Powerful t-SNE Technique Leading to Clear Separation of Type-2 AGN and H II Galaxies in BPT Diagrams

XueGuang Zhang®, Yanqiu Feng, Huan Chen, and QiRong Yuan
School of Physics and Technology, Nanjing Normal University, No. 1, Wenyuan Road, Nanjing, 210023, People’s Republic of China; sgzhang@njnu.edu.cn
Received 2020 August 23; revised 2020 October 8; accepted 2020 October 20; published 2020 December 17

Abstract

Narrow emission-line galaxies can be distinguished in the well-known BPT diagrams through narrow emission-line properties. However, there are no boundaries visible to the naked eye between type-2 active galactic nuclei (AGN) and H II galaxies in BPT diagrams, besides the extreme dividing lines expected by theoretical photoionization models. Here, based on the powerful t-SNE technique applied to the local narrow emission-line galaxies in the Sloan Digital Sky Survey Data Release 15, type-2 AGN and H II galaxies can be clearly separated in the t-SNE determined two-dimensional projected map, and then the dividing lines can be mathematically determined in BPT diagrams, leading to charming harmonization of the theoretical expectations and the actual results from real observed properties. The results not only provide an interesting and robust method to determine the dividing lines in BPT diagrams through the powerful t-SNE technique, but also lead to further confirmation on previously defined composite galaxies more efficiently classified in the BPT diagram of [O III]/Hβ versus [N II]/Hα.

Unified Astronomy Thesaurus concepts: Emission line galaxies (459); Active galactic nuclei (16); H II regions (694)

1. Introduction

BPT diagrams, named after Baldwin et al. (1981) and the pioneer work that performed a detailed study on classifications of narrow emission-line galaxies (Veilleux & Osterbrock 1987), were first proposed in the 1980s to show different physical properties of optical narrow emission lines of tens to hundreds of extragalactic emission-line objects. There are currently millions of narrow emission-line galaxies. And based on different kinds of central activities, narrow emission-line galaxies can be well classified into two main types: type-2 active galactic nuclei (AGN); narrow emission-line AGN with central AGN activities and H II galaxies without central AGN activities. Although Type-2 AGN and H II galaxies have similar optical spectral features, properties of optical narrow emission line can be commonly applied to determine classifications of type-2 AGN and H II galaxies, such as the well-known results in ongoing improved BPT diagrams (Cid Fernandes et al. 2011; Juneau et al. 2014; Kashino et al. 2017; Kewley et al. 2019).

The BPT diagrams allow the division of narrow emission-line objects in two main branches: one containing H II galaxies and the other containing type-2 AGN (Seyfert galaxies and low ionization nuclear emission-line regions (LINERs)). Unlike Seyfert galaxies, which are totally powered by central AGN activities, different mechanisms have been applied to LINERs, such as AGN activities (Terlevich & Melnick 1985; Filippenko & Terlevich 1992), photoionization by post-asymptotic giant branch (post-AGB) stars (Binette et al. 1994; Eracleous et al. 2010; Cid Cid Fernandes et al. 2011), etc. A more recent review on LINERs in Marquez et al. (2017) has shown that 60% to 90% of LINERs could be well considered as genuine AGN. Therefore, in spite of controversial mechanisms, LINERs have been accepted as a subsample of type-2 AGN in the manuscript.

In the well-known BPT diagrams, between type-2 AGN and H II galaxies, there are no clear dividing boundaries visible to the naked eye. The reported dividing lines by flux ratios of optical narrow emission lines in the BPT diagrams are estimated and determined by expected properties of extreme starbursts and/or AGNs by theoretical photoionization models, such as the results well discussed in Kauffmann et al. (2003), Groves et al. (2004), Kewley et al. (2006), Stasinska et al. (2006), Levesque et al. (2010), Melendez et al. (2014), etc. More recently, de Souza et al. (2017) studied emission-line galaxy classifications through a probabilistic Gaussian mixture model applied to spectroscopic properties from the Sloan Digital Sky Survey Data Release 7 (SDSS DR7) and SEAGal/STARLIGHT data sets, and showed the Gaussian components relative to AGN and star-forming galaxies. However, the results discussed in de Souza et al. (2017) cannot yet lead to apparent dividing lines. Thereby, it has not yet been possible to determine the dividing lines from properties of the observed narrow emission lines of real narrow emission-line galaxies. Here, based on the more recent powerful t-SNE (distributed stochastic neighbor embedding) technique (van der Maaten & Hinton 2008; van der Maaten 2014; Arora et al. 2018), which has been recently applied in astrophysics (Traven et al. 2017; Anders et al. 2018; Steinhardt et al. 2020), we will show actual results of the dividing lines in BPT diagrams from real observed properties of the narrow emission lines of local narrow emission-line galaxies to check whether there is harmonization of theoretical expectations and actual results on the dividing lines between H II galaxies and type-2 AGN in the BPT diagrams. Then, our main results and necessary discussions are shown in Section 2, and conclusions are given in Section 3.

