



Università degli Studi di Napoli Federico II

DOTTORATO DI RICERCA IN FISICA

Ciclo XXVII

Coordinatore: Prof. Raffaele Velotta

# Formation and Evolution of Galaxies and the Search for Strong Lenses within Kilo Degree Survey

Settore Scientifico Disciplinare FIS/05

**Dottorando**  
Nivya Roy

**Tutore**  
Dr. Nicola R Napolitano  
Dr. Francesco LaBarbera

Anni 2013/2017



# Contents

<b>Introduction</b>	<b>xxi</b>
<b>1 Galaxy Formation and Evolution: a brief overview</b>	<b>1</b>
1.1 The cosmological parameters of the expanding Universe . . . . .	2
1.2 Galaxy formation . . . . .	4
1.3 The Hierarchical Formation Scenario . . . . .	7
1.3.1 Galaxy Merging . . . . .	10
1.3.2 Properties of the different galaxy morphological classes . . . .	12
1.4 Scaling relations . . . . .	15
1.4.1 Early-Type Galaxies . . . . .	15
1.4.2 Late-Type Galaxies . . . . .	17
1.5 Galaxy Evolution from Observations . . . . .	17
<b>2 Kilo Degree Survey</b>	<b>21</b>
2.1 VLT Survey Telescope . . . . .	21
2.2 Kilo Degree Survey . . . . .	23
2.3 Data analysis . . . . .	24
2.3.1 Data reduction and Image Coaddition . . . . .	24
2.3.2 Masking and source flagging . . . . .	25
2.3.3 Catalog extraction and Star/Galaxy separation . . . . .	26
2.3.4 Completeness and limiting magnitude . . . . .	28
2.3.5 Multi-band catalogs and final S/G separation . . . . .	30
2.3.6 The galaxy sample . . . . .	31
2.3.7 Photometric redshifts . . . . .	32
2.3.8 Stellar masses and galaxy classification . . . . .	34
2.4 Comparison with other surveys . . . . .	36
2.5 VISTA Kilo-degree Infrared Galaxy (VIKING) Survey . . . . .	37
<b>3 Galaxy structural parameters: Sample Selection, Data Processing and validation</b>	<b>39</b>
3.1 Sample Selection . . . . .	39

3.1.1	Magnitude completeness . . . . .	41
3.1.2	Photometric Redshifts . . . . .	41
3.1.3	Spectral Energy Distribution fitting for galaxy classification . . . . .	43
3.1.4	Mass completeness as a function of the redshift . . . . .	44
3.2	Surface Photometry . . . . .	45
3.2.1	Structural Parameters from 2DPHOT . . . . .	45
3.2.2	Uncertainties on structural parameters . . . . .	53
3.2.3	Check for systematics on the estimated parameters . . . . .	54
3.3	Passive spheroids and disc-dominated galaxies classification . . . . .	58
<b>4</b>	<b>Evolution of galaxy size–mass relation from the Kilo Degree Survey</b>	<b>61</b>
4.1	Size–Mass as a function of redshift . . . . .	61
4.1.1	Passive Spheroids . . . . .	63
4.1.2	Disc-dominated . . . . .	69
4.1.3	Size–Mass relation for passive spheroids and disc dominated galaxies in g and i bands . . . . .	70
4.1.4	Passive spheroids and disc-dominated size evolution parametric fit . . . . .	72
4.2	Direct Size–Redshift relation in Mass bins . . . . .	74
4.3	Effect of the errors on the $R_e - z$ relations . . . . .	74
4.4	Discussion . . . . .	77
<b>5</b>	<b>Super-Compact massive Galaxies</b>	<b>79</b>
5.1	Introduction . . . . .	80
5.2	Sample selection . . . . .	82
5.2.1	KiDS high signal-to-noise galaxy sample . . . . .	82
5.2.2	Selection of compact galaxies . . . . .	85
5.3	Results . . . . .	88
5.3.1	The final sample . . . . .	88
5.3.2	Systematics from wrong redshifts . . . . .	89
5.3.3	Colour gradients . . . . .	91
5.3.4	Abundance vs. redshift . . . . .	92
5.4	Conclusions and future prospects . . . . .	96
<b>6</b>	<b>Strong lensing within KiDS</b>	<b>99</b>
6.1	Introduction . . . . .	99
6.2	Gravitational lensing: the theory . . . . .	100
6.3	Searching strong lenses in KiDS . . . . .	103
6.4	Visual inspection procedure . . . . .	105
6.5	Spectroscopic follow-up . . . . .	108

---

6.6 Strong Lensing as tool for systematic studies of the mass density of galaxies . . . . .	111
<b>7 Conclusions and Future</b>	<b>113</b>
<b>Appendices</b>	<b>118</b>
A Example output plots from 2DPHOT for galaxies with $\chi'^2 < 1.3$ . . .	118
B Example output plots from 2DPHOT for galaxies with $\chi'^2 > 1.3$ . . .	121
<b>Acknowledgment</b>	<b>139</b>



# List of Figures

1.1	The history of the Universe, from the <i>BigBang</i> to the current epoch.	5
1.2	<i>Left:</i> The anisotropies of the Cosmic Microwave Background (CMB) as observed by Planck. It shows tiny temperature fluctuations that correspond to regions of slightly different densities, representing the seeds of all future structure: the stars and galaxies of today. <i>Right:</i> The angular power spectrum, of the CMB temperature fluctuations in the WMAP full-sky map. The vertical axis represents the amount of anisotropy and the horizontal axis represents the scale, with large scale modes towards the left and fine scale modes to the right. The theoretical prediction of the standard $\Lambda$ CDM model is the red curve and the experimental data are the black points with error bars. The shape of this curve contains a wealth of information about the geometry and matter content of the Universe. . . . .	6
1.3	<i>Left:</i> Averaged Hubble diagram with a linear redshift scale for all SNe from a low-extinction subsample (Knop et al. 2003). Here SNe within $\Delta z < 0.01$ of each other have been combined using a weighted average in order to more clearly show the quality and behavior of the data set. The solid curve overlaid on the data represents our best-fit flat-Universe model, $(\Omega_M; \Omega_\Lambda) = (0.25; 0.75)$ . Two other cosmological models are shown for comparison: $(\Omega_M; \Omega_\Lambda) = (0.25; 0)$ and $(\Omega_M; \Omega_\Lambda) = (1; 0)$ . <i>Right:</i> $\Omega_M - \Omega_\Lambda$ plane with different methods of measurements. . . . .	7
1.4	Snapshots of the <i>Millennium simulation</i> at four different redshifts. Note that the scale is given in comoving coordinates. In the centre, a large galaxy cluster forms due to accretion of material and mergers with smaller clumps. . . . .	8
1.5	<i>Left:</i> Stellar light distributions (g,r,i bands) for a sample of galaxies at $z = 0$ arranged along the classical Hubble sequence for morphological classification, from Illustris (Credits: Illustris Team) . <i>Right:</i> Hubble sequence of EAGLE galaxies (Credits: Schaye et al. 2015). . .	10

1.6	Left: Illustration of a merger tree, depicting the growth of a dark matter halo as the result of a series of mergers. Time increases from bottom to top, and the widths of the tree branches represent the masses of the individual halos. A horizontal slice through the tree, such as that at $t = t_f$ gives the distribution of the masses of the progenitor halos at a given time. Right: shows the galaxy merger tree of a brightest-cluster galaxy from the Millennium Simulation (Springel et al. 2005a; Image courtesy, <a href="https://inspirehep.net/record/1306408/plots">https://inspirehep.net/record/1306408/plots</a> ), illustrating that galaxies grow by the hierarchical merging with other galaxies, which occurs particularly often at early cosmic times. At different epochs, the progenitor halos have merged to form a single halo. The size of each circle represents the mass of the halo. Merger histories of dark matter halos play an important role in hierarchical theories of galaxy formation. . . . .	11
1.7	Tuning-fork-style diagram of the Hubble sequence. . . . .	14
2.1	The VLT Survey Telescope(left) and OmegaCAM camera lies at the heart of the VST. This view shows its 32 CCD detectors that together create 268-megapixel images(right). . . . .	22
2.2	Layout of the KiDS survey fields, KiDS-N in the Northern galactic cap on the equator, and KiDS-S around the South Galactic Pole. The dashed contours outline the full KiDS area (observations ongoing). The solid rectangles indicate the areas observed by the GAMA spectroscopic survey. . . . .	24
2.3	Left: region file produced by the masking software Pullecenella v1.0. Star spikes, primary, secondary and tertiary haloes are evident for the bright stars, while for fainter stars spikes and primary haloes are mainly identified. Right: the corresponding multi-flag file is shown, with the different critical areas colorized with different orange grades (see also de Jong et al. 2015 for further details). . . . .	25
2.4	Left: “true star” definition on the basis of the distribution of the $S/N$ vs FWHM (this case is in $r$ -band). Sure stars are the sources with $S/N > 50$ distributed along the sequence of the smaller FWHM compatible with the point-like PSF (see also text for details). Right: Example of star/galaxy separation diagram for a $r$ -band KiDS tile. At $S/N > 50$ , the high-confidence star candidates (red dots) are used to define the blue line; at lower $S/N$ , all sources with <code>CLASS_STAR</code> $> 0.8$ are used. Sources above the blue line are classified as stars. . . . .	27

2.5	Top: procedure for the estimate of the 98% completeness magnitude. Left: <code>MAG_AUTO</code> vs. detection magnitude in a $g$ -band tile. Blue dots are true stars according to the definition as in §2.3.3, black dots all other sources with grey points showing sources with low $S/N (< 5)$ . The red interpolation line show the 98% percentile of the detection magnitude in different <code>MAG_AUTO</code> bins and the cyan line is the best linear fit to the latter. The intercept with the vertical red dashed line, indicating the <code>DETECTION_MAG</code> where sources saturates at fainter <code>DETECTION_MAG</code> , gives the 98% completeness limit. Right: distribution of the completeness limits in the KiDS-DR1/2 (DR1 with lighter colors). Bottom: procedure for the limiting magnitude. Left: $S/N$ vs. magnitude within 2 arcsec diameter. The limiting magnitude is given by the magnitude corresponding to $S/N = 5$ obtained by the fit to distribution of the values in bins of $S/N$ . Right: distribution of the limiting magnitudes in the KiDS-DR1/2 (DR1 with lighter colors).	29
2.6	Photometric redshifts. In the panel (a) we show the distribution of the full sample with spectroscopic redshifts (white bars) and the distribution of redshifts in the test sample (black bars). In the panel (b) spectroscopic redshift are plotted vs. photometric ones for the test sample galaxies. Blue line is the one-to-one relation.	33
2.7	Le Phare vs. MLPQNA comparison. In yellow, we have the photo- $z$ from Le Phare and in black from MLPQNA for the KiDS data set.	35
2.8	Comparison of SDSS $r$ -band (left) and KiDS $r$ -band (right) images.	36
2.9	Comparison of CFHTLenS $i$ -band (left) and KiDS $r$ -band (right), the weak lensing bands of these two surveys.	37
3.1	Left: Galaxy counts as a function of their magnitude in $r$ -band are compared with other literature estimates (as in the label). The match with previous literature is very good at fainter magnitude while it is not perfect at the bright due to the limited area covered. See also the discussion in the text. Right: completeness of the “high $S/N$ ” sample in $u$ , $g$ , $r$ , and $i$ -band with color code as in the legends. The completeness has been computed with respect to the 7M sample.	41
3.2	<i>Left:</i> the distribution of the spectroscopic sample of the sample adopted as knowledge base for the MLPQNA method (in red) and the photo- $z$ distribution of the “high $S/N$ ” sample (in light blue). <i>Right:</i> Comparison between spectroscopic redshift and the photometric redshift for the blind test set.	42
3.3	C-M diagram for the galaxy sample. Galaxies in red color are the passive spheroids where as the blue cloud represents the disc dominated based on the SED galaxy classification	43

- 
- 3.4 Mass completeness as a function of redshift: the ratio of the high  $S/N$  sample and the photo- $z$  sample for galaxies separated in different redshift bins are shown. See text for details. In the inset the derived completeness masses in the different redshift bins as in Table 3.1. . . . 44
- 3.5 Left: Mass completeness as a function of redshift for g band. The ratio of the high  $S/N$  sample and the photo- $z$  sample for galaxies separated in different redshift bins are shown. The high  $S/N$  sample starts to be fully incomplete in mass above that redshift 0.3 in g band, hence avoided. The sample is complete above  $\log M/M_{\odot}$ , 8.82, 9.48, and 10.58 for  $z \leq 0.1$ ,  $0.1 < z \leq 0.2$  and  $0.2 < z \leq 0.3$  respectively. Right: same as g band. The high  $S/N$  sample starts to be fully incomplete in mass above that redshift 0.4 in g band, hence avoided. The sample is complete above  $\log M/M_{\odot}$ , 9.28, 9.85, 10.39, and 10.98 for  $z \leq 0.1$ ,  $0.1 < z \leq 0.2$ ,  $0.2 < z \leq 0.3$  and  $0.3 < z \leq 0.4$  respectively. . . . . 46
- 3.6 Examples of source and mask image stamps. Source image and mask image are shown in upper and lower panel respectively. Flagged pixels are shaded in black. . . . . 47
- 3.7 PSF anisotropy within the coadd *KIDS*\_129.0\_ – 0.5 in r band. The elongation is aligned in a specific direction on the borders but random in the middle of the image. . . . . 48
- 3.8 2DPHOT output plot of a modeled galaxy. Top left panel is the source and the top right is the model subtracted residual map. Residuals after subtraction are shown in the lower panels as a function of the distance to the galaxy center, with each panel corresponding to a different bin of the polar angle. See also Appendix B and A for further examples. . . . . 49
- 3.9  $\chi^2$  distribution of g, r, and i bands from left to right respectively. . . 50
- 3.10 More examples to the 2D fit results for galaxies in r band. The left panel is the galaxies named the good fits and the right panel as bad fits. In each panel source and the model subtracted residual maps are shown for different sources. See also B and A for further examples. . . 50
- 3.11 The plot shows the Sérsic index vs.  $\chi^2$ . We note that at lower  $n$  ( $< 2.5$ ) there is an excess of large  $\chi^2$  ( $> 1.3$ ), due to the presence of substructures in the residuals. . . . . 51
- 3.12 Distribution of structural parameters in g, r, and i bands from top to bottom respectively. Panels show distribution of  $\log r_e$ ,  $n$ ,  $m_T$ , from left to right respectively. . . . . 52

- 
- 3.13 Uncertainties in the parameters  $\log r_e$ ,  $\langle \mu_e \rangle$ , and  $\log n$  as a function of the logarithm of the  $S/N$  per unit area. Different colors show different bins of  $\log r_e$  where  $r_e$  is in arcsec. For a given color the points are the uncertainties in different bins of logarithm of  $S/N/r_e^2$ . The black curve is the best fitting functional form used to model the dependence of the uncertainties on  $S/N$ . This fit is not performed for  $\log n$  as it does not shows any correlation with  $S/N/r_e^2$ . . . . . 54
- 3.14 Differences in the parameters  $r_e$ ,  $n$ , and magnitude when background is kept constant with respect to the parameter when background is subjected to change for r band. Mean values are given along with the points. Error bar shows the standard deviation in bins of parameter on x axis. In each plot x axis shows the parameter when background is a constant. . . . . 55
- 3.15 Example images shows the real KiDS (left) and mock (right) image. . . . . 56
- 3.16 Figure shows differences in the input and output parameters for  $r_e$ ,  $n$ , and magnitude, with respect to  $S/N$  in upper panel and to their input values in bottom panel. . . . . 57
- 3.17 Comparison of KiDS structural parameters with the ones derived within the SPIDER survey. The SPIDER dataset consists of passive spheroids with redshifts in the range  $0.05 < z < 0.095$ , selected from SDSS; the structural parameters are derived using 2DPHOT. We show the relationship for effective radius ( $r_e$ ), Sérsic index ( $n$ ), and model magnitude ( $mag_r$ ) respectively from left to right . Data are shown as points. Mean values and standard deviation as error bar for differences the parameters with in x axis bins are given. . . . . 58
- 4.1 Size–Mass relation for passive spheroids (top panels and left and central bottom panels). Individual galaxy values are plotted together with mean and standard deviation of the mean (boxes and errorbars). For the  $0 \leq z < 0.1$  bin we overplot some local relations from literature (solid line: Shen et al. 2003; dot dashed line: Hyde & Bernardi 2009; dashed black line: Mosleh et al. 2013; dashed black line: Baldry et al. 2012). For all other  $z$  bins we show the  $z = 0$  relation form Mosleh et al. 2013 to visually appreciate the deviation of the average relation from the local one. Bottom right panel: the stellar mass distributions in different  $z$  bins normalized to the total covolume and to the total galaxy counts in the  $0.2 \leq z < 0.3$  bin. The completeness mass limit are overplotted as vertical dashed lines. Error bars show the propagation of the photo- $z$  errors on the covolume. . . . . 62

- 
- 4.2 Mass vs. redshift plot for the passive spheroid sample. Overplotted to the individual galaxy values, we show the mean and standard deviation of the mean of the sample (error bars are comparable to the size of the boxes). Note that the steepening of the  $z \sim 0.6$  bin is due to the mass incompleteness of this bin. In the lowest  $z$  bin ( $z \sim 0.1$ ), the sample suffers some volume incompleteness (see discussion in §3.1), which produces the mean mass in the bin to be biased toward the less massive systems. . . . . 64
- 4.3 Parametric fit to the Size–Mass relation for passive spheroids. The average size mass in different bins (color code as in Fig. 4.1) is fitted with the parametric formulae as in Eqs. 4.2 (solid lines) and 4.3 (dotted lines). The  $z = 0$  fit has been reported in the subsequent  $z$  bins in blue, to visually check the divergence of the  $z > 0.1$  relations. These curves are used to define the  $R_e$  corresponding to different mass intercepts as shown in Fig. 4.4. . . . . 65
- 4.4 Size vs.  $z$  plots from the average size-mass parametric fit of passive spheroids. *Left*: the derived absolute intercept of the best fit relations as in Fig. 4.3 at mass values as in the legends. Error bars account for the  $1\sigma$  errors in the best fit. We also show the extrapolation for a small mass value ( $\text{Log}M_*/M_\odot = 9.6$ ), which is in the incompleteness range for redshift  $z > 0.2$ , as a “prediction” for higher- $z$  if our best fit relations is correct. We also overplot results from vdW+14, corresponding to  $\text{Log}M_*/M_\odot = 10.25, 10.75, 11.25$ , from bottom to the top, which well compare to our measurements in similar mass bins. *Right*: size evolution with respect to the local size at different mass intercepts. The evolution of the size with redshift becomes increasingly significant at larger masses while at  $\text{Log}M_*/M_\odot \sim 10.4$  the correlation inverts the trend as shown from our predictions for  $\text{Log}M_*/M_\odot = 9.6$ . The grey dashed line is the relation as found by Trujillo et al. (2007) for  $\text{Log}M_*/M_\odot > 11.0$ , the dotted line is the relation from Huertas-Company et al. (2013) for the  $\text{Log}M_*/M_\odot \sim 11.0$ . 66
- 4.5 Size–Mass relation for disc-dominated galaxies. Symbols have the same meaning of the passive spheroid sample in Fig. 4.1, but now data are shown with open symbols in contrast to the passive spheroid average relation also shown as full symbols. The local relation is given by a shaded area which show the range spanned by the average relation from Mosleh et al. (2013) (i.e. their disc-dominated,  $n < 2.5$ , blue samples). In the bottom right panel we summarize all results: the disc-dominated galaxies have generally larger sizes at masses, especially for  $\lesssim 10^{11}M_\odot$  and show a trend with redshift (see §4.1.1) which seems weak or absent. . . . . 68

4.6 Size vs.  $z$  plots from the average size-mass parametric fit of disc-dominated galaxies, as done for passive spheroids in Fig. 4.4. *Left*: the derived absolute intercept of the best fit relations at mass values as in the legends. Error bars account for the  $1\sigma$  errors in the best fit. We also show the extrapolation for a small mass value ( $\text{Log}M_*/M_\odot = 9.0$ ), which is in the incompleteness range for redshift  $z > 0.2$ , to take as a “prediction” to be confirmed with deeper samples. We also overplot three relations from vdW+14 in transparent colors, corresponding to  $\text{Log}M_*/M_\odot = 10.25, 10.75, 11.25$ , from bottom to the top, which nicely overlap to our measurements in similar mass bins. *Right*: size evolution with respect to the  $R_e$  at  $z = 0$  at different mass intercepts. The trend of the size with redshift of disc-dominated seems constant within the errors at all mass bins, included our predictions for  $\text{Log}M_*/M_\odot = 9.0$ . The grey dashed and dotted lines represent the same relation for passive spheroids and disc-dominated respectively (from Trujillo et al. 2007) . . . . . 69

4.7 Size–Mass relation for galaxies in g band. Individual galaxy values are plotted together with mean and standard deviation of the mean (boxes and error bars) in filled square and open square for passive spheroids and disc-dominated galaxies respectively. The local relation is given by a shaded area which show the range spanned by the average relation from Mosleh et al. (2013) (i.e. their disc-dominated,  $n < 2.5$ , blue samples). In the bottom right panel we summarize all results. . . . . 71

4.8 Size–Mass relation for galaxies in g band. Each panel and symbols follows same as 4.7 . . . . . 72

4.9 Size evolution with respect to  $z$  for passive spheroids (left) and disc-dominated galaxies (right) based on  $r$  band data. In the left panel we show the passive spheroid average relation as full squares connected by solid lines against individual galaxy values color coded according to their average mass in the mass bins. In the right panel disc-dominated values are given as squares connected by solid lines and passive spheroids are also reported as dotted lines as comparison. In each case median for redshift are given and standard deviation of mean in size is given as the error bars as filled squares. . . . . 73

- 4.10 Comparison with hydrodynamical simulations. Left panel: passive spheroid size–redshift relation is compared with the EAGLE simulations from Furlong et al. (2015) and with a modified version of GADGET-2 from Oser et al. (2012). The data points show the results obtained in §4.1.1 (solid lines) and §4.2 (dashed lines) in different mass bins as in the legends. Right panel: disc-dominated galaxies size–redshift relation with symbols as in the left panel but with the results from the average-size mass from §4.1.2 (solid lines). . . . . 75
- 4.11 Effect of the errors on the  $R_e$ - $z$  correlation as derived in Fig. 4.9. Left: average trend for the Passive sample obtained by resampling the original sample parameters by adding a shift from a Gaussian distribution using average errors (solid line) and maximum errors (dashed lines) as in the legends. The observed trend is also shown as comparison (tiny solid line). Right: the same as in the left panel for the disc dominated galaxies. . . . . 76
- 5.1  $J - K$  vs.  $g - J$  diagram. An aperture of  $2''$  of diameter is adopted here. Blue and cyan symbols are for high-confidence stars with any error and  $\delta J = \delta K < 0.05$ , respectively. Red points and open boxes are for compact candidates with any error and  $\delta J = \delta K < 0.05$ , survived after the criteria on the mass, size, 2D fit quality and visual inspection. Vega J and K magnitudes are converted to AB as  $J_{AB} = J_{Vega} + 0.930$  and  $K_{AB} = K_{Vega} + 1.834$ . We highlight the regions which are populated by stars (blue), galaxies (yellow) and QSOs (pink). We have considered as sure galaxies those objects with colors  $J - K > 0.2$  and  $g - j > 2$  (yellow shaded region). . . . . 85
- 5.2 2D fit output for 2 example candidates from 2DPHOT procedure. For each galaxy, the top panels show the galaxy image (left) and the residual after the fit (right), while the six bottom panels provide residuals after subtraction, plotted as a function of the distance to the galaxy center, with each panel corresponding to a different bin of the polar angle. Residuals are normalized with the noise expected from the model in each pixel. . . . . 86
- 5.3 Distribution of some galaxy parameters. From left to right we show: 1) photometric redshifts, 2) r-band axis ratio, 3) r-band Sérsic index, 4) r-band effective radius, 5) Chabrier IMF-based stellar mass, 6) r-band effective surface mass density. . . . . 87

- 5.4 We show the median of the  $R_e$  in the g-, r- and i- bands vs. stellar mass. Compact candidates (plotted as red points) are compared with a selection of galaxies ( $\sim 2500$  randomly extracted) from the **FULL C-Sample** (small black points). Gray shaded region highlight the region where candidates are selected (see text for details about the adopted selection criteria). We also plot the change if the spectroscopic redshift is used (dark-green lines). Finally, the four compact galaxies with  $z < 0.2$  from Trujillo et al. (2012) and the three compacts at  $z \sim 2$  from Szomoru et al. (2012), that fulfil our selection criteria, are shown as cyan stars and blue squares, respectively. . . . 90

- 5.5 We show the median of the  $R_e$  in the g-, r- and i- bands (left), r-band **MAG\_AUTO** (middle) and  $g - r$  (right) vs. redshift. The gray region in the left panel sets the region within which we have searched the compacts. Gray line in the middle panel sets the 90% completeness limit of the high- $S/N$  sample. The  $g - r$  color is calculated within  $6''$  of diameter. See further details about the symbols in Fig. 5.4. . . . 91

- 5.6 Number density of compact galaxies vs. redshift. Error bars denote  $1\sigma$  uncertainties (see the text). Vertical dashed black lines mark the completeness limit of the sample (see de Jong et al. 2015). *Left*. Number densities for all galaxies, independent of their photometric formation redshift, are plotted with open squares and solid black lines. The solid gray line with open squares takes into account possible systematics in the redshift determination (see §5.3.2). Short-dashed, long-dashed and solid violet lines are number densities of stellar-like objects from BOSS-DR10 (Damjanov et al. 2014) with  $M_{\text{dyn}} > 5 \times 10^{10}$ ,  $8 \times 10^{10}$ , and  $10^{11} M_{\odot}$ , respectively. The violet shaded region is the  $1\sigma$  error for the case with  $M_{\text{dyn}} > 10^{11} M_{\odot}$ . The violet triangle with bars also shows the abundance of galaxies at  $z \sim 0.25$ , with  $R_e < 1.5 \text{ kpc}$  and  $M_{\text{dyn}} > 8 \times 10^{10} M_{\odot}$ , from Damjanov et al. (2014). The yellow line with lighter yellow region plot abundances for compacts in the COSMOS area (Damjanov et al., 2015a), selected with the same criteria as in the present work ( $M_{\star} > 8 \times 10^{10} M_{\odot}$  and  $R_e < 1.5 \text{ kpc}$ ; I. Damjanov, private communication). Red, cyan and green points are the results for compact galaxies from Trujillo et al. (2009), Taylor et al. (2010) and Poggianti et al. (2013a), respectively. Orange boxes show the abundances for compacts in SDSS area from Saulder et al. (2015), adopting our same criteria on mass and size. Filled boxes plot the results using Sérsic profiles, while open boxes are for the de Vaucouleurs profile (note that the results for the two profiles in the lowest redshift bin are superimposed). *Right*. Black open squares, solid lines and error bars plot KiDS number densities for candidate relic compacts, defined to have photometric formation redshift  $z_f > 2$ . Gray open squares and lines take into account possible systematics in the redshift determination. Long-dashed and solid green lines are for stellar-like objects from BOSS-DR10 with  $z_f > 2$  (Damjanov et al. 2014) and  $M_{\text{dyn}} > 6 \times 10^{10}$  and  $8 \times 10^{10}$ , respectively. The results from Trujillo et al. (2009) and Taylor et al. (2010) are shown here as upper limits (see red and cyan points with arrows). The blue triangle is for the lower limit at  $z \sim 0$  provided by Trujillo et al. (2014). Orange and red lines plot abundances of relic galaxies from semi-analytical models (SAMs), based on Millennium N-body simulations (Quilis & Trujillo 2013). Relics have been defined as systems whose stellar mass has increased since  $z = 2$  to the present by less than 10% and 30%, respectively. Dashed and solid lines are for Guo et al. (2011a) and Guo et al. (2013) SAMs, respectively. Purple stars (connected by a line) are predictions from the Illustris simulations (Wellons et al., 2016). . . . . 93
- 6.1 Illustration showing how gravitational lensing works . . . . . 100

6.2	Lens event configuration. The lens and source planes and the main quantities useful to describe the event are reported. . . . .	103
6.3	Results from the visual inspection of SL candidates. At each row we show: 1) RGB color image, 2) $r$ -band image; 3) model subtracted image; 4) 2DPHOT galaxy model. Arrows show relevant features difficult to spot. . . . .	106
6.4	More examples from the visual inspection. Each panel is the same as Fig.6.3 . . . . .	107
6.5	visual inspection of 5 SL candidates of different kinds (a large arc, an Einstein rings, a faint embedded arc, a tick arc and a 2-image quasar). Top: RGB color image; middle: model subtracted image. Cyan arrows show relevant features difficult to spot. Bottom: DM fractions from Shu et al. (2015) with a state-of art sample of SL candidates (left) and the distribution of the mass and redshift of the KiDS sample proposed here (of the full 50 sources, right). Solid lines indicate the SLACS median redshifts. The KiDS sample extends more in $z$ at any given mass $\log M_*/M_\odot < 11$ and it will allow the study of DM variation (or IMF variation) with $z$ down to sub- $L_*$ systems. .	109
6.6	VIMOS-IFU confirmation of a compact galaxy selected in KiDS. Up left: color image, VIMOS field-of-view and a snapshot of the strongest emission feature found from the source at $z = 0.73$ . The galaxy centre is marked with a green circle, while two spaxel from the source used to extract source spectra are marked with red and cyan circles. Bottom left: the spectrum from the central pixel of the galaxy where typical absorption lines are indicated by blue arrows of a early type galaxy at $z = 0.4$ . Right column: two spectra obtained in two positions of the arc as marked by the red circle (corresponding to the spectrum on the top) and cyan circle (spectrum on the bottom) as in top left panel. In these spectra there are clear emissions duplicated in two different positions which indicate that the detection is genuine and correspond to a $z = 0.73$ source. . . . .	110

1	Left panels :2DPHOT output plot of a modeled galaxy. Top left plot is the source and the top right is the model subtracted residual map. Residuals after subtraction are shown in the lower plots as a function of the distance to the galaxy center, with each panel corresponding to a different bin of the polar angle. Middle panel :Surface brightness profiles of the same galaxy in the left panel. The surface brightnesses computed over different elliptical contours are plotted as a function of the ellipses of equivalent radii, in the upper plot. The error bars denote $1 \sigma$ uncertainties, computed by adding in quadrature the standard deviation of the intensity values in each ellipse with the uncertainty in the background estimate. The solid line is the best-fitting one-dimensional Sérsic model. The lower plot of each panel shows the residuals, in units of $\text{mag arcsec}^{-2}$ , obtained after subtracting the model from the data. Right Panels :Radial profiles of isophotal parameters, as derived by the isophotal fitting algorithm. From top to bottom, the panels show the profiles of ellipticity, position angle of the ellipse's major axis, and $a_4$ coefficient as a function of the equivalent radius of the fitted isophotes. Error bars mark $1 \sigma$ standard uncertainties. In the bottom panels, the minimum and maximum radii to define the global $a_4$ value are shown as vertical dashed lines. The effective radius of the galaxy is marked by a vertical solid line. The solid horizontal line denotes the mean value of $a_4$ in the selected radial range, while the dashed gray lines mark the corresponding $1 \sigma$ interval. . . . .	118
2	More examples as Fig.1 . . . . .	119
3	More examples as Fig.1 . . . . .	120
4	Example output from 2DPHOT for galaxies with $\chi'^2 > 1.3$ . Each panel is same as explained in 1 . . . . .	121
5	More examples as 4 . . . . .	122
6	More examples as 4 . . . . .	123

# List of Tables

1.1	Comparison of Planck-only and WMAP-only six parameter $\Lambda$ CDM Fits	3
3.1	90% completeness mass as a function of the photometric redshift for the high $S/N$ sample in $r$ band. . . . .	45
4.1	$R_e = B_z(1 + z)^{\beta_z}$ fit to the size-redshift relation as derived from the average size-mass fit at different redshifts (Fig. 4.4 and 4.6) and from direct fit to the size-redshift relation in different mass bins (Fig. 4.9).	67



# Introduction

Galaxies are the major structures in the universe where baryons are processed. In the current picture they have condensed in the web-like density distribution of dark matter (DM), which started to aggregate earlier than the baryons had decoupled from the radiation (Springel et al. 2005a; Boylan-Kolchin et al. 2009).

After decoupling, baryons started to collapse towards the bottom of the DM potential wells to give life to first stars and first galaxies (Rees & Ostriker 1977; Silk 1977; White & Rees 1978; Blumenthal et al. 1984), while cold gas kept streaming in the DM veins and contributed to the star formation glory, occurring in the cosmic web over-densities seeding the growth of stars and galaxies. Since then, the overall galaxy formation processes have been the consequence of the balance between the physical mechanisms that drive the condensation of the cold gas into stars (radiative cooling, infall, star formation), the feedback processes (e.g. stellar winds, Supernovae explosions, active galactic nuclei), and the gravitational pull mainly produced from the dominant DM distribution, shaping the environment where galaxy have formed.

Galaxies have been found from voids to cluster up to filaments (e.g. Colless et al. 2001b), where the density contrast can be as large as 3 dex and span stellar mass scales up to 5 dex (from  $10^7$  to  $10^{12} M_{\odot}$ ).

Despite these wide “boundary conditions”, galaxies seem to have very similar morphological and physical properties and yet a full understanding of the way galaxies have evolved across time in a few species (early-type galaxies, ETGs: ellipticals and lenticulars, late-type galaxies, LTGs: spirals and irregulars) is far to be conclusive.

In the local Universe, red, passive, more massive ETGs dominate in group and cluster regions, whilst blue, star-forming, less massive LTGs and irregulars are preferentially found in less-dense, field environments (e.g. Dressler 1980; Kauffmann et al. 2004; Balogh et al. 2004; Baldry et al. 2006; Peng et al. 2010).

Massive ETGs are found to be much more compact in the past with respect to local counterparts (Daddi et al. 2005; Trujillo et al. 2006; Trujillo et al. 2007; van der Wel et al. 2008). At redshifts  $z > 2$ , while the massive star-forming disks have effective radii of several kpc (Genzel et al. 2008), the quenched spheroids (“red nuggets”) have small effective radii of about 1 kpc. Dry-mergers are the dominant mechanisms for the size and stellar mass growth of very dense massive galaxies

(Khochfar & Silk 2006). However, minor mergers do provide a modest stellar mass accretion, but a strong evolution in size, which is what is observed (van Dokkum et al. 2010; Hilz et al. 2013; Belli et al. 2014; Tortora et al. 2014b).

In this thesis we present a project aimed at investigating the evolution of the structural properties of galaxies as derived by deep optical observations of the Kilo Degree Survey (KiDS), one of the ESO public surveys (de Jong et al. 2015), carried out with the VST/OmegaCAM on Cerro Paranal (Capaccioli & Schipani 2011; Kuijken 2011). KiDS will scan 1500 square degrees in four optical filters ( $u$ ,  $g$ ,  $r$ ,  $i$ ), and, together with its near-infrared counterpart VIKING, will provide deep photometry in nine bands in the whole area. Designed to primarily perform weak lensing studies, the exceptional image quality delivered by the combination of the high spatial resolution of VST ( $0.2''/\text{pixel}$ ) and the superb seeing of Cerro Paranal turns out to offer an ideal dataset for galaxy evolution studies, which will be the main focus of this Thesis.

We have selected an high-quality sample of galaxies from the KiDS multi-band catalog of the first and second KiDS data releases, which include about 1/10 of the final survey area. The sample has been collected after having performed an accurate star/galaxy separation and retained only those systems with high signal-to-noise ratio ( $\text{SNR} > 50$ ) for which we have derived robust structural parameters. This galaxy sample has been processed with a specialized software (2DPHOT, La Barbera et al. 2008) to perform a seeing convolved 2D fitting of the surface brightness to derive their effective radius,  $R_e$ , surface brightness  $\mu_e$ , axis ratio  $q$ , position angle, Sèrsic index  $n$ , and total magnitude. The reliability of these parameters has been assessed against mock galaxies on the KiDS images to evaluate the presence of systematics and to estimate intrinsic statistical errors of the main galaxy parameters. We have also compared our results against literature, for those galaxies in common with other dataset (i.e., SPIDER, La Barbera et al. 2010b) and found generally a good agreement.

We have then concentrated on the study of the evolution of the size–mass relation (see e.g. Shen et al. 2003), which is one of the basic relation to understand the assembly of the galaxies across the cosmic time (e.g. Guo et al. 2011a, Oser et al. 2012, Furlong et al. 2015).

We have used photometric redshifts obtained with the Multi Layer Perceptron with Quasi Newton Algorithm (MLPQNA) method (D’Abrusco et al. 2007; Brescia et al. 2013; Brescia et al. 2014). This is a machine learning technique which has been trained on a spectroscopic sample derived by SDSS and GAMA surveys.

Galaxy masses are derived using the software **Le Phare** (Arnouts et al. 1999; Ilbert et al. 2006), which performs a simple  $\chi^2$  fitting method between the stellar population synthesis (SPS) theoretical models and data.

In our catalog of structural parameters, we have identified a peculiar class of galaxies, the so-called ultra-compact massive galaxies, which are rare systems having

$R_e < 1.5 \text{ kpc}$  and masses larger than  $10^{11} M_\odot$  (see e.g. Trujillo et al. 2009, Valentinuzzi et al. 2010, Poggianti et al. 2013c,a, Damjanov et al. 2014). In the  $156 \text{ deg}^2$  of KiDS DR1 and 2 we found 92 compact systems and studied the evolution of their density distribution with redshift (Tortora et al. 2016).

Finally, in the classification of the features found in the residual images, we have identified faint arc-like structures from strong galaxy–galaxy lensing events. We have then started a spin-off project aimed at searching Strong Lensing (SL) features around KiDS galaxies. SL is an ideal tool to perform mass estimates of high and intermediate mass galaxies in non-local universe, provided that the redshift of the lens system and the source is known. The first step is to collect candidate SL systems and we have started a systematic search for such strong lens candidates in KiDS, based on the visual inspection of massive galaxies in the redshift range  $0.1 < z < 0.5$ . We have already collected a few tens of candidates and we will present some preliminary results of the spectroscopic follow-up.

**Brief thesis outline:**

In Chapter 1 we describe the current galaxy formation picture in the context of the  $\Lambda$ CDM scenario; in Chapter 2 we present the dataset used for the thesis work, i.e. the data release 1 and 2 of the KiDS survey (de Jong et al. 2015) and the galaxy catalog based on multi-band catalogs from the survey; in Chapter 3 we present the high-quality galaxy sample and the derivation of the structural parameters, while in Chapter 4 we present the results on the size-mass relation as a function of the redshift; in Chapter 5 we discuss the case of a peculiar class of ETGs, the ultra-compact massive galaxies and discuss their origin in the formation scenario; in Chapter 6 we show some preliminary results of the search of SL candidates around massive ellipticals, and discuss the potential of the KiDS survey as machinery for SL search. In Chapter 7 we draw some conclusions about the Thesis work, and discuss how the KiDS survey might work as precursor for developing scalable techniques for the Next generation surveys like LSST.



# Chapter 1

## Galaxy Formation and Evolution: a brief overview

After hundreds of years of astronomical observations and physical theories, the discovery of the extragalactic nature of galaxies during the 1920's and the cosmological consequences of the General Theory of Relativity by A. Einstein created the solid foundations of modern Cosmology. Since then, the problem of the formation and evolution of galaxies has been one of the most discussed topics in Cosmology, for more than eighty years after galaxies were recognized as “island Universes” of their own rather than mere nebulae in the Milky Way.

With the development in technology and instrumentations (e.g. the radio telescopes in the 70s and the optical and near infrared telescopes of the class of the 4m and 8m diameters between the 50s and 90s, or even the microwave and X-ray satellites in the same epochs of the last century) and, most importantly, with the advent of super computers and cosmological simulations starting from the 70s, the observational and theoretical context have become mature to elaborate a solid theory of galaxy formation and evolution.

Most of the observable properties of galaxies were accommodated within the hierarchical model of cosmic structure growth, which is rooted in the dark matter paradigm called  $\Lambda$  Cold Dark Matter ( $\Lambda$ CDM). This paradigm has emerged after the realization that galaxies and the Universe contain much more mass than can be accounted for by their visible stars and gas, e.g. from the flat rotation curves measured via the HI emission from spiral galaxy disks (e.g. Rubin et al. 1978). This form of unknown and unseen matter, whose nature is still elusive, was called Dark Matter.

At the end of 90s, another dark component, the Dark Energy, was conjectured to provide a significant contribution to the cosmic energy budget, in order to explain the evidences of an acceleration of the expansion of the universe coming from accurate measurements of high-redshift supernovae. These measurements, together with

more and more accurate measurements of the primordial imprints of the density fluctuations from the cosmic microwave background, have allowed cosmologists to obtain a robust scrutiny of the density budget of the Universe (c.f. BOOMERANG, de Bernardis et al. 2001, and later COBE, Mather et al. 1994 and WMAP Bennett et al. 2003, and finally the PLANCK mission from ESA). We know that a huge part of the Universe is made up of dark energy ( $\sim 73\%$ ) and dark matter ( $\sim 23\%$ ). The rest - everything on Earth, everything ever observed with all of our instruments, i.e. the regular or baryonic matter - adds up to  $\sim 4\%$  of the Universe density. Hence the accurate information about the cosmological parameters, the Hubble's constant ( $H_0$ ), and the density parameters, the total density parameter ( $\Omega_0$ ), the mass density parameter ( $\Omega_M$ ), and the "dark energy" parameter ( $\Omega_\Lambda$ ) is necessary to understand the theoretical framework where to build the galaxy formation theories. In this Chapter we will present a brief and not necessarily complete overview about the overall theoretical and observational scheme within which the galaxy formation and evolution is understood, and also provide some basic nomenclature and synopsis of the concepts about galaxies that will be used throughout the Thesis.

## 1.1 The cosmological parameters of the expanding Universe

By the end of 19th century, when Hubble measured the distance of Andromeda using Cepheid stars (previously discovered by Henrietta Swan Leavitt) in it, astronomers understood that the large object that were addressed as "nebulae", were in fact systems situated outside the Milky Way. Using their period-luminosity relationship he estimated a distance of 750 kpc for Andromeda, i.e. clearly farther than the current Milky Way size. This proved that Andromeda was not some small "spiral nebula" within the confines of the Milky Way, but an enormous stellar system in its own right. Hubble's discovery marked the beginning of extragalactic astronomy.

In 1929, Hubble showed that the recession velocities of galaxies are linearly related to their distances (Hubble 1929), hence demonstrating that our Universe is expanding. This relation is known as the Hubble's law

$$v = H_0 \cdot d \tag{1.1}$$

This discovery is marked as the milestone in the astronomy, where major part of the cosmological theories relies on the "Cosmological Principle", which states that on average matter-energy distribution in the Universe is homogeneous and isotropic.

In 1916, using General Relativity, Albert Einstein found that all solutions require the Universe either to expand or to contract, but he believed at that time that the Universe should be static. He introduced a cosmological constant into his field equations to obtain a static solution. In 1922 Alexander Friedmann explored both

Table 1.1: Comparison of Planck-only and WMAP-only six parameter  $\Lambda$ CDM Fits

Parameter	Plank ("CMB+Lens")	WMAP WMAP (9-year)	Difference	
			Value	WMAP $\sigma$
$\Omega_b h^2$	$0.002217 \pm 0.00033$	$0.02264 \pm 0.0005$	-0.00047	0.9
$\Omega_c h^2$	$0.1186 \pm 0.0031$	$0.1138 \pm 0.0045$	0.0048	1.1
$\Omega_\Lambda$	$0.693 \pm 0.019$	$0.721 \pm 0.025$	-0.028	1.1
$\tau$	$0.089 \pm 0.032$	$0.089 \pm 0.014$	0	0
$t_0(Gyr)$	$13.796 \pm 0.058$	$13.74 \pm 0.11$	56 Myr	0.5
$H_0(kms^{-1}Mpc^{-1})$	$67.9 \pm 1.5$	$70.0 \pm 2.2$	-2.1	1.0
$\sigma_8$	$0.823 \pm 0.018$	$0.821 \pm 0.023$	0.002	0.1
$\Omega_b$	0.0481	$0.0463 \pm 0.0024$	0.0018	0.7
$\Omega_c$	0.257	$0.233 \pm 0.023$	0.024	1.0

static and expanding solutions, known as Friedmann models. In 1998, two teams independently showed that the expansion of the Universe is accelerating at the present time using the distance–redshift relation of Type Ia supernovae (Riess et al. 1998a; Perlmutter et al. 1998). This requires an additional mass/energy component with properties very similar to those of Einstein’s cosmological constant.

The standard model of present day Cosmology describes the total density of the Universe in terms of the critical density of the Universe  $\rho_c$ . Using ordinary Newtonian gravity, one can show that a massive particle outside a spherical piece of the Universe expanding with a velocity  $v = H_0 \cdot d$  will escape the gravitational attraction only if the density of space is:

$$\rho \leq \rho_{crit} = \frac{3H_0}{8\pi G} \quad (1.2)$$

where  $G = 6.673 \times 10^{-11} m^{-3} kg^{-1} s^{-2}$  is the universal gravitational constant. This means that to stop the expansion and make the Universe contract a mass density equal or higher than this critical density is required. The value of the critical density is time-dependent, and the present estimate is

$$\rho_{0,c} = 1.87847(23) \cdot 10^{-29} h^2 g cm^{-3} \quad (1.3)$$

where  $h = H_0/100 km sec^{-1}$ ; The contribution from all kinds of matter and radiation to the present time critical density is measured via the total density parameters defined as

$$\Omega_0 = \frac{\rho_o}{\rho_{0,crit}} \quad (1.4)$$

where  $\rho_0$  is the total density of the Universe at the present time. It is also customary to express all the others present day contributions to the critical density separately as the sum of the individual density parameters

$$\Omega_0 = \Omega_M + \Omega_\Lambda + \Omega_k \quad (1.5)$$

where  $\Omega_M$ ,  $\Omega_\Lambda$ ,  $\Omega_k$  are the density of matter, dark energy and space-time curvature respectively. These density parameters together  $H_0$  and  $\rho_0$  are measurable by observations. The latest estimates for these parameters have been provided by the Planck mission<sup>1</sup> and summarized in the Table 1.1. In this Table the main cosmological parameters from Planck are compared to the same estimates from the WMAP satellite from the NASA (<https://heasarc.gsfc.nasa.gov/docs/heasarc/headates/heahistory.html>), which has collected observations until 2010.

## 1.2 Galaxy formation

The age of the Universe is estimated to be 13.8 billion years till now. Different epochs are marked in the history of Universe in this time period as depicted in Fig.1.1. According to the Initial Singularity model, Space-Time and all its mass-energy content originated from a singular state, which can be physically described only after the Planck era ( $t_P \equiv \sqrt{\hbar_P G/c^5} \simeq 5.39116 \times 10^{-44}$  s. Later the Universe evolved through a series of phase transitions from which the elementary particles behind the fundamental interactions (strong and electroweak) were generated. At the end of this phase baryonic matter was generated, and later on the Universe experienced an Inflationary epoch. At this point the Universe was filled homogeneously with a high-energy density, and the underlying space-time was rapidly expanding, inducing a global cooling of the thermal contents of the Universe. Towards the end of this period gravity started to dominated and induced the collapse of the first structures. One second after the *BigBang*, the temperature was around 10 billion degree and all we could see is the vast array of fundamental particles. As the time went on the Universe cooled down to a critical temperature at which a significant fraction of hydrogen and its isotopic species, the deuterium, were created. During this period Universe was opaque as electron scattered all the photons. Cooling process continued down to a temperature when the majority of electrons combined with nuclei to form neutral atoms, thus making the universe almost transparent. This epoch of recombination is known as the epoch when the Cosmic Microwave Background was created, and occurred around 379,000 years after big bang.

As we can deduce from the high level of isotropy of the Cosmic Microwave Background (CMB), after the big bang the Universe was almost homogeneous, and mass-energy fluctuations were very small. Primordial fluctuations are regarded as the

<sup>1</sup>[http://www.esa.int/Our\\_Activities/Space\\_Science/Planck](http://www.esa.int/Our_Activities/Space_Science/Planck)

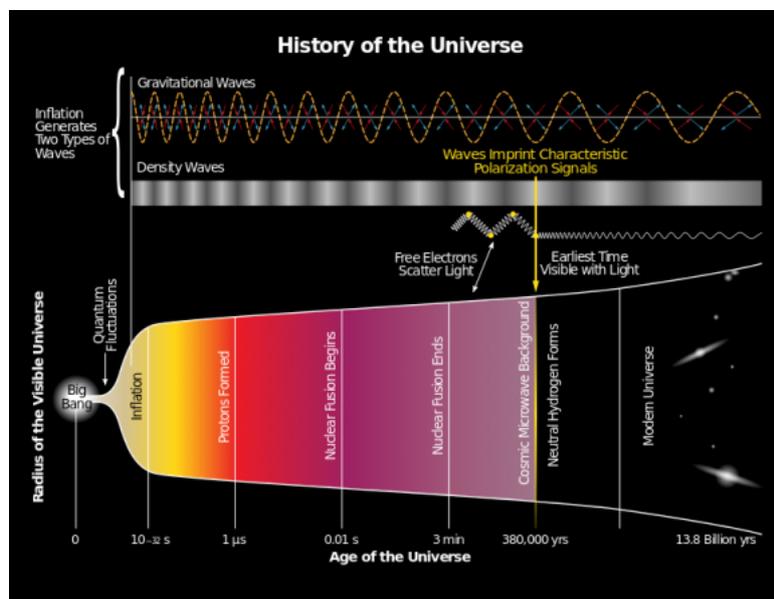


Figure 1.1: The history of the Universe, from the *BigBang* to the current epoch.

regions from which all the structures that we observe today in the Universe were originated. These fluctuations caused the gas to collapse in to the denser areas. The standard model of Cosmology explained the observed properties of the CMB. Photons were absorbed and re-emitted many times by electrons and ions, since the early Universe was dense, hot and highly ionized. As the Universe expanded and cooled and the density of ionized material dropped, photons were scattered less and less often and eventually could propagate freely to the observer from a last-scattering surface, inheriting the blackbody spectrum. Two Milestones in CMB experiments were the Cosmic Background Explorer (COBE, Mather et al. 1994), and a satellite for accurate measurements of the CMB over the entire sky called the Wilkinson Microwave Anisotropy Probe (WMAP, Bennett et al. 2003). Currently the CMB measurement are being pushed to their highest accuracy by the latest microwave satellite operating to map the anisotropies of the CMB, the Planck Space Observatory, operated by the European Space Agency.

