RAA 2016 Vol. **16** No. **3**, 43 (10pp) doi: 10.1088/1674–4527/16/3/043
http://www.raa-journal.org *http://iopscience.iop.org/raa*

***R**esearch in*
***A**stronomy and*
***A**strophysics*

A sample of galaxy pairs identified from the LAMOST spectral survey and the Sloan Digital Sky Survey

Shi-Yin Shen[1,2], Maria Argudo-Fernández[1], Li Chen[1], Xiao-Yan Chen[3], Shuai Feng[1,4], Jin-Liang Hou[1], Yong-Hui Hou[5], Peng Jiang[6], Yi-Peng Jing[7], Xu Kong[8], A-Li Luo[3], Zhi-Jian Luo[2], Zheng-Yi Shao[1,3], Ting-Gui Wang[6], Wen-Ting Wang[8], Yue-Fei Wang[5], Hong Wu[3], Xue-Bing Wu[9,10], Hai-Feng Yang[3], Ming Yang[3], Fang-Ting Yuan[1,2], Hai-Long Yuan[3], Hao-Tong Zhang[3], Jian-Nan Zhang[3] and Yong Zhang[5]

[1] Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China; *ssy@shao.ac.cn*
[2] Key Lab for Astrophysics, Shanghai 200234, China
[3] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China
[4] University of Chinese Academy of Sciences, Beijing 100049, China
[5] Nanjing Institute of Astronomical Optics & Technology, National Astronomical Observatories, Chinese Academy of Sciences, Nanjing 210042, China
[6] University of Science and Technology of China, Hefei 230026, China
[7] Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China
[8] Institute for Computational Cosmology, University of Durham, South Road, Durham, DH1 3LE, UK
[9] Department of Astronomy, Peking University, Beijing 100871, China
[10] Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

Received 2015 May 6; accepted 2015 October 12

Abstract A small fraction ($< 10\%$) of the SDSS main galaxy (MG) sample has not been targeted with spectroscopy due to the effect of fiber collisions. These galaxies have been compiled into the input catalog of the LAMOST ExtraGAlactic Surveys and named the complementary galaxy sample. In this paper, we introduce this project and status of the spectroscopies associated with the complementary galaxies in the first two years of the LAMOST spectral survey (till Sep. of 2014). Moreover, we present a sample of 1102 galaxy pairs identified from the LAMOST complementary galaxies and SDSS MGs, which are defined as two members that have a projected distance smaller than $100\,h_{70}^{-1}\mathrm{kpc}$ and a recessional velocity difference smaller than 500 $\mathrm{km\,s^{-1}}$. Compared with galaxy pairs that are only selected from SDSS, the LAMOST-SDSS pairs have the advantages of not being biased toward large separations and therefore act as a useful supplement in statistical studies of galaxy interaction and galaxy merging.

Key words: galaxies: interactions — galaxies: groups: general

1 INTRODUCTION

In the standard hierarchical structure formation model, galaxies are built up through merging processes. Numerical simulations show that galaxy mergers can trigger a starburst, feed the central supermassive black hole and transform the galaxy morphology (Springel et al. 2005). These different physical processes take place at different stages of galaxy merging. At the early stage, as two galaxies are approaching each other, they start to have interactions out to a distance of about 100 kpc. After the first passage, the galaxies start to show strong tidal tails and undergo starbursts. After a few passages, the galaxies quickly evolve into final coalescence. The whole timescale of the merging progress takes about $1 - 2$ Gyr (Torrey et al. 2012).

Using observations, the process of galaxy merging has been probed by statistical studies of galaxy pairs as a function of their separation, stellar mass, morphology, mass ratio and many other parameters (e.g. Nikolic et al. 2004; Ellison et al. 2008, 2013). In such studies, a large and unbiased sample of galaxy pairs is crucial. By far, the largest low redshift galaxy pair sample is identified from the main galaxy (MG) sample of the Sloan Digital Sky Survey (SDSS, York et al. 2000), which is a spectroscopic survey of a magnitude-limited sample down to $r < 17.77$

(Stoughton et al. 2002). The spectroscopic completeness of the MG sample in SDSS is quite high ($\sim 90\%$, Hogg et al. 2004). Based on the spectroscopic MGs in the final data release of the SDSS legacy survey (Data Release Seven, DR7, Abazajian et al. 2009), the number of galaxy pairs is over 10 000 (e.g. Ellison et al. 2011). Despite this large number, the galaxy pair sample identified from the spectroscopic MGs is far from complete. The incompleteness of galaxy pairs is mainly caused by the effect of fiber collisions in SDSS, which arises from a minimum separation of 55 arcsec between any two fibers for any given spectroscopic plate. As a result, the completeness of the galaxy pairs identified from the spectroscopic MGs is estimated to be only about 35 percent (Patton & Atfield 2008). In other words, galaxies missed by SDSS (hereafter referred to as "SDSS missed galaxies") caused by fiber collisions have a very high probability of being in galaxy pairs. Therefore, spectroscopic targeting of the SDSS missed galaxies is an efficient way to identify new galaxy pairs. Only if all these SDSS missed galaxies could be targeted by a new spectroscopic survey, a complete and unbiased galaxy pair sample could be finally made. More importantly, such a sample would be a benchmark in studies of the small-scale environment associated with low redshift galaxies.

In this article, we describe the project of observing SDSS missed MGs with the Guo Shou Jing Telescope (also named the Large Sky Area Multi-Object Fiber Spectroscopic Telescope - LAMOST)[1] and present some of its early results in the form of a new sample of galaxy pairs. This paper is organized as follows. In Section 2, we introduce the project of acquiring spectroscopic observations of the SDSS missed MGs with LAMOST. In Section 3, we present a new galaxy pair sample using new redshifts from the LAMOST survey. Finally, we provide short discussions and give a summary in Section 4.

2 LAMOST SURVEY: COMPLEMENTARY GALAXY SAMPLE

LAMOST is a special quasi-meridian reflecting Schmidt telescope located at Xinglong Station of National Astronomical Observatories, Chinese Academy of Sciences. The design of LAMOST provides an effective aperture of about 4 meters, a $\sim 5^\circ$ field of view and a spectroscopy system with ~ 4000 fibers (Wang et al. 1996; Su & Cui 2004; Cui et al. 2012). After about one year of a pilot survey (Luo et al. 2012), the LAMOST regular spectral survey started from September of 2012 and will last for five years. An overview of the LAMOST spectral survey can be found in Zhao et al. (2012). The LAMOST regular survey mainly focuses on Galactic stars, but also includes a significant fraction of extragalactic objects (e.g. Huo et al. 2013; Shi et al. 2014). One sample of extragalactic sources is the SDSS missed MGs, which is named the complementary galaxy sample in the LAMOST survey (Luo et al. 2015).

