Mem. S.A.It. Vol. 94, 227

Memorie della

Galaxy and structure formation before the cosmic noon. The perspective with galaxy spectroscopic surveys

O. Cucciati[1], G. De Lucia[2], F. Fontanot[2], G. Zamorani[1], B.C. Lemaux[3,4], and L. Lubin[4]

[1] INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, 40129 Bologna - Italy - e-mail: olga.cucciati@inaf.it
[2] INAF - Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, 34143 Trieste - Italy
[3] Gemini Observatory, NSF's NOIRLab, 670 N. A'ohoku Place, Hilo, Hawai'i, 96720, USA
[4] Department of Physics and Astronomy, University of California, Davis, 1 Shields Ave, Davis, CA 95616, USA

Received: 08-11-2022; Accepted: 20-12-2022

Abstract. Protoclusters of galaxies are interesting places where to study how environment possibly begins to shape galaxy evolution in the early Universe. Unfortunately, protoclusters are elusive targets, as they show a typical overdensity much lower than their descendant clusters, and their member galaxies are spread over a much larger area. Not only they are difficult to detect, but it is also difficult to assess whether a detected overdensity is going to evolve into a cluster, confirming it is a protocluster. Simulations of dark matter and galaxy formation and evolution can help us in this respect. By building lightcones comprising galaxy position, properties, membership to halos and their merger trees, we can mimic any real data set, and verify how well our algorithms to detect protoclusters perform. In this work, we will show how we can use simulations to understand the very nature of the proto-supercluster Hyperion, found in the COSMOS field.

Key words. galaxies: clusters: general – galaxies: high-redshift – cosmology: observations – large-scale structure of Universe

1. Introduction

It is well known that there exist several physical processes occurring in high-density environments that can affect galaxy evolution (see e.g. the reviews Boselli & Gavazzi 2006; De Lucia 2007, 2011). In the local universe we see the result of this environment-driven evolution, however galaxies are affected by the surrounding environment all along their life-time, during which they might also experience different kinds of environments, as dark matter structures build up. In the framework of the so called environmental histories of galaxies (De Lucia et al. 2012; Hirschmann et al. 2014), it is particularly interesting to isolate the on-set of environmental effects at high redshift, and find how and when they begin to play a role. For

this reason, the study of the (birth of) protoclusters of galaxies is of particular importance to understand galaxy evolution.

Several protocluster candidates have been found and analysed in the literature (see the review by Overzier 2016, although the studies on this topic have hugely evolved since then), thanks to different detections algorithms based on very different kinds of data-sets: drop-outs or photometric redshifts as in Chiang et al. 2014; Toshikawa et al. 2018; narrow-band filters surveys as in Harikane et al. 2019; Huang et al. 2022; deep spectroscopic redshift surveys as in Diener et al. 2013; Guaita et al. 2020; tomographic studies as in Cai et al. 2016; Lee et al. 2018; Newman et al. 2020, just to mention a few. Despite the vast variety of detection methods, accuracy and precision of redshift measurement are necessary to locate galaxies in the right position along the line of sight (modulo their peculiar velocities), especially when we want to detect relatively mild overdensities like protoclusters, that are easily blurred away when using photometric redshifts (see e.g. Chiang et al. 2013). For this reason, spectroscopic redshifts are especially needed, but they are more and more difficult to collect at higher and higher redshift.

Another difficulty in detecting protoclusters is that they are elusive targets. Due to the fact that they are defined by the nature of their descendants (i.e. their evolution into galaxy clusters), observational studies of protoclusters require, in synergy, an attentive comparison with simulations of structure and galaxy formation and evolution. Only simulations mimicking our observational data can tell us which is the fate of the overdensities we detect.

In this work we will show a possible way to use dedicated mock galaxy catalogues to understand our findings in the realm of protoclusters. In particular, we aim to understand the very nature of the proto-supercluster 'Hyperion' (Cucciati et al. 2018) found in the COSMOS field. This analysis is part of a work in progress (Cucciati et al., in preparation) and has just started, so we will show only preliminary results and describe the future steps.

