Systematics in the SED Fitting Parameter Estimation of Merging Galaxies

Katherine Zine\(^1\) and Samir Salim\(^1\)

\(^{1}\)Department of Astronomy, Indiana University, Bloomington, IN 47405, USA

ABSTRACT

Derivation of physical properties of galaxies using spectral energy distribution (SED) fitting is a powerful method, but can suffer from various systematics arising from model assumptions. Previously, such biases were mostly studied for individual galaxies. Here we specifically study a large sample (9000) of composite (merged) galaxies extracted from GALEX-SDSS-WISE Legacy Catalog (GSWLC) in order to check if SED fitting accurately determines the properties (stellar mass and star formation rate) of an unresolved merged galaxy. We carry out the assessment by artificially merging the galaxies that could merge in reality and running UV/optical SED fitting on them as well as on pre-merger individual galaxies (to establish ground truth). With this simple approach we know what mass and SFR to ideally expect for the merger. No biases are found when comparing the stellar mass, star formation rate (SFR), and specific star formation rate (sSFR) values from the SED fitting to the combined values from the individual galaxies, although there are tails in the SFR and sSFR merger comparisons with low (s)SFR. Furthermore, no significant residuals were found when focusing on galaxies with different mass ratios (major vs. minor mergers), different contrasts in SFR, or dust content. We conclude that the SED fitting of post-mergers does not suffer from significant biases that would arise from the fact that the merger is a composite of two galaxies with potentially very different star formation histories and with different dust properties.

Keywords: Galaxy Mergers (608) — Galaxy Properties (615) — Galaxy Masses (607) – Spectral Energy Distribution (2129)

1. INTRODUCTION

Galaxy mergers play an important role in galaxy evolution. They can lead to dramatic changes in galaxy structure and morphology. For example, when simulating the life of a galaxy undergoing a merger, the galaxies that start off as disk galaxies can become spheroids after a merger (Wuyts et al. 2009). Mergers can also augment star formation efficiency and produce starbursts (Bournaud 2011). The study of galaxy evolution requires unbiased determination of galaxy properties, such as the stellar mass and the star formation rate (SFR). Since mergers can result from the combination of different types of galaxies, each with its own distinct star formation history (SFH) prior to merging, it is important to ensure that their properties are derived without potential systematics or biases.

Spectral energy distribution (SED) fitting is an important tool in the study of galaxy evolution due to its ability to simultaneously determine many properties of a galaxy, such as star formation history and stellar mass, based on its light output (Conroy 2013). In the context of SED fitting, the SEDs are crude spectra of galaxies consisting of flux measurements at different wavelengths, usually spanning from rest frame UV to near IR (Walcher et al. 2011). In SED fitting a large number of models are created by specifying parameters relating to SFH, stellar populations, and dust attenuation (Salim et al. 2016, hereafter S16). The observed SED is compared to model SEDs to find the best fit, or more generally to build the probability distribution functions for different properties (Walcher et al. 2011). The models used in SED fitting are based on stellar population synthesis, which uses stellar evolution theory and libraries of observed stars of different spectral types or model stars with different physical properties (Conroy 2013).

Whereas the observed SEDs are the result of an underlying physics that is being modeled, the models have limitations or simplifications that may affect the accuracy of parameters derived from the SED fitting. Furthermore, the observed SEDs, which have some finite precision, may not be able to accurately distinguish between different models (the degeneracy). Possible systematics involved with SED fitting have been studied previously, but usually for the general population of
Zine & Salim

2

galaxies, which is not likely to have undergone major merging events in the recent past. For example, for the stellar mass of galaxies, Sorba & Sawicki (2018) found that in the case of photometry that was spatially unresolved, the mass found using SED fitting was under the true mass. van der Wel et al. (2006) found that there can be systematics at different redshifts depending on various factors, such as the stellar population synthesis model employed. Mitchell et al. (2013) found systematics in stellar mass caused by using only one metallicity value and depending on how mass recycling was treated. Systematics were also found by Simha et al. (2014) in colors and mass-to-light ratios when using an exponentially declining model for star formation history. S16 has discussed how two-component star formation histories, where one component is posited to be very old, are able to recover larger stellar masses and remove or reduce the outshining issue (Micha lowski et al. 2014). Yet, another set of systematics emerges from the treatment of dust. For example, the use of fixed dust attenuation curves can lead to biased SFRs and stellar masses, especially if the assumed law differs from the average law of observed galaxies (Salim & Narayanan 2020; Lo Faro et al. 2017).

