

Be concise, yet precise

A guide to scientific writing

Dr. Heloise F. Stevance

Version 0.2 - September 2020

1. Introduction	3
1.1. Who is this for?	3
1.2. Who is this from?	3
1.3. What is scientific writing?	3
2. Structuring your work	4
2.1. Some Basic Sectioning	4
2.1.1. The Abstract:	4
2.1.2. The Introduction	5
2.1.3. Methods / Observations	6
2.1.4. Results/Analysis	7
2.1.5. The Discussion	8
2.1.6. The Conclusion/Summary	9
2.2. Getting creative	10
3. Writing best practices	10
3.1. Do not use abbreviations	10
3.2. Use “We” or the passive voice	10
3.3. Use short sentences	11
3.4. Jargon and word salads	11
3.5. Information rich language	12
4. Figures, Tables and References	13
4.1. Figures	13
4.2. Tables	14
4.3. Citing your sources	14
5. Some general tips	15
5.1. How do I start?	15
5.1.1. Keep good notes	15
5.1.2. Start with bullet points	15
5.1.3. Don’t wait for inspiration	16
5.2. Writing is 90% Editing	16
5.3. Manage your time	16
5.4. Feedback time	17
5.5. When to ask for feedback:	17
6. Answers to the exercises	18
6.1. Exercise 1	18
6.2. Exercise 2	18
6.3. Exercise 3	18

1. Introduction

1.1. Who is this for?

This hand-book is aimed at late stage undergraduate students and graduate students working towards a thesis/report or journal article. My goal is to provide some of the fundamentals of scientific writing as well as practical tips that you can start implementing today.

1.2. Who is this from?

I am currently a postdoctoral astrophysicist at the University of Auckland. This guide is written from personal experience and therefore from an astrophysics perspective. If you would like to provide feedback on this material to expand it or generalise it, please email me at hfstevance@gmail.com

1.3. What is scientific writing?

Writing for a thesis or a peer-reviewed journal is a very unique exercise. If you have written a lab report before, you will have been given a taste of what it is like to “write” science. This style is very different from anything else you will have encountered.

It should be:

- **Concise**: Get to the point. At the level you are at now, nobody will be impressed by a high word count, I promise.
- **Precise**: Making sure your wording is unambiguous is key to sharing your findings with your colleagues. Reproducibility is key.
- **Impersonal**: Think about it this way - in this story the center of attention should be the data and the results, not the scientist that produced them (so avoid using “I” - but we’ll get back to that later).



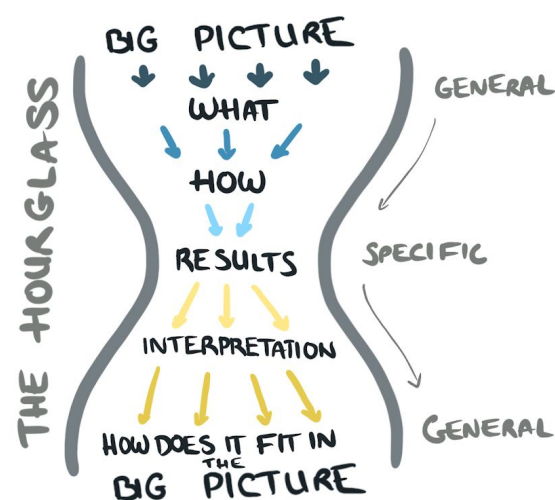
Note: This handbook will provide plenty of examples of scientific writing but it is not an example of scientific writing. I am writing in a more casual style so it’s not too dry to read.

2. Structuring your work

In the first part of this hand-book we are going to focus on one of the most important aspects of writing: **flow**. To clearly and concisely report your research - whether it is in a lab report, a thesis, or a paper- your final piece should follow a natural progression.

As a whole your work will start by providing context and explaining the relevant general concepts (wide) then transitions to detailing the work you did (narrow), before finally falling back onto the big picture and the larger implications of your research (wide). This sort of structure is sometimes referred to as an “hourglass”.

Let’s start by introducing the most fundamental structure for a paper - it is surprisingly versatile and you will want to thoroughly understand this basic form before coming up with more creative approaches for future work.



2.1. Some Basic Sectioning

2.1.1. The Abstract

The abstract of a paper (or summary of a thesis) is a window into your work. It is the first thing that your audience will read and it will tell them whether your piece is relevant to their interests. If they are very busy it may be the only part of your paper that they read. So it is very important to ensure that your key results and interpretations are included.