2. Main Results and Discussions

Based on properties of the optical narrow emission lines of main galaxies, the SQL query (http://skyserver.sdss.org/dr15/en/tools/search/sql.aspx) can be conveniently applied in SDSS DR15 (Aguado et al. 2019; Belfiore et al. 2019), leading to collected 35,857 local narrow emission-line objects with apparent narrow emission lines (line intensities at least 5
times larger than their corresponding measured errors) of Hα, Hβ, [O III]λ5007 Å, [O I]λ6300 Å, [N II]λ6583 Å, and [S II] λ6717, 6731 Å but no broad emission lines in high-quality SDSS spectra (signal-to-noise ratios larger than 10). The applied SQL query in detail is as follows:

```sql
SELECT plate, fiberid, mjd, z, snmedian, h_beta_flux, h_beta_flux_err, h_alpha_flux, h_alpha_flux_err, oii_5007_flux, oii_5007_flux_err, nii_6584_flux, nii_6584_flux_err, sii_6717_flux, sii_6717_flux_err, sii_6731_flux, sii_6731_flux_err, oii_6300_flux, oii_6300_flux_err
FROM GalSpecLine
JOIN SpecObjall ON S.specobjid = G.specobjid
WHERE S.class = 'GALAXY' AND S.SNmedian > 10
   AND S.z < 0.2 AND S.zwarning = 0
   AND G.h_beta_flux_err > 0
   AND G.h_beta_flux > 5*G.h_beta_flux_err
   AND G.h_alpha_flux > 5*G.h_alpha_flux_err
   AND G.oii_5007_flux > 0
   AND G.oii_5007_flux_err > 0
   AND G.nii_6584_flux > 5*G.nii_6584_flux_err
   AND G.nii_6584_flux_err > 0
   AND G.oii_6300_flux > 5*G.oii_6300_flux_err
   AND G.oii_6300_flux_err > 0
   AND G.sii_6717_flux > 5*G.sii_6717_flux_err
   AND G.sii_6717_flux_err > 0
   AND G.sii_6731_flux > 5*G.sii_6731_flux_err
   AND G.sii_6731_flux_err > 0
   AND S subclass = 'STARFORMING' OR S subclass = 'STARBURST'
```
Then, based on the collected properties of narrow emission lines of the 35,857 local narrow emission-line galaxies, beautiful BPT diagrams are shown in the top panels in Figure 1. It is clear there are no apparent boundaries visible to the naked eye between the expected type-2 AGN located in the top right regions and the expected H II galaxies located in the bottom left regions in the shown BPT diagrams. Therefore, it is a constructive challenge to determine the dividing lines through mathematical visualization techniques applied to the real observed narrow emission-line galaxies, besides the theoretical model determined ones.