2. The data and the dedicated lightcone

2.1. HYPERION: A very rich proto-supercluster at $z \sim 2.45$

Using the spectroscopic galaxy samples of VUDS (Le Fèvre et al. 2015) and the zCOSMOS-Deep survey, supplemented by a synergistic sample of photometric redshifts, we identified a proto-supercluster[1] in the COSMOS field (Cucciati et al. 2018). This complex structure, dubbed Hyperion, contains at least 7 density peaks within $2.4 < z < 2.5$ connected by filaments that exceed the average density of the volume by $\geq 2\sigma$. We estimated its total volume to be $\sim 10^5$ cMpc^3 and its total mass to be $M \sim 4.7 \times 10^{15} M_\odot$. Its shape is very complex and it has a maximum extension of $\sim 60 \times 60 \times 150$ cMpc^3 in RA, Dec, and redshift, respectively.

While smaller components of this proto-supercluster had previously been identified using heterogeneous galaxy samples (Diener et al. 2015; Chiang et al. 2015; Casey et al. 2015; Lee et al. 2016; Wang et al. 2016), with VUDS we obtained a panoramic view of this large structure that encompasses, connects, and considerably expands on all previous detections of the various sub-components. The characteristics of this proto-supercluster (its redshift, its richness over a large volume, the clear detection of its sub-components), together with the extensive multi-band coverage granted by the COSMOS field, provide us with the possibility to study in detail the formation of a rich supercluster and to investigate the relationship between environment and galaxy properties over an immense range in both local and global density.

The galaxy catalogue we used for our analysis is limited at $I = 25$, and down to this magnitude it comprises $\sim$ 25% of reliable spectroscopic redshifts (*spec-z* from now on) from the VUDS and zCOSMOS-Deep surveys, and $\sim$ 75% of galaxies with only photometric redshift (*photo-z* from now on).

[1] ESO press release: https://www.eso.org/public/news/eso1833/

The algorithm used to detect and characterise Hyperion is fully described in Cucciati et al. (2018). The same method, with a few minor variations, has also been applied to detect clusters and groups in the data set of the survey 'Observations of Redshift Evolution in Large-Scale Environment' (ORELSE, see Lubin et al. 2009 for the overview of the survey, and Hung et al. 2020 for the cluster detection). Briefly, the adopted method is the Voronoi tessellation Monte Carlo (VMC) mapping, which is based on the 2D Voronoi tessellation (Voronoi 1908) in overlapping redshift slices. Before selecting galaxies in a given redshift slice, we randomly extract the *photo-z* values from the *photo-z* probability distribution functions (PDF) of the galaxies that have only a *photo-z*. Galaxies with *spec-z* are always used at the measured *spec-z*. For each redshift slice, we use 100 realisations of the *photo-z* values, and for each realisation we compute, on a regular grid, a density map based on the Voronoi tessellation. The final density map in each slice is the average of the 100 density maps recovered from the 100 realisations. We finally obtain the third dimension along the line of sight by piling up the 2D density maps in the subsequent redshift slices. The final product is a data cube regularly gridded in RA, Dec and redshift, with an overdensity value assigned to each pixel.

Hyperion was defined as the contiguous volume of space in the analysed data cube that exceeds the average density of the total volume by $\geq 2\sigma$. The 7 peaks within Hyperion are the sub-volumes with an overdensity exceeding by $\geq 5\sigma$ the average density.

2.2. The GAEA semi-analytical model

We used the galaxy merger trees produced with the semi-analytic model 'GAlaxy Evolution and Assembly' (GAEA), in its implementation described in Fontanot et al. (2020). We retained only galaxies in the snapshot #30, which correspond to $z = 2.42$ in the Millennium Run cosmology, as this is the snapshot with the redshift closest to Hyperion's one.

For all galaxies in that snapshot, we retrieved from their merger trees a set of useful galaxy properties, such as x, y, z positions (in cMpc), peculiar velocities in the three directions, absolute magnitudes, K-corrections, halo membership in that snapshot, and ID and total mass of descendant group/cluster at $z = 0$.

To create a lightcone, we arbitrarily chose a line of sight, and transformed the cubic snapshot into an observed-like catalogue spanning the redshift range $2.25 \lesssim z \lesssim 2.7$. By making use also of the component of the peculiar velocities along the l.o.s, and of the K-correction, for each galaxy we derived the observed redshift z_{obs}, the I-band observed magnitude, and the position on the sky (RA and Dec).

