, 80, 84, 80, and 93 clusters at each redshift. For all clusters, the BCG stellar velocity dispersion is computed

in projection by including the ICL contribution (BCG+ICL case) with a 50 kpc aperture. At each redshift, we group clusters in bins of σ_{200} , with a minimum of 10 clusters per bin. The value of σ_{BCG}^* assigned to each bin corresponds to the median value among the clusters belonging to that bin. We report the results of this analysis in [Figure 6.7](#). Additionally, we plot the standard deviation at $z = 0$ as a function of the cluster velocity dispersion with the grey-shaded band, noting that the standard deviation is similar at all considered redshifts.

We observe that the $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation is almost redshift independent with no significant variation within its intrinsic scatter. At $z = 2$ there is a mild indication for a slightly higher normalisation in low- σ_{200} halos. This result is possibly consistent with the picture of a rapid BCG formation through dry mergers (i.e. gas-poor and negligible star formation), whose assembly anticipates that of the host halo.

As a second test of evolution, we select the most massive FoF group for the eight regions from the Hydro-10x simulation set, and for each of them, we follow the velocity dispersion as a function of redshift up to $z = 2$. [Figure 6.8](#) shows the evolution of σ_{BCG}^* versus σ_{200} both normalised at $z = 0$, with the colour-coding indicating the redshift. On average, the cluster velocity dispersion doubles from $z = 2$ to $z = 0$. The BCG velocity dispersion also increases with time, with most of the evolution occurring at $z > 1$. The evolution of the cluster velocity dispersion is rather smooth, although episodes of merger may temporarily cause a sudden increase of σ_{200} , followed by relaxation. For instance, this is the case for cluster 5 at $z = 0.2$, whose σ_{200} is higher than at $z = 0$. The same can also happen for σ_{BCG}^* , especially if these merging events affect the inner regions. Much like the BCG stellar mass accretion, the stellar velocity dispersion has very little change at late times thus confirming that, once formed through dry mergers, it remains almost independent of the cluster mass growth. In other words, we expect the late-time mass accretion in the central regions not to significantly affect the dynamics of the BCG.

6.5 VELOCITY DISPERSION PROFILES

Stellar velocity dispersion profiles of the BCGs provide an interesting diagnostic to investigate the inner core of galaxy clusters. On one hand, they allow us to explore the dynamics of the innermost cluster regions, which is expected to reflect the assembly history of the BCG. On the other hand, the combination with strong lensing measurements allows reconstructing the total mass distribution (while also yielding the DM distribution after the subtraction of the visible matter). In this section, we present the results of our analysis of the BCGs dynamics as traced by the velocity dispersion profiles. This will be performed with a two-steps approach, according to which we will derive the expected velocity dispersion profiles for the distinct tracers in galaxy clusters, while later we will assess the main features that characterise the observed BCG stellar population and quantify the ICL intake in these profiles.

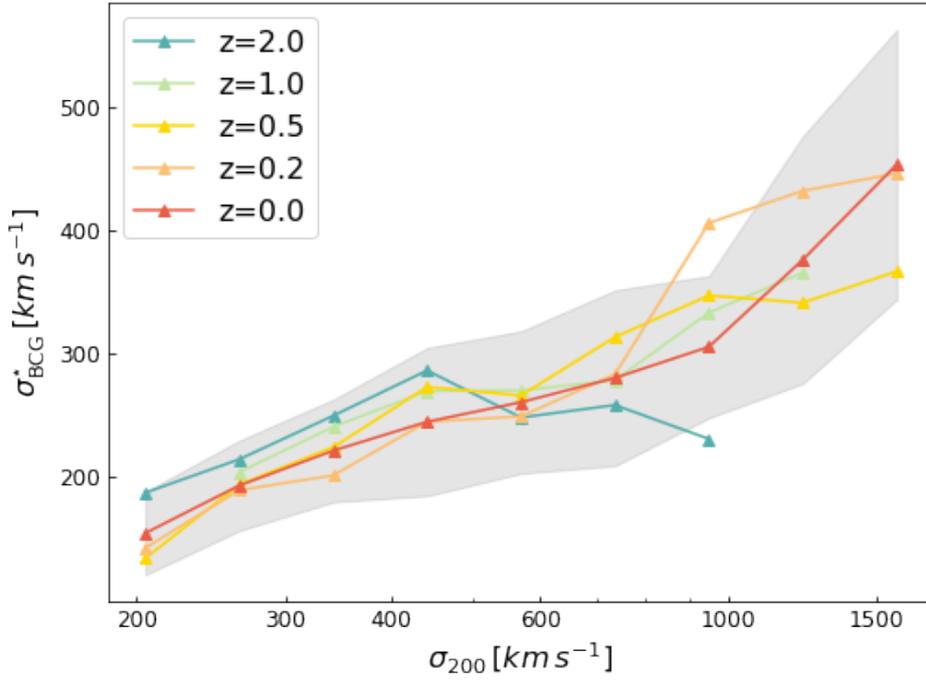


Figure 6.7: Relationship between line of sight BCG velocity dispersion as a function of cluster velocity dispersion, at five different redshifts. Clusters at each redshift are binned in σ_{200} . For each bin, we plot the corresponding median value of σ_{BCG}^* . The grey-shaded band marks the standard deviation of the clusters at $z = 0$, with comparable dispersions at different redshifts. At the different redshifts, $z = 2, 1, 0.5, 0.2, 0$, we have the following corresponding number of clusters: 113, 80, 84, 80, and 93. The value of σ_{BCG}^* is obtained by also including the contribution of ICL.

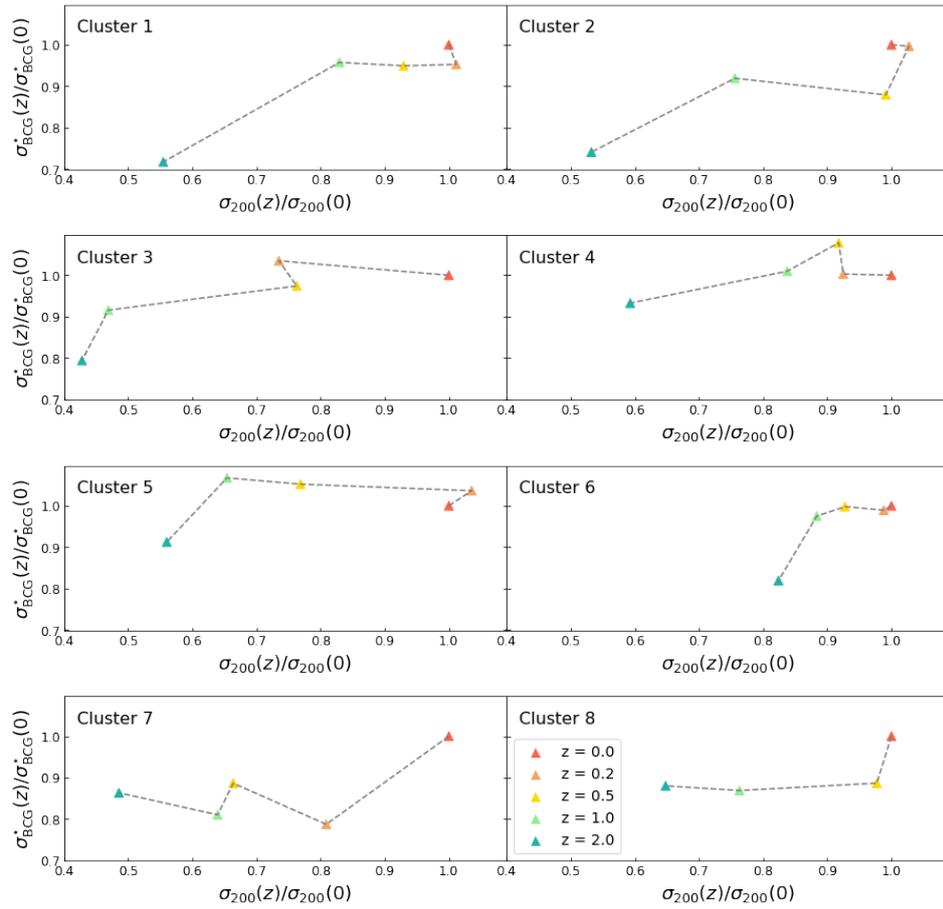


Figure 6.8: Evolution of the scatter relation between the velocity dispersion of the BCG (normalised at $z = 0$) versus the cluster velocity dispersion (normalised at $z = 0$) from 8 clusters in the Hydro-10x run. The redshift bins ($z = 2, 1, 0.5, 0.2, 0$) are labelled with different colours.

6.5.1 Dynamical properties of cluster core regions

Recently Sartoris et al. (2020) determined the full dynamical analysis of Abell S1063 ($R_{200} = 2.36$ Mpc, as inferred from weak lensing measurements) combining two different tracers: the kinematics of the cluster galaxy members and the stellar velocity dispersion profile of the BCG. For this cluster at $z = 0.3$, the observations consisted of an intensive spectroscopic campaign using the VIMOS and MUSE spectrographs at the VLT. With our analysis, we want to verify whether our simulations of clusters of comparable size reproduce the cluster dynamics as traced by the galaxy population and, in the innermost region, by the BCG. For this comparison, we extend the aperture radius to include all the stars in the main halo out to R_{200} . This has the advantage to provide us with information on the dynamics traced by both the stellar component of the main cluster halo and by the cluster galaxies at large radii.

Figure 6.9 shows the comparison between observational results from Sartoris et al. (2020) and one of the simulated clusters from the Hydro-10x runs at redshift $z = 0$. The latter was selected for its closeness in size to the observed cluster ($R_{200} = 2.34$ Mpc). Solid lines are for the sample obtained from simulations, while the filled and open circles correspond to the stars and galaxies of Abell S1063, respectively. Error bars on observational data points refer to the reported 68 per cent uncertainties. The accuracy of our simulated cluster in reproducing the observed data points is quite remarkable, especially for the velocity dispersion profile of stars. In the outer regions, we see that DM and galaxies profiles from the simulations both capture the observed negative gradients, albeit with a slightly lower normalisation for $r < 1$ Mpc. This difference can be due to several factors. Indeed, we stress that the cluster selected from simulations has been chosen for its size (the most massive in our cluster sample and the closest to the observed cluster mass), thus differences in the dynamics of its components can be ascribed to distinct dynamical states or specific cluster formation histories. A further step can be taken by extending the comparison to include the eight clusters from the Hydro-10x set of simulations to infer the general characteristics of the distinct tracers when combining all the simulated clusters. The results are shown in Figure 6.10, with shaded areas showing the 16th and 84th percentiles given by the entire cluster sample.

Except for particularly disturbed systems, as it has been discussed in Section 6.4.1, BCGs present minimal dynamical differences from cluster to cluster. On the contrary, as we move towards the outskirts, the contribution from the ICL increases. This component feels more the effects of the overall cluster potential, which is more sensitive to the global dynamical state of the cluster and the accretion pattern from the surrounding large-scale structure (e.g. Evrard et al., 2008; Saro et al., 2013), thus producing a larger scatter in the velocity dispersion profiles. Given that we are especially interested in examining the inner regions, we decided to stack the stellar profiles without any suitable rescaling of their amplitudes. Conversely, DM particles and galaxies' velocity dispersion profiles require

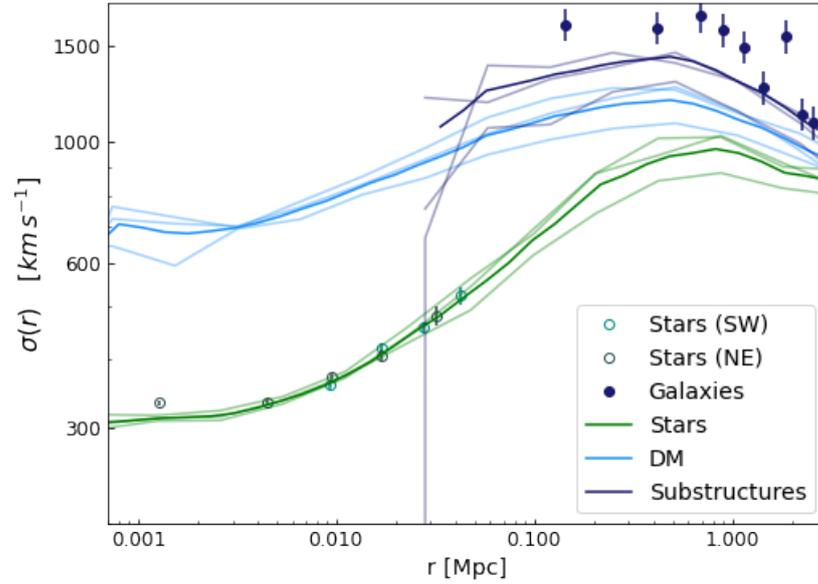


Figure 6.9: Comparison with the line of sight velocity dispersion from Sartoris et al. (2020) (points with error bars) and the projected velocity dispersion profiles from one of our simulated clusters at $z = 0$ in the Hydro-10x set. Error bars in the observational points refer to a 68 per cent confidence level. The solid lines are the mean of the three projected velocity dispersion profiles (faded lines).

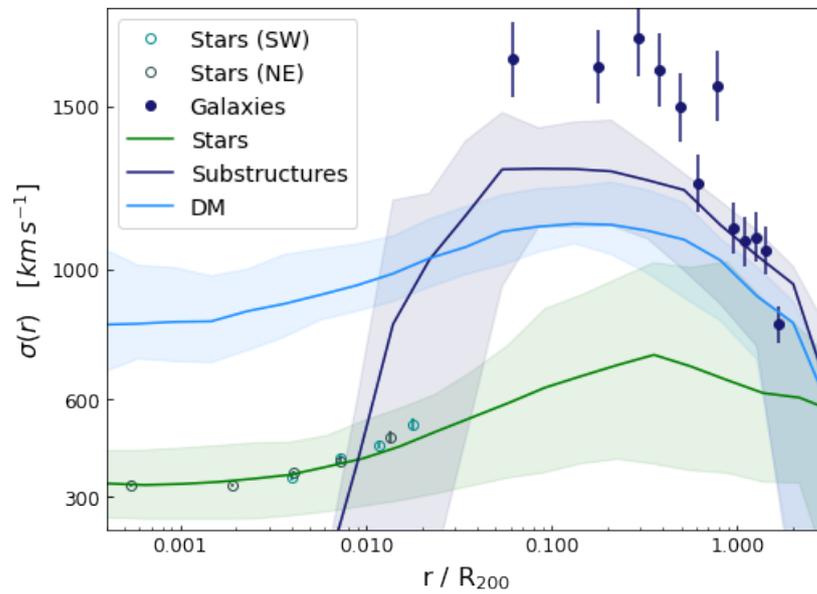


Figure 6.10: Comparison of the observed line of sight velocity dispersion from Sartoris et al. (2020), with results from the combined sample of simulated clusters at $z = 0$ from the Hydro-10x set. Error bars in the observational points refer to a 68 per cent confidence level. The shaded areas represent respectively the 16th and 84th percentiles given by the simulated cluster sample.

to be normalised to the cluster radius before stacking (provided that $\sigma(r) \propto r^{-1}$, as discussed in [Section 5.3](#)). Once we stacked the median profiles of simulated clusters, we multiply the resulting profile by the virial radius of Abell S1063, $R_{200} = 2.36$ Mpc. We see that in the internal regions, DM particles have a larger velocity dispersion than the stellar component, a consequence of the dissipative collapse which forms the BCG stars and determines their "colder" dynamics. As for substructures, since they undergo tidal effects during the merger process and are disrupted when reaching the cluster centre, the profile abruptly falls for $r/R_{200} < 0.1$. Moving away from the cluster centre, the three components converge to similar profiles. At large radii, the median profile spans values of the velocity dispersion which are lower than those in the observed cluster, similarly to what already encountered in [Figure 6.9](#). Differences at this level can be again ascribed to several factors which may include distinct dynamical states and/or projection effects. At these scales, we find the substructures' velocity dispersion profile to be consistently higher than that of the DM case. Analogously to what was discussed in [Section 6.4.1](#), we observe a velocity bias between the two tracers due to the effect of tidal stripping which is more effective in substructures with relatively low orbital velocity. The resulting mechanism is the selective removal of lower-velocity substructures, and thus an increase in the total velocity dispersion profile as traced by substructures.

6.5.2 Gradients of the stellar velocity dispersion profile

Velocity dispersion profiles of the BCGs relate to the dynamical mass profiles of clusters and are expected also to be determined by the formation history of such extremely massive galaxies. Therefore, it is instrumental to investigate and deepen our comprehension of the velocity dispersion profile and its gradient. Up to now, studies of velocity dispersion profiles of BCGs have exhibited a large variety of slopes, with a significantly larger fraction of positive slopes for BCGs, with respect to other early-type galaxies and brightest group galaxies (e.g. Von Der Linden et al., 2007; Bernardi, 2009; Huertas-Company et al., 2013).

We present here the analysis of the slopes of the velocity dispersion profiles of BCGs in our set of "Hydro-10x" simulated clusters and compare with observational results obtained by Loubser et al. (2018). We adopt their model to describe the projected velocity dispersion profiles of BCGs measured within a circular aperture of radius r_{lim} :

$$\log \left(\frac{\sigma_{\text{BCG}}^*(r < r_{\text{lim}})}{\sigma_0} \right) = \eta \log \left(\frac{r}{r_0} \right). \quad (48)$$

The BCG velocity dispersion σ_{BCG}^* is parametrised in function of a central velocity dispersion σ_0 measured within a smaller aperture with radius r_0 from the BCG centre. Loubser et al. (2018) (and Newman et al. 2013 therein) obtained the total BCG velocity dispersion within a radius $r_{\text{lim}} = 15$ kpc whose central aperture is taken as $r_0 = 5$ kpc. This choice is motivated by

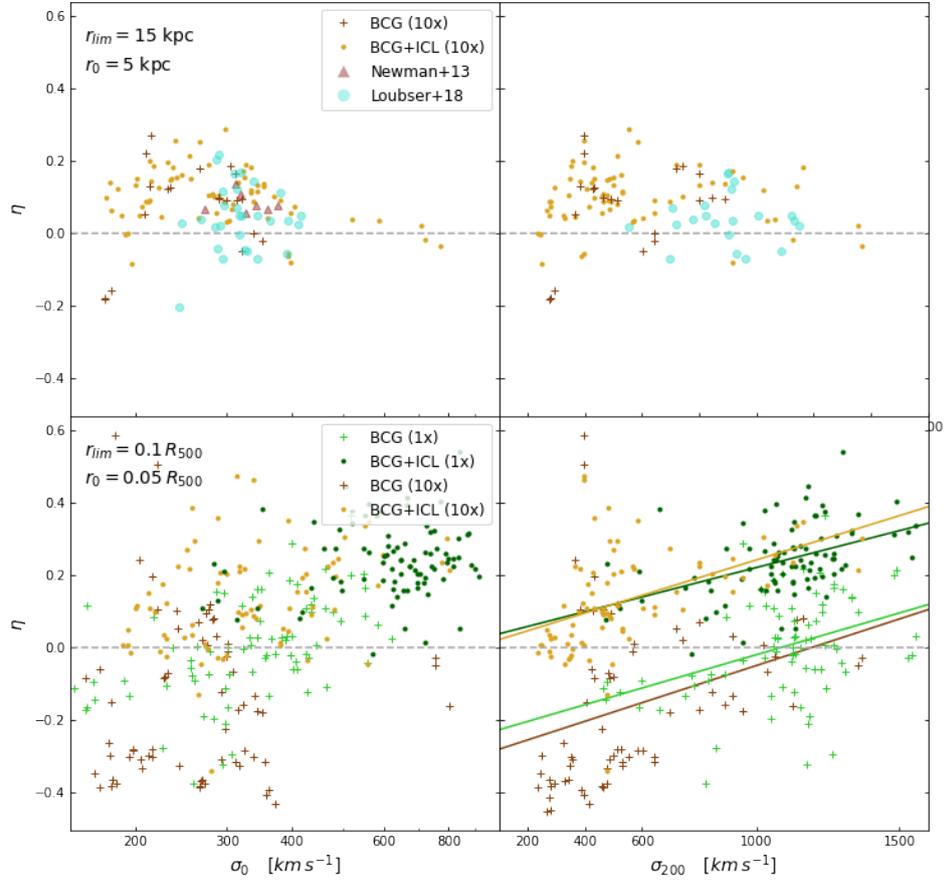


Figure 6.11: Top panels: comparison of the slope of the stellar velocity dispersion η with the observational results from Newman et al. (2013) (maroon triangles) and Loubser et al. (2018) for BCG (in light turquoise circles). The legend is common to all panels: dark green and golden points indicate the Hydro-1x and Hydro-10x's BCG+ICL samples, respectively; light green and brown crosses label the BCG-only sample from the Hydro-1x and Hydro10x runs, respectively. Following the approach adopted in Loubser et al. (2018), we report the slopes η obtained by fitting the velocity dispersion profiles over the radial range from a central aperture of $r_0 = 5$ kpc and 15 kpc. We show results only for the Hydro-10x sample. Bottom panels: the same analysis as in the top panels, but computing the slope of the velocity dispersion profiles between $0.1 R_{500}$ and $0.05 R_{500}$, with the solid lines showing the best-fit results. The left and right panels show the slope η of the velocity dispersion profiles versus the central velocity dispersion σ_0 and the cluster velocity dispersion σ_{200} , respectively.

previous studies (Graham et al., 1996) which measured a typical half-light radius R_e of 16.7 kpc in 119 Abell clusters.

As for the simulations, we decided to follow two different approaches. In one case, for the comparison with observational results, we select a bi-dimensional radial aperture $r_{\text{lim}} = 15$ kpc and compute the profiles on the Hydro-10x runs only. We exclude from this analysis the Hydro-1x clusters since their force resolution does not allow to adequately resolve the scales involved in the observational analyses. In the second case, we let r_{lim} to vary with the size of cluster by taking $r_{\text{lim}} = 0.1 R_{500}$: this value proved to include the BCG self-consistently with cluster size (Ragone-Figueroa et al., 2018, and references therein). In this case, the central aperture is also chosen to scale with the cluster radius, while being significantly larger than the softening length of both simulations with $r_0 = 0.05 R_{500}$. While this analysis cannot be directly compared with observational results, it allows us to quantify the role of the ICL in determining the profiles over fixed fractions of the characteristic cluster scale radius.

The velocity dispersion slope η as a function of the central velocity dispersion σ_0 (or cluster velocity dispersion σ_{200}) is reported in the top left (right) panel of Figure 6.11. The legend is common to all panels: in golden the Hydro-10x, in dark green the BCG+ICL sample of the Hydro-1x, and the BCG sample with light green crosses; maroon triangles and turquoise circles are the Newman et al. (2013) and Loubser et al. (2018) observational sets respectively. The top panels show the results of the analysis with $r_{\text{lim}} = 15$ kpc, while the bottom panels refer to the analysis based on the different cut on the radial aperture r_{lim} . In the former, simulations predict slightly positive slopes, a result generally consistent with observational results from Loubser et al. (2018) and Newman et al. (2013). We also note that our set of simulated BCGs shows few cases of negative slopes that often correspond to distinct projections of the same clusters having a three-dimensional negative radial gradient. These clusters do not seem to yield any distinct features with respect to the other clusters in the simulated set, except for the different gradients.

In the bottom panels, we report the results from all runs (i.e. Hydro-1x and Hydro-10x, for both BCG+ICL and BCG cases), given that they should not be contaminated by resolution effects for the chosen radial range over which profiles are analysed. On the left, we show the slope η as a function of the central velocity dispersion σ_0 : allowing r_0 to scale with R_{500} yields a different sampling of σ_0 that extends to larger values than for the observational analysis shown in the upper panels. Furthermore, we observe that the BCG-only sample generally yields lower values of the gradients with respect to their counterpart BCG+ICL. Similar results are obtained in the right bottom panel which analyses the relationship between η and the total cluster velocity dispersion σ_{200} . The generally larger values of η at fixed σ_0 (or σ_{200}) when ICL stars are included is due to the contribution of this dynamically warmer stellar component that increases the velocity dispersion when measured along the line of sight. At the same time, removing the ICL contribution still allows finding both

positive and negative η , thus indicating that positive gradients are not necessarily due to the larger velocity dispersion of the ICL, but also reflect intrinsic dynamical properties of the BCGs. Additionally, here both Hydro-1x and Hydro-10x distributions hint at a positive correlation between η and σ_{200} , suggesting that more massive halos have increasing BCG velocity dispersion slopes.

6.6 CONCLUSIONS

In this chapter, we address the reliability of simulating the physical properties of the BCGs, stressing the importance of the role played by the stellar velocity dispersion. The analysis is performed with a suite of cosmological hydrodynamical simulations, the DIANOGA set (Ragone-Figueroa et al., 2018; Bassini et al., 2020), obtained with the code GADGET-3 at two distinct resolutions (the base resolution is called Hydro-1x, the higher one is Hydro-10x). We are mostly interested in studying the stellar component associated with the main halo in the inner core (< 50 kpc) of galaxy clusters which is composed of the stars bound to the BCG and the ICL, the diffuse stellar envelope that wraps the cluster halo. According to our definitions, we name "BCG+ICL" this stellar envelope, which is mostly consistent with what is detected in observations, while we apply an unbinding procedure to disentangle the two components, that are expected to yield distinct dynamics and formation histories, to only keep the stars bound to the BCG, referring to this sample as "BCG" or "BCG-only". This separation between BCG and ICL stellar components is implemented to examine the impact of excluding the warmer stellar component, i.e. the ICL, from the BCG velocity dispersion, a step that is extremely challenging in observational analyses.

The main results of our analysis can be summarised as in the following.

- Our DIANOGA set of simulated clusters has proven to reproduce fairly well the "cold" stellar dynamics of the BCG when considering the correlation between the BCG velocity dispersion and cluster velocity dispersion, $\sigma_{\text{BCG}}^* - \sigma_{200}$, as shown in Figure 6.3: stellar velocity dispersions of simulated galaxy clusters agree with observational results from a set of local BCGs by Sohn et al. (2020). We find this agreement to hold, even if simulations tend to produce too massive BCGs when compared to observations (as shown in Figure 6.1).
- The $\sigma_{\text{BCG}}^* - \sigma_{200}$ relation is redshift-independent and features a tight distribution with σ_{BCG}^* mostly spanning $\sim 100 - 600$ km s $^{-1}$, as reported in Figure 6.7. This is further supported when following the late accretion of the most massive FoF groups in several Lagrangian regions from redshift $z = 2$ in Figure 6.8: simulations exhibit BCG velocity dispersions increasing at most by 30 per cent, against the cluster velocity dispersions that, on average, double their values over the same redshift interval. The slow growth of the central galaxy

since $z = 1$ is only occasionally interrupted by late-time mergers that impact the value of BCG and/or cluster velocity dispersion.

- Our simulations describe quite well the observed dynamics yielded by distinct tracers, i.e. stars, DM particles, and galaxies. From [Figure 6.9](#) and [6.10](#), we argue that is especially true for stars in the innermost regions which make for most of the dynamically cold stellar component in the central galaxy. At larger radii, the presence of the diffuse ICL increases causing a rise in the velocity dispersion profiles of stars which, in turn, tend to recover a shape more similar to that of the velocity dispersion traces by DM and galaxies.
- Galaxy clusters, in both simulations and observations, are characterised by a large variety of profiles in the radial stellar velocity dispersion, as seen in [Figure 6.11](#). Considering the central stars (the stellar component within a bi-dimensional aperture of 50 kpc), we find a preference for positive gradients which is only partially due to the increasing velocity dispersion from the ICL, but rather it is an intrinsic property of the dynamics of BCGs. Furthermore, simulations reproduce the observed correlation between the slope of the velocity dispersion profile η and the cluster velocity dispersion (and thus, cluster mass) when the stellar velocity dispersion is computed in projection within a circular aperture whose size scales with cluster radius (i.e. $0.1R_{500}$).

One of the main conclusions is that state-of-art simulations provide a reliable tool to interpret the dynamical processes operating at the centre of galaxy clusters and determining the projected phase-space structure of the BCG and the surrounding diffuse stellar component making up the ICL. These regions represent the main stage for the most significant evolutionary phenomena taking place (e.g. dynamical friction, mergers, galactic cannibalism) which largely impact cluster properties at these scales. To this end, studying the formation and evolution of the stellar content of the BCG (and surrounding ICL) can provide us with insights into these mechanisms and their effect on the cluster as a whole. Our results show that stellar velocity dispersions robustly connect to the cluster velocity dispersion, as seen in observational analyses (Loubser et al., 2018). Furthermore, recent measurements of stellar velocity dispersion profiles in the inner core of galaxy clusters, based on integral field spectroscopy (Sartoris et al., 2020), have enabled us to directly compare high precision measurements with the simulated stellar population at these scales, yielding remarkably good results.

Tracing the BCG and ICL dynamics in the innermost regions of galaxy clusters through detailed spectroscopic observations is expected to shed light on the processes driving star formation and leading to the assembly of the BCG since the infancy of proto-clusters at redshift $z > 2$. In addition, the reconstruction of mass profiles at small clustercentric radii should also provide information on the nature of DM, once the role of baryons is properly accounted for. In this respect, high-resolution cosmological

hydrodynamical simulations provide an ideal tool to follow the process of BCG assembly and the interplay between DM and stellar dynamics. As highlighted by the analysis presented in this paper, such simulations can account for many dynamical properties of the stellar component inside and around BCGs. On the other hand, a general limitation of the current generation of simulations, including ours, is that they tend to overpredict BCG stellar masses. This calls for the need to further improve the numerical description of feedback mechanisms, most probably related to AGN, included in such simulations for them to keep the pace of the increasing quality and quantity of data expected from the next generation of observational facilities.

IDENTIFYING INTRACLUSTER LIGHT WITH MACHINE LEARNING

Nowadays, ML techniques offer fast and efficient solutions for classification problems that would require intensive computational resources via traditional methods. In this chapter, we examine the use of a supervised Random Forest to classify stars in simulated galaxy clusters after subtracting the member galaxies. These dynamically different components are interpreted as the individual properties of the stars in the BCG and ICL. We employ matched stellar catalogues (built from the different dynamical properties of BCG and ICL) of 29 simulated clusters from the DIANOGA set to train and test the classifier. This work suggests the importance of employing ML to speed up a computationally expensive classification in simulations. We aim to answer the following scientific questions.

Can we resolve the differences in the stellar populations using ML techniques? Can we make accurate predictions compared to the expected class? Where do we observe most of the uncertainty? Can we estimate an error in our predictions? Is the ML providing a reliable classification?

The content of this chapter largely reflects what is presented in a paper published in Monthly Notices of the Royal Astronomical Society.

- **Marini I**, Borgani S, Saro A, Murante G, Granato GL, Ragone-Figueroa C, Taffoni G "Machine learning to identify ICL and BCG in simulated galaxy clusters." Monthly Notices of the Royal Astronomical Society, 2022, 514.2: 3082-3096.

7.1 SCIENTIFIC CONTEXT

In recent years, the diffuse stellar envelope observed in groups and clusters of galaxies, called ICL, has assumed a prominent place in the study of structure formation. This visible tracer exhibits properties that are rather peculiar, distinct from the other stars confined in their constituent member galaxies (e.g. Contini, 2021; Montes, 2022, and references therein). Both theoretical and observational evidences (Murante et al., 2004; Murante et al., 2007; Puchwein et al., 2010; Mihos et al., 2016; Montes and Trujillo, 2018, 2019; Spavone et al., 2020; Kluge et al., 2020, just to quote a few) have been gathered on the origin and evolution of this component. Recent findings have suggested that the ICL distribution follows the global potential well of the host galaxy cluster (e.g., Montes and Trujillo, 2019; Alonso Asensio et al., 2020; Cañas et al., 2020) and thus, it can be used as a luminous tracer for DM, highlighting its importance in the context of structure formation.

Observationally constraining the properties of the ICL is troublesome, as it requires both deep and wide observations of spatially extended low-surface brightness regions in the sky, other than a top-level data processing pipeline to avoid spurious contamination from other sources. To further complicate the scenario, the evolution of the ICL is tightly connected to the build-up of the BCG, i.e., the central galaxy in a cluster, which sits at the centre of the cluster gravitational potential. Both the spatial extent and luminosity curves of the two components smoothly merge, leaving no trace of the transition (Bender et al., 2015; Kluge et al., 2020). Therefore, the separation of the ICL from the BCG is performed in several (often laborious) ways. Some studies (e.g. Kluge et al., 2020; Spavone et al., 2020) identify the ICL as the excess of light with respect to a de Vaucouleurs profile or a double Sérsic decomposition, while often it is preferred to perform a simple cut in surface brightness (Mihos et al., 2016; Montes and Trujillo, 2018). In this regard, several studies (e.g. Contini et al., 2022; Montes, 2022, and references therein) have discussed the role of the transition radius, i.e. the clustercentric distance at which the ICL component starts dominating the stellar component. Due to the variety of methods employed to estimate the ICL contribution, the value of this transition radius may depend on the adopted method of ICL identification. From the observational side, typical values of the transition radius are around 60 – 80 kpc (Montes et al., 2021; Gonzalez et al., 2021), thus in line with results from earlier works (Zibetti et al., 2005; Gonzalez et al., 2007; Seigar et al., 2007; Iodice et al., 2016). These values slightly increase for other analyses, such as those presented by Zhang et al. (2019), who concluded that the transition from the BCG to the ICL is just outside 100 kpc, or by Chen et al. (2022) who found values ranging in the interval 70 – 200 kpc. Results based on simulations (e.g., Contini and Gu, 2021; Contini, 2021; Contini et al., 2022) agree with these observational results, and indicate that the transition radius is independent of both BCG+ICL and halo masses, with typical values of 60 ± 40 kpc, if similarly derived from profile fitting. Usually, this technique requires the assumption of a double/triple Sérsic profile (Sérsic, 1963) or a composition of different profiles such as the Jaffe profile (Jaffe, 1983, describing the BCG distribution) and NFW profile (Navarro et al., 1997) for the ICL.

To our advantage, in simulations, one can exploit the full 6D phase-space information available on star particles to investigate the properties of the ICL and BCG (Dolag et al., 2010; Remus et al., 2017). It is in this direction that Dolag et al. (2010) have invested their effort in designing a classification algorithm applicable to the star particles in the main halo of simulated clusters and groups according to their properties in phase-space. The assumptions underlying this method derive from the study of the velocity distribution of star particles which exhibit a bimodal distribution that can be associated with two distinct dynamical components. Combining this information with an unbinding procedure leads to separation into a central BCG (more compact and dynamically cold) and a diffuse ICL. Although this method should not be regarded as a procedure with outputs immediately comparable to observations, it provides us with the

dynamical information associated with each component. In other words, we expect this technique to convey information on the physical properties of both stellar components, to complement the observational data.