Recent work has enabled us to identify and understand many of the aforementioned biases, especially in the context of individual galaxies. However, when two galaxies merge such that we can no longer separate the SEDs of the original galaxies, a different set of systematics may arise in the SED fitting due to the fact that each galaxy entering a merger may have a distinct SFH and different dust properties. Wuyts et al. (2009) studied how systematics may arise during different phases of a merger. They created a simulation that progressed over time and found that during the period of enhanced star formation due to a merger the age and mass were being most underestimated, whereas the underestimation of SFR varied at different points during the merger (Wuyts et al. 2009). Hayward & Smith (2015) applied MAGPHYS SED fitting code to hydrodynamical simulations of merging galaxies and found good fits and that most of the parameters were recovered correctly, with the exception of the dust mass when the galaxy was not at near-coalescence. Lanz et al. (2014) found that SEDs cannot usually be used to indicate the interactions stage of galaxies, but this does not necessarily mean that the SEDs of merged galaxies are not at least somewhat distinct from the galaxies that have not experienced a merger (at least recently) and that such distinct SEDs would still be well modeled in the course of SED fitting (Lanz et al. 2014).

While previous studies looked at individual simulations of merging galaxies at different points in time and whether the results matched what was expected, another aspect that was so far not explored would be to examine whether the parameter values found for the merged galaxy at the end of the merger agree with what would be obtained by the addition of the values of the individual galaxies pre-merger. The advantage of this approach, which we will exploit in the current study, is that it does not require expensive simulations and that we can utilize a much larger sample of mock mergers with a wide range of mass ratios and SFR contrasts.

The paper is organized as follows. In Section 2 the data and sample used for this study are described. Section 3 describes methodology used to assess biases in the SED fitting of composite (merged) galaxies. The results are presented in Section 4 and a discussion of the results is in Section 5.

2. DATA AND SAMPLE

The parent sample for this project was taken from the GSWLC-D catalogue (S16). GSWLC is a catalog of galaxy parameters obtained from SED fitting, and consists of galaxies contained in the SDSS spectroscopic sample (redshifts between 0.01 and 0.30) that had ultraviolet (UV) coverage from GALEX (S16). The D catalog is the catalog with the longest exposure times in the UV (longer than 4000 seconds) and therefore has the highest precision of of SFR estimates (S16). It contains position in the sky, SDSS identifications, and physical parameters such as stellar mass, dust attenuation, and

![](image)

Figure 1. Physical and redshift separations of galaxies in candidate pairs. The physical separation in Mpc for the two galaxies in the 33,557 non duplicate pairs plotted against their difference in redshift. Only galaxy pairs between the two red lines were decided to be close enough to be considered potential future mergers and were kept in the analysis.

![Figure 1](image)
Figure 2. Distributions of galaxy properties before “merging” (individual galaxies in merging pairs, blue) and after “merging” (composite galaxies, orange). Properties include the stellar mass, star formation rate, specific star formation rate, and V-band attenuation.

star formation rates of 48,401 galaxies (S16). GSWLC is based on photometry in the FUV, NUV, u, g, r, i, and z bands, which will be used in the SED fitting in this work as well (S16).

One can in principle test how SED fitting performs on composite galaxies (mock mergers) by taking any two galaxies from the parent sample and combining their light. However, in order to make our assessment more pertinent, we focus on true galaxy pairs—galaxies that are close enough to each other that they could realistically merge. The added benefit of such an approach is that each galaxy in the composite will lie at the same distance, so combining their light can be accomplished by simply adding their fluxes.