Observations from these experiments showed that the CMB has a spectrum that is perfectly consistent with a blackbody spectrum, to exquisite accuracy, with a temperature  $T = 2.728 \pm 0.002\text{K}$ , the anisotropy in the CMB and small temperature fluctuations as shown in Fig 1.2.

All the experiments have measured the position of the first acoustic peak at a value consistent with  $\Omega_0 = 1$  (and  $\Omega_k = 0$ , implying either  $\Omega_M = 1$  and  $\Omega_\Lambda = 0$  or both  $> 0$ ). Baryonic matter is the most immediate component which can show us the structure of the Universe. However two different kind of measurements (primordial

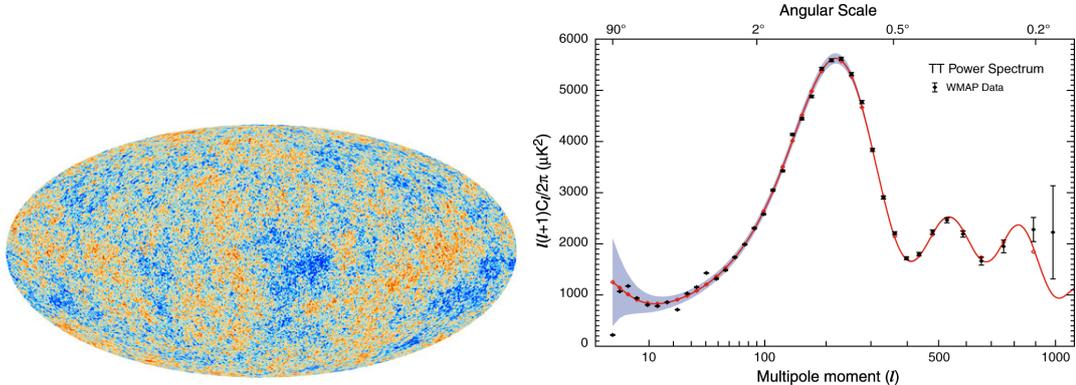


Figure 1.2: *Left:* The anisotropies of the Cosmic Microwave Background (CMB) as observed by Planck. It shows tiny temperature fluctuations that correspond to regions of slightly different densities, representing the seeds of all future structure: the stars and galaxies of today. *Right:* The angular power spectrum, of the CMB temperature fluctuations in the WMAP full-sky map. The vertical axis represents the amount of anisotropy and the horizontal axis represents the scale, with large scale modes towards the left and fine scale modes to the right. The theoretical prediction of the standard  $\Lambda$ CDM model is the red curve and the experimental data are the black points with error bars. The shape of this curve contains a wealth of information about the geometry and matter content of the Universe.

nucleosynthesis and analysis of CMB fluctuations) showed that the baryonic density of the Universe  $\Omega_B$  is only a tiny fraction ( $\Omega_B \sim 0.04$ ) of the total mass density of the Universe.

In order to fill up the empty spaces of the “missing mass problem”, i.e. the discrepancy between mass values of galaxies deduced from dynamical measurement and those deduced from the bolometric emission, some possibilities have been proposed. In 1933, Zwicky studied the velocities of galaxies in the Coma Cluster and concluded that the total mass required to hold the Cluster together is about 400 times larger than the luminous mass in stars pointing to the presence of dark matter for the first time. Even though, these observations were right, it took more than 40 years for the existence of dark matter to be generally accepted. Most of the dark baryonic matter probably resides in intracluster gas or in “dark stars” such as white dwarfs, neutron stars, black holes or objects of mass below the hydrogen burning limit. The velocity dispersion of galaxies in galaxy clusters or the flat rotation curves in spiral galaxies supports the fact that the dark matter in the Universe can be present in non-baryonic form. If the total matter density (baryonic + non-baryonic) amounts to  $\sim 27\%$  of the total mass density of the Universe it is obvious to suppose that an additional component is needed, some kind of “dark energy”. In fact a property of vacuum energy is that its equation of state is  $p = \alpha\rho$  with  $\alpha = -1$  and thus a contribution to the energy density by the vacuum energy gives a negative pressure.

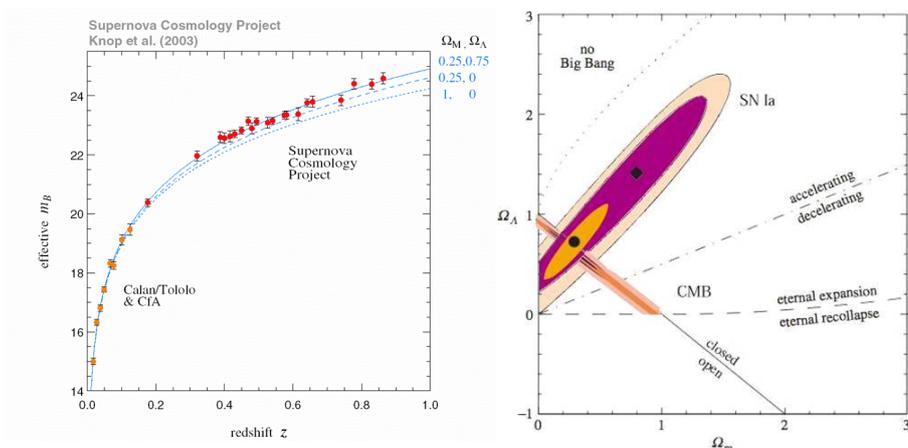


Figure 1.3: *Left* : Averaged Hubble diagram with a linear redshift scale for all SNe from a low-extinction subsample (Knop et al. 2003). Here SNe within  $\Delta z < 0.01$  of each other have been combined using a weighted average in order to more clearly show the quality and behavior of the data set. The solid curve overlaid on the data represents our best-fit flat-Universe model,  $(\Omega_M; \Omega_\Lambda) = (0.25; 0.75)$ . Two other cosmological models are shown for comparison:  $(\Omega_M; \Omega_\Lambda) = (0.25; 0)$  and  $(\Omega_M; \Omega_\Lambda) = (1; 0)$ . *Right*:  $\Omega_M - \Omega_\Lambda$  plane with different methods of measurements.

A negative pressure imply that for  $\Omega_\Lambda > 0$  the Universe must be expanding and accelerating.

$\Lambda$ CDM became the default ‘concordance’ model, when Perlmutter et al. (1998) and Riess et al. (1998b) used the distance – redshift relation of Type Ia supernovae (Knop et al. 2003) to show that the cosmic expansion is accelerating as shown in Fig. 1.3. If the expansion rate was constant, the measured distance – redshift in Fig. 1.3 should give a perfectly linear relation. The fact that distant galaxies have lower redshift and thus are moving slower than predicted by Hubble law means that the expansion is speeding up. Measurements converge to a value of  $\Omega_\Lambda = 0.73$  showing that dark energy gives the most contribution to the total density of the Universe.

### 1.3 The Hierarchical Formation Scenario

Hierarchical structure formation theories predict that massive galaxy clusters are built through the fall of matter, galaxies, groups and sub-clusters, along large scale filaments. Dissipation must have played a role in the formation of disc galaxies: uncondensed gas exists in clusters of galaxies and in individual galaxies and the characteristic mass and size of galaxies finds no natural interpretation in a purely gravitational picture because dissipationless clustering process cannot be responsible for forming both individual galaxies and galaxy clusters.

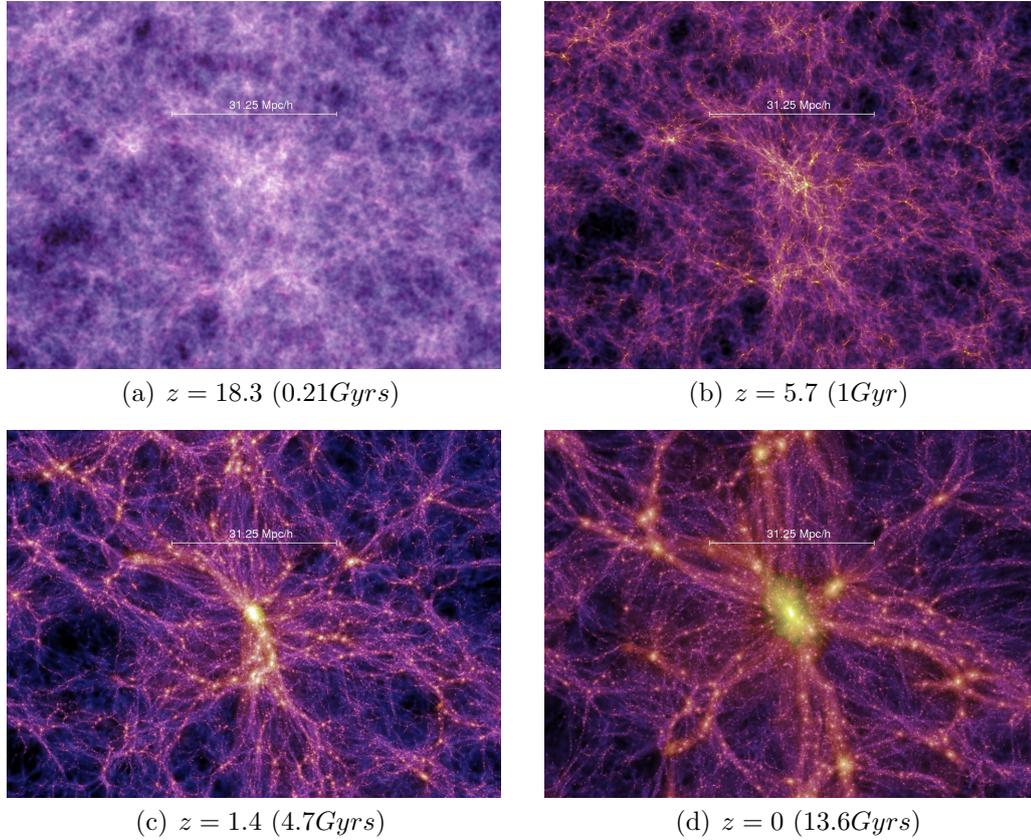


Figure 1.4: Snapshots of the *Millennium simulation* at four different redshifts. Note that the scale is given in comoving coordinates. In the centre, a large galaxy cluster forms due to accretion of material and mergers with smaller clumps.

Once fixed the background cosmological model, the distribution of the hierarchical growth of cosmic structures is well pictured by studying the DM evolution in the Universe by means of the cosmological N-body simulations (e.g. Navarro et al. 1995, Moore et al. 1999, Springel et al. 2005b).

These simulations allow one to follow the formation of the cosmological structure (from galaxies to cluster of galaxies) with great details. The concordance scenario shown from cosmological simulation is an Universe born from a “hot fireball”, an homogeneous and condensate distribution of matter and energy. The dark matter then decoupled from the primordial fluid and through the merging of small sized halos into larger one across the cosmic time the observed structures (galaxy, galaxy cluster, supercluster) which we observe today were formed (see also §1.3.1). Simulations predict that these structures are located in quasi-spherical halos of DM and connected by a more or less dense “Cosmic Web” depending from the cosmological model considered. A compelling view of structure of the Universe was given by

so called *Millennium Simulation* Springel et al. (2005b) which, using  $10^{11}$  particles shows the cosmic structures evolution in a large range of redshift (from  $z = 127$  to  $z = 0$ ).

Figure 1.4 shows the evolutionary scenario (from  $z = 18.3$  to  $z = 0$ ) of a cosmic structure on a scale of  $31.25 Mpc/h$ . At  $z = 18.3$  (Figure 1.4, top left), after some initial expansion the primordial density fluctuations are clearly recognizable: in fact there is little discernible structure and the distribution appears homogeneous and isotropic. Subsequently decreasing of  $z$  reveals the formation of filamentary structures which connect the dark matter halos. For  $z < 5$  galaxy superclusters and more massive DM halos are formed and the proto-clusters are visible along the filamentary structures. At  $z = 0$  all the galaxy clusters and the galaxies inside them are formed their structure (Figure 1.4, bottom right) reveals several hundred dark matter substructures, resolved as independent, gravitationally bound objects orbiting within the cluster halo. These substructures are the remnants of dark matter halos that fell into the cluster at earlier times and are the cocoon where galaxies have seen their birth and where their shining matter is embedded.

The physics of galaxy formation in this cosmological simulation context involves non linear processes which are difficult to handle using analytical techniques. To treat these phenomena two major approaches that have been developed. The first, *hydrodynamical simulations*, which strictly follows the N-body techniques by adding gas particles to the dark matter ones and following their evolution according to the fully nonlinear equations governing the physical processes inherent to galaxy formation in the effective potential generated by the DM field. A second one, *semi-analytical modeling*, attempts to construct a coherent set of analytic approximations which describe these same physics, in the potential and density distribution defined by the dark matter particles (see e.g. White & Frenk 1991; Cole 1991; Lacey & Silk 1991; Kauffmann 1996; De Lucia et al. 2006; Fontanot et al. 2007; Somerville et al. 2008).

Each has its strengths and weaknesses, but recently the hydrodynamical techniques, have reached a level of details that allow to produce most of the measured galaxy scaling relations in a very realistic way (see e.g. Illustris simulations: Pillepich et al. 2014, Vogelsberger et al. 2014; EAGLE simulations: Schaye et al. 2015, Crain et al. 2015). Images of simulated galaxies from Illustris and EAGLE are given in 1.5.

Indeed, hydro simulations are the most accurate computational method for solving the physics of galaxy formation via direct simulation. One way to solve them is through Smoothed Particle Hydrodynamics (SPH) in which the fundamental equations of gravitation, hydrodynamics, radiative cooling and transfer, reionization, star formation, stellar wind and several type of feedback like supernovae explosions and AGN feedback (Romeo et al. 2008; Naab et al. 2009; Schaye et al. 2010; Scannapieco & Athanassoula 2012; Oser et al. 2012; Haas et al. 2013; Vogelsberger et al. 2013) are solved for a large number of simulated particles (arranged either on a grid or

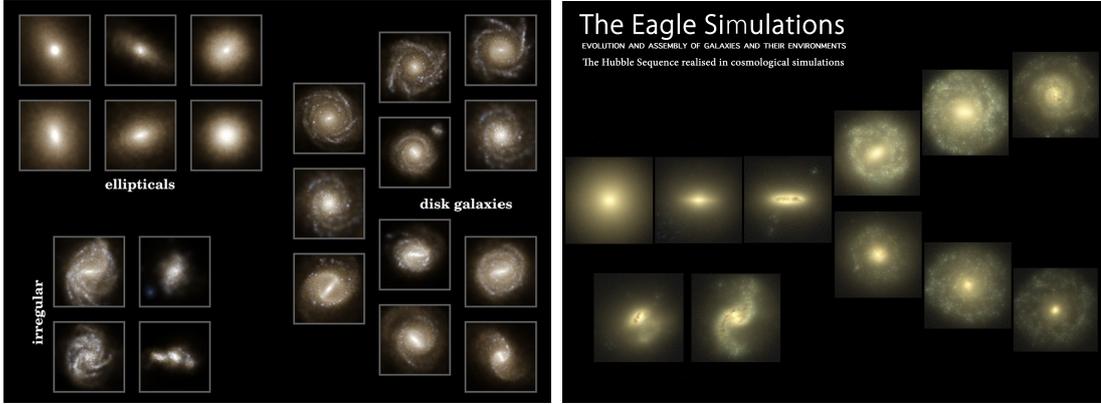


Figure 1.5: *Left:* Stellar light distributions (g,r,i bands) for a sample of galaxies at  $z = 0$  arranged along the classical Hubble sequence for morphological classification, from Illustris (Credits: Illustris Team) . *Right:* Hubble sequence of EAGLE galaxies (Credits: Schaye et al. 2015).

following the trajectories of the fluid flow).

The final result are simulated systems that are reaching nowadays an incredibly fidelity in reproducing the morphology of galaxy systems, as well as most of their basic properties like stellar mass function, the relation between stellar mass and halo mass, galaxy sizes, the relation between BH mass and stellar mass, specific star formation rates and passive fractions, Tully-Fisher relation, mass–metallicity relation (see e.g. Vogelsberger et al. 2014; Schaye et al. 2015).

### 1.3.1 Galaxy Merging

When the orbital energy is sufficiently low, close encounters between two systems can lead to a merger. In the hierarchical scenario of structure formation, mergers play an extremely important role in the assembly of galaxies and dark matter halos. Galaxy merging can happen frequently as the Universe is crowded on the galactic distance scale. A large number of the bright galaxies that we see today may have formed from the mergers of two or more smaller galaxies. The existence of large number of ellipticals in dense clusters of galaxies, supports galaxy merging. Mergers can induce new star formation, and even create gigantic black holes.

In Fig. 1.6 the formation history of a dark matter halo that can be described by a ‘merger tree’ that traces all its progenitors, is illustrated. In this figure, left panel, time increases from top to bottom, and the widths of the branches of the tree represent the masses of the individual progenitors. If we start from an early time, the mass which ends up in a halo at the present time  $t_0$  (the trunk of the tree) is distributed over many small branches; pairs of small branches merge into bigger ones as time goes on. Obviously, merger trees play a very important role in hierarchical

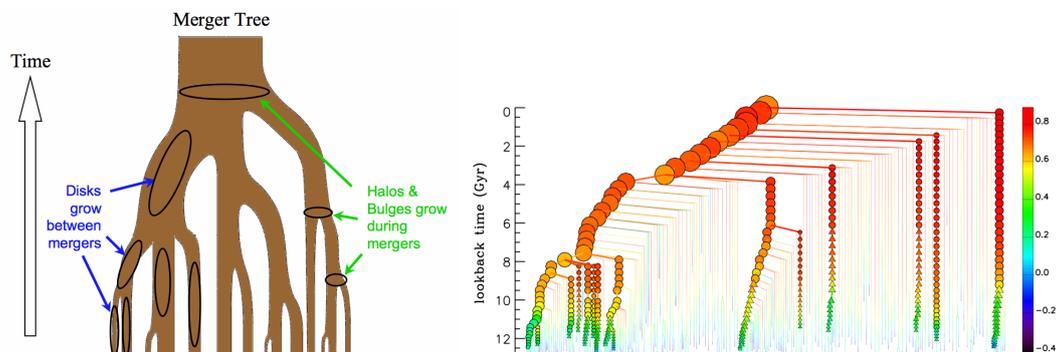


Figure 1.6: Left: Illustration of a merger tree, depicting the growth of a dark matter halo as the result of a series of mergers. Time increases from bottom to top, and the widths of the tree branches represent the masses of the individual halos. A horizontal slice through the tree, such as that at  $t = t_f$  gives the distribution of the masses of the progenitor halos at a given time. Right: shows the galaxy merger tree of a brightest-cluster galaxy from the Millennium Simulation (Springel et al. 2005a; Image courtesy, <https://inspirehep.net/record/1306408/plots>), illustrating that galaxies grow by the hierarchical merging with other galaxies, which occurs particularly often at early cosmic times. At different epochs, the progenitor halos have merged to form a single halo. The size of each circle represents the mass of the halo. Merger histories of dark matter halos play an important role in hierarchical theories of galaxy formation.

models of galaxy formation. Different simulations and the current observation based studies shows that this might be the real scenario where the large scale structures forms. Toomre & Toomre (1972), suggested that most elliptical galaxies could be merger remnants based on the numerical simulation to demonstrate that some of the extraordinary structures seen in peculiar galaxies, such as long tails, could be produced by tidal interactions between two normal spirals Toomre & Toomre 1972. Different simulations showed that the density profile of the observed ellipticals are the resultant of the merging of two spheroids (e.g. White 1978). With the addition of dark matter, hierarchical model could explain almost all the observational facts and it remains the most widely accepted model in current days. The work presented in this thesis is based on this model.

- *Minor merging*: One galaxy is larger than the other. Larger galaxy will absorb most of the gas in the smaller one. This type of merging is frequent between the main galaxy and the satellite galaxy. Our Milky Way galaxy, is in the process of absorbing gas from its own satellite galaxies like Magellanic Clouds. Recent studies shows that minor mergers are indeed important drivers of galaxy evolution. It is believed that the observed size growth of galaxies over the last 10-12 billion years is likely due to repeated minor mergers.

- *Major merging*: Mergers between equal mass progenitors making the resultant galaxy more massive. Even though major mergers have more impact on the galaxy evolution, these are less frequent than minor mergers.

Galaxy merging can be also classified as *Wet merging*, merging between two young, star forming galaxies like spirals and *Dry merging* where galaxies that are less likely to be star forming, like elliptical galaxies merge together.

The idea of hierarchical merging was emerged because of the observations of tidal tails and irregularities during merging (Toomre & Toomre 1972; Toomre 1977). Numerical N- body simulations of equal mass spiral mergers have been done, which can reproduce spheroids that share many of the physical properties of today's ETG population (White 1978). In general merging of galaxies explains the variation of galaxy sizes in cosmic time. Strong evolution in galaxy size is mainly driven by the minor merging (van Dokkum et al. 2010; Hilz et al. 2013). However, minor and major merging in a Universe where structures grow hierarchically, is also required given the strong growth in size and mass and the morphological transition from flat, disk and rotationally dominated structures to more spherical and pressure supported systems. Apart from merging, different mechanisms (Hopkins et al. 2010a) like adiabatic expansion through significant mass loss e.g. due to quasar feedback, the presence of stellar M/L gradients (which can induce differences between the measurement of the optical effective radius and the stellar mass effective radius) as well as redshift dependencies in e.g. the stellar mass estimates can also affect the evolution mechanism. In this Thesis the models considered are based on the merging scenario.

### 1.3.2 Properties of the different galaxy morphological classes

Galaxies are composed of diverse class of objects with a large number of parameters characterizing their structural properties. Any theory of galaxy formation should provide predictions on how these parameters are distributed across the galaxy types and across cosmic time. There have been many attempts to provide accurate galaxy classification depending on the details of the descriptions of the different components which emerge from their apparent properties and morphologies.

The classification provided by Hubble, albeit very simple, still catches most of the major morphological features and it is still a reference classification scheme for galaxies in use today. This Hubble's tuning fork diagram is given in Fig.1.7, see also the analogous from the EAGLE team in Fig. 1.5. In this scheme:

- *Elliptical galaxies*: Elliptical galaxies are mildly flattened, ellipsoidal systems that are mainly supported by the random motions of their stars, having a smooth and featureless light distribution. The surface brightness decreases from center to the outer side. Based on the ellipticity, they are further classified

as E0, E1, ..., E7. If the major and minor axes of an elliptical galaxy are  $a$  and  $b$ , its type is defined to be  $E_n$ , where

$$n = 10\left(1 - \frac{b}{a}\right) \quad (1.6)$$

An E0 galaxy looks circular in the sky. The shape of an E galaxy depends on the direction from which it is observed. In reality an E0 galaxy may therefore be truly spherical or it may be a circular disc viewed directly from above.

- *Spiral galaxies*: These consist of a central bulge and a thin disc of gas and other interstellar matter, forming the spiral pattern. This spiral arm is the place where young stars are being born. The spiral galaxies are sub-divided into two groups, “*normal spirals*” (Sa, Sb, Sc) and “*barred spirals*” (SBa, SBb, SBc). For barred spirals, in addition to the normal spiral structure, we can observe a bar of stars that runs through the central bulge. The classification Sa/SBa, Sb/SBb, and Sc/SBc is done according to the tightness of their spiral, the clumpiness of their spiral arms, and the size of their central bulge. Sa galaxies are with less fraction of dust and hence smaller star formation whereas Sc galaxies are with higher star formation.
- *Lenticular galaxies*: the lenticulars or S0 galaxies are placed between the elliptical and the spiral. Similar to ellipticals they contain only little interstellar matter with out any spiral arms. But they also contain a flat disc made up of stars like spiral galaxies.
- *Irregular galaxies*: These are galaxies which have odd shapes or with out any defined pattern unlike other classes.

The Hubble diagram provides only a morphology classification, but this feature results, indirectly, also from the physical mechanisms which drive the evolution of galaxies.

E.g. spiral galaxies contain a prominent disk composed of Population I stars, gas and dust and a spheroid of population II stars. The disk contain spiral arms filaments of bright O and B stars, gas, and dust, in which large number of stars are currently forming. The distribution of surface brightness (that is the integrated flux observed into a solid angle expressed in  $\text{mag arcsec}^{-2}$ ) of the disk component in a spiral galaxy is well expressed from an exponential law (*Freeman Law*):

$$I(r) = I_0 e^{-r/h}, \quad (1.7)$$

where  $h$  is a disk scale length and the central surface brightness is approximately constant at  $I_0 \simeq 140L_{\odot}pc^{-2}$  (independent of the Hubble constant).

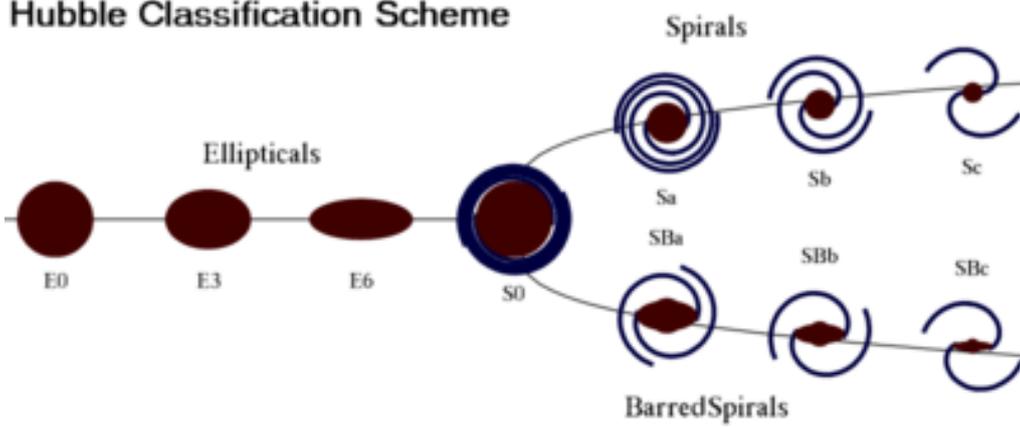


Figure 1.7: Tuning-fork-style diagram of the Hubble sequence.

In general the less star forming and old ellipticals and lenticulars are collectively called early-type galaxies (ETGs) while comparatively young, star forming spirals and irregulars are categorized as late-type galaxies (LTGs). The origin of these term ETGs and LTGs lies with the original interpretation of the Hubble tuning fork diagram Fig.1.7, with an assumption that galaxies evolved from the left (early-type ellipticals) to the right (late-type spirals) in a sequence.

Elliptical galaxies are characterized by smooth, elliptical surface brightness distributions, contain little cold gas or dust, and have red photometric colors, characteristic of an old stellar population. In this section we briefly discuss some of the main, salient observational properties. The surface brightness of an elliptical galaxy falls off smoothly with radius, and can mostly be fit with the de Vaucouleurs (1948) law, or  $r^{1/4}$  law:

$$I(r) = I_e e^{-7.67[(r/r_e)^{1/4} - 1]}, \quad (1.8)$$

where  $r_e$  is the effective radius, that is the radius at which the isophotes contain half of the total luminosity and  $I_e$  the surface brightness at  $r_e$ . Irregular galaxies are composed by a minority of spirals or elliptical that have been violently distorted by an encounter with a neighbor and a majority of simply low-luminosity gas rich systems, such as the Magellanic Clouds.

Although accurate in the majority of bright ellipticals, the  $r^{1/4}$  law can show a deviation from its ideal shapes in the internal or external galaxy regions. In particular in the inner regions deviations from the  $r^{1/4}$  law is explained with the presence of a non-thermal sources (e.g. a central black hole). On the contrary deviations from the  $r^{1/4}$  law at big radii indicates the presence of a luminous halo. These halos are interpreted as the tidal distortion of the external galaxy regions due

to an interaction with another galaxy. Sérsic (1968) proposed a generalization of the  $r^{1/4}$  law and it is nowadays known as Sérsic law or  $r^{1/n}$  law:

$$I(r) = I_e e^{-b_n[(r/r_e)^{1/n} - 1]} \quad (1.9)$$

where  $n$  is an integer,  $r_e$  is the effective radius and  $I_e$  is the SB at this distance. The parameter  $b_n$  can be approximated to  $b_n = 1.999n - 0.3271$ . The three parameters ( $r_e$ ,  $b_n$  and  $n$ ) of the equation (1.9) make this expression an ideal law to fit SBs not well reproduced from an  $r^{1/4}$  law (Caon et al. 1993). For  $n = 4$  the equation (1.9) coincides with the de Vaucouleurs' law, while for  $n = 1$  the Freeman law is obtained. Photometry of early-type galaxies reveals some geometrical properties (e.g. isophote shapes) which have important implications for the physics of these systems.

Elliptical galaxies in general have their stellar contents dominated by old, metal-rich stars and hence red in color. In addition, the colors are tightly correlated with the luminosity such that brighter ellipticals are redder. Giant ellipticals generally have low rotation velocities. As a result of the gravitational effects of the observed stellar populations, the velocity dispersion is observed to rise more strongly at the very center of most nearby ellipticals (and also spiral and S0 bulges). This points to the presence of a central supermassive black hole.

## 1.4 Scaling relations

Generally the evolution picture is studied by connecting different physical properties of galaxies. These relations are called scaling relations. These describe strong trends that are observed between important physical properties (such as mass, size, luminosity and colors) of galaxies. The trends in these relations change according to the type of the galaxies.

### 1.4.1 Early-Type Galaxies

ETGs have shown tight correlations among their structural parameters, namely their effective radius, their surface brightnesses at that radius or equivalently their total luminosity (or total stellar mass) and their central velocity dispersion. These relations, also called scaling relations, represent the foundation of their formation mechanisms and the understanding of their origin and evolution across the cosmic time. We summarize in the following the main scaling relations used to characterize the evolution of ETGs

- *The Color-Magnitude Relation (CMR)*: The colors of early-type galaxies correlated with their absolute magnitudes (Baum 1959; de Vaucouleurs 1961; Visvanathan & Sandage 1977). The CMR is typically explained by an increase of mean stellar metallicity with increasing galaxy mass as the dominant

effect. More massive galaxies have deeper potential wells, which can retain metal-enriched stellar ejecta more effectively and subsequently recycle the enriched gas into new stars (Ferreiras et al. 1999; Gallazzi et al. 2006). Another reason could be a variable integrated galactic initial mass function, with more massive stars in more massive galaxies, and thus a more substantial enrichment (Köppen et al. 2007).

- *the size–mass relation* : Galaxies exhibit a correlation between their mass or luminosity and their size or surface brightness ( Shen et al. 2003; see also Poggianti et al. 2013b). It is known that the two main classes of galaxies, early and late types, show a different dependency between size and stellar mass with LTGs having a weak, if any, dependence on the redshift, and ETGs showing a clear variation with the redshift (see e.g. Shen et al. 2003, van der Wel et al. 2014), which suggest a different evolution pattern for the two populations. However, the evolution with redshift depends strongly on the galaxy mass (see e.g. Furlong et al. 2015).
- *The Faber–Jackson Relation (FJR)*: Ellipticals exhibit a significant correlation of its residuals with galaxy size, in the sense that smaller than average galaxies have larger than average velocity dispersion (Faber & Jackson 1976. According to this relation more massive galaxies are also the more luminous. The original relation can be expressed mathematically as:  $L \propto \sigma^\gamma$  where the index  $\gamma$  is observed to be approximately equal to 4, but depends on the range of galaxy luminosities that is fitted. The Faber–Jackson relation provides a means of estimating the distance to a galaxy, which is otherwise hard to obtain, by relating it to more easily observable properties of the galaxy. In the case of elliptical galaxies, if one can measure the central stellar velocity dispersion, which can be done relatively easily by using spectroscopy to measure the Doppler shift of light emitted by the stars, then one can obtain an estimate of the true luminosity of the galaxy via the FJR. This can be compared to the apparent magnitude of the galaxy, which provides an estimate of the distance modulus and, hence, the distance to the galaxy.
- *The Kormendy Relation*: Shows a correlation between the effective radii of galaxies and their surface brightnesses at that radius (Kormendy 1977). The relation indicates that at the effective radius, large (massive) galaxies are fainter than small galaxies, hence shows large galaxies are less dense than small galaxies, an important consideration for models of galaxy formation.
- *The fundamental plane*: Usually expressed as a relationship between the effective radius, average surface brightness and central velocity dispersion of normal elliptical galaxies (Djorgovski & Davis 1987; Dressler et al. 1987; D’Onofrio

et al. 2008; D’Onofrio et al. 2016). If luminosity is proportional to mass then the virial theorem implies that  $L \propto R_{eff}\sigma^2$  where  $R_{eff}$  is the effective radius and  $\sigma$  is the line-of-sight velocity dispersion.

### 1.4.2 Late-Type Galaxies

Analogous scaling relations are found for LTGs, and we list here below

- *size–mass*; similarly to ETGs also late-type galaxies show a correlation between size and mass but the slope is different with respect to the one of the ETGs (see Shen et al. 2003), suggesting different assembly mechanisms between the two classes.
- *Tully–Fisher relation (TFR)*: Empirical relationship between the mass or intrinsic luminosity and angular velocity or emission line width of a spiral galaxy (Tully & Fisher 1977). The TFR is one of the most widely used methods of measuring extragalactic distances since spiral galaxies are relatively common and contain young, massive stars. The rotational velocity serves as a predictor of the luminosity, or absolute magnitude, of the galaxy. The distance is calculated from the distance modulus. It is analogous to the FJR relation for ETGs.

Using the above mentioned relation for different types of galaxies, one can study the evolution of galaxies from past to present. The existence of different correlations of the physical properties indicates that for all galaxies within a particular galaxy type follows a similar the formation and evolution path. The relations therefore provide insights into both the formation and evolution of galaxies, and many are also used to measure the distances to galaxies.

## 1.5 Galaxy Evolution from Observations

Early-type galaxies (ETGs) play an important role in the observational studies of galaxy formation and evolution as they contain the traces of evolution from past to present. We quantify this evolution by correlating the galaxy physical properties at different redshifts. ETGs are known to follow well-defined empirical scaling laws that relate their global observational properties: the Faber–Jackson (FJ; Faber & Jackson 1976), Kormendy (1977), fundamental plane (FP; Djorgovski & Davis 1987; Dressler et al. 1987), size–mass (Shen et al. 2003, Hyde & Bernardi 2009), color–mass), color– $\sigma$  (Bower et al. 1992), Mg2– $\sigma$  (e.g., Guzman et al. 1992; Bernardi et al. 2003),  $M_{BH} - M_*$  or  $M_{BH} - \sigma$  (de Zeeuw 2001; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002), color gradient vs mass (La Barbera et al. 2010a; Tortora et al. 2010), dynamical vs stellar mass (Tortora

et al. 2009, 2012) relations. These scaling relations provide invaluable information about the formation and evolution of ETGs, setting stringent constraints to galaxy formation models. In particular, studying the structural and mass properties of galaxies at different redshifts can thus give more insights into the formation and evolution of galaxies over time.

Early-type galaxies (ETGs) follow a steep relation between their size and the stellar mass, the so called, size-mass relation. They are also found to be much more compact in the past with respect to local counterparts (Daddi et al. 2005; Trujillo et al. 2006; Trujillo et al. 2007; Saglia et al. 2010; Trujillo et al. 2011, etc.). A simple monolithic-like scenario, where the bulk of the stars is formed in a single dissipative event, followed by a passive evolution, is not longer supported by these observations. Thus, several explanations have been offered for the dramatic size difference between local massive galaxies and quiescent galaxies at high redshift. The simplest is related to measurement systematics, since observers could have underestimated the sizes and/or overestimated the masses. However, recent studies suggest that it is difficult to change the sizes and the masses by more than a factor of 1.5, unless the initial mass function (IMF) is strongly altered (e.g., Muzzin et al. 2009; Cassata et al. 2010; Szomoru et al. 2010). Other explanations include extreme mass loss due to a quasar-driven wind (Fan et al. 2008), strong radial age gradients leading to large differences between mass-weighted and luminosity-weighted ages (Hopkins et al. 2009b; La Barbera & de Carvalho 2009), star formation due to gas accretion (Franx et al. 2008), and selection effects (e.g., van Dokkum et al. 2008; van der Wel et al. 2009).

The best candidate to represent the galaxy evolution scenario is galaxy merging. Over time, galaxies were attracted to one another by the force of their gravity, and collided together in a series of mergers. As cosmic time evolves the high- $z$  “red nuggets” are thought to merge and evolve into the present-day massive and extended galaxies. ETGs undergo mergings at different epochs, becoming metal rich, massive and red in color (Kauffmann 1996). Better than major mergers, the most plausible mechanism to explain this size and mass accretion is minor merging (e.g., Bezanson et al. 2009; Naab et al. 2009; van Dokkum et al. 2010; Hilz et al. 2013; Tortora et al. 2014b). Numerical simulations predict that such mergers are frequent (Guo & White 2008; Naab et al. 2009), leading to observed stronger size growth than mass growth (Bezanson et al. 2009). The minor merging scenario can also explain the joint observed evolution of size and central dark matter (Tortora et al. (2014b); Tortora et al. 2017, in preparation). However, interestingly, recently it has been found that a tiny fraction of the high- $z$  red nuggets might survive intact till the present epoch, without any merging experience, resulting in compact, relic systems in the nearby Universe (Tortora et al. 2016; Trujillo et al. 2012).

Investigation on size evolution of galaxies over time have been performed in the recent years based on different surveys. The most prominent surveys that have been

done in the last two decades are DEEP2 (galaxies within the redshift range  $0.75 < z < 1.4$  Davis et al. 2003); GAMA (250 square degrees with galaxies up to redshift 0.4: Driver et al. 2011; Driver et al. 2011); 2dFGRS (measuring redshifts for 250 000 galaxies; Colless et al. 2001a), and SDSS (10000 square degrees in northern sky in u, g, r, i and z bands; York et al. 2000). SDSS is one of the most successful surveys in the field of galaxy evolution studies (Kauffmann et al. 2003) in the recent years.

Better quality instrumentations and telescopes mean better database and scientific results. And some more recent surveys, like KiDS@VST, are aiming at improving the data quality, pushing the SDSS results to larger redshifts and lower masses. KiDS is a large scale optical imaging survey, which will cover 1500 square degrees over u, g, r and i bands, using VLT Survey telescope (VST) and OmegaCAM (de Jong et al. 2015; de Jong et al. 2017). The high spatial resolution of VST (0.2"/pixel), the better seeing conditions and the depth (KiDS is  $\sim 2$  magnitudes deeper than SDSS, de Jong et al. (2015)) makes the survey good for the galaxy evolution studies, and more promising for our scientific goals. Extrapolating the results presented in this paper to the whole survey, we plan to measure the structural parameters in the four KiDS wavebands for 4 million galaxies, at redshifts  $z \lesssim 0.7$  (Tortora et al. 2016).



# Chapter 2

## Kilo Degree Survey

Sky surveys serve as a basic data in the field of astronomy. They represent nowadays the largest data generators in astronomy, propelled by the advances in information and computation technology, and have transformed the ways in which astronomy is done. This trend is bound to continue, especially with the new generation of synoptic sky surveys that cover wide areas of the sky repeatedly, and open a new time domain of discovery. Surveys enable a very wide range of science, and that is perhaps their key unifying characteristic. They can be used to generate large, statistical samples of objects that can be studied as populations, or as tracers of larger structures to which they belong.

### 2.1 VLT Survey Telescope

The VLT Survey Telescope (VST) is the latest telescope added to ESO's Paranal Observatory in the Atacama Desert of northern Chile (Capaccioli & Schipani 2011; Kuijken 2011). It is placed in an enclosure adjacent to the four Very Large Telescope (VLT) Unit Telescopes on the summit of Cerro Paranal. The VST (Fig. 2.1) is a wide-field survey telescope with a field of view twice as broad as the full Moon. It is the largest telescope in the world designed to exclusively survey the sky in visible light. Like the VLT, the new survey telescope covers a wide-range of wavelengths from ultraviolet through optical to the near-infrared (0.3 to 1.0 microns). But whereas the largest telescopes, such as the VLT, can only study a small part of the sky at any one time, the VST is designed to photograph large areas quickly and deeply. With a total field view of  $1^\circ \times 1^\circ$ , twice as broad as the full Moon, the VST was conceived to support the VLT with wide-angle imaging by detecting and pre-characterising sources, which the VLT Unit Telescopes can then observe further.

The VST is comprised of two mirrors, a primary mirror (M1) with a diameter of 2.61 m and a smaller secondary mirror (M2) with a 93.8 cm diameter. The telescope

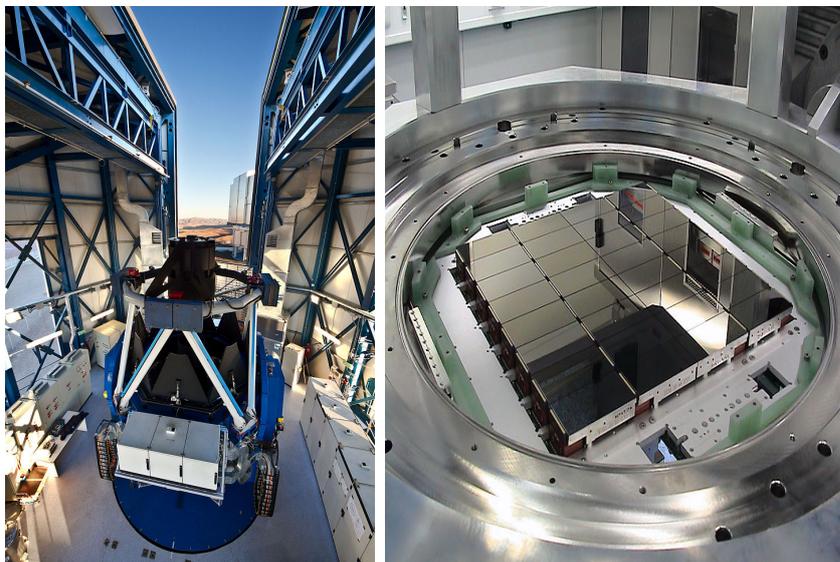


Figure 2.1: The VLT Survey Telescope(left) and OmegaCAM camera lies at the heart of the VST. This view shows its 32 CCD detectors that together create 268-megapixel images(right).

is equipped with pixel scale of  $0.2''/\text{pixel}$ , is equipped with 5 filters (the Sloan-like filters *ugriz*). OmegaCAM was built by an international consortium of five institutions: Netherlands Research School for Astronomy (NOVA), Kapteyn Astronomical Institute, Universitäts-Sternwarte München, Astronomical Observatory of Padova and ESO.

The VST project is a joint venture between the European Southern Observatory and the Capodimonte Astronomical Observatory (OAC), part of the Italian National Institute for Astrophysics (INAF). The Italian center designed the telescope while ESO is responsible for the civil engineering works at the site. The VST will be dedicated to survey programs. With its state-of-the-art camera, the quality of its optics, and the exceptional seeing conditions of Paranal, the VST is making important advances in many fields of research, from the Milky Way to cosmology. Most of the telescope time is dedicated to survey programs solicited by ESO (the ESO Public Surveys), but a fraction of the telescope time is also dedicated to GTO large programs, handled by the telescope and camera consortia, plus there is a 10% of time dedicated to Chilean programs. To give a perspective of the science currently done by the VST, we briefly summarize the three ESO Public Surveys (see also the ESO webpage: [www.eso.org/sci/observing/PublicSurveys/sciencePublicSurveys.html#VST](http://www.eso.org/sci/observing/PublicSurveys/sciencePublicSurveys.html#VST))

*VPHAS+*: The VST Photometric H- $\alpha$  Survey of the Southern Galactic Plane. PI Janet Drew (Imperial College of London). This survey will combine H- $\alpha$  and broadband *ugri* imaging over an area of 1800 square degrees capturing the whole of the

Southern Galactic Plane within the latitude range  $|b| < 5$  degrees. VPHAS+ will facilitate detailed extinction mapping of the Galactic Plane, and can be used to map the structure of the Galactic disk and its star formation history.

*VST - ATLAS*: PI Tom Shanks (Durham University). ATLAS is a 4500 square degrees survey on the Southern Sky in 5 filters ( $u, g, r, i, z$ ) down to a depth comparable to the SDSS. The aims of the survey are: 1) dark energy equation of state by examining the 'baryon wiggles' in the matter power spectrum, 2) imaging base for many other future spectroscopic survey (VLT AAOmega); study of high redshift galaxies and quasars.

*KiDS*: PI Konrad Kuijken (Leiden University). This is a 1500 square degrees in 4 bands ( $ugri + \text{NIR VISTA/VIKING survey}$ ) 2.5 magnitudes deeper than the Sloan Digital Sky Survey (SDSS). Mean science topics are: 1) weak gravitational lensing for dark matter halos and dark energy studies; 2) investigating galaxy evolution; 3) searching for galaxy clusters high redshift quasars. See also next Section.

## 2.2 Kilo Degree Survey

KiDS, the Kilo-Degree Survey, is a new, large-scale optical imaging survey in the Southern sky, designed to study some of the most fundamental questions of cosmology and galaxy formation of today (de Jong et al. 2015). Using the VLT Survey Telescope (VST) KiDS will map 1500 square degrees of the night sky in four filters ( $u, g, r, i$ ).

The two fields (Fig. 2.2) have been selected to overlap with previous, publicly available huge galaxy redshift surveys (2dF and SDSS together measured some 300,000 redshifts in the KiDS areas), which means we already have a good map of the foreground galaxy distribution, out to redshift 0.3, and the spectral diagnostics for several 100,000 galaxies. One of the major aims of KiDS is to 'weigh' these galaxies systematically as function of their type and environment. Image quality, reliability and throughput were the key design drivers, and the expectations have been fulfilled to a point unmatched by any other similarly wide-field camera.

The central science case for KiDS is mapping the matter distribution in the universe. This will be done through weak gravitational lensing and photometric redshift measurements. However, the enormous data set that KiDS will produce can also be used for many different science aims like Dark Energy equation of state, search of galaxy clusters and characterization of their properties and evolution, search of high-redshift quasars, strong lensing, search of stellar streams and satellites in the Galactic halo.

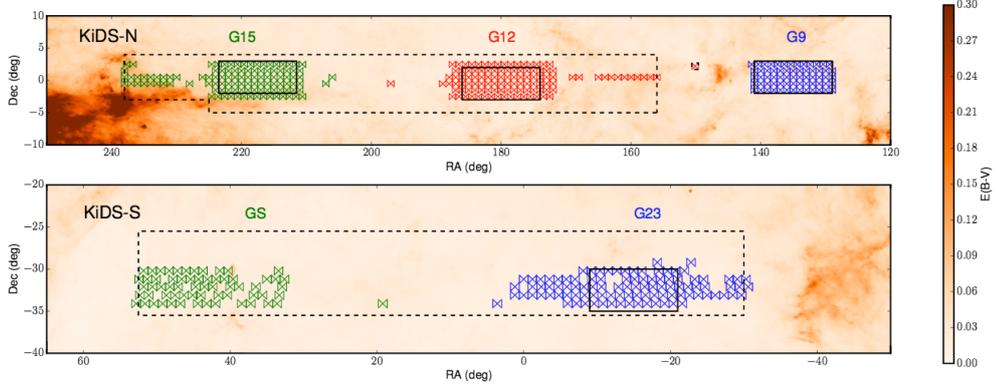


Figure 2.2: Layout of the KiDS survey fields, KiDS-N in the Northern galactic cap on the equator, and KiDS-S around the South Galactic Pole. The dashed contours outline the full KiDS area (observations ongoing). The solid rectangles indicate the areas observed by the GAMA spectroscopic survey.

## 2.3 Data analysis

The first and second data releases of the KiDS collaboration (KiDS-DR1/2) are extensively described in de Jong et al. (2015), hereafter dJ+15. This consists of 148 tiles observed in all four filters during the first year of regular operations (15 October 2011 to 31 September 2012), including data taken during Early Science Time (EST, 13 August to 15 October 2011), for a total of 50 tiles (DR1) and the second year of regular operations (1 October 2012 to 31 September 2013), for a total of 98 tiles (DR2). Here below we provide some brief details of the release products which are relevant for the work in this thesis.

### 2.3.1 Data reduction and Image Coaddition

The data processing pipeline used for KiDS-DR1/2 makes use of the Astro-WISE environment (AW hereafter, McFarland et al. 2013). Here the standard pre-reduction steps of the raw data are fully automatized and include: cross-talk correction, de-biasing and overscan correction, flat-fielding, illumination correction, de-fringing, pixel masking (cosmic-rays, hot and cold pixels, saturated pixels), satellite track removal, background subtraction.