The two well-known mathematical techniques have been accepted to do visualization and reduction of high-dimensional data, the commonly known PCA (principal component analysis) technique (Ian Jolliffe & Cadima 2016; Lever et al. 2017) and the more recent powerful t-SNE technique (van der Maaten & Hinton 2008; van der Maaten 2014; Arora et al. 2018). The PCA technique has been applied in astronomy for several decades, such as to obtain the results reported in Sodre & Cuevas (1997), Bailer-Jones et al. (1998), Warren et al. (2005), Re Fiorentin et al. (2007), and Steiner et al. (2009). The main idea behind the PCA technique through the deterministic algorithm without hyperparameters is the linear technique used to reduce the dimensionality of data that is highly correlated by transforming the original set of vectors to a new set known as principal components to preserve the global structure of the data. In the manuscript, the first two most important principal components (PCA dimension 1 and PCA dimension 2) are held by rotating the vectors for preserving variance. Meanwhile, the more recent powerful t-SNE technique involving hyperparameters is a nonlinear technique through a nondeterministic or randomized algorithm. The math behind t-SNE is quite complex but the idea is simple. It embeds the points from a higher dimension to a lower dimension trying to preserve the neighborhood of that point, preserving the local structure (cluster) of data. More specifically, the t-SNE technique minimizes the divergence between two distributions: a distribution that measures pairwise similarities of the high-dimensional data and a distribution that measures pairwise similarities of the corresponding low-dimensional points in the embedding. In the manuscript, the determined similarity of data points is shown in the corresponding reduced two-dimensional embedded space, i.e., the t-SNE dimension 1 and t-SNE dimension 2. Moreover, the commonly accepted hyperparameters applied in the t-SNE technique are perplexity, early exaggeration, learning rate, and number of steps, respectively. As we know, the powerful t-SNE technique is a particularly better technique than the PCA technique to perform visualization of high-dimensional data sets, the t-SNE technique can clearly lead to two clusters shown as blue dots for 28,688 objects and as red dots for 7169 objects in Figure 2. Meanwhile, the PCA technique cannot lead to similar results; a comparison of the results from the PCA technique and from the t-SNE technique is shown in Figures 4 and 5.

The two clusters determined by the t-SNE technique can strongly indicate intrinsic different physical properties in the BPT diagrams for the objects in the two clusters shown in Figure 2. Then, the narrow emission-line objects in the two clusters determined by the t-SNE technique are separately replotted in the BPT diagrams in bottom panels of Figure 1 with contour lines in bluish colors (for the objects shown as blue dots in Figure 2) and in reddish colors (for the objects shown as red dots in Figure 2), respectively. Then, based on the contours with contour lines in different colors in the BPT diagrams, it is interesting that the narrow emission-line objects in the two clusters shown in Figure 2 can be well distinguished in the BPT diagrams in the bottom panels of Figure 1, providing the chance to determine dividing lines between H II galaxies and type-2 AGN by independent mathematical methods.

In the BPT diagram of [O III]/Hβ versus [N II]/Hα, there are two well-accepted dividing lines between H II galaxies, composite galaxies, and type-2 AGN, one is applied to determine the outer boundary for extremely starburst galaxies, and the other one is applied to determine the pure H II galaxies without central AGN activities, the region between the two dividing lines is well defined for composite galaxies. Now, in the bottom left panel of Figure 1, the two previously defined dividing lines can be roughly compared with properties of the...
objects in the two clusters, considering the far-side outer boundaries for the objects in the two clusters. Meanwhile, the theoretical model determined dividing lines in the other two BPT diagrams can also be well compared with almost overlapped far-side boundaries of the objects determined by the t-SNE technique.

In order to construct a definite dividing line in the BPT diagram of $\text{[O III]}/\text{H}$ versus $\text{[N II]}/\text{H}$ between the objects in the two clusters determined by the t-SNE technique, a simple method has been applied. Following a series of parallel strips with directions parallel to the reference direction that linked the ridges of the two contour maps of the objects in the two clusters, the dividing line $Y_D$ is well built by the crossover points of the position dependent number ratios for the objects in the two clusters in the strips. The final dividing line $Y_D$ has been well determined and is shown in the bottom left panel of Figure 1. In the other two BPT diagrams, the reference directions are given by the direction that linked the ridge of the contour map in bluish colors and the gully of the other contour map in reddish colors, then leading to the t-SNE-technique-determined dividing lines $Y_D$. Moreover, similar results can be determined and confirmed in the BPT diagrams by giving slightly different reference directions, but following from the bottom left to the top right in the BPT diagrams. Here, Figure 3 shows an example of more detailed information on the crossover points for the objects in the two clusters in strips in the BPT diagrams.