To our disadvantage, the large volume of data to classify in state-of-the-art simulations requires intensive computational effort. To overcome this limitation, the analysis presented in this chapter aims at reproducing a similar classification method adopting ML techniques that often prove to be less computationally expensive and more efficient than traditional methods. An automated methodology for efficiently classifying the stellar components can be an essential ingredient to facilitate the use of these tools in nowadays analyses. To this end, we build a Random Forest classifier to recognise the label of a star particle solely basing the decision on the specific features of each particle. This method is widely employed in ML problems for its versatility and its performance with high-dimensional data. One essential benefit is that the computational cost of Random Forest models does not depend significantly on the size of the training set, given that it scales logarithmically. Additionally, the predictions are straightforward to interpret, while it is also extremely easy to measure the relative importance of each feature in the predictions.

The chapter is structured as follows. In [Section 7.2](#) we present the synthetic cluster set on which we train, cross-validate, and test the classifier. Furthermore, we include a description of the traditional method used to calibrate the ML model. [Section 7.3](#) describes the model and its caveats; in [Section 7.4](#), we discuss the achieved classification performance with distinct clusters and assess the reliability of the model to recover the true label, as identified by the ICL-Subfind. In [Section 7.4.5](#) we show a few examples of shell galaxies candidates found in the simulation set. Finally, we present our conclusions in [Section 7.5](#).

7.2 SIMULATIONS

The ML algorithm is trained, cross-validated, and tested on 29 clusters from a set of cosmological hydrodynamical simulations called DIANOGA. These simulations were carried out with the GADGET-3 code, a modified version of the GADGET-2 tree-PM smoothed particle hydrodynamics (SPH) public code (Springel, 2005). The major changes include a higher-order kernel function, a time-dependent artificial viscosity model, and a time-dependent artificial conduction scheme.

The 29 simulated clusters (for simplicity called D1, D2,..., D29) are the result of zoom-in simulations centred on the most massive galaxy clusters evolved in a lower-resolution N-body parent box of $1 \text{ h}^{-3} \text{ Gpc}^3$ volume with the inclusion of baryons. The cosmological model is a Λ CDM with the following parameters $\Omega_M = 0.24$, $\Omega_b = 0.04$, $n_s = 0.96$, $\sigma_8 = 0.8$ and $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$. These clusters represent the 24 most massive clusters in the parent box with masses $M_{200} \in [0.8 - 2.7] \times 10^{14} \text{ h}^{-1} M_\odot$ and 5 isolated groups with M_{200} within $[1-4] \times 10^{14} \text{ h}^{-1} M_\odot$. In the high-resolution regions, the DM particle mass is $m_{\text{DM}} = 8.3 \times 10^8$

$h^{-1} M_{\odot}$ and the initial mass of the gas particle is $m_{\text{gas}} = 3.3 \times 10^8 h^{-1} M_{\odot}$. The Plummer equivalent gravitational softening for DM particles is set to $\epsilon = 5.75 h^{-1} \text{ kpc}$. The gravitational softening lengths of gas, star, and black hole particles are $5.75 h^{-1} \text{ kpc}$, $3 h^{-1} \text{ kpc}$ and $3 h^{-1} \text{ kpc}$, respectively. Several subgrid models included in the simulations treat the unresolved baryonic physics of the simulations. Details can be found in Ragone-Figueroa et al. (2018) and references therein.

7.3 RANDOM FOREST

Given the nature of the ICL-Subfind algorithm, based on well defined properties of the stellar components, its action may be also replicated by a ML model in a faster and more efficient way. Our goal is to provide an alternative classification method for identifying stars in the main halo according to several features that are crucial in the use of the former method. To achieve this, we design a supervised classification method, based on the Random Forest classifier (based on Pedregosa et al., 2011), to which we feed a feature vector representative of the classes we are predicting (i.e., BCG and ICL). Examples of input features we tested are the potential and total energy of each particle, the particle age, mass, 3D position, and 3D velocity.

Therefore, we aim to design an adequate architecture of the Random Forest (e.g., number of trees, number of features to consider when looking for the best splits) to make the most accurate predictions.

7.3.1 Data set and training phase

We collected the data for the training set (later divided to perform cross-validation) and test set from the star particles in the simulated clusters and the output of the ICL-Subfind. The original set of simulated galaxy clusters is composed of 29 objects. We analyse the properties of these galaxy clusters to gather a fair sample of the cluster set. Besides all properties listed by Subfind (such as mass and radius), we determine the dynamical state of the host cluster (i.e. relaxed, disturbed, or intermediate), which is a good metric to derive the "thermalisation" level of the particle phase-space distributions in a cluster. Particularly disturbed objects (e.g., after halo merging events) may not have a well defined Maxwellian shape in the particle velocity distribution, thus complicating the fitting procedure used in splitting the stellar components. Thus, estimates of dynamical states are performed following the prescription described in Neto et al. (2007) based on two properties: the centre shift (identified as the distance between the minimum position of the gravitational potential x_{min} and the centre of mass x_{cm}) and the fraction of mass in substructures f_{sub} . We use the same threshold parameters as in Biffi et al. (2016). A halo is classified according to its dynamical state following Equation 32. After applying this classification to the 29 clusters at redshift $z = 0$, we find 6 relaxed, 8 disturbed, and 15 intermediate cases. The physical properties taken into

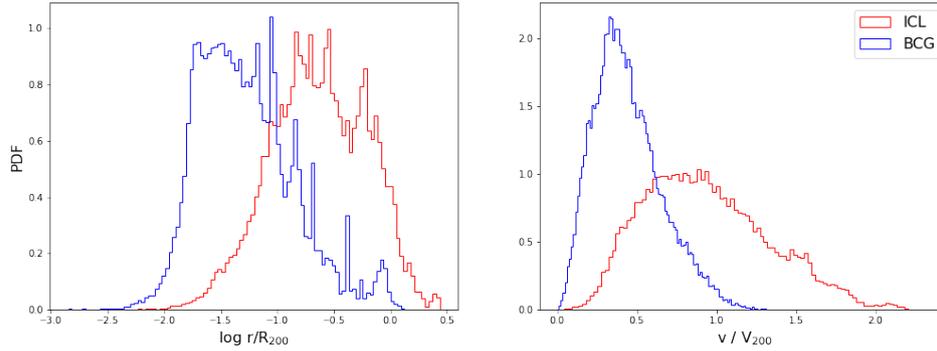


Figure 7.1: Probability density functions of the input features associated with the ICL (red) and BCG (blue) in one of the clusters. Left panel: distribution of the logarithmic cluster distance over R_{200} . Right panel: stellar rest-frame velocities scaled by the virial circular velocity V_{200} .

account for this selection (cluster mass, stellar mass of the central galaxy, and dynamical state) are listed in [Table 7.1](#).

Furthermore, we excluded a priori from our choice three clusters (that is, D7, D11, and D13) that did not reach convergence in the ICL-Subfind output while still retained as part of the test set to prove that our ML model can overcome challenging classifications for the traditional algorithm.

Out of the 26 remaining clusters, we draw 10,000 star particles each randomly selected from five clusters (i.e., D3, D9, D10, D18, and D22) for a total of 50,000 particles. ICL and BCG are represented in this sample with proportions 65 : 35. We divide the training and test sets assigning 2/3 to the former and 1/3 to the second.

7.3.2 Input features

The predicting power of a ML model heavily depends onto what extent the input features of the data set are representative of the classes one hopes to recover. Before ultimately evaluating our classifier’s performance, we infer the combination of features that best match the two classes at hand. Starting from a large parameter space, we find that the particle clustercentric distance and the module of the rest-frame particle velocity (with respect to the stellar centre of mass within R_{200}) offer most of the dynamical information needed to disentangle the two components, given that they closely relate to the particle energies employed in the ICL-Subfind unbinding procedure. [Figure 7.1](#) shows the probability density distributions of these two features drawn from one of the clusters in our simulations. To clarify the separation in the phase-space, we plot the histograms of the BCG (blue) and ICL (red) separately.

To obtain this result, we firstly examined the classifier’s performance using a larger set of possible particle properties (e.g., distance from the cluster centre and velocity relative to it, cluster mass, age, metallicity, kinetic energy and potential energy) and recorded the metric scores. Since

Table 7.1: A summary of the main characteristics of the simulated clusters used in the training and testing phases at $z = 0$. We report the given cluster name, cluster mass M_{200} , the cluster radius R_{200} , the cluster orbital velocity V_{200} , the stellar mass in the main halo $M_{*,\text{gal}}$ and the dynamical state. We add an asterisk to the clusters which are part of the training set.

centring

Name	M_{200}	R_{200}	V_{200}	$M_{*,\text{gal}}$	Dynamical state
$\star = \text{Training set}$	$[10^{15} M_{\odot}/h]$	$[\text{Mpc}/h]$	$[\text{km s}^{-1}]$	$[10^{10} M_{\odot}/h]$	
D1	1.26	1.76	1758	1402	Intermediate
D2	0.39	1.19	1188	399	Intermediate
D3*	0.49	1.28	1282	599	Intermediate
D4	0.38	1.18	1176	348	Disturbed
D5	0.14	0.84	840	177	Relaxed
D6	1.12	1.69	1687	1077	Intermediate
D7	1.10	1.68	1680	1220	Intermediate
D8	1.24	1.74	1746	843	Disturbed
D9 ¹	0.10	0.76	756	125	Relaxed
D10 ¹	1.04	1.64	1647	1342	Disturbed
D11	0.86	1.55	1547	1114	Intermediate
D12	1.58	1.89	1895	1185	Relaxed
D13	1.06	1.66	1658	1008	Disturbed
D14	1.43	1.83	1832	1372	Intermediate
D15	1.36	1.80	1803	1290	Intermediate
D16	2.74	2.28	2276	2013	Disturbed
D17	1.43	1.84	1834	972	Intermediate
D18 ¹	0.85	1.54	1542	1056	Intermediate
D19	1.14	1.70	1703	1200	Intermediate
D20	1.43	1.83	1833	1298	Intermediate
D21	1.18	1.72	1722	1174	Relaxed
D22 ¹	1.56	1.89	1887	1919	Relaxed
D23	1.06	1.66	1657	1030	Disturbed
D24	1.09	1.67	1675	1433	Intermediate
D25	0.79	1.51	1507	719	Disturbed
D26	1.26	1.76	1757	1255	Intermediate
D27	1.33	1.79	1789	1410	Relaxed
D28	1.55	1.88	1881	1457	Intermediate
D29	1.24	1.75	1749	1049	Disturbed

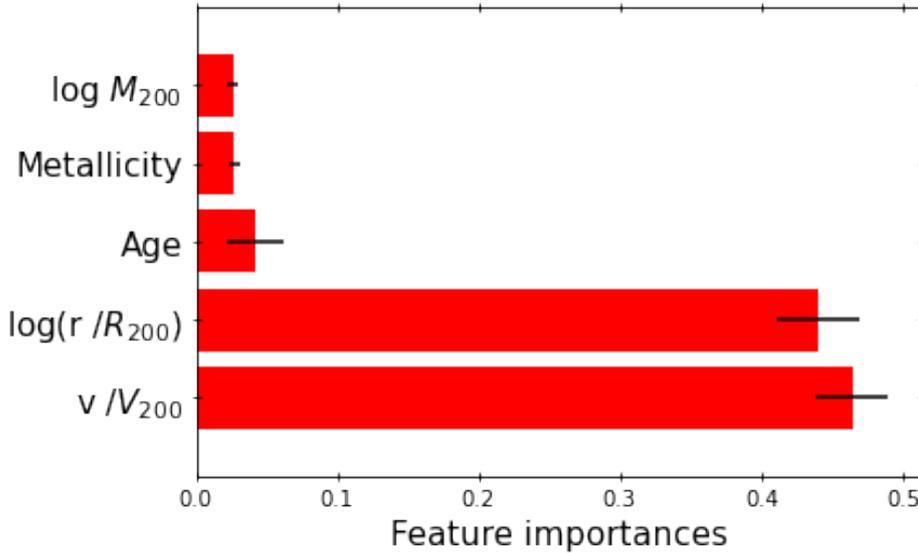


Figure 7.2: Importance of the input features including all initial input features to predict ICL and BCG components. From top to bottom: the logarithm of the cluster mass M_{200} , the stellar metallicity, age, the logarithmic cluster distance over R_{200} , and the rest-frame velocities scaled by the virial circular velocity V_{200} . The black bars encode the standard deviation when sampling the importance from the trees.

the kinetic and potential energies directly correlate with velocity and distance, we only keep the latter two. Later, we selected different subsets of these properties to assess which combinations of them provide results which are consistent with the all-features case and its metric scores. Having established that once we include distance and velocity in the features space the performance would not further significantly improve by adding other features, we decided to keep a basic parameter space and excluded such additional features. Furthermore, since our training set is composed of subsamples from different halos, we perform a scaling of each of these quantities according to the cluster virial values. We normalise clustercentric distances by R_{200} and we take its logarithm to increase the separation between the two stellar components in the parameter space. Instead, the rest frame velocity is scaled by the circular virial velocity V_{200} , which reads

$$V_{200} = \sqrt{\frac{GM_{200}}{R_{200}}}. \quad (49)$$

Assuming that the particle distributions of distance and velocity in a cluster may not be fully generalised by only scaling for the corresponding virial quantities in our final set, we provide the cluster mass M_{200} as an additional input feature to the reference model. This supplementary information may help the classifier in choosing a label over another in case of the degeneracy of the other two features if the label is somewhat still dependent on the cluster size.

Figure 7.2 reports in increasing order the importance of all the features initially included in the trained ML model and the associated standard deviations with the black bars. Although our choices were not driven by the feature importance analysis, we provide the results for our final set to point out an important aspect in the inclusion of the cluster mass as an additional feature to the parameter space. The plot shows that distance and velocity are of similar importance, whereas the other features have seemingly lower values. This is not surprising since these two features are also those that ICL-Subfind uses to perform the classification. On the other hand, the mass of the cluster M_{200} yields the lowest value among all features, provided that the training set is composed of only five clusters (and thus only five different input cluster masses (namely $M_{200} = (0.49, 0.10, 1.04, 0.85, 1.56) \times 10^{15} M_{\odot}$), whereas the other features vary from particle to particle. However, we expect that this feature could play a more important role in larger cluster sets, which is the reason why we do not exclude it from our analysis.

7.3.3 Classification performance

To tune the hyperparameters of the classifier, we use a K-fold crossing validation, with $K=5$. This involves randomly splitting the training set into K complementary subsets and repeatedly training the model on $K-1$ subsets while validating the resultant estimator when applied to the remaining subset. Each time, the classifier is trained on different combinations of the hyperparameters to obtain unbiased estimates of the classifier's average performance metrics and their uncertainty. The main hyperparameters undergoing this search are the number of features to consider when splitting a tree, the depth of the trees, and the number of trees in the Random Forest ensemble.

Results from each of these cross-validated runs are analysed with performance metrics. Notice that the estimated ICL labels are (in most of the clusters) much larger than the BCG component. For our binary classification problem, we use P, R, and FS which are independent of the imbalance nature of the classification problem. We defined them in Section 4.2.2. We point out that in the text we will refer to "true" as the labels provided by ICL-Subfind. Clearly, they are not necessarily "true" in absolute terms, but they represent our reference answer to this classification problem. In fact, there might be cases where the metric score is lowered due to a difference in the labels between the two methods, rather than due to a poor recognition of the ML classifier in the dynamical properties of stars.

7.4 RESULTS

Based on the cross-validated hyperparameter search, we find the algorithm to have a consistently good performance. Each class holds on this picture: the ICL shows $P = 95$ per cent, $R = 92$ per cent, and $FS = 93$ per cent, while the BCG class presents $P = 78$ per cent, $R = 85$ per cent and $FS = 81$ per

cent. We remind that these scores are valid for the specific subhalo finder used, Subfind, and larger differences could be found when employing other algorithms.

Besides its high accuracy, one of the benefits of employing the ML algorithm to classify star particles is its efficiency and speed up with respect to ICL-Subfind. Provided that the latter not only performs the star particle classification but also identifies the substructures in the FoF catalogue, in cases where one already has the subhalo identification for a given cluster (a standard procedure in state-of-the-art simulations of galaxy clusters to analyse galaxy populations), it is possible to bypass this step and directly obtain the labels for the stars in the main halo. Skipping this unnecessary operation can be crucial in saving run-time for large simulations whereby the subhalo identification can take several hours on different cores. For this reason, it is not straightforward to fully quantify the computational advantage of employing one technique over the other, unless one only needs the star particle classification having done the subhalo identification in previous steps. Taking this into account, a rough estimate of the run time of ICL-Subfind restricted to the sole stellar classification for a cluster at our reference resolution (considering the operational time spent by the traditional Subfind to detect subhalos) gives a speed-up by a factor of about 100. On the other hand, considering both procedures, the savings in time add up to an order of 10^5 . Furthermore, increasing the numerical resolution (and therefore the number of particles in a simulation) might entail a severe increase in the run-time, whereas no significant difference involves the ML classifier. In conclusion, we recommend the use of the ML model in cases where simulations have already undergone a subhalo identification procedure. However, further examinations are required to assess possible differences when other halo finders are used since we only examined the results by Subfind.

In the next section, we would like to quantitatively assess the resolving power of the model compared to the traditional ICL-Subfind. With this in mind, we select a random cluster in our simulation (its properties are summarised in [Table 7.1](#) under the name D5) and examine the differences between ML and ICL-Subfind in the feature distributions, phase-space profiles, and mass-weighted maps. The last part of this section will provide a more general overview of the results for the entire cluster data set.

7.4.1 *Testing on a single simulated cluster*

The first comparison is between the input features of the star particles. [Figure 7.3](#) illustrates the comparison between the number counts of the particles' logarithmic clustercentric distance (left panel) and their velocity distribution (right panel), both normalised by their virial value. True and predicted labels are marked with a contouring line and area, respectively, for both ICL (red) and BCG (blue) stars. In the bottom panels, we show the percentage residuals between the true and predicted labels over the total number of star particles in each bin to estimate where the results are

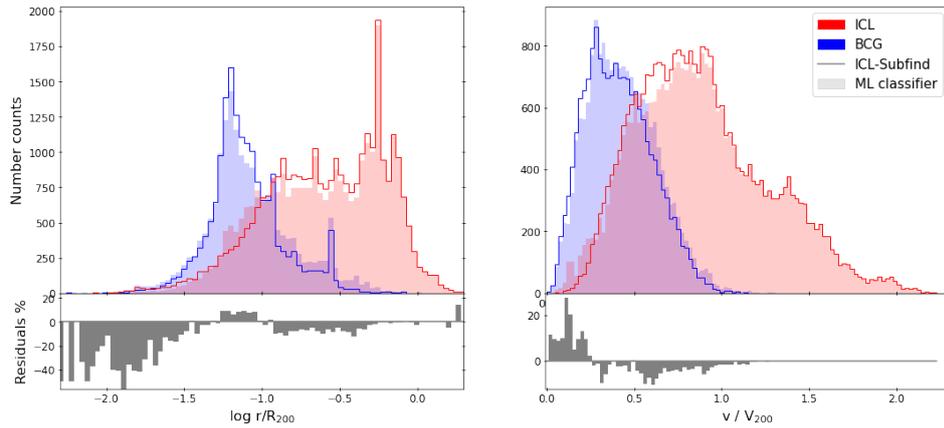


Figure 7.3: Top panel: distribution of the input features drawn from the simulated cluster. In each panel, we plot the predicted (bars) and true (line) number counts of both ICL (red) and BCG (blue) associated with the star particles. More in detail we have (from left to right) the distributions of the logarithmic clustercentric distance normalised for R_{200} and the stellar rest-frame velocities normalised for V_{200} . Bottom panel: percentage residuals measured between the true and predicted counts in each bin.

most different. This definition of residuals (in absolute value) is the same whether we consider ICL or BCG stars, given that it simply represents the excess of one class over the other, normalised by the number of particles in each bin. For this reason, we consistently choose throughout the chapter to represent the BCG excess (or deficiency, depending on the sign) of the ML prediction with respect to the model. The ICL percentage residuals will then simply correspond to the opposite number.

In both cases, we observe a generally good agreement within each predicted subgroup and its true distribution. The left panel confirms the presence of a bulk structure in the inner region, which corresponds to the BCG and the ICL, a more diffuse component that extends beyond R_{200} . The largest differences are found in the inner core of the BCG (up to 40 per cent), but they are mostly due to the low number of star particles in these bins. As we move towards the outskirts of the BCG, the distribution residuals span values around 10 per cent, which represents a more consistent estimate of the errors in the classification process at these distances. We usually find this transition region to be carrying most of the uncertainty in the labelling of stars in all clusters, as will be illustrated in the next section. In this regard, we expect both algorithms to carry uncertainties, which will sum up at the expense of the ML metric scores. In other words, the low metric scores for the ML algorithm are due to different labelling with respect to ICL-Subfind, which in turn is not necessarily always correct. The classification process suffers from finding dynamically similar particles in the ICL and BCG components that populate regions far from the centre, thus decreasing the precision of the algorithm at these distances.

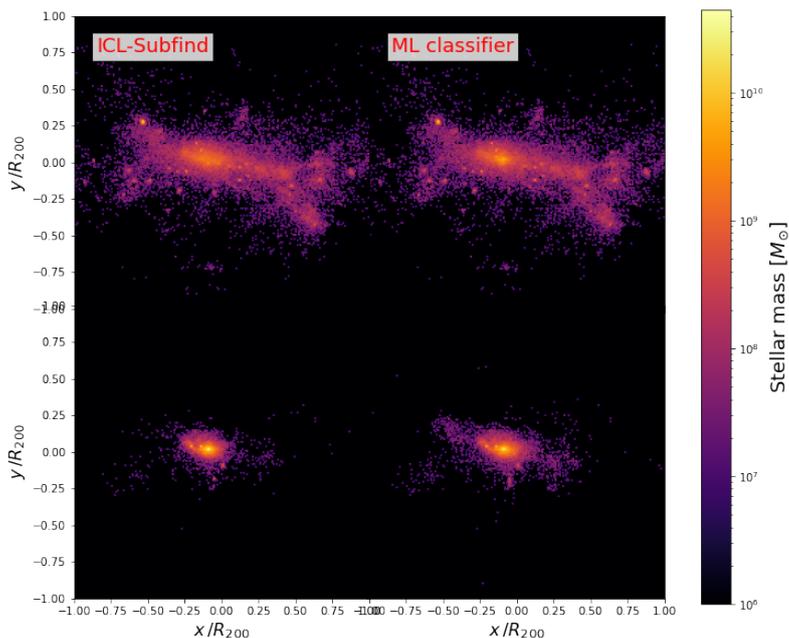


Figure 7.4: Mass-weighted maps of the stellar components (ICL in the top panels, BCG in the bottom ones) in the same cluster as in Figure 7.3. Left panels report the results from the stellar division provided by ICL-Subfind. Right panels show the mass-weighted maps for the stellar components identified with the predicted labels.

As for the velocity distribution, we confirm the presence of the two peaks which can be fit by the double Maxwellian. On this point, we highlight the closeness of the velocity distributions among the two methods, which is already a good index of the accuracy of the ML algorithm, given that this result is obtained without the need for an explicit fit. Figure 7.4 shows the mass-weighted maps of the distinct stellar components in the cluster under study. ICL in the top panels and BCG in the bottom. The panels on the left are the results from the traditional method, while on the right we illustrate the maps when employing the labels from the ML algorithm. The results are remarkably similar and the differences are mainly explained on the BCG outskirts. This can be better appreciated by comparing the 3D density profiles of both ICL (red) and BCG (blue) as computed with the true labelled stars (solid line) or with the predicted labels (dashed line) in Figure 7.5.

Another aspect to consider in evaluating the performance of the algorithm is to study the ICL fraction predicted by the ICL-Subfind and that from the ML scheme. We define the ICL fraction f_{ICL} as the number of ICL particles over the total number of stars in the main halo. The ML model yields values that are also consistent with those predicted by the traditional algorithm: in this particular cluster, we measure $f_{\text{ICL}} = 0.63$

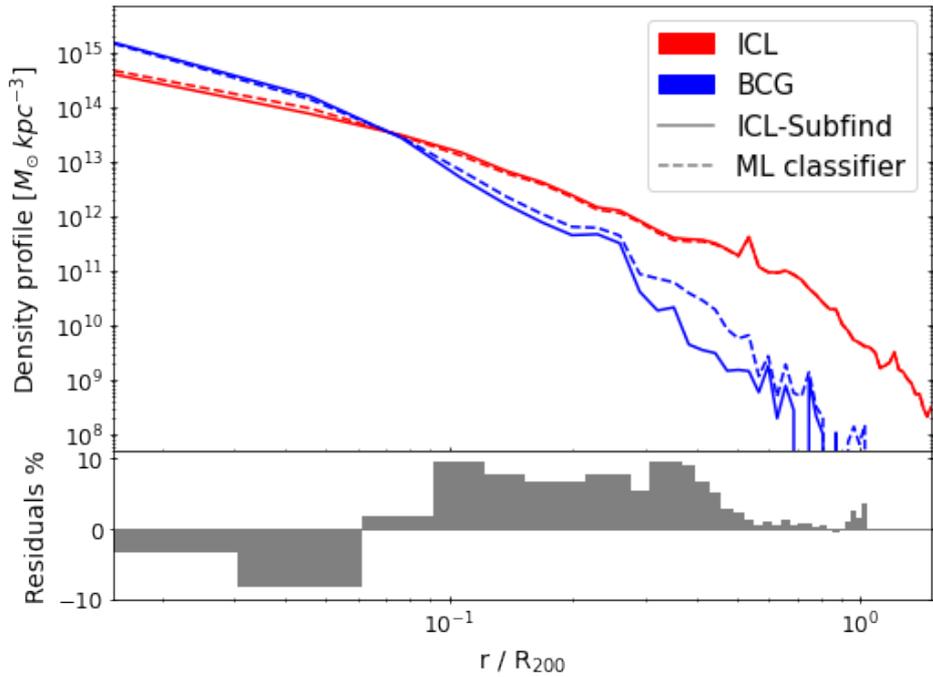


Figure 7.5: Top: density profiles of the BCG (blue) and ICL (red) in the selected cluster. The dashed lines are the profiles computed with the predicted labels, while solid lines report the profiles computed with the true labels. Bottom: residuals (in per cent) between the histograms of the labels from the ICL-Subfind and the ML classifier.

using the ML method as opposed to $f_{\text{ICL}} = 0.62$ for the traditional case when considering the ICL fraction over the stars of the main halo. We recall that this value shall not be directly compared to the observational results.

7.4.2 Testing on a simulated cluster population

In the previous section, we showed that our ML-based algorithm to separate stellar ICL and BCG populations is a robust classifier in the case of a single test cluster. We can take a step further and apply the trained classifier over all clusters not part of the training set (24 in our simulations) to present a few results which are worth discussing.

Figure 7.6 shows the comparison between the velocity dispersions derived from the fit of the double Maxwellian in Equation 24 to the stellar velocity distribution as labelled by the ICL-Subfind and ML classifications in all our clusters. In the left panel, we plot the ICL velocity dispersions; in the right panel, we present the BCG velocity dispersion. Each point marks a single cluster colour-coded for the ICL fraction as given by the ICL-Subfind sample. We report all 29 clusters, including those not converged in the ICL-Subfind (the three isolated points with low ICL fractions), while we mark the training clusters with a cross for clarity. The dashed grey line marks the 1 : 1 relation for reference. Quite remarkably the rela-

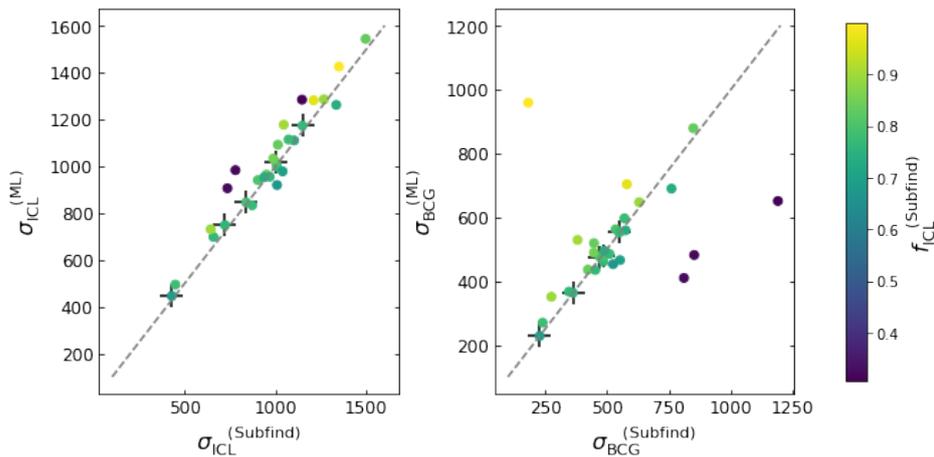


Figure 7.6: Comparison of the velocity dispersions derived from the fit of the double Maxwellian in Equation 24 in the stellar velocity distribution labelled by the ML classifier and the ICL-Subfind. Each point represents the result from the component of a single cluster, coloured according to the ICL mass fraction estimated in each cluster by the ICL-Subfind. Furthermore, we mark the clusters from the training set with a cross. The dashed grey line is the reference line to a 1 : 1 relation. Left panels: the velocity dispersions in the ICL component. Right panels: the velocity dispersions in the BCG component.

relationship between ICL velocity dispersion from ICL-Subfind and from the ML classifier (left panel of Figure 7.6) shows a small scatter around this relation, with the latter being on average 4 per cent higher. This difference increases slightly (10 – 20 per cent) in correspondence of the three groups with the lowest ICL fractions. On the contrary, for the BCG velocity dispersions (right panel of Figure 7.6), we observe a significant colour gradient, perpendicular to the reference line. Clusters hosting a larger BCG stellar fraction and a correspondingly lower ICL fraction, assigned by ICL-Subfind, have a higher BCG velocity dispersion, and vice versa. This large difference between the two methods can be traced back to the extreme values of the ICL fraction in the ICL-Subfind predictions, which do not occur in the ML case (specifically stretching for all clusters only within the range 0.60 – 0.80). Understandably, the larger/smaller the virialised system (BCG, in this case), the higher/lower the velocity dispersion.

Table 7.3 summarises several of these results for all clusters: we compare the fits of the double Maxwellian, the ICL fractions for both the ICL-Subfind and ML algorithms and the performance scores. The latter are reported for both the single classes and the means weighted with the number of the two components. As previously mentioned, the scores employed to assess the quality of the predictions by the ML classifier with respect to ICL-Subfind are P, R, and FS. We find that BCGs usually have high P and lower R scores, which expresses the capability of the ML algorithm to be generally correct when labelling BCG stars, although not returning the entire set of BCG particles compared to the true set. On the contrary, ICL has usually most of the particles assigned, thus yielding the

opposite situation. However, FS is high in most cases (a few pathological cases will be discussed in the next paragraph). A more explicit report of the classification score is provided by the corresponding weighted means in the last columns. Relaxed clusters, which typically have a well defined double Maxwellian velocity distribution, reach a mean FS of 80 – 90 per cent. This is the case for 15 clusters in the test set. On the other hand, we observe a few clusters receiving consistently low scores in the classifier’s metrics (i.e., FS is below 40 per cent) while their physical features are generally inconsistent with the results obtained by ICL-Subfind. This is the case for D7, D11, and D13 for which, as mentioned before, ICL-Subfind did not reach convergence. For these clusters, the ICL fraction is a few per cent and, indeed, there is no separation of the components. On the contrary, we believe that our ML-based method overcomes these situations by correctly identifying two dynamically distinct components, with the BCG component having a markedly smaller velocity dispersion than the ICL one. Quite interestingly we notice that a few central galaxies in our sample show tidal shell feature (see Ebrova, 2013, for a review) once we separate the BCG and ICL. This is the case for five of our clusters, two of which are part of the clusters that have not converged in the ICL-Subfind procedure. These shell-like features in the stellar distribution could be linked to past tidal shocks, associated with recent merger events. We briefly discuss this observed feature from Figure 7.14 in Section 7.4.5.