[1] *http://www.lamost.org*

2.1 The Complementary Galaxy Sample

The complementary galaxy sample is constructed from the catalog archive server of the SDSS legacy survey, where all the galaxies with r band Petrosian magnitude (Galactic reddening corrected) brighter than $r = 17.77$, but not yet having spectroscopic redshifts, are selected. The footprint of the complementary galaxies in the LAMOST survey is restricted to be in the North Galactic Cap region ($-10 < \delta < 60$ deg and $b > 0$ deg).[2] After that, we remove a small fraction of galaxy targets that might be contaminated by nearby bright stars using the spherical polygon masks in the NYU value added galaxy catalog (Blanton et al. 2005). The final number of targets in the complementary galaxy sample in the input catalog of the LAMOST spectral survey is 66 263.

In SDSS DR7, the number of MGs that have been targeted with spectroscopy in the North Galactic Cap is 639 428. That is to say, the fraction of SDSS missed MGs (66 263 of 705 691) is about 10 percent.

In Figure 1, we show the r band Petrosian magnitude distributions of the complementary galaxy sample and the SDSS spectroscopic MGs. As can be seen, the complementary galaxies have a similar magnitude distribution as the SDSS spectroscopic MGs, but are slightly biased toward faint galaxies. Most of the SDSS missed MGs are due to the fiber collision effect, therefore they are expected to be biased toward the high density region. To show this effect quantitatively, we match each complementary galaxy to the global MG sample and obtain the projected distance $\theta_{\rm min}$ (in units of arcsec) to its nearest neighbor. For comparison, we also calculate $\theta_{\rm min}$ for each SDSS spectroscopic MG. The two distributions of $\theta_{\rm min}$ are plotted in the bottom panel of Figure 1. Strong biases of two $\theta_{\rm min}$ distributions at 55 arcsec are clearly seen. This result not only shows that the complementary galaxies are biased to the high density environment, but also implies that the galaxy pair sample identified from the SDSS spectroscopic sample alone is far from complete (Patton & Atfield 2008).

To quantify the incompleteness of the galaxy pairs in SDSS, we select photometric galaxy pairs in the SDSS MG sample, which are defined as galaxies and their nearest neighbors inside radii of 100 arcsec. We match photometric pairs from both the SDSS spectroscopic MGs and from all the MGs. The fraction of galaxy pairs in these two samples gives us an estimation of the completeness of the galaxy pairs in SDSS spectroscopic MGs. We show the resultant completeness as a function of the pair angular separation in the top panel of Figure 2. Again, we see that there is a strong jump in incompleteness at the angular separation $\theta < 55$ arcsec, where the completeness is only about 30 percent and even decreases with a decrease of θ. In the bottom panel, we show the completeness of the galaxy pairs as a function of redshift. Here, we have

[2] For the south Galactic region, since the SDSS MGs are only located in three stripes, the LAMOST survey includes another independent galaxy spectroscopic survey project, which aims to acquire redshifts of a magnitude-limited sample down to $r < 18$.

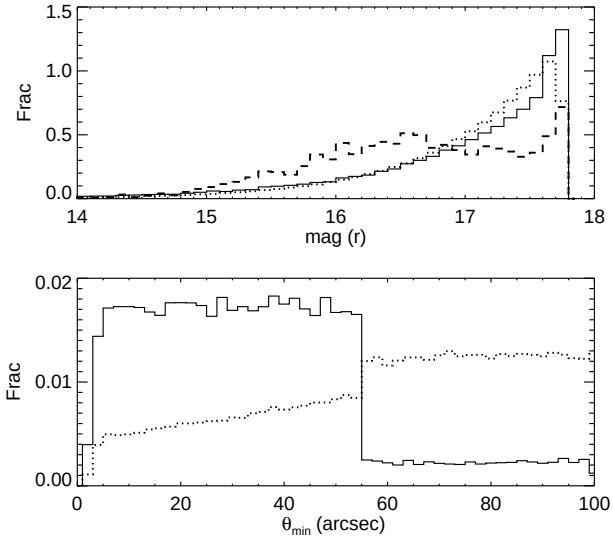

Fig. 1 The SDSS spectroscopic MGs and complementary galaxy sample. *Top*: the r band Petrosian magnitude distribution. The solid and dotted histograms show the distributions of the complementary galaxies and SDSS spectroscopic MGs respectively, while the dashed histogram shows the 3456 complementary galaxies with spectra in LAMOST DR2 (Sect. 2.2). *Bottom*: the angular separation to the nearest SDSS MG (complementary galaxies: solid, SDSS spectroscopic MGs: dotted). All the histograms have been normalized to have unit area.

defined the galaxy pairs as those with projected separation $r_{\rm p} < 100\ h_{70}^{-1}{\rm kpc}$ (see Sect. 3) and assumed that the separation follows a random distribution. In this case, the number of galaxy pairs in the SDSS spectroscopic MGs decreases with increasing redshift and reaches a plateau of ~ 30 percent at $z > 0.09$ where the 55 arcsec limit corresponds to a projected distance of 100 $h_{70}^{-1}{\rm kpc}$. Since the peak redshift of the SDSS MGs is at $z \sim 0.1$, the global completeness of the galaxy pairs in the SDSS spectroscopic MG sample is less than 40 percent.

2.2 LAMOST Observation

The complementary galaxies are mixed together with other LAMOST targets (most of them are Galactic stars) and then compiled into the LAMOST survey plates. In each plate, the number density of the complementary galaxies is very low, which are therefore assigned to fibers with higher priority than stars. In the LAMOST survey, the input sources are tiled into three different types of plates, bright (B), medium (M) and faint (F) plates, which are designed to reach the average signal-to-noise ratio (S/N$\sim$ 10) for objects down to the magnitude limits $r < 16.5$, $r < 17.8$ and $r < 18.5$ respectively. Most of the complementary galaxies have their magnitudes in the range $16.5 < r < 17.8$ (Fig. 1) and therefore are mainly tiled into the M plates. On the other hand, due to the limited number of dark nights, the amount of observing time allocated for the M plates is quite low. Because of this, only a small fraction of the complementary galaxy sample has been targeted so far.