Meanwhile, besides the t-SNE-technique-determined dividing lines $Y_D$ described by the two-degree polynomial functions shown in Figures 1 and 3, it will be very interesting to discuss properties of overlapped regions covered by both H II galaxies and type-2 AGN in the BPT diagrams. It is clear that when the objects in the two clusters are plotted in the BPT diagrams with their measured flux ratios of narrow emission lines (such as the results shown in the bottom panels of Figure 1), there are overlapped regions in the BPT diagrams for the objects in the two clusters determined by the t-SNE technique. Therefore, the overlapped regions in the BPT diagrams are very real, which will provide further clues on
properties of composite galaxies. Then, a simple mathematical idea is accepted to determine the overlapped regions between $Y_{\text{upper}} = Y_D + U$ and $Y_{\text{lower}} = Y_D - B$. Here, parameters of $U$ and $B$ are determined by the criteria that $1\%$ of H II galaxies are above $Y_{\text{upper}}$ and $1\%$ of type-2 AGN are below $Y_{\text{lower}}$ in the BPT diagrams of $[\text{N II}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$ and $[\text{S II}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$, but $0.5\%$ of H II galaxies are above $Y_{\text{upper}}$ and $0.5\%$ of type-2 AGN are below $Y_{\text{lower}}$ in the BPT diagram of $[\text{O I}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$ (mainly due to the following significantly smaller overlapped region). The well determined boundaries $Y_{\text{upper}}$ and $Y_{\text{lower}}$ are shown in Figure 1.

For the results shown in the bottom left panel of Figure 1, the overlapped region is determined between $Y_{\text{upper}} = Y_D + 0.28$ and $Y_{\text{lower}} = Y_D - 0.23$ in the BPT diagram of $[\text{N II}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$, meaning that $99\%$ of H II galaxies are locating below the upper boundary of $Y_{\text{upper}}$ and $99\%$ of type-2 AGN are located above the lower boundary of $Y_{\text{lower}}$. Therefore, we can simply say that the overlapped regions defined by the upper and lower boundaries of $Y_{\text{upper}}$ and $Y_{\text{lower}}$ represent the mixed region of H II galaxies and type-2 AGN with a significant confidence level of $99\%$. Furthermore, the region between $Y_{\text{upper}}$ and $Y_{\text{lower}}$ has interestingly similar covered space as the region between the reported dividing lines in Kewley et al. (2001, 2006) and in Kauffmann et al. (2003). However, there are no reports on composite galaxies in the BPT diagrams of $[\text{S II}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$ and $[\text{O I}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$ in the literature. We directly accepted the same criterion applied in the BPT diagram of $[\text{N II}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$. Then, the upper and lower boundaries of $Y_{\text{upper}} = Y_D + 0.075$ and $Y_{\text{lower}} = Y_D - 0.08$ have been determined and shown in the bottom-middle panel of Figure 1 in the BPT diagram of $[\text{S II}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$. Meanwhile, the overlapped region by the same criterion in the BPT diagram of $[\text{O I}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$ is too small to be shown. Therefore, a smaller percentage of $0.5\%$ rather than $1\%$ is applied to determine the upper and lower boundaries of $Y_{\text{upper}} = Y_D + 0.05$ and $Y_{\text{lower}} = Y_D - 0.07$ shown in the BPT diagram of $[\text{O I}]/\text{H}\alpha$ versus $[\text{O III}]/\text{H}\beta$ in the bottom right panel of Figure 1.

It is very interesting that there is a larger overlapped area between $Y_{\text{upper}}$ and $Y_{\text{lower}}$ for composite galaxies in the BPT...
diagram of \([\text{N II}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\), but much smaller overlapped areas in the BPT diagrams of \([\text{O I}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\) and \([\text{S II}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\). Therefore, the BPT diagram of \([\text{N II}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\) is more efficient and powerful to classify the composite galaxies than the other two BPT diagrams. Moreover, the defined composite galaxies in the BPT diagram of \([\text{N II}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\) cannot be totally in the right places for composite galaxies in the other two BPT diagrams of \([\text{O I}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\) and \([\text{S II}]/\text{H}\alpha\) versus \([\text{O III}]/\text{H}\beta\). The different properties of composite galaxies in different BPT diagrams reflect interesting but different dependence of forbidden emission lines on intrinsic different kinds of activities from AGN or from star-forming galaxies.