A common trait in the testing set is the discrepancy between predicted and true classes on the outskirts of the BCGs, where the distinct dynamical behaviours of the star particles are generally harder to discern. Figure 7.7 investigates this flaw in the performance by showing the BCG metric scores as a function of the radial distance from the centre of the subhalo. We stacked the metric score profiles for the BCG labels (P in red, R in blue, and FS in brown) of the entire cluster set to tentatively describe the expected accuracy. Shaded bands display the standard deviation given by the intrinsic distribution, while the dashed grey line marks the value 0.5 on the y-axis, below which the rate of incorrect labelling is more than one in two in the (predicted or true) BCG set. We notice that the metric score in the centre is high (> 0.7), while from $\sim 0.1R_{200}$ (~ 250 kpc) it declines very rapidly. Here, far from the central region, we expect differences between the ICL and BCG properties to become less sharp, since BCG particles will have larger entropies compared to the centre, spanning a phase space very similar to that occupied by the ICL. Furthermore, we evaluated the ML model metric scores for the fairly resolved sample of clusters (that is, excluding D7, D11, and D13) as a function of the mass of the cluster (left panels) and the dynamical state (right panels) in Figure 7.8. Dolag et al. (2010) refer to potential uncertainties in the classification process for low-mass clusters owing to the difficulty in disentangling BCG and ICL in the velocity distributions of low-mass halos, where the two Maxwellians cannot be easily discerned. On the other hand, the dynamical state and recent merger history can strongly impact the physical conditions of the stellar components in the inner regions. In the

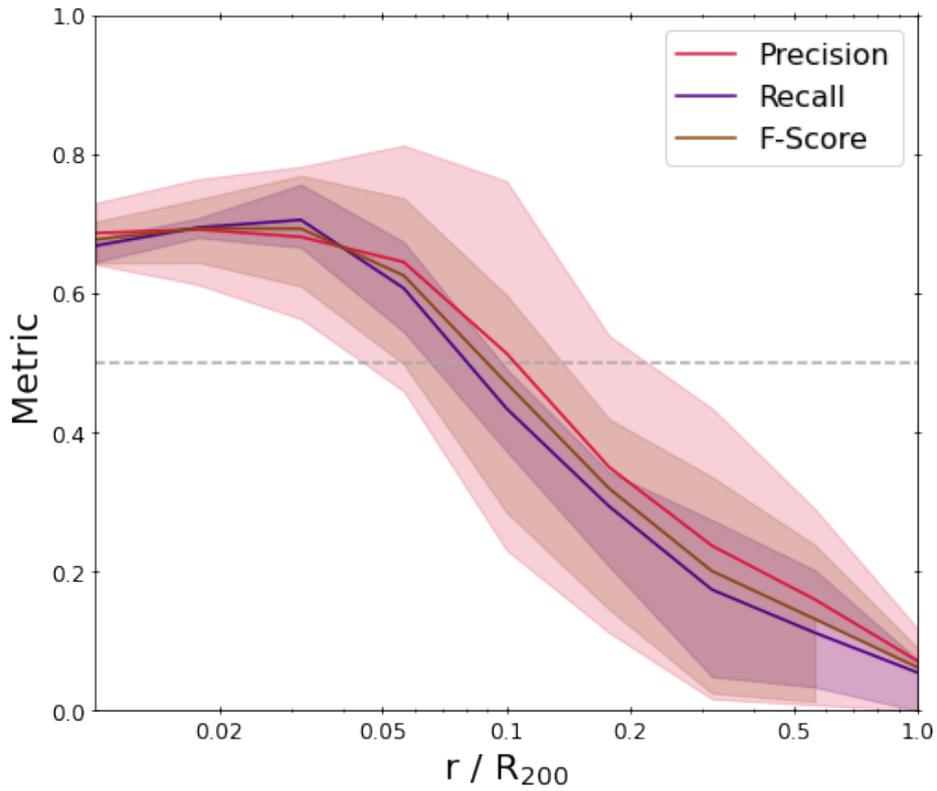


Figure 7.7: Stack of the metric score profiles relative to the BCG particles classification in the entire cluster set. The profiles are reported as a function of the cluster distance normalised by R_{200} . Solid lines refer to P (red), R (blue) and FS (brown). The dashed line marks the value 0.5 on the y-axis.

analysed simulated clusters, these merging events of massive orbiting halos with the BCG produce strangely-shaped halos with non-thermalised velocity distributions, or strongly non-spherical symmetry in the star particle distribution, which can be due either to a peculiar halo formation history or to an incorrect group identification by Subfind in the main halo. In this case, both ICL-Subfind and the ML algorithm may encounter difficulties in properly separating the two components: size estimates of the BCG are extremely sensitive to these non-thermalised distributions of star particles. However, [Figure 7.8](#) shows no significant correlation with either of the cluster properties. We plot from top to bottom the P, R, and FS for the clusters in both training (empty squares) and testing sets (filled dots). The latter are colour-coded according to the ICL fraction computed with the labels from ICL-Subfind. There seems to be a mild correlation between P and f_{ICL} , and thus in the FS, however, this shall be verified with a larger sample of clusters.

7.4.3 *Testing the robustness of the classifier*

So far, both training and testing have been described for a given set of simulated galaxy clusters that, despite their specific history of formation, share many similarities: the same sub-resolution model for star formation and feedback, the same numerical resolution, and the same redshift. In the effort to understand the real range of possible applications of our classifier – compared to what is originally obtained with ICL-Subfind – we apply our model to other simulated clusters which differ from the original cluster set in different ways. We decided to re-simulate two out of the 29 clusters (i.e., D1 and D2) in different conditions and we discuss the outcomes of these analyses in the following sections. Despite the limited statistics, we expect to obtain useful insights into the predictive power of our method from these tests.

7.4.3.1 *Changing redshifts*

We analyse here the behaviour of the ML classifier in the same simulation at redshifts different from that of $z = 0$, at which the method has been trained. We point out that both traditional and ML methods rely on the underlying physical assumption that the two stellar components can be described by a double Maxwellian early enough to label the stars consistently as for $z = 0$. This is not necessarily true if the stellar populations are still forming or evolving significantly. For this reason, we analyse our simulated clusters at two different redshifts ($z \simeq 0.5$ and $z \simeq 1$) at which most of the BCG stellar mass is already in place (Ragone-Figueroa et al., 2018). [Figure 7.9](#) illustrates the evolution of the stellar density profiles of D2 (from top to bottom panel: the redshifts are $z \simeq 1$, $z \simeq 0.5$, and $z = 0$). The legend is as before colour-coded for the stellar type (ICL in red and BCG in blue), ML results are described by a dashed line, while solid marks the ICL-Subfind output. Results and performance scores are

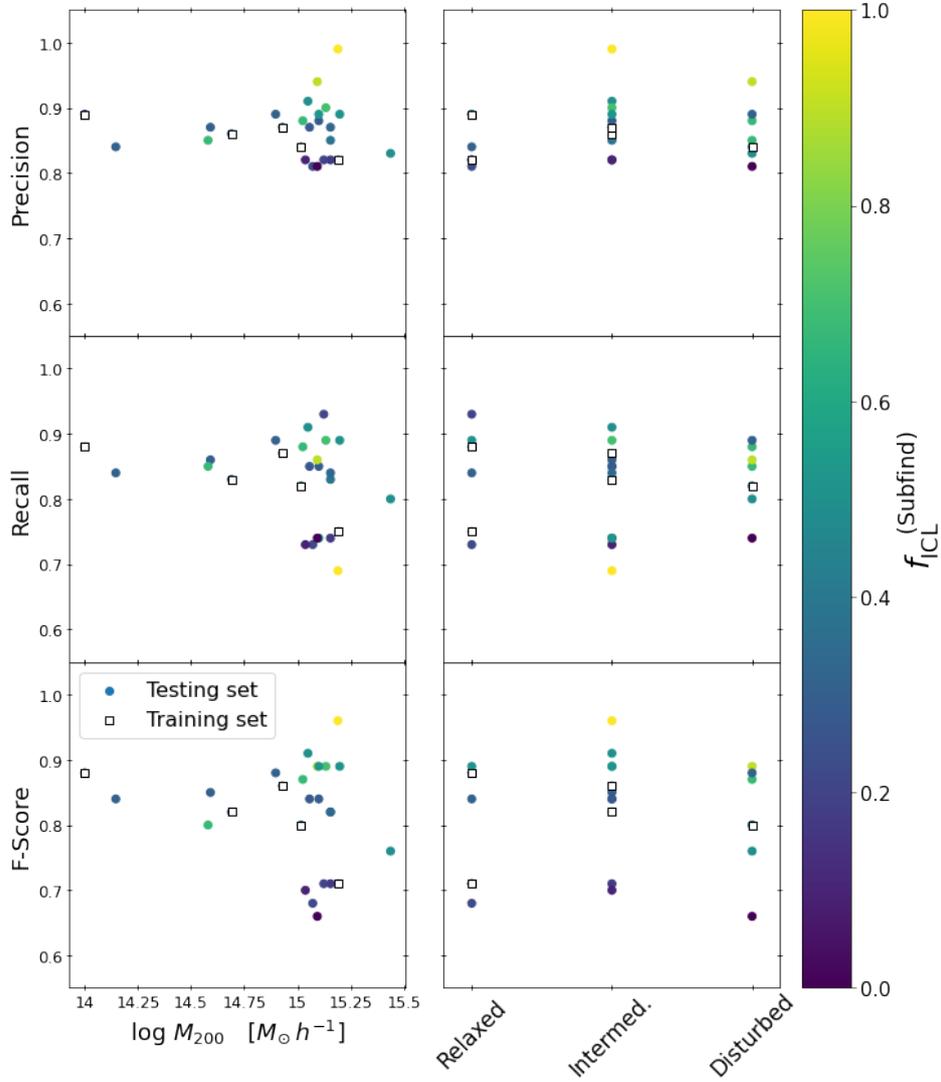


Figure 7.8: Scatter plots of the (weighted) mean metric scores (P, R, and FS from top to bottom panel) as a function of log cluster mass (left panels) and dynamical state (right panels) for all the clusters under study. We mark the pure testing set with coloured points, while the clusters in the training set are recorded with empty squares. The colour legend follows the ICL fraction estimated by ICL-Subfind. The y-axis is limited to exclude D7, D11 and D13 having very low metric scores.

largely consistent with what we found in the case of $z = 0$: no significant systematics can be detected between the ML classifier and the traditional labelling and overall the distributions are recovered. Unsurprisingly, these are slightly better in the case of $z = 0.5$, rather than in the case of a higher redshift.

7.4.3.2 *Changing numerical resolution*

An important step in understanding the quality of our predictions is to estimate the effect of the numerical resolution. This can be performed by examining the results of our classifier, trained on a cluster set at a given resolution when applied to a set at a higher resolution. Increasing the resolution in a simulation improves the description of lower-mass systems and small-scale features. In turn, this could affect the probability distribution function at the centre of the clusters. In our resolution tests, we decreased the particle mass by a factor of three with respect to the reference simulation set, yielding $m_{\text{DM}} = 2.5 \times 10^8 h^{-1} M_{\odot}$ and the initial mass of the gas particle $m_{\text{gas}} = 1.1 \times 10^8 h^{-1} M_{\odot}$ for two clusters. The performance scores are found to be quite high (e.g., $P > 0.75$, $R > 0.80$, $\text{FS} > 0.78$), with the stellar density profiles from ICL-Subfind and our method agreeing to per-cent level, as shown in [Figure 7.10](#). Further tests were performed for simulations at even higher spatial resolutions (also increasing our fiducial softening lengths three times, as in [Bassini et al. 2020](#)) giving similar high-performance scores, but they are not shown here. Therefore, our ML classifier seems to be robust when applied to simulations whose resolution is higher than that of the training set.

7.4.3.3 *Changing the feedback model*

To further test the robustness of our ML classifier, we applied it to simulations having the same resolution of the training set, but not including AGN feedback. Obviously, this is an extreme (possibly non-physical) scenario held with the only illustrative purpose of examining the consequences on the classifier performance facing underlying different physical conditions with respect to the training set. AGN feedback regulates star formation in massive galaxies, particularly impacting BCG masses (e.g. [Ragone-Figueroa et al., 2013](#)), thus this test allows us to examine the ML robustness in a conservative regime of exceedingly massive galaxies. Our analysis shows the presence of a massive BCG at the centre of the halo in both ICL-Subfind (solid) and ML (dashed line), as can be seen in [Figure 7.11](#). We notice that this causes a steepening in the density profiles with respect to what was predicted in presence of the AGN feedback mechanism (dotted lines), thus increasing the concentration of the stellar halo. In the latter, the differences in the density profiles account for up to only 3 per cent between the labels of the ML classifier and ICL-Subfind, so we decided to plot only one of them. Quite interestingly, we find that the ICL-Subfind and ML model-predicted profiles match better in the innermost part, where we would expect most of the AGN feedback to have

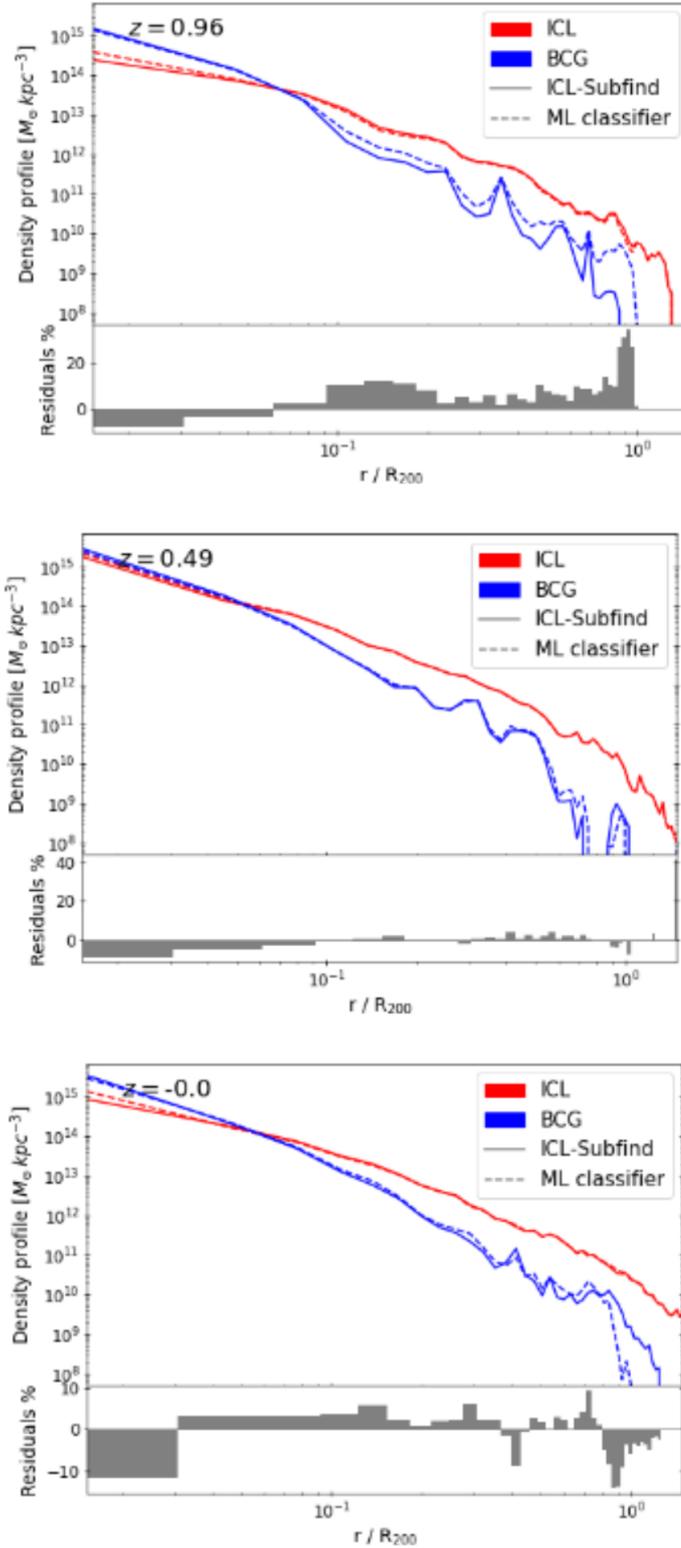


Figure 7.9: Density profiles of the stellar content in D2 taken at three different snapshots: from top to bottom the corresponding redshifts are $z \simeq 1, 0.5, 0$. For each of these plots, we show two panels. In the top panel, we report the profiles normalised at R_{200} of both ICL (red) and BCG (blue) identified by ICL-Subfind (solid line) and the ML classifier (dashed line). The lower panel shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels.

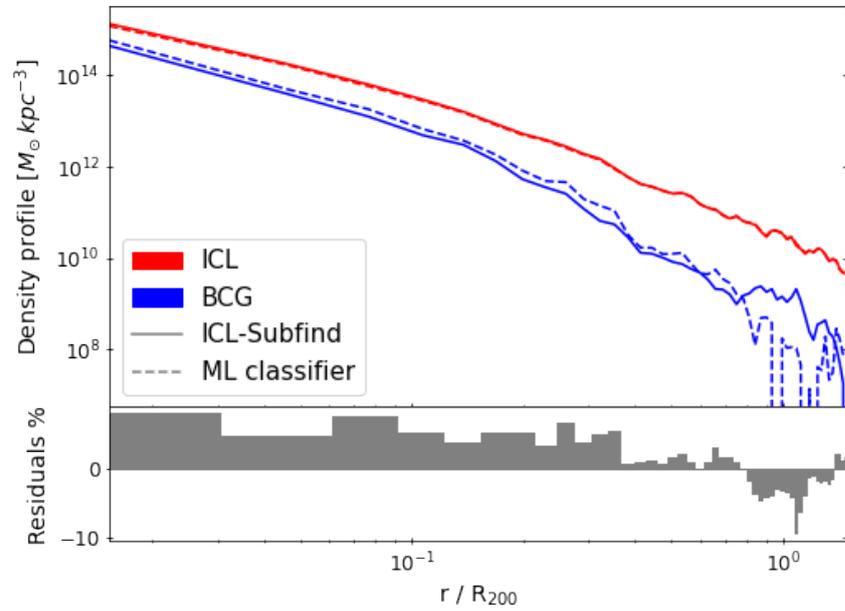


Figure 7.10: Density profiles of the stellar content in D2 simulated at higher resolution. In the top panel, we report the profiles normalised at R_{200} of both ICL (red) and BCG (blue) identified by ICL-Subfind (solid line) and the ML classifier (dashed line). The lower panel shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels. Differences account for a per cent level only between the profiles.

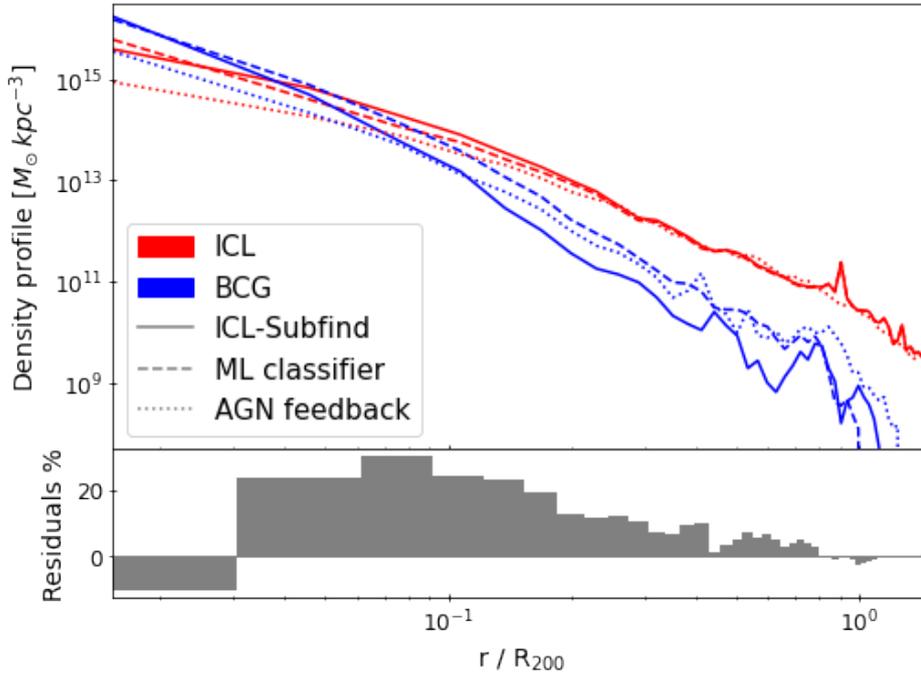


Figure 7.11: Density profiles of the stellar content in D2 simulated with no AGN feedback scheme. In the top panel, we report the profiles normalised at R_{200} of both ICL (red) and BCG (blue) identified by ICL-Subfind (solid line) and the ML classifier (dashed line). Additionally, we overplot the BCG and ICL profiles from the same cluster when including the AGN feedback model (dotted line). The lower panel shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels.

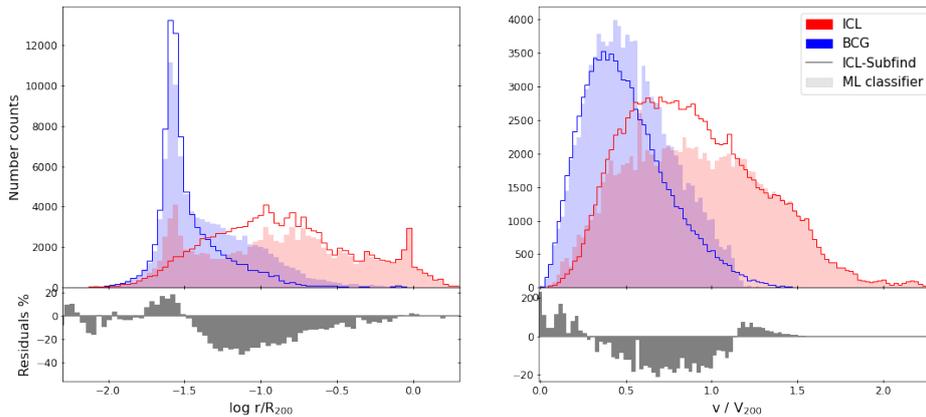


Figure 7.12: Histograms of the distributions of the features of D2 simulated with no AGN feedback scheme. In the top row, we report the number count distributions of (from left to right): the logarithmic cluster-centric distance over R_{200} and the particle rest-frame velocity over V_{200} . The lower row shows the residuals (in per cent, normalised by the number of star particles in each bin) between the BCG true and predicted labels.

a substantial effect, rather than in the BCG outskirts: here, the differences reach up to 25 – 30 per cent between the two runs. A closer look at the feature distributions in [Figure 7.12](#) shows that the phase space also shows some inconsistencies between the two methods. Yet, we can fit a double Maxwellian distribution to the star particle velocities reasonably close to the original one.

Given the good quality of our analysis to this point, we pushed our investigation to simulations when both effects (numerical resolution and subgrid physics) are different relative to the reference simulation set. Although we do not display any of the profiles, we can confirm that the performance scores for this case are in line with the previous ones, demonstrating that the ML classifier results are robust across small changes in numerical resolution, redshift (at least in the late Universe, within $z \leq 1$) and physical subgrid models.

7.4.4 Phase-space structure

To further investigate the accuracy and robustness of these methods, we examine several quantities that generally describe the phase-space structure of galaxy clusters (e.g. Marini et al., 2021a), in particular discussing it in terms of the stellar density profile $\rho(r)$, stellar velocity dispersion profile $\sigma(r)$ and phase-space density profile $Q(r) = \rho(r)/\sigma^3(r)$. We expect these quantities to provide insights into the robustness of these two methods based on the star particle distribution and dynamical information. Our ultimate goal is to detect major and/or systematic differences within the ICL and BCG subgroups as given by the two methods.

Our main findings are illustrated in [Figure 7.13](#). From the top to the bottom panel, we report the density, velocity dispersion, and phase-space density profiles of the star particles labelled as BCG (blue) and ICL (red) by the ICL-Subfind (solid lines) or the ML classifier (dashed lines). The radial distance is scaled by the virial radius R_{200} , to properly account for the different cluster sizes when stacking. The shaded bands represent the intrinsic scatter within the sample of clusters, computed as the standard deviation.

The BCG and ICL density profiles predicted by the two methods do not show significant differences. Therefore, we can provide an estimate of the transition radius, defined as the clustercentric distance at which the ICL distribution starts dominating the stellar component. Given our different approach based on dynamical criteria rather than from a fitted profile, we are able to provide an independent comparison with the values proposed in the literature. The cluster set yields an average transition radius of about 90 kpc (corresponding to $0.04 R_{200}$) which is in agreement with both theoretical and observational findings (e.g. Gonzalez et al., 2021; Contini et al., 2022). On the other hand, in the velocity dispersion profiles, we observe a systematic difference in the profiles for large radii, even beyond the expected transition radius. The ML classifier tends to prefer a dynamically hotter BCG component compared to the output of

ICL-Subfind. This is particularly highlighted in the central panel, where we compare the velocity dispersion profiles $\sigma(r)$ scaled by the velocity dispersion of the stars within the virial radius σ_{200}^* to correctly stack the distinct clusters. Velocity dispersions traced by the ICL stars are generally higher than those of the BCG stars at all radii, consistently with the results shown in [Section 7.2](#) on the velocity distributions: we find that ICL profiles are consistent within 1σ in the two methods. Conversely, BCG profiles are similar at the centre, while at large radii the ML classifier tends to include particles in the BCG with higher velocity dispersion than in the ICL-Subfind case. This results in an almost flat velocity dispersion profile. Indeed, these differences are present at large radii, where the assignment of star particles to one of the two components is less obvious, as seen in [Figure 7.7](#).

An extra step can be taken by evaluating the phase-space density which combines the density and velocity dispersion profiles to investigate the phase-space structure of halos. Both numerical and observational (just to name a few Taylor and Navarro, 2001; Dehnen and McLaughlin, 2005; Faltenbacher et al., 2005; Biviano et al., 2013, 2016; Marini et al., 2021a) studies have demonstrated that the profiles of phase-space density (or equivalently, of the pseudo-entropy $S(r) = Q(r)^{-2/3}$) have a power-law dependence on the clustercentric radius, with a rather small scatter. In this context, Marini et al. (2021a) investigated the pseudo-entropy profile of different tracers in a set of simulated clusters, including the star particles as tracers of the phase-space structure of the cluster, and demonstrated that, while BCG and ICL components separately do not produce accurate power laws for the phase-space density, the power-law profile is instead recovered when analysing together the star particles of such two components. Once again, we see that the largest differences are in the BCG outskirts.

7.4.5 *Hints for the identification of shell galaxies*

Among the test cluster set, we observe a handful of clusters that present particularly interesting features in their spatial distribution. One of the most striking cases is reported on the mass-weighted map in [Figure 7.14](#). We point out that this cluster has not reached convergence in the ICL-Subfind algorithm, therefore our next discussion will be mostly addressed to the ML output. In the central panel, we display both components (BCG+ICL), the central one illustrates the ICL population, while the right plot is for the BCG stars. The plot shows a complex stellar structure composed of spherical shells surrounding the central stellar peak in the BCG, which is not seen in the ICL-Subfind case. We assume this feature to not be caused by some numerical artefact, given that several clusters have this same symmetrical distribution in the ICL-Subfind analysis. This shell-like distribution could be due to the expansion and later disruption of stars occurring during a tidal shock, which gives origin to shell galaxies. The shells are formed as density waves induced in a thick disc population of dynamically cold stars by a weak interaction with another galaxy

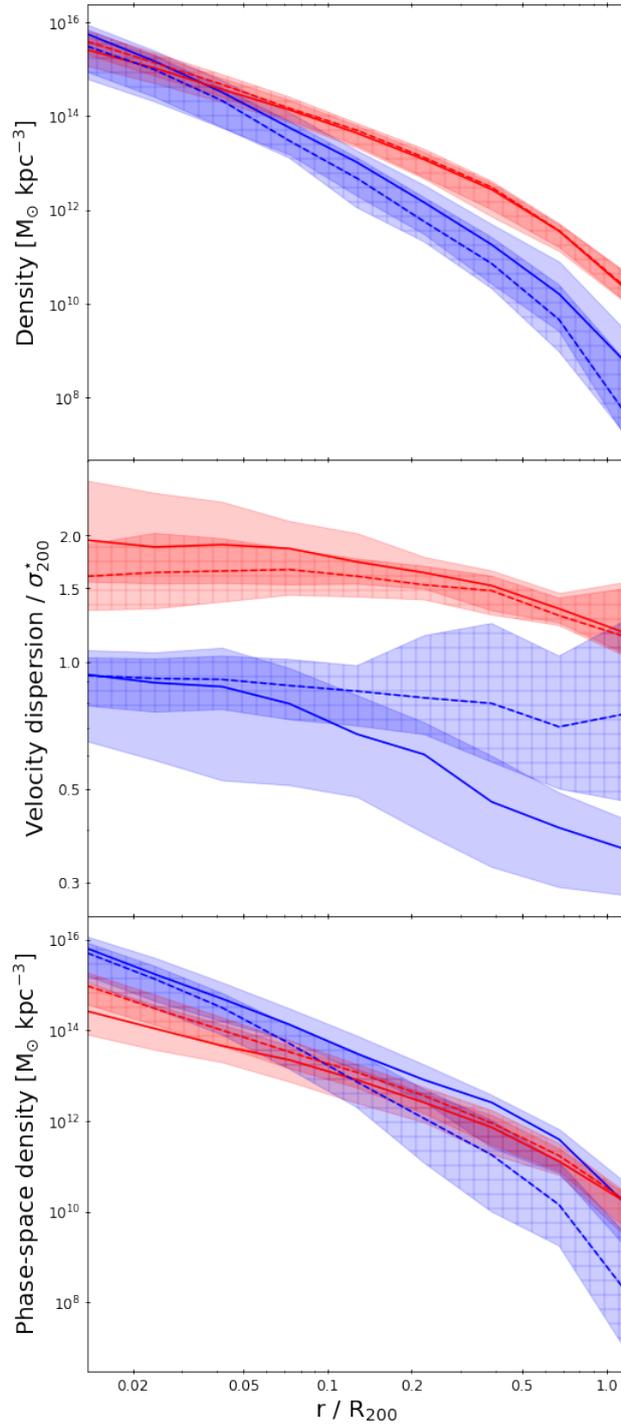


Figure 7.13: Density (top), velocity dispersion normalised for the velocity dispersion of the stars within R_{200} (central), and phase-space density (bottom) profiles of the star particles in the BCG (blue) and ICL (red). Results from the ICL-Subfind labels are given with solid lines, on the other hand, the dashed lines mark the profiles extracted from the ML classified labels. In each panel, we report the median profile of each method (dark solid line) and the uncertainty given by the intrinsic standard deviation with the shaded coloured band.

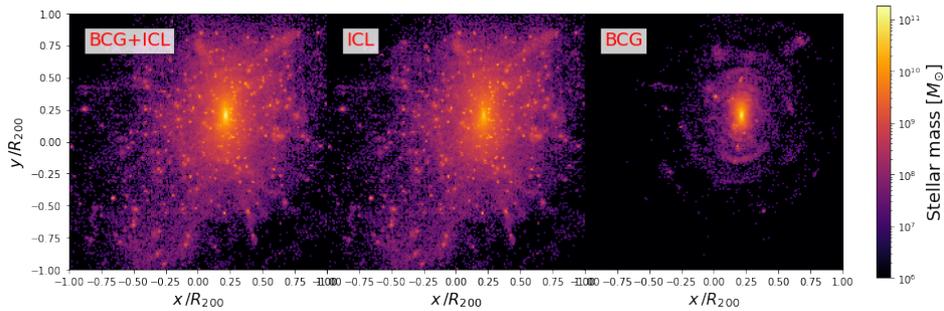


Figure 7.14: Mass-weighted map of the stellar components in cluster D7: BCG+ICL (in the left panel), ICL (central) and BCG (left) spatial distributions. We only show the results from the predicted labels, since the ICL-Subfind does not converge for this case.

(Thomson, 1991) and relatively major mergers (e.g., with a mass ratio of 1:10, see Pop et al., 2018). It is not clear whether the complexity of the shell structure may be responsible for the difficulty of ICL-Subfind converging, but it is striking to notice that the ML algorithm can detect it without specific training on this particular feature. Additionally, it is also able to spot it whenever the ICL-Subfind does. We will defer a complete analysis of this hypothesis to future work.

7.5 CONCLUSIONS

We presented a robust and efficient method to label stars in the main halo of simulated galaxy clusters as ICL or bound to the BCG based on a Random Forest classifier. The classification model is trained, cross-validated, and tested on 29 galaxy clusters simulated with cosmological hydrodynamical simulations, reaching a high level of precision. This ML method is based on a more traditional algorithm, which we call ICL-Subfind, fully described in Dolag et al. (2010). In that paper, the authors showed the existence of two dynamically distinct components in the stellar population of simulated galaxy clusters (associated with the main halo), which are identified because their velocity distributions can be fitted by a double Maxwellian distribution. Including this information in a gravitational unbinding procedure yields a spatial separation of the ICL and BCG stellar components in the central subhalo of simulated galaxy clusters. The subset of stars with the largest velocity dispersion is associated with the hottest stellar component, the ICL, while the other is assumed to be bound to the central galaxy or BCG. By applying the ICL-Subfind algorithm to the star particles in the 29 simulated clusters of the DIANOGA set, we obtain several data sets which we can use to fit a supervised model, intending to obtain consistent results with the traditional ICL-Subfind method, but far more efficiently.

To construct the classifier, we find the combination of input features that proves to best represent the two classes (labels) we are seeking, which for our specific problem are the cluster mass M_{200} , the clustercentric

distance of particles normalised by R_{200} , and the velocity of the rest frame of particles normalised to V_{200} . We use randomly selected subgroups of particles from five clusters to train and cross-validate the classifier, while the rest of the clusters are employed to further test the predicted generalisation of the algorithm.

Our results can be summarised as follows.

- Our classification method agrees to a high degree of precision with the true labels (i.e., ICL-Subfind) of the two stellar components in the cluster population. We find the existence of a central, more gravitationally bound, stellar bulk, the BCG, which is disentangled from the more diffuse ICL, that instead extends to larger distances. The fraction of ICL is also consistent with that found by ICL-Subfind and is generally higher (by a factor of about 3) than that associated with the BCG. Nevertheless, we stress that the ICL mass fraction found here shall not be regarded as immediately comparable to observations, where the separation between ICL and BCG is not performed in a dynamical analysis.
- We show that the metric scores relative to the BCG decrease steadily beyond $0.1R_{200}$, as shown in [Figure 7.7](#), in turn affecting the density and dynamical profiles. We shall recall that at these distances, both algorithms carry uncertainties to a certain extent in labelling star particles. In other words, a lower metric score at these scales may be also due to the composite effect of wrong labelling from both algorithms.
- Our model proved to be robust to changes in the numerical resolution, across different redshifts (up to $z = 1$) and with the exclusion of an AGN feedback model in the simulation. These results support the use of this method in various heterogeneous situations. It is unclear to what extent this model becomes unresponsive, and thus we conclude that further analyses are required for systems significantly different from those tested.
- Additionally, we examined the dependence of the performance of the ML algorithm on the mass of the cluster and the dynamical state. Both effects may play a role in changes in the phase-space structure of a cluster and, in turn, affect the performance. However, we did not detect any significant correlation with either of these properties.
- Finally, to gain more insight into the quality of the dynamical information retrieved by this method, we studied the phase-space structure of the two stellar components after applying both classifiers to the simulated cluster set in [Figure 7.13](#). We estimate the transition radius between the BCG-dominated and ICL-dominated regions at around $0.04 R_{200}$ (corresponding to a physical scale of 90 kpc), in line with previous observational measurements and theoretical predictions. This is particularly interesting in light of the different assumptions

used to determine its value. We identify the BCG outskirts to be the most critical region (i.e., beyond $0.1R_{200}$, corresponding to a physical distance of roughly 250 kpc) and far more prone to uncertainties in the classification process. This is due to the co-existence of the two stellar components whose physical properties overlap.

In conclusion, this method proved to be reliable and faster than the traditional method to identify ICL and BCG in the main halo of simulated galaxy clusters. Although it does not provide a new methodology for detecting ICL, it offers a robust tool to further investigate the dynamical characteristics of ICL compared to the traditional method. As a final remark, we shall refrain from claiming that this classifier will perform at this level of accuracy for simulations including significantly different astrophysical models unless these are included in the original training set. To remain a competitive alternative, when applied to very different simulations, one should resort to more advanced models, trained on vaster sets of simulations, which will need to include different cosmological and astrophysical scenarios for structure evolution. Ultimately, a dynamical analysis of the ICL should be regarded as an attempt to determine its physical properties and its origin to gain insights into the evolution of clusters and their stellar components.

