Warsaw, 04 September 2024

dr hab. Maciej Bilicki, prof. CFT PAN
Center for Theoretical Physics
Polish Academy of Sciences
al. Lotników 32/46
02-668 Warsaw, Poland

Review of the doctoral thesis
by Luis E. Suelves
entitled
Search for galaxy mergers in big sky surveys

The doctoral thesis by Mr. Luis E. Suelves, written under the supervision of Prof. Agnieszka Pollo and Dr. William J. Pearson, deals with the problem of identifying merging galaxies in today's big sky surveys. Such mergers are a stage in the evolution of some galaxies: if two or more of them interact via gravity, they can approach each other sufficiently close to start merging. In the process, their morphologies become disrupted and often the end result is a new galaxy formed from the two original ones. This mechanism is for instance believed to be responsible for the formation of large elliptical galaxies, which likely emerge from smaller ones, e.g. spiral, merging. Such fate is also expected for the Milky Way – M31 pair, as they are approaching each other. This alone shows the importance of galaxy mergers for today's astrophysics. In the era of large sky surveys that this thesis focuses on, galaxy mergers will be found in ever growing numbers, therefore new and more efficient methods to robustly identify them are needed. Several such methods are here proposed and analyzed by the Candidate.

The dissertation is devoted to practical aspects of merger identification and it presents methods to search for them in two sky surveys: the Sloan Digital Sky Survey (SDSS) and Subaru/Hyper-Suprime Cam North Eclicptic Pole (HSC NEP) survey. Various approaches of different levels of sophistication are here discussed and used for merger finding: neural networks, visual inspection, decision trees, dimensionality reduction, as well as detailed analysis of particular parameters of galaxy flux measurements and their errors. The presented results of Candidate's studies provide important contributions to the field of galaxy merger studies.

The text has the form of a stand-alone dissertation[1], it is however mentioned that Chapter 4 is based on the publication from 2023, where the Candidate is the first author and his supervisors are co-authors. Chapter 5 presents yet unpublished work, while, as declared, Chapter 6 summarizes Candidate's contribution to the publication Pearson et al. 2022, where Mr. Suelves is the second author, but also includes extension thereof that is planned to be published as well. The other Chapters cover the Introduction (#1), Data (#2), Methods (#3) and Summary (#7). The thesis also includes an appendix detailing the discussion from Ch.4.

I find this dissertation to be very technical. It presents exquisite details of image and catalog processing and analysis. While I have no doubts that the work done by the Candidate leads to advancements in the field, overall I was lacking some stronger emphasis on the underlying physics. What is more, some relevant general questions are only answered partially: Why are galaxy mergers

1 *Pl.* "monografia naukowa"

important for astrophysics? What is their physical mechanism? How to robustly identify them? What could be the consequences of incompleteness or impurity of catalogs of prospective mergers?

Chapter 1 is the introduction. This is where "Candidate's general theoretical knowledge in the discipline[2]" is normally presented, and in particular one could expect to find the physical context of the work discussed later, and discussion of such questions as I listed above. The Candidate does make an effort to cover these items in this chapter. However, generally I find it difficult to follow and in particular I had to look for these aspects somewhat "between the lines". What is more, the physical context is very much lacking and at times the discussions are scattered among the various sections. Some statements are inconsistent or even inaccurate, and the presented knowledge is not always up to date. This in particular affects sections from 1.1 to 1.3. Some examples of the issues are:

- As discussed in sec.1.1.1.1, plethora of evidence indicates that dark matter dominates over luminous matter in our Universe in terms of mass. I imagine that the presence and distribution of dark matter must be then an important, if not the dominant, factor in galaxy merger process. However, while that section is devoted to dark matter (with some passages rather unrelated to the contents of this thesis – such as detailed description of BAOs), I did not find any mention whatsoever of the possible role of dark matter in galaxy mergers. This leaves me with an impression as if galaxy mergers were entirely a baryonic process and the dark matter haloes and their structure had no influence on them whatsoever.

- In some parts I find out-of-context discussions, and some with references covering mostly those from the Candidate's research group: e.g. on SED fitting or on dust in galaxies in sec.1.1.1.2. None of these seems much related to the rest of the thesis. In fact, most of the sec.1.1.1.2 appears rather irrelevant, for instance the passages about pulsars, black holes, quasars or planets.

- In contrast, sec.1.1.2 on galaxy classification is very relevant for this dissertation. Unfortunately, I find its contents rather superficial and outdated. Are the "Hubble tuning fork" or de Vaucouleurs' classifications related to some actual galaxy evolutionary paths? If yes, how? Are these the currently adopted galaxy classifications and what physics underlies them? How are they related to galaxy mergers – the topic of this dissertation?

