The Blue-to-Red Supergiant Ratio in the Small Magellanic Cloud

Alexander Iain Dutoy

Masterarbeit in Astrophysik
angefertigt im Argelander-Institut für Astronomie

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

AUG 2023

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, 05.08.2023
Date

Signature

1. Gutachter: Prof. Dr. Norbert Langer
2. Gutachter: Prof. Dr. Frank Bigiel

Abstract

In the early Universe, hydrogen and helium were the only elements that existed until chemical enrichment brought on by stellar nucleosynthesis. We can use low-metallicity dwarf galaxies to probe the early Universe as these environments share a similar metallicity to that of the early Universe. Photometric data of these dwarf galaxies can allow us to deduce key characteristics of the individual stars in these galaxies, such as their effective temperature, $T_{\rm eff}$, and bolometric luminosity, L. Supergiant stars within these stellar populations provide a sensitive test for the internal mixing parameter efficiencies in stars, namely semiconvection, $\alpha_{\rm sc}$, and convective overshooting, $\alpha_{\rm ov}$. Both $\alpha_{\rm sc}$ and $\alpha_{\rm ov}$ make massive stellar evolution uncertain, and constraining these parameters is critical to studying the evolution of these stars and their effect on the host dwarf galaxy. We produce a blue-to-red supergiant ratio (B/R) for a neighbouring metal-poor dwarf galaxy, the Small Magellanic Cloud (SMC). We utilise bolometric correction tables to convert colour and magnitude to effective temperature and bolometric luminosity. A good agreement is found between these derived parameters and spectroscopically determined values from various spectroscopic surveys. The main sequence and asymptotic giant branch (AGB) are identified by fitting Gaussian distributions to effective temperature histograms over various luminosity bins. Once the number of main sequence stars and AGB stars are determined by integrating the areas of these Gaussian curves to 5σ, these values are subtracted from the total number of stars to produce the number of blue and red supergiants (BSGs and RSGs). The blue-to-red supergiant ratio is then determined by dividing the number of blue supergiants by the number of red supergiants. The observed B/R is compared against the B/R derived from various synthetic stellar populations. These populations are created from a previously run grid of massive stars with SMC metallicity and varying internal mixing efficiencies. The width of the main sequence in the synthetic stellar population is a direct consequence of the uncertainty in visual extinction, $A_v = 0.3 \pm 0.1$. This highlights that a smaller uncertainty in A_v could promote a more precise determination of $T_{\rm eff}$ and L, resulting in a more accurate determination of the B/R. The best correlation between the observed trend of the luminosity-dependent B/R and the one derived from the synthetic stellar population is where the internal mixing parameter efficiencies are $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$. More efficient semiconvection, $\alpha_{\rm sc} > 0.1$, results in too few RSGs at luminosities $\log(L/{\rm L}_\odot) \geq 4.25$ than observed. More efficient overshooting, $\alpha_{\rm ov} \geq 0.11$, results in longer main sequence lifetimes of the stars and a smaller B/R than observed. While there is a good agreement between the observed B/R and the B/R for $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$, it is evident that there are still data points outwith the expected uncertainty limits, implying an incomplete understanding of mixing and/or the effects of binarity.

Acknowledgements

I thank Prof Norbert Langer and Dr Abel Schootemeijer for their constant support and insightful supervision. It has been a pleasure working with the stellar research group at the University of Bonn. I want to thank Andrea Ercolino and Harim Jin for taking the time to help me understand the theory and tips on making my Python scripts more efficient.

I want to dedicate my work to my family back home in Scotland, without whom I would not have the support to push through challenging moments. I couldn't have crossed the finish line without them cheering me on.

Contents

CHAPTER 1

Introduction

When the Universe formed, hydrogen and helium were the only elements. Metals did not come to fruition in the Universe until the first stars formed and enriched the Universe through stellar feedback mechanisms such as stellar winds and supernovae. These metals were then present in the molecular clouds that formed the next generation of stars. The Sun is one such star that has formed from enriched molecular clouds and is one of the most studied objects except the Earth. As such, it is often easier to use the Sun as a standard to compare all other objects in the Universe. We typically use the physical parameters of the Sun, such as its radius, mass, and luminosity, to describe other objects. We can also compare the metallicity, [Fe/H], of other stars and galaxies to the metallicity of the Sun. Metal-poor galaxies are galaxies whose metallicity is lower than the solar value. As there were fewer metals in the early Universe, studying metal-poor galaxies can give us insightful information about the Universe.

Fortunately, we do not need to search far for a sufficiently metal-poor galaxy. Two metal-poor dwarf galaxies are near the Milky Way (MW), the Small and Large Magellanic Clouds (SMC and LMC). In this study, we focus only on the SMC, but the prospect of this work is not limited to only this dwarf galaxy. Due to its relatively close distance, $d = (62.44 \pm 1.28)$ kpc (Graczyk et al. 2020), as compared to other galaxies, we can study this dwarf galaxy in detail. The metallicity of the SMC is approximately one-fifth of that of solar metallicity ([Fe/H] = −0.7; Korn et al. 2000, Mokiem, de Koter, Vink et al. 2007). In this thesis, we will use photometric data of the SMC to develop a method to investigate how we can better use stellar evolutionary models to study the evolution of massive stars within a metal-poor dwarf galaxy. We do this by determining the effective temperature and bolometric luminosity of stars in the SMC and creating a Hertzsprung Russel diagram (HRD). We then create synthetic stellar populations by utilising a grid of stellar evolutionary models and compare the distribution of stars to that observed in the SMC. The stellar population with the closest match to observations will give us valuable insight into the physics inside stars. The method is detailed in Section 2.

Stellar evolutionary models, such as those generated with MESA[1] (Models for Experiments in Stellar Astrophysics), encounter a problem when modelling stellar populations such as the SMC. There is no consensus on the efficiency of the internal mixing parameters in stars. It is well known within the massive stellar research community that internal mixing parameters such as semiconvection and convective overshooting can alter the evolution of massive stars (Langer and Maeder 1995, Langer 2012, Schootemeijer, Langer, Grin et al. 2019). The efficiencies of these internal mixing parameters

[1] https://docs.mesastar.org/en/release-r23.05.1/index.html

determine a star's effective temperature and bolometric luminosity at a given age. Furthermore, mass-loss mechanisms such as stellar winds can affect the evolution of stars (Chiosi and Maeder 1986). However, metal-poor stars have weaker stellar winds (Mokiem, de Koter, Vink et al. 2007); therefore, the mass loss may become negligible compared to the effects of internal mixing efficiencies. Studying a metal-poor dwarf galaxy allows us to effectively neglect mass-loss mechanisms to better investigate the effects of internal mixing parameters on the evolution of stars. The internal mixing parameters should be fine-tuned to recreate the observed distribution of stars using a grid of models created by MESA. There are discrepancies in the number of blue supergiant (BSG) and red supergiant (RSG) stars between low- and high-metallicity stellar populations. No set of internal mixing parameters can accurately recreate the observed blue-to-red supergiants ratio (B/R) for both low- and high-metallicity galaxies (Langer and Maeder 1995). We aim to investigate not only if the total observed B/R of the SMC can be consistently calculated from stellar evolutionary models by the efficiency of semiconvection and overshooting, but also if we can reproduce the observed trend of the B/R over a range of bolometric luminosity intervals. Otherwise, if we cannot recreate this trend, we can comment on the need for the inclusion of rotation and binarity, which are considered non-negligible factors in determining the observed B/R in galaxies. Rotation was discussed by Schootemeijer, Langer, Grin et al. 2019. Their work showed that rotation did not significantly change the B/R in the SMC except in extreme cases. In these cases, stars may evolve as chemically homogeneous, but there are few observed stars with an extreme surface rotation, and therefore, we do not further investigate rotation. As most massive stars are born with a companion (Sana et al. 2012), binarity may be important in determining the B/R. However, binarity was out of this project's scope and will not be considered. Therefore, we take a first approach to determine if the efficiencies of the internal mixing parameters already provide observed the ratio of blue-to-red supergiants in the SMC.