Name	Conv.	$\sigma_{\text{BCG}}^{*(\text{S})}$ km s ⁻¹	$\sigma_{\text{ICL}}^{*(\text{S})}$ km s ⁻¹	$\sigma_{\text{BCG}}^{*(\text{ML})}$ km s ⁻¹	$\sigma_{\text{ICL}}^{*(\text{ML})}$ km s ⁻¹	$f_{\text{ICL}}^{(\text{S})}$	$f_{\text{ICL}}^{(\text{ML})}$	BCG			ICL			WMean		
								P	R	FS	P	R	FS	P	R	FS
D1	Yes	485	1008	461	997	0.62	0.77	0.99	0.61	0.75	0.81	1.00	0.89	0.88	0.85	0.84
D2	Yes	346	658	367	697	0.63	0.74	0.93	0.66	0.77	0.83	0.97	0.90	0.87	0.86	0.85
D3*	Yes	362	721	363	748	0.62	0.78	0.98	0.56	0.71	0.78	0.99	0.88	0.86	0.83	0.82
D4	Yes	275	643	351	730	0.82	0.97	0.86	0.15	0.26	0.85	0.99	0.92	0.85	0.85	0.80
D5	Yes	240	448	270	493	0.63	0.62	0.76	0.80	0.78	0.88	0.86	0.87	0.84	0.84	0.84
D6	Yes	423	949	436	964	0.73	0.80	0.94	0.70	0.80	0.90	0.98	0.94	0.91	0.91	0.91
D7	No	852	779	482	983	0.02	0.77	1.00	0.24	0.38	0.02	0.99	0.04	0.98	0.24	0.37
D8	Yes	580	1209	703	1281	0.94	0.83	0.26	0.78	0.39	0.98	0.87	0.92	0.94	0.86	0.89
D9 ¹	Yes	228	428	228	445	0.54	0.64	0.97	0.76	0.85	0.83	0.98	0.90	0.89	0.88	0.88
D10 ¹	Yes	484	1002	477	1015	0.67	0.83	0.94	0.49	0.64	0.80	0.98	0.88	0.84	0.82	0.80
D11	No	810	734	410	905	0.01	0.79	1.00	0.19	0.32	0.02	0.97	0.83	0.98	0.22	0.34
D12	Yes	447	1012	519	1091	0.73	0.77	0.84	0.75	0.79	0.91	0.95	0.93	0.89	0.89	0.89
D13	No	1190	1145	651	1283	0.02	0.92	1.00	0.08	0.15	0.02	0.99	0.04	0.98	0.10	0.14
D14	Yes	575	1101	559	1109	0.56	0.82	0.99	0.41	0.58	0.69	0.99	0.81	0.82	0.74	0.71
D15	Yes	381	1044	529	1176	0.84	0.79	0.61	0.78	0.69	0.96	0.90	0.93	0.90	0.89	0.89
D16	Yes	848	1495	878	1543	0.73	0.92	0.95	0.27	0.42	0.79	0.99	0.88	0.83	0.80	0.76
D17	Yes	535	902	562	940	0.67	0.81	0.91	0.69	0.78	0.81	0.97	0.89	0.85	0.83	0.82
D18 ¹	Yes	472	837	475	845	0.794	0.737	0.91	0.69	0.79	0.86	0.96	0.91	0.87	0.87	0.86

D19	Yes	453	967	435	954	0.765	0.767	0.99	0.61	0.75	0.80	1.00	0.77	0.87	0.85	0.84
D20	Yes	570	1071	596	1113	0.779	0.791	0.99	0.54	0.70	0.80	1.00	0.89	0.87	0.84	0.82
D21	Yes	760	1333	689	1261	0.738	0.817	1.00	0.39	0.57	0.70	1.00	0.83	0.81	0.73	0.68
D22 ¹	Yes	550	1147	552	1173	0.62	0.87	0.99	0.34	0.50	0.71	0.99	0.83	0.82	0.75	0.71
D23	Yes	630	1266	647	1285	0.82	0.91	0.86	0.73	0.79	0.95	0.98	0.96	0.88	0.88	0.87
D24	Yes	552	1038	466	977	0.52	0.80	1.00	0.42	0.59	0.65	1.00	0.79	0.82	0.73	0.70
D25	Yes	510	870	485	833	0.63	0.73	0.96	0.71	0.81	0.85	0.99	0.91	0.89	0.88	0.88
D26	Yes	447	986	489	1032	0.74	0.76	0.82	0.76	0.79	0.92	0.94	0.93	0.89	0.89	0.89
D27	Yes	489	937	495	952	0.57	0.84	0.98	0.40	0.57	0.70	0.99	0.82	0.82	0.74	0.71
D28	Yes	182	1348	958	1424	0.99	0.93	0.02	0.81	0.04	1.00	0.85	0.92	0.99	0.93	0.96
D29	Yes	524	1007	453	919	0.47	0.78	1.00	0.41	0.58	0.60	1.00	0.77	0.81	0.69	0.66

Table 7.3: Results from the best-fit procedure applied to the double Maxwellian in both the ICL-Subfind and ML algorithm case. We list the convergence report of ICL-Subfind in the second column. Then, we present the ICL and BCG velocity dispersions, the fraction of ICL, and in the last columns the algorithm performance scores (of the two classes and their weighted median), namely: the P, R, and FS.

¹ Training set

DETECTING INTRACLUSTER LIGHT IN MOCK IMAGES OF GALAXY CLUSTERS

Comparing observational findings of ICL with predictions from cosmological simulations is complicated by the wide range of definitions adopted to identify the ICL from observational data. In [Figure 2.3](#), we showed a handful of methods applicable to photometric observations, proving that even within a single cluster different definitions can result in very different ICL fractions. High-resolution spectroscopic information on the stellar population can ease the task but, other than being available for only a few clusters in the local Universe, currently, there are only a handful of instruments capable of attaining the resolution needed. As such, what we can hope for is to be able to quantify these differences by profiting from predictions in simulations.

In this chapter, we will take this issue one step further. According to what is discussed in [Chapter 2](#) and [Chapter 7](#), the stellar population in the ICL exhibits a fundamental property: the ICL has higher velocity dispersions (thus, higher entropy) when compared to stars bound to a galaxy, resulting in different phase-space properties. In simulations, recovering these differences is trivial since each star particle is tracked down to its position and velocity at each timestep; thus, one can classify stars in the main halo of simulated clusters into BCG and ICL according to their individual dynamical properties. Following this line of thought, in this chapter, we propose a tool for inferring similar information from observed images of galaxy clusters. This will require the creation of synthetic photometric maps of the stellar component in simulated galaxy clusters and the use of DL techniques to build a regressive model.

In the following, we aim at addressing the following issues.

Can we effectively bridge observations with simulations to study the properties of the ICL? Can we infer the dynamically different ICL in photometric maps? Can we train a simple ML-based model to perform this task? Can we understand which observable best traces the ICL composition? Which systematics may we be overlooking when using simulations to predict ICL? What can we improve in our model for future developments?

We warn the reader that the model here described represents our preliminary effort to tackle this complex problem. As such, at the current stage, it serves only to explore the fundamental challenges we will have to face in future analyses. It is better thought of as a sandbox: this work provides one with the basic units required to build a DL model tailored for the identification of the ICL in spectro-photometric maps. For this reason,

most of the discussion is related to possible future iterations and solutions to solve the issues identified in this preliminary analysis.

8.1 SCIENTIFIC CONTEXT

In the last decades, the low surface brightness optical luminosity in clusters and groups of galaxies has promised to deliver a wealth of information on the formation history of these objects. ICL holds important clues relative to structure formation and evolution since this stellar component tracks the evolution of its host cluster (or group) solely with its evolution. For example, colours and metallicities will depend on the ICL's main formation channel, and thus they can reveal whether ICL has formed mostly through tidal stripping of massive satellites or after the disruption of dwarf galaxies. As such, ICL is a fossil record of the evolutionary paths taken by the galaxies from which they formed. Discoveries in this sense have been only recently possible thanks to the significant enhancement in the sensitivity of new-generation telescopes. These have allowed us to extend the exploration to intermediate redshifts ($z \sim 1$) clusters (Zibetti et al., 2005; Burke et al., 2012; Montes and Trujillo, 2014; Ko and Jee, 2018) and into the group regimes (Da Rocha and De Oliveira, 2005; Poliakov et al., 2021). Yet, we are only scratching the surface as we are lacking large samples of galaxy clusters on which to perform in-depth studies on the ICL.

Possibly, this scarcity will change with the advent of the next generation of surveys that will allow us to reach unprecedented depths over large areas in the sky. For example, Euclid¹ is expected to observe around 10^5 galaxy clusters in a lightcone area of 15000 deg^2 up to redshift $z \sim 2$ (Sartoris et al., 2016): for the first time, we will be able to undertake a statistically significant study of the properties of the ICL across mass ranges and out to the earliest redshifts where clusters begin to virialise. In addition, the combination of multi-wavelength observations from LSST² and the Nancy Grace Roman Space Telescope³ will allow the most accurate determination of the stellar populations of the ICL. In this regard, the newly launched JWST⁴ is already granting us incredible images of this component (Pascale et al., 2022) in the innermost regions of galaxy clusters. The cameras mounted on this telescope, which can benefit from the spectroscopic data from the Integral Field Units, will particularly turn out to be a highly valuable source of information.

The quality of such data calls for theoretical predictions of comparable quality to be obtained from both hydrodynamical simulations and SAMs. One of the foremost problems faced by the scientific community at the moment is the variety of ICL definitions present in the literature. Photometry-related methods (i.e., surface brightness cuts, profile fitting, and wavelet techniques) have been discussed in Section 2.2.2, while a

¹ Euclid: <https://www.euclid-ec.org/>

² Large Synoptic Survey Telescope (LSST) <https://www.lsst.org/>

³ Nancy Grace Roman Space Telescope: <https://roman.gsfc.nasa.gov/>

⁴ James Webb Space Telescope (JWST) <https://webb.nasa.gov/>

dynamical approach has been extensively explained in [Chapter 7](#). Here, we aim at bridging these two approaches with a novel approach built upon cosmological simulations.

The project aims at developing a DL method based on CNNs to identify ICL in mock observational images of simulated galaxy clusters. We choose the U-Net as the preferred architecture since it is particularly useful for pixel-to-pixel classification/regression problems. The input images are extracted from a combination of photometric and spectroscopic properties of the stellar population in clusters. The dynamical classification of BCG and ICL (see [Chapter 7](#) for details) is used to build stellar catalogues to reconstruct the ICL fraction maps. In principle, a tool with the capability to identify the dynamically distinct components from these maps would be quite useful for application on real data.

8.2 SIMULATIONS

The cosmological hydrodynamical simulations of galaxy clusters used for this analysis are taken from the DIANOGA set (Bassini et al., 2020, and references therein) at the highest resolution (i.e., Hydro-10x). The initial gas and DM particle masses are reported in [Table 3.1](#). We select the 10 most massive central halos in the parent box of size $1 h^{-1}$ Gpc (7 Lagrangian regions), excluding only those containing contaminating particles within R_{200} . The adopted cosmology is a Λ CDM model with $\Omega_M = 0.24$, $\Omega_b = 0.037$ for the total matter and baryon density parameters, $n_s = 0.96$ for the slope of the primordial power spectrum, $\sigma_8 = 0.8$ for the normalization of the power spectrum, $h_0 = 0.72$ for the Hubble parameter in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. For each halo, we take the three orthogonal planes (i.e., XY, YZ and XZ). This set comprises 106 objects.

We constructed the DL framework with PyTorch (Paszke et al., 2017) which is an open source ML tool equipped for GPU acceleration tensor computation. We use the public library *pytorch-lr-finder*⁵ for the optimal learning rate calculation.

8.2.1 Input and output data

The purpose of this study is to demonstrate that it is possible to detect the ICL signal from synthetic observational images generated from simulated galaxy clusters. In the longer term, this would allow to adopt the same ICL and BCG dynamical definition in simulations and observations, thus allowing for a meaningful direct comparison. Ideally, we want to provide a set of mock images which include all stars in the main halo of the cluster (namely, BCG and ICL stars). The masking of the contaminants will be treated later in this section. As for describing BCG and ICL combined in a cluster, we extract primarily two classes of properties: photometric properties (i.e., magnitudes and colours in the u, g, r, i, z bandpasses from

⁵ <https://github.com/davidtvs/pytorch-lr-finder>

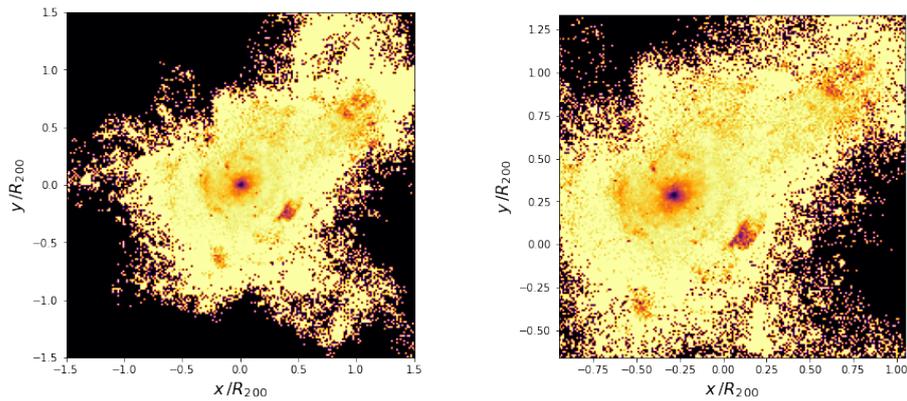
SDSS⁶) and velocity momenta of the stellar component. The former usually represents the properties that are most easily accessible in observations: as such, it represents the most common way of extracting ICL profiles in the local Universe. On the other hand, in a more complex approach, the first and second momentum of the velocity distribution along the line of sight (i.e., the mean and the r.m.s. of the particles' velocity distribution within a given pixel) provide the kinematic knowledge to disentangle the stellar components much like how it is done in simulations. In this case, observational measurements of the velocity momenta are generally estimated from the highly resolved spectral lines obtained through spectroscopy. This set of maps can be combined in different ways, as would be done when combining data from different observations. Therefore, the number of channels in the DL model will equal the number of properties we choose to use, or we can access, from observations. Thus, part of the issue is to assess which combinations of input data can provide the best results and quantify the differences between different such combinations. In our case, the output data predicted by the model is the ICL fraction present in each bin. In the following, input and output maps will have the same size, i.e. 199×199 .

In the following, we describe how clusters are framed for our dataset. To increase the volume of data for training and testing, we perform data augmentation on this set of images by following a few steps. To pick an example, in [Figure 8.1](#), we plot the maps of the ratio between the ICL mass and the total stellar mass in BCG and ICL, projected on the XY plane. Darker (brighter) colours mark lower (higher) values of this fraction. In the left panel, we present for reference the original cluster centred on the halo centre, with the size of the map corresponding to $3 R_{200}$. On this cluster, we apply five times a procedure to randomly resize the image and include the presence of contaminants as in observations. This provides us with a larger sample of clusters. In the following, we describe these steps in more detail.

Shifting and zooming. In the first step, corresponding to the right panel in [Figure 8.1](#), we randomly shift the image centre by (X_0, Y_0) , where both X_0 and Y_0 satisfy the condition $-0.4 < X_0/R_{200} < 0.4$. Since each shift is randomly extracted for one original cluster, we could potentially extract several different images according to the number of extractions. In the map in the right panel, the shift is $(-0.27, 0.34)$. The reason behind shifting the centre is twofold: (i) data augmentation and (ii) de-centring of the real images. Data augmentation is attained if we extract multiple images with different centres from the original one. A reasonable choice is to extract a small enough number of images to have a low degree of correlation within the set of images: two images should not be too similar. As for the de-centring of real images, we take into account the possibility of not perfectly centred cluster images in the observed sets. This operation changes

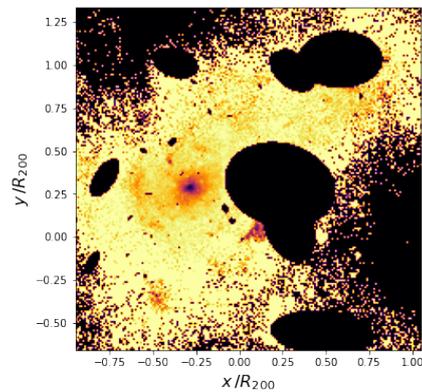
⁶ Sloan Digital Sky Survey (SDSS): <https://www.sdss.org/>

Figure 8.1: The ICL mass fraction over the ICL and BCG in one of the clusters. Darker (brighter) colours mark lower (higher) fractions.



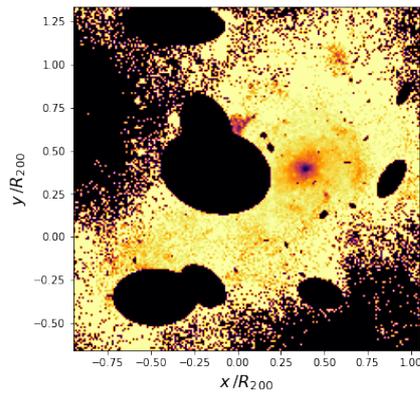
(a) One of the clusters centred on the halo centre and framed within $3 R_{200}$.

(b) The centre is shifted and the new coordinates are $(-0.27, 0.34)$, while the frame is zoomed into an image of physical size $2 R_{200}$ per side.

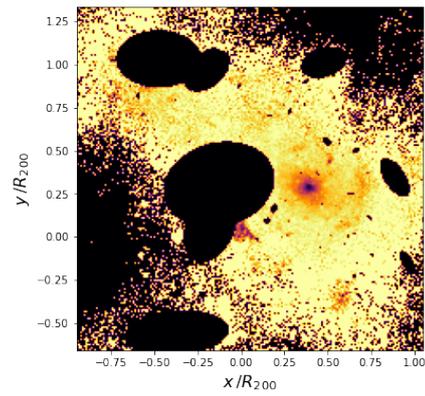


(c) We simulate the presence of interlopers (i.e., galaxies) by drawing random ellipses in the image.

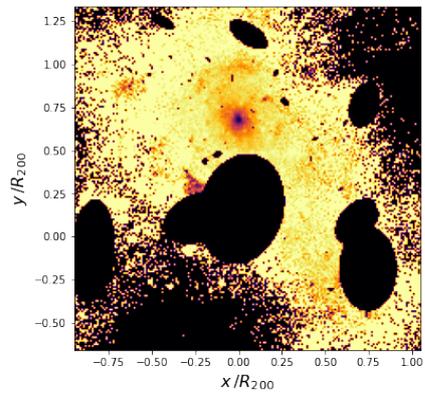
Figure 8.2: One way to provide data augmentation is to perform flips and rotations to one image, as reported in the panels.



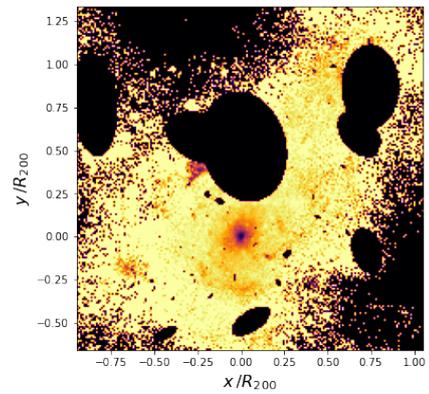
(a) 180 degrees rotation



(b) Vertical flip



(c) Horizontal flip



(d) Transpose

the reference system of coordinates, $(X, Y) \rightarrow (X', Y')$, but we will adopt this nomenclature only from the next section. At this point, we resize the physical scale of the image to $2 R_{200}$ per side.

Masking contaminants. For aperture photometry, a common operation performed on the data is to define, for a given image, a mask or array of pixel indices identifying which pixels in the image should be deleted from the data. These masks usually take into account all sources of contamination, such as foreground stars, cosmic rays, and galaxies. Common choices for masking galaxies are ellipsoids located on the contaminant which extend out to a safe radius to enclose the total light. On the other hand, in our simulations, we can exclude the contaminating galaxies from our images thanks to the halo structure provided by Subfind and what we are left with is just the star particles bound to the main halo (i.e., BCG and ICL). Therefore, to shadow the observational approach, we draw a random number of ellipsoids at random locations. For each image, we repeat this exercise five times, picking ellipsoids of varying sizes, thus for each original image we generate five new images appropriately masked. One of the images is shown in the bottom panel of [Figure 8.1](#).

Flips and rotations. Another way to increase the original number of images is to apply flips and rotations for each image. Such perturbations are reported in [Figure 8.2](#). We applied a vertical and horizontal flip, 180 degrees rotation, and the transposition of the image. Keeping the originals, we gather four new ones, for a total of 25 images per original image.

Some of the clusters contain low-resolution DM particles within R_{200} , therefore, we exclude them from the analysis. The final dataset comprises 7950 synthetic images for each channel that are then split into 70 : 20 : 10 to create training:validation:testing sets. Furthermore, for each channel in each set (i.e., training, validation, and testing), we compute the median and standard deviation to standardise the input images.

8.3 RESULTS

Choosing the best combination of input images is not a trivial task as it requires testing several configurations to assess which property (or combination of properties) carries most of the useful information. As such, we present in the following the results for all the meaningful test runs, with the following input maps: photometry, velocity, and velocity dispersion.

8.3.1 Optimisation

In [Section 4.3.2](#), we introduced the concept of the learning rate range test as a viable solution to prevent overfitting during training. Optimising the search of the minimum during the gradient descent can provide faster

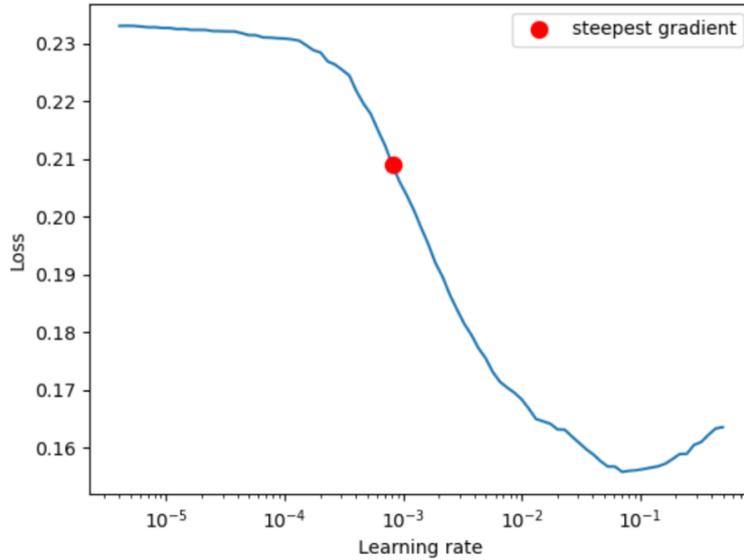


Figure 8.3: Loss function for one of the input datasets as a function of the learning rate. The red dot marks the point in the curve with the steepest gradient which corresponds to the suggested value for the learning rate.

and better results during the training phase. There are several choices one could make in this regard, we chose Adam (see [Section 4.3.1](#) for further details) for its versatility. The only hyperparameter we tune is the learning rate α , and we set the weight decay to 10^{-2} while we leave the other hyperparameter to the default values given by PyTorch (i.e., $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$) without further investigation on their parameter space. According to the Cyclical Learning Rate framework, we record the loss function for every set of data at each mini-batch (64 images) in a single epoch of training, as shown in [Figure 8.3](#). This exercise requires computing the loss function for a given range of learning rates, we choose $[1 \times 10^{-6}, 1]$ in equal bins of size 0.01 which allows us to determine the point at which the gradient in the loss function is the steepest (red dot), and therefore the learning process is more efficient. The loss function is the Mean Squared Error (MSE) which is perhaps the simplest loss function being used in ML applications. The MSE between the true label y_i and the predicted one \hat{y}_i is formally defined by the following equation:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (50)$$

where N is the number of samples we are testing against. We record the final hyperparameters in [Table 8.1](#).

8.3.2 Training phase

The training phase is approached for all the different sets similarly. We set the optimiser hyperparameters as found in the previous section. We choose

Table 8.1: A summary of the hyperparameters used for the Adam optimiser in the different input models.

Input data	Learning rate α
Magnitude	2.66×10^{-4}
Colour	2.66×10^{-4}
Velocity	6×10^{-3}
Velocity dispersion	2×10^{-4}

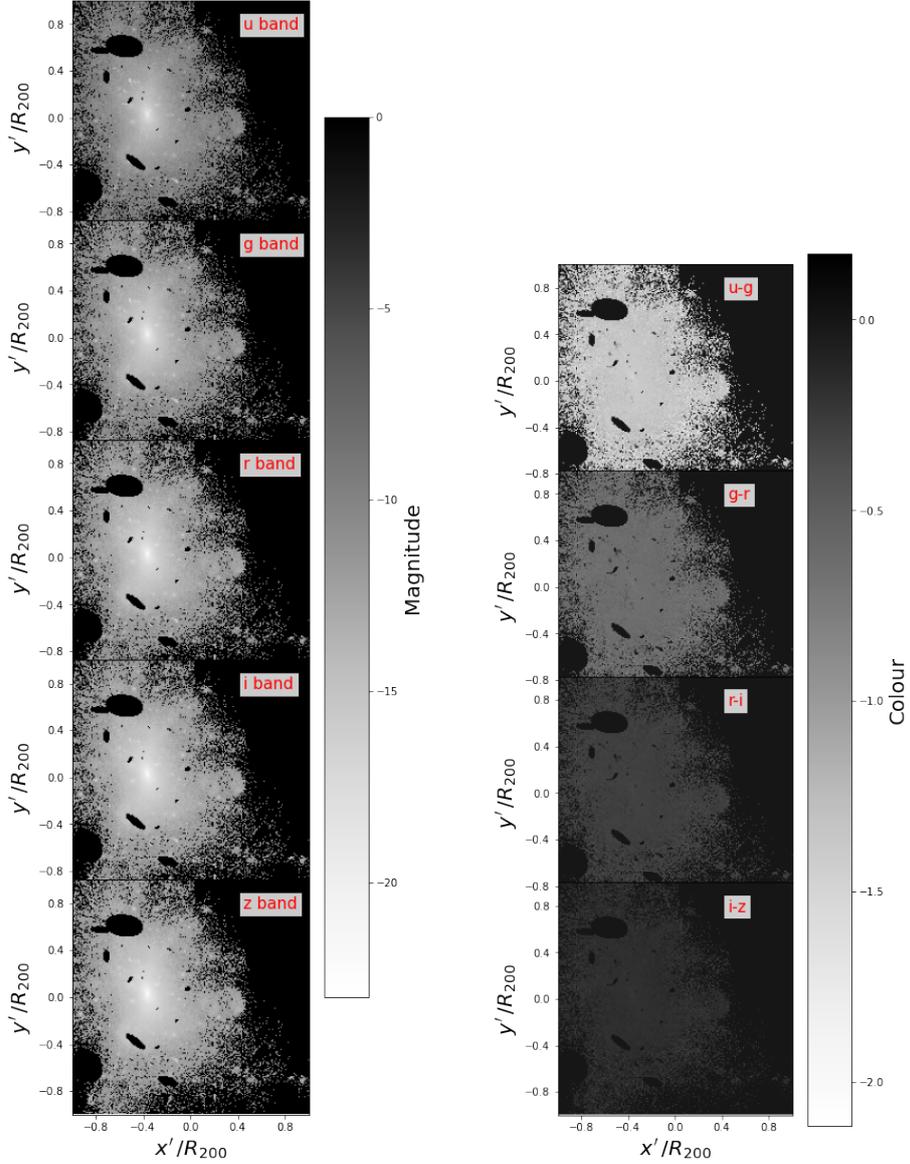
100 epochs as an initial guess for the number of times the full training data will be seen by the model. However, we apply early stopping if in 10 epochs the model does not improve its validation error over training error. In other words, we save as the best model the last iteration in which the validation error is below the training error and we stop the search if this condition does not change in 10 epochs. Most of the training phases end after ~ 20 epochs.

8.3.3 *Stellar mock images*

In the following, we present a collection of results from the use of different stellar properties (surface brightness, velocity and velocity dispersion maps) as input to the model. Gathering the bidimensional information of velocity and velocity dispersion is less trivial than the stellar photometry. The stellar population and kinematic properties have been explored through multiwavelength photometry (e.g., Montes and Trujillo, 2018, 2019), globular clusters (Alamo-Martínez and Blakeslee, 2017), and individual objects such as planetary nebulae (Arnaboldi et al., 1996) and red giant branch stars (Longobardi et al., 2018). However, only the Integral Field Spectroscopy (IFS) offers the combined spatial and spectral information to spatially resolve the kinematics of any astronomical target. These will become readily available with JWST’s launch along with the ongoing MUSE⁷’s observations.

SURFACE BRIGHTNESS. Firstly, we describe our analysis when using photometric properties since this input class currently represents the most common way to observe ICL. Surface brightness images are constructed for five different photometric bands (u, g, r, i, z band-passes from SDSS) in terms of both magnitudes and colours. The input channels of colours are four since we compute $u - g, g - r, r - i,$ and $i - z$. A representation of a cluster from this dataset is provided in Figure 8.4: magnitude and colour maps are in panels (a) and (b), respectively. For each channel, we compute the median and standard deviation of each set (i.e., training, validation, and testing) to standardise the input images. Then, we follow the procedure outlined in Section 8.3.2 to perform the training. Since the model trained with

⁷ MUSE: <http://muse-vlt.eu/science/>



(a) Magnitudes in *u*, *g*, *r*, *i*, *z* bands from the SDSS survey

(b) Colours *u-g*, *g-r*, *r-i*, *i-z*

Figure 8.4: Magnitude and colours maps

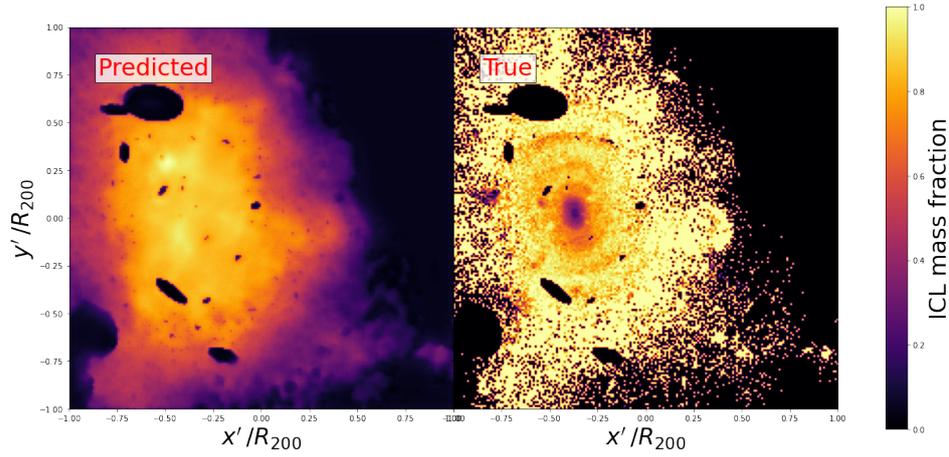
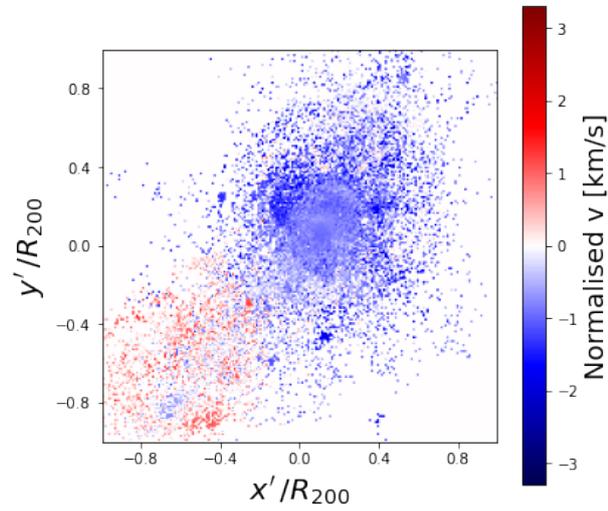


Figure 8.5: Predicted (left) and true (right panel) distribution of the ICL mass fraction in a cluster belonging to the testing set. The colour bar marks the region with the highest concentration with yellow towards purple for lower concentrations.

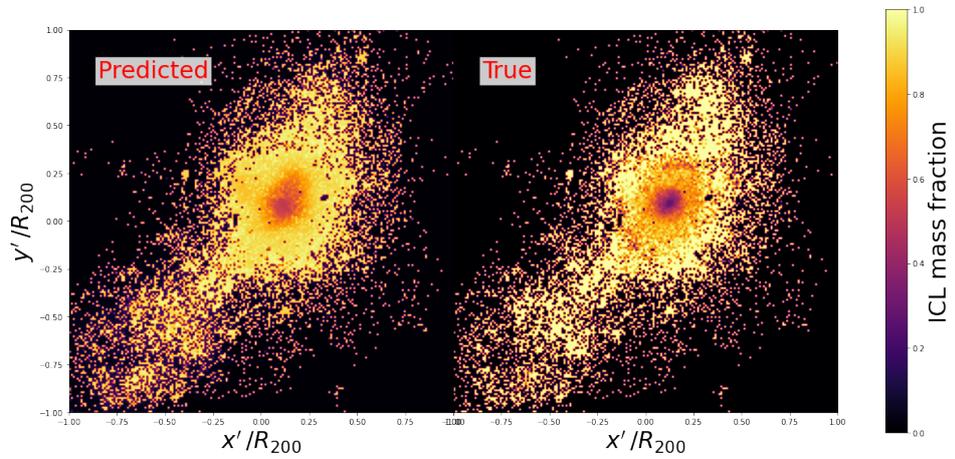
either magnitude or colour maps gives comparable predictive power, we briefly describe results coming only from the latter.

The model training phase ends due to early stopping at the 10th epoch, with a validation error 0.1 (MSE). We display the output from the previous cluster in [Figure 8.5](#): the left panel reports the predicted ICL mass fraction map, while the right panel shows the expected map. The model performance is overall inadequate for our goals. We notice that several features are not detected by the trained model: the central region has an ongoing shell structure which is absent in the predicted map, same for the BCG position which is completely missed. We suppose the shell-like features to be caused by the processes described in [Section 7.4.5](#), possibly at their earliest stage when the stellar dynamics is most disturbed and thus stars are hardly found to be bound to the BCG. We expect them to be associated with the BCG at later times, after relaxation mechanisms take place. On the other hand, the masking ellipses are distinguished, proving that simple structures should be identified by the model. Additional insight is given by the outer region (in purple on the predicted map) which shows signs of a premature end of the training phase.