The photometric calibration consists on the determination of individual zero points per CCD based on SA field observations. Basically, the calibration is made by matching fixed aperture (6.3'' diameter) not corrected for flux losses, with SDSS DR8 PSF magnitudes of stars in the SA fields as reference (AB system). The final *gri* photometry is homogenized across CCDs and dithers for each tile independently (*u*-band was not homogenized due to small source density which did not allow enough

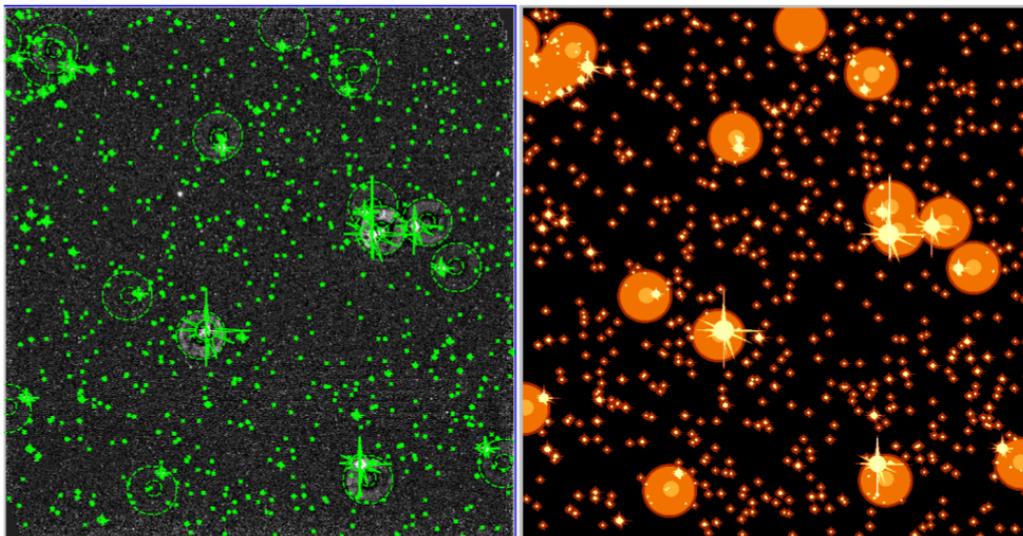


Figure 2.3: Left: region file produced by the masking software Pullecenella v1.0. Star spikes, primary, secondary and tertiary haloes are evident for the bright stars, while for fainter stars spikes and primary haloes are mainly identified. Right: the corresponding multi-flag file is shown, with the different critical areas colorized with different orange grades (see also de Jong et al. 2015 for further details).

overlaps between adjacent CCDs). The global photometric solution is derived by the overlapping sources across contiguous dithers, and minimizing the zero point differences among the CCDs with respect to an average zero point valid for the night.

The global (multi-CCD and multi-dither) astrometric calibration has been obtained using the SCAMP software (Bertin 2006) and calculated per filter per tile using 2MASS-PSC (Skrutskie et al. 2006) as astrometric reference catalog. To obtain a typical astrometric internal error of the order of  $0.01''$ , while the absolute error depends on the reference catalog adopted (typically tenths of arcsec), which are SDSS-Data Release 12 (Alam et al. 2015) for the Northern patches (KiDS-N) and the 2MASS catalogue (see Skrutskie et al. 2006) for the Southern fields (KiDS-S). The stacked images have then obtained after the software SWARP has been used to resample all dithers in a tile to the same pixel grid, with a pixel size of  $0.2''$ . After background subtraction the exposures are coadded using a weighted mean stacking procedure.

### 2.3.2 Masking and source flagging

This procedure consists in the automatic identification of the image areas affected by the presence of bright stars which produce typical saturation features (spikes) plus

star halos/ghosts. To perform this task we have used a software dubbed Pulecenella v1.0, which is also running under AW. The procedure is able to select in an automatic way different critical areas (saturation core, readout spikes, diffusion spikes, primary halo, secondary halo, tertiary halo) with no need of external star catalogs. Pulecenella makes use of S-Extractor to first analyze the input image and then process the masking products. It works in the pixel value space, by identifying all the pixels which hit the saturation level.

From the pixel mask image the following area are inferred:

- 1) saturated pixels in science image (with pixel value  $\geq$  fraction of image header SATURATION keyword value, the fraction can be set to e.g. 50%), null-value pixels (also found in science image), and zero-weighted pixels (found in weight image);
- 2) saturated star pattern with spikes and reflection features derived by the information of the image ROTATOR keyword in the image header of every single exposure of the Coadd;
- 3) the halo regions, as inferred by the star fluxes and position in the KiDS field.

The outputs are: 1) region file for visualization or possible manual change, and 2) a flag fits image with pixel values indicating for different mask regions. The flag fits image is to be used in the catalog extraction as flag input file. In Fig. 2.3 we show an example of the mask files provided by the procedure (see also dJ+15 for further details).

### 2.3.3 Catalog extraction and Star/Galaxy separation

KiDS single band source lists are based on a stand-alone procedure optimized for KiDS, dubbed KiDS-CAT, also fully implemented under AW. KiDS-CAT backbone is constituted by S-Extractor software (Bertin & Arnouts 1996) for the source detection and photometry, plus a series of steps to automatically extract catalogs of sources and per for a star/galaxy separation which are originally developed within the 2DPHOT environment (La Barbera et al. 2008, see §3.2). KiDS-CAT is ran with the following inputs: 1) a S-Extractor configuration file; 2) a KiDS coadd and the 3) corresponding image weighting map (both from AW pipeline), and 4) and external mask region (i.e. from Pulecenella masking software).

For sake of generality, KIDS-CAT philosophy is to best compromise the extraction efficiency (i.e. lower detection limit) and the undesired detection (spurious sources and objects affected by critical image regions). To this aim KIDS-CAT picks all potential real sources and then rank their reliability on the basis of their proximity of critical area (saturation features, star halos, junk pixels etc.).

Steps performed by KIDSCAT to create the single-filter sourcelists are:

- 1) run SExtractor with `SEEING_FWHM` set to 1.5";
- 2) find high-confidence star candidates;
- 3) apply the bi-weight estimators to the high-confidence star candidates and compute

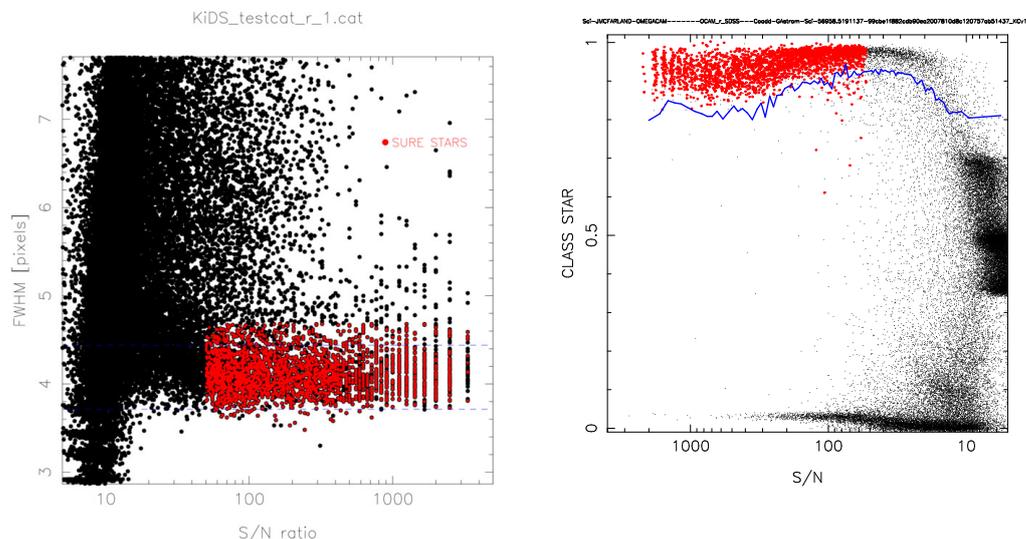


Figure 2.4: Left: “true star” definition on the basis of the distribution of the  $S/N$  vs FWHM (this case is in  $r$ -band). Sure stars are the sources with  $S/N > 50$  distributed along the sequence of the smaller FWHM compatible with the point-like PSF (see also text for details). Right: Example of star/galaxy separation diagram for a  $r$ -band KiDS tile. At  $S/N > 50$ , the high-confidence star candidates (red dots) are used to define the blue line; at lower  $S/N$ , all sources with  $\text{CLASS\_STAR} > 0.8$  are used. Sources above the blue line are classified as stars.

average the FWHM;

4) run SExtractor with `SEEING_FWHM` set to the computed average FWHM;

5) stars identified based on  $S/N$ -dependent `CLASS_STAR` criteria.

Typical spurious sources in KiDS catalog are of the order of 5% (see dJ+15 for details). In Fig. 2.4 we demonstrate the two main steps performed by KIDS-CAT for the identification of the “true stars” and for the subsequent star/galaxy (S/G, hereafter) separation.

Note that, according to the input mask files, there are two main parameters in the final catalogs that refer to the quality of the source. These are the flags `ISOFLAG_IMAGE` and `ISOAREA_IMAGE`, which correspond to the boolean sum of the critical area flags and the number of pixels affected by the same critical area, respectively. As a reference, to select sources that are not interested by any of the critical area, both parameters shall be set to zero.

From the FWHM vs  $S/N$  plots, where we have adopted  $S/N = 1.0857/\text{MAGAPER\_ERR}$  (from S-Extractor), for any given aperture, true stars are identified as (see La Barbera et al. 2008 for details):

1) no flag set on (including both S-Extractor and mask flags);

- 2)  $S/N \geq 50$ ;
- 3) S-Extractor keywords: `ISOAREA_IMAGE` > 7; `FWHM` > 0; `ELLIPTICITY` < 0.1 (or < 0.2 if number of selected objects is too small).

S/G algorithm is meant provide a general purpose separation of the two classes. The `FWHM` and `CLASS_STAR` parameters from S-Extractor as a function of the  $S/N$  are used to identify stars in two steps: a) In the  $S/N$  range where true stars are available (Fig. 2.4), the code applies the bi-weight estimator to define the `CLASS_STAR` location,  $\theta$ , and its width,  $\sigma$ ; then we define a lower envelope as  $\theta - 4\sigma$ ; b) if  $S/N$  is below that of true stars, we compute a running median trend, including only objects with `CLASS_STAR`>0.8, and shift it downwards to match the  $\theta - 4\sigma$  locus. The resulting blue curve in the plot defines the separation of stars and galaxies (i.e. objects with `CLASS_STAR` > and < the blue curve are tagged as stars and galaxies, respectively, see Fig. 2.4). This method is purely statistical, there might be some contaminants in the galaxy catalog. Note though that most of these contaminants become catastrophic sources in the photo-z procedure as they do not fit any galaxy spectroscopic template and are then excluded from our analysis

The `2DPHOT_FLAG` is the flags keeping the information on the S/G procedure. If an object is classified as a true star, the `2DPHOT` flag value is 1, while normal stars have `2DPHOT_FLAG`=4. `2DPHOT_FLAG`=0 means that the object is a galaxy. Different flag values are summed up (e.g. `2DPHOT_FLAG`=5 means the object is a true star, and is also classified as star according to the above procedure).

The catalogs of stars and galaxies will then be used to define completeness. Magnitudes are measured for all filters with S-Extractor within a large number of apertures from 2 pixels (i.e. 0.4'') to 64 pixels (i.e.  $\sim 12.8''$ ), which are ideal to make interpolation photometry and aperture correction. An approximation of total magnitude is provided by the Kron-like `MAG_AUTO`.

### 2.3.4 Completeness and limiting magnitude

The completeness algorithm is based on the procedure presented in Garilli et al. (1999), where the magnitude at which objects start to be lost in the catalog because they are below the brightness threshold in the detection cell is defined. The detection luminosity for the sources has been defined as a 2 pixels diameter magnitude from KIDS-CAT (hereafter, and in the plots, we will refer to this as `DETECTION_MAG`). This approximation should work fine for most purposes. Here the completeness magnitude is computed using the `MAG_AUTO` vs. `DETECTION_MAG` diagram, as illustrated in Fig. 2.5. In the plot stars and galaxies (as classified above) are both shown together with the objects whose `MAG_AUTO` and `DETECTION_MAG` have low  $S/N$  (<5 and 4, respectively). These latter (gray points) are excluded in the computation of the completeness mag to make the computation of detection limit more stable. All remaining sources in the catalog are also shown.

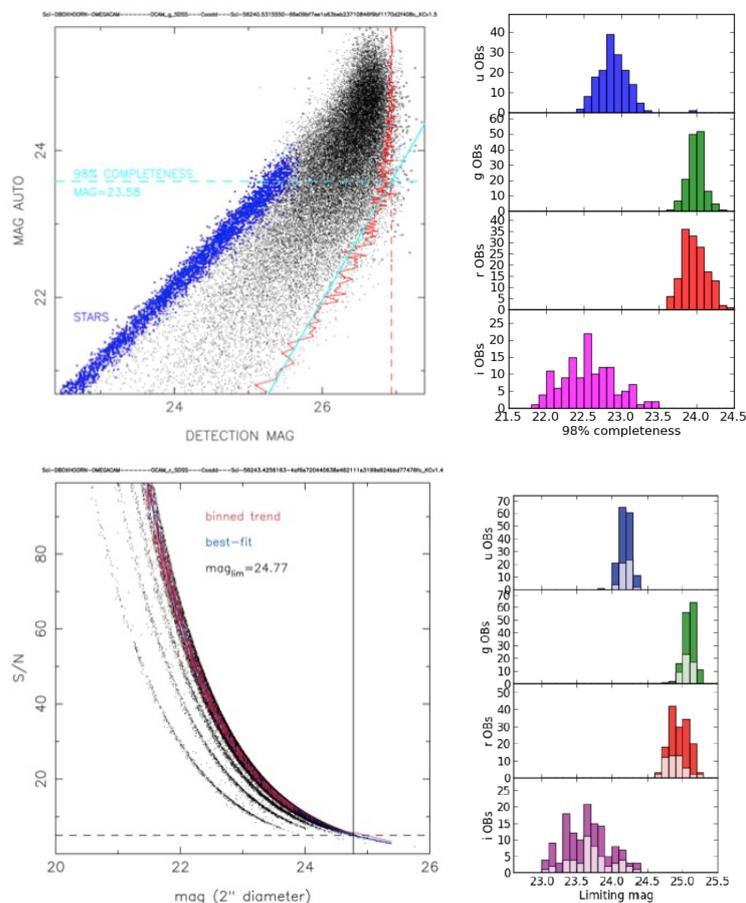


Figure 2.5: Top: procedure for the estimate of the 98% completeness magnitude. Left: MAG\_AUTO vs. detection magnitude in a  $g$ -band tile. Blue dots are true stars according to the definition as in §2.3.3, black dots all other sources with grey points showing sources with low  $S/N$  ( $< 5$ ). The red interpolation line show the 98% percentile of the detection magnitude in different MAG\_AUTO bins and the cyan line is the best linear fit to the latter. The intercept with the vertical red dashed line, indicating the DETECTION\_MAG where sources saturates at fainter DETECTION\_MAG, gives the 98% completeness limit. Right: distribution of the completeness limits in the KiDS-DR1/2 (DR1 with lighter colors). Bottom: procedure for the limiting magnitude. Left:  $S/N$  vs. magnitude within 2 arcsec diameter. The limiting magnitude is given by the magnitude corresponding to  $S/N=5$  obtained by the fit to distribution of the values in bins of  $S/N$ . Right: distribution of the limiting magnitudes in the KiDS-DR1/2 (DR1 with lighter colors).

The distribution of points in the diagram is approximately limited, by construction (S-Extractor detects objects if they are above a given  $S/N$  in the detection cell), with respect to DETECTION\_MAG. The detection limit (vertical dashed line) is defined by binning the data with respect to MAG\_AUTO, and computing in each bin

the 98-th percentile of the `DETECTION_MAG` distribution. This gives us the red, solid, curve in the diagram. A fixed part of the red curve, *below* the detection limit, is fitted by a least-squares line (cyan line).

The intersection of the cyan line with the detection limit (red, dashed line), gives you the completeness magnitude of the catalog (horizontal dashed line). By construction, the catalog should be more than 98% complete below this completeness value.

Notice that one might use a second-order polynomial fit, rather than a line, to model the solid, red, curve. After several experiments, we decided to keep the linear fit. This gives more stable results (in any case, the completeness estimates do not change by more than 0.1-0.2 mag). Notice that for *u*-band, because of the low number of galaxies with good *S/N*, the completeness estimates are somewhat uncertain, but reasonable.

The limiting magnitude procedure is based on the single plot of the *S/N* as a function of 2'' diameter magnitude. One example is shown in Fig. 2.5 for the field KIDS\_185.0-1.5, *r*-band. Here black points are all good (no masked sources of any kind) objects in the field. In red you can see the median trend, while the blue curve is the analytic fit  $p_1 * flux / (p_2 * flux + p_3)$  where  $p_1$ ,  $p_2$ , and  $p_3$  are free fitting parameters. This is essentially the analytic expression to compute *S/N* (where  $p_2$  is related to effective gain, while  $p_3$  is background noise). The analytic fit is used to compute the typical mag corresponding to `S/N_LIMIT=5` (see dashed and solid black lines in the plot) and it has the advantage that at low *S/N* this is not dominated by the many spurious detection that one expects to show up at very faint *S/N* limits. This is also useful to estimate the limiting magnitude for tiles where the binned mag-*S/N* trends do not go below  $S/N \sim 5$ .

The procedure can easily compute the same kind of limit but for different `SN_LIMIT` (i.e. 10 and 15).

### 2.3.5 Multi-band catalogs and final S/G separation

The multi-band catalog we will refer in the following, are the ones obtained for KiDS-DR1/2 and finally delivered to the community through ESO<sup>1</sup>. These contain all sources that are detected in the *r*-band sourcelists and are or are not detected in the other filters. This choice has been made ensure the re-collection of the deeper sample expected to be obtained in the survey and also the ones detected in the images with the highest image-quality, although this might exclude *r*-band drop-outs which an interested user shall find from the cross-matching of the single band sourcelists also delivered via ESO.

The multi-band catalog has been obtained by using S-Extractor in dual-image mode and make use of all the parameters that can be found in the single-band

<sup>1</sup><http://archive.eso.org/cms.html>

source lists. Indeed S-Extractor in dual-image mode takes the list of sources detected in the reference image ( $r$ -band in our case) and measures the flux corresponding to the same centroid position in the dual image (all other band in turn), according to the set-up specified for the  $r$ -band extraction. In this case, the weight map and the mask files from Pulleccenella are also referred to the reference image ( $r$ -band). This might produce an inconsistency on the number of source flagged as “reliable” (e.g. having the mask flags `ISOFLAG_IMAGE=0` and `ISOAREA_IMAGE=0` set to the  $r$ -band reference value, that might be different in other bands), however, as the pointing and pattern of the major critical area are very similar across the filters and the  $r$ -band is the deeper filter (i.e. also the star features are generally brighter), this is generally not affecting too much the final source reliability.

The aperture corrected fluxes are part of the products provided by the KiDS survey to account for the seeing differences among the different filters (e.g. for stellar science purposes). These can be used to define fluxes, which enclose the same fraction of total light in different bands. We did not use these magnitudes in our SED analysis as we needed to use total luminosities (and colors) and thus we adopted aperture large enough to account for the total light independently of the seeing (i.e. 6" diameter). The aperture corrections were calculated for each of the four filter ( $ugri$ ) by comparing the aperture fluxes with the flux in a 6" aperture, the aperture used for photometric calibration, and the aperture-corrected fluxes are included in the catalog as separate columns. Source magnitudes and fluxes are finally corrected for Galactic foreground or intergalactic extinction using the extinction map of Schlegel et al. (1998). For most of the sky region we have scanned, extinction is smaller than  $\sim 0.06$  mag in  $r$  band. An approximation of total magnitude is provided by the Kron-like `MAG_AUTO`.

In order to prevent sources in tile overlaps to appear as multiple entries in the catalog, the survey tiles have been cropped and connect seamlessly to one another. This results in slightly shallower data along the edges of the tiles, similar to the areas partially covered by CCD gaps. Overall, all included areas are covered by at least three exposures

Although one might have tried to combine the info from the different filters, for the first data releases we adopted the S/G separation, positional and shape parameters based on the  $r$ -band data.

We have also retained those sources with  $r$ -band S-Extractor `FLAGS_r < 4`, including objects that have very close and bright objects, bad pixels or was originally blended with another one.

### 2.3.6 The galaxy sample

The original multi-band catalog related to the KiDS-DR1 and 2 include  $\sim 22$  millions of sources. The Star/Galaxy separation has returned  $\sim 12$  millions of galaxies which

reduced to  $\sim 7$  millions after having removed those galaxies in the masked (critical) regions (see §2.3.3). Unfortunately for the first releases of KiDS there was a large manually masked area which, together with the automatic masks (for star haloes and spikes), ended up with a large fraction of observed area masked out. In terms of effective area this is about 33% of the original observed area ( $103/153 \text{ deg}^2$ ). This partially accounts for the fraction of the objects lost ( $1-7\text{M}/12\text{M}=42\%$ ), being the residual fraction (i.e. 9%) due to the fact that the masked regions contain generally a large fraction of spurious sources that make these regions generally more densely populated than regions with no spurious sources. The unmasked effective area adopted is of 103 square degrees. The “unmasked” sample is the one we will use as a reference for the following steps which are the estimate of the photometric redshifts and the study of the structural parameters. To perform the latter analysis, only very high-quality sources have to be taken into account (La Barbera et al. 2008, 2010b), i.e. systems with the highest  $S/N$  in  $r$ -band ( $S/N_r > 50$ ) (see more details in §3.1).

The gain in terms of morphological details and density of source is evident. In particular, with KiDS we have almost factor four more galaxies per square degree, i.e. on average  $\sim 4.7 \times 10^4/\text{deg}^2$  unique sources classified as galaxy by the S/G procedure against  $\sim 1.4 \times 10^4/\text{deg}^2$  of SDSS (based on DR10 data). The increased statistics and the highest image quality will also help in the discovery of strong gravitational events, which is part of a separate project (see, e.g., Napolitano et al. 2016, more details in Chapter 6).

### 2.3.7 Photometric redshifts

Photometric redshifts are derived from KiDS  $ugri$  photometry using two independent methods: 1) the supervised machine learning model MLPQNA: a Multi-Layer Perceptron Quasi Newton Algorithm, and 2) Spectral Energy Distribution (SED) fitting, using the software `Le Phare` (Arnouts et al. 1999).

The MLPQNA is a feed-forward neural network method, providing a general framework for representing nonlinear functional mappings between input and output variables. QNA stands for Quasi Newton Algorithm, a variable metric method used to solve optimization problems (Davidon 1991) that, when implemented as the learning rule of a MLP, can be used to find the stationary (i.e. the zero gradient) point of the learning function. The QNA implemented here is the L-BFGS algorithm (Shanno 1975). Supervised methods use an extensive set (the knowledge base or KB) of objects for which the output (in this case the redshift) is known a-priori to learn the mapping function that transforms the input data (in this case the photometric quantities) into the desired output. Usually the KB is split into three different subsets: a training set for training the method, a validation set for validating the training in particular against overfitting, and a test set for evaluating

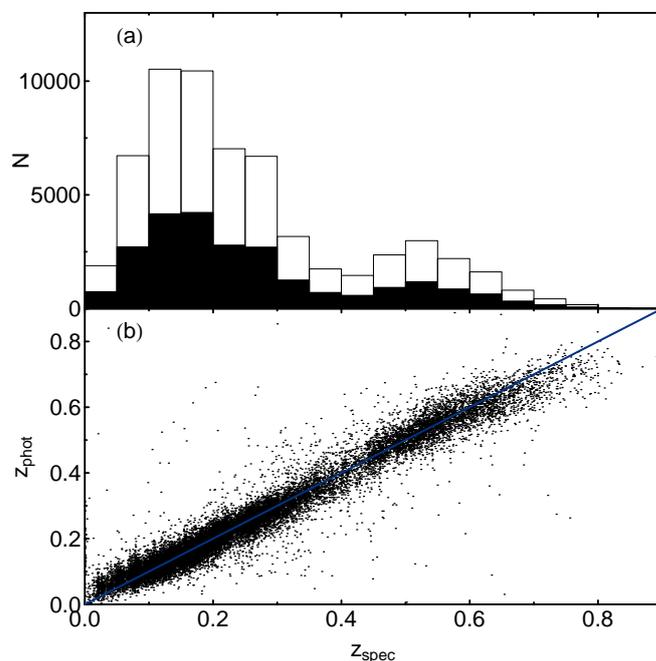


Figure 2.6: Photometric redshifts. In the panel (a) we show the distribution of the full sample with spectroscopic redshifts (white bars) and the distribution of redshifts in the test sample (black bars). In the panel (b) spectroscopic redshift are plotted vs. photometric ones for the test sample galaxies. Blue line is the one-to-one relation.

the overall performance of the model (Cavuoti et al. 2012; Brescia et al. 2013). In the method used here, the validation and the training phases are performed together, by means of the standard leave-one-out k-fold cross validation mechanism (Geisser 1975). Performances are always derived blindly, i.e. using a test set including only those objects which have never been fed to the network during either training or validation. This method has been extensively applied to different dataset/KBs, often composed through accurate cross-matching among public surveys, with good results (SDSS for galaxies: Brescia et al. 2014; UKIDSS, SDSS, GALEX and WISE for quasars: Brescia et al. 2013; CLASH-VLT data for galaxies: Biviano et al. 2013).

For the KiDS dataset, the KB of spectroscopic redshifts has been obtained by merging spectroscopic dataset from GAMA (Galaxy And Mass Assembly) data release 2 (Driver et al. 2011) and SDSS-III data release 9 (Ahn et al. 2012). For the KiDS photometry, we have adopted three different aperture magnitudes. The training set and a test set have been assembled by splitting the KB in two parts of 60% and 40%, respectively.

The disadvantage of the machine learning technique resides on the fact that  $z_{\text{phot}}$  can be accurately determined only for the sources having photometric properties in the range covered by the training sample, which implies that a cut in magnitude

had to be performed. For the sample we will discuss in this work we adopted cut on `MAGAP_6` 24.7, 24.0, 22.0, and 21.0 in *ugri* respectively, which provided us a final sample of  $\sim 1.2$  million of sources with reliable  $z_{\text{phot}}$ . Further details about experiments and the comparison with SED fitting procedures will be given in a forthcoming paper (Brescia et al., in preparation).

The whole distribution of the spectroscopic redshifts is shown in Fig. 2.6 (panel a), as the one for the test sample. The result of our reference experiments are shown in the panel (b), where the spectroscopic redshifts are plotted as a function of the estimated photometric ones. The scatter in the quantity  $|(z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}})|$  is  $\sim 0.03$ , as well as the other statistical indicators, which tell that our results are resembling quite accurately the spectroscopic ones, although at lower  $z$  the difference between photometric and spectroscopic redshifts can be of the order of 40% or higher. This experiment is finally used to produce the final catalogue of redshifts for our sample (see Cavuoti et al. 2015).

The SED fitting `Le Phare` (Arnouts et al. 1999; Ilbert et al. 2006) method performs a simple  $\chi^2$  minimization between the synthetic SEDs and data. Each theoretical SED has the redshift as a free parameter and is convolved with the KiDS filter transmission curves. The merit function  $\chi^2$  is defined as

$$\chi^2 = \sum_{f=1}^{N_f} \left( \frac{F_{\text{obs}}^f - A \times F_{\text{pred}}}{\sigma_{\text{obs}}^f} \right)^2, \quad (2.1)$$

where  $F_{\text{pred}}$  is the flux predicted for a template  $T$  at redshift  $z$ .  $F_{\text{obs}}^f$  is the observed flux and  $\sigma_{\text{obs}}^f$  the associated error. The index  $f$  refers to the considered filter and  $N_f = 4$  is the number of filter. We adopt the observed *ugri* magnitudes (and related  $1\sigma$  uncertainties) within a  $3''$  aperture, which are corrected for Galactic extinction using the map in Schlegel et al. (1998).

We have checked that the SED fitting technique performed sensibly worse than the MLPQNA as shown in Fig. 2.7, where it is seen that the scatter of the former is significantly larger than the one of the latter, and the same is also found for the bias (Longo et al., in preparation).

For this reason we decided to assume the MLPQNA estimates as our reference  $z_{\text{phot}}$  sample, although these are more limited to the property of the training sample, especially the depth (see above).

### 2.3.8 Stellar masses and galaxy classification

Stellar masses, rest-frame luminosities from stellar population synthesis (SPS) models and a galaxy spectral-type classification are obtained by means of the SED fitting with `Le Phare`, where this time the redshift is not a free parameter but it is fixed to the  $z_{\text{phot}}$  obtained with MLPQNA. We have computed the statistical errors on

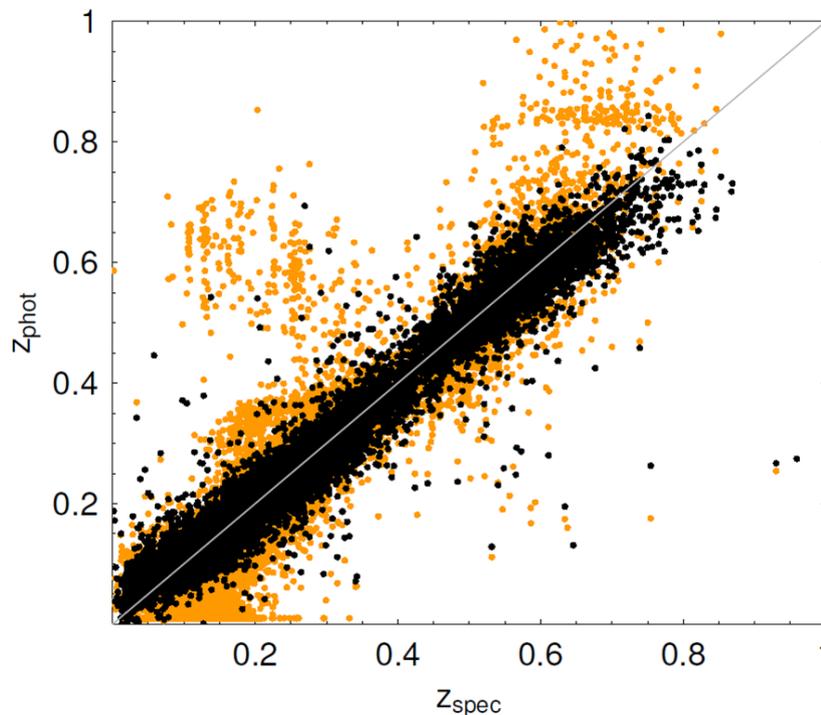


Figure 2.7: Le Phare vs. MLPQNA comparison. In yellow, we have the photo- $z$  from Le Phare and in black from MLPQNA for the KiDS data set.

masses, by generating random colors from a Gaussian with standard deviation given by the color errors and using these in the SED fitting procedure, to obtain the statistical errors of the Masses over 100 random extractions of a subsample of our galaxies. The average errors on masses are of the order of 0.2dex (maximum errors reaching 0.3dex).

This implies that the sample for which we can have classification and the full set of physical parameters and perform accurate studies of their evolutionary pattern is the one for which we also possess some information of their photo- $z$ . *In the following, we will refer to this sample, made up of  $\sim 1.2M$  of galaxies, as the “galaxy sample”.* In particular, we have adopted single burst SPS models from Bruzual & Charlot (2003), with a Chabrier (2001) IMF. Total magnitudes derived from the Sérsic fitting,  $m_S$ , (see §3.2.1) are used to correct the outcomes of Le Phare for missing flux. We have also investigated the effect of a change in the set of SPS models, allowing for a prolonged star formation.

Moreover, SED fitting can be used to provide information about the morphological type. As template set for this aim we adopt the 66 SEDs used for the CFHTLS in Ilbert et al. (2006). The set is based on the four basic templates (Ell, Sbc, Scd, Irr) in Coleman et al. (1980), and star burst models from Kinney et al. (1996). Syn-

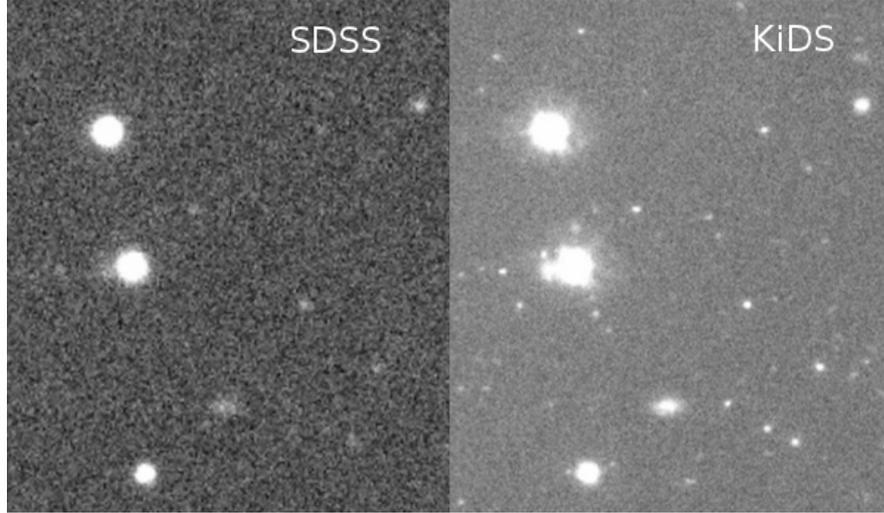


Figure 2.8: Comparison of SDSS r-band (left) and KiDS r-band (right) images.

thetic models from Bruzual & Charlot (2003) are used to linearly extrapolate this set of templates into ultraviolet and near-infrared. The final set of 66 templates (22 for ellipticals, 17 for Sbc, 12 for Scd, 11 for Im, and 4 for starburst) is obtained by linearly interpolating the original templates, in order to improve the sampling of the redshift-color space and therefore the accuracy of the redshift measurement. We do not account for internal extinction, to limit the number of free parameters. This fitting procedure provide us with a rough galaxy classification, which allows us to separate galaxies living on the red sequence and the blue cloud.

## 2.4 Comparison with other surveys

The Sloan Digital Sky Survey (SDSS) was the first “all-sky” CCD-based optical survey, setting the tone for the next generation of large-scale surveys. KiDS uses broadband filters similar to SDSS, but is considerably smaller in survey area: 1500 sq.deg for KiDS versus 20,000 sq.deg for SDSS (including SDSS II and SDSS III). However, KiDS has two major improvements that are crucial for its science goals: it is much more sensitive (by about 2 magnitudes), and it has better image quality, particularly in the r-band. Fig. 2.8 demonstrates these improvements by comparing part of a KiDS and SDSS r-band image. The CFHTLenS survey is a 155 sq.deg multi-colour optical survey, based on the CFHT Legacy Survey and optimized for weak lensing analysis. As the main science goals of CFHTLenS are the same as for KiDS and the science teams have a large overlap, this survey can be seen as a precursor to KiDS. The weak lensing analysis of CFHTLenS is based on i-band data, which is deeper than the KiDS r-band data, as shown in Fig. 2.9. However,

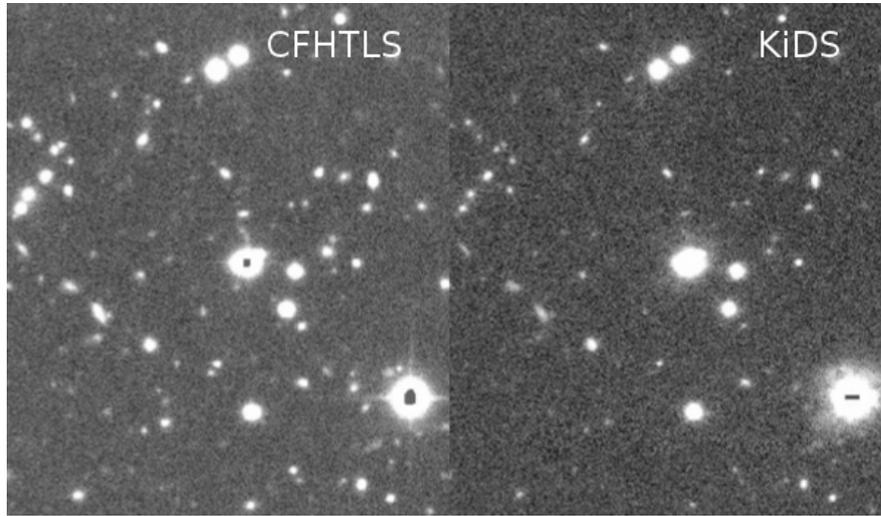


Figure 2.9: Comparison of CFHTLenS i-band (left) and KiDS r-band (right), the weak lensing bands of these two surveys.

again KiDS has two major advantages: the survey area is 10 times larger, and the PSF is much more constant over the field-of-view, simplifying the lensing analysis significantly.

## 2.5 VISTA Kilo-degree Infrared Galaxy (VIKING) Survey

VIKING is the infrared counter part of KiDS (Edge et al. 2013). It will cover 1500 square degrees of extra-galactic sky in  $z$ ,  $Y$ ,  $J$ ,  $H$ , and  $K_s$ . Accurate photometric redshifts can be determined by making use of these near-infrared measurements. Even-though, this thesis is mainly focusing on the data from optical survey KiDS, we use VIKING data for the study of compact galaxies in Chapter 5.



# Chapter 3

## Galaxy structural parameters: Sample Selection, Data Processing and validation

In this chapter, galaxy sample selection, methods to derive photometric redshift, galaxy classification, and stellar mass, extraction of structural parameters, validation of structural parameters, and classification of passive spheroids and disc-dominated galaxies are explained (Since the classification of galaxies based on the morphology always can contain some contaminations, we name ETGs and LTGs as passive spheroids and disc dominated galaxies from now on).

### 3.1 Sample Selection

The sample adopted in this analysis consists of galaxies extracted from 153 square degrees of the KiDS survey (de Jong et al. 2015) which have been already presented in Tortora et al. (2016). Details about the data reduction and calibration can be found in de Jong et al. (2015) and in Chapter 2. In the following we give a brief summary of the way the galaxy sample has been performed.

Single band source lists for the observed tiles are extracted using a stand-alone procedure named KiDS-CAT, which uses SExtractor (Bertin & Arnouts 1996) for the source detection, star galaxy separation and the catalog extraction. In particular, the star/galaxy (S/G) separation is based on the `CLASS_STAR` parameter from SExtractor measured on the r-band images, the deepest and best seeing ones for KiDS, following the procedure described in de Jong et al. (2015) (Sect. 4.5.1).

If the S/G separation is mainly based on the single  $r$ -band shape information, source colors are measured through multi-band source catalogs, which have been obtained using SExtractor in dual image mode by taking the  $r$ -band im-

ages as reference for source extraction and measuring the source fluxes in the other bands at the sky position of the  $r$ -band detection. The latter quantities have been used to perform the stellar population synthesis as described in §3.1.3. Among the sources selected as galaxies ( $\sim 11$  millions), we have retained those sources which were marked as being out of critical area from our masking procedure (see de Jong et al. 2015, Sect. 4.5.1). The effective uncritical area has been found to be 105 square deg (see Tortora et al. 2016), which contains  $\sim 7$  million galaxies. This latter sample turned out to be complete out to  $\sim 24$  mag in  $r$ -band by comparing the galaxy counts as a function of extinction corrected `MAG_AUTO`<sup>1</sup> with previous literature (e.g. Yasuda et al. 2001, Arnouts et al. 2001, McCracken et al. 2003, Capak et al. 2004, Kashikawa et al. 2004) as shown in Fig. 3.1.

There seems to be a statistically significant lack of bright galaxies in our sample with respect to other sample from literature (Yasuda et al. 2001). This is due to a combination of factors: first, the limited area we are considering in this work (105 deg<sup>2</sup>), which intrinsically limits the number of observed bright systems (mainly low- $z$  galaxies); secondly, the saturation of the galaxy centers for the very bright systems; thirdly the exclusion of deblended sources. We have estimated that all these effects made us to miss a few hundred systems in our automated selection in the bright side of the luminosity function. Among all these effects, though the effect of the area is the more severe: our preliminary tests of number counts on the recent KiDS third release (de Jong et al. in preparation) show that the effect is considerably alleviated. Finally, in order to perform accurate structural parameter measurement for these systems, we have selected galaxies with “high” signal-to-noise ( $S/N$ ), defined as  $1/\text{MAGERR\_AUTO}$  (Bertin & Arnouts 1996). Specifically, we have used  $S/N > 50$  as initial guess for reliable structural parameters (La Barbera et al. 2008). This choice of  $S/N$  will be fully checked by applying the 2D surface brightness fitting procedure (see §3.2.1) to mock galaxies in §3.2.3. We refer to the samples resulting from the  $S/N$  selection, as the “high- $S/N$ ” samples, consisting of 4712, 128906, 379937, and 144580 galaxies, in the  $u$ ,  $g$ ,  $r$ , and  $i$  bands, respectively.

---

<sup>1</sup>`MAG_AUTO` is obtained from S-Extractor, computed using a Kron’s “first moment” algorithm (Kron 1980), as it is intended to give the most precise estimate of “total magnitudes”. Kron (1980) presented the following luminosity-weighted radius,  $r_1$ , which defines the ‘first moment’ of an image,

$$r_1 = \frac{\sum rI(r)}{\sum I(r)} \quad (3.1)$$

where  $r$  is the radius of the an elliptical aperture whose elongation  $\epsilon$  and position angle  $\theta$  are defined by second order moments of the object’s light distribution. Kron showed that an aperture of radius  $2r_1$ , when  $r_1$  is obtained by integrating to a radius  $r$  that is 1% of the sky flux, contains more than  $\sim 90\%$  of an object’s total light, making it a useful tool for estimating an object’s flux. S-Extractor derive `MAG_AUTO` with  $2.5r_1$ , and the mean fraction of flux lost drops from about 10% to 6%.

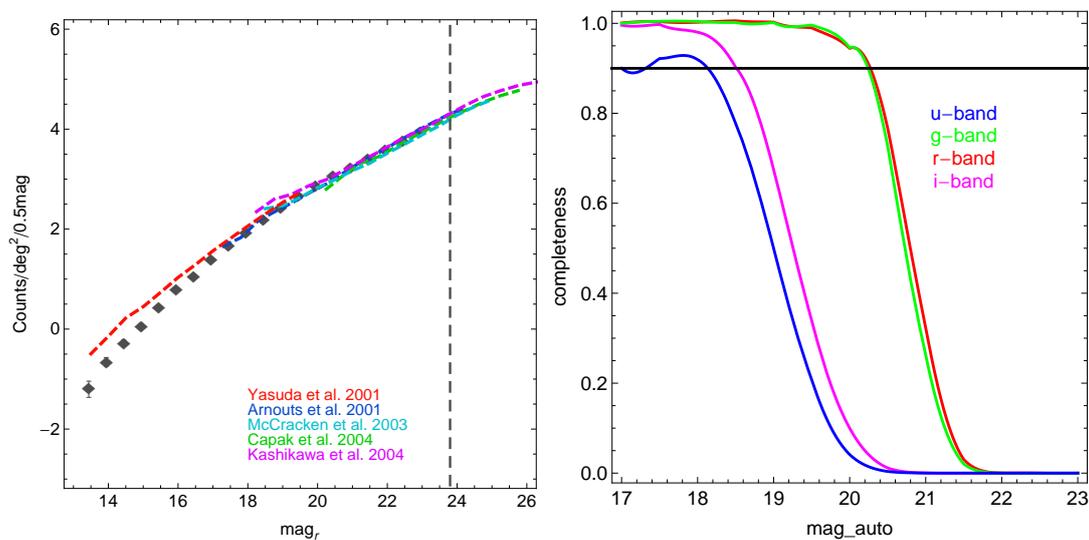


Figure 3.1: Left: Galaxy counts as a function of their magnitude in  $r$ -band are compared with other literature estimates (as in the label). The match with previous literature is very good at fainter magnitude while it is not perfect at the bright due to the limited area covered. See also the discussion in the text. Right: completeness of the “high  $S/N$ ” sample in  $u$ ,  $g$ ,  $r$ , and  $i$ -band with color code as in the legends. The completeness has been computed with respect to the 7M sample.

### 3.1.1 Magnitude completeness

The difference in counts among the different bands is due to the intrinsic depth of the different bands, being the latter a combination of exposure time and seeing, with the  $u$ -band the shallowest band and the  $r$ -band the deepest in the KiDS survey plan (de Jong et al. 2015).

In order to evaluate the completeness magnitude of our sample we have compared the galaxy counts as a function of `MAG_AUTO` of the “high  $S/N$ ” against the full 7M sample with no restriction on  $S/N$ . The ratio of their counts is shown in Fig. 3.1 from which we have derived the 90% completeness limit in each band. These were measured to be 18.2, 20.2, 20.2, and 18.5 for  $u$ ,  $g$ ,  $r$ , and  $i$ -band respectively.

### 3.1.2 Photometric Redshifts

Photometric redshifts have been derived from Multi Layer Perceptron with Quasi Newton Algorithm (MLPQNA) method (see D’Abrusco et al. 2007, Brescia et al. 2013, 2014), and fully presented in Cavuoti et al. (2015). This method makes use of an input knowledge base (KB) consisting of a sample for which both spectroscopic redshifts and multi-band integrated photometry are known and used to perform

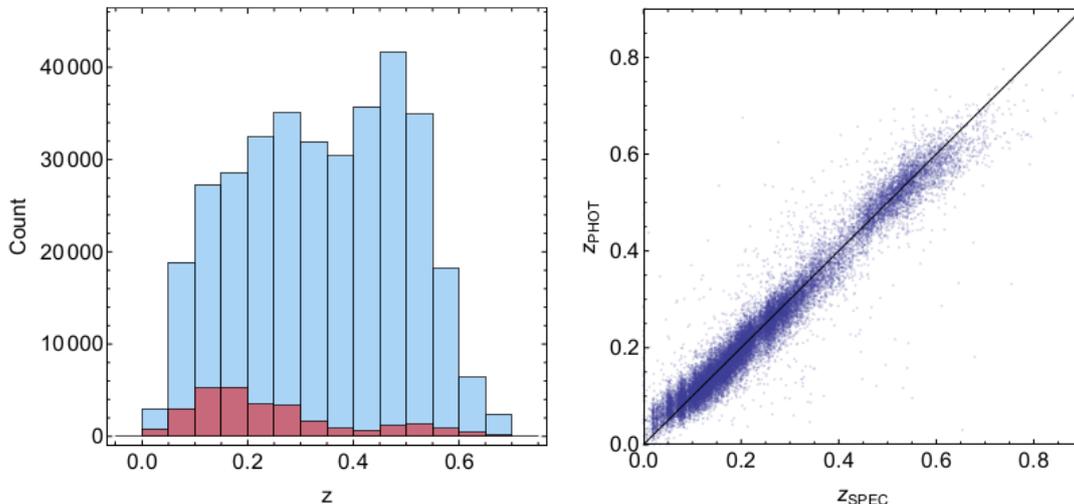


Figure 3.2: *Left*: the distribution of the spectroscopic sample of the sample adopted as knowledge base for the MLPQNA method (in red) and the photo- $z$  distribution of the “high S/N” sample (in light blue). *Right*: Comparison between spectroscopic redshift and the photometric redshift for the blind test set.

the mapping between colors and redshift. In particular we have used  $4''$  and  $6''$  diameter apertures for the magnitudes while spectroscopic redshift for the KB are given by the Sloan Digital Sky Survey data release 9 (SDSS-DR9; Ahn et al. 2012) and Galaxy And Mass Assembly data release 2 (GAMA-DR2; Driver et al. 2011), together including galaxies out to  $z \lesssim 0.8$  as shown in Fig 3.2 (left).

The final KB included  $\sim 60000$  objects, out of which 60% are used as train sample and the remaining as the blind test set. The spectroscopic redshift and the blind test sample redshifts are shown in Fig 3.2 (left), where they look good in agreement with each other. The scatter in the measurement is  $|(z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}})| \sim 0.03$  (see Cavuoti et al. 2015). These good accuracies are somehow better than the one obtained with other spectrophotometric methods on the same dataset (see e.g. Cavuoti et al. 2016 for a discussion). Indeed, the advantage of the machine learning techniques resides in the possibility of optimizing the mapping between the photometry and the spectroscopy, regardless the accuracy in the photometric calibration, but the disadvantage is the limited applicability to the volume in the parameters space covered by KB (see Cavuoti et al. 2015). In our case, for instance, out of 7 million starting systems, accurate photo- $z$  have been derived for systems down to  $r \sim 21$ , i.e.  $\sim 1.1$  million galaxies. This sample is still deeper than the typical completeness magnitude derived in §3.1.1.

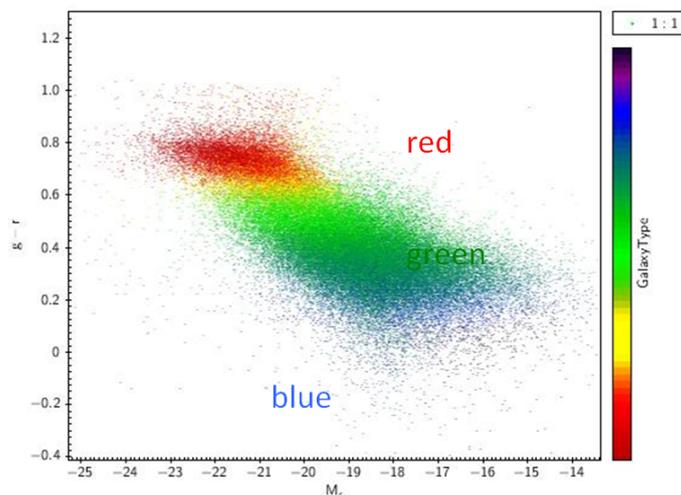


Figure 3.3: C-M diagram for the galaxy sample. Galaxies in red color are the passive spheroids whereas the blue cloud represents the disc dominated based on the SED galaxy classification

### 3.1.3 Spectral Energy Distribution fitting for galaxy classification

Stellar masses, rest-frame luminosities from stellar population synthesis (SPS) models and a galaxy spectral-type classification are obtained by means of the SED fitting with `Le Phare` software (Arnouts et al. 1999; Ilbert et al. 2006.), where the redshift is fixed to the  $z_{\text{phot}}$  obtained with MLPQNA.

In particular, we have used single burst SPS models from Bruzual & Charlot (2003), with a Chabrier (2001) IMF has been adopted. The observed *ugri* magnitudes (and related uncertainties) are the ones within a  $6''$  diameter aperture, which are corrected for Galactic extinction using the map in Schlafly & Finkbeiner (2011).

Total magnitudes derived from the Sérsic fitting,  $m_T$ , (see §3.2.1) are used to correct the total fluxes determined with `Le Phare` for missing flux. Furthermore, the k-correction is applied to derive galaxy absolute magnitude.