For our purposes, we need to mimic also the background and the foreground galaxy distribution, because we will use also photometric redshifts for the density field reconstruction. Due to their intrinsic uncertainties, *photo-z* are scattered along the l.o.s. with respect to their original true position. When we select *photo-z* galaxies within a given redshift slice, as done by the VMC, we need to take into account that a non negligible fraction of selected galaxies might come from much lower or much higher redshift, and we need to include these galaxies in our lightcone. To mimic such foreground and background galaxies, we took an already available GAEA lightcone, and attached the galaxies at $z \lesssim 2.25$ and those at $z \gtrsim 2.7$ to our lightcone, preserving their z_{obs} and I-band observed magnitude.

Finally, we apply to the lightcone the same observational biases as the VUDS sample used in Cucciati et al. (2018): $\sim$ 25% spectroscopic sampling rate, photometric redshift error of $\sigma_{pz} = 0.028(1 + z)$, cut at $I \leq 25$.

We remark that the I-band number counts in our lightcones are in very good agreement with the I-band number counts in the COSMOS field. In particular, we verified that the redshift distribution $n(z)$ of galaxies down to $I = 25$ in the lightcones in the redshift range $2.25 \lesssim z \lesssim 2.7$ underestimates by only $\sim$ 6% the $n(z)$ of the galaxies in the COSMOS field down to the same magnitude limit. By using a slightly fainter magnitude cut to select galaxies in the lightcone, to match more closely the real $n(z)$, the results that we show in Sect. 3 do not change. A more detailed comparison between the GAEA lightcone and real data for

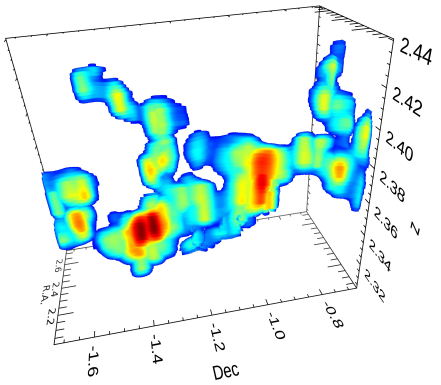

Fig. 1. Overdensity 3D map of the Hyperion-like structure found in the simulation. The colour scale ranges from an overdensity of 2σ above the average (darkest blue) to $> 5\sigma$ above average (darkest red).

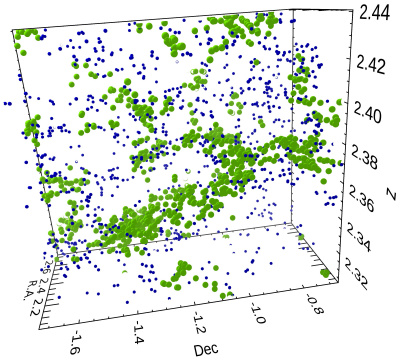

Fig. 2. 3D distribution of the galaxies down to $I = 25$ used to derive the local density in the same volume of space as in Fig. 1. The redshift is the true one (plus peculiar velocities), but the local density has been computed using only $\sim$ 25% of *spec-z*, and the remaining $\sim$ 75% of *photo-z*. Green points are galaxies in regions with a local density $\geq 2\sigma$, and blue points are galaxies in the remaining regions with a local density $< 2\sigma$.

what concerns other galaxy properties is beyond the scope of this preliminary analysis, although it will be the focus of a future work to understand how well the environmental effects on galaxy evolution are modelled in the GAEA model.

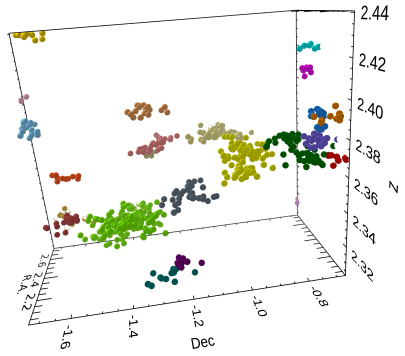

Fig. 3. Same as in Fig. 2, but here we show only galaxies that are progenitors of galaxies that by $z = 0$ will be members of clusters with $M \geq 10^{14} M_{\odot}$. Galaxies of the same colour will become members of the same cluster at $z = 0$.