The training stage is a delicate phase which indeed requires a broad parameter search and tuning. Considering that in our case, early stopping prevents the algorithm to continue its search, we may inquire why this happens. Excluding the obvious possibility of building a more complex DL architecture than the U-Net, we are left with a handful of options. One possibility could be to invest more effort in experimenting with different loss functions (e.g., mean absolute error, Huber loss) and different hyperparameters in the optimiser Adam (even attempting a different optimiser). One drawback of this

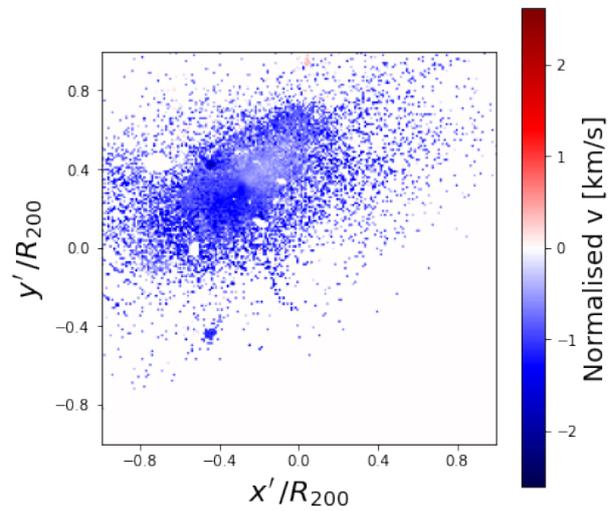


(a) Input velocity map of one of the clusters in the testing set. The velocity is scaled to the mean and standard deviation of the set of images.

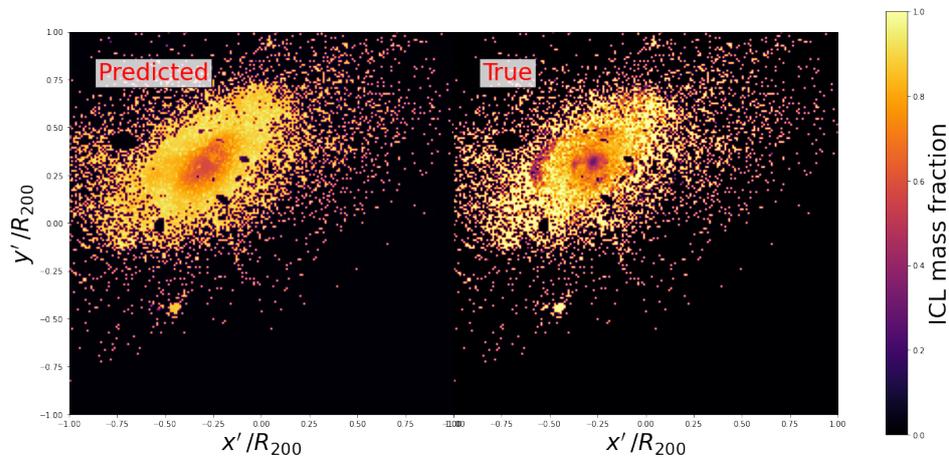


(b) Predicted (left panel) and true (right panel) ICL mass fraction maps from one cluster in the testing set. The colour bar marks the region with the highest concentration with yellow towards purple for lower concentrations.

Figure 8.6: Input (panel a) and output (panel b) of the model using velocity maps as input images.



(a) Input velocity map of one of the clusters in the testing set. The velocity is scaled to the mean and standard deviation of the set of images.



(b) Predicted (left panel) and true (right panel) ICL mass fraction maps from one cluster in the testing set. The colour bar marks the region with the highest concentration with yellow towards purple for lower concentrations.

Figure 8.7: Input (panel a) and output (panel b) of the model using velocity maps as input images.

approach would be the high computational cost required to perform a complete hyperparameter exploration analysis. Another obvious option to further investigate is the meaningfulness of the stellar photometric properties as a probe of the ICL distribution. A possibility is that our modelling is not accurate enough to capture potential features or we should not necessarily expect to gain an insight into the BCG and/or ICL distributions by merely studying the photometry. One key point would be to assert whether the predicted ICL colour gradients recover the observed ones (Yoo et al., 2021; Montes et al., 2021; Golden-Marx et al., 2022, just to name a few). Ultimately, a definitive obstacle in the training phase could be the reduced size of the input dataset which may not be large enough to allow the model to learn: early stopping may have happened because, given the size of our sample, the model could not have learnt more without being affected by overfitting. Increasing the sample size would greatly help in understanding whether this is the case. In conclusion, our work showed that, given our training set size and the chosen model, photometry maps of the stellar population alone cannot be used to probe the ICL distribution.

VELOCITY. The velocity maps are constructed taking the mean of the line of sight velocity in each pixel. We re-scale the images to the mean and standard deviation of the sample, as shown in panels (a) of [Figure 8.6](#) and [Figure 8.7](#). Notice that such re-scaling does not necessarily normalise the single image (e.g., velocities in the image are not centred on zero), but rather it homogenises the entire sample. In panels (b), we show the comparison between the true and predicted distributions of the ICL mass fraction with the trained model. Given the good results, we show two different clusters. The colour bar marks the region with the highest concentration with yellow to purple for lower concentrations. Velocity maps seem to carry most of the information needed to identify the impact of ICL in the cluster. Although the two distributions are not perfectly matched, cluster centres are in general accurately recovered (darker spots in the image corresponding to the BCG position) and also gradients in the ICL distribution follow similar physical scales. MSE in the validation set ranges around 0.02 for each image. We claim that velocity maps of the stellar population give signs of hope in recovering ICL distributions, with an accurate hyperparameter search.

VELOCITY DISPERSION. Velocity dispersion in each bin is computed as the r.m.s. of the velocity distribution in the bin. Each image is normalised with mean and standard deviation from the sample, as presented in panel (a) of [Figure 8.8](#). As before, in panel (b) of [Figure 8.8](#), we show the ICL mass distribution in both the predicted (left) and true (right) cases. Results from velocity dispersion maps seem to be in an intermediate phase between the trained model with photometry and velocity. Central regions are usually correctly

located, presenting a gradient in the ICL distribution similar to the one in the expected map. In the outskirts, we see larger uncertainties in the predictions possibly linked to failed training. Solutions to improve such results would be similar to those presented for the surface brightness case, spanning the increase in the training set size and broader parameter search. One critical point to address would be to understand why velocity maps should perform better than velocity dispersion maps as input, or whether this difference between the modelling should be connected to an incorrect training phase in the latter. With this configuration, MSE measures about 0.05 in the validation dataset.

8.3.4 Quantifying the accuracy

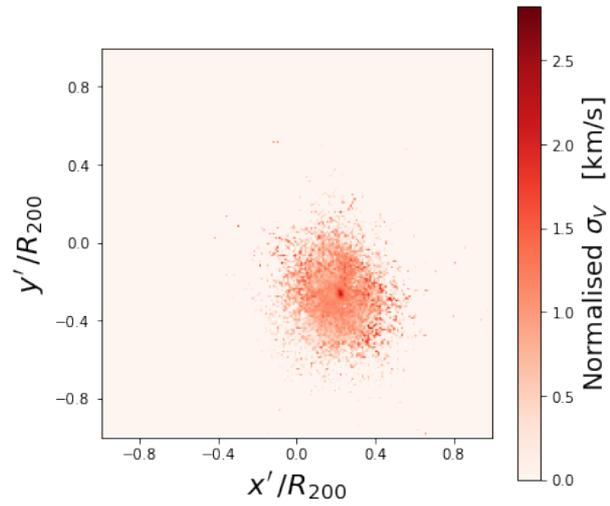
In the previous section, we have proved that velocity maps are the best input candidates to identify ICL and BCG components in synthetic (and, possibly, real) images of galaxy clusters. At this point, we wish to better quantify the errors in our predictions, particularly for our best-case scenario. In [Figure 8.9](#), we present the results for the cluster displayed in [Figure 8.6](#) (left panel) and [Figure 8.7](#) (right panel). Each panel features the Percentage Error (PE), the estimated relative error between the predicted value \hat{y}_i in the i^{th} pixel and its true value y_i in percentage, which is formally defined as:

$$\text{PE} = 100\% \frac{y_i - \hat{y}_i}{y_i}. \quad (51)$$

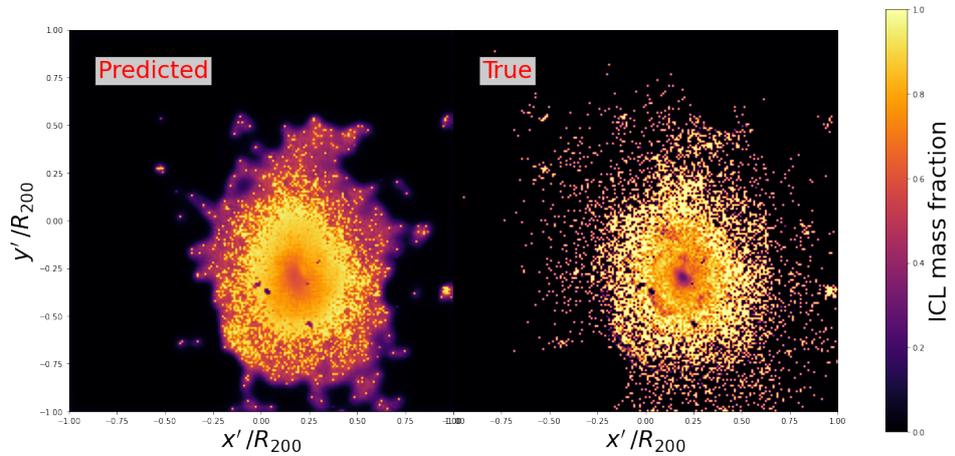
The colour bar guides the lecture: red marks areas where the predicted ICL fraction is higher than in the true case and black denotes the opposite. It is interesting to note that indeed most of the ICL over-prediction is around the cluster centre, where the BCG resides. Otherwise, the PE ranges between mid-values ($\sim 30 - 50$ per cent). We need to cure this discrepancy given that the central regions are of the most importance in our analysis. One possible alternative could be to modify the loss function, the main responsible for the parameter update during the training phase, to a customised function to provide adequate support to the model around the BCG location. In other words, we could define a loss function which scores significantly worse when a wrong guess is done in the central regions of the image than in other parts, penalising the set of model parameters at that iteration.

To put these results into perspective, we take the mean of the integrated PE along each pixel line, both vertically and horizontally). This should allow us to appreciate where most of the differences in our predictions lie in the larger cluster environment. Thus, we consider the Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N |\text{PE}|, = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (52)$$

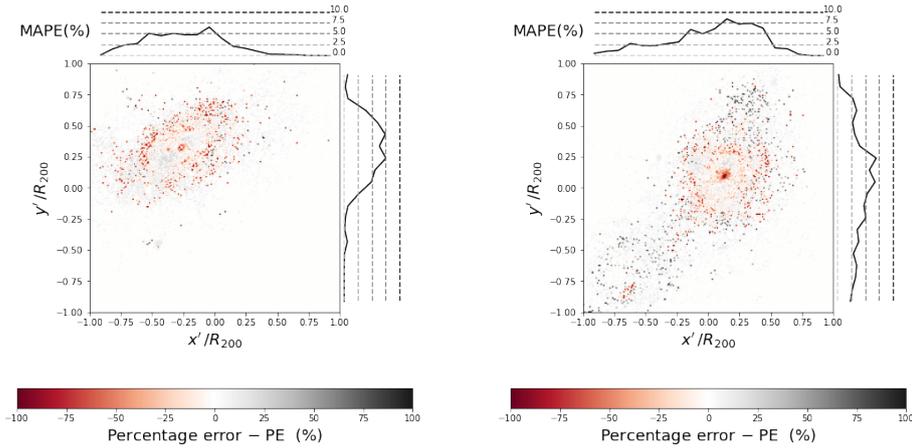


(a) Input velocity dispersion map of one of the clusters in the testing set. The velocity dispersion is scaled to the mean and standard deviation of the set of images.



(b) Predicted (left panel) and true (right panel) ICL mass fraction maps from one cluster in the testing set. The colour bar marks the region with the highest concentration with yellow towards purple for lower concentrations.

Figure 8.8: Input (panel a) and output (panel b) of the model using velocity dispersion maps as input images.



(a) Error map for the cluster in Figure 8.6

(b) Error map for the cluster in Figure 8.7.

Figure 8.9: Percentage error maps for the clusters with velocity maps as the input dataset. In each central panel, we display the PE between the predicted and true ICL mass fraction for pixel. Generalising these results, we compute the MAPE given by each line of pixels. We report these errors on the top histogram (i.e., MAPE along the vertical direction) and right histogram (MAPE along the horizontal direction) with respect to the central panel. The grey dashed lines mark the necessary scale.

where N is the number of pixels along the line. Taking the absolute sum allows us to take into account errors both in the direction of under and over-prediction. We report these results in the histograms at the top and the right of each panel. Additionally, we include dashed solid lines to mark the scale of the curves. Quite surprisingly, the MAPE is distributed almost independently of the distance from the cluster centre. Approximately, we measure errors at most of 10 per cent. This estimate also reassures us that the highest errors (close to the BCG) do not extend over the whole map, as they are diluted within the direction of integration of each pixel in the MAPE. In other words, the highest errors registered (close to the BCG location we measure $PE \simeq 98$ per cent) are restricted to small regions and do not affect the large scale scenario, estimated by the MAPE ($\simeq 8$ per cent).

8.3.5 Increasing the numerical resolution

In this chapter, we presented a novel way to determine the ICL distribution in clusters from mock images of several stellar properties (i.e., photometry and dynamics). However, one caveat within the use of this technique is the underlying assumption that our simulation set must correctly resolve both ICL and BCG stellar populations. In other words, to rely on the DL model predictions which have been tuned on the simulated stellar properties, we have to ensure that the stellar population is a realistic representation of the observed BCG and ICL. In this regard, a first check can be done to assess

the dependence of the simulation predictions on numerical resolution. Restricted to our case, this test would monitor to what extent the BCG and ICL build-up is dependent on the given simulation. Ultimately, our goal is to build a model consistent at all resolutions, and possibly redshifts. To this end, we gather three simulation runs from the DIANOGA set corresponding to three numerical resolutions (i.e., 1x, 10x, and 25x, as described in Table 3.1) of one cluster at $z = 0$ and we analyse the ICL distribution within the virial radius R_{200} . Provided that initial conditions and included physics are the same, we should expect to obtain similar clusters at different resolutions, unless other factors are playing a role.

Our first result is presented in Figure 8.10. We define the ICL fraction as the ratio of ICL stars to the stellar population of the main halo (i.e., BCG and ICL). Notice that this definition is not weighted by the mass of star particles, and therefore does not directly relate to observed ICL fractions. In the plot, the solid lines represent the ICL fractions as a function of the normalised distance from the cluster centre for the three numerical resolutions. Each cluster has a low fraction in the inner regions, where BCG is usually located, while at larger radii the fraction increases to ~ 1 . Most of the tension is found at the centre: as the resolution increases profiles become shallower. This is caused by a larger presence of stars associated with ICL at the centre. This result suggests that the boost in resolution may have eased the formation of a larger dynamically warm component at the centre, unable to bind with the central galaxy's potential.

In Figure 8.11, we plot the cumulative stellar mass of BCG (blue) and ICL (red) as a function of the clustercentric distance normalised by R_{200} . The three resolutions are reported with different line styles: dashed (1x), dash-dotted (10x) and dotted (25x). With vertical grey lines, we mark the transition radii. We remind that the transition radius (as described in Chapter 7) determines the radius where the transition between the BCG-dominated and ICL-dominated regions takes place. Contrarily to what is shown in Figure 8.10, here we can appreciate quantitatively the differences among the different resolutions. This allows us to estimate the stellar mass growth of both BCG and ICL with increasing resolution: generally speaking, more stars form in the main halo of the simulated cluster. Secondly, we notice that such an increase does not take place homogeneously. This is highlighted by the vertical lines which show that as the resolution increases, the transition radius goes towards smaller radii. In other words, we expect to encounter an excess of ICL stars at smaller radii. This is further confirmed by the BCG and ICL change between 1x and 25x in the innermost region: from a dominating central galaxy in the matter density distribution at the 1x resolution ($M_{\text{BCG}}^* \simeq 4M_{\text{ICL}}^*$), we note a much smaller difference between the two components in the 25x cluster ($M_{\text{BCG}}^* \simeq 2M_{\text{ICL}}^*$). This is in line with the ICL fraction shown in Figure 8.10.

In summary, we determined that the build-up of the stellar component in the main halo of our simulated galaxy clusters is not independent of numerical resolution. Generally, we find at $z = 0$ added BCG and ICL stars

as the resolution increases, causing the formation of more massive stellar halos. This result can be explained in light of multiple coexisting factors that can be affected by numerical resolution: star formation processes, disruption of small substructures, and AGN feedback.

Star formation processes. The subgrid model implemented to compute star formation in the simulation depends on the resolved densities of gas particles; this means that, as the resolution improves, smaller scales are resolved, causing more stars to form unless feedback energy from SNe and AGN provides a self-regulation of the star formation process. Therefore, an enhancement in the numerical resolution may produce an enhancement in star formation.

Disruption of small structures. Another possible explanation can be given by the larger number of substructures which are formed and dissolved to assemble ICL. ICL is preferentially built from the stars stripped from the orbiting galaxy members. With increasing numerical resolution, simulations can form smaller galaxies. Ultimately, these galaxies may contribute to the ICL build-up, with their disruption caused by the tidal field in the cluster environment, thus contributing to an increase of the total stellar mass associated with the diffuse stellar halo.

AGN feedback. The additional impact provided by the AGN feedback should not be ignored at these scales. It is established that the AGN feedback affects the growth and structure of ICL, but most importantly of BCG, and therefore should be a key ingredient in determining the stellar content of galaxy clusters. Understanding the impact of AGN feedback on the cluster's stellar halo can unveil important connections between the two. We point out that across the different resolutions the AGN feedback model does not change, however since increasing resolutions leads to higher maxima of the gas density we should expect variations in the SMBH accretion mechanism.

Only a complete analysis of the formation mechanisms of these stellar components as a function of the numerical resolution may be able to disentangle these effects. As this detailed investigation goes beyond the scope of this work, we will defer this analysis to future research.

8.4 CONCLUSIONS AND FUTURE DEVELOPMENTS

Nowadays, DL techniques represent invaluable frameworks to efficiently extract key information from the large volume of data in both cosmological simulations and observations. As the challenges we are posed with are also structurally becoming computationally more expensive, these tools allow us to ease the increasing computational overhead. Their added benefit is the invaluable predictive power they can provide in regression problems of different natures. This is the task that we addressed in this chapter, as traditional models are unable to compete.

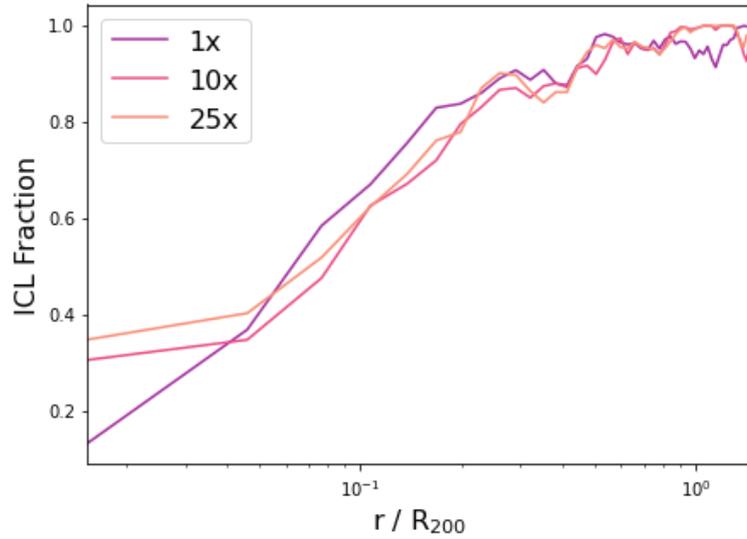


Figure 8.10: ICL fractions for one cluster as a function of the normalised clustercentric distance. The ICL fraction is determined as $\text{ICL}/(\text{BCG}+\text{ICL})$. Lines are colour-coded for the three different numerical resolutions (i.e., 1x, 10x, 25x). The dotted horizontal lines report the median ICL fraction from the entire sample of clusters in each resolution run.

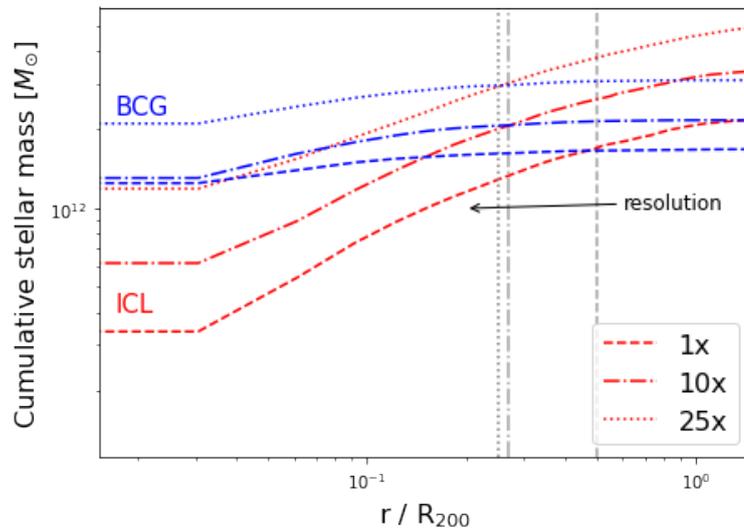


Figure 8.11: Cumulative stellar mass of the ICL (red) and BCG (blue) as a function of the normalised clustercentric radius for the same cluster simulated at three different numerical resolutions (i.e., 1x, 10x, 25x). Each resolution is marked with a different line style, as reported in the legend. The vertical grey lines represent the "transition radius" for each pair of simulations: these monitor the change between the BCG-dominated area to the ICL-dominated one. As the resolution increases, the shift to the ICL-dominated region moves to smaller radii.

The challenge of unambiguously identifying ICL in galaxy clusters and groups has been going on for some time and it has become a matter of increasing interest, as current and future surveys gather larger and larger volumes of data to be processed. One example is provided by the recent observations carried out by JWST, which is providing deep and high spatial resolution images to study ICL with a high signal-to-noise ratio up to a radial distance of ~ 400 kpc (Montes and Trujillo, 2022), twice as far with respect to previous HST⁸ studies of intermediate redshift clusters (e.g. DeMaio et al., 2018). This opens up the possibility of exploring the rich mixture of processes that are building the ICL.

In this chapter, we discussed a method to connect observations of the ICL and predictions from simulations using DL techniques. Observationally, we can determine the ICL distribution in a cluster, once we do several assumptions on the light profile. However, not only do they often provide different results according to the assumption chosen, but they also do not necessarily detect the dynamical differences we expect to be present in this component. In this chapter, we argued the importance of identifying ICL by dynamical means both in observational analyses and simulations. In principle, the scatter present in the ICL measurements from the literature (e.g. Kluge et al., 2021) could be alleviated by standardising the methods used to define this component, although we are still far from having settled the matter. In this study, we selected a sample of clusters (among the most massive) from the DIANOGA zoom-in cosmological hydrodynamical simulations, which we used to create mock images (199×199) of galaxy clusters. From the original dataset, we performed data augmentation by modifying (e.g., resizing, zooming, rotating) the images and we included a simplified treatment of contaminants (i.e., interlopers) by randomly masking elliptic regions in the image. We created these images for different observables of the stellar components: namely, photometry (magnitudes and colours from the SDSS bandpasses), line of sight velocity and velocity dispersion. We trained three different models for the three observables, using the same DL architecture (a U-Net model). Training is done by employing early stopping and Adam optimiser to prevent overfitting, preceded by a search for the best learning rate hyperparameter. The analysis has shown the velocity map to be the best candidate as input feature in the task. We discussed the accuracy of the results, promoting our work as a proof of concept of how one can include DL techniques to effectively advance in the identification of a dynamically-defined ICL from photometric and spectroscopic maps.

In its present form, the project presented here has taken us possibly further than what originally we had hoped. It has shown us a way forward to connect the predictive power of cosmological zoom-in simulations to observed galaxy clusters in a novel fashion. Additionally, it has proved that nowadays DL techniques are accessible to a large audience thanks to the effort of a community that has endlessly worked in the past years to provide useful online guides and tutorials: today, it is possible to incorporate a DL

Future steps

⁸ Hubble Space Telescope (HST): <https://hubblesite.org/>

framework in the standard astronomer’s toolbox. Nevertheless, we should keep in mind that a few relevant points still need to be addressed in our work for a satisfactory completion of this project.

Firstly, we need to increase the resolution of our simulations to probe the central regions of galaxy clusters (\sim kpc scale). In this work, we framed our images to extend to the virial radius to be substantially robust against resolution effects. However, from one standpoint, we still need to fully assess whether a trained DL model can detect consistently the transition (i.e., gradient) between the BCG-dominated region in the inner core to the ICL-dominated one when including a much smaller field-of-view. Such a requirement is needed since current high-resolution spectrographs (such as the Integral Field Unit mounted on MUSE) have field of views of $\sim 1 \times 1$ arcmin which corresponds to observing a cluster’s innermost regions in the local Universe (e.g., at $z = 0.3$, 1 arcmin corresponds to a physical scale of ~ 260 kpc). This step is possibly the most critical we will have to face. Preliminary analyses have shown that the modelling of ICL and BCG build-up in our simulated galaxy clusters depends on the numerical resolution, as we have discussed in [Section 8.3.5](#).

Secondly, we aim to refine our procedure to generate mock images of simulated galaxy clusters. A detailed treatment of interlopers, instrument sensitivity, background noise, and the inclusion of a point spread function are just a few examples of how the dataset might significantly improve in capturing realistic features commonly present in observational data. Ultimately, we should set up a collection of mock observations from a chosen instrument in order to calibrate our model to infer ICL from its measurements, once we have demonstrated that we can effectively extract meaningful information on the stellar population from a DL framework. Therefore, our current work should be taken as a mere proof of concept for future studies on ICL detection, rather than as a stand-alone project. Following this line of thought, in [Section 3.6](#), we highlighted the benefits of smoothing a synthetic image by employing a kernel that technically translates the distribution of particles to a continuous field. The operating space is large and open to the possibilities for an advanced degree of realism in our set of synthetic images of galaxy clusters.

In conclusion, we have demonstrated that DL techniques (especially CNNs) are powerful tools for further research in constructing a full pipeline of ICL detection. However, the results shown here highlight that there is still room for improvement in the current implementation and as such we should strive to work on future progress with the available computational resources.

Part IV

FINAL REMARKS

CONCLUSIONS

This chapter will conclude the study by summarising the key findings and discussing the value and contribution thereof. It will also review the limitations of the study and propose cues for future research.

9.1 THIS WORK

In this thesis, we resorted to advanced numerical cosmological simulations to study the internal dynamics of galaxy clusters. According to the standard cosmological scenario, structure formation occurs hierarchically, whereby the smallest perturbations collapse foremost to form larger objects. Galaxy clusters (presented in [Chapter 2](#)), being the largest gravitationally bound systems in the Universe, are thought to have assembled in recent times and thus are expected to still retain traces of the main mechanisms involved in the gravitational collapse. In this study, we addressed a number of scientific questions on the topic by resorting to the analysis of advanced hydrodynamical cosmological simulations (presented in [Chapter 3](#)) and the help of ML and DL techniques (described in [Chapter 4](#)).

In the first part of the thesis (including [Chapter 5](#) and [Chapter 6](#)), we tested several predictions from our hydrodynamical cosmological simulation sets, and compared them to the most recent observational results in the local Universe. Conversely, [Chapter 7](#) and [Chapter 8](#) focused on developing the numerical methods in support of this new working framework. The simulated galaxy cluster set is sampled from 29 zoom-in simulations of massive galaxy clusters (DIANOGA) in the redshift range $0 \leq z \leq 2$.

More in detail, we addressed the following aspects of the cluster galaxy population.

On the phase-space structure of simulated galaxy clusters

In [Chapter 5](#), we presented a study on the properties of the phase-space structure of galaxy clusters. The physical mechanisms shaping the universal density profiles at the largest mass scales are still unknown, but it has been suggested that they may be the result of the evolution of a dynamical attractor whose fingerprint is the universal radial profiles of the so-called pseudo-entropy (or phase-space density). Pseudo-entropy is defined from the velocity dispersion σ and the density profile ρ of a given collisionless tracer as $\sigma^2/\rho^{2/3}$ and it is usually very well described by a simple power-law. By exploring different resolutions and physics (i.e.,

N-body and hydrodynamical simulations), we study pseudo-entropy profiles for the collisionless tracers in galaxy clusters. We compare observed and simulated pseudo-entropy profiles and find good agreement in both normalisation and slope. Finally, we investigate the pseudo-entropy traced by the stars focusing our interest on the dynamical distinction between ICL and the stars bound to the BCG: the combination of these two pseudo-entropy profiles is well described by a single power-law out to the cluster virial radius. The results presented in this chapter have been published in Marini et al. (2021a).

Velocity dispersion of the brightest cluster galaxies

In [Chapter 6](#), the analysis focused on characterising the dynamical properties of the BCGs which is a peculiar family of objects: being the most luminous (and most massive) galaxies in the Universe, they are often located at the bottom of the gravitational potential of galaxy clusters. Because of their privileged position, their properties are severely influenced by, and in turn heavily affect, the extreme environmental conditions of galaxy cluster centres, the latter being sites of interesting evolutionary phenomena. This is one of the many reasons why these systems are attractive targets to benchmark models of galaxy formation. To attain this scope, we employed the central galaxies in the DIANOGA set. This has required an accurate analysis of the dynamical properties describing the simulated cluster population, such as density profiles, but particularly velocity dispersion profiles which can be directly compared to the observational data at our disposal (i.e., Sartoris et al., 2020; Sohn et al., 2020, 2022). Recent studies have demonstrated that stellar velocity dispersion profiles tend to increase at large radii, approximating their dynamics to that of galaxies in the external regions. A direct comparison of our findings with the observational results has shown that state-of-art cosmological simulations are close to mimicking the stellar dynamics in these regions, although there is still room for further investigations. The content of this chapter has been published in Marini et al. (2021b).

Identifying intracluster light with machine learning

Given the dynamical distinction between BCG and ICL, in [Chapter 7](#) we developed a ML method based on a supervised Random Forest to classify stars in simulated galaxy clusters into these two classes. Nowadays ML techniques offer fast and efficient solutions for many problems that would require intensive computational resources via traditional methods. This project aimed at providing a classification method to automatically label stellar particles in simulated central galaxies solely basing the decision on a set of specific features for each star particle. After subtracting cluster member galaxies, it is possible to identify two dynamically distinct stellar components in galaxy clusters. These are interpreted as the individual distinct properties of the stars bound to the BCG and the ICL, a diffuse

stellar envelope bound to the entire cluster’s potential. Both the spatial extent and luminosity curves of the two components smoothly merge, leaving no trace of the transition. Therefore, the separation of the ICL from the BCG can be performed in several (but particularly laborious) ways. We trained and tested a supervised Random Forest algorithm based on the matched catalogues of the star particles in the simulated cluster population from the DIANOGA set. These catalogues were constructed using an already developed algorithm (Dolag et al., 2010) which identifies the two populations on the account of the stellar particle dynamical features. One essential benefit of using the Random Forest is that the estimated computational cost does not significantly depend on the dataset size. Additionally, the predictions are straightforward to interpret while it is also extremely easy to measure the relative importance of each feature in the predictions. This project turned out to be instrumental in the finalisation of the study on the BCG velocity dispersion (see Chapter 6) since it allowed us to characterise the stellar population of the central regions of galaxy clusters in different fashions: namely, by including or not the ICL. The model is found to correctly identify most of the stars, while the larger errors are exhibited at the BCG outskirts, where the differences between the physical properties of the two components are less obvious. We find that our classifier provides consistent results in simulations for clusters at redshift $z < 1$, using different numerical resolutions and with implementation for the feedback from SNe and AGNs. Our findings are reported in Marini et al. (2022).

Detecting intracluster light in mock images of galaxy clusters

Chapter 8 was centred on studying the dynamics of the central galaxy (i.e., BCG) and the diffuse stellar light (i.e., ICL) in clusters to develop an algorithm to identify the latter in observed images of galaxy clusters. Recent observational studies have gathered a large volume of photometric (and spectroscopic) data on the stellar component in the innermost region of clusters. Nevertheless, the techniques to identify this component are significantly limited to modelling light profiles, masking the BCG or the use of a flux limit threshold. These methods are not only often inaccurate, but they also may provide results on the ICL fraction which generally differ from one another. On the contrary, in simulations, we know the full six-dimensional phase-space information for each particle. By exploiting this advantage, in Chapter 7 we showed how to classify particles bound to the main halo (i.e., BCG) or not (ICL) solely basing the decision on their dynamical properties. Such an objective classification is used to construct synthetic photometric and spectroscopic maps of the stellar distribution in the main halo which can be employed as input for a DL algorithm to identify in a real image such distributions. To this purpose, we selected a sample of clusters (among the most massive) from the DIANOGA zoom-in cosmological hydrodynamical simulations, which we used to create mock images of galaxy clusters. From the original dataset, we performed data augmentation by modifying (e.g., resizing, zooming, rotating) the images

and we included a basic treatment of contaminants (i.e., interlopers) by randomly masking elliptic regions in the image. We created these images for different observables of the stellar components (magnitudes/colours from the SDSS bandpasses, line of sight velocity and velocity dispersion). We trained three different models for the three observables, using the same DL architecture (a U-Net model, Ronneberger et al. 2015). Training is done by employing early stopping and Adam optimiser to prevent overfitting, preceded by a search for the best learning rate parameter. The analysis has shown the velocity map to be the best candidate for input parameters in the task. We found a significant tension in the ICL/BCG build-up as a function of the numerical resolution of the simulation: higher resolutions correspond to larger stellar halos. Possible factors responsible for the change in ICL and BCG assembly are most probably higher star formation rates and a larger number of substructures, whereas a potential smaller contribution may come from the AGN feedback modelling. Results will be published in Marini et al. (2023, in preparation).

9.2 FUTURE PERSPECTIVES

There is no doubt that the comparison with observational data represents a necessary step when analysing theoretical results as it helps benchmark our predictions. We expect to further constrain models of structure formation as provided by cosmological simulations disposing of more observational measurements. As far as this concerns, in this thesis we addressed several points of tensions existing between observations and simulations: e.g., the phase-space of member galaxies and the BCG assembly (also considering its correlation with ICL). Although observational measurements of the diffuse ICL components are quite challenging, due to the low-surface brightness and contamination from other sources, if we could have easy access to this information, we would be able to further assess the outcomes of our analyses. This would also lead to a better understanding of how BCG and ICL stellar components form and constrain their dynamical features more stringently.