The spectroscopic data used in this study are from the LAMOST Data Release 2 (DR2), i.e. the data till June 2014. In LAMOST DR2, there are 3456 complementary galaxies that have been targeted with spectroscopy and whose spectra have been released. The magnitude distribution of these 3456 LAMOST targeted galaxies is plotted as the dashed histogram in the top panel of Figure 1. Compared with the 66 263 complementary galaxies, the LAMOST targets are evidently biased toward bright galaxies. There are two reasons for this bias. First, some of the bright complementary galaxies ($r < 16.5$) are compiled into the target list of LAMOST B plates, which far outnumbers the M plates. The other reason for the bias against faint galaxies is the failure to acquire spectroscopic measurements because of low S/N (see the next section).

2.3 Redshift Measurements

In the catalog of LAMOST DR2, only 1951 of the 3456 complementary galaxies have redshifts measured from the LAMOST 1-D pipeline and are published in the LAMOST catalog. The failure of the redshift measurements is mainly due to the low S/N in LAMOST spectra of faint galaxies ($r > 16.5$). We show the mean S/N in the r-band wavelength range versus the r-band magnitude of the 3456 LAMOST targeted galaxies in the top panel of Figure 3. The distribution of S/N is shown in the bottom panel. As can be seen, when $r > 16.5$, the median S/N of the LAMOST spectra becomes lower than 10, which makes the number of spectra listed in the LAMOST catalog with measured z decrease significantly (see the dotted histogram in the bottom panel of Fig. 3).

To further improve the rate of successful redshift measurements, we developed an independent redshift measurement algorithm based on principal component analy-

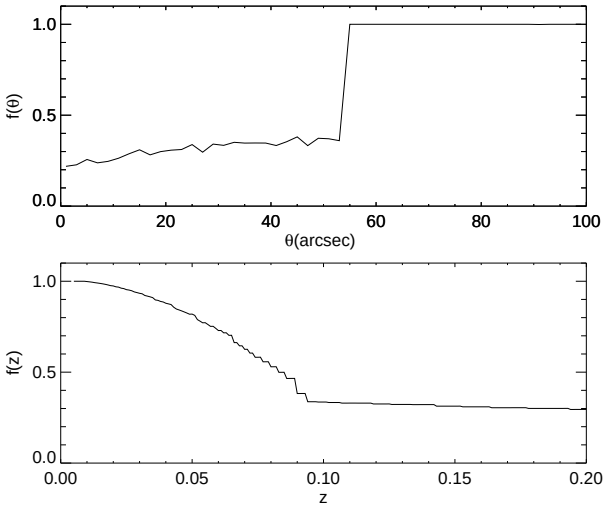

Fig. 2 The completeness of galaxy pairs in SDSS spectroscopic MGs. The top panel shows the completeness as a function of angular separation between the pair members. The bottom panel shows the completeness of the galaxy pairs (defined as those with projected distance $r_{\rm p} < 100\,h_{70}^{-1}{\rm kpc}$) as function of redshift.

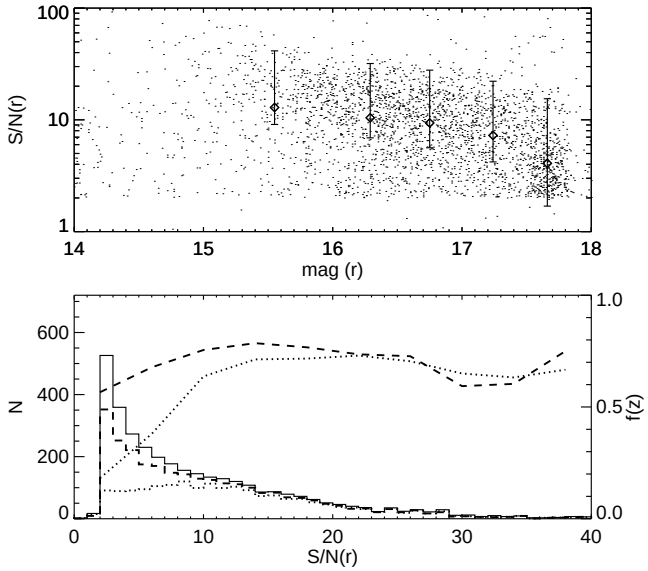

Fig. 3 The S/N in LAMOST spectra of the complementary galaxies. The top panel shows the magnitude as a function of S/N, while the bottom panel shows the histograms of S/N. *Top*: the small dots represent all 3456 complementary galaxies. The diamonds with error bars show the median and 16/84 percentiles of the S/N distribution in magnitude bins. *Bottom*: the solid histogram shows all the 3456 complementary galaxies, while the dashed and dotted histograms show the sub-samples of galaxies with redshifts measured from the PCA algorithm and from the LAMOST 1-D pipeline respectively. The fractions of galaxies with redshift measured (right y-axis) for the spectra at different S/N are shown by the dotted line (PCA algorithm) and dashed line (1-D pipeline).

sis (PCA) following the pipeline used for the 'CMASS' galaxies in the Baryon Oscillation Spectroscopic Survey (BOSS) that uses data from SDSS III (Bolton et al. 2012). The details of the pipeline will be presented in an upcoming paper. Here, we outline the basic routines in Appendix A. With this algorithm, we obtained 2796 redshift measurements.

In Figure 3, the dashed line shows the fraction of spectra with PCA redshifts measured. It is clear that our new PCA algorithm significantly improves the redshift measurement of the LAMOST spectra, especially at the low S/N end.

To check the quality of redshift measurements from the LAMOST 1-D pipeline and our PCA algorithm, we compare these redshifts with other independent measurements. After the SDSS legacy survey, some of the SDSS missed MGs were analyzed using BOSS spectroscopy in SDSS III (Dawson et al. 2013)[3]. We compared the 3456

[3] These galaxies typically have photo flag BRIGHT_GAL in SDSS III.

