

Statistical assessment of the relation between the inferred morphological type and the emission-line activity type of a large sample of galaxies

R. A. Ortega-Minakata¹, J. P. Torres-Papaqui¹, H. Andernach¹ and J. M. Islas-Islas^{1,2}

¹Departamento de Astronomía, Universidad de Guanajuato, 36000, Guanajuato, Mexico

²Universidad Tecnológica de Tulancingo, 43642, Hidalgo, Mexico

Abstract. We quantify the statistical evidence of the relation between the inferred morphology and the emission-line activity type of galaxies for a large sample of galaxies. We compare the distribution of the inferred morphologies of galaxies of different dominant activity types, showing that the difference in the median morphological type between the samples of different activity types is significant. We also test the significance of the difference in the mean morphological type between all the activity-type samples using an ANOVA model with a modified Tukey test that takes into account heteroscedasticity and the unequal sample sizes. We show this test in the form of simultaneous confidence intervals for all pairwise comparisons of the mean morphological types of the samples. Using this test, scarcely applied in astronomy, we conclude that there are statistically significant differences in the inferred morphologies of galaxies of different dominant activity types.

Keywords. methods: statistical, galaxies: statistics, galaxies: active

1. Introduction and sample selection and classification

There is anecdotal evidence of galaxies with different morphologies having different dominant emission-line activity types (Heckman 1980; Ho, Filippenko & Sargent 1997; Sabater *et al.* 2012).

We inferred the morphological type of 800,685 galaxies of the spectroscopic sample of the SDSS-DR7 (Abazajian *et al.* 2009) using the relations between the broad-band colours $u-g$, $g-r$, $r-i$ and $i-z$ in the SDSS system and a concentration index of their brightness profile, $R_{90}(r)/R_{50}(r)$ (where $R_{90}(r)$ and $R_{50}(r)$ are the Petrosian radii containing 90% and 50% of the Petrosian flux in the r band, respectively), of subsamples of these galaxies and their morphology classified by eye (Shimasaku *et al.*, 2001; Fukugita *et al.*, 2007). The numerical scale of our inferred morphology (1–6) is that of Fukugita *et al.*'s (2007) visual classification, so galaxies of earlier types have an inferred morphology closer to 0, while later types have one closer to 6. Fig. 1 (left) shows the distribution of our inferred morphology.

Also, using a standard diagnostic diagram based on emission-line ratios (Baldwin, Phillips, & Terlevich 1981, BPT) we classified our large sample of galaxies according to their dominant activity type (Fig. 1, right). We used the STARLIGHT code (Cid Fernandes *et al.* 2005, www.starlight.ufsc.br) to measure the emission-line flux of all the lines seen in the spectra of the entire sample of galaxies. For our analysis, we only considered 216,510 galaxies with redshifts 0.03–0.30, with good quality spectra according to relevant flags in the spectroscopic catalogue of the SDSS-DR7, and with the lines

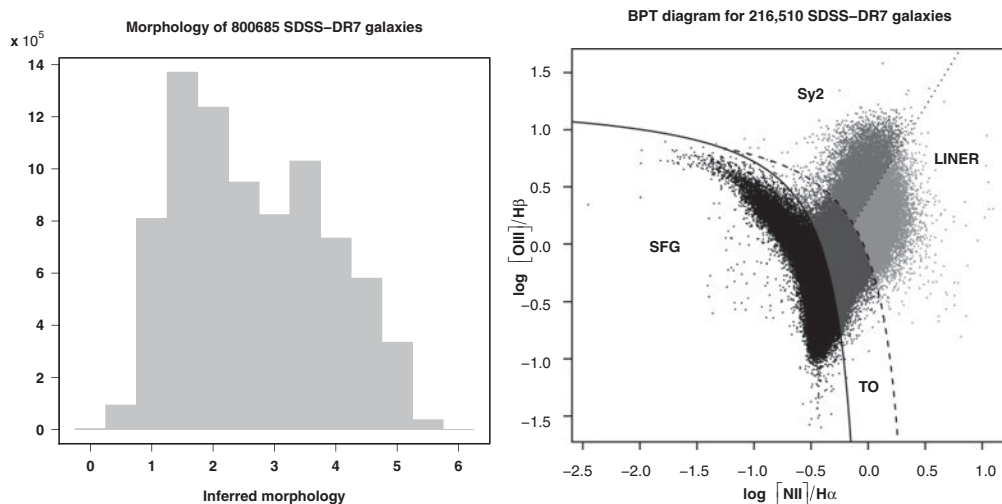


Figure 1. *Left:* Distribution of our inferred morphology for 800,685 galaxies.
Right: BPT diagnostic diagram for 216,510 galaxies and samples defined from it.

relevant to the diagnostic diagram having $S/N > 3$. The criterion to separate star-forming galaxies (SFGs) from AGN-dominated galaxies is that of Kauffmann *et al.* (2003); to separate transition objects (TOs) from AGN-dominated galaxies we use the criterion proposed by Kewley *et al.* (2001), and to distinguish between Sy2s and LINERs we use the method explained by Torres-Papaqui *et al.* (2012).

2. Analysis and conclusions

Using the R suite of statistical analysis (www.r-project.org/), we compared the distribution of our inferred morphology for samples of galaxies with different dominant emission-line activity (Fig. 2, left). The samples are those obtained from the BPT diagram (see previous section) and a sample of 16,733 passive galaxies, i.e. those in which none of the emission lines relevant for the BPT diagram are detected, also in the 0.03–0.30 redshift range. The vertical extent of each box is the inter-quartile range (IQR), while the whiskers length is 1.5 times the IQR from the extremes of the box in each direction. The median of each sample is the thick, black bar, while the mean is the black dot.