At the current stage, there is an increasing urgency to address at least some of these issues. In light of the work done during this PhD thesis, we outline a few key points we consider of major importance.

Tension with cluster galaxies dynamics. Galaxy cluster substructures are notoriously some of the best indicators of the dynamical state (Biffi et al., 2016) and accretion history (Dolag et al., 2009) of their host. Far more than that, they provide a direct link between observables in simulated DM substructures with those of the galaxies that orbit in clusters, the galaxy cluster population. In this work, we provided many examples to show the existing tensions among several observed and predicted properties (e.g., phase-space properties, quenching efficiencies). To these must be added the vast literature on the topic which has recently raised debate even on the small scale distribution of DM (Meneghetti et al., 2020). We encourage a joint

effort between observers and simulators in addressing the intrinsic systematic each method carries and the enforcement of fair comparisons (e.g., the inclusion of projection effects, assessing the diversity of sub-resolution physical processes included in simulations and numerical resolution). To this end, few scientific questions shall be urged.

How can we competitively describe the phase-space properties of simulated galaxy clusters compared to the observed ones? Which are the physical processes that drive the differences across the galaxy population properties? Why are quenched galaxies overly predominant in simulations? What is the fate of orbiting substructures? Can the orbit (more radiant or tangential) of a galaxy determine its evolution?

The stellar halo build-up. As we have shown, the assembly histories of BCG and ICL are intertwined, as these two components often share similar histories. In our studies, the dynamics predicted in the central regions is fairly comparable to the observed one, however, we still encounter many difficulties in truly recovering other combined properties, such as the stellar mass (in [Chapter 5](#)) or the colour distribution (see [Chapter 8](#)). To further complicate the scenario, we found that numerical resolution affects significantly our findings on the stellar build-up. All combined it presses us for further research on the topic.

How can we recover the observed values of stellar masses of BCG and ICL in simulations? Are we taking into account all observational effects (e.g., profile fitting to recover masses, interlopers)? Can we quantify the impact of numerical resolution in our simulations? Can we exploit the ICL properties to reconstruct the cluster history? Can the ICL be used to measure the mass accretion history in clusters? Which formation mechanism is main driver in the ICL formation?

Modelling of the photometric properties. There are still several uncertainties in the modelling of the photometric properties of the stellar population in galaxy clusters. Additionally, given the significant fraction of quenched galaxies, there is nearly an absence of a star forming blue galaxy population in clusters. While this indicates that our model of galaxy formation in simulations needs to be improved, in any case we need to ensure that the computation of the photometric properties is accurate. This point will aim to answer some of the following scientific questions.

Why can we not recover the photometric properties of a statistical sample of galaxies in clusters? Why are the simulated galaxies so much redder and more passive than the observed ones? Why at the same time the simulated BCGs are too blue and with an excess of star formation? What is the impact of dust? Can we recover the colour gradients observed in the ICL?

In conclusion, in the long quest of outlining the main ingredients for the formation and evolution of structures in the Universe, high-resolution cosmological simulations take us one step further ahead. Indeed, it is key to grasp the intrinsic caveats they carry and understand what improvements we may apply to resolve part of the possible tensions. Currently, we are still far from having a complete description of the key physical and astrophysical processes determining the observational processes of galaxy clusters. From the observational side, we expect significant advancements in the forthcoming years from the upcoming ground-based (e.g. LSST) and space (e.g. Euclid and eROSITA) telescopes which will provide observational data of unprecedented quality and quantity.

In the standard cosmological model, structure formation is addressed in terms of the growth of the small perturbations in the initial cosmic density field. In this picture, structures – such as galaxies and clusters of galaxies – are thought to have formed in an expanding Universe due to the gravitational force acting on small density perturbations generated in the primordial epochs (Mo et al., 2010). Although the scientific community has not yet arrived at a common agreement on how these fluctuations may have been generated, the most accredited model is based on the idea that if the very early Universe went through an inflationary phase (namely an exponential expansion – we refer to the following studies Guth, 1998; Linde, 2005), which may have generated quantum fluctuations that, in turn, led to some energy density fluctuations. The latter is usually described by a Gaussian random field completely characterised by its second momentum, the power spectrum $P(k)$. Thanks to the precise measurements of the temperature anisotropies obtained with the Planck telescope (Tauber et al., 2004), as of today we know the initial conditions of the Universe very precisely until the recombination epoch. In [Figure A.1](#) it is shown the latest temperature (i.e. CMB) map obtained with the Planck telescope, from which one can measure the power spectrum: the different colours trace differences in temperature as reported by the legend.

In this chapter, we will describe the evolution of cosmological perturbations. When the density fluctuations are small, or in other words when the density contrast relative to the background $\delta = \delta\rho/\rho$ is small, we can adopt a Newtonian perturbation approach which will be extensively discussed in [Section A.2](#). The latter is valid when considering density variations on scales smaller than the Hubble length $c/H(t)$ and for a weak gravitational field. When $\delta(\vec{x}, t)$ is comparable to unity, the linear theory breaks down and it is required to use non-linear methods which are difficult to handle analytically, in this regard we will briefly examine the case of the spherical collapse in [Section A.4](#). The ultimate stage of such collapse is a configuration of equilibrium in which matter settles: this is the time when structures are governed by collisionless dynamics. It can be demonstrated that galaxies and DM halos are collisionless systems, in [Section A.5](#) will then present a few consequences of this statement.

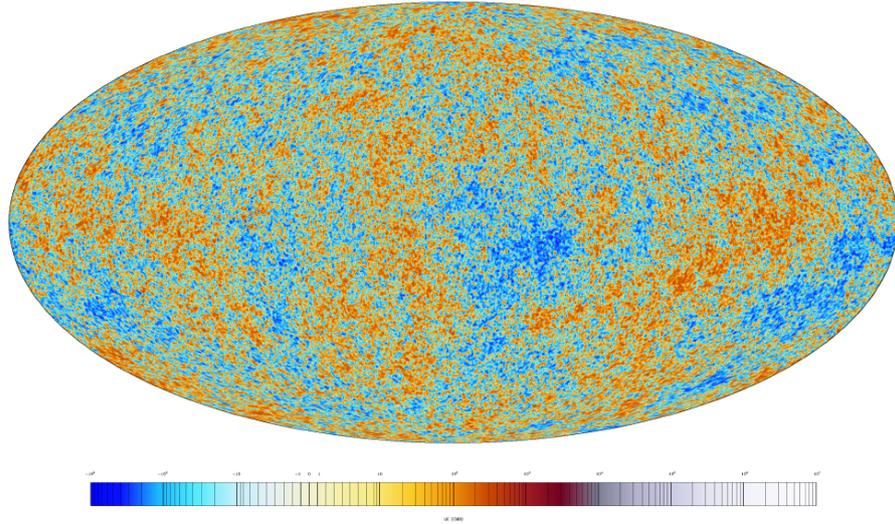


Figure A.1: Map of the CMB from the Planck telescope. Colour differences correspond to differences in temperature. Credit: ESA and Planck collaboration

A.1 STATISTICAL PROPERTIES

Let $\rho(\vec{x}, t)$ be the density and $\bar{\rho}(t)$ the average density of the universe. Structures are thought to form via peaks collapse in the $\rho(\vec{x}, t)$ characterised by the *density contrast field*:

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}. \quad (53)$$

In the linear regime, when the equations governing the evolution of the perturbations are all linear in perturbation quantities, the latter can be recast as:

$$\begin{aligned} \delta(\vec{x}, t) &= \sum_{\vec{k}} \delta_{\vec{k}}(t) \exp(i\vec{k} \cdot \vec{x}) \\ \delta_{\vec{k}}(t) &= \frac{1}{V_u} \int \delta(\vec{x}, t) \exp(-i\vec{k} \cdot \vec{x}) \end{aligned} \quad (54)$$

where V_u is the volume of a large box on which perturbations are assumed periodic and $\delta_{\vec{k}}$ are the Fourier coefficients. It is straightforward to see the impossibility of specifying all the infinite number of Fourier modes \vec{k} defining the cosmic density field. Yet, we can approach the problem with statistics. For a given cosmology, the cosmic density field is believed to be a uniform and isotropic Gaussian field generated by some random processes, thus model tests should be based on seeking statistical properties of the cosmic density field rather than on matching the predicted and observed field point by point. Upon this prescription, such field is fully characterised by its second momenta $P(k)$,

$$P(k) \equiv V_u \langle |\delta_{\vec{k}}|^2 \rangle \equiv V_u \langle \delta_{\vec{k}} \delta_{-\vec{k}} \rangle \quad (55)$$

notoriously known as the *power spectrum*. We will consider δ to be the realisation of a random process which is described by a probability distribution function (we then seek its properties, rather than the exact form of δ).

A.2 LINEAR EVOLUTION OF DENSITY PERTURBATIONS

To explicitly treat the case of evolution in an expanding FLRW universe, let us describe the system in comoving coordinates \vec{x} . Physical coordinates will transform upon the following:

$$\vec{r} = a(t)\vec{x}, \quad (56)$$

where $a(t)$ is the scale factor; accordingly the physical velocity \vec{u} will transform:

$$\vec{u} = \dot{a}(t)\vec{x} + \vec{v} \quad (57)$$

where \vec{v} is the peculiar velocity of a fluid element that detached from the Hubble flow for an observer comoving with the background and the dot labels the time derivative. Therefore, we can safely say that the contributions to the velocity of a fluid element come from the expansion of the universe (Hubble flow) and from the deviations to this. We thus find the *evolution of cosmological perturbations* to be described by:

$$\frac{\partial^2 \delta}{\partial t^2} + 2 \frac{\dot{a}}{a} \frac{\partial \delta}{\partial t} = 4\pi G \bar{\rho} \delta - \frac{c_s^2}{a^2} \nabla^2 \delta + \frac{2}{3} \frac{T}{a^2} \nabla^2 S. \quad (58)$$

On the right-hand side of the equation, there are listed all the terms contributing to the growth of the fluctuations: the gravitational force accountable for the gravitational collapse and two pressure terms. If we assume the fluid to be an ideal non-relativistic gas, the equation of state becomes $P \propto \rho^{5/3} \exp(kS)$, with k constant, and it will depend only on the density ρ and the specific entropy S of the fluid. Thus, the pressure term carries the contributions of two possible types of perturbations: *adiabatic* and *entropic*. The first one corresponds to fluctuations in the form of the local equation of state of the system, e.g. fluctuations in the relative number densities of the different particle types present in the system, while the second is associated with fluctuations in its energy density. Moreover, the entropy perturbations are also called isocurvature, given that the total density of the system remains homogeneous. In contrast, the adiabatic ones are also known as curvature perturbations, as they induce inhomogeneities in the spatial curvature.

The second term on the left-hand side is the Hubble drag term which is responsible for suppressing the growth due to the expansion of the universe. To have a general idea of the phenomena, let us imagine to closely follow a perturbation in the background density field from when it originated. The perturbation will go through an instability driven by the self-gravity of the region and the tendency to collapse is resisted by

the internal pressure gradient. Moreover, depending on the cosmology of the universe, the growth will experience slow damping due to the Hubble drag term: the future events will then depend on how effective is the damping force on the fluctuations. To indulge further in the subject we refer to Mo et al. (2010).

Linear regime

In the linear regime (i.e., $\delta \simeq 1$), all the equations can be expanded to first order quantities. Moreover, if the curvature of the universe can be neglected, which is true for a flat universe or when the scales of interest are much smaller than the horizon size, then we can select the mode functions to be plane waves and the perturbation fields can be written by means of their Fourier transform as in Equation 54. This will also allow us to convert the partial temporal derivative in total. Therefore, the time evolution of $\delta_{\vec{k}}$, at temperature \bar{T} , is:

$$\frac{d^2\delta_{\vec{k}}}{dt^2} + 2\frac{\dot{a}}{a}\frac{d\delta_{\vec{k}}}{dt} = \left[4\pi G\bar{\rho} - \frac{k^2 c_s^2}{a^2}\right]\delta_{\vec{k}} - \frac{2}{3}\frac{\bar{T}}{a^2}k^2 S_{\vec{k}}. \quad (59)$$

The advantage of expressing it this way is that all modes \vec{k} grow at the same rate albeit independently from one another.

The description of how these fluctuations grow, according to Equation 59, is given by the *linear growth factor* $D_+(a)$ when we can neglect the contributions from baryons, specifically for perturbations in a collisionless fluid or for isentropic perturbations well above the Jeans scale (introduced in Section A.3),

$$\delta(a) \propto D_+(a) = \frac{5\Omega_M}{2} E(a) \int_0^a \frac{da'}{[a'E(a')]^3} \quad (60)$$

where we assume

$$E(a) = \frac{H(a)}{H_0} = [\Omega_M a^{-3} + (1 + \Omega_M - \Omega_\Lambda)a^{-2} + \Omega_\Lambda]^{1/2}. \quad (61)$$

Here, all the contributions coming from neutrinos and relativistic particles are neglected (for other and more general forms of $D_+(a)$ see Percival, 2005). The growth factor $D_+(a)$ is commonly normalised either to unity at $a = 1$, or such that it rises proportional to a for $a \ll 1$. We stress the effective dependence on the cosmological parameters for the evolution of the fluctuation. In Figure A.2 is plotted the growth factor as a function of redshift z , normalised to unity today and divided by a . In an Einstein-de Sitter model, $D_+(a) = a$, thus the curves show how much earlier structures grow in the five assumed model universes compared to an Einstein-de Sitter.

A.3 GRAVITATIONAL INSTABILITY

The gravitational collapse follows the action of gravity on a self-gravitating system. To investigate on which scales the final collapse cannot be prevented anymore, let us restrict the study to the event of adiabatic perturbations only: the initial perturbations are isentropic ($k^2 S_{\vec{k}} = 0$) and far more the

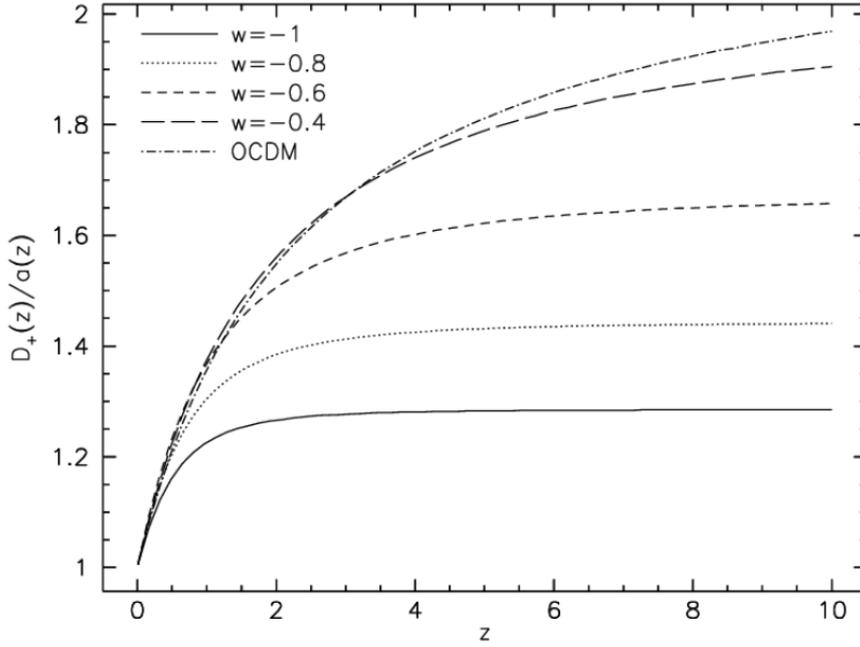


Figure A.2: Growth factor $D_+(z)$ as a function of redshift for five different cosmological models as indicated. The growth factor is normalised to unity at the present epoch, and divided by the scale factor to emphasise the differences between the models. With increasing ω , the growth factor increases towards its value for the open model with Ω_M . Credits: Bartelmann et al. (2002)

evolution is adiabatic. The differential equation for gravitational instability in a static medium is then obtained by setting $\dot{a} = 0$ in Equation 59. In this case, the evolution of perturbations is described by:

$$\frac{d^2\delta_{\vec{k}}}{dt^2} = \left[4\pi G\bar{\rho} - \frac{k^2 c_s^2}{a^2} \right] \delta_{\vec{k}} \quad (62)$$

which resembles the equation for a harmonic oscillator with a frequency $\omega^2 = \frac{k^2 c_s^2}{a^2} - 4\pi G\bar{\rho}$, so we can write:

$$\ddot{\delta}_{\vec{k}} = -\omega^2 \delta_{\vec{k}}.$$

According to the sign of ω^2 , the solution of this equation can be very different:

$\omega^2 > 0$ The solution is given by $\delta_{\vec{k}} \propto \exp(\pm i\omega t)$ which is an acoustic wave that propagates with the speed of sound;

$\omega^2 < 0$ The solution is a non-propagating stationary wave with an exponential decreasing (or increasing) mode with time, such as $\delta_{\vec{k}} \propto \exp(\pm\omega t)$.

The boundary of the two regimes is for $\omega^2 = 0$ which defines a characteristic proper length, the *Jeans length*,

$$\lambda_J = \frac{2\pi a}{k_J} = c_s \sqrt{\frac{\pi}{G\bar{\rho}}} \quad (63)$$

expressing the distance a sound wave can travel in a gravitational free-fall time $t_{\text{ff}} \sim (G\bar{\rho})^{-1/2}$. Any perturbation with a proper length larger than λ_J will belong to the second scenario listed above ($\omega^2 < 0$) where the pressure can no longer support gravity and the growing mode will reflect the gravitational (or Jeans) instability.

A.4 SPHERICAL COLLAPSE MODEL

Non-linear regime

The description of structure formation when $\delta \gg 1$ (which is the case for galaxies and clusters of galaxies) requires going beyond the linear approximation and introducing new tools to carry on the computation. The simplest model one can adopt is the spherical collapse model: the initial perturbation $\delta_i > 0$ is shaped as a sphere with radius R_i and enclosed mass $M_i = \frac{4\pi}{3}(1 + \delta_i)\bar{\rho}R_i^3$. In this scenario, the radius of the shell initially expands with the Hubble flow, but the expansion will slow down with time, turn over and eventually collapse due to gravitational instability. Therefore, we can distinguish two different phases in the evolution of perturbations: a decelerated expansion that ends at the time of turn around t_{ta} – which also coincides with the epoch of maximum expansion for the sphere – and the inevitable collapse into the virialised configuration at t_{vir} .

To do the maths, let us consider the mentioned perturbation to evolve with the background universe (for simplicity, we will consider an Einstein-de Sitter model where $\Omega_M = 1$ and $\Omega_\Lambda = 0$) in such a way that the peculiar velocity at the surface of the sphere vanishes. The linear growth of perturbations is a superposition of a growing δ_+ and a decaying δ_- mode. The combination of growing and decreasing modes is necessary to satisfy the correct boundary condition on the velocity: $v_i = 0$ (given by $v_i = i\delta'_k/k$) which requires that $\delta_+(t_i) = \frac{3}{5}\delta_i$.

The time interval between the initial time t_i and the epoch of turn around t_{ta} only depends on the free-fall time scale which is the characteristic time for a body to collapse under its gravitational potential. The perturbation will then have a density equal to:

$$\rho_p(t_{\text{ta}}) = \frac{3\pi}{32Gt_{\text{ta}}^2}.$$

On the other hand, assuming an Einstein-de Sitter model, the unperturbed background universe has density defined as:

$$\rho(t) = \frac{1}{6\pi Gt^2} \quad (64)$$

from which it follows that the density contrast Δ_{ta} at the epoch of turn-around reads

$$\Delta_{\text{ta}} = 1 + \delta_{\text{ta}} = \frac{\rho_p(t_{\text{ta}})}{\rho(t_{\text{ta}})} = \left(\frac{3\pi}{4}\right)^2 \simeq 5.6. \quad (65)$$

Interestingly enough, we can produce the same calculation if we consider the linear case. The linear growth of perturbations has a growing δ_+ and a decaying δ_- mode. However, after a short time, one can assume the decaying mode to be negligible, thus we are left with

$$\begin{aligned}\delta_+(t_{\text{ta}}) &= \delta_+(t_i) \left(\frac{t_{\text{ta}}}{t_i} \right)^{2/3} \\ &= \delta_+(t_i) \left(\frac{3\pi}{4} \right)^{2/3} \frac{\Omega_p(t_i)}{\delta_i} \\ &\simeq \frac{3}{5} \left(\frac{3\pi}{4} \right)^{2/3} \\ &\simeq 1.07\end{aligned}$$

The virial theorem states a selected ratio between the radius at the time of turn around R_{ta} and that of the time at the collapse R_{vir} , namely $R_{\text{ta}} = 2R_{\text{vir}}$. From this, it follows $\rho_p(t_{\text{vir}}) = 8\rho_p(t_{\text{ta}})$, which can be used to compute the density contrast Δ_{vir} once the structure is formed:

$$\Delta_{\text{vir}} = \frac{\rho_p(t_{\text{vir}})}{\rho(t_{\text{ta}})} = 8 \left(\frac{3\pi}{4} \right)^2 = 18\pi^2 \simeq 178. \quad (66)$$

Other cosmologies give different results, but the calculation is estimated similarly. For lower Ω_M , the density contrast is larger: this is because given the same initial radius R_i , the mass enclosed is smaller and so is the physical density, which induces longer times to collapse: e.g., a Λ CDM model has a final density contrast of roughly 350. Bryan and Norman (1998) found an expression for Δ_{vir} in function of the cosmology ($\Omega_K = 0$) given by:

$$\Delta_{\text{vir}}(x) = 18\pi^2 + 82x - 39x^2 \quad (67)$$

where we have

$$x = \Omega_M(z) - 1.$$

Curiously enough, we can carry out the same analysis for the initial density contrast δ_i as if it could grow only at the linear rate $D_+(a)$. In this case, the density contrast would be hundreds of times smaller at the turnaround epoch reaching the value of $\delta_{\text{ta}} = 1.062$ for an Einstein-de Sitter universe. This partially motivates why it is commonly said that the linear theory breaks down when $\delta \sim 1$. Furthermore, we see that for a mass shell to collapse at time t_{ta} , the overdensity within it, linearly extrapolated to the collapse time, must be $\delta_{\text{ta}} = 1.68 = \delta_c$. This defines the critical overdensity δ_c that gives us powerful insights on the potential formation of structures: a region with such overdensity against the background will most probably form a virialised object. Often used definitions in practice are R_{vir} – the halo radius that includes an average contrast density of Δ_{vir} for a given cosmology – and M_{vir} which corresponds to the mass encompassed by R_{vir} . In literature, these definitions, and other similar ones, are used as a

proxy to define virialised regions. More generally, we denote R_Δ and M_Δ , respectively the radius and mass enclosed in a region of density contrast Δ with respect to $\delta_c(z)$. In Figure A.3 is shown the density contrast in function of the cosmological parameter Ω_M .

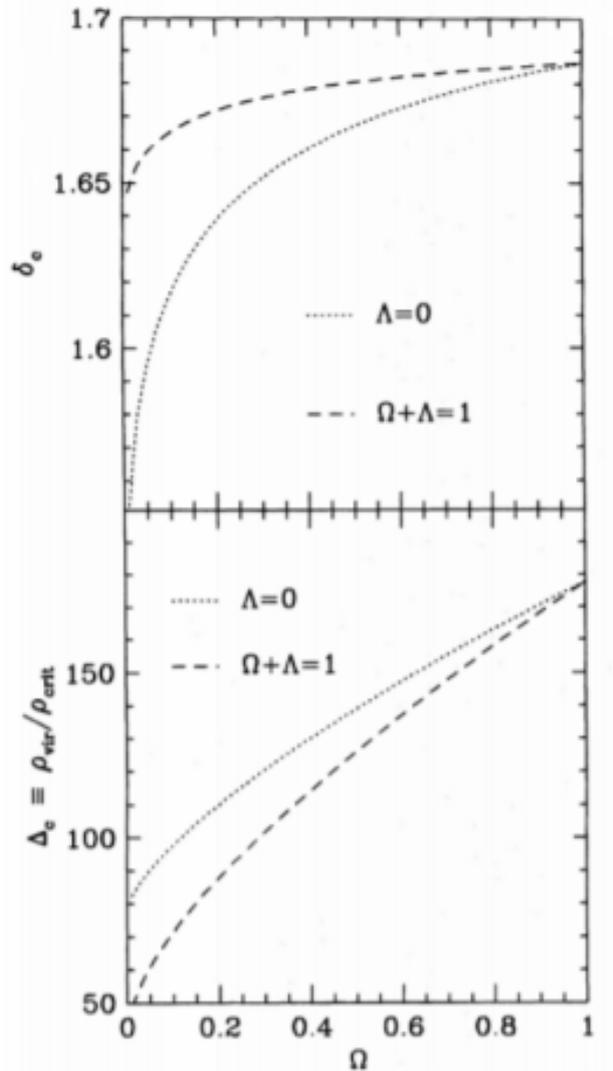


Figure A.3: Upper panel: critical threshold for collapse, δ_c as a function of Ω_M in the spherical collapse model. Results are plotted for open models with $\Omega_\Lambda = 0$ (dotted line) and flat models with $\Omega_M + \Omega_\Lambda = 1$ (dashed line). Lower panel: the virial density of collapsed objects in units of the critical density; the dotted and dashed lines are as in the upper panel. Credits: Eke et al. (1996)

A.5 BASICS ON DYNAMICS OF A COLLISIONLESS SELF-GRAVITATING SYSTEM

The final stage of gravitational collapse is a system in dynamical equilibrium. It can be demonstrated that the dynamics governing galaxies and

DM inside clusters of galaxies can be approximated to a fluid consisting of collisionless particles fully described by the Boltzmann and the Poisson equations. This is true for groups of particles that fulfil certain dynamical characteristics written in terms of time scales. A proto-cluster is said to have reached a state of equilibrium if the following chain is satisfied:

$$t_{\text{coll}} \gg t_{\text{relax}} \gg t_{\text{H}} \gg t_{\text{cross}}$$

where:

- t_{coll} time for which a particle experiences a collision;
- t_{relax} time scale for a particle, having encountered many others, to feel a cumulative effect on its velocity;
- t_{H} Hubble time;
- t_{cross} expected time for a particle to cross entirely the halo.

A.5.1 The Vlasov equation

Let us consider a large number of particles of the same mass m moving under the influence of a smooth potential field $\phi(\vec{x}, t)$. A complete description of the system is given by its continuous distribution function $f(\vec{x}, \vec{v}, t)$, which describes the number density of particles in the phase-space $\vec{\omega} = (\vec{x}, \vec{v})$. In this scheme, the density structure reads:

$$\rho = m \int f(\vec{x}, \vec{v}, t) d\vec{v} \quad (68)$$

To find the dynamical equation for the distribution function, we assume that the flow of matter through the phase-space is governed by the smooth six-dimensional vector field:

$$(\dot{\vec{x}}, \dot{\vec{v}}) = (\vec{v}, -\vec{\nabla}\phi) = \dot{\vec{\omega}}. \quad (69)$$

With collisionless matter, we can apply the mass conservation principle which is illustrated in the Liouville theorem by the following expression

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial}{\partial \omega_i} (\dot{\omega}_i f) = 0, \quad (70)$$

that can be recast in explicit coordinates

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla} f - \vec{\nabla} \phi \cdot \frac{\partial f}{\partial \vec{v}} = 0. \quad (71)$$

This is the *Vlasov equation*, also known as the *collisionless Boltzmann equation*. In (Equation 70) we note that df/dt expresses the Lagrangian derivative of the distribution function, so if it is equal to 0, the flow of objects through phase-space is incompressible, moreover, the phase-space density around

a given particle remains constant. To recover the dynamics of the system, one needs to associate the Vlasov equation with the Poisson

$$\nabla^2\phi(\vec{x}, t) = 4\pi G\rho(\vec{r}, t) \quad (72)$$

which links the gravitational potential ϕ to the underlying matter distribution ρ .

A.5.2 Jeans equation

To obtain some physical information from equation [Equation 71](#), let us consider the velocity moments of the distribution function f . If we integrate it over the three-dimensional velocity \vec{v} then

$$\int \frac{\partial f}{\partial t} d\vec{v} + \int v_i \frac{\partial f}{\partial x_i} d\vec{v} - \int \frac{\partial \phi}{\partial x_i} \frac{\partial f}{\partial v_i} d\vec{v} = 0. \quad (73)$$

The integration on the velocity does not affect the terms that only depend on the time t and the spatial \vec{x} coordinates as they are independent of the phase-space. This implies that we can extract the time derivative from the first term on the left-hand side and the spatial derivative on the second, as follows:

$$\frac{\partial}{\partial t} \int f d\vec{v} + \frac{\partial}{\partial x_i} \int v_i f d\vec{v} - \frac{\partial \phi}{\partial x_i} \int \frac{\partial f}{\partial v_i} d\vec{v} = 0. \quad (74)$$

On the other hand, we can apply the divergence theorem (i.e. $\int_{\Omega} \vec{\nabla} f dV = \oint_{\partial\Omega} f \vec{n} \cdot \vec{S}$, where \vec{n} is a versor normal to the surface S and the considered volume Ω is determined in the velocity space) on the last term. This term vanishes as there are no particles with infinite velocity. Thus, if we define the spatial and velocity density respectively as

$$\nu = \int f d\vec{v} \quad (75)$$

$$\bar{v}_i = \int f v_i d\vec{v} \quad (76)$$

the final expression for [Equation 71](#) reads:

$$\frac{\partial \nu}{\partial t} + \frac{\partial(\nu \bar{v}_i)}{\partial x_i} = 0 \quad (77)$$

Multiplying the latter for v_j , integrating over the velocities and applying the divergence theorem once again, we obtain:

$$\frac{\partial(\nu \bar{v}_j)}{\partial t} + \frac{\partial(\nu \bar{v}_i \bar{v}_j)}{\partial x_i} + \nu \frac{\partial \phi}{\partial x_j} = 0 \quad (78)$$

where $\bar{v}_i \bar{v}_j = \Omega^{-1} \int v_i v_j f d\vec{v}$. With the change ([Equation 78](#)) $-\nu_j$ ([Equation 78](#)), we finally obtain the *Jeans equation*:

$$\frac{\partial(\bar{v}_j)}{\partial t} + \bar{v}_j \frac{\partial(\bar{v}_j)}{\partial x_i} = -\frac{\partial \phi}{\partial x_j} - \frac{1}{\nu} \frac{\partial(\nu \sigma_{ij}^2)}{\partial x_i}. \quad (79)$$

The stress tensor $\sigma_{ij}^2 = \overline{(v_i - \bar{v}_i)(v_j - \bar{v}_j)} = \overline{v_i v_j} - \bar{v}_i \bar{v}_j$ introduced in the equation expresses an anisotropic pressure and it is the equivalent of the pressure term in the Euler equation. Therefore, the velocity field of collisionless matter provides support to the system against the collapse in the same way the pressure does in a collisional system.

A.5.2.1 A special case: spherical symmetry

Let us consider the case of a spherical, steady-state dynamical model which causes the following terms in Equation 79 to vanish:

$$\bar{v}_r = \bar{v}_\theta = \bar{v}_\phi = 0 \quad (80)$$

$$\overline{v_\theta^2} = \overline{v_\phi^2} \quad (81)$$

$$\overline{v_r v_\theta} = \overline{v_\phi v_\theta} = \overline{v_r v_\phi} = 0 \quad (82)$$

up to second-order approximation. The Jeans equation transforms,

$$\vec{\nabla}(\nu \vec{v} \cdot \vec{v}) + n \vec{\nabla} \phi = 0 \quad (83)$$

which can be recast in spherical coordinates to read,

$$\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \overline{v_r \vec{v} \cdot \nu}) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \overline{v_\theta \vec{v} \cdot \nu}) = -\nu \vec{\nabla} \phi. \quad (84)$$

Adopting the symmetries that hold in this particular case and using the properties of unit vectors, the only non-vanishing equation is:

$$\frac{1}{\rho} \frac{d}{dr} (\rho \overline{v_r^2}) + 2\beta \frac{\overline{v_r^2}}{r^2} = -\frac{d\phi}{dr} \quad (85)$$

where β is the orbital anisotropy profile defined as $\beta = 1 - \overline{v_\theta^2}/\overline{v_r^2}$. Equation 85 carries still two unknowns which means that to solve it, assumptions on β or $\overline{v_r^2}$ are required. By taking the gradient of the gravitational potential as $\frac{d\phi}{dr} = \frac{GM(r)}{r^2}$, we get the mass M enclosed within a radius r :

$$M(< r) = -\frac{r\sigma_r^2}{G} \left(\frac{d \ln \nu}{d \ln r} + \frac{d \ln \sigma_r^2}{d \ln r} + 2\beta \right) \quad (86)$$

The parameter β describes the particles' orbit in the system:

$\beta < 0$ tangentially anisotropic;

$\beta = 0$ perfectly isotropic orbits;

$0 < \beta < 1$ radially anisotropic;

$\beta = 1$ all orbits are radial.

It has been observed that in relaxed clusters, the profile of the anisotropy parameters β is radially anisotropic (Biviano et al., 2013; Capasso et al., 2019).