We have also used the spectrophotometric classes from `Le Phare` to derive a classification for our galaxy. As template set for this aim we adopt the 66 SEDs used for the CFHTLS in Ilbert et al. (2006). The set is based on the four basic templates (Ell, Sbc, Scd, Irr) in Coleman et al. (1980), and star burst models from Kinney et al. (1996). Synthetic models from Bruzual & Charlot (2003) are used to linearly extrapolate this set of templates into ultraviolet and near-infrared. The final set of 66 templates (22 for ellipticals, 17 for Sbc, 12 for Scd, 11 for Im, and 4 for starburst) is obtained by linearly interpolating the original templates, in order to improve the sampling of the redshift-color space and therefore the accuracy of

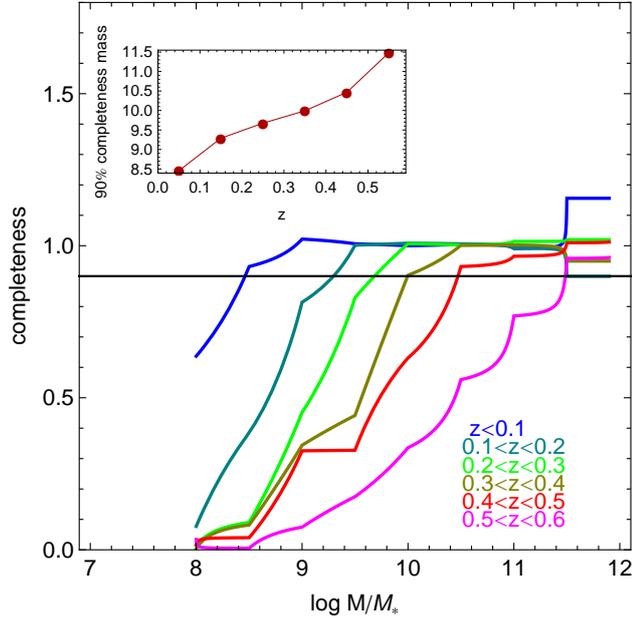


Figure 3.4: Mass completeness as a function of redshift: the ratio of the high  $S/N$  sample and the photo- $z$  sample for galaxies separated in different redshift bins are shown. See text for details. In the inset the derived completeness masses in the different redshift bins as in Table 3.1.

the redshift measurement. We do not account for internal extinction, to limit the number of free parameters. This fitting procedure provided us with a rough galaxy classification, which allows us to separate passive spheroids from disc-dominated galaxies. This is visible in Fig. 3.3, where galaxy  $g - r$  colors are plotted against the absolute magnitude obtained with *Le Phare*: passive spheroids populate the red part of the color-magnitude diagram (i.e. the so-called red sequence), while the later galaxy type (disc-dominated and starbursts) are located at the bluer colors and progressively populate the green valley and the blue cloud.

### 3.1.4 Mass completeness as a function of the redshift

In the following we will study the behavior of the galaxy properties as a function of the redshift. It is well known that some of the galaxy physical quantities (e.g. size, Sérsic index, color, etc.) correlate with mass. Hence it is important to define a mass complete sample in all redshift bins. We have then computed the mass completeness in the same way we have computed the magnitude completeness in §3.1.1, i.e. by comparing the high  $S/N$  counts this time against the photo- $z$  sample, once galaxies have been separated in different photo- $z$  bins. Results are shown in Fig. 3.4 and completeness masses are reported in Table 3.1. The table stops at  $z = 0.6$  because

Table 3.1: 90% completeness mass as a function of the photometric redshift for the high  $S/N$  sample in  $r$  band.

photo- $z$ bin	No. of galaxies in bin	90% compl.	$\log M/M_{\odot}$
$z \leq 0.1$	21536		8.5
$0.1 < z \leq 0.2$	55564		9.3
$0.2 < z \leq 0.3$	67369		9.7
$0.3 < z \leq 0.4$	62102		10.0
$0.4 < z \leq 0.5$	77143		10.5
$0.5 < z \leq 0.6$	53050		11.5

the high  $S/N$  sample starts to be fully incomplete in mass above that redshift. The mass completeness plots for  $g$  and  $i$  bands are given in 3.5. Note that the redshift bins, where 90% completeness is not achieved for both the bands are avoided.

## 3.2 Surface Photometry

In this section we present the measurement of structural parameters for the galaxy sample described above, using 2DPHOT (La Barbera et al. 2008). We evaluate parameter uncertainties and determine the reliability of the fitting procedure against mock galaxies on KiDS images. We finally compare the results obtained with KiDS for galaxies in common with literature analysed on SDSS data (i.e. La Barbera et al. 2010b).

### 3.2.1 Structural Parameters from 2DPHOT

Surface photometry of the high  $S/N$  sample has been performed using 2DPHOT (La Barbera et al. 2008), an automated software environment that allows 2D fitting of the light distribution of galaxies on astronomical images. The main steps performed by 2DPHOT are: 1) Catalog extraction and star-galaxy separation; 2) PSF modelling; 3) 2D fitting. We will details all these phases here below. Here we just stress that there are also other tools that we might have used, like Galfit (Peng et al. 2011) and GASPHOT (D’Onofrio et al. 2014) which has been extensively applied to galaxy in the WINGS survey (Fasano et al. 2006; Moretti et al. 2014; Gullieuszik et al. 2015). We have decided to use 2DPHOT as a first choice because it has been optimized for subarcsec galaxies and offered a good compromise between performances and computing time.

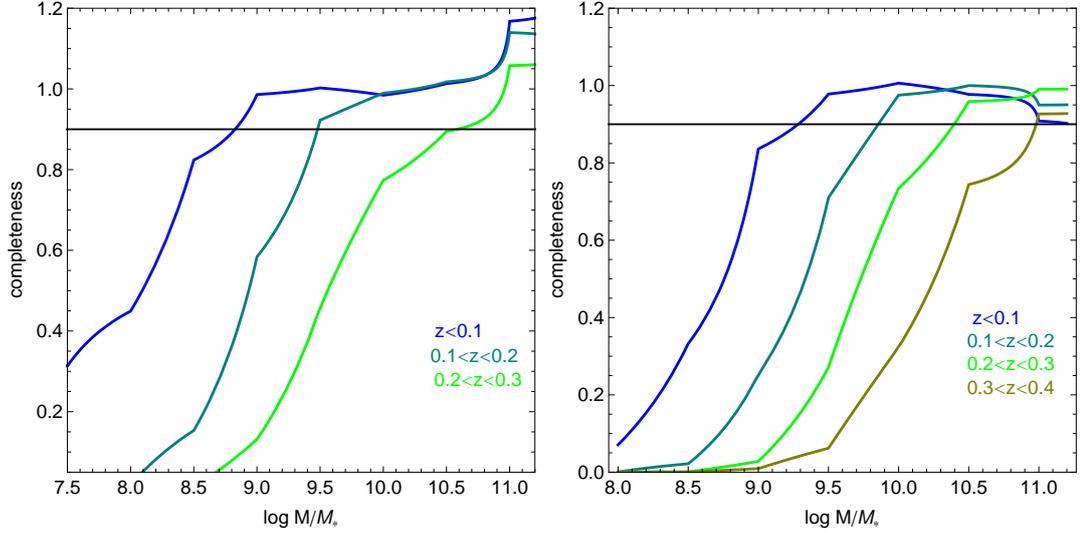


Figure 3.5: Left: Mass completeness as a function of redshift for g band. The ratio of the high  $S/N$  sample and the photo- $z$  sample for galaxies separated in different redshift bins are shown. The high  $S/N$  sample starts to be fully incomplete in mass above that redshift 0.3 in g band, hence avoided. The sample is complete above  $\log M/M_{\odot}$ , 8.82, 9.48, and 10.58 for  $z \leq 0.1$ ,  $0.1 < z \leq 0.2$  and  $0.2 < z \leq 0.3$  respectively. Right: same as g band. The high  $S/N$  sample starts to be fully incomplete in mass above that redshift 0.4 in g band, hence avoided. The sample is complete above  $\log M/M_{\odot}$ , 9.28, 9.85, 10.39, and 10.98 for  $z \leq 0.1$ ,  $0.1 < z \leq 0.2$ ,  $0.2 < z \leq 0.3$  and  $0.3 < z \leq 0.4$  respectively.

### Catalog extraction and star-galaxy separation

2DPHOT makes use of S-Extractor (Bertin & Arnouts 1996) to create the source catalog from the input image. Then it does a star-galaxy separation based on the `SEEING_FWHM` and the effective radius parameter  $r_e$ . In first run, S-Extractor detect the sources from the input image and calculate their Kron magnitude, FWHM, and `ELLIPTICITY` in order to select the preliminary star candidates. From the sure star locus, 2DPHOT estimates the `SEEING_FWHM` and feed this to S-Extractor to obtain a reliable stellarity index. Using the input `SEEING_FWHM` S-Extractor makes a second run. Sure stars are identified through this two step procedures.

2DPHOT extracts image sections with the detected sources in the center. The stamp is large enough to include a significant sky region around each source. This allows obtaining a reliable estimate of the local background from the two-dimensional fitting program. Size of the stamps depends on the `ISOAREA` parameter from S-Extractor. Image sections of mask images correspond to each detected sources is also created by flagging all the other sources in the stamp (eg. see Fig.3.6).

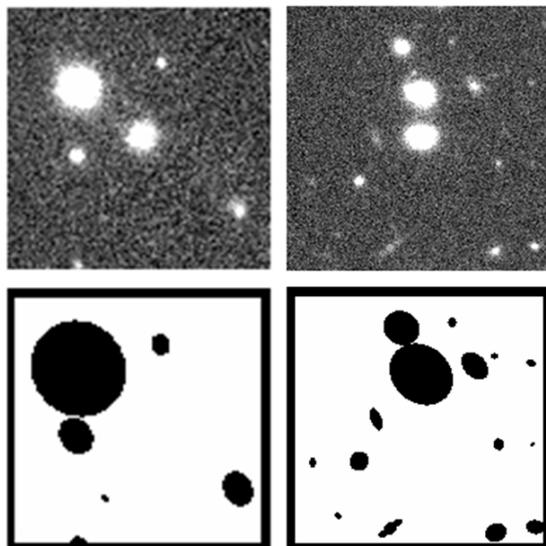


Figure 3.6: Examples of source and mask image stamps. Source image and mask image are shown in upper and lower panel respectively. Flagged pixels are shaded in black.

### PSF modeling

2DPHOT has been optimized to perform a Point Spread Function (PSF) convolved Sérsic modelling of galaxies. In large field detectors, the PSF is somehow a strong function of the position across the field-of-view: in Fig. 3.7 we show a typical PSF pattern in VST/OmegaCAM images where the solid lines show the amplitude of the elongation and orientation (anisotropy) of the PSF. Especially in the image borders, the orientation of PSFs is strongly aligned, while in the center the PSF tend to be more randomly orientated (isotropic), with smaller elongations. The PSF strongly affects the measurement of the surface brightness profile of galaxies by anisotropically redistributing the light from the inner brighter regions to the outer haloes (see e.g. de Jong 2008), hence altering the inferred galaxy structural parameters (e.g. effective radius, axis ratio, slope of the light profile, etc.). The PSF is modeled with two or more Moffat profiles (three in our case; see La Barbera et al. 2008 for details).

### 2D Fitting

A discrete, coarse two-dimensional fitting algorithm is performed in order to obtain a set of initial parameters for the sources. 2DPHOT creates 120 discrete models by varying effective radius  $r_e$  (10 values derived based on the pixel scale and the

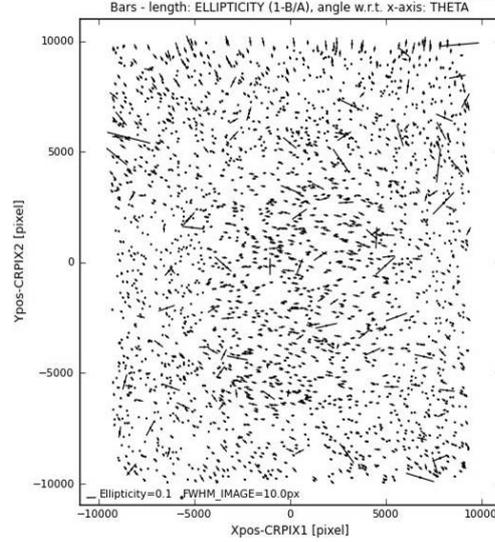


Figure 3.7: PSF anisotropy within the coadd *KIDS\_129.0-0.5* in r band. The elongation is aligned in a specific direction on the borders but random in the middle of the image.

SEEING\_FWHM of the image), total magnitude  $m_T = \{m_k, m_k - 0.2, m_k - 0.4\}^2$ , where  $m_k$  is the Kron magnitude (MAG\_AUTO from S-Extractor), and the Sérsic index  $n = \{1, 3, 5, 7\}$ . Each model is compared with the object image using a full  $\chi^2$  minimization algorithm, where all the pixels, that are not flagged within, the image stamp is considered. Parameters correspond to the lowest  $\chi^2$  model is taken as the input for the final 2D fitting.

The best-fit parameters are found by  $\chi^2$  minimization where the function to match with the 2D distribution of the surface brightness values is the convolved function given by

$$M(BG, \{p_k\}) = BG + B(\{p_k\}) \circ S \quad (3.2)$$

where B is the galaxy brightness distribution, which is described by a set of parameters  $\{p_k\}$ ; S is the PSF model; BG is the value of the local background; and the symbol  $\circ$  denotes convolution. For the Sérsic models, the parameters  $\{p_k\}$  are the effective major semiaxis  $r_m$ , the central surface brightness  $I_0$ , the Sérsic index  $n$ , the axial ratio  $b/a$ , the position angle PA, the coordinates of the photometric center, and the local value of the background. The modeled PSF is convolved with a 2D Sérsic profiles with the form

$$B(r, r_m, n) = I_0 + \frac{2.5b_n}{\ln(10)} [(r/r_m)^{1/n} - 1] \quad (3.3)$$

<sup>2</sup>Kron algorithm tend to over estimate the true magnitude, so we set a grid which is lower than the Kron magnitude

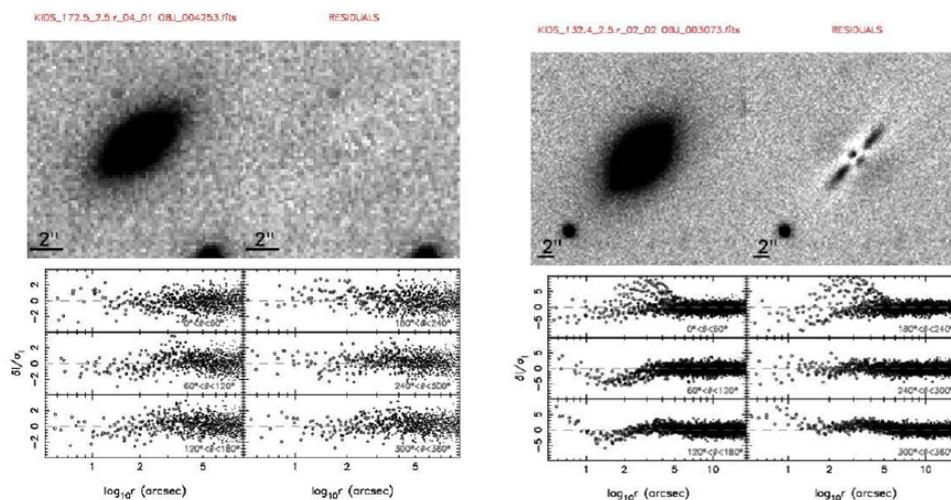


Figure 3.8: 2DPHOT output plot of a modeled galaxy. Top left panel is the source and the top right is the model subtracted residual map. Residuals after subtraction are shown in the lower panels as a function of the distance to the galaxy center, with each panel corresponding to a different bin of the polar angle. See also Appendix B and A for further examples.

In Fig. 3.8 two illustrative examples of two-dimensional fit results for galaxies in  $r$  band is given. Our single Sérsic fit provides, for each galaxy, the following parameters: the effective (half-light) radius,  $r_e$ , the mean surface brightness within that radius,  $\langle \mu \rangle_e$ , the Sérsic index (shape parameter),  $n$ , the axial ratio  $b/a$  and the position angle of the major axis, PA. More in details,  $r_e$  is computed as the equivalent radius of the ellipse that encloses half of the total galaxy light, i.e.,  $r_e = (b/a)^{1/2} r_m$ . The total (apparent) magnitude,  $m_T$ , is, by the definition,

$$m_T = -2.5 \log(2\pi) - 5 \log(r_e) + \langle \mu \rangle_e . \quad (3.4)$$

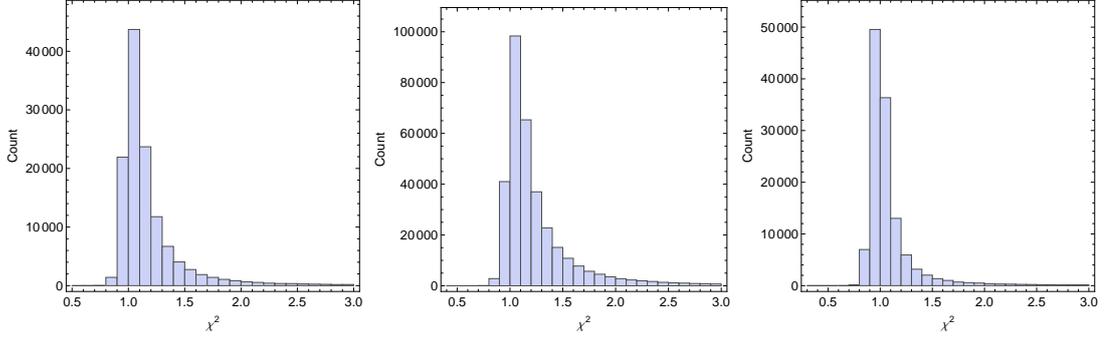


Figure 3.9:  $\chi'^2$  distribution of g, r, and i bands from left to right respectively.

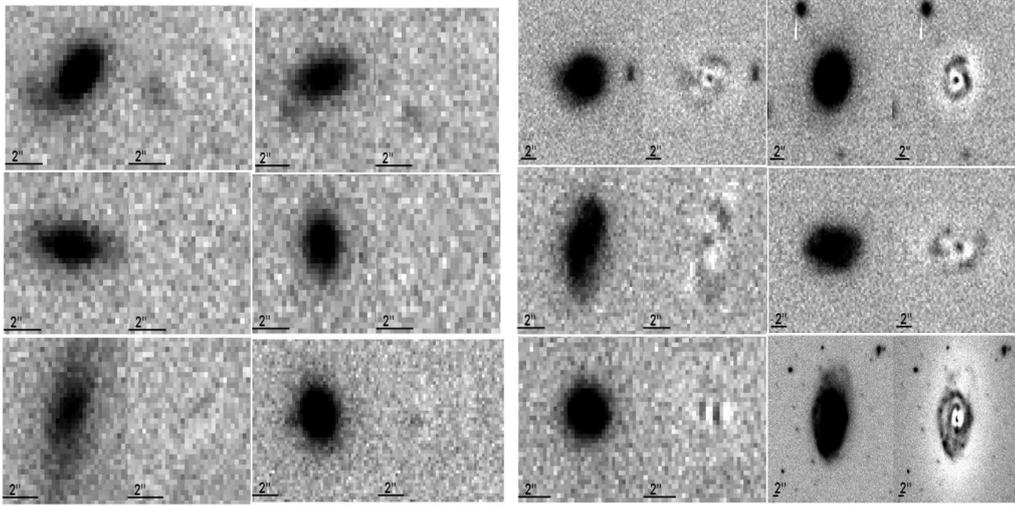


Figure 3.10: More examples to the 2D fit results for galaxies in r band. The left panel is the galaxies named the good fits and the right panel as bad fits. In each panel source and the model subtracted residual maps are shown for different sources. See also B and A for further examples.

$\chi'^2$  values obtained from the two-dimensional fitting of galaxies in each band by considering only the central region of the image stamp. For each galaxy, all pixels  $1\sigma$  above the local sky value background value are selected and the 2D model intensity value of each pixel is computed from the two dimensional seeing convolved Sérsic model as in Eq. 3.2. For the selected pixels, the  $\chi'^2$  is computed as the rms of residuals between the galaxy image and the model. The distribution of the  $\chi'^2$  for the whole high  $S/N$  sample in the  $g$ ,  $r$ , and  $i$  bands are shown in Fig. 3.9. This  $\chi'^2$  computation is different from the two-dimensional fitting procedure, where the sum of square residuals over all the galaxy stamp image is minimized. In Fig. 3.10 few illustrative examples of galaxy fitting is given: in one case we have a galaxy for which 2DPHOT has found a best-fit model with  $\chi'^2 < 1.3$  (middle panel) and in a second

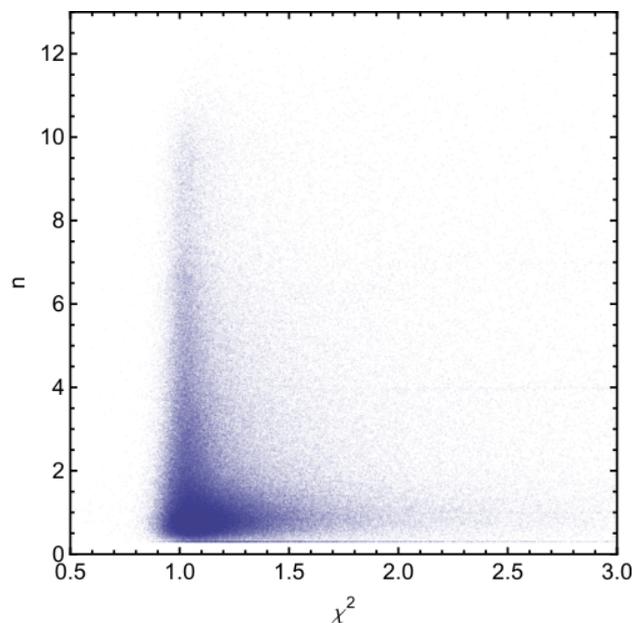


Figure 3.11: The plot shows the Sérsic index vs.  $\chi^2$ . We note that at lower  $n$  ( $< 2.5$ ) there is an excess of large  $\chi^2$  ( $> 1.3$ ), due to the presence of substructures in the residuals.

case the  $\chi'^2 > 1.3$  (right panel) due to a patterns in the residual by an unaccounted disk. We substantiate our argument using Fig. 3.11 where we plot the  $n$ -index vs.  $\chi'^2$  which shows that for lower Sérsic index ( $n < 2.5$ ) there is an excess of large  $\chi'^2$ , i.e. worse fit, due the fact that at these low- $n$  late-type systems are predominant (La Barbera et al. 2002, Ravindranath et al. 2002, Trujillo et al. 2007) and tend to have significant substructures. The use of a single Sérsic profile is indeed the more general choice we could make, as it is well know that galaxies generally host more than one photometric component (see e.g. Kormendy et al. 2009). This is not only true for late-type systems, showing a bulge+disk structure but also for some large ellipticals, now systematically found to have extended (exponential) haloes (e.g. Iodice et al. 2016). Looking at the  $\chi'^2$  distribution in Fig. 3.9, the fraction of galaxies with  $\chi'^2 > 1.3$  is not negligible, and amounts to  $\sim 30\%$  in  $r$ -band.

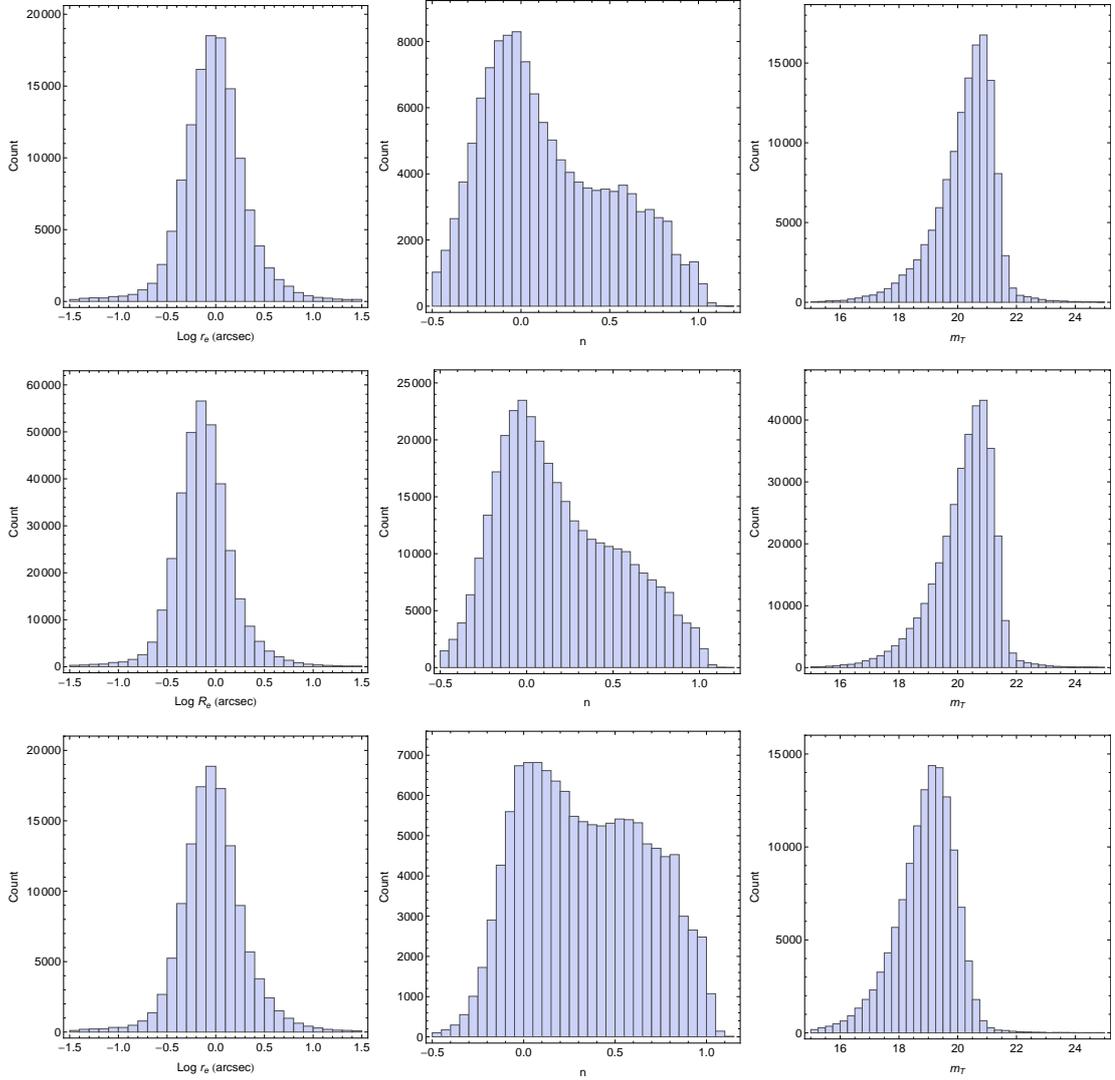


Figure 3.12: Distribution of structural parameters in g, r, and i bands from top to bottom respectively. Panels show distribution of  $\log r_e$ ,  $n$ ,  $m_T$ , from left to right respectively.

Maximum size we opted for each galaxy image stamps<sup>3</sup> is  $\sim 105$  arcsec by side, i.e. 500 pixels due to the resolution of telescope of  $0.21''/\text{pix}$ . We have also discarded galaxies  $r_e < 0.21''$  as it is significantly affected by the lower resolution limit of the telescope and  $r_e > 20''$  ( $\sim 1/5$  of the image size), because of the maximum stamp

<sup>3</sup>Size of the stamps is proportional to ISOAREA parameter from S-Extractor. By the term maximum size we say the stamp size can not go beyond this pixel limit. When smaller than the maximum size value, the stamp size is proportional to the ISOAREA parameter measured by S-Extractor, and it is always smaller than the maximum size value. Both the proportionality factor and the maximum stamp size can be varied.

size we opted. In some cases this choice of maximum stamp size is seemed to be a little restrictive for large nearby galaxies. We have checked that these were just a few cases which are not going to impact the overall sample completeness.

However, the adoption of multi-component models has two main disadvantages: the degeneracies among parameters and the higher computing time due to the higher dimensionality of the parameter space. In particular, the amount and the quality of the information (e.g. the number of pixels across which typically galaxies are distributed on CCDs) makes very hard to obtain reliable modeling of multi-component features in galaxies, especially when the ratio between the two component is unbalanced toward one (see e.g. the case in Fig. 3.8, right, where the inner disk represents a minor component of the dominant bulge). On the other hand, the single Sérsic fit can efficiently and more reliably catch most of the structure of the galaxy, leaving the signature of secondary components in the residuals (see e.g. the plots of the residuals along different azimuthal angles) and reflected in the bad  $\chi^2$  (typically  $> 1.3$ ).

The distribution of the best-fit structural parameters the effective (half-light) radius,  $r_e$ , the Sérsic index,  $n$ , the total magnitude,  $m_T$ , the mean effective surface brightness,  $\langle\mu_e\rangle$  in Fig. 3.12.

### 3.2.2 Uncertainties on structural parameters

Uncertainties on structural parameters are estimated by comparing the differences in  $\log r_e$ ,  $\langle\mu_e\rangle$ , and  $\log n$  between contiguous wavebands. The comparison is performed on r and i bands. The basic assumption is that the r and i bands are close enough that the variation of the galaxy properties from one band to other is dominated by the measurement errors (La Barbera et al. 2010b).

For the selection of the sample to use for the surface photometry analysis, we made use of the integrated  $S/N$ . However error on the estimate will strongly depend on the  $S/N$  per resolution element as the 2D fitting is performed by comparing the flux from the image and from the (convolved) model pixel by pixel. Hence the integrated  $S/N$  is diluted in the procedure depending on the FWHM and also on the intrinsic area covered by the galaxy: the larger is the galaxy size (and the area covered) the more is the number of pixels over which the  $S/N$  is redistributed. For this reason we have evaluated the differences in the  $\log r_e$ ,  $\langle\mu_e\rangle$ , and  $\log n$  between  $r$  and  $i$  bands with respect to the  $S/N$  per unit area of the galaxy image,  $S/N/r_e^2$ . We have also binned the same logarithmic differences in terms of the logarithm of the mean effective radius  $\log r_e$  to check the systematic variation of the uncertainties as a function of the galaxy size. The  $S/N$  is defined as the mean value, of the inverse of the uncertainties on Kron magnitudes (`MAGERR_AUTO` from S-Extractor), between two bands. Bins are made such that the number of galaxies in each bin is same. Measurement errors on  $\log r_e$ ,  $\langle\mu_e\rangle$ , and  $\log n$  are computed from the mean absolute

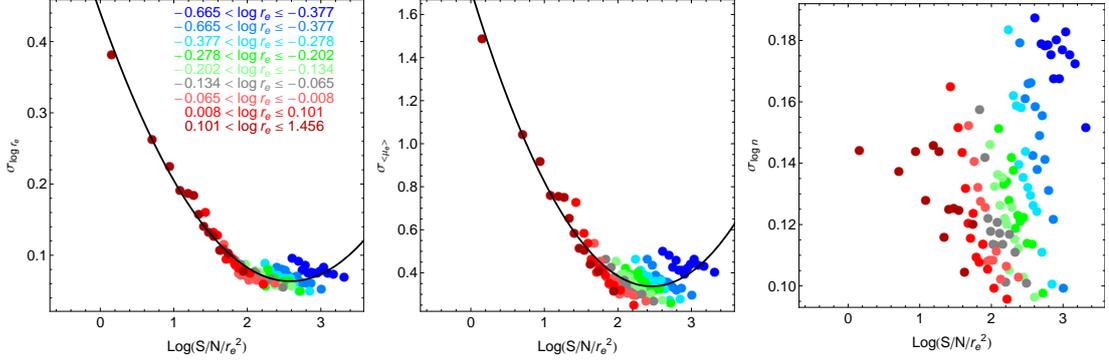


Figure 3.13: Uncertainties in the parameters  $\log r_e$ ,  $\langle \mu_e \rangle$ , and  $\log n$  as a function of the logarithm of the  $S/N$  per unit area. Different colors show different bins of  $\log r_e$  where  $r_e$  is in arcsec. For a given color the points are the uncertainties in different bins of logarithm of  $S/N/r_e^2$ . The black curve is the best fitting functional form used to model the dependence of the uncertainties on  $S/N$ . This fit is not performed for  $\log n$  as it does not show any correlation with  $S/N/r_e^2$ .

deviation of the corresponding differences in that bin. The results are shown in Fig 3.13. The errors on the parameters depend on the  $S/N$  per unit area: as the value of  $S/N$  per area decreases ( $\text{Log}(S/N/r_e^2) < 2$ ), the errors tend to increase. This is due to the combined effect of the  $S/N$  and the number of pixels where the signal is distributed. Low  $S/N/r_e^2$  are sources with large  $r_e$  and small  $S/N$  where as high  $S/N/r_e^2$  are systems that might have large  $S/N$ , but due to the small number of pixels induces the uncertainty on parameters. Most of the galaxies are in the range  $-0.5 < \log r_e < 0.2$ , where the errors on the parameters are less than 0.1dex for  $r_e$  and less than 0.4dex for  $\langle \mu_e \rangle$ , but the errors on  $n$  are more randomly distributed and do not show particular trends. However, also in this case, they stay remarkably contained below 0.2dex.

### 3.2.3 Check for systematics on the estimated parameters

In this section we proceed with a series of validation tests to check the presence of biases in the parameter estimates and to have insight on the actual statistical errors. To do that we will make use of mock galaxies generated on KiDS images and literature sample having an overlap with our KiDS galaxy sample. However we will start with a basic check on the effect of the background evaluation on the parameter estimates.

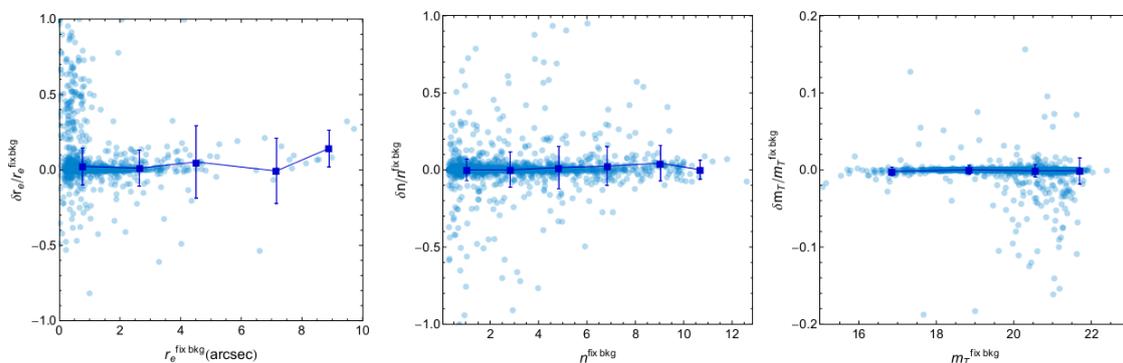


Figure 3.14: Differences in the parameters  $r_e$ ,  $n$ , and magnitude when background is kept constant with respect to the parameter when background is subjected to change for r band. Mean values are given along with the points. Error bar shows the standard deviation in bins of parameter on x axis. In each plot x axis shows the parameter when background is a constant.

### Effect of sky background

We have discussed in §3.2.1 that the background is a free parameter in our fitting procedure (see e.g. Eq. 3.2). However it is well known that the contemporary fit of the background and the photometric laws can be degenerate and might produce some systematics.

We decided to check the differences of the output parameters by fitting background fixed to background set as a free parameter. We measured the background value far from the galaxy (local background value calculated from the galaxy stamp images, which is 1.5 times the S-Extractor ISOAREA parameter, see La Barbera et al. 2008 for more details) and entered as the initial guess in the fitting procedure. Here we fix this value of background for the modeling of galaxies.

We randomly selected  $\sim 3000$  galaxies from our high  $S/N$  galaxy sample and again extracted the structural parameters by fixing the sky background value. We compare the two set of structural parameters we obtained with the standard procedure and the one with fixed background. We show the differences in structural parameters in 3.14.

Most of the galaxies in our selection is having difference in the measurement of parameters nearly zero. The error bars represent the mean and standard deviation of the scattered plot which show that the background fit does not introduce systematics and that the error associate to the background measurement are of the order of 10-25% in  $r_e$ , less than 13% in  $n$ , and less than 2% in the total magnitude, which are in line with the estimate in §3.2.2.

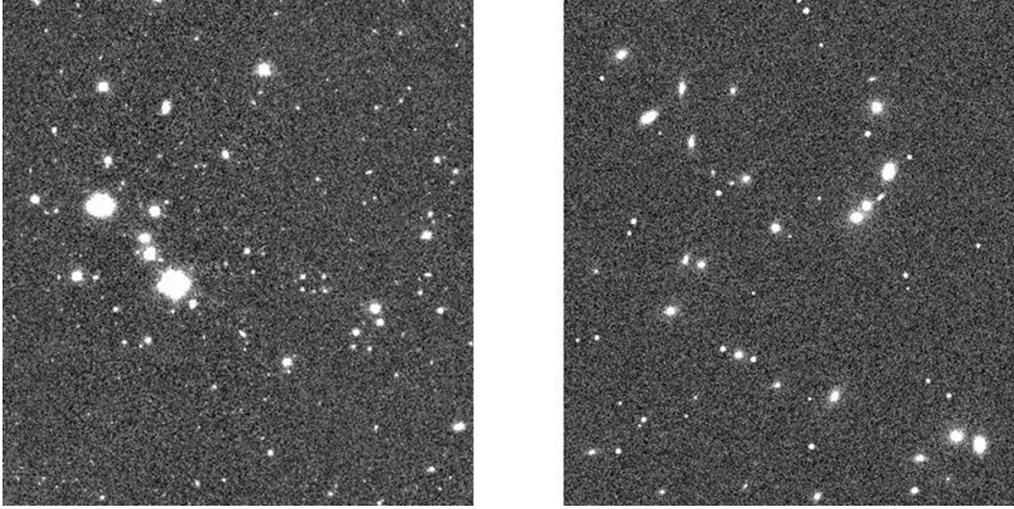


Figure 3.15: Example images shows the real KiDS (left) and mock (right) image.

### Galaxy simulations

To assess the reliability of the parameters from the fitting procedure and estimate their intrinsic statistical errors, we have generated mock observations of galaxies in simulated KiDS images (Note that we already calculated the uncertainty on parameters for real KiDS images using *r* and *i* band data in §3.2.2, which is more realistic than using mock galaxies). We first recreated the KiDS image background by reproducing the local rms of the image background obtained by SExtractor. Then we added artificial galaxies for which pre-defined values for the physical properties magnitude, Sérsic index, effective radius, and axis ratio were taken in order to reproduce the real galaxy parameter space. In particular we have uniformly sample the parameters in the following intervals:  $0.2 \leq r_e \leq 20$  arcseconds,  $0.6 \leq n \leq 10$ ,  $0.5 \leq b/a \leq 1$ ,  $16 \leq m_T \leq 24$  mag and  $0 \leq P.A. \leq \pi/2$ . We have simulated about 1800 galaxies on image chunks of 3000 pixels by side in order to reproduce the same galaxy density observed in KiDS images. We have generated such mock observations in different bands and in different seeing conditions. In Fig. 3.15 we show an example of simulated *r*-band image compared with a real one. We have then applied blindly the fitting procedure by 2DPHOT as the mock data were real and derived the structural parameters of the simulated images performing the standard 2D seeing convolved Sérsic fit, after having modeled the PSF in the same way it is performed on real images (see §3.2.1). The relative differences between the measured quantities and the input ones adopted to generate the simulated galaxies are shown in Fig. 3.16 as a function of the  $S/N$ .

The figure shows that the input and output values are well in agreement with one another except for the lower  $S/N < 40$  regime where we start observing a systematic

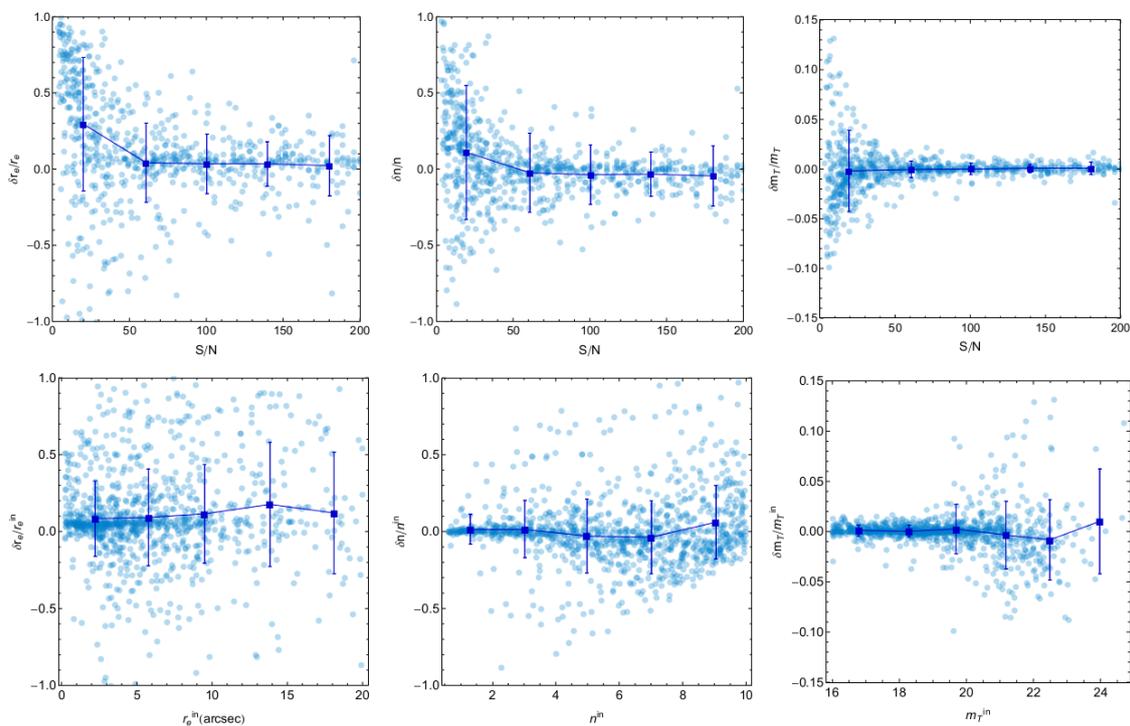


Figure 3.16: Figure shows differences in the input and output parameters for  $r_e$ ,  $n$ , and magnitude, with respect to  $S/N$  in upper panel and to their input values in bottom panel.

deviation of the measured values from the real ones. This is *a posteriori* confirmation that our choice of  $S/N \gtrsim 50$  for robust structural parameter studies was correct. In the same Figure we show the relative differences of the same observables against the input values (bottom row): in this case there is no trend in the derived quantities. We have found that these remarkable accuracies are independent of the band and of the seeing, as long as we restrict to galaxies with  $S/N > 50$  in any given bands.

### Comparison against a SPIDER-SDSS common sample

Another straightforward check to test the accuracy of our structural parameter estimates was to compare our catalog with the SPIDER catalog of galaxies (La Barbera et al. 2010b) having common sky area with KiDS DR1 and 2. The SPIDER sample includes 39,993 spheroids with SDSS optical imaging and UKIDSS NIR imaging, with redshifts in the range  $0.05 \lesssim z \lesssim 0.1$  (i.e. a more local sample). This sample has been analyzed with 2DPHOT in La Barbera et al. (2010b) which we refer the reader for details. By matching the KiDS data with SPIDER we found 344 galaxies in common for which we can have a direct comparison of the derived parameters. This is a very useful test to check the relative differences among the structural pa-

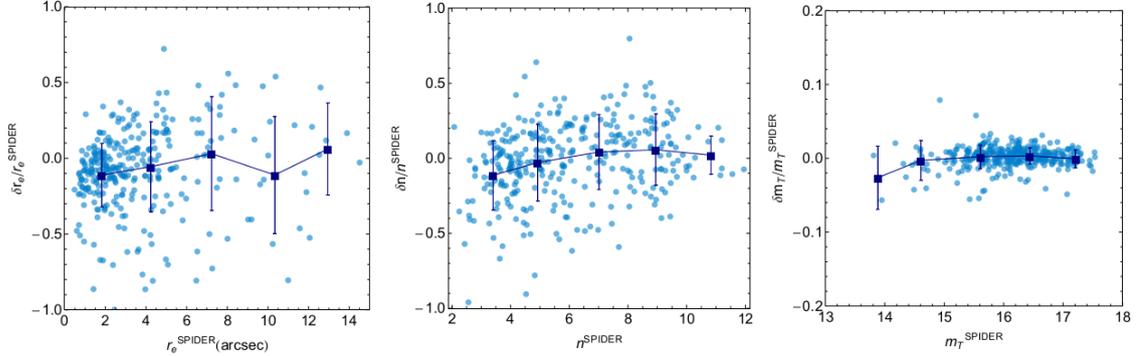


Figure 3.17: Comparison of KiDS structural parameters with the ones derived within the SPIDER survey. The SPIDER dataset consists of passive spheroids with redshifts in the range  $0.05 < z < 0.095$ , selected from SDSS; the structural parameters are derived using 2DPHOT. We show the relationship for effective radius ( $r_e$ ), Sérsic index ( $n$ ), and model magnitude ( $mag_r$ ) respectively from left to right. Data are shown as points. Mean values and standard deviation as error bar for differences the parameters with in x axis bins are given.

rameters with two different datasets in terms of depth (KiDS is two magnitudes deeper than SDSS) and image quality (both pixel scale and seeing are about twice smaller in KiDS).

The results are shown in Fig. 3.17, where we can see a good agreement among the parameters from two datasets with the scatter in line with the statistical errors discussed in §3.2.3.

### 3.3 Passive spheroids and disc-dominated galaxies classification

We are interested to characterize the internal structure of galaxies as a function of the redshift and derive insight on their formation mechanisms. Early- and Late-type galaxies are expected to follow different evolution channels and show different scaling relations (e.g. Faber-Jackson, Tully-Fisher). In §3.1.3 we have seen that we have a criterion to separate passive spheroids from disc dominated based on their color information, however the advantage of having a structural analysis performed and in particular the fact that we have seen that the Sérsic index seems to be sensitive to the galaxy substructures (§3.2.1), suggests that we can have an improved morphological selection using the combined information on SED properties of galaxies (based on SED\_CLASS parameter which is obtained from the SED fitting procedure explained in §2.3.8) and their light profiles. We have separated passive spheroids and disc dominated using the following criteria: 1) passive spheroids are all galaxies with

$n > 2.5$  and SED classification (Tortora et al. 2016),  $\text{SED\_CLASS} \leq 22$ ; 2) disc dominated are all galaxies with  $n < 2.5$  and SED classification,  $22 < \text{SED\_CLASS} \leq 56$  (See §3.1.3);

As mentioned in §3.2.1 our final sample of galaxies excludes galaxies with constraints given by the pixel size and the image stamp size and for which a worse model fit was obtained. To summarize, we have retained only those galaxies with  $r_e > 0.21$  (i.e. larger than the telescope resolution, 0.21 arcsec/pixel),  $\log r_e < 1.7$  (i.e. 1/5 of the stamp size of 105 arcsec) and  $\chi'^2 < 1.3$ . The final sample consists of 49,972 passive spheroids and 144859 disc dominated galaxies in r band. We just remark that there are a number of galaxies (13403) which turned out to have a  $\text{SED\_CLASS} > 51$ , which we have also excluded from our analysis.



# Chapter 4

## Evolution of galaxy size–mass relation from the Kilo Degree Survey

In this Chapter we present results about the evolution across cosmic time of galaxy sizes and size–mass relations. The correlation between these two quantities is strictly related to the way the galaxies have been assembled. It is known that the two main classes of galaxies, early and late types, show a different dependency between size and stellar mass with disc-dominated galaxies having a weak, if any, dependence on the redshift, and passive spheroids showing a clear variation with the redshift (see e.g. Shen et al. 2003, van der Wel et al. 2014, Poggianti et al. 2013b), which suggest a different evolution pattern for the two populations. However, the evolution with redshift depends strongly on the galaxy mass (see e.g. Furlong et al. 2015) and with our sample we expect to expand previous analyses toward lower mass regime and, in general, fine sampling the parameter space (luminosity, mass, redshift etc.). For the results presented in this chapter we adopted the following cosmology:  $H_0 = 75$  km/s/Mpc,  $\Omega_m = 0.29$  and  $\Omega_\Lambda = 0.71$ .

### 4.1 Size–Mass as a function of redshift

We will start by investigating the size–mass relation as a function of the redshift and compare this with previous literature data and simulations. We split the sample in passive spheroids and disc-dominated galaxies as described in §3.3 and discuss the size–mass relations separately.