Here is a basic recipe for an abstract:

- Why is this work relevant to the field or recent events?
- What did you do?
- What did you find?
- What does it mean?
- How does that fit into the big picture?

Notice something? This is the hour-glass structure we just introduced. That’s not surprising - your abstract is basically a very condensed version of your paper.

Exercise 1:

Read this abstract from [Eldridge et al. 2020](#), and try to identify the different “ingredients” in the abstract recipe - what “color” (as presented above) would each sentence belong to? (Answers at the end of the hand-book)

“The recent identification of a candidate very massive ($70 M_{\odot}$) black hole (BH) is at odds with our current understanding of stellar winds and pair-instability supernovae. We investigate alternate explanations for this system by searching the BPASS v2.2 stellar and population synthesis models for those that match the observed properties of the system. We find binary evolution models that match the LB-1 system, at the reported Gaia distance, with more moderate BH masses of 4-7 M_{\odot} . We also examine the suggestion that the binary motion may have led to an incorrect distance determination by Gaia. We find that the Gaia distance is accurate and that the binary system is consistent with the observation at this distance. Consequently, it is highly improbable that the BH in this system has the extreme mass originally suggested. Instead, it is more likely to be representative of the typical BH binary population expected in our Galaxy.”

2.1.2. The Introduction

The introduction will start by reiterating and expanding the context (a.k.a big picture) you have given in the abstract. Instead of one (or 2) sentence you now have room to provide more details: What’s the current understanding of the field? What blind spots are there? What main question is your study trying to answer? Why is it relevant/important? What is this work going to cover?

Your goal is to give the reader with an overview of the field and where your work fits in. There is unfortunately no “fail-safe” recipe for introductions as they can vary in length and substance depending on the paper that you are writing and the audience that you expect will read it (e.g. is it for a niche of experts or is it likely to be read by non-specialist astronomers?).

As a general rule, the introduction tends to have a “triangular” shape (as opposed to an hour-glass): It goes from very generic to very specific. A good idea is to start with a statement that is widely accepted in the field and introduces the subject - for example in the field of supernovae you will see countless papers with a first sentence similar to:

“Core collapse supernovae (CCSNe) are the result of the death of a massive star ($M_{\text{ZAMS}} > 8 M_{\text{sun}}$)”. ([Stevance et al. 2020](#))

Another common feature of introductions is that you will often see at the very end a paragraph detailing what each Section is going to cover - here is an example from [Eldridge et al. 2020](#):

“The outline of this article is as follows: We first describe our fitting mechanism and present the results of our model search in Section 2. We investigate the impact of the astrometric wobble due to binary motion on Gaia’s parallax measurement in Section 3. We then discuss in Section 4 the expected binary population of BH binaries from BPASS and suggest that future Gaia discoveries of these systems will allow us to rigidly constrain the formation of BH at the point of core collapse within stellar interiors.”



Untold Rule #1: In astronomy introductions are rarely split into subsections - use paragraphs to separate ideas and concepts.

It is worth noting that the introduction is one of the hardest sections to write, alongside the discussion. This is where leveraging the experience of your advisor or senior peers will become necessary - they will be able to guide you to write an introduction that makes sense for the study you are writing.

2.1.3. Methods / Observations

The second section of your paper will most likely be where you explain your methods or how you obtained your data / reduced the observations. The keyword you want to keep in mind when writing this is **“reproducibility”**. This section will often be skipped by casual readers but will be scrutinized by experts trying to use your results for their own science. Here are some of the key information you will want to put in this section depending on the type of data you are using:

Observational Data (reported for the first time):

- When the observations were taken and with what telescope/instrument; PI
- Observing condition (seeing)
- Observation log in a table
- Steps of data reduction and any software you used
- Any issues with the data that should be reported?

Observational Data (pre-existing):

- Where does the data come from - reference and link to database if available.
- Quick summary of the data: e.g what instrument, which sources you are using or how many (if you are using a catalogue of 10,000 stars you will obviously not name them individually).

Synthetic “Data” (Models):

- Give a succinct summary of the model/simulations you are using: focus on the information that will be relevant to your analysis and make sure to cite the appropriate references so that an interested reader can go look for more details.
- If available, provide the download link to the models

Note that you can obviously use both synthetic data and observational data in the same paper, in which case you might want to have two subsections: one describing the observed data and one describing the model. That being said, it is quite common in observational papers to see synthetic data/models being used directly in the analysis/discussion with just a citation to the relevant literature. This is particularly true when the models play a partial role (as opposed to a central one) in a much larger analysis work.