- In sec.1.1.1.2, how is this general separation of galaxies into "red" and "blue" related to galaxy mergers, where to locate irregular or peculiar galaxies in diagrams such as fig.1.4?

- By the end of sec.1.1.3 we find the passages:
"...one of the other main influences in galaxy evolution come from the actual collapse of galaxies with each other, from galaxy mergers."
and
"Thus, galaxy mergers are a crucial stage in mass assembly of galaxies, although it is not yet known to what extended. Furthermore, they have also shown to have a significant effect in the star formation, morphological transformation, and many other important attributes of galaxies."
Unfortunately, no further discussion on these statements is offered nor references are given, it is therefore difficult to appreciate how galaxy mergers are important for galaxy growth and evolution.

- Sec.1.2 "Galaxy Mergers" is arguably the most important in the Introduction. However, the merging process is presented here with very minimal physical background and supported by a few

2 *"Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie"*

very old references. Has there been no progress in understanding the mergers since Toomre & Toomre 1972 or Quinn 1984? What is the role and importance of the dark matter and baryonic components in galaxy mergers? Last but not least, why are galaxy mergers important for astrophysics in general?

- I find the conclusion “galaxy interactions are less likely to result in a merger within environments more dense in number of galaxies, where the galaxies have a higher relative velocity” somewhat counter-intuitive. In dense environments, the potential wells (e.g. of central galaxies) should be deeper than in the field. This should lead to more frequent galaxy infall into these wells and hence, more frequent galaxy mergers. What has led to a different conclusion?

- Fig.1.7 and elsewhere where visual appearance is discussed: how do we know that these are examples of true mergers? More generally, regarding sec.1.3.1, how safe is visual inspection to find genuine galaxy mergers? How to avoid the risk of finding only chance projections of not interacting galaxies? These questions also apply more generally to the entire thesis. How to identify true mergers needed for instance to train machine-learning models or to test the performance of merger identification approaches?

- Sec.1.3.2 attempts to answer the latter question: we need to know the distances, or at least redshifts, to and between the two galaxies to know if they are physically close to each other. This simple statement is however accompanied by a long discussion of only vaguely relevant aspects, which in addition are presented in a hard-to-follow and often inaccurate manner, and with just a few and not always relevant references. For instance, the distance-velocity comparison from Hubble should be the seminal 1929 paper; it was Friedman in 1920s rather than Lemaitre in 1930s who first came up with solutions indicating universal expansion. Then, a (somewhat superficial) discussion of difficulties with distances in cosmology follows. Why is it relevant for this section? How do galaxy peculiar velocities (called here “proper motions”) matter in this context? Why do we need to deal with not only spectroscopic, but also photometric redshifts in galaxy merger studies?
Some details (missing here) of the ‘cross-pairs’ method are provided only much later in sec.3.2.1, but also there it is not clearly stated how the two quantities – relative distance Δr and relative velocity Δv – are related to the discussion from sec.1.3.2.

- How are the morphological parameters discussed in sec.1.3.3 related to and used for finding mergers?

- Why is the discussion on large language models and ChatGaia in sec.1.3.4 relevant for the rest of the thesis? Such methods do not seem to be used in the presented work.
An additional comment on the contents of this section is that while methods such as CNN are introduced, it is not clarified how they work and in particular how they are used for merger finding.

Chapter 2 presents the data used in the thesis and how they were prepared and processed for merger finding. This is a short chapter of only 7 pages. It does describe the key aspects of the relevant data, but perhaps some more details could have been added for a better overview of the datasets and their properties.

- Among the data sets used in the dissertation, we find SDSS DR6 released back in 2008. Since then, SDSS had more then 10 further data releases. Why this particular data release was therefore employed for the presented study? Especially that it is also mentioned that background modeling in SDSS DR6 had issues which were corrected later in DR8.

- In sec.2.2.1, what is the relation between GAMA-KiDS and HSC-NEP?

- In sec.2.2.2, it would be good to know in a bit more detail what the contribution of the Candidate to the data preparation was in the work Pearson et al. 2022. For more information on how the relevant dataset was created we are redirected to the original paper, but I think these are important details which should have been given here as well.

- Where do the redshifts come from in the HSC-NEP dataset?

Chapter 3 presents the methodology used. Here we find the relevant details which are key to understand the results presented afterwards in Ch. 4–6.

- Regarding sec.3.1 and related discussion later, including Ch.4. The task of finding galaxy mergers with a supervised learning algorithm seems to be more adequate for an image-based analysis, for instance with convolutional neural networks processing the images directly. This is indeed discussed in sec.3.3 and ch.6, but here “shallow” networks were used instead, operating on already pre-defined photometric features. Why such a choice was made?

- In sec.3.1.1, what were the input data for the neural network?