1.1 Internal mixing parameters

The interiors of stars are obscured due to the scattering of photons in a high-density environment. Therefore, it is difficult for observers to study exactly what occurs inside a star. Convective and radiative energy transport are the two main mechanisms by which energy is transported throughout stellar interiors. Convective and radiative energy transport occurs in different regions of stars depending on their mass. For low-mass main-sequence (MS) stars, the star is entirely convective. Intermediate stars have radiative cores and convective envelopes. High-mass have convective cores and radiative envelopes. The boundary region at the interface where convective and radiative energy transport meet is the Schwarzchild boundary.

1.1.1 Convection

Convection is where the motion of particles is determined by the radiative temperature gradient $\nabla_{\rm rad}$ in a given region of the star, where the gas fulfils the criterion $\nabla_{\rm rad} > \nabla_{\rm ad}$. Where $\nabla_{\rm ad} = \left(\frac{\rm dln(T)}{\rm dln(P)}\right)_{\rm ad}$ is the adiabatic temperature gradient. Here, T is the temperature, and P is the pressure, both taken at the mass coordinate m/M. A gas element within the star will move, rising from hot layers, exchanging its thermal energy with the cooler layers above and then sinking again. As the core of the star can reach high enough temperatures such that fusion can take place, the composition of the core is different from the cooler outer layers. This produces a chemical gradient. A gas element can contain a different chemical composition than its surroundings after they sink/rise. Therefore, processed material can be mixed throughout the star, and unprocessed material can be brought to the core to be further processed,

extending the star's lifetime by providing more fuel.

1.1.2 Radiation

Radiative energy transport is where photons carry energy from the stellar interior towards the star's surface via photons. A photon emitted from one particle is absorbed by a neighbouring particle, which further emits another photon. This occurs when $\nabla_{\rm rad} < \nabla_{\rm ad}$. As there is no particle motion here, processed material from the star's interior is not mixed into the layers above and unprocessed material is not brought to active burning regions where fusion occurs.

These two mechanisms, convection and radiation, are the dominant processes in stars. However, two other processes may further aid the mixing of material within the stars. These are semiconvection and convective overshooting.

1.1.3 Semiconvection

In massive stars with a convective core, as the core fuses hydrogen during the MS, the hydrogen fraction in the core decreases. This decreases the core's opacity, κ, which in turn decreases the radiative temperature gradient as $\nabla_{\rm rad} \propto \kappa$. Eventually, in a region where the radiative temperature gradient falls below $\nabla_{\rm ad}$, the region becomes stable against convection. Therefore, radiative energy transport begins to dominate. Over the course of the MS, the convective core shrinks due to the decrease in the core hydrogen fraction and consequent decrease in κ and $\nabla_{\rm rad}$. Thereby introducing a mean molecular gradient $\nabla_\mu = \bar{m}/m_H$, where $\bar{m}$ is the mean mass of particles in the star, and m_H is the mass of hydrogen.

In the existence of a stabilising mean molecular gradient, semiconvection takes place. Semiconvection occurs when $\nabla_{\rm ad} < \nabla_{\rm rad} < \nabla_{\rm ad} + f\nabla_\mu$, where f is defined as in Schootemeijer, Langer, Grin et al. 2019. Limited mixing occurs in regions where this condition is valid until the mean molecular weight gradient is flattened and radiative energy transport dominates. This mixing may bring more fuel to the active burning regions below and nuclear-processed material to the envelope, affecting the star's stellar structure and the burning phases' lifetime. With more semiconvective mixing, the star's core increases, resulting in more blue supergiant stars.

1.1.4 Convective overshooting

The material in convective zones can end up in non-convective regions through the Schwarzschild boundary due to a process called convective overshooting. Overshooting occurs when the convective motion of a gas element causes the gas element to reach out of the convective zone. This can be directed towards the core or into the envelope, depending on the location of the convective zone. Unprocessed nuclear material can be brought to the core, providing more fuel for the star and increasing the star's main-sequence lifetime. This further results in the star having a different stellar structure than if overshooting did not exist by mixing nuclear-processed material. In the later evolutionary stage, where helium burning occurs in the core, overshooting also fuels the helium core, prolonging the helium burning phase. However, the stellar evolutionary tracks we show here only implement convective overshooting during the main sequence. Overshooting enriches the chemical composition of the star's envelope, changing the star's effective temperature and bolometric luminosity by altering the opacity of the envelope. In cases where overshooting is efficient, semiconvective regions do not develop as

mean molecular weight gradients are washed out. Therefore, overshooting can modify the star's physical properties, and more efficient overshooting can result in more red supergiant stars.

1.2 Supergiants

Supergiant stars are sensitive probes on the internal mixing efficiencies. Supergiant stars are post-main sequence massive stars with masses typically over $7\,M_{\odot}$. These stars have finished core hydrogen burning and have started core helium burning. Due to their high mass, these stars are very luminous, with luminosities above $1 \times 10^3\,L_{\odot}$. These luminous stars are easily observed in the SMC (Dufton et al. 2019, Ramachandran et al. 2019, Davies, Crowther and Beasor 2018, Bouret et al. 2013, Hunter et al. 2008, Mokiem, de Koter, Evans et al. 2006 , Trundle, D. J. Lennon et al. 2004, Venn 1999), and so they provide a useful tool to study the internal mixing efficiencies in a metal-poor dwarf galaxy. Supergiant stars typically follow the same path after the main sequence; the star contracts after core hydrogen exhaustion until hydrogen shell burning begins. After an active shell-burning region is created, the mirror principle takes effect, and as the core further contracts until core helium burning is achieved, the envelope expands. This results in an increase in luminosity and a decrease in effective temperature. The effective temperature determines whether we observe this star as blue or red, where blue stars have a higher effective temperature than red stars. We will see in this work that modifying the internal mixing parameters controls how many blue and red supergiant stars will form. We will compare the ratio of blue-to-red supergiants to the observed ratio by changing these internal mixing parameters in Section 4.2. It may not appear significant at first, but whether the star appears blue or red can give us insight into the internal stellar structure, which affects the star's end phase. The compactness of the core of these stars decides whether the star explodes as a supernova or collapses into a black hole or neutron star (Sukhbold and Woosley 2014). If, in the latter case where a black hole forms, the star is also in a binary system with another black hole, a merger event is possible. This could result in the release of gravitational waves. Therefore, supergiant stars could be the key to further understanding the evolution of our Universe.

CHAPTER 2

Methods

Recent access to GAIA DR3 data (Gaia Collaboration, Vallenari et al. 2023) paved the way to use a more complete set of observational data. We can now determine the physical parameters of a larger volume of stars than before, allowing stronger and more valid conclusions to be drawn from these observations. The following section details the methods taken to convert photometric data into physical parameters of the stars and determine the number of blue and red supergiant stars in the SMC. The photometric data is collected and reduced to contain only relevant sources. Assumptions were made on the visual extinction and surface gravity of these sources. Together with the known metallicity of the SMC and bolometric corrections from synthetic stellar spectra, the effective temperature and bolometric luminosity, along with their associated errors, were calculated for each source to produce a Hertzsprung-Russell diagram (HRD).

2.1 Photometry

The following photometric data was accessed via the GAIA archive as this gave the most recent and most complete data set of the SMC[1]. The catalogue used was GAIA DR3. The filters we used are G, G_{bp}, and G_{rp}, the coverage of these filters can be found on the Spanish Virtual Observatory (SVO)[2] (Rodrigo, Solano and Bayo 2012, and Rodrigo and Solano 2020).

The search parameters were initially set to include all stars in a boxed region from (RA, DEC) = (3.0, -75.5) to (25.0, -70.0). A magnitude limit of $G \leq 18$ was used to limit the number of faint sources. The resulting data set contained 335,337 sources. This region was selected as it encompasses the entire SMC. However, the contamination of this data set from stars that do not belong to the SMC is presumed to be significant due to the lack of further search conditions. We, therefore, explore methods to reduce the contamination of non-SMC stars.

2.1.1 Foreground removal

To counter this contamination problem, we investigated the proper motions and parallaxes of the sources within our catalogue, similar to the work of Yang et al. 2019. Their research determined likely members

[1] https://gea.esac.esa.int/archive/

[2] http://svo2.cab.inta-csic.es/svo/theory/fps/

of the SMC by producing histograms of the source's parallax and proper motion in right ascension (RA) and declination (DEC). Similar steps were taken here to identify potential members of the SMC. Firstly, histograms in the parameter and respective parameter error plane were created as shown in Figures 2.1, 2.2, and 2.3. Figure 2.1 is for the proper motion $\mu_{\rm RA}$ in right ascencsion, Figure 2.2 is for the proper motion $\mu_{\rm DEC}$ in declination, and Figure 2.3 is for the parallax π. The distribution of proper motions and parallaxes followed a somewhat Gaussian distribution arising from the statistical uncertainty of these measurements and the spread of proper motions. The peaks of the distributions were identified and compared to the results from Yang et al. 2019 and presented in Table 2.1. As our peak values lie close to their work, we continued with this approach to model a Gaussian fit. A selection of sources within 5σ of this peak were selected as likely members of the SMC. This was done to keep the number of sources chosen high enough to produce one of the most extensive photometric data sets to deduce the blue-to-red supergiant ratio.