Untold Rule #2: The “Methods” section will typically only contain the methods by which the data/main findings were obtained. If you do some specific fitting or statistics in the analysis, it is expected that you will describe those methods where they are relevant to the results.

2.1.4. Results/Analysis

The next step is to present your findings - this can be done in one section, but is often split. How that split is applied depends on what makes sense for your study. When writing it up, you once again want to keep “reproducibility” in mind - your colleagues need to understand **how** you obtained your results.

How you will actually structure and present your findings is something you will have to figure out individually with the help of your advisor. To give you some inspiration here are a few case studies.

Case studies:

Note you do not need to read the following papers fully - just skim through the sections to get an idea of the structure.

→ Single result/analysis section:

This is pretty self-explanatory, here is a recent example: [Perego et al. 2020](#). This formula is most appropriate for short pieces that have one core result and little supplementary analysis.

→ Multiple analysis/result sections:

In [Stevance et al. 2020b](#) the analysis and results are not distinguished but two types of analyses are split into two sections: Section 3 answers the main focus of this paper whilst Section 4 offers supplementary analysis supporting and improving the results of Section 3. The opposite structure is seen in [Dobbels et al 2019](#), where their Section 3 presents results that will be used as reference and comparison for their main analysis presented in their Section 4.

→ Separating the result and analysis section: [Stevance et al. 2017](#)

This paper presents new observations so a “Results” section (Section 3) is first provided with a description of the new data including some key quantities. The following section (Section 4) provides further analysis with more advanced methods and some assumptions. Whilst the analysis of Section 4 could be improved in the future as the field develops better methods/understanding, Section 3 would remain mostly unchanged.

2.1.5. The Discussion

Now that you have presented your key findings and explained how you obtained your results, it is time to **interpret** those results and **place them in the context of the literature**. Here are some general questions to get you started:

- Do your conclusions agree with previous findings (observational data or theoretical models)?
- Do they differ?
- What are the implications of your conclusions?
- Are there any major caveats the reader should keep in mind?

Like the introduction, this section requires a good understanding of the existing literature which makes it very difficult to write. It is quite likely that you will need significant input from your advisor or senior people in your field before you can finalise your discussion. It is also often the place where caveats raised by the referee can be addressed.



Top tip: Before you start writing your discussion, take a good look at the discussions in some of the key papers you have been reading during your analysis work. What ideas are commonly discussed? What comparisons have previous authors made? (make a note of the relevant references)

2.1.6. The Conclusion/Summary

The final science section of your paper is a summary of your results and how they fit in the big picture. People who don't have time to read your whole paper will typically skip the middle sections and go directly to this one.

A common practice is to start your conclusion with a one sentence summary of what your paper accomplished. For example:

"In this work we presented an ageing method based on the comparison of BPASS models to observation using hoki." - [Stevance et al. 2020b](#)

"In this work we estimated the far-infrared SED, and its corresponding dust properties, via machine learning techniques"- [Dobbels et al 2019](#)

Something you may see is key findings being enumerated with bullet points (see refs above). This is a very impactful way to quickly convey the main message of your work to busy readers who have little time and attention to dedicate to keeping up with the literature.

Finally you will want to end your conclusion on some broader considerations: re-iterate the context that your results fit in and expand on future prospects or work that should be undertaken. In a sense, your conclusion is also a "triangle" (but the reverse of the introduction) - starting with a narrow, specific, summary of what your study did and finishing with a wide view of the bigger picture. It helps complete the hourglass shape you are aiming to achieve for your piece as a whole.



Untold Rule #3: Usually the conclusion doesn't contain any references (e.g Stevance et al 2020). The relevant literature will have been mentioned before.

2.2. Getting creative

The classic structure presented above will work a lot of the time but you may encounter papers that do not follow a strict *Intro -> Methods -> Results/Analysis -> Discussion -> Conclusion* structure.

The introduction, discussion and conclusion will always be there in some form but depending on the focus of the study and the work done, the classic structure may be awkward: Whatever you do, “flow” should prevail - each section should lead to the next in a sensible way.

A personal example where I had to diverge from the classic structure was [Stevance et al. 2020a](#): the paper included a review of a specific aspect of the field, a re-work of a previous study and then a full new analysis. The unusual content called for a more specialised structure.

When you are just starting out, it is most likely that you will be able to stick with the basics, but as you grow in your research expertise and writing experience, do not hesitate to steer away from the classics when you need to.