Finally, there have been more detailed discussions of robust results obtained through the t-SNE technique. On the one hand, the results shown in Figures 1, 2, and 4 show results similar to those confirmed through the t-SNE technique applied with different hyperparameters to the same four-dimensional data of narrow line ratios of the collected 35,857 narrow emission-line galaxies. Here, the accepted hyperparameters reapplied in the t-SNE technique are \([50, 10, 80, 1000]\) for perplexity, early exaggeration, learning rate, and number of steps, respectively. Two definite data clusters can be reconfirmed and similar final results can be reconfirmed in the BPT diagrams as the results shown in Figure 1. On the other hand, in order to reduce computational complexity, especially to reduce probable effects of the much different number densities of the objects in different regions of BPT diagrams, the BPT diagrams have been well smoothed leading to the almost even number densities in different regions with a total of 14,554 objects collected from the 35,857 objects, which are shown in the bottom panels of Figure 5. Then, the similar t-SNE and PCA techniques are applied to perform the data reduction and visualization of the four-dimensional data of the smoothed 14,554 narrow emission-line galaxies. Here, the accepted hyperparameters applied in the t-SNE technique are \([150, 5, 100, 800]\) for perplexity, early exaggeration, learning rate, and number of steps, respectively. The top panels of Figure 5 show the 2D projections by the t-SNE and PCA techniques. Two clusters can be well confirmed by the t-SNE technique. The bottom panels of Figure 5 show properties of the objects in the
two clusters and then the dividing lines applied in the BPT diagrams for the 14,554 narrow emission-line galaxies.

Before the end of the section, there is one point we should note. As we have known, there are two subsamples included in type-2 AGN: Seyfert galaxies and LINERs. Moreover, dividing lines have been reported between LINERs and Seyfert galaxies in the BPT diagrams, such as the results shown in Kewley et al. (2006) and in Schawinski et al. (2007). However, as the results discussed in de Souza et al. (2017), the Gaussian mixture model cannot provide statistical evidence for the existence of a Seyfert/LINER dichotomy for the emission-line galaxies from SDSS DR7. Moreover, as shown by the results of the t-SNE technique in Figure 2, in the top left panel of Figure 4, and in the top left panel of Figure 5, we cannot yet find a third data cluster determined by the t-SNE technique. More efforts are necessary to construct new high-dimensional data including more valuable information. Therefore, at the current stage, we cannot provide clear clues on the dividing lines between LINERs and Seyfert galaxies through the t-SNE technique, and we do not show any further discussion on dividing lines between LINERs and Seyfert galaxies in the manuscript.

3. Conclusions

BPT diagrams are powerful tools for classifying narrow emission-line galaxies with different central activity properties. However, how to define the dividing lines in the BPT diagrams between different kinds of narrow emission-line galaxies is always an interesting challenge. Here, we can find well-defined dividing lines through the pure mathematical t-SNE technique applied to the local narrow emission-line galaxies in SDSS DR15. The results not only show the charming harmonization of the theoretical expectations and the actual results from real observed properties through the powerful t-SNE technique, but also provide further confirmation on classification of the composite galaxies more efficiently in the BPT diagram of [O III]/Hβ versus [N II]/Hα.

The authors gratefully acknowledge the anonymous referee for providing constructive comments and suggestions that greatly improved the paper. X.-G.Z. is grateful for the grant support from Nanjing Normal University and the support from grant NSFC-11973029. Q.-R.Y. is grateful for the support from grant NSFC-11873032. This paper has made use of data from the SDSS projects. The SDSS-III website is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration.