It is possible that some of the sources are not members of the SMC but chance alignment with the proper motion and parallax of the sources in the SMC. As this foreground removal method tests the member status of sources with two independent parameters, we expect the number of chance alignments to have an insignificant effect on the determined ratio of blue-to-red supergiants. Therefore, the probability of a chance alignment was not studied in more detail. Furthermore, the SMC lies out of the plane of the Milky Way, where the majority of the foreground stars are. This helps reduce the number of stars, which could contaminate our source list.

The number of sources that remain after foreground removal was determined to be 217, 400, a reduction of our sample by approximately one-third.

	$\mu_{\alpha,peak}$ [mas/yr]	$\sigma(\mu_\alpha)$ [mas/yr]	$\mu_{\delta,peak}$ [mas/yr]	$\sigma(\mu_\delta)$ [mas/yr]	π_{peak} [mas]	$\sigma(\pi)$ [mas]
This work	0.693	0.201	-1.239	0.119	-0.007	0.066
Yang et al. 2019	0.695	0.240	-1.206	0.140	-0.009	0.066

Table 2.1: Resulting fit parameters of the Gaussian distributions for proper motions and parallax in columns 2-7. Results from this work are in the first row, and results from Yang et al. 2019 are in the second row.

Figure 2.4 shows the colour-magnitude diagram (CMD) of the stars in our catalogue, where green data points highlight the final selection of sources after foreground removal. From Figure 2.4, we see the bright yellow sources are removed from the final selection. These stars are at a shorter distance than the SMC and appear very bright in the CMD. There are also a collection of faint red stars not selected in our catalogue; these are faint foreground stars.

Figure 2.5 is a map of sources in the boxed region we downloaded from the GAIA archive. Again, green data points are for the final selection of sources after foreground removal. From Figure 2.5, it is clear that 47 Tucanae, which resides within (RA, DEC) = (4.5, -72.5) to (7.0, -71.6) and is represented by a grey cluster of sources, is not part of the final selection. 47 Tucane is a globular cluster at a distance of $d = (4.00 \pm 0.34)$ kpc (McLaughlin et al. 2006), and therefore should not be included in our catalogue.

Removing 47 Tucanae indicates that this method of foreground removal can remove foreground sources.

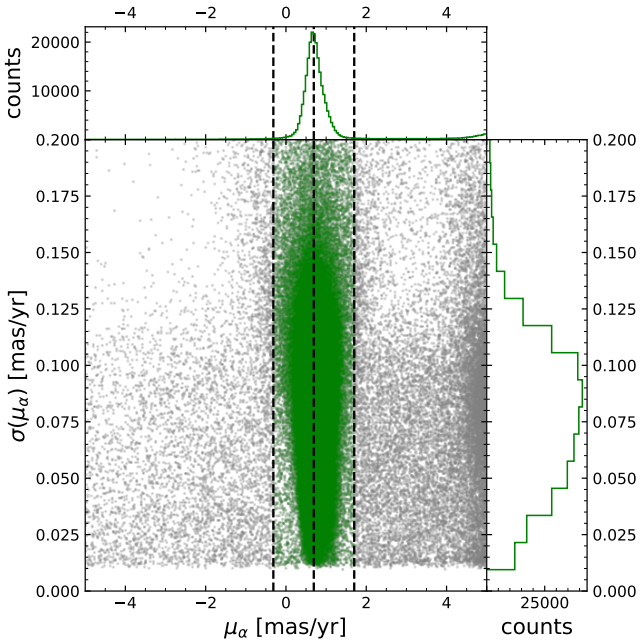


Figure 2.1: Proper motions in RA of sources, green indicates the selected sources after foreground removal.

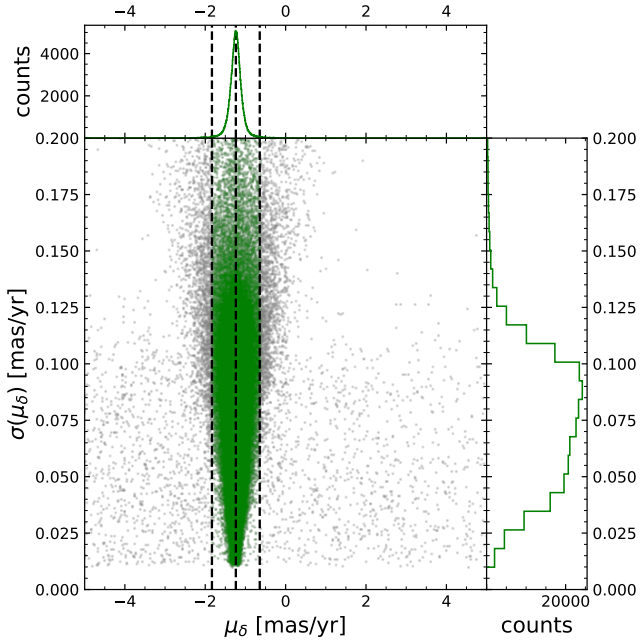


Figure 2.2: Proper motions in DEC of sources, green indicates the selected sources after foreground removal.

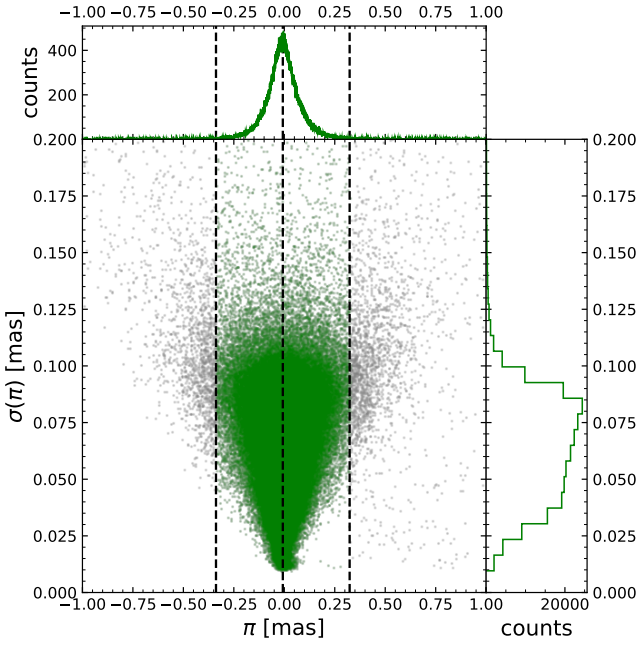


Figure 2.3: Parallax of sources, green indicates the selected sources after foreground removal.

2.1.2 Background sources

Background sources are trickier to remove as a small uncertainty in parallax results in a large uncertainty in the distance, seen by Equation 2.1 adapted from Bailer-Jones et al. 2021. Equation 2.2 arises from Gaussian error propagation. The uncertainties in parallax are given in Table 3 of Gaia Collaboration, Brown et al. 2021 and are still referred to as the current uncertainty values in Gaia Collaboration, Vallenari et al. 2023. For fainter sources, there is a larger uncertainty on their parallax. Hot blue stars are typically fainter than cooler red stars in the G band, as they emit most of their light in bluer wavelengths. Thus the sources we are interested in, blue supergiants, fall in the regime of having a large uncertainty in parallax.

$$\bar{\pi} \sim \frac{1}{d} \tag{2.1}$$

$$\sigma(\bar{\pi}) = \left| -\bar{\pi}\frac{\sigma(d)}{d} \right| \tag{2.2}$$

We must look at the fractional parallax uncertainty (fpu = $\sigma(\pi)/\pi$) to derive distances from parallax. For large fpu, the distance determination is unclear because the denominator can be very close to 0 (Luri et al. 2018). For the SMC with a distance of $d = (62.44 \pm 1.28)$ kpc (Graczyk et al. 2020), $\pi = (1.600 \pm 0.033) \times 10^{-2}$ mas. However, Gaia Collaboration, Brown et al. 2021 reports a minimum uncertainty of $\sigma(\pi) = 0.02$ mas to 0.03 mas for $G < 15$ mag. If we take the value of $\pi = 0.016$ mas and $\sigma(\pi) = 0.066$ mas from Table 2.1, we can see that we have a high fpu. Therefore a distance measurement is subject to very large uncertainty, and we can't determine which sources are within the SMC and which are background stars. If we use this parallax and uncertainty to calculate the distance of the star using Equation 2.2, we find $d = (62.44 \pm 257.57)$ kpc, thus proving that we cannot use this as a method to remove background sources.