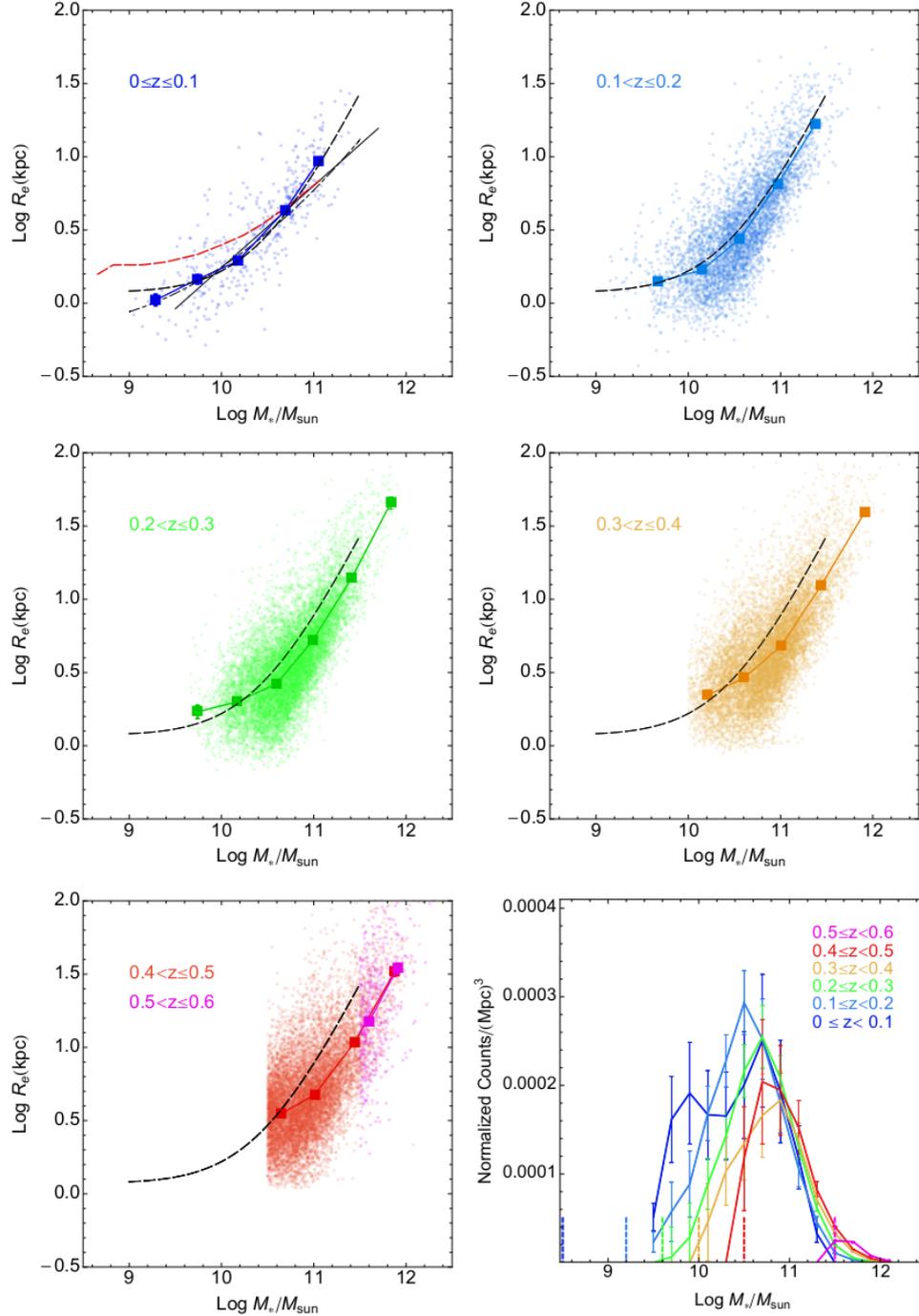


Figure 4.1: Size–Mass relation for passive spheroids (top panels and left and central bottom panels). Individual galaxy values are plotted together with mean and standard deviation of the mean (boxes and errorbars). For the  $0 \leq z < 0.1$  bin we overplot some local relations from literature (solid line: Shen et al. 2003; dot dashed line: Hyde & Bernardi 2009; dashed black line: Mosleh et al. 2013; dashed black line: Baldry et al. 2012). For all other  $z$  bins we show the  $z = 0$  relation form Mosleh et al. 2013 to visually appreciate the deviation of the average relation from the local one. Bottom right panel: the stellar mass distributions in different  $z$  bins normalized to the total covolume and to the total galaxy counts in the  $0.2 \leq z < 0.3$  bin. The completeness mass limit are overplotted as vertical dashed lines. Error bars show the propagation of the photo- $z$  errors on the covolume.

### 4.1.1 Passive Spheroids

In Fig. 4.1 we start by showing the size–mass of passive spheroids in different redshift bins with overplotted the mean as boxes and the standard deviation of the mean as errorbars. In the Figure only the 90% complete sample is shown, and this becomes clear in particular at  $z > 0.3$  where the sample starts to be severely incomplete at  $\log M_*/M_\odot < 10$ . The two bins at  $z > 0.4$  are shown together as the contribution of galaxies in the bin  $0.5 < z \leq 0.6$  is minimal and limited to the very high mass end. The mean of the latter redshift bin are fully consistent with the ones derived for the lower  $z$  bin,  $0.5 < z \leq 0.6$ , hence we decided to cumulate the two samples.

In the figure we have also plotted some relevant literature trends obtained at  $z = 0$  (i.e. Shen et al. 2003, S+03 hereafter, Hyde & Bernardi 2009, HB+09 hereafter, Mosleh et al. 2013, M+13 hereafter), after having scaled all masses to the Chabrier IMF, which is our reference choice.

There seems to be a striking consistency of our mean values with the size-mass from M+13, which have been shown by a double power-law best fit (their Eq. 4) to their data as dashed line in our Figure<sup>1</sup>. In particular, we clearly see in our data a flattening of the relation at masses below  $\text{Log} M_*/M_\odot \sim 10.0$  in the lowest  $z$  bin. The  $z = 0$  relation from M+13 nicely matches also the average trend in our next  $z$  bin ( $0.1 < z \leq 0.2$ ), where the flattening of the relation is even more evident.

Differently from M+13, S+03 results are represented as their best fit single power-law,  $R_e \propto M^\alpha$ , while HB+09 have performed a parabolic in the Log-Log plane to reproduce the curvature they have observed in their data too and that it is reported in Fig. 4.1. Both S+03 and HB+09 have a good agreement with our data at the intermediate mass scales, while they diverge at the lower masses, where they do not seem to catch the flattening of the average size-mass relation, and also at the higher masses ( $\text{Log} M_*/M_\odot > 10.8$ ), where they suffer systematics due to the adoption of circularized radii (see Dutton et al. 2011). To conclude our comparison with previous literature, we also show the average relation obtained by Baldry et al. (2012) with GAMA galaxies, where we see also a flattening of the relation at  $\text{Log} M_*/M_\odot \sim 10.0$ , but the overall relation seems tilted with respect to our average relation.

We can use for the moment the M+13 results as a  $z = 0$  reference to compare the size-mass relations in the other redshift bins and visually evaluate the evolution of the mass–size with lookback time. Going toward higher  $z$ , in Fig. 4.1 we show that the mean correlation (boxes connected by the solid lines) starts to deviate from the  $z = 0$  relation after  $z = 0.2$ . The divergence is significant within the errors at stellar masses  $\text{Log} M_*/M_\odot \gtrsim 10.5$ , while at lower masses there is an hint of an inversion of the trend with respect to the  $z = 0$  relation, which might be partially due to the incompleteness of the lower mass bins. To check the significance of the difference of

<sup>1</sup>Note that, the M+13 effective radii are obtained from a single Sérsic profile, likewise us.

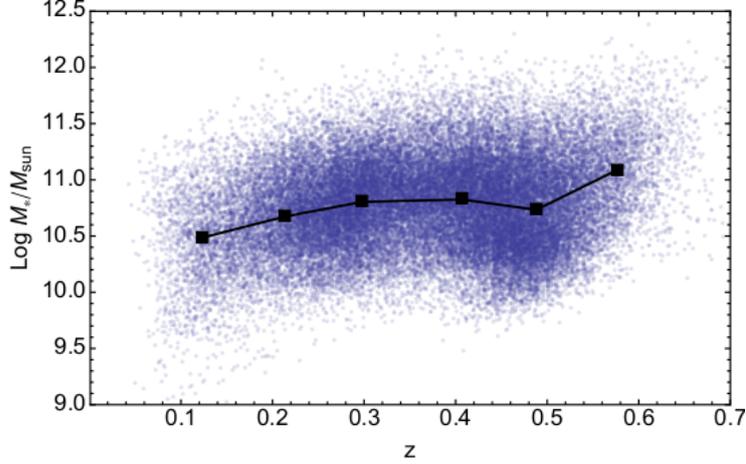


Figure 4.2: Mass vs. redshift plot for the passive spheroid sample. Overplotted to the individual galaxy values, we show the mean and standard deviation of the mean of the sample (error bars are comparable to the size of the boxes). Note that the steepening of the  $z \sim 0.6$  bin is due to the mass incompleteness of this bin. In the lowest  $z$  bin ( $z \sim 0.1$ ), the sample suffers some volume incompleteness (see discussion in §3.1), which produces the mean mass in the bin to be biased toward the less massive systems.

the data with respect to the  $z = 0$  relation, we have performed a simple  $\chi^2$  test and computed the probability that the observed data points are consistent within the errors with the M+13 fitting formula. The  $\chi^2$  is defined as

$$\tilde{\chi}_{obs}^2 = \sum_i^{N_{z_j}} \frac{(R_{e,i} - M13[M_{*,i}])^2}{N_{z_j} \delta R_{e,i}^2}, \quad (4.1)$$

where  $R_{e,i}$  and  $M_{*,i}$  are the individual values in the  $j$ -th  $z$  bin,  $z_j$ ,  $M13[x]$  is the M+13 relation and  $\delta R_{e,i}$  are the statistical errors on the effective as derived in §3.2.3, which have been doubled to take into account also the uncertainties on the M+13 relation.

We have found that for the first two redshift bins, the reduced  $\tilde{\chi}_{obs}^2$  are 0.42 and 0.89 respectively, which are smaller than the ones giving a 95% significance (1.17 for 396 and 1.05 for 2246 degrees of freedom respectively) of the M+13 relation to correctly describing the data. On the other hand, for  $z > 0.3$  the  $\tilde{\chi}_{obs}^2$  are always larger than the one giving the 5% significance, thus we can estimate that the data at  $z > 0.3$  differ from the  $z = 0$  relation at more than 95% significance. On the higher mass side, the sample does not suffer any particular incompleteness, as shown by the mass distribution in the  $z$ -bins in the bottom-right panel of the same Fig. 4.1 (except possibly for the low- $z$  bin, see also below). Here, the counts have been normalized to the total area of the mass function at  $0.2 \leq z < 0.3$  and error bars mainly reflects

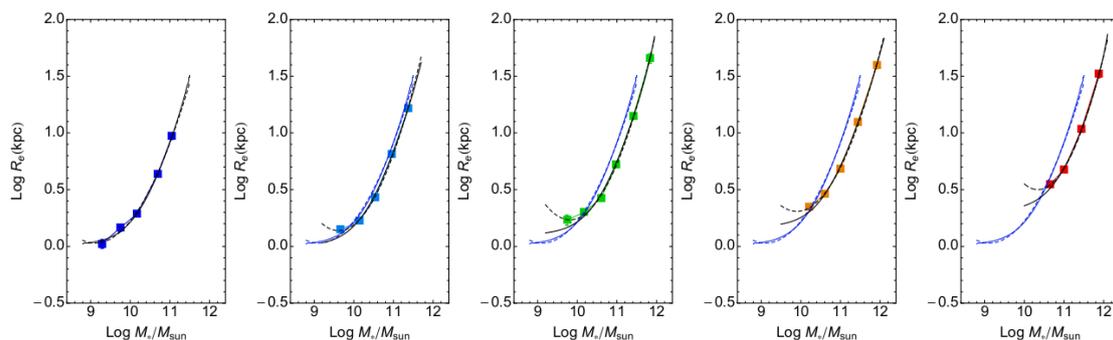


Figure 4.3: Parametric fit to the Size–Mass relation for passive spheroids. The average size mass in different bins (color code as in Fig. 4.1) is fitted with the parametric formulae as in Eqs. 4.2 (solid lines) and 4.3 (dotted lines). The  $z = 0$  fit has been reported in the subsequent  $z$  bins in blue, to visually check the divergence of the  $z > 0.1$  relations. These curves are used to define the  $R_e$  corresponding to different mass intercepts as shown in Fig. 4.4.

the photo- $z$  errors on the determination of the comoving volume in the different  $z$ -bins. The drop of the counts after the first peak at  $\text{Log}M_*/M_\odot = 10.5 - 11.0$  going towards lower masses, is typical of the ETG mass function measured at all redshifts (see e.g. Kelvin et al. 2014) and does not reflect an intrinsic incompleteness of the sample. We conclude that the observed trend with  $z$ , which moves the passive spheroids sample progressively away from the  $z = 0$ , is genuine and has to be related to an evolution pattern in the galaxy structural parameters.

This should not be due to an evolution of the stellar mass, as the average stellar mass of our sample does not show any significant trend with the redshift. This is demonstrated in Fig. 4.2, where the average masses stay almost constant in the range  $\text{Log}M_*/M_\odot \sim 10.7 - 10.8$  as a function of  $z$  although a steepening is observed only at the  $z \sim 0.6$  bin, which is due to the mass incompleteness of this bin (below  $\text{Log}M_*/M_\odot = 11.5$ ). A possible selection effect is also present in the lowest  $z$  bin ( $z \sim 0.1$ ), due to the volume incompleteness discussed in §3.1), which causes the average mass in the bin to be biased toward the less massive systems.

We conclude that the responsible of the evolution of the size-mass relation is mainly the effective radius. Visually, this means that galaxies more massive than  $\text{Log}M_*/M_\odot \sim 10.5$  have sizes (i.e.  $R_e$ ) that decrease with increasing redshift at any given mass. To better quantify this effect and to estimate also the amount of the size variation in the different mass intervals, we have performed a fit to the average size–mass at different redshifts and then evaluated the  $R_e$  corresponding to different mass intercepts (see also van der Wel et al. 2014, hereafter vdW+14).

To fit the size–mass we have used the two fitting formula used in M+13 and HB09 (as showed in Fig. 4.1), which we report here below for clarity:

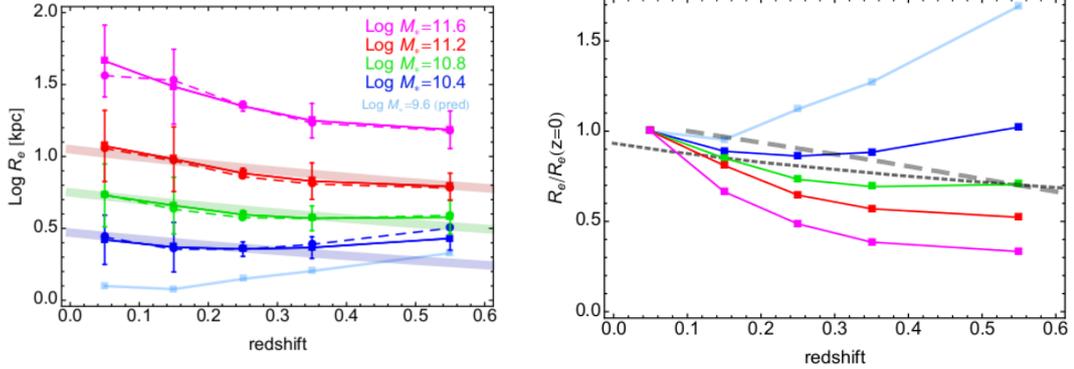


Figure 4.4: Size vs.  $z$  plots from the average size-mass parametric fit of passive spheroids. *Left*: the derived absolute intercept of the best fit relations as in Fig. 4.3 at mass values as in the legends. Error bars account for the  $1\sigma$  errors in the best fit. We also show the extrapolation for a small mass value ( $\text{Log}M_*/M_\odot = 9.6$ ), which is in the incompleteness range for redshift  $z > 0.2$ , as a “prediction” for higher- $z$  if our best fit relations is correct. We also overplot results from vdW+14, corresponding to  $\text{Log}M_*/M_\odot = 10.25, 10.75, 11.25$ , from bottom to the top, which well compare to our measurements in similar mass bins. *Right*: size evolution with respect to the local size at different mass intercepts. The evolution of the size with redshift becomes increasingly significant at larger masses while at  $\text{Log}M_*/M_\odot \sim 10.4$  the correlation inverts the trend as shown from our predictions for  $\text{Log}M_*/M_\odot = 9.6$ . The grey dashed line is the relation as found by Trujillo et al. (2007) for  $\text{Log}M_*/M_\odot > 11.0$ , the dotted line is the relation from Huertas-Company et al. (2013) for the  $\text{Log}M_*/M_\odot \sim 11.0$ .

$$R_e = \gamma(M_*)^\alpha (1 + M_*/M_0)^{(\beta-\alpha)} \text{ [from M + 13]}, \quad (4.2)$$

where  $R_e$  is in kpc,  $M_*$  in solar units, and  $\alpha, \beta, M_0$  are free parameters, and

$$Y = p_0 + p_1 X + p_2 X^2 \text{ [from HB09]} \quad (4.3)$$

where  $Y = \text{Log}R_e/\text{kpc}$ ,  $X = \text{Log}M_*/M_\odot$  and  $p_0, p_1, p_3$  are free parameters to be adjusted to best fit the data points. The best fit relations for both cases are shown in Fig. 4.3. The fit is generally very good for both fitting function across the data points, however Eq. 4.3 seems to predict a very strong up-turn of the trend, right outside the first data point, which we cannot confirm with our current dataset.

In Fig. 4.4 we show the  $R_e$  vs.  $z$  relation obtained as intercept of the best fit relations as in Fig. 4.3 at different mass values. The errors on the individual estimate take into account the  $1\sigma$  errors in the best fit. There is an evident trend of the sizes to decrease with redshift in all mass values except  $\text{Log}M_*/M_\odot = 10.4$ .

Table 4.1:  $R_e = B_z(1+z)^{\beta_z}$  fit to the size-redshift relation as derived from the average size-mass fit at different redshifts (Fig. 4.4 and 4.6) and from direct fit to the size-redshift relation in different mass bins (Fig. 4.9).

LogM <sub>*</sub>	Indirect from size-mass				Direct Fit			
	passive spheroids		disc-dominated		passive spheroids		disc-dominated	
	LogB <sub>z</sub>	β <sub>z</sub>	LogB <sub>z</sub>	β <sub>z</sub>	LogB <sub>z</sub>	β <sub>z</sub>	LogB <sub>z</sub>	β <sub>z</sub>
10.0	–	–	0.59 ± 0.09	–0.12 ± 0.07	–	–	0.01 ± 0.17	0.35 ± 0.13
10.4	0.33 ± 0.13	0.04 ± 0.10	0.84 ± 0.03	–0.22 ± 0.02	–0.07 ± 0.08	0.35 ± 0.06	0.53 ± 0.14	0.02 ± 0.11
10.8	1.00 ± 0.14	–0.29 ± 0.11	1.04 ± 0.06	–0.21 ± 0.04	0.74 ± 0.20	–0.11 ± 0.15	1.01 ± 0.21	–0.18 ± 0.15
11.2	1.63 ± 0.14	–0.57 ± 0.11	1.04 ± 0.13	–0.03 ± 0.10	1.67 ± 0.11	–0.62 ± 0.08	0.68 ± 0.04	0.17 ± 0.03
11.6	2.58 ± 0.24	–0.94 ± 0.19	–	–	1.88 ± 0.04	–0.51 ± 0.03	–	–

This trend seems nicely consistent with a similar analysis performed by vdW+14 on HST data for CANDELS (Koekemoer et al. 2011) and shown in the same figure, where we show their results for  $\text{Log}M_*/M_\odot = 10.25, 10.75, 11.25$ , from bottom to the top (see also the color code, as in the legends). We result significantly discrepant from CANDELS at higher- $z$  ( $>0.3$ ) for the lowest mass value for which our sample is complete out to  $z \sim 0.5$  ( $\text{Log}M_*/M_\odot = 10.4$ ).

We further note that the steepest variation of the sizes is found in our highest mass intercept ( $\text{Log}M_*/M_\odot = 11.6$ ), for which we measure a reduction of the size with respect to the value at  $z = 0$  that reaches about 70% at  $z > 0.5$  as shown in the bottom panel of Fig. 4.4.

In the same Fig. 4.4, we show the size evolution with respect to the  $R_e$  at  $z = 0$  for all masses as done in the upper panel. The evolution of the galaxy size over cosmic time becomes increasingly significant at larger masses, while at  $\text{Log}M_*/M_\odot \sim 10.4$  the correlation possibly inverts the trend as shown also from our estimates for  $\text{Log}M_*/M_\odot = 9.6$ . The latter is the extrapolation of our best fit to the data to a small mass value, which is in the incompleteness mass range for redshift  $z > 0.2$ . This represents a “prediction” of what we might expect at the higher- $z$  if our best fit relations will be confirmed by deeper samples with lower completeness mass. In the same figure we show the trend from Trujillo et al. (2007), based on HST observations, which shows a similar slope of the trend with the redshift of our more massive bins ( $\text{Log}M_*/M_\odot > 11$ ), but a higher normalization. We also overplot the results found for the ETG sample by Huertas-Company et al. (2013) for the only mass bin  $\text{Log}M_*/M_\odot \sim 11$ , which is consistent with our results at  $\text{Log}M_*/M_\odot = 10.8$ . This consistency has not been found with the other mass bins from the same authors.

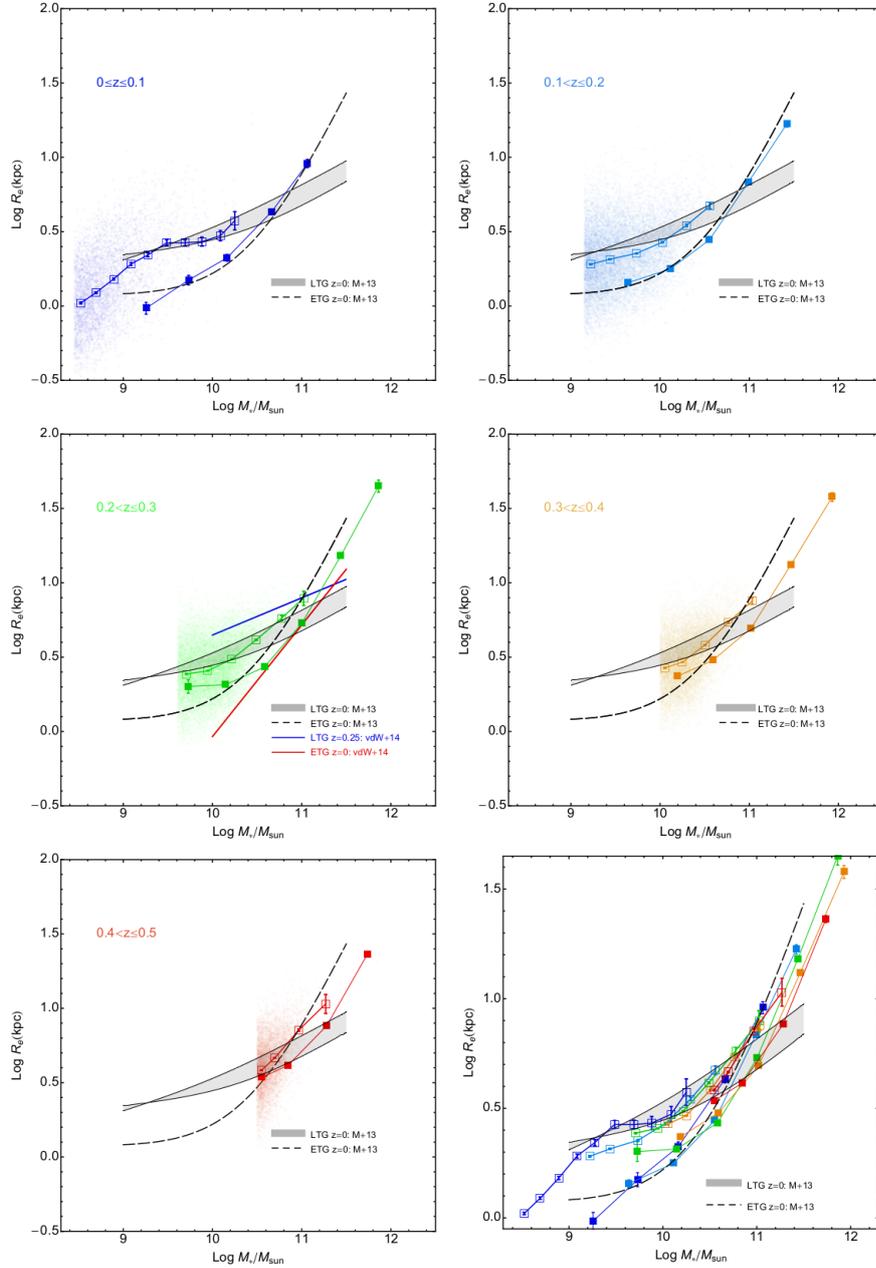


Figure 4.5: Size–Mass relation for disc-dominated galaxies. Symbols have the same meaning of the passive spheroid sample in Fig. 4.1, but now data are shown with open symbols in contrast to the passive spheroid average relation also shown as full symbols. The local relation is given by a shaded area which show the range spanned by the average relation from Mosleh et al. (2013) (i.e. their disc-dominated,  $n < 2.5$ , blue samples). In the bottom right panel we summarize all results: the disc-dominated galaxies have generally larger sizes at masses, especially for  $\lesssim 10^{11} M_{\odot}$  and show a trend with redshift (see §4.1.1) which seems weak or absent.

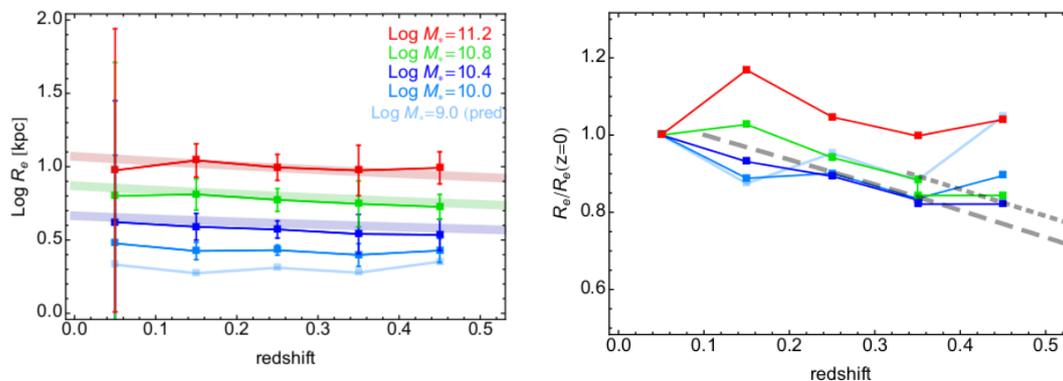


Figure 4.6: Size vs.  $z$  plots from the average size-mass parametric fit of disc-dominated galaxies, as done for passive spheroids in Fig. 4.4. *Left*: the derived absolute intercept of the best fit relations at mass values as in the legends. Error bars account for the  $1\sigma$  errors in the best fit. We also show the extrapolation for a small mass value ( $\text{Log } M_*/M_\odot = 9.0$ ), which is in the incompleteness range for redshift  $z > 0.2$ , to take as a “prediction” to be confirmed with deeper samples. We also overplot three relations from vdW+14 in transparent colors, corresponding to  $\text{Log } M_*/M_\odot = 10.25, 10.75, 11.25$ , from bottom to the top, which nicely overlap to our measurements in similar mass bins. *Right*: size evolution with respect to the  $R_e$  at  $z = 0$  at different mass intercepts. The trend of the size with redshift of disc-dominated seems constant within the errors at all mass bins, included our predictions for  $\text{Log } M_*/M_\odot = 9.0$ . The grey dashed and dotted lines represent the same relation for passive spheroids and disc-dominated respectively (from Trujillo et al. 2007)

We could not track back these discrepancies in the  $R_e(z)/R_e(z = 0)$  in the original samples from the two analyses mentioned above as the galaxy selection are somehow different from ours (e.g. Trujillo et al. 2007 use systems with  $n > 2.5$ , Huertas-Company et al. 2013 distinguish group and field galaxies) and also the local values adopted by them are different. We just report here that there is some tension between our normalization of the  $R_e(z)/R_e(z = 0)$  and other literature results.

### 4.1.2 Disc-dominated

The mass size relation of late-type systems is shown in Fig. 4.5 as open symbols and compared with the ones of passive spheroids from Fig. 4.1. In all panels we show again the local relation, by M+13, but here represented as a shaded area which reproduces the larger spanning of their inferences, depending on the different selections made (disc-dominated,  $n < 2.5$ , blue samples, etc.). Our  $z \sim 0$  results (top left) are again very well consistent with literature, and we can see a change in the overall slope at  $\text{Log } M_*/M_\odot < 9$  which is not reported in previous data. In all

other redshift bin, we see that the size–mass data tend to tilt with respect to the local relation, around a fixed mass scale ( $\text{Log}M_*/M_\odot \sim 10.5$ ).

In our sample, disc-dominated galaxies have always larger sizes at masses  $\lesssim 10^{11.0}M_\odot$ , which seems in contrast with the results from vdW+14 that found disc-dominated galaxies having smaller sizes than passive spheroids at  $M_* \gtrsim 10^{11.0}M_\odot$  (see e.g. the shaded area in 4.5, crossing the passive spheroid mean relation at  $z = 0$ ). We note though that we do not have a significant sample of disc-dominated galaxies at masses  $M_* \gtrsim 10^{11.0}M_\odot$  and we cannot exclude that the passive spheroids might in fact have larger sizes at that mass range (we expect to investigate more this issue with the next KiDS data release). We finally see that disc-dominated galaxies do not show a clear trend with redshift as seen clearly for passive spheroids.

As done for passive spheroids in §4.1.1, we have quantified the dependence on the redshift by fitting the  $R_e - M_*$  relations at the different redshifts and determining the intercept at different mass values. In this case we have used only the double power law formula (Eq. 4.2), since the data do not show any signature of the inversion of their trend at low masses. The results are shown in Fig. 4.6, for the highest mass bins for which the sample is complete at  $z < 0.5$ . Disc-dominated sizes show a very flat trend with redshifts, much flatter than the passive spheroids. This is consistent with the results from vdW+14, also shown as thick shaded lines, using the same intercept approach. We have again extrapolated the estimates to a very low mass value ( $\text{Log}M_*/M_\odot = 9.0$ ), to see whether the correlation changes with the mass, and we can confirm that this flat trend is almost independent of the mass. In the same Fig. 4.6 (bottom panel) we have also estimated the trend with redshift of the size normalized to the local value and our results seem to have a trend which is noisy but clearly flatter than the ones obtained by, e.g., Trujillo et al. (2007), especially at  $z > 0.3$  where in their sample there is a clear decreasing slope of the relation, which we cannot confirm with our data.

### 4.1.3 Size–Mass relation for passive spheroids and disc dominated galaxies in g and i bands

Here we show the Size–Mass relation for passive spheroids and disc dominated galaxies in g and i bands with a comparison to M+13 is shown in Fig. 4.7 and 4.8. The method follows same as explained in §4.1.1 and §4.1.2.

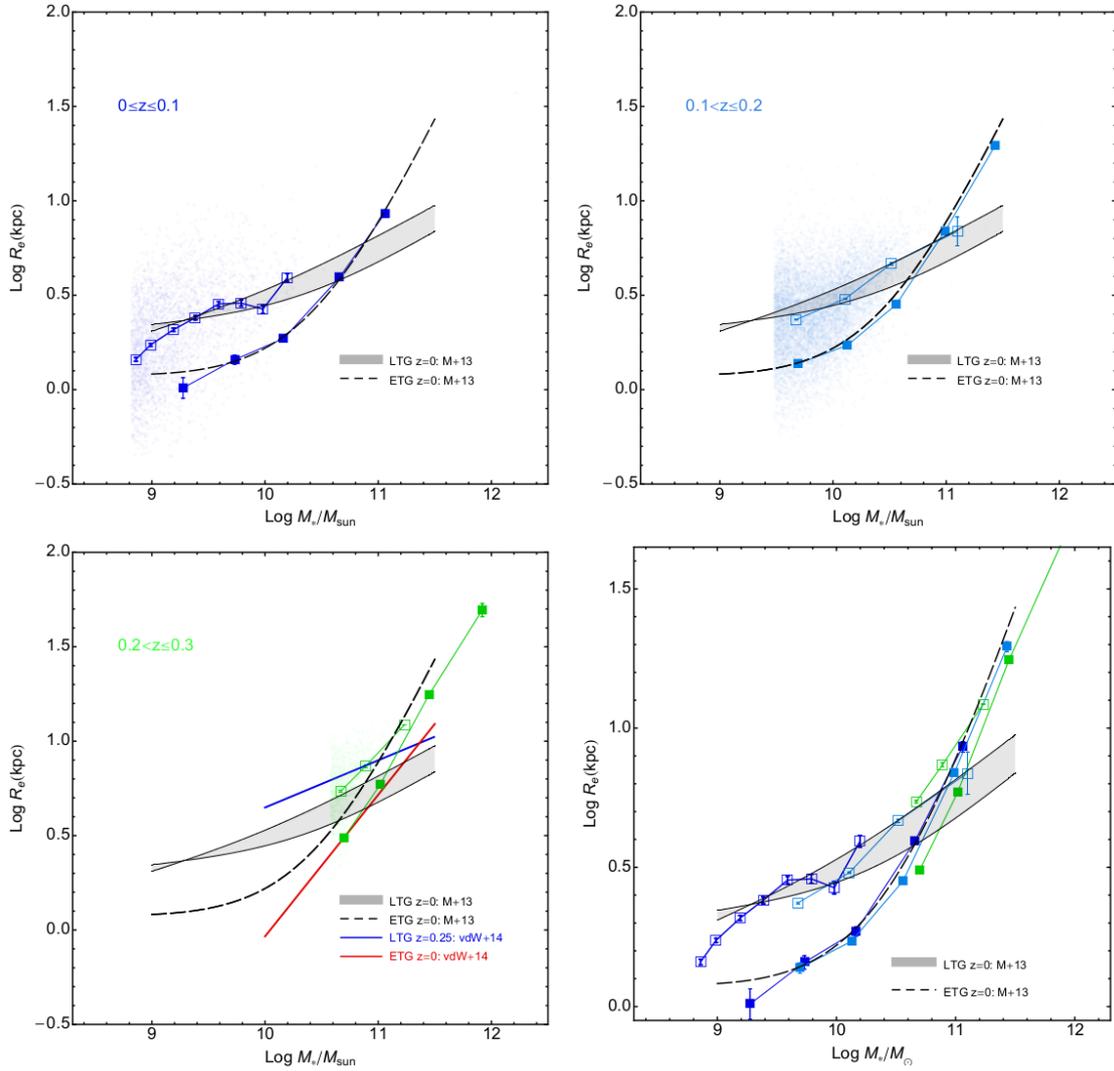


Figure 4.7: Size–Mass relation for galaxies in g band. Individual galaxy values are plotted together with mean and standard deviation of the mean (boxes and error bars) in filled square and open square for passive spheroids and disc-dominated galaxies respectively. The local relation is given by a shaded area which show the range spanned by the average relation from Mosleh et al. (2013) (i.e. their disc-dominated,  $n < 2.5$ , blue samples). In the bottom right panel we summarize all results.

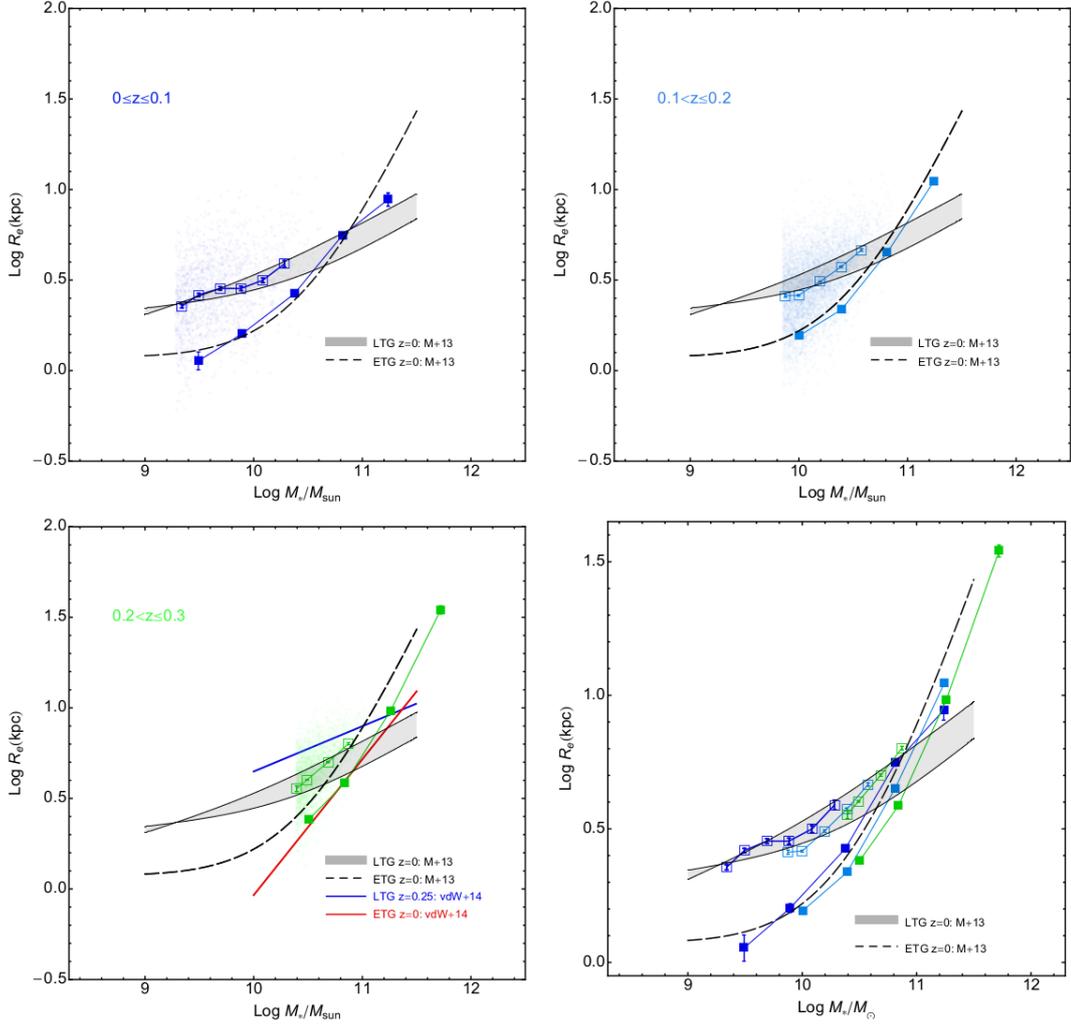


Figure 4.8: Size–Mass relation for galaxies in  $g$  band. Each panel and symbols follows same as 4.7

We found the trend in the relation is same for passive spheroids and disc-dominated as Fig. 4.1 and 4.5 for  $r$  band.

#### 4.1.4 Passive spheroids and disc-dominated size evolution parametric fit

To evaluate more quantitatively the dependence of the size with redshift, we have performed a parametric fit using the standard relation  $R_e/kpc = B_z(1+z)^{\beta_z}$  (see vdW+14) and reported the best fit parameters in Table 4.1. The results based on the average size mass relation as a function of the redshift for passive spheroids and

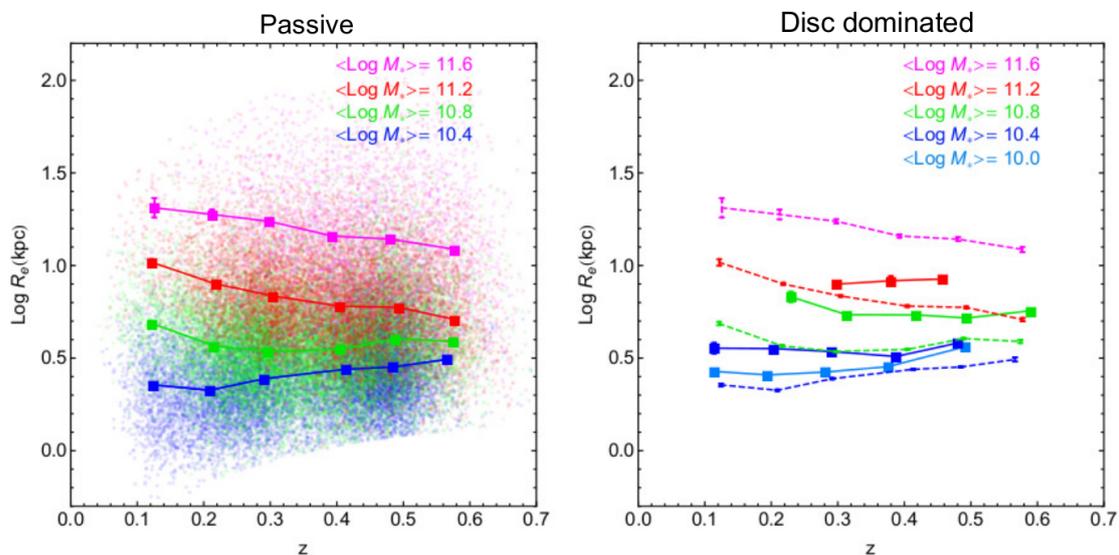


Figure 4.9: Size evolution with respect to  $z$  for passive spheroids (left) and disc-dominated galaxies (right) based on  $r$  band data. In the left panel we show the passive spheroid average relation as full squares connected by solid lines against individual galaxy values color coded according to their average mass in the mass bins. In the right panel disc-dominated values are given as squares connected by solid lines and passive spheroids are also reported as dotted lines as comparison. In each case median for redshift are given and standard deviation of mean in size is given as the error bars as filled squares.

disc-dominated galaxies are the ones in the first four columns. In the same Table we also report the results from the direct fit to the average size redshifts in different mass bins as derived in §4.2.

Looking at the values of the slope  $\beta_z$  for disc-dominated galaxies as compared to the one of the passive spheroids for the “indirect” method, it is evident that, for a given mass bin, the trend of the size with redshift is much steeper for the passive spheroids than for the disc-dominated galaxies. We will discuss this evidence in the context of the galaxy assembly in the §4.4. The overall normalization of the relation, though, given by the  $B_z$  parameter, show that disc-dominated galaxies have generally larger sizes for  $\text{Log}M_* \lesssim 11$ , while at  $\text{Log}M_* \gtrsim 11$  passive spheroids are bigger at almost all  $z$ . This seems to suggest that the tension found in the previous section about the size of disc-dominated galaxies being always larger than the passive spheroids is alleviated if we extrapolate the results of the  $R_e - z$  fit of disc-dominated galaxies and passive spheroids at masses  $\text{Log}M_* \gtrsim 11$ . We will return to these global properties of the two galaxy families in the next section.

## 4.2 Direct Size–Redshift relation in Mass bins

In the previous section we have derived the evolution of the galaxy size with redshift, starting from the evidence that the size-mass evolves with  $z$ . We have seen that at any given mass above  $\text{Log}M_* \gtrsim 10.5$  the sizes of passive spheroids decrease with increasing  $z$  (i.e. the overall size-mass normalization is decreasing with  $z$ ), while this trend was much reduced in magnitude for disc-dominated galaxies.

In this section we offer a complementary analysis of the size evolution by directly inferring the  $R_e - z$  relation in different mass bins. Being this inference non parametric, it provides a more realistic estimate of the actual dependence of the size from the redshift, although it is directly affected by the incompleteness of the sample in the different redshift and mass bins.

The results of the passive spheroids and disc-dominated galaxies are presented face-to-face in Fig. 4.9, where we show the average  $R_e - z$  dependence in different mass bins, following the mass grouping and color code adopted in the previous section, for a direct comparison of the two approaches. For the passive spheroids, we also show the individual values with the same color code to better evaluate the spread of the relation. For the disc-dominated galaxies we have omitted the individual values because, being their relative normalization in the different mass bins smaller than the passive spheroid case, it was too crowded to appreciate any difference among the different mass bins. We have performed also for these estimate the  $R_e = B_z(1+z)^{\beta_z}$  fit and best fit parameters are reported in Table 4.1.

Both the data and the parametric fit show the same features discussed for the size-redshift obtained for the “indirect” relations in the previous Section. Namely, the passive spheroids show steeper decreasing trends with  $z$  for mass bins  $\text{Log}M_* \gtrsim 10.5$  while they almost flatten out at lower masses. Disc-dominated galaxies show shallower slopes (see Table 4.1) than passive spheroids and, at masses  $\text{Log}M_* \lesssim 11.0$ , they show larger sizes than the passive spheroids (see the comparison between passive spheroids and disc-dominated galaxies in Fig. 4.9, right panel). We will interpret these different variations of the size with  $z$  in the next paragraph.

## 4.3 Effect of the errors on the $R_e - z$ relations

We want to check the effect of the uncertainties on the different quantities entering into the size-redshift trends discussed in §4.2. The trend found can indeed be affected from the intrinsic scatter of the mass, effective radius estimate and photometric redshift. In principle, the covariance among the individual errors might spuriously generate a correlation from the observed quantities. On the other way around, the observed trend can be even shallower than the intrinsic one for the scatter due to the different quantities that move objects from one bin to another, hence diluting the

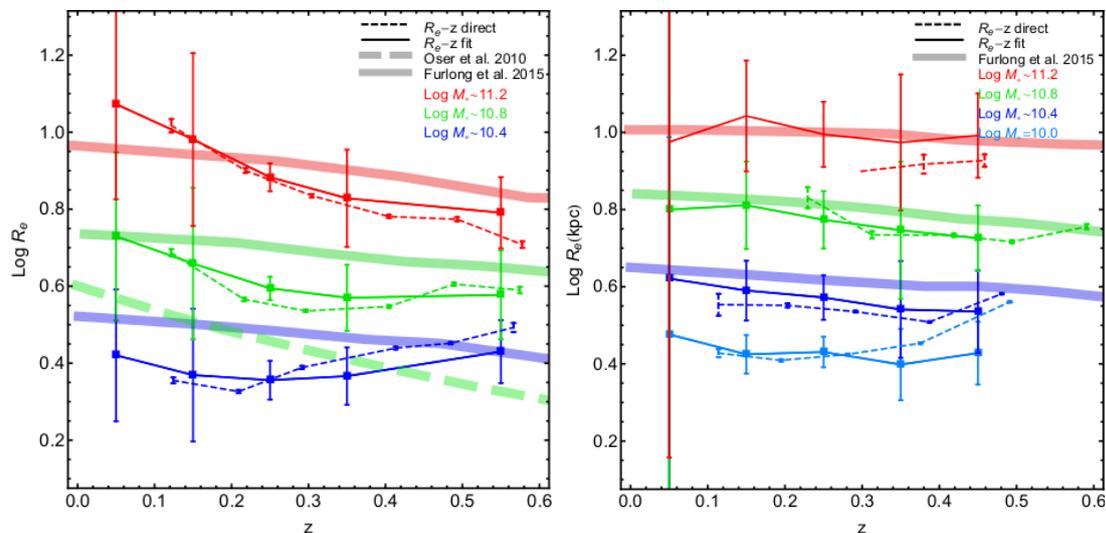


Figure 4.10: Comparison with hydrodynamical simulations. Left panel: passive spheroidal size–redshift relation is compared with the EAGLE simulations from Furlong et al. (2015) and with a modified version of GADGET-2 from Oser et al. (2012). The data points show the results obtained in §4.1.1 (solid lines) and §4.2 (dashed lines) in different mass bins as in the legends. Right panel: disc-dominated galaxies size–redshift relation with symbols as in the left panel but with the results from the average-size mass from §4.1.2 (solid lines).

real trends. In order to check for the presence of these effects, and evaluate in which direction the correlations that we have derived in §4.1.1 and §4.1.2 and reported in Table 4.1 can be biased by the intrinsic scatter of the individual parameters, we have performed a series of bootstrap experiments to obtain random resamplings of our datasets. We have perturbed all galaxy mass,  $R_e$  and  $z_{\text{phot}}$  by randomly adding a offset extracted from a Gaussian distribution having zero mean and a constant standard deviation equal to the average error of the different quantities (namely  $\sigma_{\text{Log}M_*}$ ,  $\sigma_{R_e}$ ,  $\sigma_{z_{\text{phot}}}$  for the mass, size and  $z_{\text{phot}}$  respectively), hence resampling the same observed relations, but adding the effect of random errors on the individual parameters.

In §3.1.3 we have mentioned that average errors on masses are of the order of 0.2dex (maximum errors reaching 0.3dex), while the relative errors on  $R_e$  are of the order of 15% (20% maximum, see e.g. Figs. 3.14, 3.16, 3.17), while the scatter for the  $z_{\text{phot}}$  has been reported to be of the order of  $0.03(1 + z_{\text{spec}})$  (see §2.3.7). We have re-extracted the catalog values 100 times and obtained, at every extraction, a correlation like Fig. 4.9, which we have finally averaged to obtain the average trend in each mass bin.

In Fig. 4.11 we show the “bootstrap” results for Passive and Disc dominated galaxies obtained both for the average errors (solid lines) and for the maximum

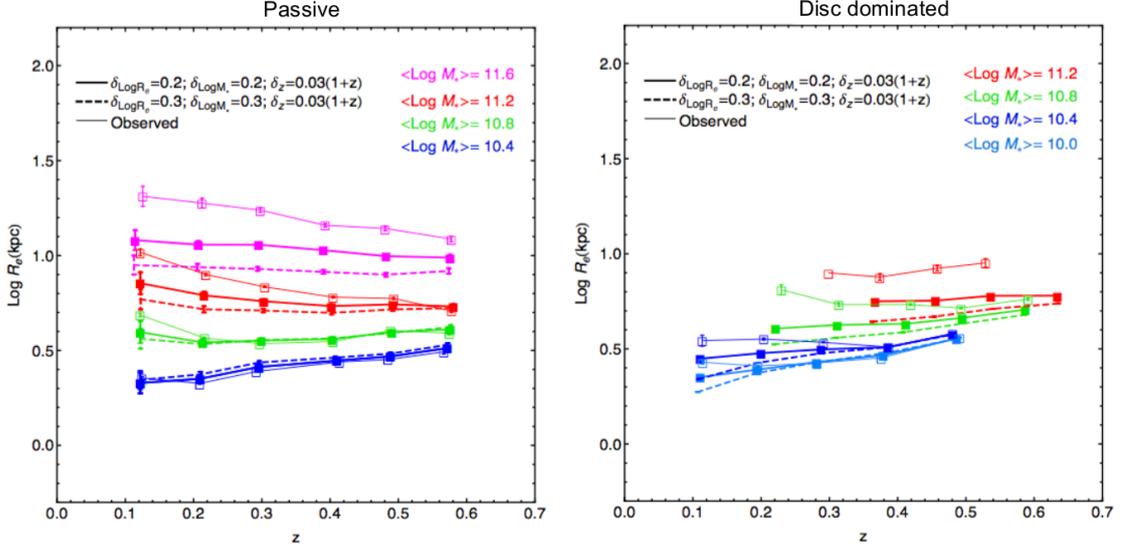


Figure 4.11: Effect of the errors on the  $R_e$ - $z$  correlation as derived in Fig. 4.9. Left: average trend for the Passive sample obtained by resampling the original sample parameters by adding a shift from a Gaussian distribution using average errors (solid line) and maximum errors (dashed lines) as in the legends. The observed trend is also shown as comparison (tiny solid line). Right: the same as in the left panel for the disc dominated galaxies.

errors (dashed lines). We can clearly see that for the passive spheroids, the larger the errors assumed the flatter is the final trend obtained. This demonstrates that the effect of the uncertainties on the quantities is statistically to reduce the steepness of the observed trends (tiny lines in Fig. 4.11) rather than to introduce a spurious slope. The same effect is also seen for the Disc dominated galaxies although, for the lower mass bin, we see that errors produce a steepening of the correlation in the lower redshift bins.