ORCID iDs

XueGuang Zhang @ https://orcid.org/0000-0001-5458-8354

References

Aguado, D. S., Ahumada, R., Almeida, A., et al. 2019, ApJS, 240, 23
Anders, F., Chiappini, C., Santiago, B. X., et al. 2018, A&A, 619, 125
Arora, S., Hu, W., & Kothari, P. K. 2018, in Proc. Machine Learning Research 75, Proc. 31st Conf. on Learning Theory, 1455, http://proceedings.mlr.press/v75/arora18a.html
Bailer-Jones, C. A. L., Irwin, M., & von Hippel, T. 1998, MNRAS, 298, 361
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Belfiore, F., Westfall, K. B., Schaefer, A., et al. 2019, AJ, 158, 160
Binette, L., Magris, C. G., Stasińska, G., & Bruzual, A. G. 1994, A&A, 292, 13
Cid Fernandes, R., Stasińska, G., Mateus, A., et al. 2011, MNRAS, 413, 1687
de Souza, R. S., Dantas, M. L. L., Costa-Duarte, M. V., et al. 2017, MNRAS, 472, 280
Dopita, M. A., & Sutherland, R. S. 1995, ApJ, 455, 468
Dopita, M. A., & Sutherland, R. S. 1996, ApJS, 102, 161
Erelementous, M., Hwang, J. A., & Flihop, H. M. L. G. 2010, ApJ, 711, 796
Ferland, G. J., & Netzer, H. 1983, ApJ, 264, 105
Filippenko, A. V., & Terlevich, R. 1992, ApJL, 397, 79
Groves, B. A., Dopita, M. A., & Sutherland, R. S. 2004, ApJS, 153, 75
Halpern, J. P., & Steiner, J. E. 1983, ApJL, 269, 37
Heckman, T. M. 1980, A&A, 87, 152
Jolliffe, I. T., & Cadima, J. 2016, RSPTA, 374, 20150202
Juneau, S., Bournaud, F., Charlot, S., et al. 2014, ApJ, 788, 88
Kashino, D., Silverman, J. D., Sanders, D., et al. 2017, ApJ, 835, 88
Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055
Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121
Kewley, L. J., Groves, B., Kauffmann, G., et al. 2006, MNRAS, 372, 961
Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, ARA&A, 57, 511
Lever, J., Krzywinski, M., & Altman, N. 2017, NatMe, 16, 641
Levesque, E. M., Kewley, L. J., & Larson, K. L. 2010, A&ARv, 17, 105
Levesque, E. M., Weaver, J. R., Maxfield, J., et al. 2005, ApJ, 634, 376
Juneau, S., Bournaud, F., Charlot, S., et al. 2014, ApJ, 788, 88
Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055
Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121
Kewley, L. J., Groves, B., Kauffmann, G., et al. 2006, MNRAS, 372, 961
Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, ARA&A, 57, 511
Lever, J., Krzywinski, M., & Altman, N. 2017, NatMe, 16, 641
Levesque, E. M., Kewley, L. J., & Larson, K. L. 2010, A&ARv, 17, 105
Marquez, I., Masegosa, J., Gonzalez-Martín, O., et al. 2017, FrASS, 4, 34
Melendez, M., Heckman, T. M., Martinez-Paredes, M., Kruener, S. B., & Mendoza, C. 2014, MNRAS, 443, 1358
Re Fiorentin, P., Bailer-Jones, C. A. L., Lee, Y. S., et al. 2007, A&A, 467, 1333
Schawinski, K., Thomas, D., Sarzi, M., et al. 2007, MNRAS, 382, 1415
Sodre, L., & Cuevas, H. 1994, A&A, 292, 13
Stasinska, G., Cid Fernandes, R., Mateus, A., et al. 2006, MNRAS, 371, 972
Steiner, J. E., Menezes, R. B., Ricci, T. V., & Oliveira, A. S. 2009, MNRAS, 395, 64
Steinhardt, C. L., Weaver, J. R., Maxfield, J., et al. 2020, ApJ, 891, 136
Terlevich, R., & Melnick, J. 1985, MNRAS, 213, 841
Traven, G., Matijevic, G., Zwitter, T., et al. 2017, ApJS, 228, 24
van der Maaten, L. 2014, J. Mach. Learn. Res., 15, 3221
van der Maaten, L., & Hinton, G. 2008, J. Mach. Learn. Res., 9, 2579
Terlevich, R., & Melnick, J. 1985, MNRAS, 213, 841
van der Maaten, L. 2014, J. Mach. Learn. Res., 15, 3221
van der Maaten, L., & Hinton, G. 2008, J. Mach. Learn. Res., 9, 2579
Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295
Warren, J. S., Hughes, J. P., Badenes, C., et al. 2005, ApJ, 634, 376

Zhang et al.