The number of bright background stars that could be in the magnitude range we are interested in is not expected to be significant. Most of the background stars would come from nearby field stars or very distant stars in other galaxies. The stars in distant galaxies will not be bright enough to affect our determination of the blue-to-red supergiant ratio. The nearby field stars should not compromise the quality of this data set as field stars are thought to originate from local star-forming regions, such as the SMC, rather than forming in isolated regions in space. Therefore, background stars could be SMC stars with an underestimated visual extinction.

A parallax limit was, therefore, not implemented in the search parameters of the GAIA archive, as this could result in a loss of SMC sources. The number of sources in our catalogue remains at 217, 400.

2.1.3 Final selection

We have now determined 217, 400 sources that most likely belong to the SMC. The selected sources are shown in Figure 2.4 and Figure 2.5.

We can now use the photometric data to determine the effective temperatures and bolometric luminosities of the SMC sources to produce a Hertzsprung-Russell diagram and compare the results with theoretical models.

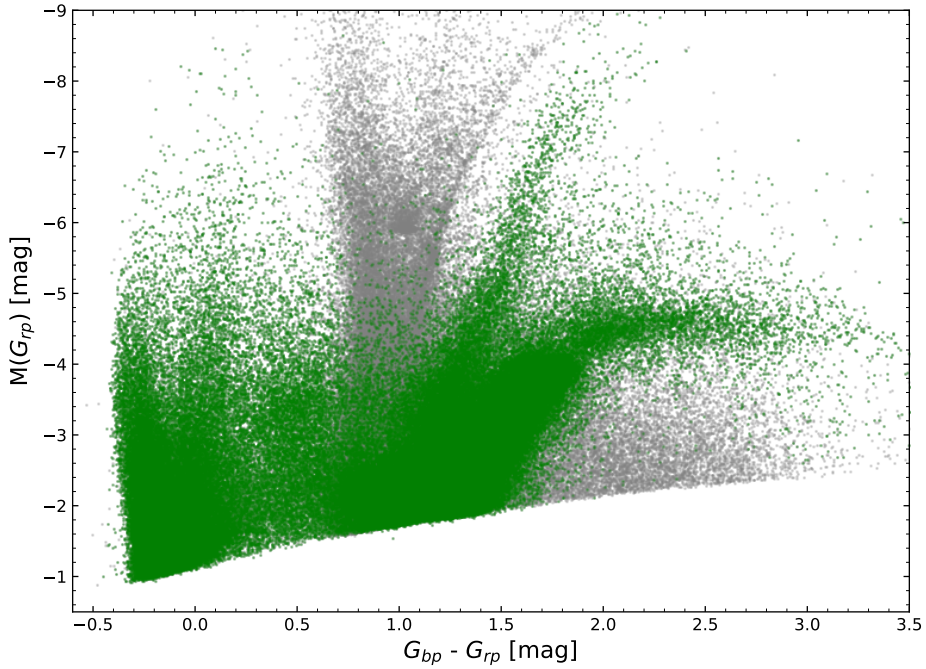


Figure 2.4: Colour-Magnitude diagram, green indicates the selected sources after foreground removal.

2.1.4 A note on OBe stars

Blue helium-burning stars occupy a specific region of the HRD close to the terminal-age main sequence (TAMS). If a main sequence star develops a significant circumstellar disc, the disc emits longer wavelengths and alters the photometric colour. Additionally, but less significantly, the disc can absorb light from the star and remit redder wavelengths. If we consider photometrically determined effective temperatures, a star with a decrection disc will appear redder and produce a cooler determined effective temperature than if there were no circumstellar disc. OBe stars are main sequence stars that can create a decretion disc due to their rapid rotation (Schootemeijer, D. J. Lennon et al. 2022). A hydrogen-burning star with a redder colour than expected will appear in the region of the HRD, where we expect helium-burning stars. This provides an issue for determining the blue-to-red supergiant ratio by artificially inflating the number of blue supergiants, as these OBe stars have no red supergiant counterpart. To determine the number of OBe stars in the SMC, Schootemeijer, D. J. Lennon et al. 2022 investigated stars with known spectral types. Schootemeijer, D. J. Lennon et al. 2022 used four filters and observed a region of a colour-colour diagram that OBe stars occupied. They determined that the majority of these OBe stars occupied a narrow of magnitudes and colour (see Figure 1 of their paper), which is represented in Figure 2.6. Our method here is thus to take the efforts performed before us and randomly discard sources within this region. We randomly discard these sources to avoid biases towards a type of star. Boxes were created in the CMD, and every second or third source was discarded depending on the ratio of OBe stars to other stars from Figure 1 of Schootemeijer, D. J. Lennon et al. 2022. This resulted in 3114 sources being removed from our selection, represented in Figure 2.7, which shows the absolute magnitude $M(G_{\rm rp})$ against colour $G_{\rm rp} - G_{\rm bp}$ for only the bluest region of the CMD. We remove

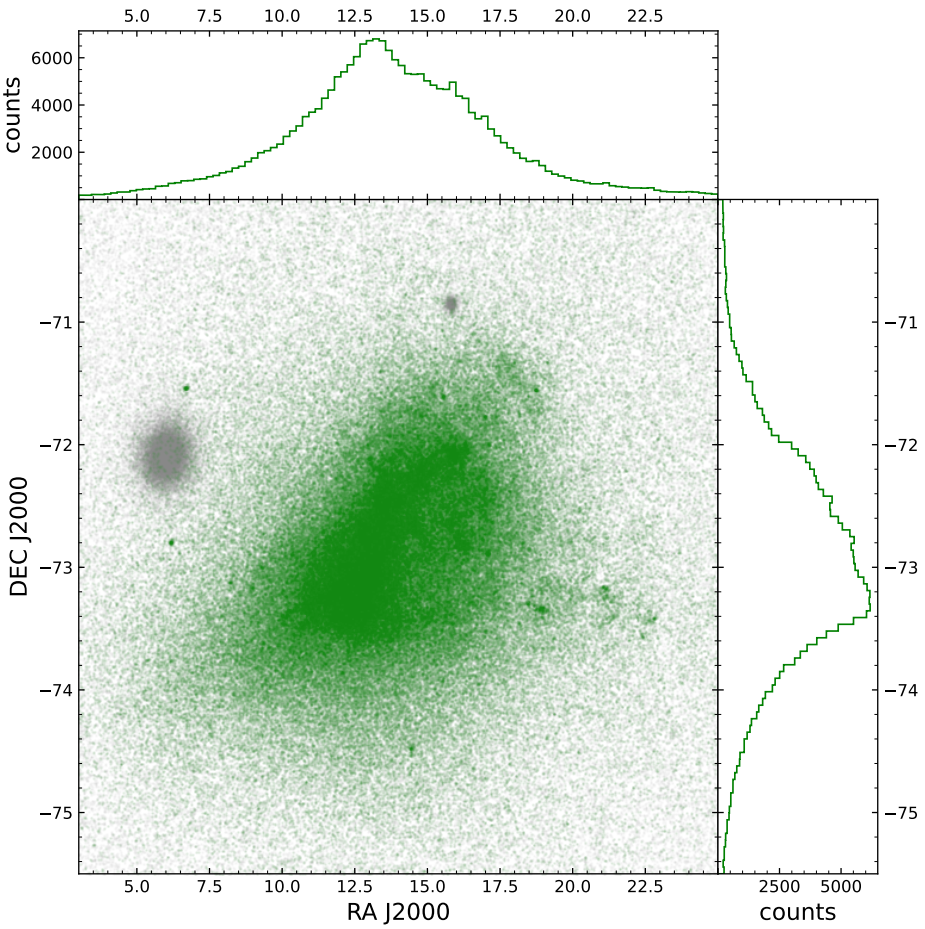


Figure 2.5: Map of sources, green indicates the selected sources after foreground removal.

these stars later in the blue-to-red supergiant ratio determination to preserve a smooth CMD and HRD; otherwise, an artificial terminal-age main sequence appears in the HRD (see Section 2.3).