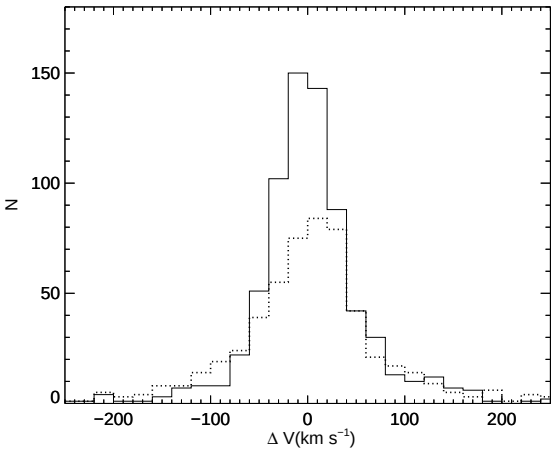

Fig. 4 Histogram showing the differences in redshift (in terms of recessional velocity) between the LAMOST and SDSS measurements. The solid histogram shows the differences between SDSS redshifts and our PCA redshift measurements, while the dotted histogram represents the differences between the SDSS redshifts and LAMOST 1-D redshifts.

complementary galaxies with the SDSS DR12 spectroscopic catalog and obtained 1056 matches. For these 1056 galaxies, the LAMOST catalog lists 604 redshifts while our PCA algorithm provides 923 redshift measurements.

In Figure 4, we show histograms of the differences in the redshifts (in terms of recessional velocity difference ΔV) of these galaxies with both LAMOST and SDSS redshifts. The LAMOST redshifts from the 1-D pipeline and PCA algorithm both show good consistency with SDSS values. For the LAMOST catalog z, the standard deviation of ΔV is about $80\,\mathrm{km\,s^{-1}}$. For the PCA z, the scatter of ΔV is even smaller, $\sim 58\,\mathrm{km\,s^{-1}}$. Given the better consistency with the SDSS redshifts of the PCA redshift measurements, we use the PCA redshifts for these galaxies with both PCA redshifts and catalog redshifts. For the 1056 galaxies with SDSS redshifts, we take their redshifts from the SDSS catalog. As we will show in the next section, the criterion of the velocity difference we adopted to identify galaxy pairs is $|\Delta V| < 500\,\mathrm{km\,s^{-1}}$. Therefore, the scatter between SDSS and LAMOST redshift measurements has few impacts on pair identification.

3 THE GALAXY PAIR SAMPLE

In this section, we combine the redshifts of the complementary galaxies with the SDSS spectroscopic MGs to identify new galaxy pairs. For the 3456 complementary galaxies, we have obtained 3137 redshifts. Among them, 1056 redshifts come from SDSS DR12, 1906 from the PCA algorithm, and 175 from the LAMOST 1-D pipeline.

In terms of observation, a galaxy pair is typically defined from the projected distance r_p and recessional velocity difference $|\Delta V|$ of two neighboring galaxies. However, there is no consensus on the critical values of r_p and $|\Delta V|$. For example, based on SDSS DR7, Liu et al. (2011) defined an AGN pair sample with $|\Delta V| < 600\,\mathrm{km\,s^{-1}}$ and $r_\mathrm{p} < 100\,h_{70}^{-1}\mathrm{kpc}$, whereas Patton et al. (2011) searched for galaxy pairs using $|\Delta V| < 1000\,\mathrm{km\,s^{-1}}$ and $r_\mathrm{p} < 80\,h_{70}^{-1}\mathrm{kpc}$ (see also Scudder et al. 2012; Argudo-Fernández et al. 2015). For the projected distance r_p, there is evidence that galaxies show interactions with their neighbors at $r_\mathrm{p} > 80\,h_{70}^{-1}\mathrm{kpc}$ (Scudder et al. 2012). In this study, we defined a critical value $r_\mathrm{p} < 100\ h_{70}^{-1}\mathrm{kpc}$. For ΔV, a large critical value (for example, $|\Delta V| < 1000\,\mathrm{km\,s^{-1}}$) might introduce a significant fraction that represents contamination from a high density environment (e.g. galaxy groups and clusters). In this study, we select galaxy pairs using $|\Delta V| < 500\,\mathrm{km\,s^{-1}}$ (see Appendix B for more discussions).

We match the redshifts of 3137 complementary galaxies with the spectroscopic MGs in SDSS DR7 using the criteria $r_\mathrm{p} < 100\ h_{70}^{-1}\mathrm{kpc}$ and $|\Delta V| < 500\,\mathrm{km\,s^{-1}}$, and obtain 1141 galaxy pair candidates. In a few cases, a complementary galaxy may have more than one SDSS spectroscopic MGs matched. In this case, we choose the galaxy with the smallest r_p to be the member of the matched pair and mark this pair as being in a multiple system. We will describe multiple systems in Section 3.1. Moreover, to have better quality control in the pair sample, we visually inspect the SDSS images of all the pair candidates. In a few cases, the imperfect SDSS pipeline de-blends large galaxies into several small children, resulting in 39 fake pairs. Thus, our final sample includes 1102 galaxy pairs.

We show histograms of r_p and $|\Delta V|$ for the final 1102 galaxy pairs in the bottom two panels of Figure 5, where the distributions of the angular separation Θ (in arcsec) between the pair members and their average redshifts are shown in the top two panels. To have a better understanding of the statistical properties of the new pair sample, we show the distributions of the SDSS only pairs as the dotted histograms in each panel of Figure 5 for comparison. The SDSS only pairs are selected from the SDSS DR7 group catalog of Yang et al. (2008) using the same criteria above. The number of SDSS pairs is 16 973. Because of this large number, we have not visually inspected this sample.