- In sec.3.1.4, only accuracy is defined, while I believe equally important to quantify model performance are recall and specificity, mentioned only later in sec.3.2.3.4. An additional comment is that such statistics should have been defined in a more general context, and not only when either NN or decision trees are discussed. They apply to any classification problem, and in particular to the merger identification problem.

- I find sec.3.2 “Decision Tree” rather hard to follow. It would have been much more convenient if an illustration of the DT was provided in the first place, to better understand the order of the steps taken and decisions made in the process. Even more confusingly, the subsection 3.2.3 of this section is also called “Decision Tree”, while Fig.3.1, which apparently illustrates decision trees, seems to be only relevant for sec.3.2.2 “Visual inspection”.

- The detailed section 3.4.2 seems to describe an external pipeline. What was Candidate’s contribution to what is discussed here?

- A general note on this chapter: in several places we are redirected to details presented only later, e.g. to sec.4.1.1 (p.34), 4.2.3 (p.37), 5.1.1 (p.37), 5.1.2 (p.39), etc. This contributes to the difficulty of following the narrative of the thesis and I believe some of these details could have been discussed where they are first mentioned.

Chapter 4 describes the study published in Suelves et al. 2023, where a neural network model was used to search for galaxy mergers in SDSS. Detailed analysis of the features most relevant for merger identification led to an interesting and somewhat surprising finding that of most importance are not really properties of galaxies themselves, but associated sky background errors (“skyErr”). The proposed interpretation of this finding is that “higher values of the sky background error reflect the traces of merging processes – for example, faint tidal tails – otherwise missed by the sky background measurement due to the dominance of the signal from a galaxy itself”. Let me however note that the analysis of deeper, better quality data from Stripe 82 (sec.4.2.4.1) does not seem to confirm the importance of the skyErr parameters for merger finding there. It then remains an open question if the results of this chapter can be generalized beyond the specificity of the SDSS DR6 data.

- My general comment on this chapter is that while we are offered much detailed discussion, what seems to be somewhat lacking are conclusions and interpretation of the findings. For instance, not much is commented upon the fact that in sec.4.2.4.1, where deeper imaging is analyzed, the resulting accuracies are not much larger than 50%, which suggests that the approach has not really worked well. More importantly, sec.4.3 called "Conclusions" is mostly devoted to summarizing the contents of this chapter, while the actual conclusions and interpretation are provided only in the final paragraph. Also here, no comment is offered on the fact that the method did work for SDSS DR6 general data, but not much so for the deeper Stripe 82.

- A comment on Fig.4.11 and related discussion. Such a classification problem where one looks for a boundary between different classes seems very adequate for dedicated approaches as for instance Support Vector Machines. Would it be applicable here?

Chapter 5 – also entitled "Decision Tree" – presents unpublished work on detailed analysis of the skyErr diagrams for Galaxy Zoo Data Release 1. This relatively short chapter includes many details, but, as stated in its beginning, it is still a proof of concept. In the final paragraph we additionally find "We considered that the implementation of this project had a fundamental complication, the separation of the data by magnitudes and by GZ DR1 morphologies. […] However, it complicated the generalization. In future work, we would like to separate the more general decision tree from a more specific analysis of the magnitude and merger votes, which would provide insight on what are the best galaxy images that the sky error method could trace." This confirms that this approach is currently only at an early stage. What is more, the text does not seem to provide a definitive outcome and conclusions from the study.

- From the beginning of this chapter I could not identify what the outline and scope of the analysis is. We do find the statement "This chapter presents a proof of concept of the method which we plan to fine tun further before its submission for publication" [original spelling], however it is not clarified what "the method" is. What follows is more an exploration of various aspects of the skyErr-g vs. skyErr-r diagram than some clear methodological approach.

- It would be good to know more what the next steps are planned for this analysis to go beyond a proof of concept. Some of the results quoted do not indicate that the method works too well – for instance, specificity can be as low as 55% only (table 5.3).

- Also the percentages of mergers quoted below table 5.1 are mostly below 50%, which does not seem to be in line with the statement that "This not only confirms the potential of the decision diagram for using the sky background error to find galaxy mergers of any types, but also provides a potential recipe to locate multiple type of mergers by studying different regions of the diagram."

- In more detail, throughout the chapter it was not clear to me why equally much attention is given to non-mergers as to the mergers themselves. The thesis is devoted to merger identification, therefore shouldn't the goal be to maximize various relevant metrics for them and to treat anything else as "the rest"?

- A comment on Figs.5.1, 5.3 & 5.4. The horizontal and vertical features respectively at the bottom and at the left indicate some minimum values of the skyErr parameters, namely their lower limits. How safe is to use them in the analysis?