3. Writing best practices

Now that we have gone over how to structure our work, let's dive into some writing conventions that you will want to observe.

3.1. Do not use abbreviations

Scientific writing is a formal endeavour. Instead of “can't” use “cannot”; instead of “isn't” use “is not”; etc...

3.2. Use “We” or the passive voice

It is extremely rare to use “I” in a paper - that is kind of like core file systems on your computer: do not mess with it unless you really know what you are doing. The “passive voice” is a term that may be new to you but you will have used it. Here are a few examples:

The subject The verb The complements
(thing that does the action) (the action) (thing that the action is done to)

<u>Active Voice</u>	<u>Passive Voice</u>
We modelled 100 stars	100 stars <u>were</u> modelled
We heated the test tubes to 50C	The test tubes <u>were</u> heated to 50C
We observed WR102 with the VLT	WR102 <u>was</u> observed with the VLT

Notice that the subject (red) disappears in the passive voice? That is entirely the point. In science the focus is on the methods and the findings, not on the people. Using the passive voice constantly can be a little awkward and heavy though (it can make some sentences unnecessarily complicated). Using a mixture of the passive voice and active voice where “we” is the subject keeps your writing impersonal whilst alleviating some of that awkwardness.

Exercise 2:

Rephrase the following sentences using the passive voice:

“We applied a gaussian filter to smooth the data”

“We binned our spectrum to improve the signal to noise ratio”

3.3. Use short sentences

In novels, long, windy sentences can bring some interesting character or tone. We are not interested in that: make the information easy to process by keeping sentences relatively short. If your sentence is 3 or 4 lines long there is probably a good way to split it up.

Exercise 3:

Split and shorten the following sentence - make sure you retain all of the information.

“At the end of the life of massive stars (those that are born with more than 8 times the mass of the sun), their core runs out of fuel and as a result they are no longer supported against gravity causing core collapse and subsequently their explosion as supernovae whose will enrich the interstellar medium with new elements that are necessary to form planets and the building blocks of life as we known it.”

3.4. Jargon and word salads

When you first start reading papers you may find it really challenging to follow along, especially because of “jargon”, i.e. special words or expressions that have a very specific meaning in a given field.

The point of jargon is not to sound more scientific, or professional, or clever. The point of jargon is **specificity**. For example, instead of saying “a star born with more than 8 solar masses” we say “a massive star”. Everyone who has been in the field for some time knows what this means and it keeps our sentences short but still informative: Be concise, yet precise.

Conversely, a “word salad” is a sentence filled with jargon in a way that makes no sense and carries virtually no information. You will see this a lot in pseudo-science, and you occasionally encounter it in writing where the author tries too hard to “sound” like a scientist.



How To Avoid Word Salads:

- Only use the jargon you understand.
- If you have a choice between a simple and a complicated word (and they deliver the same level of information) use the simple word.

3.5. Information rich language

Our goal is to provide our reader with as much information as possible, so you should avoid vague words and find a specific counterpart. For example “big” carries much less information than “high amplitude”. Additionally, you should quantify all of the important information whenever possible. Instead of “Star 1 is brighter than Star 2” you could say “Star 1 is 2.5 magnitudes brighter than Star 2 in the V-band”.

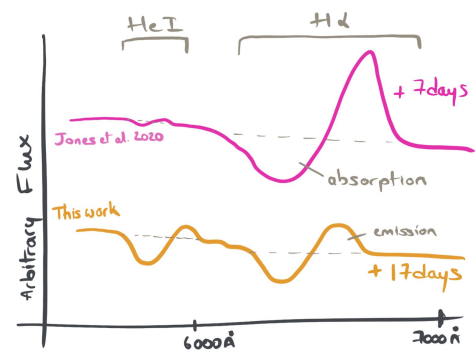
Below is an example of uninformative writing compared to what we would expect in a publication. The “data” and the reference in the figure are both fake by the way.

Poor description

The helium line is bigger in our work whilst the hydrogen line is smaller, and the ejecta velocities are the same because the absorption minima are the same.

Good description

The He I 5785 P Cygni profile is hardly noticeable in the spectrum obtained by Jones et al. 2020 at +7 days, but becomes more pronounced by +17 days, as seen in our data. The H alpha line, by contrast, weakens significantly, particularly the emission profile. The absorption minima of both He I 5785 and H alpha remain constant at around 5650Å and 6300Å, corresponding to ejecta velocities of 7000 km/s and 12,000km/s, respectively.