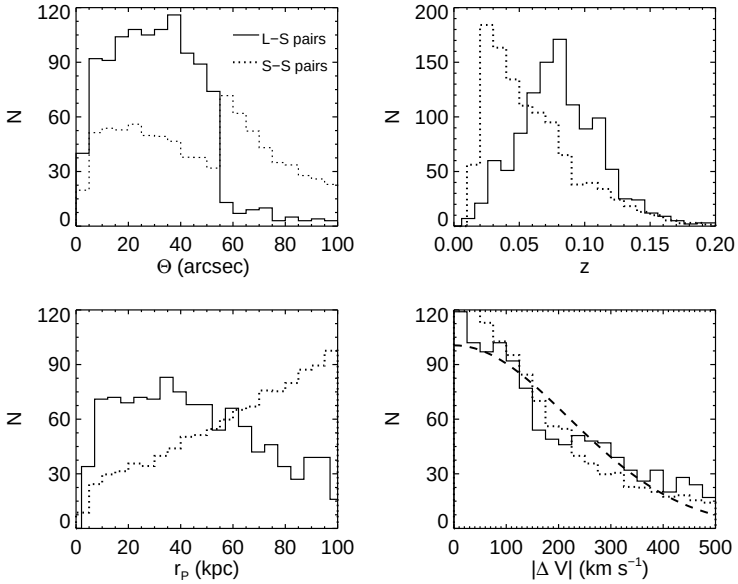

Fig. 5 Basic statistical properties of the 1102 galaxy pairs identified from the LAMOST complementary galaxies and SDSS MGs (*solid histograms*). The statistical properties of the SDSS selected galaxy pairs are plotted as the dotted histograms in each panel for comparison (normalized to 1102). The top left, top right, bottom left and bottom right panels show the histograms of the angular separation, redshift, projected distance and recessional velocity differences respectively. In the bottom right panel, the dashed curve shows a Gaussian distribution function with a standard deviation of $218\,\mathrm{km\,s^{-1}}$.

However, because there are only 39 of 1141 LAMOST-SDSS pairs that are fake, we expect the impact of the fake pairs on its statistical properties to be quite small. For convenience in the following discussion, we abbreviate the LAMOST-SDSS pairs as the LS pairs and the SDSS only pairs as SS pairs.

As can be seen from Figure 5, except for the distribution of $|\Delta V|$, the LS pairs show a significant difference from the SS pairs. The angular separation of the SS pairs is clearly biased toward large values ($\Theta > 55$ arcsec, top left panel). This, as we mentioned, is because of collisions among the 55 arcsec fibers used in SDSS. The SS pairs are also biased toward lower redshifts (top right panel), which again is because of the effect of fiber collisions. Although the low z selection effect compensates the high Θ bias, as a combination, the SS pairs are still biased toward large separations ($r_{\rm p}$) as shown in the bottom left panel. Compared with the SS pair sample, our new LS pairs only increase the number of pairs by a few percent ($\sim 7\%$). However, for the close pairs with more significant interaction ($r_{\rm p} < 30\ h_{70}^{-1}\mathrm{kpc}$), the number of LS pairs enlarges the SS pairs by more than 14 percent (358 versus 2527). Therefore, the LS pair sample could be a useful supplement to the current SS pair sample, especially for close pairs, which are valuable in statistical studies of galaxy interaction and merging.

For the velocity difference (bottom right panel), the standard deviation of the LS and SS pairs is 218 and 198 $\mathrm{km\,s^{-1}}$ respectively. Considering the fact that the dispersion of the redshift differences between the LAMOST and SDSS measurements could be as large as $\sim 80\,\mathrm{km\,s^{-1}}$ (Fig. 4), these two results are consistent with each other. In this panel, we also plot a Gaussian distribution with standard deviation $\sigma = 218\,\mathrm{km\,s^{-1}}$ for comparison. As we can see, the distribution of $|\Delta V|$ for LS pairs deviates from a Gaussian distribution, especially at the tails (see Appendix B for a more detailed discussion).

The catalog of LS pairs is listed in Table 1. In addition to basic parameters (e.g. RA, Dec, redshift) for each pair member, we also list two extra flags ($\mathrm{M\_Flag}$ and $\mathrm{O\_Flag}$) for each pair, which characterize the multiplicity and the degree of image overlap for the pair members respectively. In the following two sub-sections, we give brief descriptions of these two flags.

3.1 Multiple System

Our galaxy pair sample is defined from two simple observational criteria ($r_{\rm p} < 100\,h_{70}^{-1}\mathrm{kpc}$ and $|\Delta V| < 500\,\mathrm{km\,s^{-1}}$). Besides the pair members, the other neighboring galaxies have not been taken into consideration. However, to study the galaxy interaction and galaxy merging using galaxy pairs, it is better to refine a sample of $physical$ pairs, where the interaction between two pair members dominates the effects from other neighbors (e.g. Argudo-Fernández et al. 2015). For this purpose, we mark the galaxy pairs in a multiple system with a multiple flag '$\mathrm{M\_Flag}$'. Our aim is to indicate that these pairs may not be suitable for studying galaxy interactions by only considering their members.

We define a galaxy pair in a multiple system when either of its members has another MG neighbor ($r < 17.77$) that also satisfies the pair definition ($r_{\rm p} < 100\,h_{70}^{-1}\mathrm{kpc}$ and $|\Delta V| < 500\,\mathrm{km\,s^{-1}}$). Within the current data, 197 of the

Table 1 The catalog of galaxy pairs identified from the LAMOST complementary galaxies and SDSS MGs. For each pair, α_1, δ_1 and z_1 are the Right Accession, Declination and redshift of the LAMOST complementary galaxy respectively, while α_2, δ_2 and z_2 are those of the SDSS galaxy. $\mathrm{M\_Flag}$ and $\mathrm{O\_Flag}$ are the multiplicity and overlapping flags respectively (see Sections 3.1 and 3.2 for details). The table is sorted in ascending order of α_1. The complete table is available at *http://www.raa-journal.org/docs/Supp/2306Tab1.dat*.

ID	α_1[deg]	δ_1 [deg]	z_1	α_2[deg]	δ_2[deg]	z_2	M_Flag	O_Flag
1	112.23653	36.91987	0.06006	112.24269	36.91610	0.05988	0	0
2	114.54001	28.13682	0.07852	114.54659	28.12799	0.07942	0	0
3	116.04450	23.99016	0.07545	116.05284	23.99939	0.07533	0	0
4	116.46624	26.47179	0.12311	116.46654	26.47719	0.12373	0	0
5	119.86285	23.97182	0.09218	119.85204	23.98273	0.09318	0	0
6	120.09847	39.83033	0.01320	120.17009	39.87050	0.01326	1	0
7	120.40500	15.70968	0.01545	120.36157	15.74989	0.01637	1	0
8	120.83544	23.96435	0.05784	120.84409	23.96933	0.05725	0	0
9	120.89831	28.54465	0.14178	120.89262	28.55000	0.14096	0	0
10	121.02204	31.44029	0.07302	121.03594	31.43697	0.07303	0	0
11	121.37741	22.13299	0.13852	121.38057	22.12452	0.14003	0	0
12	123.44543	8.38181	0.11409	123.44723	8.38925	0.11324	0	0
13	123.98402	8.28538	0.14295	123.97318	8.28571	0.14438	0	0
14	124.05299	3.85966	0.08668	124.05083	3.85807	0.08777	0	1
15	124.46973	7.57755	0.12465	124.46754	7.57477	0.12484	0	0