Chapter 6 presents Candidate's contribution to Pearson et al. 2022, and extends the sky error method to deeper HSC NEP images, with the results planned to be published soon. However, similarly as in Ch.5, this short Chapter (9 pages) presents in my opinion only a proof of concept, which will require further refinements to reach more solid scientific results. In particular, I am left with the impression that the presented methodology has not really worked for the deeper HSC data, taking into account the various statistics presented. Some more detailed discussion on that would be in place.

- First of all, I am wondering if the deep learning (DL) model used originally to find mergers can be considered successful and reliable at all. The visual inspection of the pre-assigned merger candidates confirmed only ~20% of them as genuine. It is proposed that this might be due to either "contamination in the DL classification" or to "the performance of our inspection". Any of these options would be worrying – as either the DL model was not trained and tested properly, or the visual inspection cannot be trusted. The text does not seem to provide a clear suggestion which of the options is more likely and what would be the consequences for the proposed methodology if any of these two would cause such low statistics.

- It is claimed that "the parameter distributions indicate that the LSB pixels around mergers and non-mergers are not statistically the same". I understand that this refers to the distributions shown in figs.6.3 & 6.4. Has this been quantified e.g. with a test such as Kolmogorov-Smirnov? My visual judgment is that the differences between the respective distributions are rather small, but this is only a qualitative assessment.

- Red rectangle in fig.6.5 includes 62% mergers, which is interpreted as successful performance of the approach. Yet this is not much larger than 50%, and indicates 38% contamination. I am not convinced these are promising statistics.

- I find the figures in this chapter a bit less informative as either they do not have the units of the greyscale bars provided (figs. 6.1 & 6.2), headers are not properly explained ("iqr", "frac1" etc., figs 6.3. & 6.4), or the x-y axes are not labeled (fig. 6.5).

- What could be possibly improved for this method to provide better statistics of merger identification, in view of the plans to apply it to future LSST data?

Chapter 7 provides a very short summary of the thesis. In my opinion it would benefit from having added some more context how the presented research is placed in the field of merger finding, especially with respect to studies done by other groups. Although it is stated that "I have built methods for finding galaxy mergers that can be faster and more accurate than previous methods", I have not found a relevant comparison with these "previous methods" to be clearly presented in the thesis. Moreover, what are the future prospects starting off from the current studies, if and how the discussed methods can be improved?

On a more general note, the thesis includes a non-negligible number of typos and stylistic deficiencies, which could have been avoided if more careful proofreading had been done. We find these even in the Abstract. Some examples are: “Large” rather than “Legacy” (Survey of Space and Time) [also in sec.1.5], as well as “striped” rather than “stripped”; mentioning “the diagram” not previously introduced, and a repetition “For that, I worked on dimensionality reduction methods, on the image calibration, and on Machine Learning (ML) techniques such as Neural Networks (NN) or dimensionality reduction methods.” Many more such issues are in the main text, which unfortunately adversely affects the readability of the dissertation.[3]

Conclusions:
Despite these shortcomings that I have identified, my overall evaluation of this dissertation is **positive**. While the work could have been put in a better physical context, and some more critical assessment of the presented results could have been made, I have no doubts that the findings of this thesis provide valuable contributions to the field of galaxy merger studies in particular, and to astrophysics in general. Furthermore, I believe that in this dissertation Mr. Luis E. Suelves has demonstrated his ability to individually conduct research.

I therefore conclude that the presented dissertation meets the formal requirements for a PhD thesis and recommend admission of the Candidate to the subsequent stages of the procedure, including the public defense.

Konkluzja:
Pomimo zidentyfikowanych przeze mnie niedociągnięć, moja ogólna ocena rozprawy jest **pozytywna**. Praca mogłaby być opatrzona kontekstem fizycznym w większym stopniu, a przedstawione wyniki mogłyby być poddane bardziej krytycznej ocenie, nie mam jednak wątpliwości, że przedstawione w rozprawie rezultaty są cennym wkładem w badania złączeń galaktyk w szczególności, i w astrofizykę w ogólności. Co więcej, uważam że w niniejszej rozprawie Pan Luis E. Suelves dowiódł umiejętności samodzielnego prowadzenia pracy naukowej.

Stwierdzam zatem, że przedstawiona mi do recenzji rozprawa spełnia wymagania ustawowe stawiane rozprawom doktorskim i wnoszę o dopuszczenie jej do dalszej części postępowania, w tym publicznej obrony.

Maciej Bilicki

Podpisany elektronicznie przez
Maciej Andrzej Bilicki
04.09.2024
8:26:00 +02'00'

3 An interesting detail are somewhat peculiar credits to two of the figures:
Fig.1.4 “Wikipedia Joshua (b) Addapted from (Baldry et al., 2004) by Prof. Schroeder Francesco M. Valentino.”
Fig.1.5 “Courtesy of dr. Wojciech A. Hellwing, in the lecture Introduction to Cosmology.”