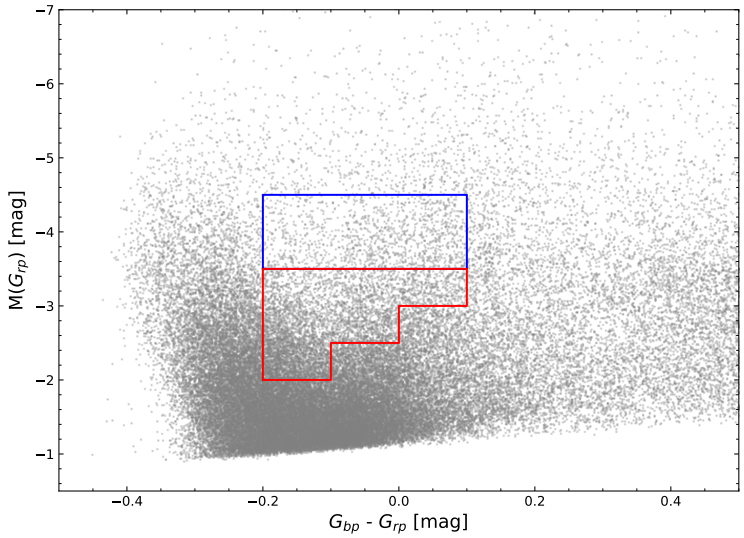


Figure 2.6: CMD showing the regions where OBe stars were randomly selected to counter the inflated blue supergiant problem. Every third star in the blue region was discarded, and every second star in the red region was discarded. The region of the CMD and how many stars to discard was decided in reference to Schootemeijer, D. J. Lennon et al. 2022.

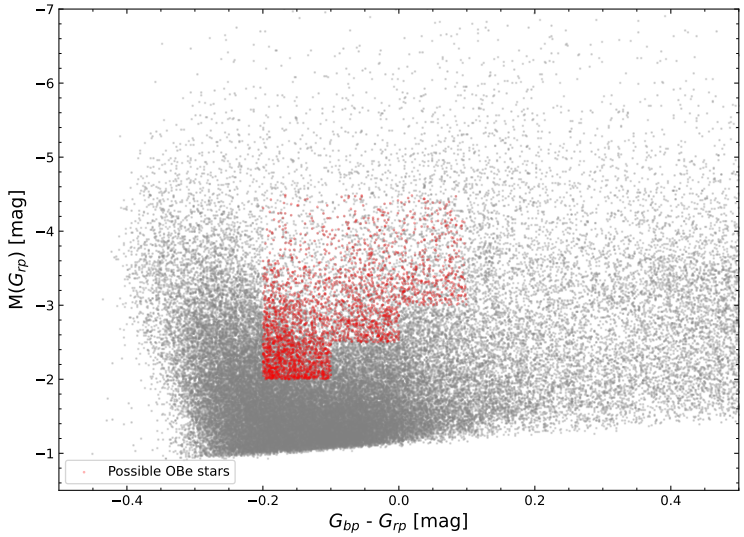


Figure 2.7: CMD highlighting the resulting 3114 possible OBe stars in the regions shown in Figure 2.6.

2.2 Calculating effective temperature and bolometric luminosity

To convert colours and magnitudes into physical parameters of the sources, we rely on previously computed theoretical stellar atmosphere model spectra. There exist several different catalogues we can work from, and two were investigated in detail: MIST[3] and PARSEC[4]. We proceeded with MIST due to the extensive parameter space of MIST with regard to visual extinction, surface gravity and metallicity.

MIST, or MESA Isochrones & Stellar Tracks, is a grid of single-star evolutionary models that are computed for a range of masses and metallicities (Dotter 2016, Choi et al. 2016). Also available on their website is a table of bolometric corrections, derived from stellar atmosphere models and blackbody assumptions, that allow the conversion of photometric magnitude to bolometric magnitude. This can, in turn, be used to determine the bolometric luminosity of the star by knowing the distance. We can also use the bolometric corrections in two filters, the same filters as our photometric data mentioned in Section 2.1, to derive a relation between colour and effective temperature for a given metallicity [Fe/H], visual extinction A_v and surface gravity $\log(g / cms^{-2})$.

The method to convert the colour and magnitude of the source into effective temperature and bolometric luminosity shall be the following.

1. We will make assumptions, specified in more detail below, on the visual extinction and surface gravity to create a colour-effective temperature relation to obtain the effective temperature for each star.

2. Once we obtain the effective temperature, we will create an effective temperature-bolometric correction relation to obtain the bolometric correction in G_{rp}.

3. We will use the bolometric correction of the star with the distance and visual extinction of the SMC to determine the bolometric luminosity of the star.

4. This process is done for a visual extinction $A_v = 0.3 \pm 0.1$ to quantify errors in this method; this is detailed more in Section 2.2.6.

2.2.1 Extinction

There have been many previous studies of the visual extinction of the SMC (Schootemeijer, Langer, D. Lennon et al. 2021, Dobashi et al. 2009, Gordon et al. 2003). Extinction will significantly affect the stars' derived effective temperatures and luminosities, especially in the luminous blue stars. Shorter wavelengths interact with interstellar dust more than longer wavelengths. Hence the effect of extinction is more significant for bluer stars. The SMC is an active star-forming region that contains regions of dense molecular clouds and also diffuse regions. An extinction map, such as the work of Dobashi et al. 2009, could be used to identify these regions of low and high extinction. However, these maps use low-mass stars which do not trace high-mass star-forming regions. Therefore, the visual extinction determined through extinction maps is typically lower than we expect for high-mass stars. As we are interested in supergiant stars, we consider using a single value of extinction with an associated uncertainty $A_v = 0.3 \pm 0.1$. This value is also determined as the peak of the distribution of extinction for luminous sources in the SMC by Schootemeijer, Langer, D. Lennon et al. 2021 (see Figure 6). An uncertainty

[3] https://waps.cfa.harvard.edu/MIST/
[4] http://stev.oapd.inaf.it/YBC/

of ±0.1 is considered a conservative range for the visual extinctions in the SMC. This is confirmed by Schootemeijer, Langer, D. Lennon et al. 2021, who show the second and third most common visual extinction in the SMC is $A_v = 0.2$ and 0.4. The bolometric correction tables from MIST give bolometric corrections with a visual extinction ranging from 0.0 to 6.0 in varying step sizes. We, therefore, are limited in the preciseness of which visual extinction we can use. We use the bolometric corrections for $A_v = 0.3 \pm 0.1$.

We need to determine the extinction in the G_{bp} and G_{rp} filters to correct the photometric data. To do this, we used the effective wavelength of these filters from Rodrigo, Solano and Bayo 2012 and Rodrigo and Solano 2020, which is the wavelength at which the filter is most sensitive. The effective wavelengths are $\lambda_{\text{eff, bp}} = 5\,041.61$ Å and $\lambda_{\text{eff, rp}} = 7\,690.74$ Å. Gordon et al. 2003 provided a table of sample averages from the SMC bar extinction curves (Table 4 of their paper). The data in this table was linearly interpolated to find the extinction in G_{bp} and G_{rp}. The extinction values $A_{G_{bp}} = 1.156 A_v$ and $A_{G_{rp}} = 0.637 A_v$ were determined.

We expect some stars whose extinction is overestimated. This will result in stars with temperatures hotter than the zero-age main sequence (ZAMS), which is not physically possible. Bluer helium-burning stars whose extinction was overestimated will appear in the main sequence region of the HRD and not be counted in the blue-to-red supergiant ratio. However, massive stars with low metallicity have weaker stellar winds than high-metallicity stars. Thus the possibility that one of these massive metal-poor stars has not cleared out the surrounding gas could be high (Vink and Sander 2021, Hainich et al. 2015). Therefore, overestimating visual extinction may not cause a significant problem in determining the B/R. Another possible drawback is that some stars will have their extinction underestimated, and thus main sequence stars will appear where we expect helium-burning stars in the HRD. Massive stars form in collapsing dark molecular clouds where the chance of these newly formed massive stars having high visual extinctions is non-negligible, and underestimating the visual extinction is very possible.