The error of the median inferred morphology of each activity type is shown by the width of the notch (a double-cone-shaped region drawn around the median), which is proportional to IQR/\sqrt{N} , where N is the number of galaxies in each sample of different activity type (shown below each). Comparing any two samples, no overlapping notches imply significantly different medians. Note that a large number of galaxies or a small dispersion within each sample can make the notches small, sometimes barely visible.

The figure shows that there are significant differences in the median morphology for different dominant activity types, going from later to earlier morphologies along a sequence of SFGs, TOs, Sy2s, LINERs and passive galaxies. This suggests a strong relationship between the dominant activity type and the morphology of galaxies, with SFGs being predominantly later types, and LINERs and passive galaxies being mostly earlier types.

Fig. 2 (right) shows a test of the significance of the difference in the mean morphological type between all pairs of activity-type samples. This test uses an ANOVA model with a modified Tukey test. It compares all possible pairs of means simultaneously; the null

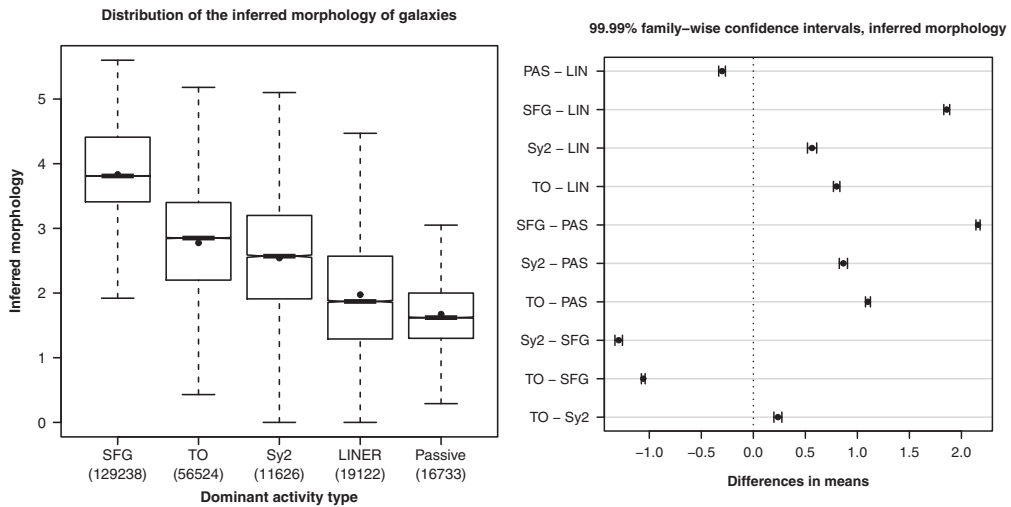


Figure 2. *Left:* Distribution of our inferred morphology for samples of galaxies with different dominant emission-line activity. Sample sizes are given in parentheses. *Right:* Test of the significance of the difference in the mean morphological type between all the activity-type samples.

hypothesis is that all means being compared are equal. Simultaneous confidence intervals for all pairwise comparisons are formed by adding or subtracting the standard error multiplied by the value of a *t*-distribution at the level of confidence considered (here 99.99%). For this test, the standard error is calculated considering the unequal sample sizes and their heteroscedasticity (different variance for different samples). Details of this test are explained in Herberich, Sikorski & Hothorn (2010) and Hothorn, Bretz & Westfall (2008).

The figure shows these confidence intervals for all the differences of the mean morphological types of the samples defined using the BPT diagram and the sample of passive galaxies. Any confidence intervals including zero would indicate no significant difference for that pair of means. For all our activity-type samples, the differences in the mean inferred morphology are significant at least at the 99.99% confidence level.

J. P. Torres-Papaqui acknowledges DAIP-UG for support grant 432/14. R. A. Ortega-Minakata thanks DAIP-UG for support for travel expenses to attend the Symposium.

References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., *et al.* 2009, *ApJS*, 182, 543
 Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
 Cid Fernandes, R., Mateus, A., Sodré Jr., L., *et al.* 2005, *MNRAS*, 358, 363
 Fukugita, M., Nakamura, O., Okamura, S., *et al.* 2007, *AJ*, 134, 579
 Heckman, T. M. 1980, *A&A* 87, 152
 Herberich, E., Sikorski, J., & Hothorn, T. 2010, *PLoS ONE*, 5, 1 (Plos ONE 5(3): e9788.doi:10.1371/journal.pone.0009788)
 Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1997, *ApJ*, 487, 568
 Hothorn, T., Bretz, F., & Westfall, P. 2008, *Biom. J.*, 50, 346
 Kauffmann, G., Heckman, T. M., Tremonti, C., *et al.* 2003, *MNRAS*, 346, 1055
 Kewley, L. J., Dopita, M. A., Sutherland, R. S., *et al.* 2001, *ApJ*, 556, 121
 Sabater, J., Verdes-Montenegro, L., Leon, S., *et al.* 2012, *A&A*, 545, A15
 Shimasaku, K., Fukugita, M., Doi, M., *et al.* 2001, *AJ*, 122, 1238
 Torres-Papaqui, J. P., Coziol, R., Andernach, A., *et al.* 2012, *Rev. Mexicana AyA*, 48, 275