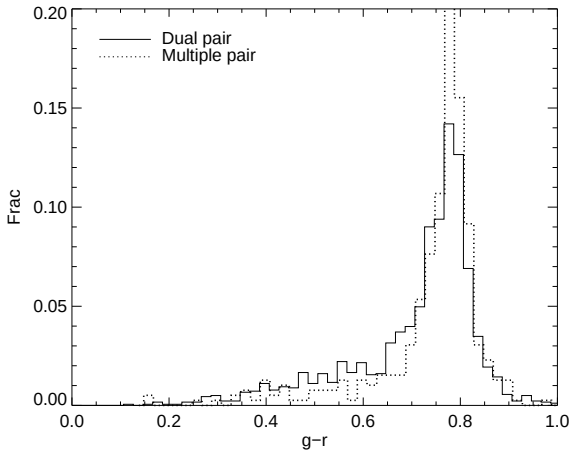

Fig. 6 Distribution of the the $g-r$ color of the pair members. The solid and dotted histograms (normalized to unit area) show the galaxies in the pairs with $\mathrm{M\_Flag} = 0$ (dual) and $\mathrm{M\_Flag} = 1$ (multiple) respectively.

1102 LS pairs are found in multiple systems. These pairs are marked with flag $\mathrm{M\_Flag} = 1$ in Table 1 and denoted as 'multiple' pairs below. For comparison, the pairs with $\mathrm{M\_Flag} = 0$ are denoted as 'dual' pairs.

To check the possible distinction between the multiple and dual pairs, we compare the $g-r$ color (taken from the SDSS model magnitudes with K-correction applied) distributions of their members in Figure 6. The galaxies in multiple pairs are on average redder than those in dual ones. Specifically, the red galaxy fractions ($g-r > 0.7$) in the dual and multiple pairs are 0.52 and 0.66 respectively. This color bias is caused by the fact that multiple pairs are biased toward a high density environment (i.e. galaxy groups and clusters).

We note that the $\mathrm{M\_Flag}$ set for the case of multiple systems is quite a preliminary parameter. Since the pairs are selected from a magnitude-limited sample, both the pair sample and the multiplicity flag have strong redshift dependence. For example, a triplet system with $M_{\mathrm{r},i} = -20, -21$ and –22 mag would be identified as a triplet, a pair or a single galaxy in the SDSS MG sample ($r < 17.77$) at redshift $z = 0.05,\ 0.1$ and 0.15 respectively. On the other hand, the multiplicity flag is also not complete, which is because of high incompleteness in the SDSS spectroscopic MGs at small scales and also because of the small fraction of complementary galaxies that have been targeted by LAMOST so far. For the dual pairs with $\mathrm{M\_Flag} = 0$, they are still possibly in multiple systems, i.e. having another companion galaxy but without a measured redshift yet.[4] That is to say, the fraction of pairs in multiple systems (197 of 1102) is actually a lower limit. Given the significant fraction of galaxy pairs in multiple systems, studies of

[4] The pair with ID=52 in Table 1 is actually an isolated and compact galaxy triplet, which is found to be a triplet candidate during our visual inspection and then spectroscopically confirmed by a follow-up observation with the 2.16m telescope at Xinglong Station (Feng et al. 2015).

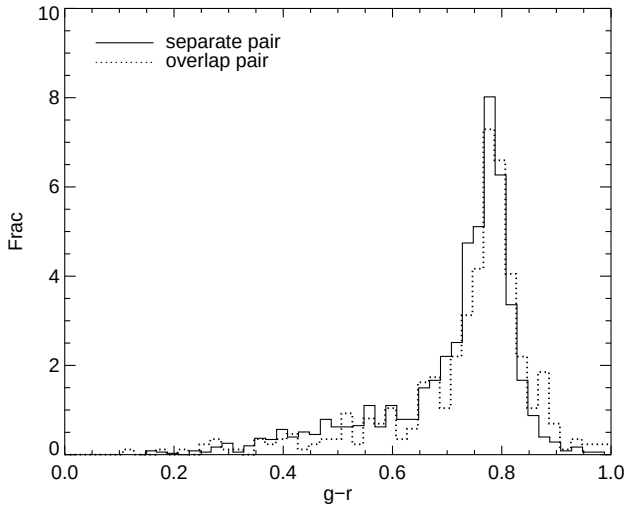

Fig. 7 Distribution of the $g-r$ color of the pair members. The solid and dotted histograms (normalized to unit area) show the galaxies in the pairs with $\mathrm{O\_Flag} = 0$ and $\mathrm{O\_Flag} = 1$ respectively.

galaxy interactions with galaxy pairs should take the multiplicity of the galaxy systems into consideration (Shen et al. 2016, in preparation).

3.2 Overlapping Pairs

The luminosity/mass ratio of the pair members plays an important role in galaxy merging (Jiang et al. 2014). However, photometry of a very close galaxy pair is a non-trivial task (Simard et al. 2011). During the visual inspection of the galaxy pairs, we noticed some of the galaxy pairs whose images of their members significantly overlap each other. We mark such pairs ($N = 216$) with a flag $\mathrm{O\_Flag} = 1$. For these overlapping pairs, the uncertainties in their photometry might introduce uncertainties and possible biases into estimation of the stellar mass and mass ratio.

We show the $g-r$ color distribution of the overlapping pairs as the dotted histogram in Figure 7, where the distribution of the other pairs is shown as the solid histogram for comparison. As can be seen, although the overlapping pairs have a much closer projected distance, their colors are biased toward redder values. The fraction of red galaxies ($g-r > 0.70$) in the overlapping pairs is 0.62, which is significantly higher than the fraction 0.52 for non-overlapping pairs. Unless there are large uncertainties in the photometry of these overlapping pairs, it is hard to explain the color bias we see in Figure 7. Actually, Patton et al. (2011) have already shown that the poor photometry of the official pipeline used by SDSS is largely responsible for the suspiciously large fraction of extremely red galaxies (e.g. $g-r > 0.9$) in the very close galaxy pairs. Therefore, to get a better estimation of the mass ratio of these overlapping pairs, more detailed photometry is required (e.g. Simard et al. 2011). Here, we set this flag to be a cautionary indicator and leave the detailed photometry for future work.