The effect of extinction is more significant for blue stars than red stars, as the peak of their emitted flux is at shorter wavelengths than for red stars. The red supergiant branch in the HRD is a clearly defined structure. Due to the weaker effect of extinction on redder stars, the risk of contaminating the red supergiant branch with red giants or asymptotic giant branch (AGB) stars whose extinction was underestimated is low. The result of a single value for the visual extinction hence affects blue stars more than red stars. If we underestimate the visual extinction, this could lead to an inflated blue-to-red supergiant ratio because we consider more blue main sequence stars to be burning helium. Therefore, we have adopted an uncertainty in the visual extinction of $\sigma_{A_v} = \pm 0.1$, and we will investigate how extinction changes the determined blue-to-red supergiant ratio. Visual extinction is the largest uncertainty in our effective temperature and bolometric luminosity determination. Visual extinction changes the source's colour significantly enough to lead to significant uncertainties in the effective temperature. This is then propagated through to bolometric luminosity by using the effective temperature to determine the bolometric correction. The results of this are shown in Section 4.

2.2.2 Surface gravity

As this project aims to determine the ratio of blue-to-red supergiants, we have made two assumptions on surface gravity. For redder stars, we assume a surface gravity of $\log(g / \text{cms}^{-2}) = 0.0$ and for bluer stars, we assume a surface gravity of $\log(g / \text{cms}^{-2}) = 3.0$. These selections were made based on the evolutionary tracks run by Schootemeijer, Langer, Grin et al. 2019 with $\alpha_{sc} = 10.0$ and $\alpha_{ov} = 0.23$. These mixing parameter efficiencies were suggested in the work of Schootemeijer, Langer, Grin et al.

2019. In Figure 2.8, we see that for hot blue stars, the surface gravity log(g / cms^{-2}) = 3.0 is a fair assumption, and for cooler red stars, the surface gravity log(g / cms^{-2}) = 0.0 is also a fair assumption. Now we need to determine which star can be classified as red and which one as blue. To do this, a boundary condition on photometric colour was implemented. A star whose colour exceeds this 'transition colour' would be classified as a red star with log(g / cms^{-2}) = 0.0, and a star whose colour is less than this transition colour would be a blue star with log(g / cms^{-2}) = 3.0. The transition colour between the two surface gravity regimes was selected based on when the assumption of a blackbody model atmosphere spectra was used (see Figure 6 of Choi et al. 2016). For A_v = 0.3 and [Fe/H] = −0.75, this transition colour was $G_{\rm bp} - G_{\rm rp}$ = 0.872 mag. This is more clearly represented later in the colour-effective temperature and effective temperature-bolometric correction relations. The surface gravity parameter space could be investigated further by using other surface gravities between log(g / cms^{-2}) = 0.0 and log(g / cms^{-2}) = 3.0. Still, these two surface gravities were deemed sufficient to determine a blue-to-red supergiant ratio.

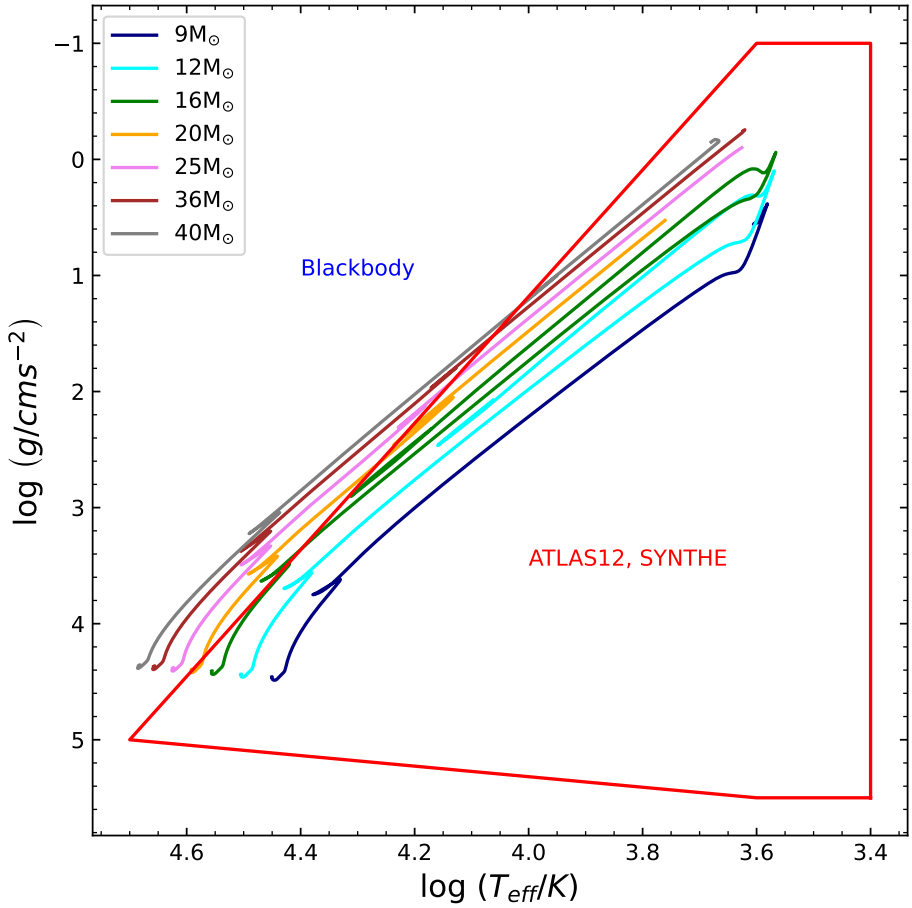


Figure 2.8: Evolutionary tracks in the log($T_{\rm eff}$/K) - log(g / cms^{-2}) plane for a range of intermediate-mass to high-mass stars from Schootemeijer, Langer, Grin et al. 2019. The red area is taken from Choi et al. 2016, where bolometric corrections were calculated from synthetic spectra using ATLAS12 and SYNTHE codes and a blackbody assumption was used outside this area.

2.2.3 Colour-effective temperature relation

To determine an effective temperature of a source, we need to create a colour-temperature relation. As mentioned in Section 2.2, we can use the bolometric corrections to calculate a 'bolometric colour' using Equation 2.2.3. We use this colour and effective temperature for $A_v = 0.3 \pm 0.1$, [Fe/H] = −0.75 and $\log(g\ /\ cms^{-2}) = [0.0, 3.0]$ to create colour-effective temperature relations.

$$G_{bp} - G_{rp} = BC(G_{rp}) - BC(G_{bp}) \tag{2.3}$$

No known function currently exists that can describe how bolometric colour varies over the range of effective temperatures from $\log(T_{eff}/K) = 3.4$ to $\log(T_{eff}/K) = 6.0$. Creevey et al. 2022 created a function[5] that takes $\log(g\ /\ cms^{-2})$, $\log(T_{eff}/K)$, [Fe/H], and A_v as input to determine the bolometric correction. The function provided is nothing more than interpolation between points in their bolometric correction tables. They did not provide a function to determine the effective temperature from colour. As we are interested in determining the effective temperature first and then the bolometric correction, we were required to build a different method. Linear interpolation was tested as a potential method; however, polynomial fitting of varying degrees was opted for instead as it provides smooth functions that did not cause sharp artefacts in the resulting HRD. Linear interpolation further resulted in several stars hotter than the ZAMS. A double-decaying exponential function was tested, but the function failed to fit the data points at low effective temperatures, resulting in large residuals between the data and the fit.

In Figure 2.9, the effective temperatures and bolometric colours are calculated using the bolometric corrections from MIST, for a metallicity of [Fe/H]= −0.75, surface gravity of $\log(g\ /\ cms^{-2}) = [0.0, 3.0]$, and visual extinction of $A_v = 0.3$ are plotted. It is clear from this figure that for high temperatures, the bolometric corrections are the same for the two different surface gravities, as the red and blue data points overlap. This is the direct consequence of implementing a blackbody assumption (see Figure 2.8). The function plateaus at high temperatures due to the Rayleigh-Jeans tail of the Planck function. As the peak of the Planck function moves towards smaller wavelengths for higher temperatures, the difference of the integrated flux between the Gaia filters G_{bp} and G_{rp} remains relatively constant for increasing temperatures. Due to the plateau of the colour-effective temperature relation, a slight colour change can lead to a significant difference in the determined effective temperature and, thus, the star's position on the HRD.