3. Preliminary analysis and conclusions

To find overdense structures in the lightcone, we applied the same VMC method as applied in the real data.

Although the lightcone covers a much larger area than the COSMOS field, at first we performed our analysis on a sub-area of $1.2 \times 1.2\text{deg}^2$. Considering also the redshift range explored, the total analysed volume in the lightcones is more than 8 times the volume of $\sim 60 \times 60 \times 150$ cMpc comprising Hyperion.

Fig. 1 shows the 3D density map of a structure similar to Hyperion that we found in the analysed volume. This structure is similar to Hyperion in its volume ($\sim 10^5$ cMpc^3), its total mass ($M \sim 6 \times 10^{15} M_{\odot}$) and maximum extension along the three axes ($\sim 70 \times 100 \times 150$ cMpc^3).

Thanks to the use of the mock catalogue, we can give a closer look to this structure. Fig. 2 shows the 3D distribution of all galaxies in the lightcone in the same volume as Fig. 1. We show only galaxies with $I \leq 25$, and their position along the l.o.s is their cosmological redshift plus peculiar velocities. This means that in the figure the information along the l.o.s. is not degraded by the photometric redshift error. From the comparison of Fig. 1

and Fig. 2 it is evident that the Hyperion-like structure truly comprises galaxies that are more clustered than the other galaxies in the surroundings, proving that the VMC is very effective in finding overdense structures even in case of $\sim$ 75% of galaxies with *photo-z*. The comparison of the two figures also shows that the most dense peaks in Fig. 1 are elongated along the l.o.s. with respect to the true galaxy distribution shown in Fig. 2. As discussed in Cucciati et al. (2018), this elongation is mainly due to the photometric redshift error and to the chosen depth of the redshift slices, and it hampers the measurement of the volume (which appears larger than reality) and of the density (which appears lower) of the structure under analysis. It is not trivial to assess by how much these measurements are off. A very compact and dense peak would mostly remain above the detection threshold even if the effect of the elongation is to lower its density, and so its elongation would be visible an possibly measurable. In contrast, a lower density peak, or the outskirts of a large-scale structure as Hyperion, would quickly fall below the detection threshold, and their elongation along the l.o.s. would be less evident and measurable. In Cucciati et al. (2018) we made an attempt to measure the elongation of Hyperion's seven peaks, based on the questionable assumption of spherical shape. Those measurements showed that the volume of the seven peaks appeared larger by a factor comprised between $\sim$ 2 and $\sim$ 6, depending on the peak shape. We have not attempted yet to estimate this factor for the entire Hyperion, and we plan to implement this kind of analysis on the lightcone under analysis.

Finally, by using the galaxy and halo merger trees available for the lightcone, we can study whether Hyperion can be the progenitor of $z = 0$ cluster(s). Fig. 3 shows, in the same volume as Fig. 1 and Fig. 2, the galaxies that are progenitors of galaxies that at z=0 will be members of clusters with total mass $M \geq 10^{14} M_{\odot}$. All galaxies that are progenitors of the same z=0 cluster are plotted with the same colour. By definition, each set can be considered a protocluster. It is clear that in the explored volume the richest protoclusters are indeed part of the Hyperion-like structure. At the same time, this structure comprises several protoclusters of different richness.

These preliminary results show that this kind of analysis is very useful to better understand our findings on protoclusters. Fig. 2 shows that indeed the VMC can robustly identify real overdense structures. The next step will be to statistically quantify this robustness. Fig. 3 tells us that Hyperion can be considered as a filamentary structure comprising several protoclusters. This analysis can be extended in several way. First, we plan to understand how rare such Hyperion-like structures are, by investigating the remaining volume of the lightcone, which covers a total area of $> 30\mathrm{deg}^2$. Then it will be interesting to study whether the protoclusters comprised in one of such structures are all at the same evolutionary stage. Finally, we will link these Hyperion-like structures at $z \sim 2.4$ to their descendants down to $z = 0$ (but see also Ata et al. 2022 for a different way of studying Hyperion's fate). Namely, we will focus our studies at those redshifts where we know from the literature there exist superclusters of galaxies that might be consistent with the possible descendants of Hyperion (see e.g. the supercluster SC1604, Gal et al. 2008).