4 CONCLUSIONS

In this paper, we describe the spectroscopic survey of SDSS missed MGs ($r < 17.77$) using LAMOST. The SDSS missed MGs are identified as a complementary galaxy sample in the LAMOST survey. In the first two years of the LAMOST survey, due to the limited survey time allotted to the medium (M) plates, only a small fraction (3456 of 66 263) of the complementary galaxies have obtained LAMOST spectra. The majority of these spectra have quite low S/N, which is mostly due to poor seeing conditions. We developed a PCA algorithm to improve the redshift measurements of these low S/N spectra. Together with the SDSS DR12 match and LAMOST 1-D pipeline results, we finally obtained 3137 redshifts of the 3456 complementary galaxies.

Considering the fact that the SDSS missed galaxies are mainly caused by the effect of fiber collisions, the spectroscopy of the complementary galaxies has great potential in identifying new galaxy pairs. We present such a catalog of galaxy pairs identified from the first two years data from the LAMOST survey. From the redshifts of 3137 complementary galaxies, we obtained a sample of 1102 galaxy pairs after a careful visual inspection. Compared with the galaxy pairs only selected from SDSS data, our pair sample includes a larger fraction of close pairs ($r_{\rm p} < 30\,h_{70}^{-1}{\rm kpc}$). Because of such advantages, our sample increased the current SDSS close pair sample by about ~ 15 percent. Like in other studies, our pairs are selected using two simple observational criteria ($r_{\rm p}$ and $|\Delta V|$). Whether they are physically bound systems has not been taken into consideration. We find that at least ~ 20 percent of the pairs are actually located in multiple systems ($\mathrm{M\_Flag} = 1$). Moreover, during the visual inspection, we find, for about 20 percent of

the pairs, the images of their members overlap each other. Therefore, the photometry and stellar mass estimation of the galaxies in these overlapping pairs ($\mathrm{O\_Flag} = 1$) should be used with caution.

With the continued progress of the LAMOST survey in targeting complementary galaxies in current and future seasons (from Sep. of 2014), we expect that the completeness of the galaxy pairs will be further improved. Once the LAMOST survey has finished acquiring the spectroscopies of the complementary galaxy sample, great advancements in studies of galaxy pairs, galaxy interactions, galaxy merging and the small-scale environmental effects of galaxies are expected.

Acknowledgements The Guo Shou Jing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope, LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by National Astronomical Observatories, Chinese Academy of Sciences.

This work is supported by the National Basic Research Program of China (973 Program, 2014 CB845705), Strategic Priority Research Program "The Emergence of Cosmological Structures" of the Chinese Academy of Sciences (CAS; grant XDB09030200) and the National Natural Science Foundation of China (Nos. 11573050 and 11433003).

Appendix A: REDSHIFT MEASUREMENTS WITH PCA EIGEN-TEMPLATES

We developed an independent algorithm to measure the redshifts of the LAMOST complementary galaxies using PCA eigen-templates and chi-squared fitting.

The eigen-templates are derived from the 1056 complementary galaxies with DR12 redshifts. Specifically, we first divide 1056 galaxies into 20 $g-r$ color bins with similar numbers. Then, using the SDSS DR12 redshifts, we stack their LAMOST spectra and build the LAMOST composite spectra for galaxies in each $g-r$ bin. We apply a PCA decomposition to these 20 composite spectra and find that the first four eigenvectors can recover most of the spectroscopic features.

We explore the redshift of each galaxy in the range $0.005 < z < 0.5$ by taking trial values that are advanced in steps representing the pixels of each spectrum. For each trial redshift, we fit the observed spectrum with the error-weighted least-square linear combination of the four 'eigen-spectra' and a fourth-order polynominal. The polynominal is introduced to compensate the calibration of the LAMOST spectrum. The reduced χ^2 value (the resulting χ^2 divided by the number of fitted pixels) for each trial redshift defines a $\chi^2(z)$ curve in the probed redshift range. The best redshift estimation is then defined by the minimum of the $\chi^2(z)$ curve. The error is evaluated at the location where the χ^2 is increased by one at each side of the minimum values. During the fitting, besides the pixels that are part of the ANDMASK set, we have also masked the pixels at wavelengths where the sky-subtraction residuals are three times more than the average noise.

Appendix B: THE PAIR-WISE PECULIAR VELOCITY DISTRIBUTION OF GALAXY PAIRS

The distribution of the recessional velocity difference between the galaxy pair members, also known as the pair-wise peculiar velocity distribution function (PVDF), plays a very important role in large scale structure studies. On non-linear scales ($r < 1$ Mpc), both observations and theoretical models suggest the PVDF has an exponential form, which can be explained and approximated by a weighted integral of Gaussian distributions of subunits (e.g. Diaferio & Geller 1996; Sheth 1996). However, the PVDF of the pairs in our analysis is on a very small scale, even smaller than the viral radius of galaxies ($r < 100$ kpc), which has not drawn much attention in related studies.

We show the pair-wise peculiar velocity (ΔV) distribution of the SDSS selected pairs in Figure B.1. The SDSS pairs are selected with criteria $r_{\rm p} < 100$ kpc and $|\Delta V| < 800\,\mathrm{km\,s^{-1}}$. As we can see, the ΔV distribution shows long tails out to $\pm 800\,\mathrm{km\,s^{-1}}$ and cannot be fitted well by a Gaussian function. We first fit the observed ΔV distribution with the usually adopted exponential profile,

$$f(\Delta V) = \frac{1}{\sqrt{2}\sigma}\exp\left(-\frac{\sqrt{2}\Delta V}{\sigma}\right), \tag{B.1}$$

where σ characterizes the pair-wise peculiar velocity dispersion. The best fitting of the exponential profile has $\sigma = 324\,\mathrm{km\,s^{-1}}$ and is shown as the blue line in Figure B.1. Although better than the Gaussian profile fitting, the exponential profile still cannot fit the fine structure of the observed ΔV distribution well, especially at ranges $|\Delta V| > 200\,\mathrm{km\,s^{-1}}$.