Different degree polynomials were used to fit the data points for the low- and high-temperature regimes. A degree 3 polynomial was used for the red stars, and for the blue stars, a degree 8 polynomial. These polynomials provided the smoothest functions and best-fit functions to the data points. The fits for different visual extinction values $A_v = 0.3$, 0.2, and 0.4 can be seen in Figures 2.9, A.1, and A.2 respectively. The residuals of these fits for $A_v = 0.3$ are shown in Section 2.2.6. The residuals are approximately 0.05 in $\log(T_{eff}/K)$ for effective temperature and 0.1 mag in bolometric correction, but these are insignificant compared to the uncertainty in the visual extinction.

The star's colour was supplied to the polynomial functions to determine the effective temperature. The colour of the star determined which function to use. If $G_{bp} - G_{rp} \geq 0.872$, then the red 3rd degree polynomial function was used, and if $G_{bp} - G_{rp} < 0.872$, then the blue 8th degree polynomial function was used. The bluest colour limit was set at $G_{bp} - G_{rp} = -0.41$ for $A_v = 0.3$ and [Fe/H] = -0.75; an

[5] https://www.cosmos.esa.int/web/gaia/dr3-bolometric-correction-tool

effective temperature could not be determined for stars bluer than this colour. As seen in the CMD (Figure 2.4), the number of stars with $G_{\rm bp} - G_{\rm rp} > 3.25$ drops significantly.

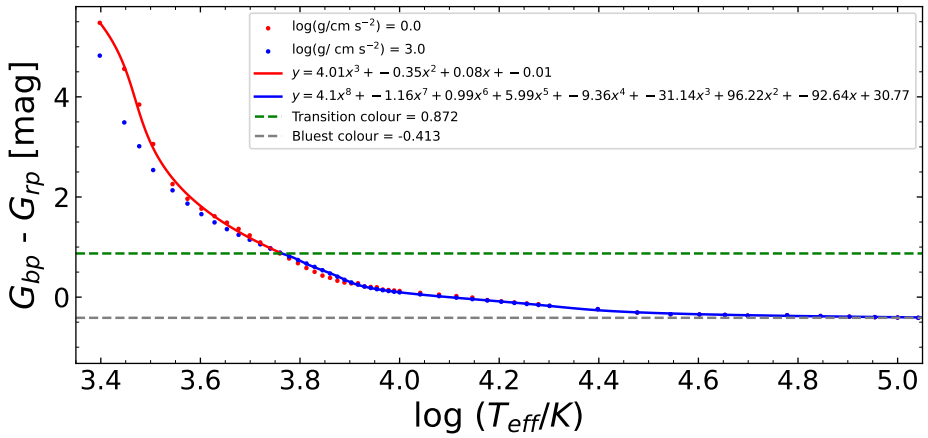


Figure 2.9: Colour-effective temperature relation for $A_v = 0.3$ using bolometric corrections and effective temperatures from MIST. The red points are for a log(g / cms^{-2}) = 0.0, and the blue points for log(g / cms^{-2}) = 3.0. Red stars are those with colours above the green line, and blue stars are those with colours below the green line.

2.2.4 Effective temperature-bolometric correction relation

A similar approach for the colour-effective temperature relations was used for the effective temperature-bolometric correction relations. Polynomial fits were used to create a smooth function to determine the bolometric correction in the $G_{\rm rp}$ filter for the previously calculated effective temperatures. Figures 2.10, A.3, and A.4 show the relationship between bolometric correction and effective temperature for the two surface gravities and the two fits used for $A_v = 0.3$, 0.2, and 0.4 respectively. The corresponding transition effective temperature was $\log(T_{\rm eff}/{\rm K}) = 3.758$ and was calculated using the transition colour for [Fe/H] = −0.75, log(g / cms^{-2}) = 0.0, and $A_v = 0.3$. Figure 2.10 represents it as a green line. Bolometric corrections were calculated for all sources for the three different extinction values. As we can see from Figure 2.8, this transition temperature matches well with the effective temperature at which the blackbody approximation begins for the MIST bolometric corrections for log(g / cms^{-2}) = 0.0.

To determine the bolometric correction, the star's effective temperature was supplied to the polynomial functions. The effective temperature of the star determined which function to use. If $\log(T_{\rm eff}/{\rm K}) \leq 3.763$, then the red 2$^{\rm nd}$ degree polynomial function was used, and if $\log(T_{\rm eff}/{\rm K}) > 3.763$, then the blue 4$^{\rm th}$ degree polynomial function was used.

2.2.5 Bolometric luminosity

The bolometric corrections were used together with the extinction and distance to the SMC to determine the bolometric magnitude and bolometric luminosity. This was done using the following equations (Schootemeijer, Langer, D. Lennon et al. 2021):

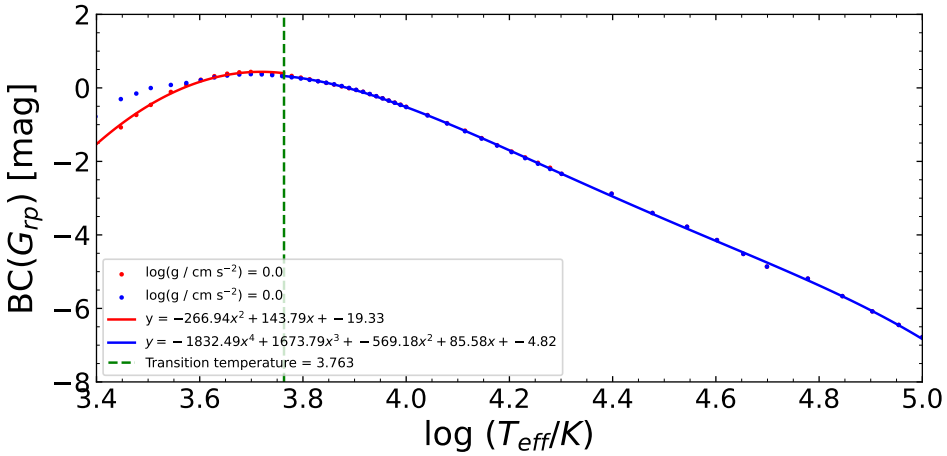


Figure 2.10: Temperature-bolometric correction relation for $A_v = 0.3$ using bolometric corrections and effective temperatures from MIST. The red points are for a log(g / cms^{-2}) = 0.0, and the blue points are for log(g / cms^{-2}) = 3.0. Red stars are those with colours below the green line, and blue stars are those with colours above the green line.

$$DM = 5 \cdot \log_{10}\left(\frac{d}{10\text{pc}}\right) \tag{2.4}$$

$$M_{\text{bol}} = G_{\text{rp}} + BC(G_{\text{rp}}) - DM - A_{G_{\text{rp}}} \tag{2.5}$$

$$\log_{10}\left(\frac{L}{L_\odot}\right) = -0.4 \cdot (M_{\text{bol}} - M_{\text{bol},\odot}) \tag{2.6}$$

Where DM is the distance modulus, d is the distance to the SMC in parsecs, G_{rp} is the apparent magnitude in the G_{rp} filter, BC(G_{rp}) is the bolometric correction in the G_{rp} filter determined in Section 2.2.4, $A_{G_{\text{rp}}}$ is the extinction in the G_{rp} filter, L is the bolometric luminosity, M_{bol} is the absolute bolometric magnitude, and $M_{\text{bol},\odot}$ is the absolute bolometric magnitude of the Sun which was taken as 4.74 (Mamajek et al. 2015). The bolometric luminosity was calculated for $A_v = (0.2, 0.3, 0.4)$.

2.2.6 Input Errors

To quantify the uncertainties in the derived effective temperature, bolometric correction, and bolometric luminosity, we investigated three different sources of errors. Firstly, we checked the uncertainties in the photometric measurements from GAIA; these are shown in Figure 2.11 and Figure 2.12. For G_{rp}, the uncertainty peaks at a value of $\sigma(G_{\text{rp}}) = 0.01$ and for the $G_{\text{bp}} - G_{\text{rp}}$ the uncertainty peaks at a value of $\sigma(G_{\text{bp}} - G_{\text{rp}}) = 0.015$ mag.