Another fundamental future step is a statistical generalisation of this analysis to large protoclusters samples. This is of particular importance for future missions and surveys, that will be based on large galaxy data-sets, where it will be difficult to mimic in detail every single protocluster candidate. We will be able to exploit, for instance, new ground-based multi-object spectrographs like the Subaru Prime Focus Spectrograph (PSF, see *e.g.* Takada et al. 2014) and the Multi-Object Optical and Near-infrared Spectrograph at VLT (MOONS, see *e.g.* Cirasuolo et al. 2020), and space missions like Euclid and the Nancy Grace Roman Space Telescope (Laureijs et al. 2011 and Spergel et al. 2015, respectively) that will provide both photometry and spectroscopy. In particular for what concerns Euclid, we have already shown in a previous work (Cucciati et al. 2016) the robustness of the identification of the highest density peaks in the Euclid Deep survey up to $z = 1.5$, exploiting the synergy of photo-

metric and spectroscopic redshifts. It will be an intriguing challenge to expand those results to higher redshift into the realm of protoclusters, upgrading our current tools to work with smaller and smaller percentage of spectroscopic redshifts.

References

Ata, M., Lee, K.-G., Vecchia, C. D., et al. 2022, Nature Astronomy, 6, 857

Boselli, A. & Gavazzi, G. 2006, PASP, 118, 517

Cai, Z., Fan, X., Peirani, S., et al. 2016, ApJ, 833, 135

Casey, C. M., Cooray, A., Capak, P., et al. 2015, ApJ, 808, L33

Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2013, ApJ, 779, 127

Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2014, ApJ, 782, L3

Chiang, Y.-K., Overzier, R. A., Gebhardt, K., et al. 2015, ApJ, 808, 37

Cirasuolo, M., Fairley, A., Rees, P., et al. 2020, The Messenger, 180, 10

Cucciati, O., Lemaux, B. C., Zamorani, G., et al. 2018, A&A, 619, A49

Cucciati, O., Marulli, F., Cimatti, A., et al. 2016, MNRAS, 462, 1786

De Lucia, G. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 379, Cosmic Frontiers, ed. N. Metcalfe & T. Shanks, 257

De Lucia, G. 2011, in Astrophysics and Space Science Proceedings, Vol. 27, Environment and the Formation of Galaxies: 30 years later, 203

De Lucia, G., Weinmann, S., Poggianti, B. M., Aragón-Salamanca, A., & Zaritsky, D. 2012, MNRAS, 423, 1277

Diener, C., Lilly, S. J., Knobel, C., et al. 2013, ApJ, 765, 109

Diener, C., Lilly, S. J., Ledoux, C., et al. 2015, ApJ, 802, 31

Fontanot, F., De Lucia, G., Hirschmann, M., et al. 2020, MNRAS, 496, 3943

Gal, R. R., Lemaux, B. C., Lubin, L. M., Kocevski, D., & Squires, G. K. 2008, ApJ, 684, 933

Guaita, L., Pompei, E., Castellano, M., et al. 2020, A&A, 640, A107

Harikane, Y., Ouchi, M., Ono, Y., et al. 2019, ApJ, 883, 142

Hirschmann, M., De Lucia, G., Wilman, D., et al. 2014, MNRAS, 444, 2938

Huang, Y., Lee, K.-S., Cucciati, O., et al. 2022, arXiv e-prints, arXiv:2206.07101

Hung, D., Lemaux, B. C., Gal, R. R., et al. 2020, MNRAS, 491, 5524

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193

Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, A&A, 576, A79

Lee, K.-G., Hennawi, J. F., White, M., et al. 2016, ApJ, 817, 160

Lee, K.-G., Krolewski, A., White, M., et al. 2018, ApJS, 237, 31

Lubin, L. M., Gal, R. R., Lemaux, B. C., Kocevski, D. D., & Squires, G. K. 2009, AJ, 137, 4867

Newman, A. B., Rudie, G. C., Blanc, G. A., et al. 2020, ApJ, 891, 147

Overzier, R. A. 2016, A&A Rev., 24, 14

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757

Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1

Toshikawa, J., Uchiyama, H., Kashikawa, N., et al. 2018, PASJ, 70, S12

Voronoi, G. F. 1908, J. Reine Angew. Math., 134, 198

Wang, T., Elbaz, D., Daddi, E., et al. 2016, ApJ, 828, 56