We have also checked that the quantities that is affecting more the trend is the mass as the scattered quantities move galaxies from the central mass bins to the contiguous (small and larger mass) ones, hence making all relations to converge toward the ones of the intermediate bins, as shown from the case of maximum errors.

Overall this test demonstrates that the trends discussed in Table 4.1 are real and possibly shallower than the ones that we had measured if we could reduce the uncertainties on the observed quantities. The only exception is for low-mass Disc dominated systems (i.e.  $\text{Log} M_*/M_\odot \lesssim 10.2$ ), that at lower redshift ( $z < 0.2$ ) might have a steeper trend with respect to the almost flat trend observed in Fig. 4.9.

As discussed above, the major source of uncertainties in the derived trends is the one on the mass, which we plan to reduce in the future by adding NIR bands in our SED estimates.

## 4.4 Discussion

The main result of this work is that the two main classes of galaxies, passive and disc-dominated, show different relations between size and stellar mass and size and redshift, which are well consistent with previous literature (Shen et al. 2003, Baldry et al. 2012, van der Wel et al. 2014). Our sample, complete in mass down to  $\text{Log}M_* \lesssim 9.0$  at  $z < 0.2$  and down to  $\text{Log}M_* \gtrsim 10.0$  at higher- $z$ , has allowed us to highlight some features that were not clearly assessed in previous datasets. First a curvature in the  $R_e - M_*$ , which seems present at almost all  $z$ -bins for both passive spheroids and disc-dominated galaxies, but becomes less clear at  $z > 0.4$ , mainly because of the mass incompleteness. The disc-dominated galaxies size-mass relation also presents a knee in the relation at the very low masses ( $\text{Log}M_* \lesssim 9.5$ ) at  $z < 0.1$ , which was not reported earlier and needs to be checked against galaxy simulations in order to be interpreted.

These trends are partially anticipated by previous literature, where the size and mass-size studies have been performed up to very large  $z$  (e.g. Trujillo et al. 2006, Huertas-Company et al. 2013, van der Wel et al. 2014), however this is the first time that such a large dataset has been used to investigate intermediate redshift with unprecedented fine coverage in mass and redshift.

The results found for our passive spheroid and disc-dominated samples are consistent with the expectation of the galaxy growth from recent hydrodynamical simulations (Furlong et al. 2015) from the EAGLE set-up (Schaye et al. 2015), as demonstrated in Fig. 4.10. Overall, the predictions from simulations nicely match our trends at all mass scales, although the match of the passive spheroids is slightly more discrepant with respect to the striking agreement found for disc-dominated galaxies, especially for the higher mass values. The sizes predicted for passive spheroids from simulations are slightly larger than the data, possibly revealing a too strong feedback recipe to prevent the simulated systems to collapse too much. This is a well known effect of hydrodynamical simulations (e.g. Scannapieco & Athanassoula 2012) as a consequence of the so-called angular momentum catastrophe (Katz & Gunn 1991; Navarro & White 1994) consisting in a too large angular momentum transfer into the galaxy haloes which cannot retain the collapse of the cold gas into stars toward the galaxy center. The effect is balanced by the inclusion of feedback mechanisms in the centers, which balance the gas collapse (e.g. Governato et al. 2004; Sales et al. 2010; Hopkins et al. 2014 etc.), but whose recipes are still under refinement. To demonstrate directly the effect caused by an insufficient energy injection in simulations, in Fig. 4.10 we also show the predictions of the  $R_e - z$  for  $\sim 10^{11} M_\odot$  passive spheroid from Oser et al. 2012 (note that they do not provide explicitly disc-dominated predictions) with a modified version of the parallel TreeSPH code GADGET-2 (Springel 2005) and no AGN feedback, ending up in more compact systems for a given mass bin. Our data are bracketed by the predictions of the

two sets of simulations (EAGLE and GADGET-2) with different feedback set-up, indicating that there is still something missing into the passive spheroid formation pattern.

The remarkable result that emerges from this comparison is that the observed sizes are naturally explained in the context of the galaxy assembly described in the cosmological simulations. In particular, the size growth is interpreted in Furlong et al. (2015) as the consequence of the accreted mass fraction since  $z = 2$ . The more stellar mass is accreted from sources other than the main progenitor at a given time the more the final size of a galaxy are found to increase. This does not take into account the type of mergers that contribute to the size growth, but clearly establish that size growth and accreted mass fraction are inherently related (see their Fig. 5). Summary of these results present in this chapter is given in Chapter 7.

# Chapter 5

## Super-Compact massive Galaxies

The understanding of the physical processes which drive the galaxy mass built up and size accretion are among the most timely topics in galaxy evolution studies. Massive early-type galaxies (ETGs) are found to be much more compact in the past than in the present Universe (Daddi et al. 2005; Trujillo et al. 2006; Trujillo et al. 2007; van der Wel et al. 2008). At redshifts  $z > 2$ , while the massive star-forming disks have effective radii of several kpc (Genzel et al. 2008), the quenched spheroids (“red nuggets”) have small effective radii of about 1 kpc. Such red nuggets are thought to form through a chain of different processes: a) accretion-driven violent disc instability, b) dissipative contraction resulting in the formation of compact, star-forming “blue nuggets”, c) quenching of star formation (Dekel & Burkert 2014). After these processes occur, a gradual expansion in size of the red nuggets may take place, leading to the formation of the massive ETGs we observe in the nearby Universe. Theoretical studies point to dry-mergers as the dominant mechanism for the size and stellar mass growth of very dense massive galaxies (Khochfar & Silk 2006). In particular, minor mergers would provide a modest stellar mass accretion, but a strong evolution in galaxy size (van Dokkum et al. 2010; Hilz et al. 2013; Belli et al. 2014; Tortora et al. 2014b). Mergers are believed to be common for very massive systems at high redshifts, with major merger rates (mergers per galaxy per Gyr) in the range  $0.3 - 1 \text{ Gyr}^{-1}$  at  $z \sim 2$  and smaller than  $0.2 \text{ Gyr}^{-1}$  at  $z \lesssim 0.5$  (Hopkins et al. 2010b). An alternative scenario explains the size evolution as the result of quasar feedback, rather than merging, making galaxies to puff up after a loss of large amounts of (cold) gas (Fan et al. 2008, 2010). In this chapter we present the KiDS data sample and the selection of our photometrically selected compact galaxies. The main results, and in particular the evolution of number density as a function of redshift, are presented in §5.3.

## 5.1 Introduction

Over cosmic time, one may expect that high- $z$  compact galaxies evolve into present-day, massive, big galaxies. However, a fraction of these objects might survive intact till the present epoch, resulting in compact, relic systems in the nearby Universe characterized by old stellar populations. Recently there have been some efforts to search for massive compact galaxies at low redshifts (Trujillo et al. 2009; Taylor et al. 2010; Valentinuzzi et al. 2010; Shih & Stockton 2011; Trujillo et al. 2012; Poggianti et al. 2013a; Poggianti et al. 2013c; Damjanov et al. 2013, 2014, 2015a; Saulder et al. 2015) and investigate further their dynamical and stellar population properties, as well as the role of environment on their properties (e.g. Ferré-Mateu et al. 2012; Läsker et al. 2013; Ferré-Mateu et al. 2015; Trujillo et al. 2014; Yıldırım et al. 2015; Damjanov et al. 2015b; Stringer et al. 2015; Wellons et al. 2016).

In some theoretical models, that include the effect of galaxy mergers, the fraction of massive objects that would survive without undergoing any significant transformation since  $z \sim 2$  to the present could reach a fraction of about 1 – 10% (Hopkins et al. 2009a; Quilis & Trujillo 2013). At “low” redshifts ( $z \lesssim 0.2$ ), theoretical models predict a density of relic remnants in the range  $10^{-7} - 10^{-5}$ , which means that, in large surveys like the Sloan Digital Sky Survey (SDSS), we might expect to find a few candidates in this redshift range. However, recent observational works have shown the paucity of old, super-compact ( $R_e \lesssim 1.5$  kpc), massive ( $M_* \gtrsim 10^{11} M_\odot$ ) galaxies in the local Universe ( $z \lesssim 0.2$ ; Trujillo et al. 2009; Taylor et al. 2010; Ferré-Mateu et al. 2012). Indeed, NGC1277, a nearby lenticular galaxy in the Perseus cluster, is actually the only well-characterized, old system at  $z \sim 0$ , that might be a true relic galaxy (Trujillo et al. 2014; see also Martín-Navarro et al. 2015). Other candidates have been recently detected by Saulder et al. (2015), although only a few of them fulfill the above (restrictive) size and mass criteria ( $R_e \lesssim 1.5$  kpc and  $M_* \gtrsim 10^{11} M_\odot$ ), none of them being at  $z < 0.05$ . Larger numbers of old compact systems have been found at lower masses ( $< 10^{11} M_\odot$ ), when relaxing the compactness selection criteria (Valentinuzzi et al. 2010; Poggianti et al. 2013a). The (almost) lack of nearby super-compact, relic systems may represent a challenge for the current paradigm of galaxy formation; in particular, we have to understand whether this lack is due to some observational bias, e.g. the limited spatial resolution of photometric data at  $z \sim 0$ ; a failure of theoretical predictions; and/or an environmental effect, for which, as suggested by NGC 1277, relic galaxies might be more frequent in high-density cluster regions.

In the intermediate redshift range ( $0.2 \lesssim z \lesssim 0.7$ ), compacts have been recently investigated in detail by Damjanov et al. (2014), who selected  $\sim 200$  massive compacts from a sample of stellar-like objects within the 6373.2 sq. deg. of the BOSS survey; 20% of these galaxies are dominated by old stellar populations, which make them reliable candidates to be the product of the unperturbed evolution of compact

high- $z$  systems. However, 93% of these galaxies do not have measured  $R_e$  which hampers the selection of such systems as truly compact objects. More recently, Damjanov et al. (2015a), have analyzed F814W HST images for the COSMOS field, providing robust size measurements for a sample of 1599 compact systems in the range  $0.2 \lesssim z \lesssim 0.8$ . Other studies have performed detailed analysis of stellar populations and morphology of small samples of compact galaxies at these redshifts (Hsu et al. 2014; Stockton et al. 2014). The population of dense passively evolving galaxies in this intermediate redshift range possibly represents a link between compact systems, dominating the massive quiescent galaxy population at high  $z$ , and their descendants in the nearby Universe. Indeed, large samples of compacts, with high-quality photometry (to derive reliable structural parameters), and spectroscopic data, are actually necessary to better understand the formation and evolution of these systems.

The Kilo Degree Survey (KiDS; de Jong et al. 2015) is one of the ESO public surveys carried out with the VLT Survey Telescope (VST; Capaccioli & Schipani 2011) equipped with the one square degree field of view and high angular resolution ( $0.2''/\text{pixel}$ ) OmegaCAM camera (Kuijken et al. 2004; Kuijken 2011). KiDS is mainly designed for weak lensing studies, providing deep imaging in four optical bands ( $ugri$ ), over a 1500 square degree of the sky with excellent seeing (e.g.  $0.65''$  median FWHM in  $r$ -band). The high image quality and deep photometry are ideal to investigate massive compact systems.

According to predictions from simulations (Guo et al. 2011a, 2013), we can expect to find  $\sim 0.3 - 3.5$  relic per square degree, at redshift  $z < 0.5$ . This prediction does critically depend on the physical processes shaping size and mass evolution of galaxies, such as the relative importance of major and minor galaxy merging.

Several compact galaxy definitions have been adopted in the literature (Trujillo et al. 2009; Taylor et al. 2010; Poggianti et al. 2013a; Damjanov et al. 2015a). In the present work, we present the properties of a sample of dense massive galaxy candidates in KiDS, defining as massive super-compact galaxies (MSCGs) those early-type systems with  $M_\star > 8 \times 10^{10} M_\odot$  and  $R_e < 1.5$  kpc (Trujillo et al. 2009). These selection criteria are rather conservative, providing the ideal benchmark for galaxy evolution theories. This chapter is organized as follows. In §5.2 we present the KiDS data sample and the selection of our photometrically selected compact galaxies. The main results, and in particular the evolution of number density as a function of redshift, are presented in §5.3. A discussion of the results and future prospects are outlined in §5.4. We adopt a cosmological model with  $(\Omega_m, \Omega_\Lambda, h) = (0.3, 0.7, 0.70)$ , where  $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (Komatsu et al. 2011).

## 5.2 Sample selection

The galaxy sample presented in this work is based on the data included in the first and second data releases of KiDS presented in de Jong et al. (2015), which we address the interested reader for details. The total dataset includes 156 KiDS pointings (133 from the KiDS data release 2), in which we have identified about 22 million sources including  $\sim 7$  millions which have been classified as high quality extended sources (mostly galaxy-like, see below). A full description of the galaxy sample is given in Napolitano et al. (2016, in prep.). In the following section, we summarise the main steps for the galaxy selection procedure and the determination of galaxy physical quantities as structural parameters, photometric redshifts and stellar masses.

### 5.2.1 KiDS high signal-to-noise galaxy sample

We start from the KiDS multi-band source catalogs, where the photometry has been obtained with S-Extractor (Bertin & Arnouts 1996) in dual image mode, using as reference the positions of the sources detected in the  $r$ -band images. While magnitudes are measured for all of the filters, the star/galaxy separation, positional and shape parameters are based on the  $r$ -band data. The choice of  $r$ -band is motivated by the fact that it typically has the best image quality and thus provides the most reliable source positions and shapes. Critical areas as saturated pixels, star spikes, reflection halos, satellite tracks, etc. have been masked using both a dedicated automatic procedure and visual inspection. The total area after removing the borders of the individual KiDS pointings is 163.9 square degrees, while the unmasked effective area adopted is of 105.4 square degrees.

Star/galaxy separation is based on the distribution of the S-Extractor parameters `CLASS_STAR` and  $S/N$  (signal-to-noise ratio) of a number of sure stars (see La Barbera et al. 2008 and de Jong et al. 2015 for further details). We have further retained those sources which were marked as being out of critical area from our masking procedure. From the original catalog, the star/galaxy separation leaves  $\sim 11$  million galaxies, of which  $\sim 7$  millions have high quality photometry being non-deblended sources located out of the masked area.

To perform accurate surface photometry and determine reliable structural parameters, the highest-quality sources have been further selected (La Barbera et al. 2008, 2010b). Thus, we have finally gathered those systems with the highest  $S/N$  in the  $r$ -band images,  $S/N_r \equiv 1/\text{MAGERR\_AUTO\_r} > 50$  (La Barbera et al. 2008, 2010b; Roy et al. 2016, in preparation). This sample consists of  $\sim 380\,000$  galaxies.

Relevant properties for each galaxy are derived as described here below:

- *Photometry.* As a standard KiDS catalogs' parameters, we have derived S-Extractor aperture photometry in the four bands ( $ugri$ ) within several radii.

For our analysis we have adopted aperture magnitudes `MAGAP_2`, `MAGAP_4` and `MAGAP_6`, measured within circular apertures of 2", 4" and 6" of diameter, respectively, while a first probe of the total magnitude is provided by the Kron-like `MAG_AUTO`.

- *Structural parameters.* Surface photometry is performed using the 2DPHOT environment. 2DPHOT is an automatic software designed to obtain both integrated and surface photometry of galaxies in wide-field images. The software first produces a local PSF model from a series of identified *sure stars*. This is done, for each galaxy, by fitting the four closest stars to that galaxy with a sum of three two-dimensional Moffat functions. Then galaxy snapshots are fitted with PSF-convolved Sérsic models having elliptical isophotes plus a local background value (see La Barbera et al. 2008 for further details). The fit provides the following parameters for the four wavebands: surface brightness  $\mu_e$ , circularized effective radius,  $R_e$ , Sérsic index,  $n$ , total magnitude,  $m_S$ , axis ratio,  $q$ , and position angle. As it is common use in the literature, in this section we use the circularized effective radius,  $R_e$ , defined as  $R_e = \sqrt{q}R_{e,\text{maj}}$ , where  $R_{e,\text{maj}}$  is the major-axis effective radius.
- *Photometric redshifts.* Redshifts are determined with the Multi Layer Perceptron with Quasi Newton Algorithm (MLPQNA) method (Brescia et al. 2013, 2014), and fully presented in Cavuoti et al. (2015), which we refer for all details. Both apertures of 4" and 6" of diameter are used. In machine learning supervised methods, a knowledge sample is needed to train the neural network performing the mapping between magnitudes and redshift. The knowledge base consisted on spectroscopic redshift from the Sloan Digital Sky Survey data release 9 (SDSS-DR9; Ahn et al. 2012) and Galaxy And Mass Assembly data release 2 (GAMA-DR2; Driver et al. 2011) which together provide redshifts up to  $z \lesssim 0.8$ . This knowledge base includes  $\sim 60000$  objects, 60% of which are used as training sample, and the remaining ones are used for the blind test set. The redshifts in the blind test sample resemble the spectroscopic redshifts quite well. The scatter in the quantity  $\Delta z \equiv (z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}})$  is  $\sim 0.03$ . After these experiments, we have finally produced the final catalogue of redshifts for our sample. The cut operated in the fitting magnitudes to resemble the luminosity ranges in the knowledge base will impact the completeness in the faint regime. But our high- $S/N$  sample is not affected by this source of incompleteness.
- *Stellar masses.* We have used the software **Le Phare** (Arnouts et al. 1999; Ilbert et al. 2006), which performs a simple  $\chi^2$  fitting method between the stellar population synthesis (SPS) theoretical models and data. Single burst models from Bruzual & Charlot (2003) and a Chabrier (2001) IMF are used.

Models are redshifted using the photometric redshifts. We adopt the observed *ugri* magnitudes (and related  $1\sigma$  uncertainties) within a  $6''$  aperture of diameter, which are corrected for Galactic extinction using the map in Schlafly & Finkbeiner (2011). Total magnitudes derived from the Sérsic fitting,  $m_S$ , are used to correct the outcomes of **Le Phare** for missing flux. The single burst assumption is suitable to describe the old stellar populations in the compact galaxies we are interested in (Thomas et al. 2005; Tortora et al. 2009). The estimated photometric ages are used to check if galaxies are compatible with being relic remnants of systems formed at  $z \sim 2$ . The degeneracy between age and metallicity in the stellar population analysis can be solved only with forthcoming spectroscopic follow-ups.

- *Galaxy classification.* Using **Le Phare**, we have also fitted the observed magnitudes **MAGAP\_6** with a set of 66 empirical spectral templates used in Ilbert et al. (2006). The set is based on the four basic templates (Ell, Sbc, Scd, Irr) described in Coleman et al. (1980), and star burst models from Kinney et al. (1996). GISSEL synthetic models (Bruzual & Charlot, 2003) are used to linearly extrapolate this set of templates into ultraviolet and near-infrared. The final set of 66 templates (22 for ellipticals, 17 for Sbc, 12 for Scd, 11 for Im, and 4 for starburst) is obtained by linearly interpolating the original templates, in order to improve the sampling of the color space. We have selected the ETGs by choosing the galaxies which are best fitted by one of the elliptical templates (see Napolitano et al. 2016, in preparation).
- *Near Infrared photometry from VIKING-DR1.* As a complementary dataset for our selection of compact galaxies, we have used the data from the first release (DR1) of the ESO VISTA Kilo Degree Infrared Galaxy (VIKING) survey. The VIKING survey is the KiDS twin survey, and provides the near-IR coverage of the same sky region in the 5 wavebands Z, Y, J, H and Ks. VIKING-DR1 consists of 108 observed tiles (Edge et al. 2014). We have found that it overlaps with  $\sim 58\%$  of our 156 sq. deg. in the KiDS survey. For uniformity, we have used photometry within the same aperture as the optical data: **APERMAG\_3** for VIKING and **MAGAP\_2** for KiDS, which correspond to an aperture of  $2''$  of diameter. In particular, we have concentrated on J and Ks passbands, which we do not use for SPS fitting, but only to improve the star-galaxy separation criterion (Maddox et al. 2008; Muzzin et al. 2013) as will be discussed in details in the next section.

The sample of high- $S/N$  galaxies is complete down to a magnitude of **MAG\_AUTO\_r**  $\sim 21$ , which correspond to stellar masses  $\gtrsim 5 \times 10^{10}$  up to redshift  $z = 0.5$  (see see Napolitano et al. 2016, in prep., for further details).

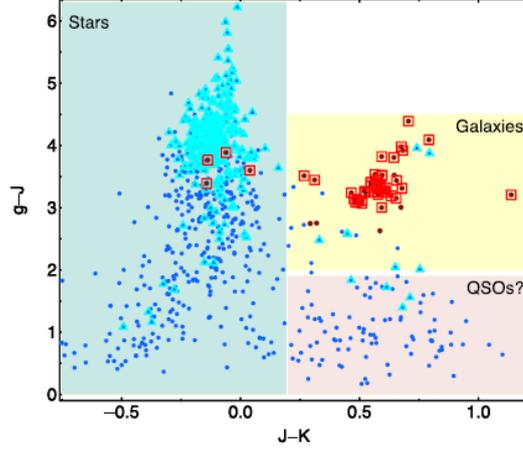


Figure 5.1:  $J - K$  vs.  $g - J$  diagram. An aperture of  $2''$  of diameter is adopted here. Blue and cyan symbols are for high-confidence stars with any error and  $\delta J = \delta K < 0.05$ , respectively. Red points and open boxes are for compact candidates with any error and  $\delta J = \delta K < 0.05$ , survived after the criteria on the mass, size, 2D fit quality and visual inspection. Vega J and K magnitudes are converted to AB as  $J_{\text{AB}} = J_{\text{Vega}} + 0.930$  and  $K_{\text{AB}} = K_{\text{Vega}} + 1.834$ . We highlight the regions which are populated by stars (blue), galaxies (yellow) and QSOs (pink). We have considered as sure galaxies those objects with colors  $J - K > 0.2$  and  $g - j > 2$  (yellow shaded region).

### 5.2.2 Selection of compact galaxies

MSCGs have been selected using the following criteria:

1. *Massiveness.* The most massive galaxies with  $M_{\star} > 8 \times 10^{10} M_{\odot}$  are taken (Trujillo et al. 2009), reducing the original sample of  $\sim 380000$  galaxies to  $\sim 30600$  massive galaxies.
2. *Compactness.* We select the densest galaxies by following recent literature (Trujillo et al. 2009). We get galaxies with median circularized radius,  $R_e$ , among the  $g$ -,  $r$ - and  $i$ -bands, less than 1.5 kpc. About 1300 compact candidates remain after this selection.
3. *Best-fitted effective radii.* The best fitted structural parameters are considered, taking those systems with a reduced  $\chi^2_{2D}$  from 2DPHOT smaller than 1.5 in  $g$ ,  $r$  and  $i$  filters (La Barbera et al. 2010b). To avoid any accidental wrong fits we have also removed galaxies with unreasonable best-fitted parameters, applying a minimum value for the size ( $R_e = 0.05$  arcsec), the Sérsic index ( $n > 0.5$ ) and the axial ratio ( $q = 0.1$ ) in all the bands. The minimum value in the Sérsic index has also allowed us to possibly remove misclassified stars, which

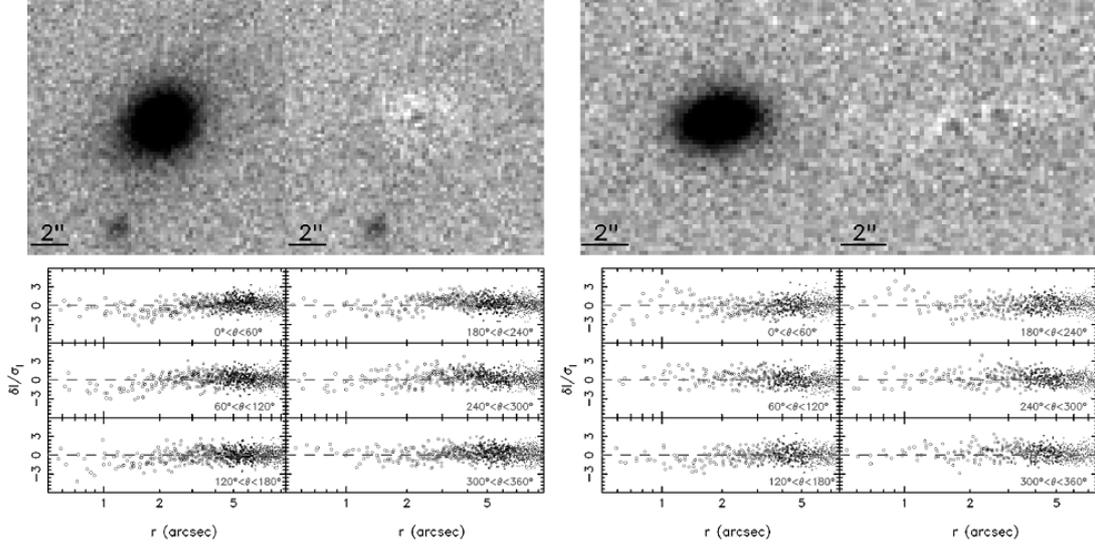


Figure 5.2: 2D fit output for 2 example candidates from 2DPHOT procedure. For each galaxy, the top panels show the galaxy image (left) and the residual after the fit (right), while the six bottom panels provide residuals after subtraction, plotted as a function of the distance to the galaxy center, with each panel corresponding to a different bin of the polar angle. Residuals are normalized with the noise expected from the model in each pixel.

are expected to be fitted by a box-like profile (mimicked by a Sérsic profile with  $n \rightarrow 0$ ). These last criteria on the quality of the structural parameters reduce further the sample to the 106 highly reliable candidates.

4. *Eye-ball check.* We have made an eyeball inspection of the images and residuals from Sérsic fitting of the candidates, with the aim of removing problematic objects or possible misclassified stars. To reduce subjectivity, three of the authors have independently checked the images and graded them according to the following scheme: grade-2 are sure galaxies, grade-1 are uncertain galaxies lacking a well-defined elliptical shape, and grade-0 are misclassified objects (either stars or corrupted fits). The mean of the three classifications has been adopted as the final grade. In this way, we have retained those galaxies with a grade larger than 1, to include only systems graded by at least one of the observers as a sure galaxy. The candidates are further reduced to 96, after removing objects with significant contamination from neighbours.
5. *Optical+NIR star-galaxy separation.* We have adopted a morphological criterion to perform the star-galaxy classification (Bertin & Arnouts 1996; La Barbera et al. 2008) and used visual inspection as ultimate check of the galaxy

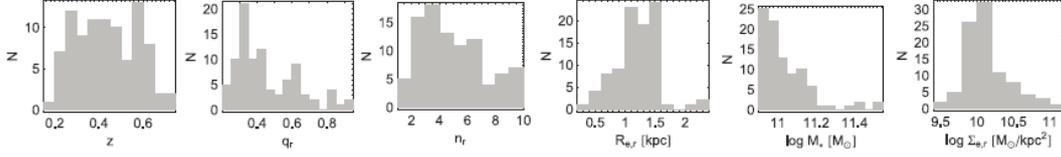


Figure 5.3: Distribution of some galaxy parameters. From left to right we show: 1) photometric redshifts, 2) r-band axis ratio, 3) r-band Sérsic index, 4) r-band effective radius, 5) Chabrier IMF-based stellar mass, 6) r-band effective surface mass density.

classification. However, based on optical data only a star can be still misclassified as a galaxy on the basis of its morphology, and this issue can be highly dramatic for very compact objects (generally with size comparable or smaller than the seeing). In absence of spectroscopic information, optical+NIR color-color diagrams can provide a strong constraint on the galaxy nature of the candidates (e.g. Maddox et al. 2008; Muzzin et al. 2013). In particular,  $g$ ,  $J$  and  $Ks$  magnitudes within  $2''$  of diameter are adopted for this purpose on those fields with coverage by the two surveys. The Vega VISTA magnitudes are converted to AB using the conversion formulae from **Le Phare**,  $J_{AB} = J_{Vega} + 0.930$  and  $K_{AB} = K_{Vega} + 1.834$ . Stars and galaxies with the best  $J$  and  $Ks$  photometry are also considered ( $\delta J, \delta K < 0.05$ ). The results are shown in Fig. 5.1. Stars have blue  $J - K$  colors (i.e.,  $J - K \lesssim 0.2$ , see light blue shaded region in Fig. 5.1). We have also found four objects with  $J - K \lesssim 0.2$  and, indeed, are erroneously classified as galaxies. We take as compact candidates those systems with  $J - K > 0.2$  and  $g - J > 2$  (see light-yellow shaded region in Fig. 5.1). For those galaxies with available VIKING photometry, this selection allows our set of candidates from the previous classification steps to be refined. Out of 46 galaxies with VIKING photometry, we remain with 42 high-confidence candidates.

In Fig. 5.1 we have also highlighted the locus populated by point sources with red  $J - K$  ( $\gtrsim 0.2$ ), but blue  $g - J$  ( $\lesssim 2$ ) colors, classified as stars by S-Extractor, which are presumably quasars (Maddox et al. 2008, see pink region in Fig. 5.1). However, the analysis of this class of objects is beyond the scope of this chapter.

To perform an homogeneous comparison of the sample of our compact candidates with a sample of "normal" galaxies, using the same original catalog described in §5.2, we have created two control samples (**C-Samples**) of galaxies with similar stellar masses ( $M_* > 8 \times 10^{10} M_\odot$ ), the same lower limits on Sérsic index ( $n > 0.5$ ), axis ratio ( $q > 0.1$ ) and  $R_e$  ( $R_e > 0.05$  arcsec) as those for the compact systems, but without applying any compactness criterion. The first **FULL C-Sample** consists of all

galaxies, with no restriction on the galaxy type, while the **ETGs C-Sample** consists of ETGs, classified as described in §5.2.

## 5.3 Results

### 5.3.1 The final sample

We start by summarizing the sample we are left with for our analysis: after the first three criteria we have selected 106 candidates, which are reduced to 96 after the eye-ball check in (iv). The matching with VIKING included 46 candidates and, after applying the criterion (v), we are left with a sample of 42 out of 46 galaxy candidates (i.e. 4/46 objects are likely stars according to their optical-NIR colors, corresponding to a  $\sim 10\%$  contamination). Updating the total number, we are left with 92 candidates, of which  $\sim 4$  further sources from the region that does not overlap with VIKING might be stars. This corresponds to a density of  $\sim 0.9$  compact galaxies per square degree, cumulatively for  $z \lesssim 0.7$ , while we do not find objects at  $z \lesssim 0.2$ . In contrast, Trujillo et al. (2009) find 29 secure **MSCGs** at  $z < 0.2$  within the SDSS area, using our same criteria. However, only one of these is at very low redshift ( $z < 0.1$ ), and none of them features old stellar populations (Ferré-Mateu et al. 2012), and thus they cannot actually be classified as relics. The absence of compact systems at  $z < 0.2$  in our datasample might be related to environmental issues, since in the current KiDS area we are missing nearby clusters of galaxies, where a larger fraction of compact galaxies is predicted (Stringer et al. 2015), and seems to be observed indeed (Poggianti et al. 2013a; Valentinuzzi et al. 2010). A similar environment effect would be present in the Trujillo et al. (2009) sample. Examples of the 2D fitting results for four candidates are presented in Fig. 5.2, where postage stamps of the galaxy images and the residuals are shown.

Fig. 5.3 shows the distributions of redshift, axis ratio, Sérsic index, effective radius, stellar mass and surface mass density for our sample of compact candidates. The median redshift of the sample is  $\sim 0.44$ , with an rms of  $\sim 0.18$ , larger than the median redshifts of the **FULL** and **ETGs C-Samples** ( $\sim 0.36$  and  $\sim 0.34$ , respectively). Compact candidates have a median r-band effective radius of  $\sim 1.2$  kpc (rms=0.31 kpc), a Sérsic index of  $\sim 4.3$  (rms=2.3), an axis ratio of 0.40 (rms=0.14) and median stellar mass of  $\sim 10.99$  dex (rms=0.09). On the contrary, for the **FULL C-Sample** the median size, Sérsic index, axis ratio and  $M_*$  are  $\sim 7.7$  kpc, 4.7, 0.73 and  $\sim 11.07$  dex, respectively. Finally, if the **ETGs C-Sample** is considered, the median size, Sérsic index, axis ratio and  $M_*$  are  $\sim 8.4$  kpc, 5.5, 0.75 and  $\sim 11.10$  dex, respectively. Thus, smaller sizes translate into shallower light profiles (smaller  $n$ ) when compacts are compared with normal galaxies (consistently with the galaxy merging framework in Hilz et al. 2013). Compact candidates also present lower axis ratios than normal

galaxies. These elongated shapes are common in massive compact galaxies, both at high- $z$  (van der Wel et al. 2011; Buitrago et al. 2013), and at lower redshifts (Trujillo et al. 2012, 2014).

In our compact sample, the average surface mass density within  $1 R_e$  is  $\sim 1.2 \times 10^{10} M_\odot/\text{kpc}^2$  (rms of  $\sim 5.6 \times 10^9 M_\odot/\text{kpc}^2$ ), about 2 orders of magnitude larger than the average surface mass density of the **C-Samples**. This median density is a bit larger than the range of values of  $7 - 15 \times 10^9 M_\odot/\text{kpc}^2$  found in 4 compact galaxies at  $z < 0.2$  by Trujillo et al. (2012), but the values are consistent within the scatter. Taking the only 3 galaxies from Szomoru et al. (2012), which satisfy the same criteria we are adopting in this chapter, we find a median density of  $\sim 1.1 \times 10^{10} M_\odot/\text{kpc}^2$ , which is fully consistent with our result.

The  $r$ -band  $R_e$  as a function of  $M_*$  is shown in Fig. 5.4. The compact candidates are plotted with a random subset of  $\sim 2500$  galaxies from the **FULL C-Sample**. As the compact candidates are selected to have small sizes, they lie in a region of the size-mass diagram where very few objects are detected, which provides a further evidence in favor of the rarity of these compact objects at  $z \lesssim 0.7$  (Fig. 5.4; Trujillo et al. 2009). In Fig. 5.4 we also overplot the  $z < 0.2$  galaxies in Trujillo et al. (2012) and the three  $z \sim 2$  galaxies in Szomoru et al. (2012) which fulfill our selection criteria. The evolution with redshift of the sizes, magnitudes and colors are shown in Fig. 5.5. The  $R_e$ s look quite constant as a function of redshift for the compact systems. At all the redshifts, almost all the compacts have faint **MAG\_AUTO** when compared with the control sample population (middle panel in Fig. 5.5). We find that 75 candidates ( $\sim 82\%$  of the whole sample) lie below the  $r$ -band magnitude completeness limit of  $\sim 21$ , implying that our sample is complete up to redshift  $z \sim 0.5$ . Finally, most of the galaxies populate the red-sequence, and are the best candidates to be remnants of high- $z$  red nuggets (right panel in Fig. 5.5).

58 out of 92 galaxies have old ages consistent with a formation redshift  $z_f \gtrsim 2$ , and so could be the remnants of the compact galaxies observed at  $z > 1$  (Gargiulo et al. 2011, 2012; Szomoru et al. 2012). 82 out of 92 galaxies (i.e.  $\sim 89\%$  of the whole sample) are best fitted by an elliptical template, and are classified as ETGs (see §5.2). The rest of the galaxies have redshifts  $z > 0.5$  and have colors which are best-fitted by Sbc models. However, particularly at these redshifts, one should take this color-based classification with caution, as a spectroscopic follow-up is actually needed to perform an accurate stellar population analysis.

### 5.3.2 Systematics from wrong redshifts

We have finally cross-matched our sample of candidates with SDSS and GAMA spectroscopy catalogs, finding spectra for 9 of the selected candidates.

One of the main systematics in our selection of compact galaxies is induced by wrong redshift determinations, which can affect both the (linear) effective radii and

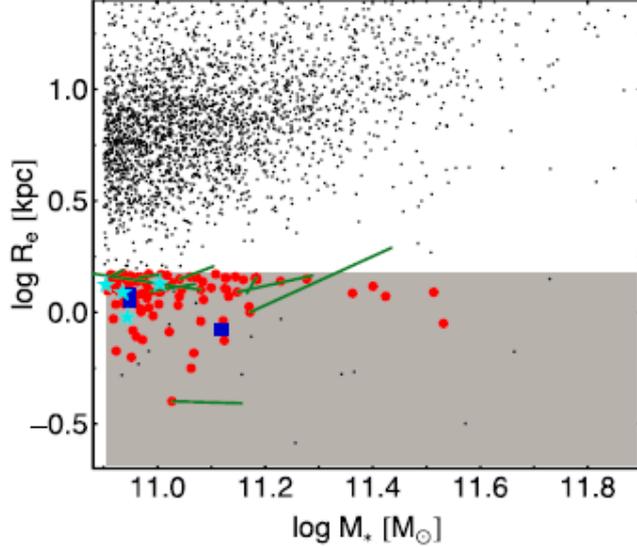


Figure 5.4: We show the median of the  $R_e$  in the g-, r- and i- bands vs. stellar mass. Compact candidates (plotted as red points) are compared with a selection of galaxies ( $\sim 2500$  randomly extracted) from the FULL C-Sample (small black points). Gray shaded region highlight the region where candidates are selected (see text for details about the adopted selection criteria). We also plot the change if the spectroscopic redshift is used (dark-green lines). Finally, the four compact galaxies with  $z < 0.2$  from Trujillo et al. (2012) and the three compacts at  $z \sim 2$  from Szomoru et al. (2012), that fulfil our selection criteria, are shown as cyan stars and blue squares, respectively.

stellar masses, moving the compact out of our selection criteria (e.g., see Figs. 5.4 and 5.5). 7 out of 9 of these systems are included in the SDSS+GAMA datasample used for photometric redshift determinations (see §5.2): 5 in the training sample and 2 in the test sample. Although the photometric redshifts are shown to approximate quite well the spectroscopic ones (see §5.2 and Cavuoti et al. 2015), we note that almost all of the photometric redshifts are underestimated with respect to the spectroscopic value. The median difference  $\Delta z \equiv z_{\text{phot}} - z_{\text{spec}}$  is  $-0.07$ , which increases to  $-0.1$  if the galaxies not in the training sample are considered. This systematics can be probably related to the under-sampling of this galaxy population, just 5 galaxies in the training sample, which can fail the optimum training of the network, and the quality of photometric redshift outcomes is degraded.

As first test, we have re-computed the sizes and stellar masses of these 9 objects using the  $z_{\text{spec}}$ . Only 4 galaxies survive to the size and mass selection criteria, i.e.  $\sim 44\%$  (1 not in the training sample).

Despite the small sample available for this test, we cannot exclude the presence of a systematics of the photometric redshift determination for compact sources which

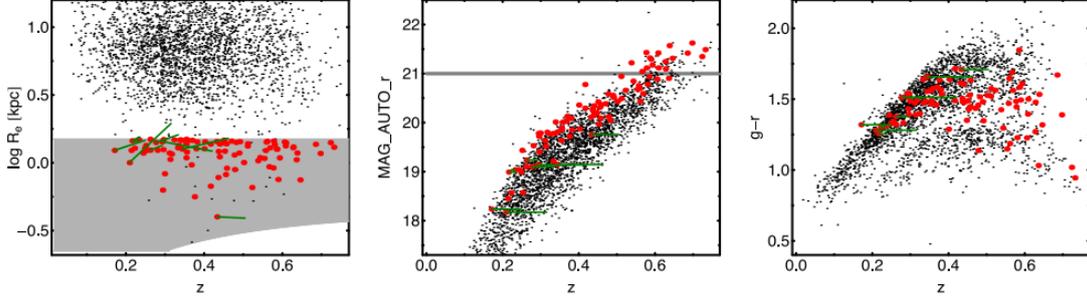


Figure 5.5: We show the median of the  $R_e$  in the  $g$ -,  $r$ - and  $i$ - bands (left),  $r$ -band  $\text{MAG\_AUTO}$  (middle) and  $g - r$  (right) vs. redshift. The gray region in the left panel sets the region within which we have searched the compacts. Gray line in the middle panel sets the 90% completeness limit of the high- $S/N$  sample. The  $g - r$  color is calculated within  $6''$  of diameter. See further details about the symbols in Fig. 5.4.

we might quantify of the order of  $\Delta z = -0.1$ . Qualitatively, if the spectroscopic redshift is larger than photometric one, then  $R_e$  in physical scale gets larger. The effect on  $M_*$  is not as simple as the one on  $R_e$ . As an exercise, we have investigated the impact of a redshift error of  $\Delta z = 0.1$  on the 83 galaxies in the sample without spectroscopic redshift. Using re-calculated values for  $R_e$  and stellar mass, we find that 26 of the sources (31%) survive the cuts.

Therefore, a spectroscopic follow-up will be needed to finally confirm the nature of these galaxies, allowing us to classify these candidates as compact or very-compact galaxies. Increasing the knowledge base and the population of compact systems with measured spectroscopic redshifts will improve the quality of the photometric redshift estimates.

### 5.3.3 Colour gradients

A big improvement of our analysis with respect to previous works on compact galaxies at similar redshift ranges (e.g. Damjanov et al. 2014) is the high- $S/N$  photometry which allows us to derive robust structural parameters and obtain the first determination of color profiles for these systems. Therefore, following Tortora et al. (2010), we define the color gradient as the angular coefficient of the relation  $X - Y$  vs.  $\log R$ ,  $\nabla_{X-Y} = \frac{\delta(X - Y)}{\delta \log R}$ , measured in mag/dex (omitted in the following unless needed for clarity). The fit of each color profile is performed in the range  $R_1 = 0.1R_e \leq R \leq R_2 = R_e$ , where the effective radius is measured in the  $r$ -band. By definition, a positive CG,  $\nabla_{X-Y} > 0$ , means that a galaxy is redder as  $R$  increases, while it is bluer outward for a negative gradient. PSF-convolved Sérsic profiles in  $g$ -,  $r$ - and  $i$ -band are used and, in particular,  $g - r$ ,  $r - i$  and  $g - i$  colors

are discussed. We omit detailed analysis in terms of redshift and stellar mass, but we pinpoint what is the range of color gradients in our galaxy sample, comparing these results with the ETGs C-Sample and some previous literature.

On average, compact population has  $\nabla_{g-r} = -0.21$  (rms 0.52),  $\nabla_{r-i} = -0.07$  (rms 0.59) and  $\nabla_{g-i} = -0.30$  (rms 0.73) which are substantially consistent with the gradients for the control population of ETGs which are  $\nabla_{g-r} = -0.17$  (rms 0.33),  $\nabla_{r-i} = -0.05$  (rms 0.23) and  $\nabla_{g-i} = -0.21$  (rms 0.40). Hence, compact candidates look quite similar to normal ETGs, except for the scatter, which is partially or totally related to the larger uncertainties on the structural parameters in our small objects.

These results agree with previously reported ranges of color gradients of passively evolving massive galaxies at low- or intermediate- $z$  (Tamura et al. 2000; Tamura & Ohta 2000; Wu et al. 2005; Tortora et al. 2010; La Barbera et al. 2012) or at high- $z$  (Guo et al. 2011b; Gargiulo et al. 2011, 2012) and with simulations (e.g., Tortora et al. 2013a).

Finally, we find that 30 compact candidates have all negative color gradients, 9 have positive gradients, while the rest of the sample have at least one of the three gradients with a different sign with respect to the others. This wide range of behaviors demonstrate that these objects can be formed in quite different initial conditions (Gargiulo et al. 2012; Damjanov et al. 2014).

### 5.3.4 Abundance vs. redshift

The number density of compact massive galaxies as a function of redshift is an important constraint on models of galaxy assembly. In recent years there have been different efforts to produce a census of such systems in different redshift bins (e.g. Trujillo et al. 2009; Taylor et al. 2010; Cassata et al. 2013; Poggianti et al. 2013c,a; Damjanov et al. 2014, 2015a, Saulder et al. 2015). Sample size is a crucial aspect to increase the constraining power. If the compact galaxies found in the present work are a representative subsample of the whole population of compacts over the whole area of 1500 sq. deg. that will be mapped by the KiDS survey, we expect to increase the present sample by a factor of ten in the next few years.

For what concerns our current sample, we have binned galaxies with respect to redshift and normalized to the comoving volume corresponding to the observed KiDS sky area<sup>1</sup>. The errors on number counts take into account fluctuations due to Poisson noise, as well as those due to large-scale structure (i.e. the cosmic variance): they are calculated with the online `CosmicVarianceCalculator`<sup>2</sup> tool (Trenti & Stiavelli

<sup>1</sup>This is obtained by multiplying the number of candidates by  $f_{\text{area}} = A_{\text{sky}}/A_{\text{survey}}$ , where  $A_{\text{sky}}$  (= 41253 sq. deg.) is the full sky area and  $A_{\text{survey}} = 105.4$  is the effective KiDS-DR2 area. Then, the density is derived by dividing by the comoving volume corresponding to each redshift bin.

<sup>2</sup><http://casa.colorado.edu/~trenti/CosmicVariance.html>

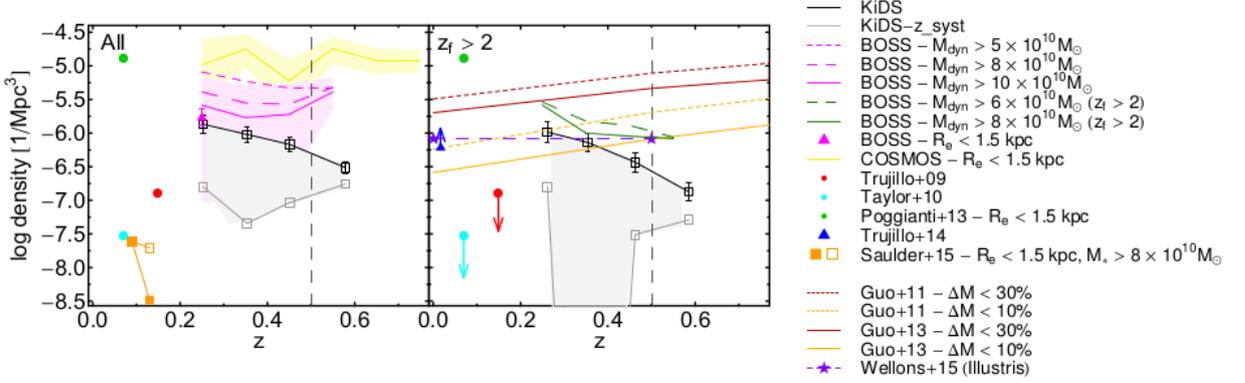


Figure 5.6: Number density of compact galaxies vs. redshift. Error bars denote  $1\sigma$  uncertainties (see the text). Vertical dashed black lines mark the completeness limit of the sample (see de Jong et al. 2015). *Left.* Number densities for all galaxies, independent of their photometric formation redshift, are plotted with open squares and solid black lines. The solid gray line with open squares takes into account possible systematics in the redshift determination (see §5.3.2). Short-dashed, long-dashed and solid violet lines are number densities of stellar-like objects from BOSS-DR10 (Damjanov et al. 2014) with  $M_{\text{dyn}} > 5 \times 10^{10}$ ,  $8 \times 10^{10}$ , and  $10^{11} M_{\odot}$ , respectively. The violet shaded region is the  $1\sigma$  error for the case with  $M_{\text{dyn}} > 10^{11} M_{\odot}$ . The violet triangle with bars also shows the abundance of galaxies at  $z \sim 0.25$ , with  $R_e < 1.5$  kpc and  $M_{\text{dyn}} > 8 \times 10^{10} M_{\odot}$ , from Damjanov et al. (2014). The yellow line with lighter yellow region plot abundances for compacts in the COSMOS area (Damjanov et al., 2015a), selected with the same criteria as in the present work ( $M_{\star} > 8 \times 10^{10} M_{\odot}$  and  $R_e < 1.5$  kpc; I. Damjanov, private communication). Red, cyan and green points are the results for compact galaxies from Trujillo et al. (2009), Taylor et al. (2010) and Poggianti et al. (2013a), respectively. Orange boxes show the abundances for compacts in SDSS area from Saulder et al. (2015), adopting our same criteria on mass and size. Filled boxes plot the results using Sérsic profiles, while open boxes are for the de Vaucouleurs profile (note that the results for the two profiles in the lowest redshift bin are superimposed). *Right.* Black open squares, solid lines and error bars plot KiDS number densities for candidate relic compacts, defined to have photometric formation redshift  $z_f > 2$ . Gray open squares and lines take into account possible systematics in the redshift determination. Long-dashed and solid green lines are for stellar-like objects from BOSS-DR10 with  $z_f > 2$  (Damjanov et al. 2014) and  $M_{\text{dyn}} > 6 \times 10^{10}$  and  $8 \times 10^{10}$ , respectively. The results from Trujillo et al. (2009) and Taylor et al. (2010) are shown here as upper limits (see red and cyan points with arrows). The blue triangle is for the lower limit at  $z \sim 0$  provided by Trujillo et al. (2014). Orange and red lines plot abundances of relic galaxies from semi-analytical models (SAMs), based on Millennium N-body simulations (Quilis & Trujillo 2013). Relics have been defined as systems whose stellar mass has increased since  $z = 2$  to the present by less than 10% and 30%, respectively. Dashed and solid lines are for Guo et al. (2011a) and Guo et al. (2013) SAMs, respectively. Purple stars (connected by a line) are predictions from the Illustris simulations (Wellons et al., 2016).

2008). The cosmic variance increases the Poissonian error budget by  $\sim 5 - 30\%$ . The total relative error on abundances (i.e. number densities) amounts to  $\sim 25 - 30\%$ .