We perform galaxy pair selection using a two-step process. First, we identify 33,557 candidate pairs, which are defined as a pair with the smallest angular separation from each other not exceeding one degree. In order to establish if the candidate pair is physically associated, we examine their redshifts. In Figure 1 we show the physical projected separation of candidate pairs where the galaxies have similar redshifts. We see an overdensity of pairs around $\Delta z = 0$. Informed by the appearance of this distribution, for our final sample we require the difference in redshifts to be less than 0.003, corresponding to a maximum velocity of 900 km $s^{-1}$. The pairs with close redshifts are all within about 2 Mpc projected in the sky (the size of galaxy groups), and often much closer, so we do not impose any cuts on the projected separation. The final sample consists of 9,032 pairs with 15,736 unique galaxies in the pairs, or 32.5% of the parent sample. The number of unique galaxies is somewhat smaller than twice the number of pairs because a given galaxy can be a member of more than one pair.
3. SED FITTING

For the 9,032 pairs of close galaxies we combine the fluxes for each pair to imitate what would happen if the galaxies were recently merged. The fluxes were simply added together so their light output would be combined. For the fitting, the observed errors were added in quadrature with a calibration error ranging from 1-2% for SDSS bands and 3-5% for GALEX (S16). These errors were used instead of 10% calibration errors assumed by default in CIGALE. Finally, the redshift used in the fitting was taken to be the average redshift of the two galaxies.

SED fitting was run with the 2020 version of the Code Investigating GALaxy Emission, or CIGALE (Boquien et al. 2019). We perform SED fitting using Bruzual & Charlot (2003) stellar population synthesis models calculated for four stellar metallicities and for the Chabrier IMF. Nebular emission lines are included in the models as described in S16. Importantly, we use several attenuation curves, again following S16. In particular, we use the Noll et al. (2009) modification of the Calzetti et al. (2000) attenuation curve, with two slopes each steeper than the original Calzetti curve. To this curve we add the UV bump, up to twice the strength of the UV bump in the MW. Note that this SED fitting does not include IR constraints as the more recent GSWLC-2 does (Salim et al. 2018). The principle difference between the SED fitting employed here and in S16 regards the parameterization of the star formation history. Namely, prior to running the final fits, we tested the performance of the SED fitting in order to find the parameters and variables that would potentially improve fitting over S16 judging by average $\chi^2$. We took 1000 galaxies from GSWLC-D and performed the fitting with different SFH parameter combinations and we found that some improvements were possible, as explained here.

S16 uses a double exponential, where the first (older) exponential commences at a fixed age (10 Gyr before the epoch of observation) whereas the second component is essentially flat in the SFR (with a very long decline rate) but varies in age and intensity. Here we used the delayed exponential model from Appendix E of Boquien et al. (2019), but with the same age as the old exponential component previously (10 Gyr) and without the extremely short decline models of 1Myr. To this smooth delayed exponential we add the second, exponentially declining, component, again after the Appendix E. Unlike the second component in S16, this one includes various short decline rates and ages. We find that the new SF parameterization has somewhat smaller $\chi^2$ values for the best fitting models compared to S16, although the improvement is not dramatic.

The fluxes of the original 48,401 galaxies were run through CIGALE in order to obtain the values that the CIGALE SED fitting gave for the galaxies individually. The combined fluxes of the 9,032 pairs were also run through CIGALE with the same parameters to obtain the merged values.