4. Figures, Tables and References

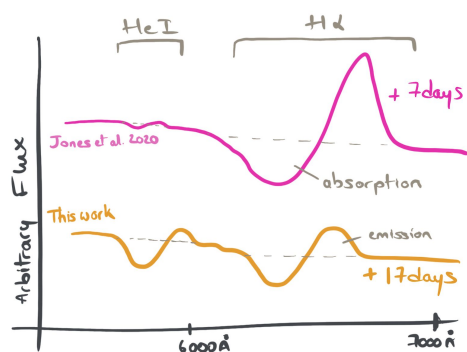
4.1. Figures

An image is worth a thousand words and figures are one of the things (alongside abstract and conclusion) that busy readers will focus on when going through your work. Even to someone who reads your whole piece, the figures will stand out as emphasising the main messages and results of your work.

Consequently, you need to create and pick your graphs strategically: don't go and put every plot you have made during your analysis in your final report/thesis/paper. Choose key figures that will illustrate the main messages of your paper.

Best Practices

- **Legibility:** Make sure the axis labels are a reasonable size, it can be easy when using default values in matplotlib to end up with comically large or small labels.
- **Legends and titles:** although in lab reports legends and plot titles may have been mandatory, in science papers these descriptions are often left to the caption. For more complex plots involving colour bars or a large number of lines or marker types, legends may be appropriate - but for most plots they are not necessary. Similarly, plot titles are often superfluous as the first thing a caption will do is describe the plot - if a large number of graphs are present in one figure, subtitles can be provided for each plot if necessary.



- Example of a caption:

Spectra of fake supernova at +7 days (magenta) and +17 days (orange) after explosion between 5300 and 7050 Angstrom. The flux was scaled and offset for comparison purposes. The +7 days data were obtained from Jones et al. 2020.

Note: The caption should be self-contained

4.2. Tables

Tables are not only containers for your data (e.g. Table 1 [Bestenlehner et al. 2020](#)), they can also be a fantastic way to summarise information (e.g. most tables in [Stevance et al. 2018](#)) or offer a tabular comparison to the literature (e.g. Table 2 in [Stevance et al. 2017](#)).



Python to LaTeX: If you have giant numerical tables to fill, you don't have to do it by hand! Check out [this python module](#): it can create LaTeX code from your numpy tables.

4.3. Citing your sources

An inherent part of writing science in a responsible way is to cite your sources. That does not only include references to the literature: If you have used images from a textbook or a website you **must cite the source**. **Unless otherwise stated, it is assumed that everything you present is your own work**, so if you use an image or copy a table from someone else without giving a source that is **plagiarism**. Same for captions and any written text associated with images.



Use a citation manager: You don't have to do your bibliography by hand! Latex has a great reference manager system called BibTeX which will do the heavy lifting for you. For a quick intro check [this page](#) - if you are a complete LaTeX beginner I recommend you ask a peer or your advisor to get you started.

5. Some general tips

Now that you know the theory of what to write and how to write it, you need to actually do it, and this is often the tricky bit. So here is a collection of tips to help you put all that theory into practice.

5.1. How do I start?

5.1.1. Keep good notes

Writing does not start when you first open your report/thesis/paper template - it begins long before that, during the data analysis. Remember “reproducibility”? How are you going to communicate to your colleagues how you obtained your results if you cannot remember what you did? Take good notes as you go, they will form the basis of your Results/Analysis Sections.

5.1.2. Start with bullet points

Do not start writing without a plan. Before you start writing full sentences and paragraphs, go ahead and brainstorm. It can help at first to just dump all your ideas in a document without thinking too much about it and then go through that list and start organising your thoughts.

Start with the main sections and subsections that you want to write, then populate these subsections with the messages you want to get across and in what order. **You will not get this right the first time around** - this is something you can iterate on and that can change overtime.

Here is an example of how I would organise some bullet point ideas before starting to write an introduction:

1. Introduction
 - a. What’s a supernova and supernova classification
 - b. Why type IIb SNe are important
 - c. What is spectropolarimetry and what it can tell us
 - d. Past studies of Type IIb specpol
 - e. Introduce our supernova briefly - state we are providing and analysing newly obtained data
 - f. Describe paper organisation
2. Methods
3. Results
4. etc...



Untold Rule#3: No one writes a paper in the order that it is read, or even in the order that you write your bullet points. Here is the most common writing sequence: Methods and results first, then the discussion and introduction, followed by the conclusion and finally, the abstract.