In Fig. 5.6 we plot the redshift evolution of the abundance of compact galaxies (left panel) and that for the subsample of systems with old photometric ages (i.e. formation redshift  $z_f \geq 2$ ; right panel). We consider these potentially old MSCGs, as candidate remnants of the compact ETGs found at high redshift by several studies (see §5.1). We also re-determine the abundances by accounting for possible systematics in the photometric redshifts (see §5.3.2; see grey symbols in Fig. 5.6). We remind the reader that our sample starts to be incomplete at  $z \gtrsim 0.5$  (see vertical dashed lines).

In the left panel of Fig. 5.6, we plot number densities for the whole sample of KiDS compacts, independent of the galaxy formation redshift. Our number densities are compared with estimates from Damjanov et al. (2014), for the number density of stellar-like objects having spectroscopic redshifts from BOSS-DR10 (Ahn et al. 2014) and with three different cuts in total dynamical mass,  $M_{\text{dyn}}$  ( $> 5 \times 10^{10}$ ,  $8 \times 10^{10}$ , and  $10^{11} M_{\odot}$ , respectively; see Fig. 5.6). Notice that objects in the Damjanov et al. (2014) sample are not classified according to either morphology or galaxy age, nor do they have an accurate estimate of the intrinsic  $R_e$ . Hence, their selection might miss compact systems that are actually spatially resolved in BOSS-DR10. On the other hand, since Damjanov et al. (2014) also include compacts with  $R_e > 1.5$  kpc and are selected with respect to  $M_{\text{dyn}}$ , rather than  $M_{\star}$ , it is not surprising, perhaps, that those abundance estimates are larger than ours. In fact, if we consider the abundance estimate of massive BOSS targets with  $R_e < 1.5$  kpc and  $M_{\text{dyn}} > 8 \times 10^{10} M_{\odot}$  from Damjanov et al. (2014) (see purple triangle in the left panel of Fig. 5.6), the number density of compacts in the lowest redshift bin ( $z \sim 0.25$ ) is  $(1.7 \pm 0.5) \times 10^{-6} \text{ Mpc}^{-3}$ , in much better agreement with our density estimate of  $(1.6 \pm 0.4) \times 10^{-6} \text{ Mpc}^{-3}$  for the same redshift bin. On the other hand, selecting galaxies with  $R_e < 2.5$  kpc in KiDS, would lead to abundances  $\sim 0.8$  dex larger than for  $R_e < 1.5$  kpc, still in good agreement with estimates for stellar-like objects from Damjanov et al. (2014). The yellow region in the left panel of Fig. 5.6 plots number densities for galaxies in the COSMOS survey (Damjanov et al., 2015a)<sup>3</sup>. Remarkably, no evolution with redshift is found, for both KiDS and COSMOS samples, although, surprisingly, abundances for COSMOS (on average  $\sim 10^{-5} \text{ Mpc}^{-3}$ ) are about one order of magnitude larger than our KiDS estimates. Since Damjanov et al. (2014) claim to find consistent density estimates between COSMOS and BOSS (the latter having an area 4000 times larger than COSMOS), cosmic variance seems not to be responsible for the above discrepancy. However, we cannot exclude that galaxy environment, which might be the actual driver of the number density of compact relics at  $z \sim 0$  (see, e.g. Trujillo et al. 2014; Poggianti et al. 2013a; Valentinuzzi et al. 2010), may be

<sup>3</sup>These data are kindly computed for us by I. Damjanov (private communication) by applying the same size and mass selection criteria as in the present work.

different for galaxies in the KiDS-DR2 and COSMOS areas – an issue that will be addressed in forthcoming extensions of the present work.

The results for photometrically old MSCGs (right panel of Fig. 5.6) are first compared with estimates for old compact galaxies which have a formation redshift larger than 2 ( $z_f > 2$ ) compact galaxies from Damjanov et al. (2014), who selected samples with two cuts in  $M_{\text{dyn}}$  ( $> 6 \times 10^{10}$  and  $8 \times 10^{10}$ , respectively; see Fig. 5.6). As for the left panel, there are differences in the selection criteria that make a direct comparison not trivial. However, the trends for both BOSS and KiDS samples look qualitatively similar, with a very mild evolution with redshift in the range where our KiDS sample is complete. Moreover, as noticed above, for BOSS compacts with  $R_e < 1.5\text{kpc}$ , the Damjanov et al. (2014) abundance estimates are fairly consistent with ours, within the uncertainties.

Finally, we compare MSCGs number densities with predictions from semi-analytical models<sup>4</sup> (SAMs). Quilis & Trujillo (2013) have determined the evolution of the abundance of compact galaxies from SAMs based on Millennium N-body simulations (Guo et al. 2011a, 2013), where relic compacts are defined as galaxies which have barely increased their stellar mass between  $z \sim 2$  and  $z \sim 0$ . Operatively, they selected from the merger tree those objects that have increased their mass since  $z = 2$  by less than 10% and 30%, respectively, i.e. galaxies whose mass at  $z \sim 2$  is larger than 90% and 70% of the mass limit applied to select compacts. Our results are consistent in the lowest redshift bin with Guo et al. (2013) for simulated galaxies which have increased their mass at most by 10%, while we estimate lower densities than in simulations at higher redshift. However, similarly to what discussed for the comparison to BOSS estimates, theoretical predictions should be actually considered as upper limits, as Quilis & Trujillo (2013) did not apply any precise selection in size, since the resolution in the simulations does not allow reliable estimates of galaxy effective radii to be obtained. We also compare our findings to results from the hydrodynamical Illustris simulation of Wellons et al. (2016) (see also Wellons et al. 2015). Wellons et al. (2016) select 35 massive and compact galaxies at  $z = 2$  and follow their evolution to  $z = 0$ . Only 1 out of these 35 systems evolves into a galaxy that satisfies our mass and size criteria at  $z \leq 0.5$ . This corresponds to a number density of  $8.28 \times 10^{-7} \text{Mpc}^{-3}$ , which is consistent with the abundances of compact galaxies that have accreted less than 10% of their final mass from the Millennium simulations. As a further caveat here, we point out that in our selection we adopt the same mass cutoff value at all redshifts, while simulations perform the mass selection at  $z = 0$ , implying that at high redshifts, simulated galaxies with masses smaller than the  $z = 0$  cutoff value are actually included in the analysis. This is another reason why one may expect that predicted number densities for compacts at high redshift are actually higher than the observed ones.

---

<sup>4</sup>We caution the reader that stellar masses and sizes are measured in a different way between simulations and observations, hampering a straightforward comparison of the two.

At redshifts  $z \lesssim 0.2$ , in the left panel of Fig. 5.6, we see a lack of candidates. This seems to contrast the results of Trujillo et al. (2009) who found 29 secure **MSCGs** at  $z < 0.2$  fulfilling our same criteria, all of them having young ages  $\lesssim 4$  Gyr (see also Ferré-Mateu et al. 2012). However, one should notice that out of the 29 **MSCGs** of Trujillo et al. (2009), only one is at redshift  $< 0.1$ , still pointing to the extreme paucity of such systems in the nearby Universe, consistent with our result. Similarly, Taylor et al. (2010) found one possible old **MSCG** at low redshift, using a more relaxed criterion for the size, than the one we adopt here.

Saulder et al. (2015) have found a sample of 76 compact galaxies from SDSS at  $0.05 < z < 0.2$ , which resemble quiescent galaxies at high- $z$ , i.e. systems with small effective radii and large velocity dispersions. If we consider their compacts with  $R_e < 1.5$  kpc and  $M_* > 8 \times 10^{10} M_\odot$ , 1 galaxy at  $z < 0.1$  and 6 galaxies at  $z > 0.1$  are left when the effective radius from a de Vaucouleurs profile is used. Instead, these numbers change to 1 galaxy at  $z < 0.1$  and only 1 at  $z > 0.1$  if a Sérsic profile is fitted. These numbers correspond to abundances of  $2.4 \times 10^{-8} \text{ Mpc}^{-3}$  in the redshift range  $0.05 < z < 0.1$ , and  $2 \times 10^{-8} \text{ Mpc}^{-3}$  and  $3.3 \times 10^{-9} \text{ Mpc}^{-3}$  at  $0.1 < z < 0.2$ , if de Vaucouleurs or Sérsic profile are fitted, respectively. As mentioned in §5.1, these findings seem to trouble the current hierarchical paradigm of galaxy formation, where some relic systems at  $z \sim 0$  are actually expected to be found. In contrast, Poggianti et al. (2013a) have found 4 galaxies fulfilling our same criteria (corresponding to 1.4% of their sample galaxies with masses larger than  $8 \times 10^{10} M_\odot$ ), and all of these galaxies are old, with mass-weighted ages older than 8 Gyr. These numbers translate into a very large abundance of  $\sim 10^{-5} \text{ Mpc}^{-3}$ . Recently, Trujillo et al. (2014) added a new brick to the story, finding one relic compact in the Perseus cluster, i.e. NGC 1277, reconciling the observations at  $z \lesssim 0.2$  with predictions from simulations. All these results point to an overabundance of **MSCGs** in dense cluster regions, that are actually under-represented over the area currently mapped by KiDS.

As discussed in §5.3.2, the sample of **MSCGs** would be reduced significantly in size by systematics in the redshift determination. Fig. 5.6 “corrects” the abundance estimates for the possible systematics in redshift. The corrected abundances (see gray boxes and lines in Fig. 5.6) should be seen as our current lower limit on number densities of **SMCGs** in KiDS. This issue will be addressed by on-going spectroscopic follow-up of **MSCGs**, and discussed in details in a future paper.

## 5.4 Conclusions and future prospects

Thanks to the large area covered, high image quality, excellent spatial resolution and seeing, the Kilo Degree Survey (KiDS) provides a unique dataset to study the properties of super-compact massive galaxies (**MSCGs**) – a family of systems which plays a key role into our understanding of galaxy formation and evolution.

In this chapter, we present a sample of candidate **MSCGs**, based on 156 sq. deg. of KiDS, in four optical bands ( $u$ ,  $g$ ,  $r$  and  $i$ ). We start from a sample of  $\sim 0.4$  million galaxies with high- $S/N$ , measured photometry and structural parameters in all the four bands. For a subsample of these galaxies, we have used the KiDS photometry to estimate: 1) photometric redshifts based on machine learning techniques (Brescia et al. 2014; Cavauoti et al. 2015); 2) structural parameters using the software 2DPHOT (La Barbera et al. 2008); 3) stellar masses, fitting colors with SPS model predictions. The resulting sample is  $> 90\%$  complete down to an  $r$ -band magnitude  $\sim 21$ , and down to a stellar mass of  $3 - 5 \times 10^{10} M_{\odot}$ , up to a redshift  $z \sim 0.5$ . We select the most massive ( $M_{\star} > 8 \times 10^{10}$ ) and most compact ( $R_e < 1.5$  kpc) galaxies with (photometric) redshift  $z \lesssim 0.7$ . We remove star contaminants by performing a visual inspection of the final sample of candidates and then, for galaxies with available near-IR photometry from VIKING-DR1, we combine optical+NIR photometry to reduce the fraction of contaminants.

The final sample consists of 92 compact candidates, with a number density of  $\sim 0.9$  compact galaxies per square degree, at  $z < 0.7$ . Nine candidates have spectroscopic information from SDSS and GAMA surveys, that is used to assess the systematics in the redshift determination of our sample. On average, compact galaxies have negative color gradients which are similar to the ones for normal passively evolving galaxies (e.g., Tamura et al. 2000; Tamura & Ohta 2000; Tortora et al. 2010; Tortora et al. 2013a; La Barbera et al. 2012). However, the variety of gradients, to be confirmed with a spectroscopic follow-up of the present sample, seems to suggest that compact galaxies formed under a wide range of initial conditions (Gargiulo et al. 2012). We also discuss the evolution with redshift of the number density of compact systems, and in particular that for the oldest galaxies, which are possibly remnants of the high-redshift ( $z_f \gtrsim 2$ ) compact population detected by many studies. Remarkably, we do not find any **MSCGs** candidate at  $z \lesssim 0.2$ . This finding, which is consistent with the recent studies, might be related to the effect of galaxy environment on the abundance of compact systems (see Sec. 1). Although observational studies at intermediate- and high- $z$  do not point to a clear picture of how environment affects galaxy sizes (see Damjanov et al. 2015b and references therein), recent cosmological simulations predict a larger fraction of massive compact systems in high- than in low-density regions (Stringer et al. 2015). This prediction is supported by the fact that NGC 1277 – the only well-characterized compact, massive, ETG at  $z \sim 0$  – resides in the core region of a nearby rich cluster of galaxies (Trujillo et al. 2014; see also Valentinuzzi et al. 2010; Poggianti et al. 2013a). Thus, the absence of compact galaxies at  $z < 0.2$  could be related to the smaller fraction of dense structures in the area currently mapped by KiDS.

At the end of the KiDS survey we will collect  $\sim 1000$  SMCGs at  $z \lesssim 0.5$ , to be spectroscopically validated. This massive effort can be faced only using a multi-facility approach in the North and South hemisphere, to cover the whole KiDS

area during the entire solar year. To carefully set-up such a program we have carried out a pilot multi-object spectroscopy (MOS) at TNG to have a preliminary estimate of the purity of our sample and to best calibrate the telescope time to the science aims. Despite we have observed only half of the planned sample, due to bad weather conditions, the TNG results have been very encouraging, especially for the confirmation rate ( $\gtrsim 70\%$ ). Motivated by the TNG results, we have set a multi-site/instrument strategy, including NTT and VLT from the South and INT, WHT, TNG and GRANTECAN from the North. The multisite will allow to cover the two KiDS patches (the KiDS-North from La Palma and KiDS-South from Chile), while the multi-facility will allow to optimize the exposure time according to the target brightness (ranging from  $r \sim 18$  to  $r \sim 21$ ). Time for observing about 100 candidates at INT (6 nights), TNG (3 nights), NTT (3 nights) and GTC (28 hours) has been awarded in the last semesters. Data for about 80 objects are obtained from these observations and are under analysis. The determination of the spectroscopic redshifts of our candidates will put stringent constraints on the SMCGs cosmic density up to  $z \sim 0.5$  and will reduce the impact of the cosmic variance on their number counts. The study of the local environment of the compact galaxies is crucial. This could eventually explain the discrepancy observed in the literature.

# Chapter 6

## Strong lensing within KiDS

In Chapter 3, we have described the structural parameter extraction using 2DPHOT and shown the model subtracted (residual) images of some of the analyzed systems. We have seen that for the galaxies which were not perfectly fitted by a single Sérsic spheroidal component there was a variety of features in the residual images. The most spectacular cases were spiral arms, or even elongated streams from interacting systems. Among this variegated zoology of substructures, we have identified faint arc-like structures from strong galaxy-galaxy lensing events. We have then started a spin-off project aimed at systematically searching Strong Lensing (SL) features around KiDS galaxies. Based on the visual inspection of massive galaxies in the redshift range  $0.1 < z < 0.5$ , we have already collected a few tens of candidates. In this chapter we present some preliminary results based on the visual inspection of residual images and the future steps towards the spectroscopic follow-up for the studies and future applications in next generation ground based high-image quality surveys.

### 6.1 Introduction

According to General Relativity, a strong gravitational field manifests itself through a curvature of the space-time that can modify the path of a light ray. This process is commonly named Gravitational Lensing. It is working exactly as a normal lens, such as the ones in a magnifying glass or a pair of spectacles, bending light rays that pass through them in a process known as refraction, in order to focus the light somewhere (such as in your eye). Similarly, in the astrophysical context a large mass bends the path of background light. The gravitational field of a massive object cause light rays passing close to that object to be bent and refocused somewhere else. The more massive the object, the stronger its gravitational field and hence the greater the bending of light rays. This effect is manifested through spectacular astrophysical events, like multiple images of quasars, Einstein's rings, and arcs and

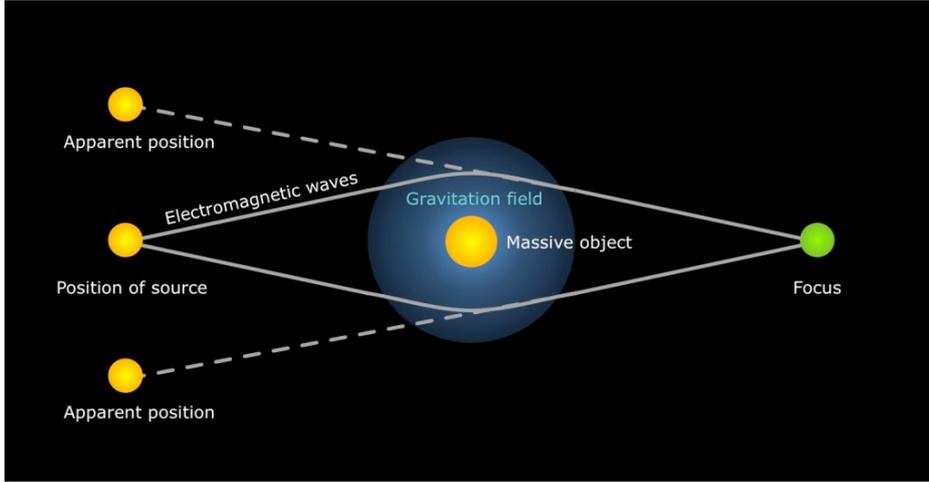


Figure 6.1: Illustration showing how gravitational lensing works

arclets in clusters of galaxies. The schematic of the gravitational lensing is given in Fig. 6.1.

The most common way to derive the total mass (stars and dark matter) of galaxies consists to measure rotational velocities or velocity dispersions from the spectral lines in measured spectra. In particular, elliptical galaxies are hot systems dominated by non-ordered motion, and measuring the velocity dispersion within an aperture will allow to constrain the total mass profile within that aperture. The study of the lensed images or arcs in galaxies provides an alternative tool, for mass determination, to this more traditional approach. However, using both lensing, dynamics and stellar mass from spectra or colors, we can provide a unique view on the stellar and dark matter distribution in galaxies at intermediate redshift (i.e.  $0.1 \lesssim z \lesssim 0.5$ ). Nowadays, about 600 galaxy-scale lenses are known, and surveys as KiDS or EUCLID will allow to enormously increase this number, These large numbers will allow to, e.g., greatly improve the precision in the mass density slope measurements (Barnabè et al., 2011), in better estimate the presence of substructures (Vegetti & Koopmans, 2009) and to put constraints on the nature of dark matter (Li et al., 2016). In this chapter, I will discuss a spin-off project of the structural parameter analysis, which I have started within this PhD thesis project, aimed at visually detecting new gravitational lenses in KiDS.

## 6.2 Gravitational lensing: the theory

In order to obtain a complete description of the phenomena linked to gravitational lensing, we need to study the path of the light rays in a curved space-time. We are interested in those cases in which the space-time curvature is “small” (so neglecting

the gravitational fields generated near neutron stars and black holes). A correct and complete derivation of the lens equation within the framework of the General Relativity can be found in Schneider et al. (1992). In this section we will use a more heuristic approach, starting a light ray propagation in a classical Newtonian space, and generalizing the results to the correct formulae from General Relativity.

To write the equation of *Lensing*, we use a simple calculation of geometrical optics. We consider a point mass which affect the trajectory of a light ray which passes close to it. Thus, let's consider a light ray moving in the direction  $z$  at a distance  $b$ , called *impact parameter*, from the point mass. The light ray experiments a Newtonian acceleration with an orthogonal component:

$$g_{\perp} = \frac{GMb}{(b^2 + z^2)^{\frac{3}{2}}} \quad (6.1)$$

if the deflection is small, then the orthogonal component of the total velocity is found integrating the formula  $v_{\perp} = \int g_{\perp} dt = \int g_{\perp} \frac{dz}{c} = \frac{2GM}{bc}$ , and the resultant deflection angle will be:

$$\alpha = \frac{v_{\perp}}{c} = \frac{2GM}{bc^2} \quad (6.2)$$

A more accurate calculation, however, requires General Relativity, resulting in a deflection angle which is exactly twice that of Equation (6.2), i.e., the correct equation for the deflection angle for a point mass is

$$\alpha = \frac{4GM}{bc^2} \quad (6.3)$$

Bearing in mind that gravitational lenses are extended mass distributions, it is possible to derive a similar equation for the deflection angle in a concrete case, starting from the Newtonian potential, which is defined as:

$$\psi(\vec{r}) = -G \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' \quad (6.4)$$

Since  $\vec{a} = -\nabla\psi$ , by integrating over the whole travel time and changing the variable of integration ( $dt \approx dz/c$ ), it is possible to obtain the trajectory of the light ray. As long as one does not concern oneself with large scale structures, it is well justified to take the thin lens approximation, that is to say, to make the ansatz, that all lateral acceleration happens "almost at once", before there is any appreciable deviation from the unlensed path. In the thin lens limit, one obtains:

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = 2 \int \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} \psi \frac{dz}{c} \quad (6.5)$$

where a factor of 2 is needed for our calculation to agree with General Relativity. It is therefore possible to define a *projected potential*  $\Psi(x, y) = \int \psi dz$  in terms of which to express our *deflection angle*  $(\alpha_x, \alpha_y) = \frac{(v_x, v_y)}{c}$  as

$$\vec{\alpha} = \frac{2}{c^2} \vec{\nabla} \Psi(x, y) \quad (6.6)$$

The *projected potential* satisfied a bi-dimensional Poisson equation

$$\vec{\nabla}^2 \Psi(x, y) = 4\pi G \Sigma(x, y) \quad (6.7)$$

where  $\Sigma(x, y) = \int \rho(x, y, z) dz$  is the projected mass density of the lens. This equation allows to measure the deflection angle of a light ray passing through the lens plane at a position  $(x, y)$ .

By applying the sine theorem, we can derive the ray-trace equation, which establishes a relation between deflection angle  $\alpha$ , apparent angle  $\theta$  and "true" source angle  $\beta$ , all measured with respect to the lens, provided that  $\alpha, \theta, \beta \ll 1$ :

$$\beta = \theta - \frac{D_{ds}}{D_s} \alpha(D_d \theta) \quad (6.8)$$

where  $D_{ds}$  is the distance between the source and the deflector body, while  $D_s$  is the distance between the source and the observer (Figure 6.2).  $D_{ds}$  and  $D_s$  are angular diameter distances, depending on the cosmological background (through the cosmological parameters) and the redshift of the lens  $z_d$  and of the source  $z_s$  (Schneider et al. 1992). We can distinguish three regimes in gravitational lensing:

- *strong lensing*, that consists in the formation of multiple (magnified) images of the source, and time delayed,
- *microlensing*, as a case of strong lensing, when the images are unresolvable, and, finally,
- *weak lensing*, characterized only by a deformation and a light amplification of the images without formation of different images.

We are particularly interested to the strong lensing (SL, hereafter). By solving the lens equation in the strong lensing regime, it is possible to show that, typically, gravitational lenses form 2 or 4 images, and assume different configurations on the sky. When the source is extended, then the images take the form of spectacular tangential or radial arcs. If galaxies act as lenses, then, point-like objects as quasars are doubly or quadruply imaged, while if the source is a galaxy, then, we observe tangential arcs. A peculiar lensing configuration is realized when lens and source are aligned, and a spectacular ring, called Einstein ring, is formed around the lens galaxy. Thus, measuring the position of these images or the radius of the arcs, it is possible to constrain the mass distribution of the lensing galaxy.

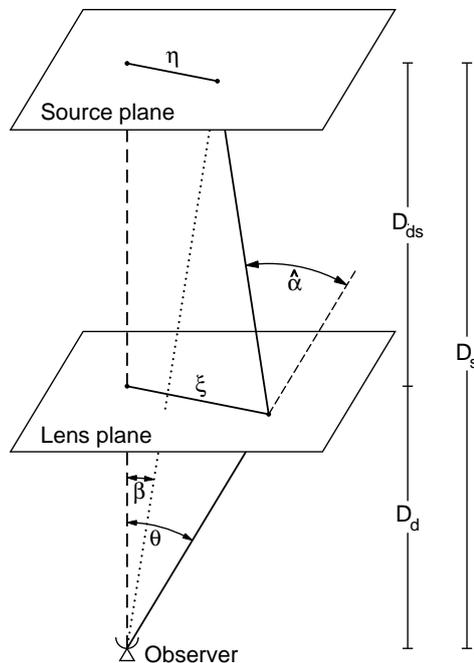


Figure 6.2: Lens event configuration. The lens and source planes and the main quantities useful to describe the event are reported.

### 6.3 Searching strong lenses in KiDS

Upcoming telescopes, such as Euclid (Laureijs et al., 2011) and the Large Synoptic Survey Telescope (LSST; LSST Science Collaboration et al. 2009), will increase the rate of discovery of new lenses, reaching the number of  $\sim 10^5$  new strong lensing systems (Oguri & Marshall, 2010; Pawase et al., 2012; Collett, 2015). Also, the number of lenses that will be observed by the Square Kilometer Array is expected to be of the same order of magnitude (McKean et al., 2015). The ongoing optical wide surveys, such as the Kilo Degree Survey (KiDS), the Dark Energy Survey (DES; The Dark Energy Survey Collaboration 2005) and the Subaru Hyper Suprime-Cam Survey (Miyazaki et al. 2012) are expected to find samples of lenses of the order of  $\sim 10^3$  (see, e.g, Collett 2015). Sub-mm observations from Herschel (Negrello et al., 2010) and the South Pole Telescope (Carlstrom et al. 2011), together with deeper, high resolution observations from the Atacama Large Millimeter/sub-millimeter Array, are expected to provide several hundred new lenses as well.

KiDS survey might work as precursor for developing scalable techniques for the Next generation telescopes (like LSST and Euclid). Thanks to the superb seeing, the pixel scales and the image quality of the VST (which is comparable to the one expected, e.g., for LSST, i.e.  $0.2''$  arcsec/pix angular scale and seeing  $\sim 0.6''$ ),

KiDS is particularly suitable for a systematic census of lensing galaxies based on the identification of arc-like structure from intermediate to massive galaxies (i.e., with velocity dispersion of the order of  $200 \text{ km s}^{-1}$  or larger), to galaxy groups and galaxy clusters, thus covering a wide range of lens masses.

If we assume that  $\beta = 0$  in Eq. 6.8, i.e. lens and source are aligned, then the equation is providing the radius of the forming Einstein ring, if a mass model is adopted and the distances fixed. If a spherical isothermal sphere, with density  $\rho = \sigma_v^2/(2\pi Gr^2)$ , is adopted, then the Einstein ring writes easily as a function of the model velocity dispersion,  $\sigma_v$ , as

$$R_E \sim 1'' \times (\sigma_v/220 \text{ km s}^{-1})^2. \quad (6.9)$$

For a typical  $FWHM \sim 0.7''$  in  $r$ -band, we can expect to observe lensing arcs of gravitational structures of  $\sigma_v > 180 \text{ km s}^{-1}$ .

The search for SL in KiDS is expected to provide up to thousands galaxy rings (e.g. by extrapolating CFHT-LS numbers, Oguri et al. 2014). This can be performed only with (semi-)automated techniques (e.g., Alard 2007; Gavazzi et al. 2014; Petrillo et al., in preparation), as potential candidates in KiDS can be of the order of thousands per square degree, making visual inspection inapplicable. However, lensing finding techniques still suffer from a number of issues yet to be solved, affecting the completeness and reliability of their catalogs and need to be trained over visual samples.

Thus, within KiDS this search is performed using two approaches:

- 1) the visual inspection of SL features (arcs or multiple-images);
- 2) the use of automatic tools for pattern recognition specialised to the SL features search.

This latter can be performed with (semi-)automated techniques (e.g. Gavazzi et al. 2014) as potential candidates in KiDS (depending on the adopted cuts) can be of the order of tens per square degree. Within our project we are also developing convolutional network techniques to automatically find arc-like features or potential multiple quasar images (Petrillo et al. 2017, in preparation). However, the visual inspection is still needed to double check the validity of the automatic selections (see Petrillo et al. 2017) and it remains the 0-order approach to initiate a systematic SL search in imaging surveys. The strategy adopted for KiDS take advantage from accurate model subtracted images where the contribution of the galaxy light has been modeled with 2DPHOT (La Barbera et al. 2008), and the signal-to-noise of the (even faint) arc-like and multiplet features in the residual image is highly increased.

## 6.4 Visual inspection procedure

Due to the yet limited extension of the area (150 sq. deg.) included in the KiDS ESO-Data Release 2 (ESO-DR2), we have decided to start a visual search based on simple criteria:

- 1) sources with spectroscopic redshifts between  $z = 0.1$  and  $z = 0.5$
- 2) sources brighter than  $r = 20$ ,

which together maximize the SL probability for the depth allowed by KiDS images. These criteria were restrictive and biased toward larger  $R_E$  easier to spot by visual inspection. In order to quantify the performance of our search, we decided to start on some control fields, overlapping with the SDSS, and selected  $\sim 600$  lens candidates in  $\sim 100$  sq.deg, including a ten of SLACS lenses.

The visual inspection was based on

- a) the galaxy  $r$ -band image ( $r$ -band is chosen as reference since it is the deepest and highest quality among KiDS wavebands),
- b) a three-color combined image to find a gradient in color between the candidate lens galaxy and the arc,
- c) the 2DPHOT model (La Barbera et al. 2008) subtracted image to detect residual arc features.

We have inspected about 2500 lens candidates and graded them with a 4 level rank from A to D, with A grading being very likely real candidates. A/B grades were found for about 2% of the candidates, i.e.  $\sim 50$  good candidates (27 A-grade and 23 B-grade). About 25 of those (mainly A) have brighter features and/or the highest contrast to be suitable for this follow-up. Note that with large photometric sample one can still attempt a DM analysis (e.g. Shu et al. 2015), but without the lens redshift this would provide only upper limits on the DM fraction with little constraints of the assembly on mass and DM with time.

Examples of lens candidates are in Fig. 6.3 and Fig. 6.4.

We have then moved to other galaxy subsamples and applied the same criteria to search for more candidates. We have selected about

- 600 candidates in the KiDS area overlapping with the GAMA survey (Driver et al. 2011), for which we had spectroscopical redshift and stellar masses (Liske et al. 2015);
- 1260 candidates among the massive ultra-compact galaxies discussed in Chapter 4, for which we had photometric redshifts and stellar masses from the KiDS photometry.

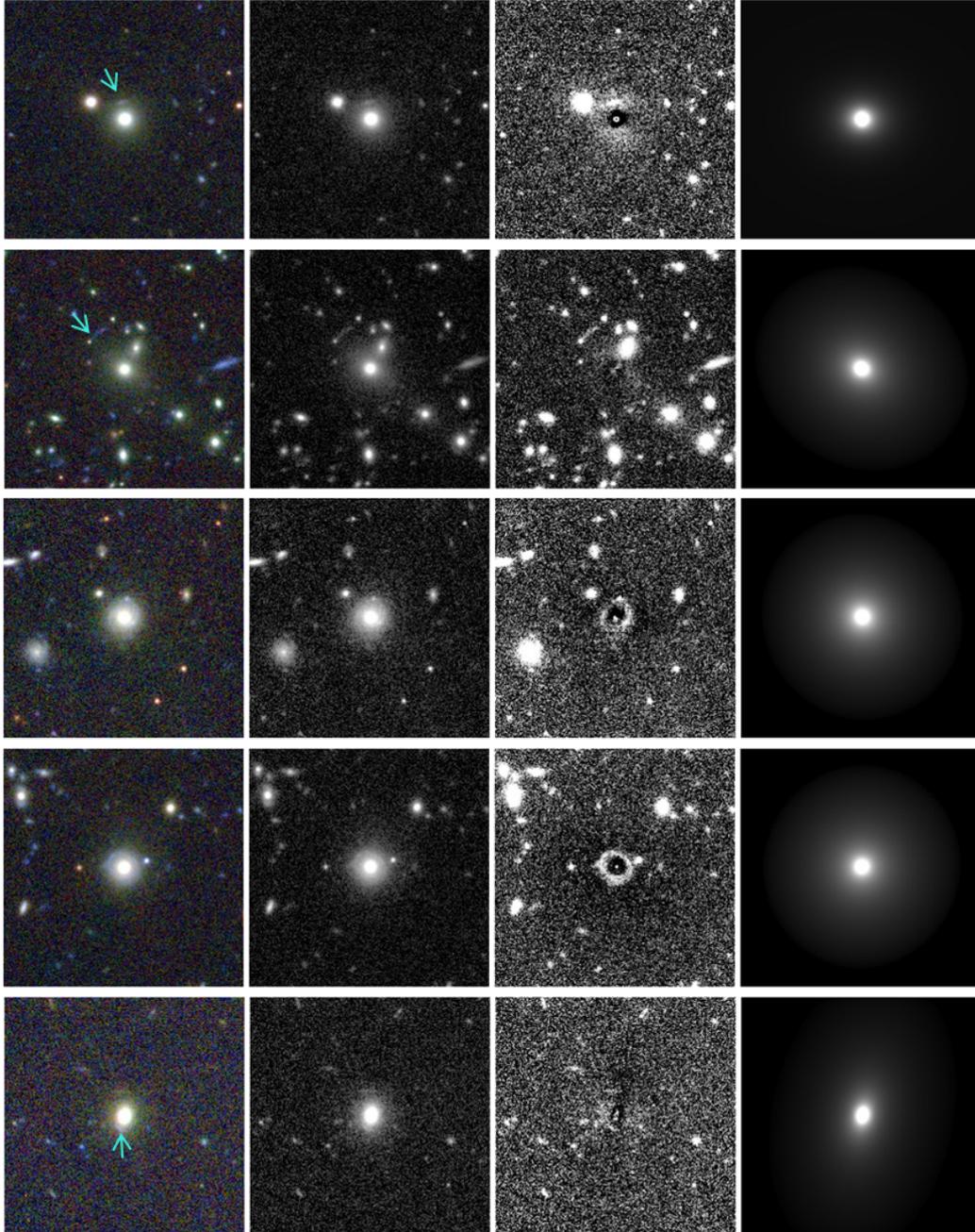


Figure 6.3: Results from the visual inspection of SL candidates. At each row we show: 1) RGB color image, 2)  $r$ -band image; 3) model subtracted image; 4) 2DPHOT galaxy model. Arrows show relevant features difficult to spot.

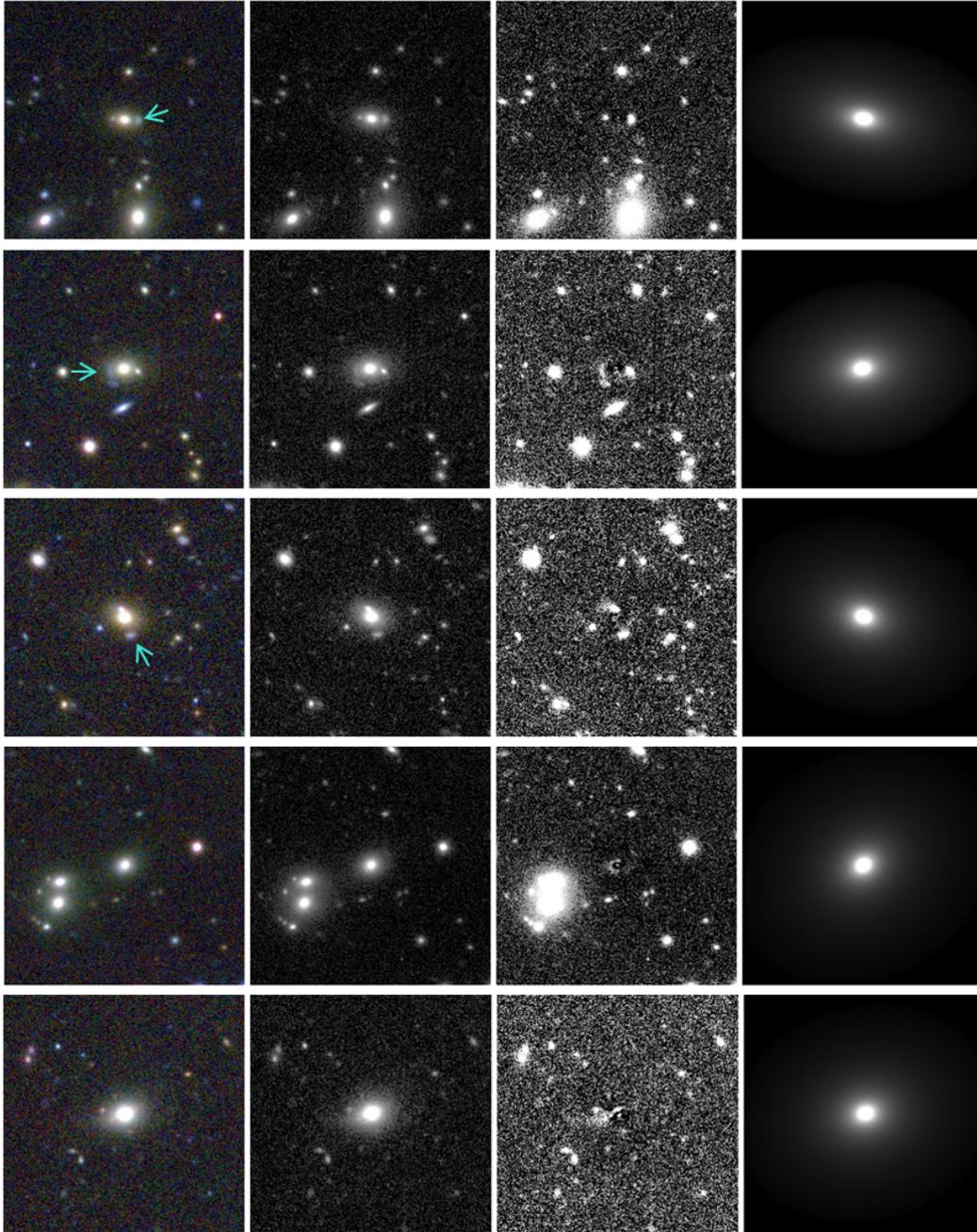


Figure 6.4: More examples from the visual inspection. Each panel is the same as Fig.6.3

From the GAMA sample we have found 10 A-grade candidates and 5 B-grade ones, while for the Compact galaxy sample 13 A-grade and 7 B-grade candidates. All together the totality of very likely candidates (A-grade) amount to 50 systems.

## 6.5 Spectroscopic follow-up

The next step to confirm the genuineness of the lensing nature of the candidates is to perform a spectroscopic follow-up. The measure of the redshift of the arc, which confirms that the detected arc is the image of a background source, is a further step to obtain a robust sample of lens galaxies.

We started applying for observing time with VIMOS@VLT. VIMOS has been used in its Integral field unit (IFU) configuration, in a medium resolution ( $R=720$ ) filter-free mode, a scale of  $0.67''$  fiber and a wavelength range from 490 to 1015nm. With the field-of-view of  $27'' \times 27''$  we can optimally integrate the arc-like emissions of the lensed galaxies and simultaneously acquire spectra of the central galaxy to measure the velocity dispersion, with  $<1$ h integration time per system. The main emission from the arcs we will look for in the spectra are produced by star-forming regions. We expect to observe the [OII] doublet emission ( $3726\text{-}3729\text{\AA}$ ) that we might dominate our  $r/i$ -band fluxes. Indeed, our lens candidates cover the range  $z \sim [0.1, 0.5]$  and the background lensed galaxies shall have a  $z_s$  predominantly in the range  $\sim [0.3, 1.7]$  (we have checked that  $> 60\%$  of the CASTLES sample with  $z_{\text{lens}} < 0.5$  have  $z_s < 1.7$ . We also stress that being selected in the optical band only we are confident that our visually selected sample is biased toward lower- $z_s$  and  $60\%$  is a very conservative fraction), which implies that the [OII] doublet should be observed at  $\lambda < 1000$  nm. Using the grism MR-708nm we can perform mid-resolution ( $R \sim 700$ ) spectroscopy over the useful wavelength range (i.e., 490-1015 nm) and obtain both the redshift of the lensed galaxy and an independent classification of this background galaxy together with the estimate of its star formation rate. Other useful lines would be [OIII] emission ( $5007\text{\AA}$ ) out to  $z \sim 1$ ;  $H_\beta$  ( $4861\text{\AA}$ ) out to  $z \sim 1.1$ ; CaHK lines out to  $z \sim 1.4$  (e.g. for Lick indexes) and gravity-sensitive features of the lens system to constrain the IMF slope (Spiniello et al. 2014).

The main aims of this project are: 1) to confirm the SL nature of the candidates; 2) to estimate the total projected mass within  $R_E$  to be used with 3) to measure the spatially resolved  $\sigma$  of the lens galaxy to derive the shape of DM profile, the DM fraction and stellar masses from stellar population synthesis (SPS) models; 4) to perform the SPS analysis on VIMOS spectroscopy to infer the IMF of the lens galaxy (dynamically: e.g. Tortora et al. 2013b; Tortora et al. 2014a; and using gravity-sensitive lines: Spiniello et al. 2014; La Barbera et al. 2013). The pilot sample that we propose to observe is complementary to literature SL catalogs as they cover a different parameter space, especially extending more in redshift in any

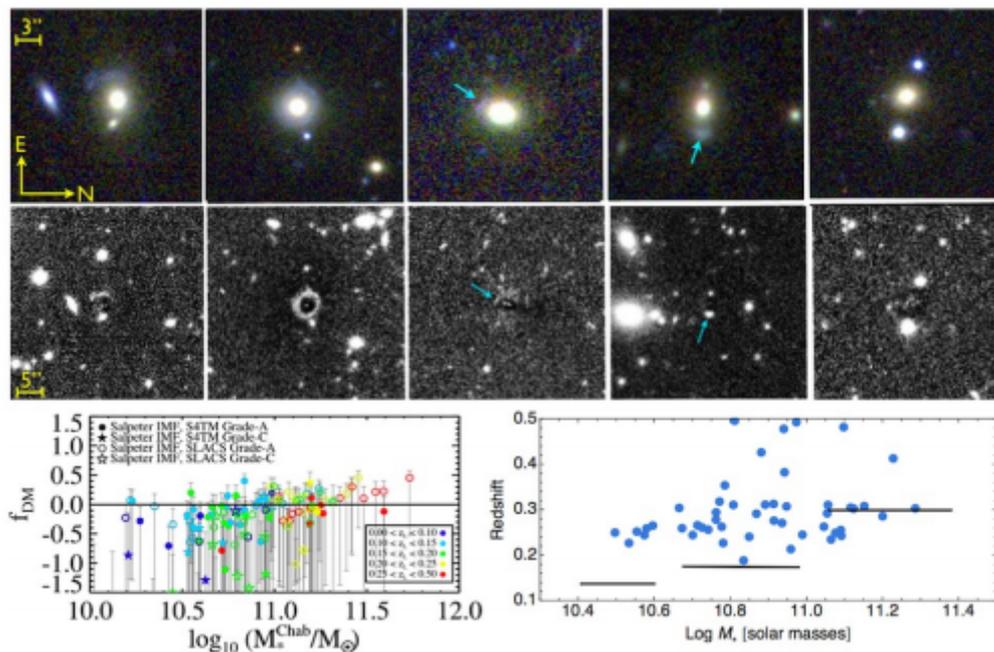


Figure 6.5: visual inspection of 5 SL candidates of different kinds (a large arc, an Einstein rings, a faint embedded arc, a tick arc and a 2-image quasar). Top: RGB color image; middle: model subtracted image. Cyan arrows show relevant features difficult to spot. Bottom: DM fractions from Shu et al. (2015) with a state-of art sample of SL candidates (left) and the distribution of the mass and redshift of the KiDS sample proposed here (of the full 50 sources, right). Solid lines indicate the SLACS median redshifts. The KiDS sample extends more in  $z$  at any given mass  $\log M_*/M_\odot < 11$  and it will allow the study of DM variation (or IMF variation) with  $z$  down to sub- $L_*$  systems.

given mass bin (See Fig. 6.5).

*This new SL sample will allow a significant leap in the study of DM fractions at intermediate- $z$ , almost doubling the same of confirmed SL at  $z > 0.1$  in the lower mass bins where one expect strong transition of the systems from stellar dominated cores to DM dominated ones or alternatively from low to high-IMF normalization in the scenario of the non-universal IMF.*

In Fig. 6.6, we show a very preliminary example of a confirmed strong lens in our sample. The system (ID KIDS 1748825) comes from the compact galaxy sample. The observations were carried out in service mode between October 2016 and January 2017 as part of the ongoing observational proposal 098.A-0620(A) "Spectroscopic follow-up of new strong lens candidates in the KiDS public survey" (PI N. Napolitano).

Observations have been taken in a single observing block made of 3 single scien-

tific exposures of  $\sim 683$ sec each, which led to a total exposure time of 43 minutes for the system. Seeing was  $< 0.9''$  and sky conditions were clear.

Fig. 6.6 shows the color image of the candidate and the corresponding field-of-view (FOV) of the VIMOS-IFU stacked observations, i.e. the image where all pixels contain the integrated flux of the spectrum obtained in the single spaxel (the pixel corresponding to the fiber gathering the spectra in that detector position). From the VIMOS-IFU FOV the central galaxy is clearly recognized (and marked by a green circle), while the arc shows up clearer in the snapshot of the data-cube at  $\lambda = 6462$  which is the wavelength of the peak emission from the lensed source which stick out from the galaxy continuum.

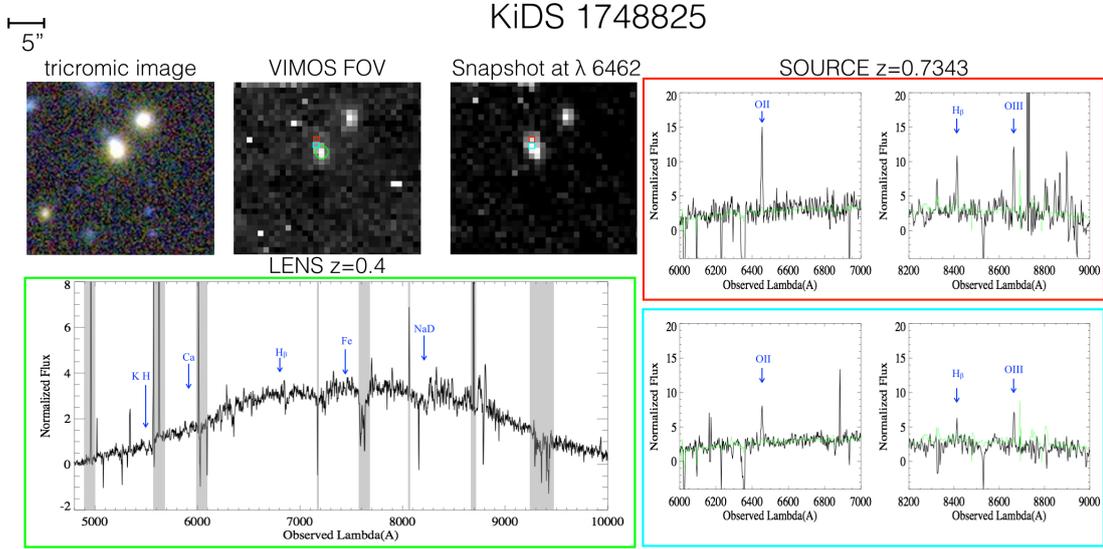


Figure 6.6: VIMOS-IFU confirmation of a compact galaxy selected in KiDS. Up left: color image, VIMOS field-of-view and a snapshot of the strongest emission feature found from the source at  $z = 0.73$ . The galaxy centre is marked with a green circle, while two spaxel from the source used to extract source spectra are marked with red and cyan circles. Bottom left: the spectrum from the central pixel of the galaxy where typical absorption lines are indicated by blue arrows of a early type galaxy at  $z = 0.4$ . Right column: two spectra obtained in two positions of the arc as marked by the red circle (corresponding to the spectrum on the top) and cyan circle (spectrum on the bottom) as in top left panel. In these spectra there are clear emissions duplicated in two different positions which indicate that the detection is genuine and correspond to a  $z = 0.73$  source.

We have then checked the presence of additional emission lines in the area expected to host the arc emission (e.g. the red and cyan circle in the FOV) and compared the spectra obtained in these spaxels with the central one of the galaxy. This is shown in the unidimensional spectra in the same Fig. 6.6, where we compare

face-to-face the spectra from the central pixel from the galaxy (green frame), and the stacked signal from 5 spaxels in the two different positions of the FOV that have to host the arc according to the position computed on the color image. These “arc” spectra are shown in the figure in red and cyan frames. In the galaxy spectrum we see typical absorption lines (indicated by blue arrows) of a early type galaxy at  $z = 0.4$ . In both “arc” spectra there are strong emissions (e.g. [OIII] and  $H\beta$ ) that are not present in the galaxy spectrum and that made us to estimate the redshift of the emitting source at  $z = 0.73$ , i.e. larger than the redshift of the foreground galaxy at  $z = 0.39$ , and fully compatible with a lensing configuration.

A larger statistical sample of confirmed lenses from KiDS it will be possible to produce accurate analyses of the dark matter content at different redshifts (see Fig. 6.5) and possibly break the “degeneracy” between DM fraction and IMF and trace the epoch where these transition start to occur (see e.g. Tortora et al. 2014b).

## **6.6 Strong Lensing as tool for systematic studies of the mass density of galaxies**

SL is an ideal tool to perform accurate studies of the DM distribution of intermediate and massive galaxies (e.g. with velocity dispersions,  $\sigma_v > 180 \text{ km s}^{-1}$ ) and the unique technique which provides with model-independent mass estimates within the arc features. SL studies, often in combination with stellar dynamics in the galaxy central regions, have found that the light and halo profiles conspire to have a total mass density profile which is nearly isothermal (Treu & Koopmans 2004; Koopmans et al. 2006; Bolton et al. 2006; Auger et al. 2010; Chae et al. 2014; Oguri et al. 2014), i.e. a total-mass density slope of  $\alpha \sim -2$ . This “conspiracy”, once thought to be universal, has been recently found not to be so, as lower mass galaxies might have steeper  $\alpha$  (Dutton & Treu 2014; Tortora et al. 2014a). The physical reasons for this systematic variation of the observed slope resides in the way baryons and DM are mixed, after having been assembled over time (Remus et al. 2013). Moreover, the combination of SL mass measurements and (3D) stellar dynamics allows us to put tight constraints on the mass profile of the lens galaxy and the measurement of the fraction of dark matter enclosed within a certain aperture become highly accurate. When also a precise measurement of the galaxy’s luminosity is added, lensing and dynamics can provide a measurement of the stellar mass-to-light ratio, constraining the Initial Mass Function (Treu et al. 2010).