4. RESULTS

4.1. Physical properties of parent sample and galaxy pair sample

We first explore how the properties of composite ("merged") galaxies compare to the sum of the two galaxies before they were merged. This is done for the star formation rates (SFR), the specific star formation
rates (sSFR), and the V-band attenuation and is shown in Figure 2. For the stellar mass the distribution shifts by some 0.5 dex as expected if most mergers are major. For SFR values, the values obtained from the fitting for the merged galaxies have a pronounced high-SFR peak, whereas the low SFR is suppressed, suggesting that there are fewer mergers among the quiescent galaxies (dry mergers). For the sSFR however, the merged values have a peak between the two peaks for the values of the pre-merged galaxies, which is also as expected given that many of the galaxy pairs are a combination of a passive plus star forming galaxy, so the result of a merger would move the galaxy toward the green valley. Indeed it has been proposed that many of the galaxies in the green valley are the result of such mergers (Kaviraj et al. 2009). The two histograms for the V-band attenuation look similar and don’t have much of an offset.

The stellar mass ratios of the galaxies involved in each mock merger are shown in Figure 3. There are many more major mergers than minor mergers, which is expected because most of the pairs are at redshifts where only massive galaxies (log $M_\ast > 10$) are detectable in the SDSS, so the difference between the galaxies in a given pair cannot be very high. Having a large fraction of major mergers is a welcome aspect of this sample because it is the major mergers in which any biases would be accentuated.

The plot of the stellar mass versus the specific star formation rate was also examined (Figure 4). The plot shows all 48,401 galaxies from the parent sample. On top of that are points for the merged galaxy pairs. The merged pairs have fewer low mass galaxies than the parent sample (low mass galaxies live in underdense environments and are unlikely to merge) and have a tail to lower specific star formation rates. We will see subsequently that these lower values are biased.

4.2. Biases in parameters of composite (“merged”) galaxies

In order to explore potential biases in parameters recovered for mock mergers, expected and recovered values were compared to see how well SED fitting on the composite galaxies agreed with the expected values. The expected values (“ground truth” in this comparison) were obtained by adding the parameter values for the individual galaxies, except in the case of the sSFR, where the sum of the SFRs divided by the sum of stellar masses is used. Figure 5 shows comparisons of the expected versus recovered stellar masses, star formation rates, and specific star formation rates for the merged pairs of only the major mergers. The left most plot compares the expected mass value for the merged pair to the value recovered from the CIGALE fitting on the merged pair. The plot of the mass comparison shows that the data points follow the one to one line very well, demonstrating that there is no overall bias in the stellar mass. There is also very little scatter (0.04 dex), comparable to or even smaller than the formal error in individual masses as estimated from the width of the probability distribution functions. For the star formation rate comparison histogram in the middle, there does not seem to be a bias for log SFR $> 0$, but there is a tail around log SFR (expected) = −1, in the sense that the recovered SFRs are lower. The histogram on the right for specific star formation rate is mostly along the one to one line but there is a tail around log sSFR (expected) = −12.5 where again the recovered values are somewhat lower.

Whereas the direct comparison does not reveal any major biases (except for low SFRs and sSFRs), it is
still of interest to explore if any residuals emerge for particular subsets of mock mergers. For example, one might expect that the biases may increase for mergers that have more comparable mass ratios (major mergers). In Figure 6 we show stellar mass residuals between expected and recovered mass as a function of the contrast in $V$-band attenuation, the reduced $\chi^2$ value, the stellar mass, and the SFR. The median value of the residuals is shown by a red line. Residuals stay very close to zero (white line) in all four plots, showing that there are no systematic stellar mass residuals even for cases where one galaxy in the pair has a very different dust attenuation from the other, where mergers involve very equal mass galaxies, or where one fit is much worse than the other.

The second set of figures in Figure 7 shows residuals as a function of the contrast between the same parameters used in Figure 6, but now for SFR. The red line again shows the median value of the residuals. We see that the residuals are not driven by increasing mass contrast or the difference in the quality of the fits (upper right and lower left panels), but they do increase somewhat when one galaxy has much less dust than another or when a passive and active star forming galaxy are combined. However, even in those cases the residuals are around 0.1 dex, which is comparable to or smaller than a typical SFR random error (0.1 to 0.6 dex, Figure 6 in S16).