Alternatively, as a preliminary test, we fit the observed ΔV distribution with a multi-component model. The observed ΔV distribution might be contributed by different components and each component corresponds to different physical circumstances. The ideal case of a galaxy pair is that the pair members form a gravitationally bound system. In this case, ΔV represents the orbital velocity of the pair members. A more common case of an observed galaxy pair is when the galaxy pairs have other close neighbors. That is to say, the pair members and their neighbors are located together in a larger galaxy system, e.g. galaxy groups or clusters. In this case, members of the pair follow individual orbits in the gravitational potential of the host halo and ΔV represents its velocity dispersion. Finally, ΔV may be dominated by the difference in the Hubble flow. In this case, the radial distance of the members ($\sim$ Mpc) is much larger than the tangential distance (< 100 kpc). That is to

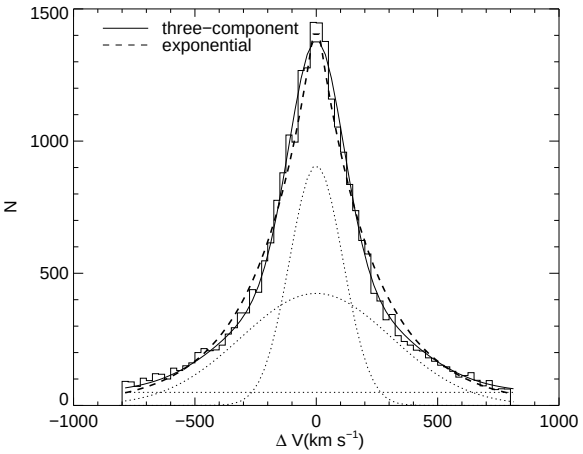

Fig. B.1 Distribution of difference in recessional velocity ΔV of the galaxy pairs in SDSS. The best exponential model fitting is shown as the blue line. The best three-component model fitting is shown as the red line, where the contributions of each component are shown as the dotted lines.

say, the 'pair' members actually have few physical interactions with each other and such pairs could be considered as contaminations from projection. Motivated by the above scenario, we fit the observed ΔV distribution with three components, one narrow Gaussian profile (orbital velocities of ideal pairs), one broad Gaussian profile (velocity dispersion of host halos) and a flat component (projection contamination). The best fitting of this three-component model is shown as the solid line in Figure B.1. As we can see, this model yields a reasonably good fitting to the observed ΔV distribution. The ΔV profiles of the three components are shown as the dotted lines in Figure B.1. The narrow Gaussian component has a standard deviation of $\sim 110\,\mathrm{km\,s^{-1}}$ and contributes about 40 percent of the galaxy pairs with $|\Delta V| < 800\,\mathrm{km\,s^{-1}}$. The broad component has a standard deviation $\sim 300\,\mathrm{km\,s^{-1}}$ and contributes half. The constant component contributes the last 10 percent.

In Section 3.2, we select the LAMOST-SDSS galaxy pairs with criteria $r_\mathrm{p} < 100$ kpc and $|\Delta V| < 500\,\mathrm{km\,s^{-1}}$. Our motivation for choosing a smaller critical value of $|\Delta V|$ is to reduce the contribution from galaxy groups and clusters, i.e. the broad component in Figure B.1, while including all possible conditions related to ideal pairs. A more detailed study of the ΔV distribution of the galaxy pairs in different environments is in preparation.

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Argudo-Fernández, M., Verley, S., Bergond, G., et al. 2015, A&A, 578, A110

Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562

Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, AJ, 144, 144

Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, RAA (Research in Astronomy and Astrophysics), 12, 1197

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

Diaferio, A., & Geller, M. J. 1996, ApJ, 467, 19

Ellison, S. L., Mendel, J. T., Patton, D. R., & Scudder, J. M. 2013, MNRAS, 435, 3627

Ellison, S. L., Patton, D. R., Mendel, J. T., & Scudder, J. M. 2011, MNRAS, 418, 2043

Ellison, S. L., Patton, D. R., Simard, L., & McConnachie, A. W. 2008, AJ, 135, 1877

Feng, S., Shao, Z.-Y., Shen, S.-Y., et al. 2015, arXiv:1512.02439

Hogg, D. W., Blanton, M. R., Brinchmann, J., et al. 2004, ApJ, 601, L29

Huo, Z.-Y., Liu, X.-W., Xiang, M.-S., et al. 2013, AJ, 145, 159

Jiang, C. Y., Jing, Y. P., & Han, J. 2014, ApJ, 790, 7

Liu, X., Shen, Y., Strauss, M. A., & Hao, L. 2011, ApJ, 737, 101

Luo, A.-L., Zhang, H.-T., Zhao, Y.-H., et al. 2012, RAA (Research in Astronomy and Astrophysics), 12, 1243

Luo, A.-L., Zhao, Y.-H., Zhao, G., et al. 2015, RAA (Research in Astronomy and Astrophysics), 15, 1095

Nikolic, B., Cullen, H., & Alexander, P. 2004, MNRAS, 355, 874

Patton, D. R., & Atfield, J. E. 2008, ApJ, 685, 235

Patton, D. R., Ellison, S. L., Simard, L., McConnachie, A. W., & Mendel, J. T. 2011, MNRAS, 412, 591

Scudder, J. M., Ellison, S. L., Torrey, P., Patton, D. R., & Mendel, J. T. 2012, MNRAS, 426, 549

Sheth, R. K. 1996, MNRAS, 279, 1310

Shi, Z.-X., Luo, A.-L., Comte, G., et al. 2014, RAA (Research in Astronomy and Astrophysics), 14, 1234

Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11

Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485

Su, D.-Q., & Cui, X.-Q. 2004, ChJAA (Chin. J. Astron. Astrophys.), 4, 1

Torrey, P., Cox, T. J., Kewley, L., & Hernquist, L. 2012, ApJ, 746, 108

Wang, S.-G., Su, D.-Q., Chu, Y.-Q., Cui, X., & Wang, Y.-N. 1996, Appl. Opt., 35, 5155

Yang, X., Mo, H. J., & van den Bosch, F. C. 2008, ApJ, 676, 248

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, RAA (Research in Astronomy and Astrophysics), 12, 723