Such detailed studies have been performed over samples of nearby early-type galaxies and a similar systematic study is ongoing at higher- $z$  SL systems (e.g., by SLACS, Bolton et al. 2008; SL2S, Ruff et al. 2011), where SL is much more effective of standard dynamical studies. For this reason it is becoming of fundamental importance to enlarge the sample of SL systems either in the galaxy parameter space

(e.g. toward less massive systems, like sub- $L_*$ ) and in redshift space. To enlarge the number of known lenses greatly not only will improve the precision in the mass density slope measurements (Barnabè et al., 2011), but it could also allow to provide better estimates of the amount of substructures (Vegetti & Koopmans, 2009).

# Chapter 7

## Conclusions and Future

KiDS is a large scale optical survey with a high quality data which can be used for the galaxy evolution studies. We made a galaxy sample with high  $S/N$  from the source catalogs. For the selected galaxies we extracted the structural parameters using 2DPHOT and we made the validation checks for the data and did the comparisons with external catalogs. Using the Sérsic index provided by 2DPHOT and the galaxy classification based on the SED fitting, we made a sample of passive spheroids and disc-dominated with redshift up to 0.5 for the galaxy evolution studies we are going to carry out.

We have demonstrated the large potential of the KiDS dataset for the structural parameter analysis of galaxies at least up to  $z = 0.6$ . We have analyzed a sample of  $\sim 380,000$  galaxies with signal-to-noise ratio large enough ( $SNR_r > 50$ ) to derive accurate structural parameters. We have showed that our estimates are robust against mock galaxies and external overlapping datasets. We have used in particular the size and stellar masses to investigate the evolution of the size-mass relation up to  $z \sim 0.6$  and compared the results with hydrodynamical simulations for galaxy assembly. The main results of our analysis can be summarized as follows:

- The size–mass–redshift show a very good agreement with the size–mass and the size–redshift correlations obtained either in local analyses (e.g. Shen et al. 2003, Baldry et al. 2012, Mosleh et al. 2013) or at higher- $z$  (e.g. Trujillo et al. 2007, van der Wel et al. 2014). The size–mass relation of passive spheroids shows a clear evolution of the average quantities with redshift which we have interpreted as a consequence of the size decreases with increasing redshift at masses larger than  $\text{Log}M_*/M_\odot \sim 10.5$ , while the evolution of the sizes for the disc-dominated is very weak, which produces no appreciable evolution of their size–mass relations;
- We have derived the  $R_e$  vs.  $z$  evolution using two approaches: 1) by fitting the size–mass relation at different redshift bins and then estimating the  $R_e - z$

evolution along different mass intercepts (see §4.1.1 and §4.1.2) and 2) by direct fitting the measured  $R_e$  vs.  $z$  in different mass bins. The results of the two methods consistently show a substantial evolution of sizes with redshift, with passive spheroids having a steeper decrease of their sizes with increasing redshifts with respect to disc-dominated galaxies. The normalization and slope of the  $R_e$  vs.  $z$ , parameterized using the standard  $R_e/kpc \propto (1+z)^{\beta z}$  (see Table 4.1), are consistent with recent analyses using accurate size measurement with single Sérsic profiles (van der Wel et al. 2014).

- We have compared the data with suites of recent hydrodynamical simulations of galaxy assembly with a full treatment of galaxy feedback (including supernovae and AGN feedback) and show that also in this case our results are nicely reproduced by the simulated galaxy trends. In particular, we have directly compared the expectations on the  $R_e - z$  relations in different mass bins for both the passive spheroid and disc-dominated systems. We have checked that the trends of the latter are strikingly in good agreement with the ones from simulation, while simulations for passive spheroids are slightly offset, although consistent with the data within the errors. We have also checked that simulations with no AGN feedback (e.g. from Oser et al. 2012) show a much larger discrepancy, showing that the an insufficient feedback recipe produce a tension with data, due to the too compact sizes in simulated galaxies.
- The two main classes of galaxies, disc-dominated and passive spheroids, show very different dependencies between size and stellar mass (Shen et al. 2003). The weak dependence between size and mass for disc-dominated galaxies implies that the high-mass disc-dominated, on average, have higher surface mass densities than low-mass disc-dominated. In contrast, passive spheroids show a more complex relationship between stellar mass and density, with the density peaking for systems with stellar masses around  $M_* = 4 \times 10^{10} M_\odot$  and decreasing toward both lower and higher masses, as reflected in the classical Kormendy (1977) relation. This fundamental difference does not depend on whether classification of passive spheroids and disc-dominated is based on star formation activity, bulge dominance (*Srsicindex*), or visual inspection, and it implies that the two types have very different evolutionary and assembly histories.
- The VST image quality provides the opportunity to search for massive compact galaxies, which can provide a useful benchmark for galaxy evolution models. Massive passive spheroids are found to be much more compact in the past with respect to local counterparts (Daddi et al. 2005; Trujillo et al. 2006; Trujillo et al. 2007; van der Wel et al. 2008). Over cosmic time, the high- $z$  compact galaxies are thought to evolve into the present-day massive, large, galaxies.

However, a fraction of these systems might survive intact having stellar populations with old ages (Trujillo et al. 2009; Taylor et al. 2010; Valentinuzzi et al. 2010; Poggianti et al. 2013a; Poggianti et al. 2013c; Damjanov et al. 2014; Tortora et al. 2016; Tortora et al. 2016). With KiDS we have started the census of the most compact galaxies at redshifts  $z \lesssim 0.7$ , with sizes  $< 1.5$  kpc and masses  $> 10.8$  dex (Tortora et al. 2016). The oldest among these systems are the relic of those compact galaxies at redshifts  $z \gtrsim 2$ , and can constrain the role of merging in the last billion years of the galaxy evolution history. We have selected 92 candidates after visual inspection and adding near-IR information for a subsample of these systems, planning to reach a number 10 more at the end of the survey. Only 9 out of 92 candidates have a spectroscopic counterpart, thus a spectroscopic follow-up is needed to validate the selection, determine the redshifts, more accurate stellar mass and age determination, and velocity dispersion to perform dynamical studies.

- KiDS is also suitable for a systematic census of lensing galaxies based on the identification of arc-like structure from intermediate to massive galaxies (i.e., with velocity dispersion of the order of  $200 \text{ km s}^{-1}$  or larger), to galaxy groups and galaxy clusters, thus covering a wide range of lens masses. We have started a visual search based on simple criteria: 1) sources with SDSS spectroscopic redshifts between  $z = 0.1$  and  $z = 0.5$  and 2) sources brighter than  $r = 20$ , which together maximize the SL probability for the depth allowed by KiDS images. These criteria were restrictive and biased toward larger Einstein radii, easier to spot by visual inspection. In order to quantify the performance of our search, we decided to start on some control fields, overlapping with the SDSS, and selected  $\sim 600$  lens candidates in  $\sim 100$  sq.deg, including a ten of SLACS lenses. The visual inspection was performed by six observers and based on a) the galaxy  $r$ -band image ( $r$ -band is chosen as reference since it is the deepest and highest quality among KiDS wavebands), b) a three-color combined image to find a gradient in color between the candidate lens galaxy and the arc, c) the 2DPHOT model subtracted image to detect residual arc features. We ended up with a list of 18 new potential lens candidates, half of which have high significance (de Jong et al. 2015; Napolitano et al. 2016). The spectroscopic follow-up will provide us with the confirmation of the lensing nature of the lens candidates and the independent kinematic information to constrain the mass profile of the lensing galaxies.

The large sample expected from the KiDS and the image quality will allow us an unprecedented details in the evolution of the galaxy size and mass over the cosmic time to compare with expectation from simulations. We expect to expand considerably the analysis presented in this thesis with the next KiDS data releases,

both in terms of size and depth of the sample, as we will gather statistics toward higher redshift to confirm our trends.



# APPENDIX

## Appendix A. Example output plots from 2DPHOT for galaxies with $\chi^2 < 1.3$

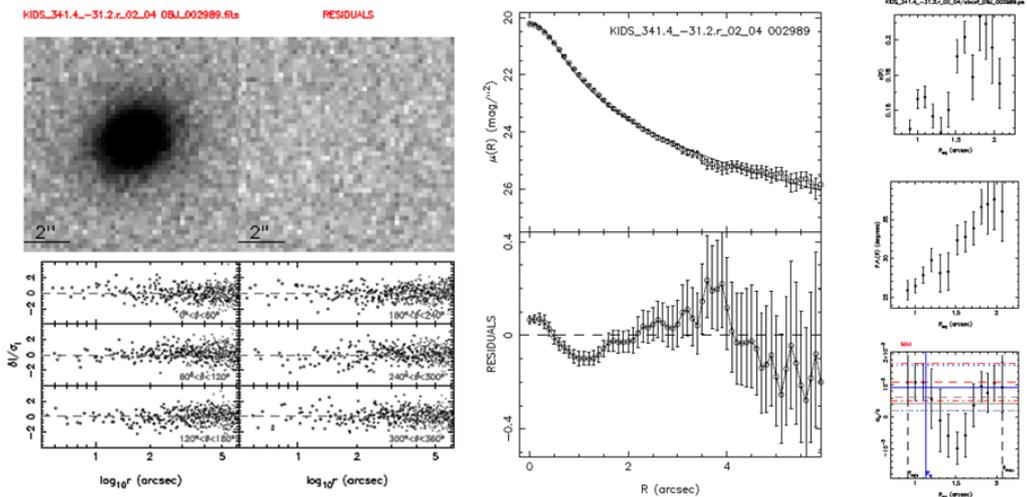
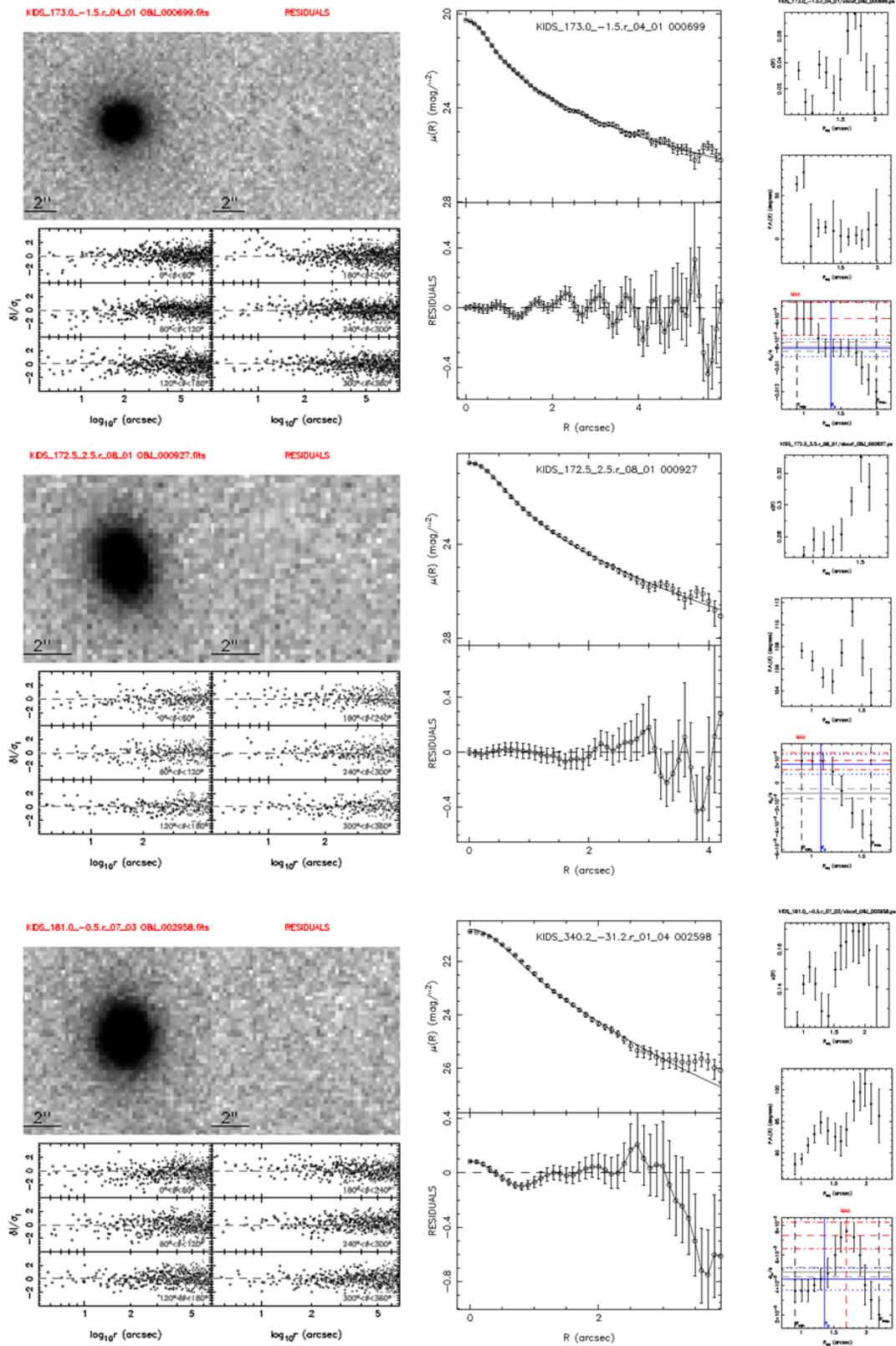


Figure 1: Left panels :2DPHOT output plot of a modeled galaxy. Top left plot is the source and the top right is the model subtracted residual map. Residuals after subtraction are shown in the lower plots as a function of the distance to the galaxy center, with each panel corresponding to a different bin of the polar angle. Middle panel :Surface brightness profiles of the same galaxy in the left panel. The surface brightnesses computed over different elliptical contours are plotted as a function of the ellipses of equivalent radii, in the upper plot. The error bars denote  $1\sigma$  uncertainties, computed by adding in quadrature the standard deviation of the intensity values in each ellipse with the uncertainty in the background estimate. The solid line is the best-fitting one-dimensional Sérsic model. The lower plot of each panel shows the residuals, in units of  $\text{mag arcsec}^{-2}$ , obtained after subtracting the model from the data. Right Panels :Radial profiles of isophotal parameters, as derived by the isophotal fitting algorithm. From top to bottom, the panels show the profiles of ellipticity, position angle of the ellipse's major axis, and  $a_4$  coefficient as a function of the equivalent radius of the fitted isophotes. Error bars mark  $1\sigma$  standard uncertainties. In the bottom panels, the minimum and maximum radii to define the global  $a_4$  value are shown as vertical dashed lines. The effective radius of the galaxy is marked by a vertical solid line. The solid horizontal line denotes the mean value of  $a_4$  in the selected radial range, while the dashed gray lines mark the corresponding  $1\sigma$  interval.



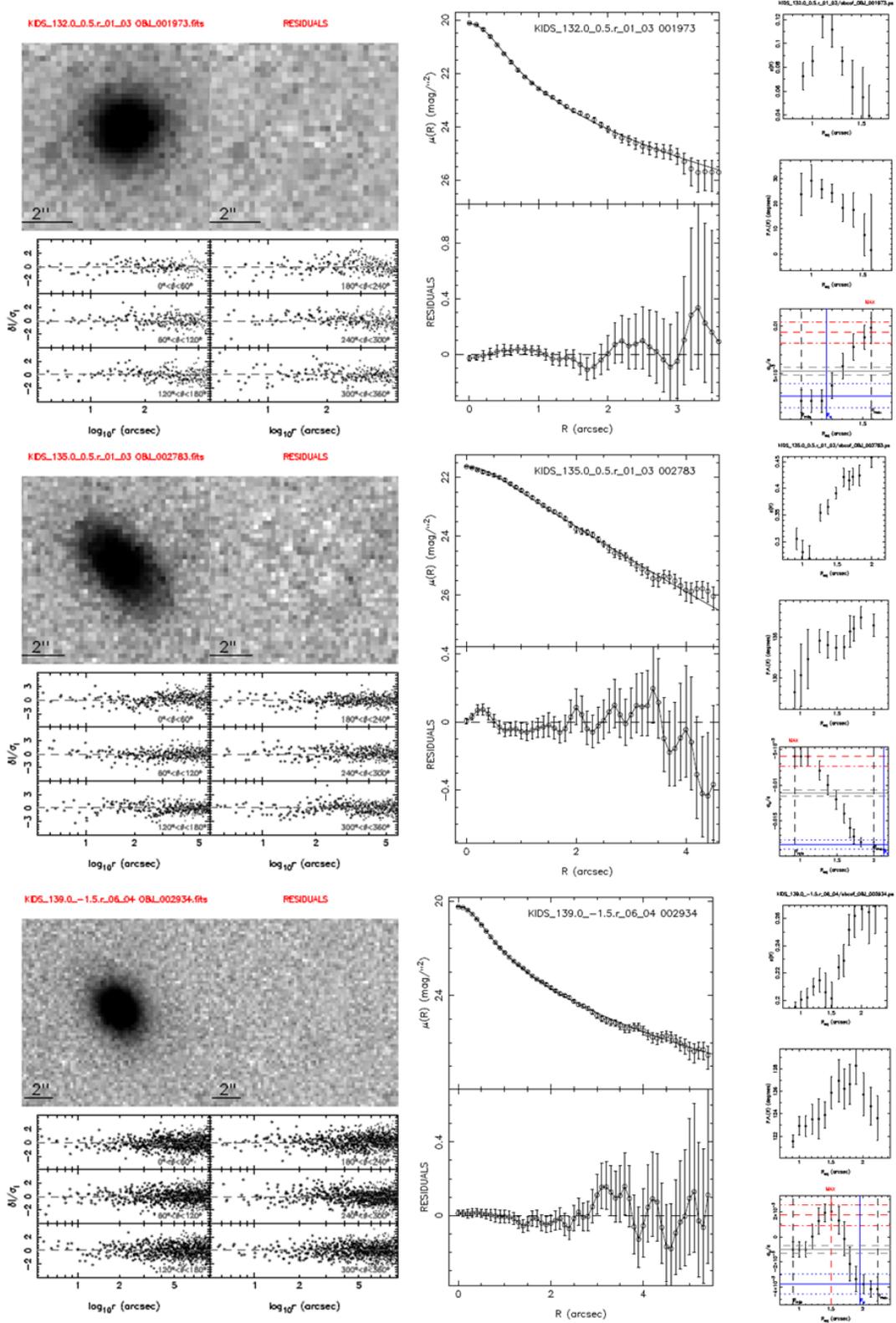


Figure 3: More examples as Fig.1

## Appendix B. Example output plots from 2DPHOT for galaxies with $\chi^2 > 1.3$

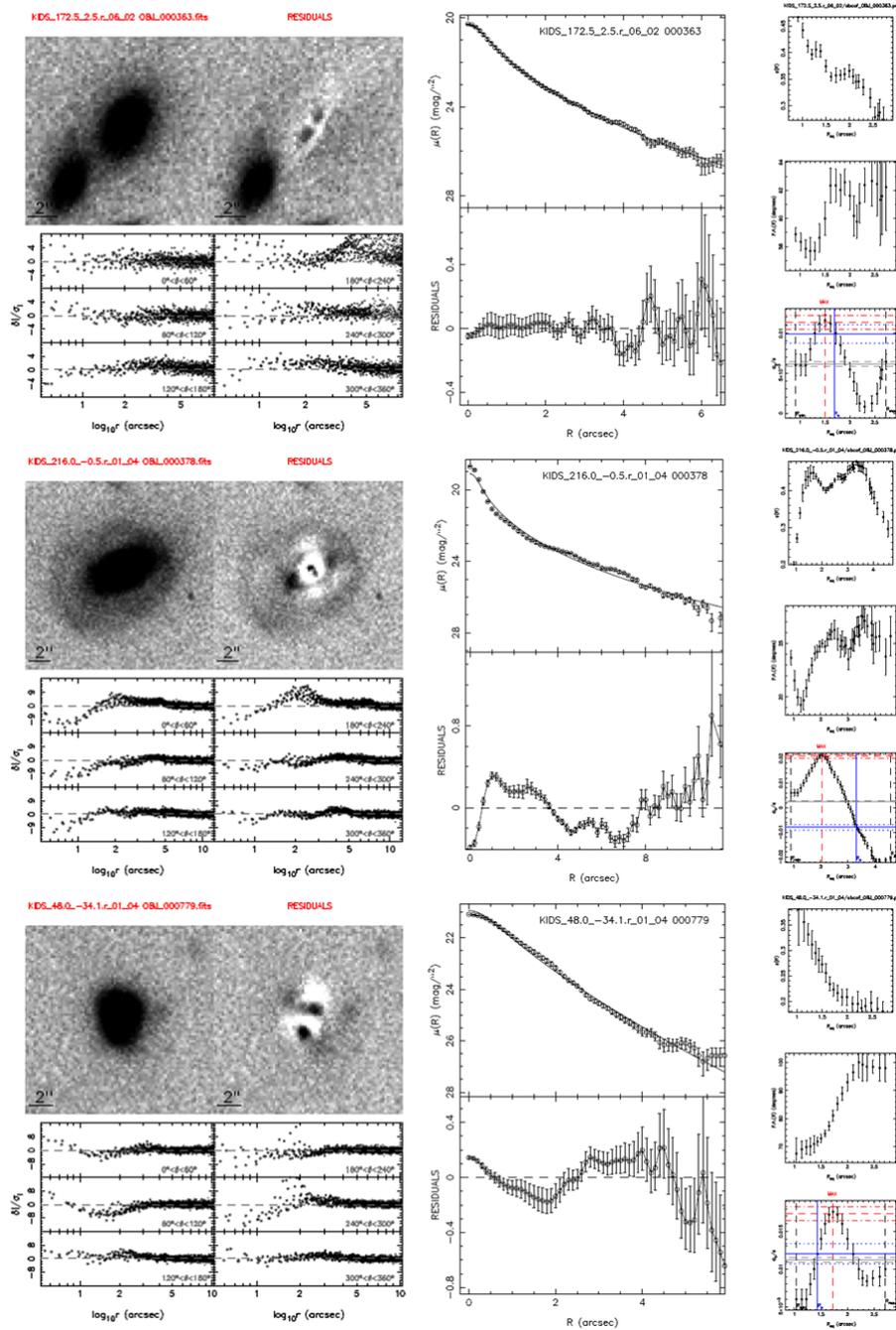
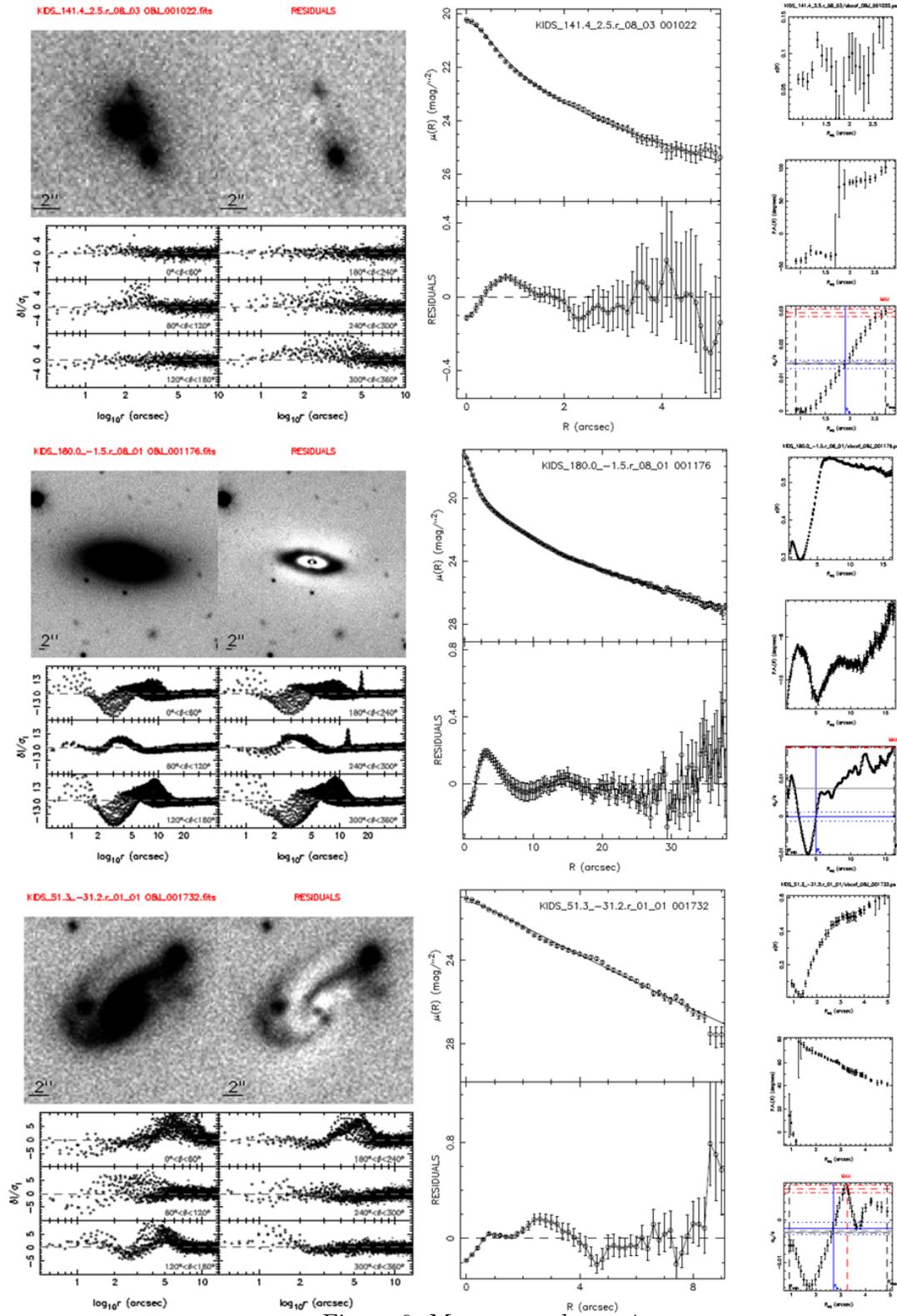


Figure 4: Example output from 2DPHOT for galaxies with  $\chi^2 > 1.3$ . Each panel is same as explained in 1





---

# Bibliography

Ahn C. P. et al., 2014, ApJS, 211, 17

Ahn C. P. et al., 2012, ApJS, 203, 21

Alam S. et al., 2015, ApJS, 219, 12

Alard C., 2007, MNRAS, 382, L58

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540

Arnouts S. et al., 2001, A&A, 379, 740

Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010, ApJ, 724, 511

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469

Baldry I. K. et al., 2012, MNRAS, 421, 621

Balogh M. L., Baldry I. K., Nichol R., Miller C., Bower R., Glazebrook K., 2004, ApJ, 615, L101

Barnabè M., Czoske O., Koopmans L. V. E., Treu T., Bolton A. S., 2011, MNRAS, 415, 2215

Baum W. A., 1959, PASP, 71, 106

Belli S., Newman A. B., Ellis R. S., 2014, ApJ, 783, 117

Bennett C. L. et al., 2003, ApJS, 148, 1

Bernardi M. et al., 2003, AJ, 125, 1882

- 
- Bertin E., 2006, in *Astronomical Society of the Pacific Conference Series*, Vol. 351, *Astronomical Data Analysis Software and Systems XV*, Gabriel C., Arviset C., Ponz D., Enrique S., eds., p. 112
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Bezanson R., van Dokkum P. G., Tal T., Marchesini D., Kriek M., Franx M., Coppi P., 2009, *ApJ*, 697, 1290
- Biviano A. et al., 2013, *A&A*, 558, A1
- Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, *Nature*, 311, 517
- Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Gavazzi R., Moustakas L. A., Wayth R., Schlegel D. J., 2008, *ApJ*, 682, 964
- Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, *ApJ*, 638, 703
- Bower R. G., Lucey J. R., Ellis R. S., 1992, *MNRAS*, 254, 601
- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150
- Brescia M., Cavuoti S., D'Abrusco R., Longo G., Mercurio A., 2013, *ApJ*, 772, 140
- Brescia M., Cavuoti S., Longo G., De Stefano V., 2014, *A&A*, 568, A126
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Buitrago F., Trujillo I., Conselice C. J., Häußler B., 2013, *MNRAS*, 428, 1460
- Caon N., Capaccioli M., D'Onofrio M., 1993, *MNRAS*, 265, 1013
- Capaccioli M., Schipani P., 2011, *The Messenger*, 146, 2
- Capak P. et al., 2004, *AJ*, 127, 180
- Carlstrom J. E. et al., 2011, *PASP*, 123, 568
- Cassata P. et al., 2010, *ApJ*, 714, L79
- Cassata P. et al., 2013, *ApJ*, 775, 106
- Cavuoti S., Brescia M., De Stefano V., Longo G., 2016, *ArXiv e-prints*
- Cavuoti S., Brescia M., Longo G., Mercurio A., 2012, *A&A*, 546, A13

- Cavuoti S. et al., 2015, MNRAS, 452, 3100
- Chabrier G., 2001, ApJ, 554, 1274
- Chae K.-H., Bernardi M., Kravtsov A. V., 2014, MNRAS, 437, 3670
- Cole S., 1991, ApJ, 367, 45
- Coleman G. D., Wu C.-C., Weedman D. W., 1980, ApJS, 43, 393
- Colless M. et al., 2001a, MNRAS, 328, 1039
- Colless M., Saglia R. P., Burstein D., Davies R. L., McMahan R. K., Wegner G., 2001b, MNRAS, 321, 277
- Collett T. E., 2015, ApJ, 811, 20
- Crain R. A. et al., 2015, MNRAS, 450, 1937
- D'Abrusco R., Staiano A., Longo G., Brescia M., Paolillo M., De Filippis E., Tagli-  
ferri R., 2007, ApJ, 663, 752
- Daddi E. et al., 2005, ApJ, 626, 680
- Damjanov I., Chilingarian I., Hwang H. S., Geller M. J., 2013, ApJ, 775, L48
- Damjanov I., Geller M. J., Zahid H. J., Hwang H. S., 2015a, ApJ, 806, 158
- Damjanov I., Hwang H. S., Geller M. J., Chilingarian I., 2014, ApJ, 793, 39
- Damjanov I., Zahid H. J., Geller M. J., Hwang H. S., 2015b, ArXiv e-prints
- Davidon W., 1991, SIAM Journal on Optimization, 1, 1
- Davis M. et al., 2003, in Proc. SPIE, Vol. 4834, Discoveries and Research Prospects  
from 6- to 10-Meter-Class Telescopes II, Guhathakurta P., ed., pp. 161–172
- de Bernardis P. et al., 2001, in American Institute of Physics Conference Series, Vol.  
555, Cosmology and Particle Physics, Durrer R., Garcia-Bellido J., Shaposhnikov  
M., eds., pp. 85–94
- de Jong J. T. A. et al., 2015, A&A, 582, A62
- de Jong J. T. A. et al., 2017, ArXiv e-prints
- de Jong R. S., 2008, MNRAS, 388, 1521

---

De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499

de Vaucouleurs G., 1948, *Annales d'Astrophysique*, 11, 247

de Vaucouleurs G., 1961, *ApJS*, 5, 233

de Zeeuw T., 2001, in *Black Holes in Binaries and Galactic Nuclei*, Kaper L., Heuvel E. P. J. V. D., Woudt P. A., eds., p. 78

Dekel A., Burkert A., 2014, MNRAS, 438, 1870

Djorgovski S., Davis M., 1987, *ApJ*, 313, 59

D'Onofrio M. et al., 2014, *A&A*, 572, A87

D'Onofrio M., Cariddi S., Chiosi C., Chiosi E., Marziani P., 2016, ArXiv e-prints

D'Onofrio M. et al., 2008, *ApJ*, 685, 875

Dressler A., 1980, *ApJ*, 236, 351

Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, *ApJ*, 313, 42

Driver S. P. et al., 2011, MNRAS, 413, 971

Dutton A. A. et al., 2011, MNRAS, 416, 322

Dutton A. A., Treu T., 2014, MNRAS, 438, 3594

Edge A., Sutherland W., Kuijken K., Driver S., McMahon R., Eales S., Emerson J. P., 2013, *The Messenger*, 154, 32

Edge A., Sutherland W., The Viking Team, 2014, *VizieR Online Data Catalog*, 2329, 0

Faber S. M., Jackson R. E., 1976, *ApJ*, 204, 668

Fan L., Lapi A., Bressan A., Bernardi M., De Zotti G., Danese L., 2010, *ApJ*, 718, 1460

Fan L., Lapi A., De Zotti G., Danese L., 2008, *ApJ*, 689, L101

Fasano G. et al., 2006, *A&A*, 445, 805

Ferrarese L., Merritt D., 2000, *ApJ*, 539, L9

- Ferré-Mateu A., Mezcua M., Trujillo I., Balcells M., van den Bosch R. C. E., 2015, *ApJ*, 808, 79
- Ferré-Mateu A., Vazdekis A., Trujillo I., Sánchez-Blázquez P., Ricciardelli E., de la Rosa I. G., 2012, *MNRAS*, 423, 632
- Ferreras I., Charlot S., Silk J., 1999, *ApJ*, 521, 81
- Fontanot F., Monaco P., Silva L., Grazian A., 2007, *MNRAS*, 382, 903
- Franx M., van Dokkum P. G., Förster Schreiber N. M., Wuyts S., Labbé I., Toft S., 2008, *ApJ*, 688, 770
- Furlong M. et al., 2015, *MNRAS*, 450, 4486
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., 2006, *MNRAS*, 370, 1106
- Gargiulo A., Saracco P., Longhetti M., 2011, *MNRAS*, 412, 1804
- Gargiulo A., Saracco P., Longhetti M., La Barbera F., Tamburri S., 2012, *MNRAS*, 425, 2698
- Garilli B., Maccagni D., Andreon S., 1999, *A&A*, 342, 408
- Gavazzi R., Marshall P. J., Treu T., Sonnenfeld A., 2014, *ApJ*, 785, 144
- Gebhardt K. et al., 2000, *ApJ*, 539, L13
- Geisser S., 1975, *J. Amer. Statist. Assoc.*, 70, 320
- Genzel R. et al., 2008, *ApJ*, 687, 59
- Governato F. et al., 2004, *ApJ*, 607, 688
- Gullieuszik M. et al., 2015, *A&A*, 581, A41
- Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, *MNRAS*, 428, 1351
- Guo Q. et al., 2011a, *MNRAS*, 413, 101
- Guo Q., White S. D. M., 2008, *MNRAS*, 384, 2
- Guo Y. et al., 2011b, *ApJ*, 735, 18
- Guzman R., Lucey J. R., Carter D., Terlevich R. J., 1992, *MNRAS*, 257, 187

---

Haas M. R., Schaye J., Booth C. M., Dalla Vecchia C., Springel V., Theuns T., Wiersma R. P. C., 2013, MNRAS, 435, 2931

Hilz M., Naab T., Ostriker J. P., 2013, MNRAS, 429, 2924

Hopkins P. F., Bundy K., Hernquist L., Wuyts S., Cox T. J., 2010a, MNRAS, 401, 1099

Hopkins P. F. et al., 2010b, ApJ, 724, 915

Hopkins P. F., Hernquist L., Cox T. J., Keres D., Wuyts S., 2009a, ApJ, 691, 1424

Hopkins P. F., Lauer T. R., Cox T. J., Hernquist L., Kormendy J., 2009b, ApJS, 181, 486

Hsu L.-Y., Stockton A., Shih H.-Y., 2014, ApJ, 796, 92

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huertas-Company M. et al., 2013, MNRAS, 428, 1715

Hyde J. B., Bernardi M., 2009, MNRAS, 396, 1171

Ilbert O. et al., 2006, A&A, 457, 841

Iodice E. et al., 2016, ApJ, 820, 42

Kashikawa N. et al., 2004, PASJ, 56, 1011

Katz N., Gunn J. E., 1991, ApJ, 377, 365

Kauffmann G., 1996, MNRAS, 281, 487

Kauffmann G. et al., 2003, MNRAS, 341, 54

Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713

Kelvin L. S. et al., 2014, MNRAS, 444, 1647

Khochfar S., Silk J., 2006, ApJ, 648, L21

Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38

Knop R. A. et al., 2003, ApJ, 598, 102

Koekemoer A. M. et al., 2011, ApJS, 197, 36

- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Koopmans L. V. E., Treu T., Bolton A. S., Burles S., Moustakas L. A., 2006, *ApJ*, 649, 599
- Köppen J., Weidner C., Kroupa P., 2007, *MNRAS*, 375, 673
- Kormendy J., 1977, *ApJ*, 218, 333
- Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, *ApJS*, 182, 216
- Kron R. G., 1980, *ApJS*, 43, 305
- Kuijken K., 2011, *The Messenger*, 146, 8
- Kuijken K. et al., 2004, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 5492, *Ground-based Instrumentation for Astronomy*, Moorwood A. F. M., Iye M., eds., pp. 484–493
- La Barbera F., Busarello G., Merluzzi P., Massarotti M., Capaccioli M., 2002, *ApJ*, 571, 790
- La Barbera F., de Carvalho R. R., 2009, *ApJ*, 699, L76
- La Barbera F., De Carvalho R. R., De La Rosa I. G., Gal R. R., Swindle R., Lopes P. A. A., 2010a, *AJ*, 140, 1528
- La Barbera F., de Carvalho R. R., de La Rosa I. G., Lopes P. A. A., Kohl-Moreira J. L., Capelato H. V., 2010b, *MNRAS*, 408, 1313
- La Barbera F., de Carvalho R. R., Kohl-Moreira J. L., Gal R. R., Soares-Santos M., Capaccioli M., Santos R., Sant’anna N., 2008, *PASP*, 120, 681
- La Barbera F., Ferreras I., de Carvalho R. R., Bruzual G., Charlot S., Pasquali A., Merlin E., 2012, *MNRAS*, 426, 2300
- La Barbera F., Ferreras I., Vazdekis A., de la Rosa I. G., de Carvalho R. R., Trevisan M., Falcón-Barroso J., Ricciardelli E., 2013, *MNRAS*, 433, 3017
- Lacey C., Silk J., 1991, *ApJ*, 381, 14
- Läscher R., van den Bosch R. C. E., van de Ven G., Ferreras I., La Barbera F., Vazdekis A., Falcón-Barroso J., 2013, *MNRAS*, 434, L31
- Laureijs R. et al., 2011, *arXiv.org*

- 
- Li R., Frenk C. S., Cole S., Gao L., Bose S., Hellwing W. A., 2016, *MNRAS*, 460, 363
- Liske J. et al., 2015, *MNRAS*, 452, 2087
- LSST Science Collaboration et al., 2009, ArXiv e-prints
- Maddox N., Hewett P. C., Warren S. J., Croom S. M., 2008, *MNRAS*, 386, 1605
- Magorrian J. et al., 1998, *AJ*, 115, 2285
- Martín-Navarro I., La Barbera F., Vazdekis A., Ferré-Mateu A., Trujillo I., Beasley M. A., 2015, *MNRAS*, 451, 1081
- Mather J. C. et al., 1994, *ApJ*, 420, 439
- McCracken H. J. et al., 2003, *A&A*, 410, 17
- McFarland J. P., Verdoes-Kleijn G., Sikkema G., Helmich E. M., Boxhoorn D. R., Valentijn E. A., 2013, *Experimental Astronomy*, 35, 45
- McKean J. et al., 2015, *Proceedings of Science*, 84
- Miyazaki S. et al., 2012, in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics, pp. 84460Z–84460Z
- Moore B., Quinn T., Governato F., Stadel J., Lake G., 1999, *MNRAS*, 310, 1147
- Moretti A. et al., 2014, *A&A*, 564, A138
- Mosleh M., Williams R. J., Franx M., 2013, *ApJ*, 777, 117
- Muzzin A. et al., 2013, *ApJS*, 206, 8
- Muzzin A., van Dokkum P., Franx M., Marchesini D., Kriek M., Labbé I., 2009, *ApJ*, 706, L188
- Naab T., Johansson P. H., Ostriker J. P., 2009, *ApJ*, 699, L178
- Napolitano N. R. et al., 2016, *The Universe of Digital Sky Surveys*, 42, 129
- Navarro J. F., Frenk C. S., White S. D. M., 1995, *MNRAS*, 275, 56
- Navarro J. F., White S. D. M., 1994, *MNRAS*, 267, 401
- Negrello M. et al., 2010, *Science*, 330, 800
- Oguri M., Marshall P. J., 2010, *MNRAS*, 405, 2579

- Oguri M., Rusu C. E., Falco E. E., 2014, MNRAS, 439, 2494
- Oser L., Naab T., Ostriker J. P., Johansson P. H., 2012, ApJ, 744, 63
- Pawase R. S., Faure C., Courbin F., Kokotanekova R., Meylan G., 2012, ArXiv e-prints
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2011, GALFIT: Detailed Structural Decomposition of Galaxy Images. Astrophysics Source Code Library
- Peng Y.-j. et al., 2010, ApJ, 721, 193
- Perlmutter S. et al., 1998, Nature, 391, 51
- Pillepich A. et al., 2014, MNRAS, 444, 237
- Poggianti B. M. et al., 2013a, ApJ, 762, 77
- Poggianti B. M. et al., 2013b, in IAU Symposium, Vol. 295, The Intriguing Life of Massive Galaxies, Thomas D., Pasquali A., Ferreras I., eds., pp. 151–154
- Poggianti B. M., Moretti A., Calvi R., D’Onofrio M., Valentinuzzi T., Fritz J., Renzini A., 2013c, ApJ, 777, 125
- Quilis V., Trujillo I., 2013, ApJ, 773, L8
- Ravindranath S. et al., 2002, in Bulletin of the American Astronomical Society, Vol. 201, American Astronomical Society Meeting Abstracts, p. 1099
- Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541
- Remus R.-S., Burkert A., Dolag K., Johansson P. H., Naab T., Oser L., Thomas J., 2013, ApJ, 766, 71
- Riess A. G. et al., 1998a, AJ, 116, 1009
- Riess A. G. et al., 1998b, AJ, 116, 1009
- Romeo A. D., Napolitano N. R., Covone G., Sommer-Larsen J., Antonuccio-Delogu V., Capaccioli M., 2008, MNRAS, 389, 13
- Rubin V. C., Thonnard N., Ford, Jr. W. K., 1978, ApJ, 225, L107
- Ruff A. J., Gavazzi R., Marshall P. J., Treu T., Auger M. W., Brault F., 2011, ApJ, 727, 96
- Saglia R. P. et al., 2010, A&A, 524, A6

- 
- Sales L. V., Navarro J. F., Schaye J., Dalla Vecchia C., Springel V., Booth C. M., 2010, MNRAS, 409, 1541
- Saulder C., van den Bosch R. C. E., Mieske S., 2015, A&A, 578, A134
- Scannapieco C., Athanassoula E., 2012, MNRAS, 425, L10
- Schaye J. et al., 2015, MNRAS, 446, 521
- Schaye J. et al., 2010, MNRAS, 402, 1536
- Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
- Schneider P., Ehlers J., Falco E., 1992, Gravitational Lenses. Springer-Verlag Berlin Heidelberg New York. Also Astronomy and Astrophysics Library
- Sérsic J. L., 1968, Atlas de galaxias australes. ""
- Shanno D., 1975, Math. Comp., 24, 647
- Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978
- Shih H.-Y., Stockton A., 2011, ApJ, 733, 45
- Shu Y. et al., 2015, ApJ, 803, 71
- Silk J., 1977, ApJ, 211, 638
- Skrutskie M. F. et al., 2006, AJ, 131, 1163
- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
- Spiniello C., Trager S., Koopmans L. V. E., Conroy C., 2014, MNRAS, 438, 1483
- Springel V., 2005, MNRAS, 364, 1105
- Springel V. et al., 2005a, Nature, 435, 629
- Springel V. et al., 2005b, Nature, 435, 629
- Stockton A., Shih H.-Y., Larson K., Mann A. W., 2014, ApJ, 780, 134
- Stringer M., Trujillo I., Dalla Vecchia C., Martinez-Valpuesta I., 2015, MNRAS, 449, 2396

- Szomoru D., Franx M., van Dokkum P. G., 2012, *ApJ*, 749, 121
- Szomoru D. et al., 2010, *ApJ*, 714, L244
- Tamura N., Kobayashi C., Arimoto N., Kodama T., Ohta K., 2000, *AJ*, 119, 2134
- Tamura N., Ohta K., 2000, *AJ*, 120, 533
- Taylor E. N., Franx M., Glazebrook K., Brinchmann J., van der Wel A., van Dokkum P. G., 2010, *ApJ*, 720, 723
- The Dark Energy Survey Collaboration, 2005, ArXiv e-prints
- Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, *ApJ*, 621, 673
- Toomre A., 1977, in *Evolution of Galaxies and Stellar Populations*, Tinsley B. M., Larson D. Campbell R. B. G., eds., p. 401
- Toomre A., Toomre J., 1972, *ApJ*, 178, 623
- Tortora C., La Barbera F., Napolitano N. R., de Carvalho R. R., Romanowsky A. J., 2012, *MNRAS*, 425, 577
- Tortora C., La Barbera F., Napolitano N. R., Romanowsky A. J., Ferreras I., de Carvalho R. R., 2014a, *MNRAS*, 445, 115
- Tortora C. et al., 2016, *MNRAS*, 457, 2845
- Tortora C., Napolitano N. R., Cardone V. F., Capaccioli M., Jetzer P., Molinaro R., 2010, *MNRAS*, 407, 144
- Tortora C., Napolitano N. R., Romanowsky A. J., Capaccioli M., Covone G., 2009, *MNRAS*, 396, 1132
- Tortora C., Napolitano N. R., Saglia R. P., Romanowsky A. J., Covone G., Capaccioli M., 2014b, *MNRAS*, 445, 162
- Tortora C., Pipino A., D'Ercole A., Napolitano N. R., Matteucci F., 2013a, *MNRAS*, 435, 786
- Tortora C., Romanowsky A. J., Napolitano N. R., 2013b, *ApJ*, 765, 8
- Tremaine S. et al., 2002, *ApJ*, 574, 740
- Trenti M., Stiavelli M., 2008, *ApJ*, 676, 767
- Treu T., Auger M. W., Koopmans L. V. E., Gavazzi R., Marshall P. J., Bolton A. S., 2010, *ApJ*, 709, 1195

---

Treu T., Koopmans L. V. E., 2004, ApJ, 611, 739

Trujillo I., Carrasco E. R., Ferré-Mateu A., 2012, ApJ, 751, 45

Trujillo I., Cenarro A. J., de Lorenzo-Cáceres A., Vazdekis A., de la Rosa I. G., Cava A., 2009, ApJ, 692, L118

Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, MNRAS, 382, 109

Trujillo I., Ferré-Mateu A., Balcells M., Vazdekis A., Sánchez-Blázquez P., 2014, ApJ, 780, L20

Trujillo I., Ferreras I., de La Rosa I. G., 2011, MNRAS, 415, 3903

Trujillo I. et al., 2006, ApJ, 650, 18

Tully R. B., Fisher J. R., 1977, A&A, 54, 661

Valentinuzzi T. et al., 2010, ApJ, 712, 226

van der Wel A., Bell E. F., van den Bosch F. C., Gallazzi A., Rix H.-W., 2009, ApJ, 698, 1232

van der Wel A. et al., 2014, ApJ, 788, 28

van der Wel A., Holden B. P., Zirm A. W., Franx M., Rettura A., Illingworth G. D., Ford H. C., 2008, ApJ, 688, 48

van der Wel A. et al., 2011, ApJ, 730, 38

van Dokkum P. G. et al., 2008, ApJ, 677, L5

van Dokkum P. G. et al., 2010, ApJ, 709, 1018

Vegetti S., Koopmans L. V. E., 2009, MNRAS, 400, 1583

Visvanathan N., Sandage A., 1977, ApJ, 216, 214

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M. et al., 2014, MNRAS, 444, 1518

Wellons S. et al., 2016, MNRAS, 456, 1030

Wellons S. et al., 2015, MNRAS, 449, 361

White S. D. M., 1978, MNRAS, 184, 185

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wu H., Shao Z., Mo H. J., Xia X., Deng Z., 2005, ApJ, 622, 244

Yasuda N. et al., 2001, AJ, 122, 1104

Yildirim A., van den Bosch R. C. E., van de Ven G., Husemann B., Lyubenova M.,  
Walsh J. L., Gebhardt K., Gültekin K., 2015, MNRAS, 452, 1792

York D. G. et al., 2000, AJ, 120, 1579

---

# Acknowledgement

First and foremost, I thank my PhD supervisor, Dr. Nicola R Napolitano and Dr. Francesco LaBarbera for the patient guidance, encouragement and advice, which made my thesis work possible. I thank them for their patience to listen and answer even my silly questions.

I sincerely thank Dr. Crescenzo Tortora for his encouragement, tolerance, contribution and mainly his willingness to answer all my questions. Without his help, I could not have completed this thesis. I would like to thank our group member Fedor Getman for all his suggestions and being always there to solve the technical issues.

I extend my sincere gratitude to Dr. Mauro D'Onofrio and Dr. Vincenzo Antonuccio for spending your time on my thesis. Your comments and suggestions helped to improve the thesis.

I am grateful to Prof. Massimo Capaccioli, Dr. Giovanni Covone and Prof. Maurizio Paolillo for their valuable advises and encouraging words. Special thanks to our PhD coordinator Prof. Raffaele Velotta for all his support.

I acknowledge financial support from INAF for the PhD grant of my thesis project.

I greatly appreciate Guido Celentano for all his help and kind words throughout all these four years. I would like to thank you for always being available to solve the official problems. Many thanks to my friends, Anoop, Jijil, Sam, Mahak, Payal, Tahmina, and Akif for making my stay in Naples was always fun.

Once again I would like to thank all the people who contributed in some way to the work described in this thesis.

Finally, I would also like to say a heartfelt thank you to my parents, brother, sister my other family members, and my friends in India for always believing in me and encouraging me whenever I needed it. Thank you for making it possible for me to complete what I started.