5. DISCUSSION

As shown in Section 4, no major systematics were found in the determination of principle galaxy properties of post-mergers. The pre- and post-merger figures in Figure 2 agree well with the the current understanding of galaxy mergers and agree with the fact that most of the mergers in the sample were major mergers. In Figure 5, the recovered values match the expected values well, with median values close to zero and with the largest standard deviations being 0.55. No biases are seen for the stellar mass comparison, however the SFR and sSFR comparisons both have biased tails at log SFR (expected) = $-1$ and log sSFR (expected) = $-12.5$. The mass residuals do not depend much on the contrast of the $V$-band attenuation, reduced $\chi^2$ value, stellar mass, and SFR, whereas the SFR residuals are somewhat sensitive to dust and SFR contrast. Overall, regardless of the different dust properties and SFH, the parameters of the composite merger do not suffer from significant biases in the SED fitting.
Figure 7. Difference in the recovered and expected SFR for major mergers (the residuals) as a function of the contrast between the members of the merging pair, in $V$-band attenuation, stellar mass, reduced $\chi^2$, and star formation rate. The white line is zero residual and the red line is the median value of the residuals.

Only the main physical parameters were explored in this study and there may be other parameters which would be more affected by biases in the SED fitting of recently merged galaxies. On the other hand, here we explore just the more usual SED fitting involving stellar emission. The inclusion of dust emission in the infrared, as is done in the so called "energy balance" SED fitting, may help further reduce any systematics (da Cunha et al. 2008; Salim et al. 2018). The approach presented in this study, using artificial mergers with empirically determined ground truth, offers new opportunities for perfecting SED fitting as a research tool.

The construction of GSWLC was funded through NASA awards NNX12AE06G and 80NSSc20K0440.

6. REFERENCES

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103. doi:10.1051/0004-6361/201834156
Bournaud, F. 2011, EAS Publications Series, 51, 107. doi:10.1051/eas/1151008
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000. doi:10.1046/j.1365-8711.2003.06897.x
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682. doi:10.1086/308692
Conroy, C. 2013, ARA&A, 51, 393. doi:10.1146/annurev-astro-082812-141017
Hayward, C. C. & Smith, D. J. B. 2015, MNRAS, 446, 1512. doi:10.1093/mnras/stv2195
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595. doi:10.1111/j.1365-2966.2008.13535.x
Kaviraj, S., Peirani, S., Khochfar, S., et al. 2009, MNRAS, 394, 1713. doi:10.1111/j.1365-2966.2009.14403.x
Lanz, L., Hayward, C. C., Zezas, A., et al. 2014, ApJ, 785, 39. doi:10.1088/0004-637X/785/1/39
Lo Faro, B., Buat, V., Roehlly, Y., et al. 2017, MNRAS, 472, 1372. doi:10.1093/mnras/stx1901
Michałowski, M. J., Hayward, C. C., Dunlop, J. S., et al. 2014, A&A, 571, A75. doi:10.1051/0004-6361/201424174
Mitchell, P. D., Lacey, C. G., Baugh, C. M., et al. 2013, MNRAS, 435, 87. doi:10.1093/mnras/stt1280
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793. doi:10.1051/0004-6361/200912497
Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, ApJS, 227, 2. doi:10.3847/0067-0049/227/1/2
Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11. doi:10.3847/1538-4357/aabf3c
Salim, S. & Narayanan, D. 2020, ARA&A, 58, 529. doi:10.1146/annurev-astro-032620-021933
Simha, V., Weinberg, D. H., Conroy, C., et al. 2014, arXiv:1404.0402
Sorba, R. & Sawicki, M. 2018, MNRAS, 476, 1532. doi:10.1093/mnras/sty186
van der Wel, A., Franx, M., Wuyts, S., et al. 2006, ApJ, 652, 97. doi:10.1086/508128
Walcher, J., Groves, B., Budavári, T., et al. 2011, Ap&SS, 331, 1. doi:10.1007/s10509-010-0458-z
Wuyts, S., Franx, M., Cox, T. J., et al. 2009, ApJ, 696, 348. doi:10.1088/0004-637X/696/1/348