A.5.3 Virial Theorem

The virial theorem has many applications in physics as it states the energy exchange in time of a system, e.g. in a galaxy or a cluster of galaxies. For our purposes, it can be derived from the equations of a collisionless system, like the one described above. From [Equation 78](#), let us consider the mass density as in [Equation 68](#) and integrate over space to read:

$$\int x_k \frac{\partial(\rho \bar{v}_j)}{\partial t} d\vec{x} = - \sum_i \int x_k \frac{\partial(\rho \bar{v}_j \bar{v}_i)}{\partial x_i} d\vec{x} - \int \rho x_k \frac{\partial \phi}{\partial x_j} d\vec{x}. \quad (87)$$

Applying the divergence theorem on the first term of the right-hand side of the latter, we can write:

$$\begin{aligned} \sum_i \int x_k \frac{\partial(\rho \bar{v}_j \bar{v}_i)}{\partial x_i} d\vec{x} &= \sum_i \int \left[\frac{\partial(x_k \rho \bar{v}_j \bar{v}_i)}{\partial x_i} + \rho \bar{v}_j \bar{v}_i \frac{\partial(x_k)}{\partial x_i} \right] d\vec{x} \\ &= \int \left[\sum_i \frac{\partial(x_k \rho \bar{v}_j \bar{v}_i)}{\partial x_i} + \rho \bar{v}_j \bar{v}_k \right] d\vec{x} \\ &= \oint \sum_i x_k \rho \bar{v}_j \bar{v}_i dS_i + \int \rho \bar{v}_j \bar{v}_k d\vec{x} \\ &= \oint \sum_i x_k \rho \bar{v}_j \bar{v}_i dS_i + \int \rho \bar{v}_j \bar{v}_k d\vec{x} + \int \rho \sigma_{jk}^2 d\vec{x} \\ &= \Sigma_{jk} + 2T_{jk} + \Pi_{jk} \end{aligned}$$

in the latter, we explicitly write the symmetrical and asymmetrical terms arising from $\bar{v}_j \bar{v}_i$. Each term is a different contribution to the energy:

$$K_{jk} = T_{jk} + \Pi_{jk}/2 \quad (88)$$

is the *kinetic energy tensor*, and

$$\Sigma_{jk} = - \oint \sum_i x_k \rho \bar{v}_j \bar{v}_i dS_i \quad (89)$$

is the *surface pressure term*. The last term on the right-hand side of [Equation 87](#) is the *Chandrasekhar potential energy tensor* W_{jk} . Therefore, it reads:

$$\int x_k \frac{\partial(\rho \bar{v}_j)}{\partial t} d\vec{x} = 2T_{jk} + \Pi_{jk} + \Sigma_{jk} + W_{jk}. \quad (90)$$

Using that the tensors on the right side of the latter are all symmetric with respect to their subscript, it is straightforward:

$$\frac{1}{2} \frac{d}{dt} \int \rho [x_k \bar{v}_j + x_j \bar{v}_k] d\vec{x} = 2T_{jk} + \Pi_{jk} + \Sigma_{jk} + W_{jk}. \quad (91)$$

The *moment of inertia tensor* is defined such as:

$$I_{jk} = \int \rho x_j x_k d\vec{x} \quad (92)$$

and by taking the time derivative

$$\begin{aligned}\frac{dI_{jk}}{dt} &= \int \frac{\partial \rho}{\partial t} x_j x_k d\vec{x} \\ &= - \int \frac{\partial(\rho \bar{v}_i)}{\partial x_i} x_j x_k d\vec{x} \\ &= \int \rho [x_k \bar{v}_j + x_j \bar{v}_k] d\vec{x}\end{aligned}$$

where we have used the continuity equation. Finally, we obtain:

$$\frac{1}{2} \frac{d^2 I_{jk}}{dt^2} = 2K_{jk} + \Sigma_{jk} + W_{jk} \quad (93)$$

which is the *virial theorem* in tensorial form. By taking the trace on both sides, we reduce to the more common form of the scalar virial equation:

$$\frac{1}{2} \frac{d^2 I}{dt^2} = 2K + \Sigma + W \quad (94)$$

where:

$$I = \text{Tr}(I_{jk}) = \int \rho r^2 d\vec{x}; \quad (95)$$

$$K = \text{Tr}(K_{jk}) = \frac{1}{2} \int \rho \bar{v}^2 d\vec{x}; \quad (96)$$

$$W = \text{Tr}(W_{jk}) = - \int \rho \vec{x} \cdot \vec{\nabla} \phi d\vec{x}; \quad (97)$$

$$\Sigma = \text{Tr}(\Sigma_{jk}) = - \int \rho \bar{v}^2 \vec{x} \cdot d\vec{S}. \quad (98)$$

If we consider a system for which the moment of inertia does not change across time (i.e. static), the term on the left-hand side vanishes, which leaves us with the contribution from kinetic and potential energies only plus the possible work done by an external pressure W ,

$$2K + W + \Sigma = 0. \quad (99)$$

In case there are no external forces, the expression reduces to:

$$E = -K = \frac{W}{2}. \quad (100)$$

A.5.3.1 A special case: spherical symmetry

Let us consider the case of a universe with a non-vanishing cosmological constant Λ , in which a spherical perturbation of mass M collapses to a point. At the moment of maximum expansion, the total energy retained by the system will be only potential,

$$E_{\text{TOT}} = -\frac{3GM^2}{5r_{\text{max}}}. \quad (101)$$

However, during the collapse, the system does not dissipate energy, since we are taking into consideration only collisionless particles, and the collapse will end once a quasi-static equilibrium is reached. If a uniform object is formed then the potential energy is simply

$$W = -\frac{3GM^2}{5r_{\text{vir}}} \quad (102)$$

therefore the virial theorem in [Equation 100](#) implies $r_{\text{max}} = 2 r_{\text{vir}}$.

ACKNOWLEDGEMENTS

This work would be incomplete without acknowledging the people that made it possible in the first place.

First and foremost, my greatest appreciation goes to my supervisors Stefano Borgani and Alex Saro. Their unwavering guidance and tireless patience guided me in carrying out this project. Far more than that, I am particularly grateful for the invaluable discussions and their constant support in following my own ideas during the realisation of this project: one can only hope to come across them both during their research path. Also, I would like to extend my sincere thanks to the large body of incredible researchers at the Observatory of Trieste who offered their help and knowledge when in need: Giuseppe Murante, Elena Rasia, Luca Tornatore, Cinthia Ragone-Figueroa, Gian Luigi Granato, Giuliano Taffoni, Luigi Bassini, Veronica Strazzullo and Maurilio Pannella. I wish to individually acknowledge Matteo Boi who helped me develop the machine learning part. He has been a kind proofreader and tutor to my scientific curiosities in this field, I hope our scientific paths will cross again.

Furthermore, I would like to express my deep gratitude to Klaus Dolag and his research team at the Ludwig Maximilians University in Munich where I have spent 5 months during the last year. His kindness and attention have not gone unnoticed, along with many cups of good coffee which I am grateful for. In this regard, special thanks go to his amazing research group that has patiently taught me many technical aspects connected with cosmological simulations, between one laugh and another.

Aside from the strict work environment, my sincere gratitude goes to Sam, my family, flatmates and friends for their unflinching support. This milestone would not be worth celebrating without them, in fact, it may have not even been reached. Each one of them has helped and encouraged me to passionately pursue this big starry dream, whether this was through a phone call, a message or in front of a pizza. In this regard, I would like to acknowledge the PhD crew in Trieste: they have unwittingly taught me to not underestimate the support and understanding that one student can provide to another just by being there day by day. Also, I know they might be the only ones who will read this acknowledgement. ;-)

Last but not least, a final comment on the drafting of the thesis. Most of the writing has taken place during an amazing Interrail tour around the United Kingdom in August 2022. I wish to individually thank my fellow travellers from the GB team (Sam, Veronica and Isa) for being there and definitely not trying to distract me while working. Admittedly, you made



Figure A.4: Recap of the itinerant writing of this thesis.

the writing a much better and more interesting experience.

For the interested party, you can find here a sketch of the route undertaken during the Interrail and some extracted statistics (after all, we do like some math, no?). Yes, I know it is in Spanish. And no, I do not know why. The gap between Glasgow and Fort William was covered by bus, due to an intermediate stop (Glencoe) not covered by the railway system. We also visited the Isle of Skye (top-left) hereby not marked.

BIBLIOGRAPHY

- Abbott, TMC et al. (2018). 'Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing'. In: *Physical Review D* 98.4, p. 043526.
- Adhikari, Susmita, Neal Dalal and Robert T Chamberlain (2014). 'Splashback in accreting dark matter halos'. In: *Journal of Cosmology and Astroparticle Physics* 2014.11, p. 019.
- Aguerri, JAL et al. (2017). 'Deep spectroscopy in nearby galaxy clusters—III. Orbital structure of galaxies in Abell 85'. In: *Monthly Notices of the Royal Astronomical Society* 468.1, pp. 364–377.
- Aguirre, A et al. (2001). 'Metal enrichment of the intergalactic medium at $z = 3$ by galactic winds'. In: *The Astrophysical Journal* 560.2, p. 599.
- Alamo-Martínez, KA and JP Blakeslee (2017). 'Specific frequencies and luminosity profiles of cluster galaxies and intracluster light in Abell 1689'. In: *The Astrophysical Journal* 849.1, p. 6.
- Allen, SW, AE Evrard and AB Mantz (2011). 'Cosmological parameters from observations of galaxy clusters'. In: *Annual Review of Astronomy and Astrophysics* 49, pp. 409–470.
- Alonso Asensio, Isaac et al. (2020). 'The intracluster light as a tracer of the total matter density distribution: a view from simulations'. In: *Monthly Notices of the Royal Astronomical Society* 494.2, pp. 1859–1864.
- Alzubaidi, Laith et al. (2021). 'Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions'. In: *Journal of big Data* 8.1, pp. 1–74.
- Annunziatella, Marianna et al. (2016). 'CLASH-VLT: Environment-driven evolution of galaxies in the $z = 0.209$ cluster Abell 209'. In: *Astronomy & Astrophysics* 585, A160.
- Arabadjis, JS, MW Bautz and G Arabadjis (2004). 'Extracting the dark matter profile of a relaxed galaxy cluster'. In: *The Astrophysical Journal* 617.1, p. 303.
- Armitage, Thomas J et al. (2018). 'The Cluster-EAGLE project: velocity bias and the velocity dispersion–mass relation of cluster galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 474.3, pp. 3746–3759.
- Arnaboldi, M et al. (1996). 'The kinematics of the planetary nebulae in the outer regions of NGC 4406'. In: *The Astrophysical Journal* 472.1, p. 145.
- Arnouts, Stephane et al. (2013). 'Encoding of the infrared excess in the NUVrK color diagram for star-forming galaxies'. In: *Astronomy & Astrophysics* 558, A67.
- Ascasibar, Yago et al. (2004). 'On the physical origin of dark matter density profiles'. In: *Monthly Notices of the Royal Astronomical Society* 352.4, pp. 1109–1120.

- Bahé, Yannick M et al. (2017). 'The Hydrangea simulations: galaxy formation in and around massive clusters'. In: *Monthly Notices of the Royal Astronomical Society* 470.4, pp. 4186–4208.
- Baldry, Ivan K et al. (2012). 'Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at $z < 0.06$ '. In: *Monthly Notices of the Royal Astronomical Society* 421.1, pp. 621–634.
- Balogh, Michael L, Julio F Navarro and Simon L Morris (2000). 'The origin of star formation gradients in rich galaxy clusters'. In: *The Astrophysical Journal* 540.1, p. 113.
- Barnes, Eric I et al. (2007). 'Density profiles of collisionless equilibria. II. Anisotropic spherical systems'. In: *The Astrophysical Journal* 654.2, p. 814.
- Barnes, J and P Hut (1986). 'A hierarchical O (N log N) force-calculation algorithm'. In: *nature* 324.6096, p. 446.
- Bartelmann, M, F Perrotta and C Baccigalupi (2002). 'Halo concentrations and weak-lensing number counts in dark energy cosmologies'. In: *Astronomy & Astrophysics* 396.1, pp. 21–30.
- Bassini, L et al. (2020). 'The DIANOGA simulations of galaxy clusters: characterising star formation in protoclusters'. In: *Astronomy & Astrophysics* 642, A37.
- Beers, Timothy C, Kevin Flynn and Karl Gebhardt (1990). 'Measures of location and scale for velocities in clusters of galaxies-A robust approach'. In: *The Astronomical Journal* 100, pp. 32–46.
- Bell, Eric F et al. (2003). 'The optical and near-infrared properties of galaxies. I. Luminosity and stellar mass functions'. In: *The Astrophysical Journal Supplement Series* 149.2, p. 289.
- Bellstedt, Sabine et al. (2016). 'The evolution in the stellar mass of brightest cluster galaxies over the past 10 billion years'. In: *Monthly Notices of the Royal Astronomical Society* 460.3, pp. 2862–2874.
- Bender, Ralf et al. (2015). 'Structure and Formation of cD Galaxies: NGC 6166 in ABELL 2199'. In: *The Astrophysical Journal* 807.1, p. 56.
- Benz, W (1990). 'Smooth particle hydrodynamics: a review'. In: *The numerical modelling of nonlinear stellar pulsations*. Springer, pp. 269–288.
- Bernardi, Mariangela (2009). 'Evolution in the structural properties of early-type brightest cluster galaxies at small lookback time and dependence on the environment'. In: *Monthly Notices of the Royal Astronomical Society* 395.3, pp. 1491–1506.
- Bertschinger, E (1985). 'Self-similar secondary infall and accretion in an Einstein-de Sitter universe'. In: *The Astrophysical Journal Supplement Series* 58, pp. 39–65.
- Biffi, Veronica et al. (2016). 'On the nature of hydrostatic equilibrium in galaxy clusters'. In: *The Astrophysical Journal* 827.2, p. 112.
- Binney, J and S Tremaine (1987). 'Galactic dynamics, ed'. In: *Binney, J. & Tremaine, S.*
- Bishop, Christopher M and Nasser M Nasrabadi (2006). *Pattern recognition and machine learning*. Vol. 4. 4. Springer.

- Biviano, A and M Girardi (2003). 'The Mass Profile of Galaxy Clusters out to $\sim 2r_{200}$ '. In: *The Astrophysical Journal* 585.1, p. 205.
- Biviano, A et al. (2013). 'CLASH-VLT: The mass, velocity-anisotropy, and pseudo-phase-space density profiles of the $z = 0.44$ galaxy cluster MACS J1206.2-0847'. In: *Astronomy & Astrophysics* 558, A1.
- Biviano, A et al. (2016). 'The dynamics of $z \sim 1$ clusters of galaxies from the GCLASS survey'. In: *Astronomy & Astrophysics* 594, A51.
- Biviano, A et al. (2021). 'The GOGREEN survey: Internal dynamics of clusters of galaxies at redshift 0.9–1.4'. In: *Astronomy & Astrophysics* 650, A105.
- Bondi, HJ (1952). 'On spherically symmetrical accretion'. In: *Monthly Notices of the Royal Astronomical Society* 112.2, pp. 195–204.
- Bondi, HJ and F Hoyle (1944). 'On the mechanism of accretion by stars'. In: *Monthly Notices of the Royal Astronomical Society* 104.5, pp. 273–282.
- Borgani, S and AV Kravtsov (2011). 'Cosmological simulations of galaxy clusters'. In: *Advanced Science Letters* 4.2, pp. 204–227.
- Borgani, S et al. (2008). 'The chemical enrichment of the ICM from hydrodynamical simulations'. In: *Space Science Reviews* 134.1-4, pp. 379–403.
- Borgani, Stefano et al. (1999). 'Velocity dispersions of CNOC clusters and the evolution of the cluster abundance'. In: *The Astrophysical Journal* 527.2, p. 561.
- Bosch, Frank C van den and Go Ogiya (2018). 'Dark matter substructure in numerical simulations: a tale of discreteness noise, runaway instabilities, and artificial disruption'. In: *Monthly Notices of the Royal Astronomical Society* 475.3, pp. 4066–4087.
- Bose, Sownak and Abraham Loeb (2021). 'Measuring the mass and concentration of dark matter halos from the velocity dispersion profile of their stars'. In: *The Astrophysical Journal* 912.2, p. 114.
- Boselli, Alessandro, Matteo Fossati and Ming Sun (2022). 'Ram pressure stripping in high-density environments'. In: *The Astronomy and Astrophysics Review* 30.1, pp. 1–127.
- Breiman, Leo (2001). 'Random forests'. In: *Machine learning* 45.1, pp. 5–32.
- Brighenti, F and WG Mathews (2006). 'Stopping cooling flows with jets'. In: *The Astrophysical Journal* 643.1, p. 120.
- Bruzual, G and Stephane Charlot (2003). 'Stellar population synthesis at the resolution of 2003'. In: *Monthly Notices of the Royal Astronomical Society* 344.4, pp. 1000–1028.
- Bryan, GL and ML Norman (1998). 'Statistical properties of X-ray clusters: Analytic and numerical comparisons'. In: *The Astrophysical Journal* 495.1, p. 80.
- Burke, Claire et al. (2012). 'Measurement of the intracluster light at $z \sim 1$ '. In: *Monthly Notices of the Royal Astronomical Society* 425.3, pp. 2058–2068.
- Burles, Scott and David Tytler (1998). 'The Deuterium abundance toward Q1937–1009'. In: *The Astrophysical Journal* 499.2, p. 699.

- Cañas, Rodrigo et al. (2020). 'From stellar haloes to intracluster light: the physics of the Intra-Halo Stellar Component in cosmological hydrodynamical simulations'. In: *Monthly Notices of the Royal Astronomical Society* 494.3, pp. 4314–4333.
- Capasso, R et al. (2019). 'Galaxy kinematics and mass calibration in massive SZE-selected galaxy clusters to $z = 1.3$ '. In: *Monthly Notices of the Royal Astronomical Society* 482.1, pp. 1043–1061.
- Cappellari, Michele and Eric Emsellem (2004). 'Parametric recovery of line-of-sight velocity distributions from absorption-line spectra of galaxies via penalized likelihood'. In: *Publications of the Astronomical Society of the Pacific* 116.816, p. 138.
- Chabrier, Gilles (2003). 'Galactic stellar and substellar initial mass function'. In: *Publications of the Astronomical Society of the Pacific* 115.809, p. 763.
- Chandrasekhar, S (1943). 'Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction.' In: *The Astrophysical Journal* 97, p. 255.
- Chen, Mandy C et al. (2018). 'A Likely Supermassive Black Hole Revealed by Its Einstein Radius in Hubble Frontier Fields Images'. In: *The Astrophysical Journal* 863.2, p. 135.
- Chen, Xiaokai et al. (2022). 'The sphere of influence of the bright central galaxies in the diffuse light of SDSS clusters'. In: *Monthly Notices of the Royal Astronomical Society* 514.2, pp. 2692–2706.
- Chiu, I et al. (2018). 'Baryon content in a sample of 91 galaxy clusters selected by the South Pole Telescope at $0.2 < z < 1.25$ '. In: *Monthly Notices of the Royal Astronomical Society* 478.3, pp. 3072–3099.
- Clowe, D et al. (2006). 'A direct empirical proof of the existence of dark matter'. In: *The Astrophysical Journal Letters* 648.2, p. L109.
- Cole, Shaun et al. (2001). 'The 2dF galaxy redshift survey: near-infrared galaxy luminosity functions'. In: *Monthly Notices of the Royal Astronomical Society* 326.1, pp. 255–273.
- Cole, Shaun et al. (2005). 'The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications'. In: *Monthly Notices of the Royal Astronomical Society* 362.2, pp. 505–534.
- Collaboration, Planck et al. (2020). 'Planck 2018 results. I. Overview and the cosmological legacy of Planck'. In:
- Contini, E, HZ Chen and Q Gu (2022). 'The Transition Region between Brightest Cluster Galaxies and Intracluster Light in Galaxy Groups and Clusters'. In: *The Astrophysical Journal* 928.2, p. 99.
- Contini, E, SK Yi and X Kang (2018). 'The different growth pathways of brightest cluster galaxies and intracluster light'. In: *Monthly Notices of the Royal Astronomical Society* 479.1, pp. 932–944.
- Contini, E et al. (2013). 'On the formation and physical properties of the intracluster light in hierarchical galaxy formation models'. In: *Monthly Notices of the Royal Astronomical Society* 437.4, pp. 3787–3802.
- Contini, Emanuele (2021). 'On the Origin and Evolution of the Intra-Cluster Light: A Brief Review of the Most Recent Developments'. In: *Galaxies* 9.3, p. 60.

- Contini, Emanuele and Q Gu (2020). 'On the Mass Distribution of the Intracluster Light in Galaxy Groups and Clusters'. In: *The Astrophysical Journal* 901.2, p. 128.
- (2021). 'Brightest Cluster Galaxies and Intracluster Light: Their Mass Distribution in the Innermost Regions of Groups and Clusters'. In: *The Astrophysical Journal* 915.2, p. 106.
- Contini, Emanuele et al. (2014). 'On the formation and physical properties of the intracluster light in hierarchical galaxy formation models'. In: *Monthly Notices of the Royal Astronomical Society* 437.4, pp. 3787–3802.
- Da Rocha, Cristiano and C Mendes De Oliveira (2005). 'Intragroup diffuse light in compact groups of galaxies: HCG 79, 88 and 95'. In: *Monthly Notices of the Royal Astronomical Society* 364.3, pp. 1069–1081.
- Daddi, E et al. (2004). 'A new photometric technique for the joint selection of star-forming and passive galaxies at $1.4 \leq z \leq 2.5$ '. In: *The Astrophysical Journal* 617.2, p. 746.
- De Lucia, G and J Blaizot (2007). 'The hierarchical formation of the brightest cluster galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 375.1, pp. 2–14.
- De Lucia, Gabriella, Michaela Hirschmann and Fabio Fontanot (2019). 'Nature versus nurture: what regulates star formation in satellite galaxies?' In: *Monthly Notices of the Royal Astronomical Society* 482.4, pp. 5041–5051.
- De Propriis, Roberto et al. (2021). 'Brightest cluster galaxies: the centre can (not?) hold'. In: *Monthly Notices of the Royal Astronomical Society* 500.1, pp. 310–318.
- DeMaio, Tahlia et al. (2018). 'Lost but not forgotten: intracluster light in galaxy groups and clusters'. In: *Monthly Notices of the Royal Astronomical Society* 474.3, pp. 3009–3031.
- DeMaio, Tahlia et al. (2020). 'The growth of brightest cluster galaxies and intracluster light over the past 10 billion years'. In: *Monthly Notices of the Royal Astronomical Society* 491.3, pp. 3751–3759.
- Dehnen, W and JI Read (2011). 'N-body simulations of gravitational dynamics'. In: *The European Physical Journal Plus* 126.5, p. 55.
- Dehnen, Walter and Dean E McLaughlin (2005). 'Dynamical insight into dark matter haloes'. In: *Monthly Notices of the Royal Astronomical Society* 363.4, pp. 1057–1068.
- Di Matteo, T, V Springel and L Hernquist (2005). 'Energy input from quasars regulates the growth and activity of black holes and their host galaxies'. In: *nature* 433.7026, p. 604.
- Di Matteo, T et al. (2008). 'Direct cosmological simulations of the growth of black holes and galaxies'. In: *The Astrophysical Journal* 676.1, p. 33.
- Diaferio, Antonaldo and Margaret J Geller (1997). 'Infall regions of galaxy clusters'. In: *The Astrophysical Journal* 481.2, p. 633.
- Diemand, J, B Moore and J Stadel (2004). 'Velocity and spatial biases in cold dark matter subhalo distributions'. In: *Monthly Notices of the Royal Astronomical Society* 352.2, pp. 535–546.

- Diemer, B and AV Kravtsov (2015). 'A universal model for halo concentrations'. In: *The Astrophysical Journal* 799.1, p. 108.
- Doherty, M et al. (2009). 'The edge of the M87 halo and the kinematics of the diffuse light in the Virgo cluster core'. In: *Astronomy & Astrophysics* 502.3, pp. 771–786.
- Dolag, K, G Murante and Stefano Borgani (2010). 'Dynamical difference between the cD galaxy and the diffuse, stellar component in simulated galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 405.3, pp. 1544–1559.
- Dolag, K et al. (2005). 'The imprints of local superclusters on the Sunyaev-Zel'dovich signals and their detectability with Planck'. In: *Monthly Notices of the Royal Astronomical Society* 363.1, pp. 29–39.
- Dolag, K et al. (2008). 'Simulation techniques for cosmological simulations'. In: *Space science reviews* 134.1-4, pp. 229–268.
- Dolag, K et al. (2009). 'Substructures in hydrodynamical cluster simulations'. In: *Monthly Notices of the Royal Astronomical Society* 399.2, pp. 497–514.
- Donnari, Martina et al. (2019). 'The star formation activity of IllustrisTNG galaxies: main sequence, UVJ diagram, quenched fractions, and systematics'. In: *Monthly Notices of the Royal Astronomical Society* 485.4, pp. 4817–4840.
- Donnari, Martina et al. (2021). 'Quenched fractions in the IllustrisTNG simulations: the roles of AGN feedback, environment, and pre-processing'. In: *Monthly Notices of the Royal Astronomical Society* 500.3, pp. 4004–4024.
- Doroshkevich, AG (1970). 'Spatial structure of perturbations and origin of galactic rotation in fluctuation theory'. In: *Astrophysics* 6.4, pp. 320–330.
- Dressler, A (1980). 'Galaxy morphology in rich clusters - Implications for the formation and evolution of galaxies'. In: *The Astrophysical Journal* 236, pp. 351–365.
- Dressler, Alan (1979). 'The dynamics and structure of the cD galaxy in Abell 2029'. In: *The Astrophysical Journal* 231, pp. 659–670.
- Dubinski, J (1998). 'The origin of the brightest cluster galaxies'. In: *The Astrophysical Journal* 502.1, p. 141.
- Dubinski, John (1996). 'A parallel tree code'. In: *New Astronomy* 1.2, pp. 133–147.
- Ebrova, Ivana (2013). 'Shell galaxies: kinematical signature of shells, satellite galaxy disruption and dynamical friction'. In: *arXiv preprint:1312.1643*.
- Edwards, Louise OV et al. (2016). 'Stellar populations of BCGs, close companions and intracluster light in Abell 85, Abell 2457 and IIZw108'. In: *Monthly Notices of the Royal Astronomical Society* 461.1, pp. 230–239.
- Eke, VR, S Cole and CS Frenk (1996). 'Cluster evolution as a diagnostic for Ω '. In: *Monthly Notices of the Royal Astronomical Society* 282.1, pp. 263–280.
- Ellien, A et al. (2021). 'DAWIS: a detection algorithm with wavelets for intracluster light studies'. In: *Astronomy & Astrophysics* 649, A38.

- Erfanianfar, G et al. (2019). 'Stellar mass–halo mass relation for the brightest central galaxies of X-ray clusters since $z \sim 0.65$ '. In: *Astronomy & Astrophysics* 631, A175.
- Evrard, AE et al. (2008). 'Virial scaling of massive dark matter halos: why clusters prefer a high normalization cosmology'. In: *The astrophysical journal* 672.1, p. 122.
- Fabian, AC et al. (2022). 'Hidden cooling flows in clusters of galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 515.3, pp. 3336–3345.
- Fabian, Andrew C (2012). 'Observational evidence of active galactic nuclei feedback'. In: *Annual Review of Astronomy and Astrophysics* 50, pp. 455–489.
- Fall, S Michael and George Efstathiou (1980). 'Formation and rotation of disc galaxies with haloes'. In: *Monthly Notices of the Royal Astronomical Society* 193.2, pp. 189–206.
- Fall, S Michael and Aaron J Romanowsky (2013). 'Angular momentum and galaxy formation revisited: effects of variable mass-to-light ratios'. In: *The Astrophysical Journal Letters* 769.2, p. L26.
- (2018). 'New perspectives on galactic angular momentum, galaxy formation, and the Hubble Sequence'. In: *arXiv preprint arXiv:1812.06144*.
- Faltenbacher, A and J Diemand (2006). 'Velocity distributions in clusters of galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 369.4, pp. 1698–1702.
- Faltenbacher, A et al. (2007). 'Entropy of gas and dark matter in galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 376.3, pp. 1327–1334.
- Faltenbacher, Andreas et al. (2005). 'Supersonic motions of galaxies in clusters'. In: *Monthly Notices of the Royal Astronomical Society* 358.1, pp. 139–148.
- Feldmeier, JJ et al. (2004). 'Intracluster planetary nebulae in the Virgo Cluster. III. Luminosity of the intracluster light and tests of the spatial distribution'. In: *The Astrophysical Journal* 615.1, p. 196.
- Fixsen, DJ et al. (1996). 'The cosmic microwave background spectrum from the full COBE* FIRAS data set'. In: *The Astrophysical Journal* 473.2, p. 576.
- Foreman-Mackey, Daniel et al. (2013). 'emcee: the MCMC hammer'. In: *Publications of the Astronomical Society of the Pacific* 125.925, p. 306.
- Fraser-McKelvie, Amelia, Michael JI Brown and Kevin A Pimbblet (2014). 'The rarity of star formation in brightest cluster galaxies as measured by WISE'. In: *Monthly Notices of the Royal Astronomical Society: Letters* 444.1, pp. L63–L67.
- Furlong, M et al. (2015). 'Evolution of galaxy stellar masses and star formation rates in the EAGLE simulations'. In: *Monthly Notices of the Royal Astronomical Society* 450.4, pp. 4486–4504.
- Geller, MJF and JP Huchra (1983). 'Groups of galaxies. III-The CfA survey'. In: *The Astrophysical Journal Supplement Series* 52, pp. 61–87.
- Giallongo, Emanuele et al. (2014). 'Diffuse optical intracluster light as a measure of stellar tidal stripping: The cluster CL0024+ 17 at $z \sim 0.4$

- observed at the large binocular telescope'. In: *The Astrophysical Journal* 781.1, p. 24.
- Gingold, RA and JJ Monaghan (1977). 'Smoothed particle hydrodynamics: theory and application to non-spherical stars'. In: *Monthly notices of the royal astronomical society* 181.3, pp. 375–389.
- Giodini, S et al. (2009). 'Stellar and total baryon mass fractions in groups and clusters since redshift 1'. In: *The Astrophysical Journal* 703.1, p. 982.
- Golden-Marx, Jesse B et al. (2022). 'Characterising the Intracluster Light over the Redshift Range $0.2 < z < 0.8$ in the DES-ACT Overla'. In: *arXiv preprint arXiv:2209.05519*.
- Gonzalez, Anthony H, Dennis Zaritsky and Ann I Zabludoff (2007). 'A census of baryons in galaxy clusters and groups'. In: *The Astrophysical Journal* 666.1, p. 147.
- Gonzalez, Anthony H et al. (2021). 'Discovery of a possible splashback feature in the intracluster light of MACS J1149.5+ 2223'. In: *Monthly Notices of the Royal Astronomical Society* 507.1, pp. 963–970.
- Goodfellow, Ian, Yoshua Bengio and Aaron Courville (2016). *Deep learning*. MIT press.
- Graham, Alister et al. (1996). 'Brightest cluster galaxy profile shapes'. In: *The Astrophysical Journal* 465, pp. 534–547.
- Green, Sheridan B and Frank C van den Bosch (2019). 'The tidal evolution of dark matter substructure—I. subhalo density profiles'. In: *Monthly Notices of the Royal Astronomical Society* 490.2, pp. 2091–2101.
- Greggio, L and A Renzini (1983). 'The binary model for type I supernovae - Theoretical rates'. In: *Astronomy and Astrophysics* 118, pp. 217–222.
- Gu, Meng et al. (2020). 'Spectroscopic constraints on the buildup of intracluster light in the Coma Cluster'. In: *The Astrophysical Journal* 894.1, p. 32.
- Gunn, James E and J Richard Gott III (1972). 'On the infall of matter into clusters of galaxies and some effects on their evolution'. In: *The Astrophysical Journal* 176, p. 1.
- Güsten, R and PG Mezger (1982). 'Star formation and abundance gradients in the galaxy'. In: *Vistas in Astronomy* 26, pp. 159–224.
- Guth, AH (1998). *The inflationary universe: the quest for a new theory of cosmic origins*. Random House.
- Habib, Md et al. (2020). 'A comparative study of classifiers in the context of papaya disease recognition'. In: *Proceedings of International Joint Conference on Computational Intelligence*. Springer, pp. 417–429.
- Hahn, Oliver et al. (2017). 'RHAPSODY-G simulations—I. The cool cores, hot gas and stellar content of massive galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 470.1, pp. 166–186.
- Hanany, Shaul et al. (2000). 'MAXIMA-1: a measurement of the cosmic microwave background anisotropy on angular scales of $10'$ - $5'$ '. In: *The Astrophysical Journal* 545.1, p. L5.
- Hansen, SH and B Moore (2006). 'A universal density slope–velocity anisotropy relation for relaxed structures'. In: *New Astronomy* 11.5, pp. 333–338.

- He, Qiuhan et al. (2020). 'Constraining the inner density slope of massive galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 496.4, pp. 4717–4733.
- Henriksen, RN (2006). 'Isolated and non-isolated dark matter haloes and the Navarro, Frenk and White profile'. In: *Monthly Notices of the Royal Astronomical Society* 366.2, pp. 697–704.
- Henriques, Bruno MB et al. (2013). 'Simulations of the galaxy population constrained by observations from $z = 3$ to the present day: implications for galactic winds and the fate of their ejecta'. In: *Monthly Notices of the Royal Astronomical Society* 431.4, pp. 3373–3395.
- Hernquist, L and N Katz (1989). 'TREESPH-A unification of SPH with the hierarchical tree method'. In: *The Astrophysical Journal Supplement Series* 70, pp. 419–446.
- Hinshaw, G et al. (2007). 'Three-year wilkinson microwave anisotropy probe (wmap*) observations: Temperature analysis'. In: *The Astrophysical Journal Supplement Series* 170.2, p. 288.
- Hiotelis, N and A Del Popolo (2006). 'On the reliability of merger-trees and the mass-growth histories of dark matter haloes'. In: *Astrophysics and Space Science* 301.1-4, pp. 167–177.
- Hirschmann, Michaela, Gabriella De Lucia and Fabio Fontanot (2016). 'Galaxy assembly, stellar feedback and metal enrichment: the view from the GAEA model'. In: *Monthly Notices of the Royal Astronomical Society* 461.2, pp. 1760–1785.
- Hockney, RW and JW Eastwood (1988). *Computer simulation using particles*. crc Press.
- Hopkins, Philip F (2015). 'A new class of accurate, mesh-free hydrodynamic simulation methods'. In: *Monthly Notices of the Royal Astronomical Society* 450.1, pp. 53–110.
- Hoyle, F (1951). 'The Origin of the Rotations of the Galaxies'. In: *Problems of cosmical aerodynamics*, p. 195.
- Hoyle, F and RA Lyttleton (1941). 'On the accretion theory of stellar evolution'. In: *Monthly Notices of the Royal Astronomical Society* 101, p. 227.
- Hubble, E (1929). 'A relation between distance and radial velocity among extra-galactic nebulae'. In: *Proceedings of the national academy of sciences* 15.3, pp. 168–173.
- Hubble, Edwin and Milton L Humason (1931). 'The velocity-distance relation among extra-galactic nebulae'. In: *The Astrophysical Journal* 74, p. 43.
- Huertas-Company, Marc et al. (2013). 'The evolution of the mass–size relation for early-type galaxies from $z \sim 1$ to the present: dependence on environment, mass range and detailed morphology'. In: *Monthly Notices of the Royal Astronomical Society* 428.2, pp. 1715–1742.
- Huss, A, B Jain and M Steinmetz (1999). 'The formation and evolution of clusters of galaxies in different cosmogonies'. In: *Monthly Notices of the Royal Astronomical Society* 308.4, pp. 1011–1031.