5.1.3. Don't wait for inspiration

If you wait until the perfect sentence comes to you to write things down you will end up with a blank piece of paper. Just write. It will be bad at first, regardless of how much experience you have. **Good writers don't write perfectly on their first draft, they just revise more than bad writers.** So just write, ANYTHING, don't worry about it being formal, or following the guidelines I gave you in Section 3 - all of that can be fixed later. Just focus on getting some words down, because **you can't edit an empty page.**

5.2. Writing is 90% Editing

Once you have some material down it is time to start editing. I cannot emphasise enough how iterative this process is. You will revise and revise and revise again, then get feedback, then revise again, repeat a few times after which you will go through the refereeing process and then - that's right! - you will revise your manuscript again.

Just like a good night's sleep is the best debugger, **a rested brain is the best editor.** When you are done with a draft (or a piece of it), leave it alone for a day or two and then read it again with a fresh mind. You would be surprised how many things jump at you as being either unclear or needing polishing.

Here are a few things to look out for when you read over your work:

- Does it flow or do some sections and paragraphs seem to come out of nowhere? Maybe you need to move things around or add transitional sentences.
- Are some sentences too long? Can you break them up?
- Are you repeating some words too often? Use a [thesaurus!](#)
- Typos can be hard to spot in your own writing, a good trick is to read your work out loud - it helps you focus on the task at hand.

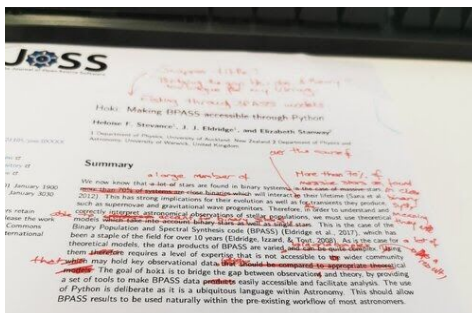
5.3. Manage your time

You know what I am about to say but it is always worth repeating - DO NOT start writing at the last minute. It will not go well. Your work will be sub-par and you will be unnecessarily stressed. The off-time between writing sessions is important: an essay written in 6 hours in one sitting will never be as good as an essay written in 6 hours in two sessions with rest and breaks.

You can never start too early, even if it's just bullet points.

5.4. Feedback time

The next step is to get feedback on your manuscript. There is no stage in your career when you will not get feedback and criticism - don't be nervous. This is not an exam, this is not a pass or fail. You and your advisor/colleagues are working together towards a common



goal. You will receive a lot of notes, and if your annotated draft looks like this picture don't panic! It's perfectly normal. In fact, as you become a better writer and a better editor you will give yourself the red pen treatment: this is an image of the draft of one of my papers I personally annotated 3 days after writing those very words. Trust the process.

5.5. When to ask for feedback:

As a general rule, you should not give your very first draft for feedback. Allow yourself some time to review and edit your own work.

That being said, don't wait months before you get the opinion of another human! You don't have to wait until you have a complete draft to ask for feedback, especially when you are just getting started. As soon as you have a few subsections that you have somewhat polished, get your advisor or a senior peer to read over it and give you some advice.

Good Luck!

That's all the tips I have for now, I hope they will help you start your writing journey.

6. Answers to the exercises

6.1. Exercise 1

The recent identification of a candidate very massive ($70 M_{\odot}$) black hole (BH) is at odds with our current understanding of stellar winds and pair-instability supernovae. We investigate alternate explanations for this system by searching the BPASS v2.2 stellar and population synthesis models for those that match the observed properties of the system. We find binary evolution models that match the LB-1 system, at the reported Gaia distance, with more moderate BH masses of 4-7 M_{\odot} . We also examine the suggestion that the binary motion may have led to an incorrect distance determination by Gaia. We find that the Gaia distance is accurate and that the binary system is consistent with the observation at this distance. Consequently, it is highly improbable that the BH in this system has the extreme mass originally suggested. Instead, it is more likely to be representative of the typical BH binary population expected in our Galaxy.

6.2. Exercise 2

“A gaussian filter was applied to smooth the data”

“The spectrum was binned to improve the signal to noise ratio”

6.3. Exercise 3

Note that there is **not just one** right answer for this sort of exercise of course.

Here is how I would rephrase the original sentence.

“Massive stars ($M_{\text{zams}} > 8M_{\text{sun}}$), when they reach the end of their lives, will run out of fuel in their core. No longer supported against gravity, their core will collapse resulting in a Supernovae explosion. The ejecta will then enrich the interstellar medium with the elements necessary to form planets and life as we know it.”