- Iodice, E et al. (2016). 'The fornax deep survey with VST. I. The extended and diffuse stellar halo of NGC 1399 out to 192 kpc'. In: *The Astrophysical Journal* 820.1, p. 42.
- Jaffe, Walter (1983). 'A simple model for the distribution of light in spherical galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 202.4, pp. 995–999.
- Jaffé, Yara L et al. (2015). 'BUDHIES II: a phase-space view of HI gas stripping and star formation quenching in cluster galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 448.2, pp. 1715–1728.
- Jaffé, Yara L et al. (2016). 'BUDHIES–III: the fate of HI and the quenching of galaxies in evolving environments'. In: *Monthly Notices of the Royal Astronomical Society* 461.2, pp. 1202–1221.
- Jiménez-Teja, Y and R Dupke (2016). 'Disentangling the ICL with the CHEFs: Abell 2744 as a Case Study'. In: *The Astrophysical Journal* 820.1, p. 49.
- Joffre, M et al. (2000). 'Weak Lensing Mass Maps of Very Nearby Clusters of Galaxies'. In: *APS April Meeting Abstracts*, pp. V16–003.
- Kaiser, Nick (1986). 'Evolution and clustering of rich clusters'. In: *Monthly Notices of the Royal Astronomical Society* 222.2, pp. 323–345.
- (1987). 'Clustering in real space and in redshift space'. In: *Monthly Notices of the Royal Astronomical Society* 227.1, pp. 1–21.
- (1992). 'Evolution of clusters of galaxies'. In: *Clusters and Superclusters of Galaxies*. Springer, pp. 323–330.
- Karakas, AI (2010). 'Updated stellar yields from asymptotic giant branch models'. In: *Monthly Notices of the Royal Astronomical Society* 403.3, pp. 1413–1425.
- Katz, Neal, David H Weinberg and Lars Hernquist (1996). 'Cosmological Simulations with TreeSPH'. In: *The Astrophysical Journal Supplement Series* 105, p. 19.
- Kelson, Daniel D et al. (2002). 'Determination of the dark matter profile of A2199 from integrated starlight'. In: *The Astrophysical Journal* 576.2, p. 720.
- Kingma, Diederik P and Jimmy Ba (2014). 'Adam: A method for stochastic optimization'. In: *arXiv preprint arXiv:1412.6980*.
- Kluge, M et al. (2021). 'Photometric dissection of intracluster light and its correlations with host cluster properties'. In: *The Astrophysical Journal Supplement Series* 252.2, p. 27.
- Kluge, Matthias et al. (2020). 'Structure of brightest cluster galaxies and intracluster light'. In: *The Astrophysical Journal Supplement Series* 247.2, p. 43.
- Knebe, A et al. (2013). 'Structure finding in cosmological simulations: the state of affairs'. In: *Monthly Notices of the Royal Astronomical Society* 435.2, pp. 1618–1658.
- Ko, Jongwan and M James Jee (2018). 'Evidence for the Existence of Abundant Intracluster Light at $z = 1.24$ '. In: *The Astrophysical Journal* 862.2, p. 95.

- Kravtsov, AV and S Borgani (2012). 'Formation of galaxy clusters'. In: *Annual Review of Astronomy and Astrophysics* 50, pp. 353–409.
- Kravtsov, AV, AA Vikhlinin and AV Meshcheryakov (2018). 'Stellar mass—halo mass relation and star formation efficiency in high-mass halos'. In: *Astronomy Letters* 44.1, pp. 8–34.
- Kravtsov, Andrey V, Alexey Vikhlinin and Daisuke Nagai (2006). 'A new robust low-scatter X-ray mass indicator for clusters of galaxies'. In: *The Astrophysical Journal* 650.1, p. 128.
- Kuhlen, Michael, Mark Vogelsberger and Raul Angulo (2012). 'Numerical simulations of the dark universe: State of the art and the next decade'. In: *Physics of the Dark Universe* 1.1-2, pp. 50–93.
- Kuhn, Max, Kjell Johnson et al. (2013). *Applied predictive modeling*. Vol. 26. Springer.
- Lapi, A and A Cavaliere (2011). 'Self-similar Dynamical Relaxation of Dark Matter Halos in an Expanding Universe'. In: *The Astrophysical Journal* 743.2, p. 127.
- Larson, RB, BM Tinsley and C Nelson Caldwell (1980). 'The evolution of disk galaxies and the origin of So galaxies'. In: *The Astrophysical Journal* 237, pp. 692–707.
- Lau, ET, D Nagai and AV Kravtsov (2009). 'Effects of baryon dissipation on the dark matter virial scaling relation'. In: *The Astrophysical Journal* 708.2, p. 1419.
- Lesgourgues, J and S Pastor (2006). 'Massive neutrinos and cosmology'. In: *Physics Reports* 429.6, pp. 307–379.
- Lin, Yen-Ting and Joseph J Mohr (2004). 'K-band properties of galaxy clusters and groups: brightest cluster galaxies and intracluster light'. In: *The Astrophysical Journal* 617.2, p. 879.
- Linde, A (2005). 'Prospects of inflation'. In: *Physica Scripta* 2005.T117, p. 40.
- Longair, MS (2007). *Galaxy formation*. Springer Science & Business Media.
- Longobardi, A et al. (2015). 'The outer regions of the giant Virgo galaxy M87 Kinematic separation of stellar halo and intracluster light'. In: *Astronomy & Astrophysics* 579, A135.
- Longobardi, A et al. (2018). 'Kinematics of the outer halo of M 87 as mapped by planetary nebulae'. In: *Astronomy & Astrophysics* 620, A111.
- Lotz, Marcel et al. (2019). 'Gone after one orbit: How cluster environments quench galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 488.4, pp. 5370–5389.
- Loubser, SI et al. (2018). 'Diversity in the stellar velocity dispersion profiles of a large sample of brightest cluster galaxies $z \leq 0.3$ '. In: *Monthly Notices of the Royal Astronomical Society* 477.1, pp. 335–358.
- Loubser, SI et al. (2020). 'Dynamical masses of brightest cluster galaxies I: stellar velocity anisotropy and mass-to-light ratios'. In: *Monthly Notices of the Royal Astronomical Society* 496.2, pp. 1857–1880.
- Lucy, LB (1977). 'A numerical approach to the testing of the fission hypothesis'. In: *The astronomical journal* 82, pp. 1013–1024.

- Ludlow, AD et al. (2010). 'Secondary infall and the pseudo-phase-space density profiles of cold dark matter haloes'. In: *Monthly Notices of the Royal Astronomical Society* 406.1, pp. 137–146.
- Lynden-Bell, Donald (1967). 'Statistical mechanics of violent relaxation in stellar systems'. In: *Monthly Notices of the Royal Astronomical Society* 136, p. 101.
- Madau, P, A Ferrara and MJ Rees (2001). 'Early metal enrichment of the intergalactic medium by pregalactic outflows'. In: *The Astrophysical Journal* 555.1, p. 92.
- Madau, P, F Haardt and MJ Rees (1999). 'Radiative transfer in a clumpy universe. III. The nature of cosmological ionizing sources'. In: *The Astrophysical Journal* 514.2, p. 648.
- Magorrian, J et al. (1998). 'The demography of massive dark objects in galaxy centers'. In: *The Astronomical Journal* 115.6, p. 2285.
- Mahajan, Smriti, Gary A Mamon and Somak Raychaudhury (2011). 'The velocity modulation of galaxy properties in and near clusters: quantifying the decrease in star formation in backplash galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 416.4, pp. 2882–2902.
- Mamon, GA et al. (2019). 'Structural and dynamical modeling of WINGS clusters-II. The orbital anisotropies of elliptical, spiral, and lenticular galaxies'. In: *Astronomy & Astrophysics* 631, A131.
- Mamon, Gary A, Andrea Biviano and Gwenaël Boué (2013). 'MAMPOSSt: Modelling anisotropy and mass profiles of observed spherical systems—I. Gaussian 3D velocities'. In: *Monthly Notices of the Royal Astronomical Society* 429.4, pp. 3079–3098.
- Mamon, Gary A, Andrea Biviano and Giuseppe Murante (2010). 'The universal distribution of halo interlopers in projected phase space-Bias in galaxy cluster concentration and velocity anisotropy?' In: *Astronomy & Astrophysics* 520, A30.
- Marini, I et al. (2021a). 'On the phase-space structure of galaxy clusters from cosmological simulations'. In: *Monthly Notices of the Royal Astronomical Society* 500.3, pp. 3462–3480.
- Marini, I et al. (2021b). 'Velocity dispersion of brightest cluster galaxies in cosmological simulations'. In: *Monthly Notices of the Royal Astronomical Society* 507.4, pp. 5780–5795.
- Marini, I et al. (2022). 'Machine learning to identify ICL and BCG in simulated galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 514.2, 3082–3096.
- Matteucci, F (2003). 'What determines galactic evolution?' In: *The Evolution of Galaxies*. Springer, pp. 245–254.
- Matteucci, F and L Greggio (1986). 'Relative roles of type I and II supernovae in the chemical enrichment of the interstellar gas'. In: *Astronomy and Astrophysics* 154, pp. 279–287.
- Melchiorri, Alessandro et al. (2000). 'A measurement of Ω from the North American test flight of BOOMERanG'. In: *The Astrophysical Journal* 536.2, p. L63.

- Meneghetti, Massimo et al. (2020). 'An excess of small-scale gravitational lenses observed in galaxy clusters'. In: *Science*.
- Merluzzi, P et al. (2010). 'ACCESS: NIR luminosity function and stellar mass function of galaxies in the Shapley supercluster environment'. In: *Monthly Notices of the Royal Astronomical Society* 402.2, pp. 753–766.
- Merritt, D (1987). 'The distribution of dark matter in the coma cluster'. In: *The Astrophysical Journal* 313, pp. 121–135.
- Mihos, J Christopher (2004). 'Interactions and mergers of cluster galaxies'. In: *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, p. 277.
- Mihos, J Christopher et al. (2016). 'The Burrell Schmidt deep Virgo survey: tidal debris, galaxy halos, and diffuse intracluster light in the Virgo cluster'. In: *The Astrophysical Journal* 834.1, p. 16.
- Mihos, JC et al. (2005). 'Diffuse light in the Virgo cluster'. In: *The Astrophysical Journal Letters* 631.1, p. L41.
- Miller, Christopher J et al. (2003). 'The environment of active galactic nuclei in the sloan digital sky survey'. In: *The Astrophysical Journal* 597.1, p. 142.
- Mo, H, F Van den Bosch and S White (2010). *Galaxy formation and evolution*. Cambridge University Press.
- Monaco, P et al. (2006). 'Diffuse Stellar Component in Galaxy Clusters and the Evolution of the Most Massive Galaxies at $z \leq 1$ '. In: *The Astrophysical Journal Letters* 652.2, p. L89.
- Monaghan, Joseph J and John C Lattanzio (1985). 'A refined particle method for astrophysical problems'. In: *Astronomy and astrophysics* 149, pp. 135–143.
- Montes, M and I Trujillo (2014). 'Intracluster light at the frontier: A2744'. In: *The Astrophysical Journal* 794.2, p. 137.
- Montes, Mireia (2022). 'The faint light in groups and clusters of galaxies'. In: *Nature Astronomy*, pp. 1–9.
- Montes, Mireia and Ignacio Trujillo (2018). 'Intracluster light at the Frontier-II. The Frontier Fields Clusters'. In: *Monthly Notices of the Royal Astronomical Society* 474.1, pp. 917–932.
- (2019). 'Intracluster light: a luminous tracer for dark matter in clusters of galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 482.2, pp. 2838–2851.
- (2022). 'A new era of intracluster light studies with JWST'. In: *arXiv preprint arXiv:2209.00043*.
- Montes, Mireia et al. (2021). 'The Buildup of the Intracluster Light of A85 as Seen by Subaru's Hyper Suprime-Cam'. In: *The Astrophysical Journal* 910.1, p. 45.
- More, Surhud, Benedikt Diemer and Andrey V Kravtsov (2015). 'The splashback radius as a physical halo boundary and the growth of halo mass'. In: *The Astrophysical Journal* 810.1, p. 36.
- More, Surhud et al. (2016). 'Detection of the splashback radius and halo assembly bias of massive galaxy clusters'. In: *The Astrophysical Journal* 825.1, p. 39.

- Morishita, Takahiro et al. (2017). 'Characterizing intracluster light in the Hubble Frontier Fields'. In: *The Astrophysical Journal* 846.2, p. 139.
- Muldrew, Stuart I, Frazer R Pearce and Chris Power (2011). 'The accuracy of subhalo detection'. In: *Monthly Notices of the Royal Astronomical Society* 410.4, pp. 2617–2624.
- Munari, E et al. (2013). 'The relation between velocity dispersion and mass in simulated clusters of galaxies: dependence on the tracer and the baryonic physics'. In: *Monthly Notices of the Royal Astronomical Society* 430.4, pp. 2638–2649.
- Murante, G et al. (2004). 'The diffuse light in simulations of galaxy clusters'. In: *The Astrophysical Journal Letters* 607.2, p. L83.
- Murante, Giuseppe et al. (2007). 'The importance of mergers for the origin of intracluster stars in cosmological simulations of galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 377.1, pp. 2–16.
- Muzzin, Adam et al. (2013). 'The evolution of the stellar mass functions of star-forming and quiescent galaxies to $z = 4$ from the COSMOS/UltraVISTA survey'. In: *The Astrophysical Journal* 777.1, p. 18.
- Nath, BB and N Trentham (1997). 'On the enrichment of the intergalactic medium by galactic winds'. In: *Monthly Notices of the Royal Astronomical Society* 291.3, pp. 505–516.
- Navarro, JF (1996). 'The structure of cold dark matter halos'. In: *Symposium-international astronomical union*. Vol. 171. Cambridge University Press, pp. 255–258.
- Navarro, Julio F, Vincent R Eke and Carlos S Frenk (1996). 'The cores of dwarf galaxy haloes'. In: *Monthly Notices of the Royal Astronomical Society* 283.3, pp. L72–L78.
- Navarro, Julio F, Carlos S Frenk and Simon DM White (1997). 'A universal density profile from hierarchical clustering'. In: *The Astrophysical Journal* 490.2, p. 493.
- Nelson, Dylan et al. (2018). 'First results from the IllustrisTNG simulations: the galaxy colour bimodality'. In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. 624–647.
- Neto, Angelo F et al. (2007). 'The statistics of Λ CDM halo concentrations'. In: *Monthly Notices of the Royal Astronomical Society* 381.4, pp. 1450–1462.
- Newman, Andrew B et al. (2013). 'The density profiles of massive, relaxed galaxy clusters. I. The total density over three decades in radius'. In: *The Astrophysical Journal* 765.1, p. 24.
- Nierenberg, AM et al. (2016). 'The missing satellite problem in 3D'. In: *Monthly Notices of the Royal Astronomical Society* 462.4, pp. 4473–4481.
- Oemler Jr, Augustus (1974). 'The systematic properties of clusters of galaxies. Photometry of 15 clusters'. In: *The Astrophysical Journal* 194, pp. 1–20.
- Oman, Kyle A and Michael J Hudson (2016). 'Satellite quenching time-scales in clusters from projected phase space measurements matched to simulated orbits'. In: *Monthly Notices of the Royal Astronomical Society* 463.3, pp. 3083–3095.

- Omma, H et al. (2004). 'Heating cooling flows with jets'. In: *Monthly Notices of the Royal Astronomical Society* 348.4, pp. 1105–1119.
- Oppenheimer, Benjamin D et al. (2010). 'Feedback and recycled wind accretion: assembling the $z = 0$ galaxy mass function'. In: *Monthly Notices of the Royal Astronomical Society* 406.4, pp. 2325–2338.
- Padovani, P and F Matteucci (1993). 'Stellar mass loss in elliptical galaxies and the fueling of active galactic nuclei'. In: *The Astrophysical Journal* 416, p. 26.
- Pascale, Massimo et al. (2022). 'Unscrambling the Lensed Galaxies in JWST Images behind SMACS 0723'. In: *The Astrophysical Journal Letters* 938.1, p. L6.
- Paszke, Adam et al. (2017). 'Automatic differentiation in PyTorch'. In: *NIPS-W*.
- Pedregosa, F. et al. (2011). 'Scikit-learn: Machine Learning in Python'. In: *Journal of Machine Learning Research* 12, pp. 2825–2830.
- Peebles, PJE (1969). 'Origin of the angular momentum of galaxies'. In: *The Astrophysical Journal* 155, p. 393.
- Peng, Ying-jie et al. (2010). 'Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function'. In: *The Astrophysical Journal* 721.1, p. 193.
- Penzias, AA and RW Wilson (1965). 'A measurement of excess antenna temperature at 4080 Mc/s.'. In: *The Astrophysical Journal* 142, pp. 419–421.
- Percival, WJ (2005). 'Cosmological structure formation in a homogeneous dark energy background'. In: *Astronomy & Astrophysics* 443.3, pp. 819–830.
- Percival, Will J et al. (2007). 'Measuring the baryon acoustic oscillation scale using the sloan digital sky survey and 2dF galaxy redshift survey'. In: *Monthly Notices of the Royal Astronomical Society* 381.3, pp. 1053–1066.
- Pillepich, Annalisa et al. (2018). 'First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. 648–675.
- Poliakov, Denis et al. (2021). 'Quantified diffuse light in compact groups of galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 503.4, pp. 6059–6077.
- Pop, Ana-Roxana et al. (2018). 'Formation and incidence of shell galaxies in the Illustris simulation'. In: *Monthly Notices of the Royal Astronomical Society* 480.2, pp. 1715–1739.
- Postman, M and TR Lauer (1995). 'Brightest cluster galaxies as standard candles'. In: *The Astrophysical Journal* 440, pp. 28–47.
- Pratt, GW et al. (2019). 'The galaxy cluster mass scale and its impact on cosmological constraints from the cluster population'. In: *Space Science Reviews* 215.2, pp. 1–82.
- Presotto, Valentina et al. (2014). 'Intracluster light properties in the CLASH-VLT cluster MACS J1206. 2-0847'. In: *Astronomy & Astrophysics* 565, A126.

- Press, WH et al. (1992). 'Numerical recipes in C++'. In: *The art of scientific computing* 2, p. 1002.
- Price, Daniel J (2012). 'Smoothed particle hydrodynamics and magneto-hydrodynamics'. In: *Journal of Computational Physics* 231.3, pp. 759–794.
- Puchwein, Ewald et al. (2010). 'Intracluster stars in simulations with active galactic nucleus feedback'. In: *Monthly Notices of the Royal Astronomical Society* 406.2, pp. 936–951.
- Quinlan, J Ross (1993). 'Combining instance-based and model-based learning'. In: *Proceedings of the tenth international conference on machine learning*, pp. 236–243.
- Ragone-Figueroa, C et al. (2018). 'BCG mass evolution in cosmological hydro-simulations'. In: *Monthly Notices of the Royal Astronomical Society* 479.1, pp. 1125–1136.
- Ragone-Figueroa, Cinthia et al. (2013). 'Brightest cluster galaxies in cosmological simulations: achievements and limitations of active galactic nuclei feedback models'. In: *Monthly Notices of the Royal Astronomical Society* 436.2, pp. 1750–1764.
- Ragone-Figueroa, Cinthia et al. (2020). 'Evolution and role of mergers in the BCG–cluster alignment. A view from cosmological hydro-simulations'. In: *Monthly Notices of the Royal Astronomical Society* 495.2, pp. 2436–2445.
- Rasia, Elena, Giuseppe Tormen and Lauro Moscardini (2004). 'A dynamical model for the distribution of dark matter and gas in galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 351.1, pp. 237–252.
- Rasia, Elena et al. (2006). 'Systematics in the X-ray cluster mass estimators'. In: *Monthly Notices of the Royal Astronomical Society* 369.4, pp. 2013–2024.
- Remus, Rhea-Silvia, Klaus Dolag and Tadziu L Hoffmann (2017). 'The Outer Halos of Very Massive Galaxies: BCGs and their DSC in the Magneticum Simulations'. In: *Galaxies* 5.3, p. 49.
- Riess, Adam G et al. (1999). 'BVRI light curves for 22 type Ia supernovae'. In: *The Astronomical Journal* 117.2, p. 707.
- Rines, K et al. (2003). 'CAIRNS: the cluster and infall region nearby survey. I. Redshifts and mass profiles'. In: *The Astronomical Journal* 126.5, p. 2152.
- Rodrigo, Carlos and Enrique Solano (2020). 'The SVO Filter Profile Service'. In: *XIV. o Scientific Meeting (virtual) of the Spanish Astronomical Society*, p. 182.
- Rodrigo, Carlos, Enrique Solano, Amelia Bayo et al. (2012). 'SVO Filter Profile Service Version 1.0'. In: *IVOA Working Draft* 15.
- Romano, Donatella et al. (2010). 'Quantifying the uncertainties of chemical evolution studies-II. Stellar yields'. In: *Astronomy & Astrophysics* 522, A32.
- Romanowsky, Aaron J and S Michael Fall (2012). 'Angular momentum and galaxy formation revisited'. In: *The Astrophysical Journal Supplement Series* 203.2, p. 17.

- Ronneberger, Olaf, Philipp Fischer and Thomas Brox (2015). 'U-net: Convolutional networks for biomedical image segmentation'. In: *International Conference on Medical image computing and computer-assisted intervention*. Springer, pp. 234–241.
- Rubin, Vera C, W Kent Ford Jr and Norbert Thonnard (1978). 'Extended rotation curves of high-luminosity spiral galaxies. IV-Systematic dynamical properties, SA through SC'. In: *The Astrophysical Journal* 225, pp. L107–L111.
- (1980). 'Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605/R= 4kpc/to UGC 2885/R= 122 kpc'. In: *The Astrophysical Journal* 238, pp. 471–487.
- Rudick, CS, JC Mihos and C McBride (2006). 'The formation and evolution of intracluster light'. In: *The Astrophysical Journal* 648.2, p. 936.
- Rudick, Craig S, J Christopher Mihos and Cameron K McBride (2011). 'The quantity of intracluster light: comparing theoretical and observational measurement techniques using simulated clusters'. In: *The Astrophysical Journal* 732.1, p. 48.
- Sampaio-Santos, H et al. (2021). 'Is diffuse intracluster light a good tracer of the galaxy cluster matter distribution?' In: *Monthly Notices of the Royal Astronomical Society* 501.1, pp. 1300–1315.
- Saro, Alex et al. (2013). 'Toward unbiased galaxy cluster masses from line-of-sight velocity dispersions'. In: *The Astrophysical Journal* 772.1, p. 47.
- Saro, Alexandro et al. (2006). 'Properties of the galaxy population in hydrodynamical simulations of clusters'. In: *Monthly Notices of the Royal Astronomical Society* 373.1, pp. 397–410.
- Sartoris, B et al. (2020). 'CLASH-VLT: a full dynamical reconstruction of the mass profile of Abell S1063 from 1 kpc out to the virial radius'. In: *Astronomy & Astrophysics* 637, A34.
- Sartoris, Barbara et al. (2016). 'Next generation cosmology: constraints from the Euclid galaxy cluster survey'. In: *Monthly Notices of the Royal Astronomical Society* 459.2, pp. 1764–1780.
- Scannapieco, C et al. (2005). 'Feedback and metal enrichment in cosmological smoothed particle hydrodynamics simulations—I. A model for chemical enrichment'. In: *Monthly Notices of the Royal Astronomical Society* 364.2, pp. 552–564.
- Schechter, P (1976). 'An analytic expression for the luminosity function for galaxies.' In: *The Astrophysical Journal* 203, pp. 297–306.
- Schmidt, M (1959). 'The rate of star formation.' In: *The Astrophysical Journal* 129, p. 243.
- Seigar, Marc S, Alister W Graham and Helmut Jerjen (2007). 'Intracluster light and the extended stellar envelopes of cD galaxies: an analytical description'. In: *Monthly Notices of the Royal Astronomical Society* 378.4, pp. 1575–1588.
- Serra, Ana Laura and Antonaldo Diaferio (2013). 'Identification of members in the central and outer regions of galaxy clusters'. In: *The Astrophysical Journal* 768.2, p. 116.

- Sérsic, JL (1963). 'Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy'. In: *Boletín de la Asociación Argentina de Astronomía La Plata Argentina* 6, pp. 41–43.
- Shakura, N and RA Sunyaev (1973). 'Black holes in binary systems. Observational appearance.' In: *Astronomy and Astrophysics* 24, pp. 337–355.
- Shin, Tae-hyeon et al. (2019). 'Measurement of the splashback feature around SZ-selected Galaxy clusters with DES, SPT, and ACT'. In: *Monthly Notices of the Royal Astronomical Society* 487.2, pp. 2900–2918.
- Siudek, M et al. (2022). 'Shaping physical properties of galaxy subtypes in the VIPERS survey: environment matters'. In: *arXiv preprint arXiv:2205.14736*.
- Smith, Leslie N (2017). 'Cyclical learning rates for training neural networks'. In: *2017 IEEE winter conference on applications of computer vision (WACV)*. IEEE, pp. 464–472.
- Sohn, Jubee et al. (2020). 'Velocity dispersions of brightest cluster galaxies and their host clusters'. In: *The Astrophysical Journal* 891.2, p. 129.
- Sohn, Jubee et al. (2022). 'HectoMAP: The Complete Redshift Survey (Data Release 2)'. In: *arXiv preprint arXiv:2210.16499*.
- Solanes, José M et al. (2002). 'The three-dimensional structure of the Virgo cluster region from Tully-Fisher and HI data'. In: *The Astronomical Journal* 124.5, p. 2440.
- Spavone, M et al. (2020). 'The Fornax Deep Survey with VST-VIII. Connecting the accretion history with the cluster density'. In: *Astronomy & Astrophysics* 639, A14.
- Springel, V et al. (2005). 'Simulations of the formation, evolution and clustering of galaxies and quasars'. In: *nature* 435.7042, p. 629.
- Springel, Volker (2005). 'The cosmological simulation code GADGET-2'. In: *Monthly notices of the royal astronomical society* 364.4, pp. 1105–1134.
- (2010). 'Smoothed particle hydrodynamics in astrophysics'. In: *Annual Review of Astronomy and Astrophysics* 48, pp. 391–430.
- (2016). 'High performance computing and numerical modelling'. In: *Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality*. Springer, pp. 251–358.
- Springel, Volker and Lars Hernquist (2003). 'Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation'. In: *Monthly Notices of the Royal Astronomical Society* 339.2, pp. 289–311.
- Springel, Volker et al. (2001). 'Populating a cluster of galaxies—I. Results at $z = 0$ '. In: *Monthly Notices of the Royal Astronomical Society* 328.3, pp. 726–750.
- Stevens, Eli, Luca Antiga and Thomas Viehmann (2020). *Deep learning with PyTorch*. Manning Publications.
- Talbot, R, WD Arnett et al. (1973). 'Evolution Of Galaxies. 2. Chemical Evolution Coefficients'. In.
- Tauber, JA, ESA, the Planck Scientific Collaboration et al. (2004). 'The Planck mission'. In: *Advances in Space Research* 34.3, pp. 491–496.

- Taylor, James E and Julio F Navarro (2001). 'The phase-space density profiles of cold dark matter halos'. In: *The Astrophysical Journal* 563.2, p. 483.
- Teklu, Adelheid F et al. (2015). 'Connecting angular momentum and galactic dynamics: the complex interplay between spin, mass, and morphology'. In: *The Astrophysical Journal* 812.1, p. 29.
- Thielemann, F-K et al. (2003). 'Supernova nucleosynthesis and galactic evolution'. In: *From Twilight to Highlight: The Physics of Supernovae*. Springer, pp. 331–343.
- Thomson, RC (1991). 'Shell formation in elliptical galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 253.2, pp. 256–278.
- Tiret, O et al. (2007). 'Velocity dispersion around ellipticals in MOND'. In: *Astronomy & astrophysics* 476.1, pp. L1–L4.
- Torbet, E et al. (1999). 'A measurement of the angular power spectrum of the microwave background made from the high Chilean Andes'. In: *The Astrophysical Journal* 521.2, p. L79.
- Tormen, Giuseppe and Edmund Bertschinger (1996). 'Adding Long-wavelength Modes to an N-body Simulation'. In: *The Astrophysical Journal* 472.1, p. 14.
- Tornatore, L, A Ferrara and R Schneider (2007a). 'Population III stars: hidden or disappeared?' In: *Monthly Notices of the Royal Astronomical Society* 382.3, pp. 945–950.
- Tornatore, Luca et al. (2007b). 'Chemical enrichment of galaxy clusters from hydrodynamical simulations'. In: *Monthly Notices of the Royal Astronomical Society* 382.3, pp. 1050–1072.
- Tozzi, Paolo and Colin Norman (2001). 'The evolution of X-ray clusters and the entropy of the intracluster medium'. In: *The Astrophysical Journal* 546.1, p. 63.
- Umetsu, Keiichi and Tom Broadhurst (2008). 'Combining lens distortion and depletion to map the mass distribution of A1689'. In: *The Astrophysical Journal* 684.1, p. 177.
- Umetsu, Keiichi and Benedikt Diemer (2017). 'Lensing Constraints on the Mass Profile Shape and the Splashback Radius of Galaxy Clusters'. In: *The Astrophysical Journal* 836.2, p. 231.
- Valentini, Milena et al. (2017). 'On the effect of galactic outflows in cosmological simulations of disc galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 470.3, pp. 3167–3193.
- Verdugo, M, Bodo L Ziegler and B Gerken (2008). 'The galaxy populations from the centers to the infall regions in $z \sim 0.25$ clusters'. In: *Astronomy & Astrophysics* 486.1, pp. 9–24.
- Vikhlinin, A et al. (2006). 'Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation'. In: *The Astrophysical Journal* 640.2, p. 691.
- Voit, G Mark et al. (2003). 'On the origin of intracluster entropy'. In: *The Astrophysical Journal* 593.1, p. 272.

- Voit, GM, ST Kay and GL Bryan (2005). 'The baseline intracluster entropy profile from gravitational structure formation'. In: *Monthly Notices of the Royal Astronomical Society* 364.3, pp. 909–916.
- Von Der Linden, Anja et al. (2007). 'How special are brightest group and cluster galaxies?' In: *Monthly Notices of the Royal Astronomical Society* 379.3, pp. 867–893.
- Wang, J and SDM White (2009). 'Are mergers responsible for universal halo properties?' In: *Monthly Notices of the Royal Astronomical Society* 396.2, pp. 709–717.
- Weinberg, David H et al. (2008). 'Baryon dynamics, dark matter substructure, and galaxies'. In: *The Astrophysical Journal* 678.1, p. 6.
- Weinmann, Simone M et al. (2012). 'A fundamental problem in our understanding of low-mass galaxy evolution'. In: *Monthly Notices of the Royal Astronomical Society* 426.4, pp. 2797–2812.
- Werner, Stephane V et al. (2022). 'Satellite quenching was not important for $z \sim 1$ clusters: most quenching occurred during infall'. In: *Monthly Notices of the Royal Astronomical Society* 510.1, pp. 674–686.
- Wetzel, Andrew R, Jeremy L Tinker and Charlie Conroy (2012). 'Galaxy evolution in groups and clusters: star formation rates, red sequence fractions and the persistent bimodality'. In: *Monthly Notices of the Royal Astronomical Society* 424.1, pp. 232–243.
- Wetzel, Andrew R et al. (2013). 'Galaxy evolution in groups and clusters: satellite star formation histories and quenching time-scales in a hierarchical Universe'. In: *Monthly Notices of the Royal Astronomical Society* 432.1, pp. 336–358.
- White, DA, C Jones and W Forman (1997). 'An investigation of cooling flows and general cluster properties from an X-ray image deprojection analysis of 207 clusters of galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 292.2, pp. 419–467.
- White, SIMON DM (1984). 'Angular momentum growth in protogalaxies'. In: *The Astrophysical Journal* 286, pp. 38–41.
- Williams, Rik J et al. (2009). 'Detection of quiescent galaxies in a bicolor sequence from $z = 0 - 2$ '. In: *The Astrophysical Journal* 691.2, p. 1879.
- Wolf, Joe et al. (2010). 'Accurate masses for dispersion-supported galaxies'. In: *Monthly Notices of the Royal Astronomical Society* 406.2, pp. 1220–1237.
- Woosley, SE and Thomas A Weaver (1995). *The evolution and explosion of massive Stars II: Explosive hydrodynamics and nucleosynthesis*. Tech. rep. Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States).
- Xu, Guohong (1995). 'A New Parallel N-Body Gravity Solver: TPM'. In: *The Astrophysical Journal Supplement Series* 98, p. 355.
- Yepes, G et al. (1997). 'Hydrodynamical simulations of galaxy formation: effects of supernova feedback'. In: *Monthly Notices of the Royal Astronomical Society* 284.1, pp. 235–256.

- Yoo, Jaewon et al. (2021). 'Intracluster light properties in a fossil cluster at $z = 0.47$ '. In: *Monthly Notices of the Royal Astronomical Society* 508.2, pp. 2634–2649.
- Zhang, Y-Y et al. (2011). 'HIFLUGCS: Galaxy cluster scaling relations between X-ray luminosity, gas mass, cluster radius, and velocity dispersion'. In: *Astronomy & Astrophysics* 526, A105.
- Zhang, Y et al. (2016). 'Galaxies in X-ray selected clusters and groups in dark energy survey data. I. Stellar mass growth of bright central galaxies since $z \sim 1.2$ '. In: *The Astrophysical Journal* 816.2, p. 98.
- Zhang, Yuanyuan et al. (2019). 'Dark Energy Survey Year 1 Results: Detection of Intracluster Light at Redshift ~ 0.25 '. In: *The Astrophysical Journal* 874.2, p. 165.
- Zibetti, S (2007). 'Statistical properties of the intracluster light from SDSS image stacking'. In: *Proceedings of the International Astronomical Union* 3.S244, pp. 176–185.
- Zibetti, Stefano et al. (2005). 'Intergalactic stars in $z \sim 0.25$ galaxy clusters: systematic properties from stacking of Sloan Digital Sky Survey imaging data'. In: *Monthly Notices of the Royal Astronomical Society* 358.3, pp. 949–967.
- Zürcher, Dominik and Surhud More (2019). 'The splashback radius of Planck SZ clusters'. In: *The Astrophysical Journal* 874.2, p. 184.
- Zwicky, F (1937a). 'Nebulae as gravitational lenses'. In: *Physical Review* 51.4, p. 290.
- (1937b). 'On the Masses of Nebulae and of Clusters of Nebulae'. In: *The Astrophysical Journal* 86, p. 217.

COLOPHON

Internal Dynamics of Galaxy Clusters from Cosmological Hydrodynamical Simulations,

©Ilaria Marini, Tesi di Dottorato

Università degli Studi di Trieste Ciclo XXXV

Final Version as of 17th January 2023 ¶.