Secondly, we investigated the uncertainties due to the residuals of the polynomial fitting functions. The residuals can be found in Figures 2.13 and 2.14. For log(T_{eff}/K), the uncertainties are less than $\sigma(\log(T_{\text{eff}}/\text{K})) = 0.1$ and for BC(G_{rp}), the uncertainties are less than $\sigma(\text{BC}(G_{\text{rp}})) = 0.15$.

Lastly, and most importantly, we investigated the uncertainty in visual extinction $\sigma(A_v) = 0.1$ as discussed in Section 2.2.1. To quantify the uncertainty in effective temperature, we corrected the

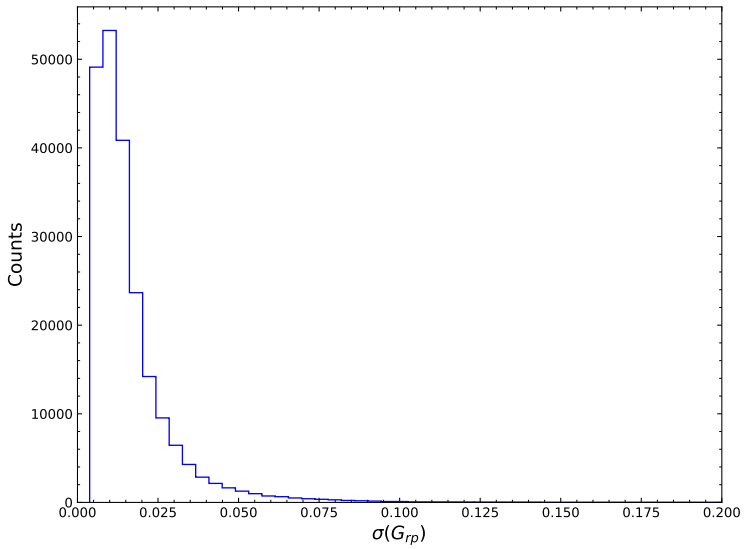


Figure 2.11: Histogram of the uncertainties in G_{rp} of the 217,400 selected sources.

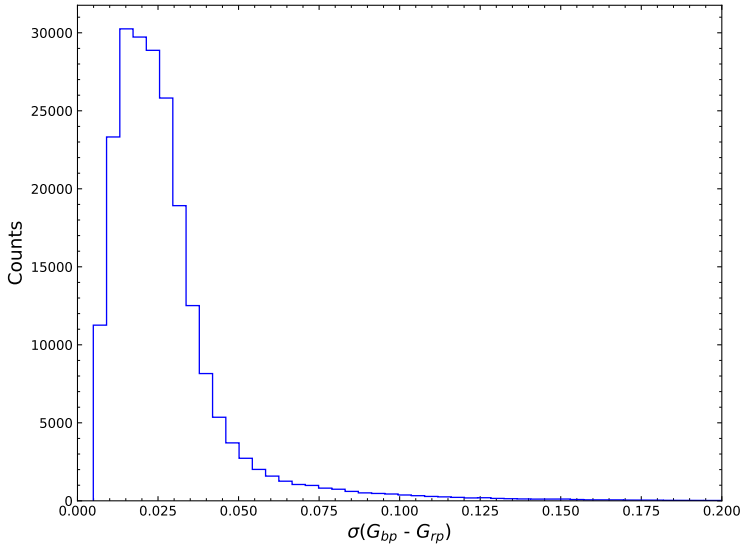


Figure 2.12: Histogram of the uncertainties in $G_{\mathrm{bp}} - G_{\mathrm{rp}}$ the 217,400 selected sources.

photometric data for a visual extinction of $A_v = 0.2$, 0.3, and 0.4. This led to three different colours per source. Then we solved the colour-effective temperature relation for each colour. This resulted in three effective temperatures, one for each visual extinction. The upper bound on the effective temperature was with a visual extinction of $A_v = 0.4$ as we assume the star is bluer than it appears. The lower bound on the effective temperature was for $A_v = 0.4$. The three effective temperatures were then supplied to the effective temperature-bolometric correction relation to derive three bolometric corrections. The three bolometric corrections were then paired with their corresponding visual extinction to calculate three

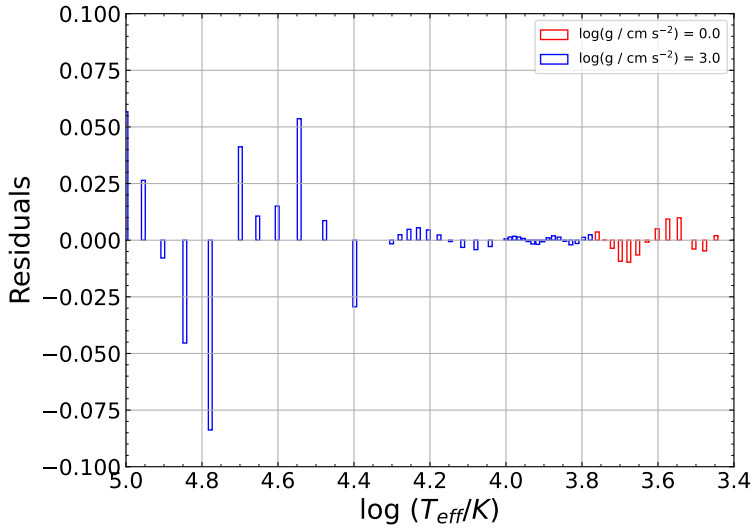


Figure 2.13: Residuals between the MIST data and the polynomial fits for the colour-effective temperature relation.

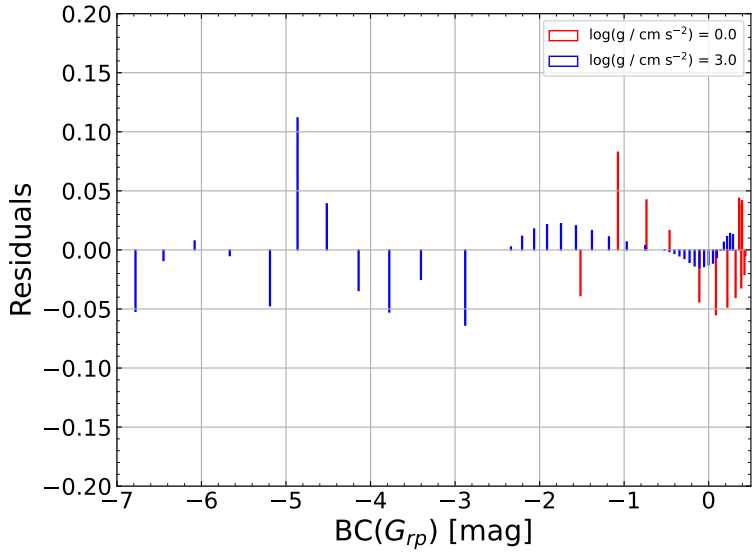


Figure 2.14: Residuals between the MIST data and the polynomial fits for the effective temperature-bolometric correction relation.

bolometric luminosities according to the equations in Section 2.2.5. Gaussian error propagation was also used for the distance modulus and bolometric magnitude. The effective temperature and bolometric luminosity uncertainties can be found in Section 2.3.1.

2.3 Creating a Hertzsprung-Russell diagram

The resulting Hertzsprung-Russell diagram, including the OBe stars, was created and is shown in Figure 2.15. To correct for OBe star contamination, the OBe stars will be removed when considering the blue-to-red supergiant ratio. However, removing these sources from the HRD before calculating the blue-to-red supergiant ratio introduced the terminal-age main sequence's apparent but artificial location (see Figure A.5). Therefore, the TAMS would be hotter than observed, resulting in our method of counting blue and red helium-burning stars to produce more blue helium-burning stars. Figure 2.16 shows the location of the OBe stars in the HRD. We can identify some structures in the HRD that are expected. At $\log(T_{\rm eff}/{\rm K}) \leq 3.7$, we see two branches, the red giant branch (RGB), which further leads into the asymptotic giant branch (AGB) at lower temperatures and luminosities, and the red supergiant (RSG) branch which extends to higher luminosities. At $\log(T_{\rm eff}/{\rm K}) \geq 4.0$, we can see a dense region that represents the main sequence of intermediate- and high-mass stars. Due to the magnitude limit in our survey, we do not see the low-mass main sequence.

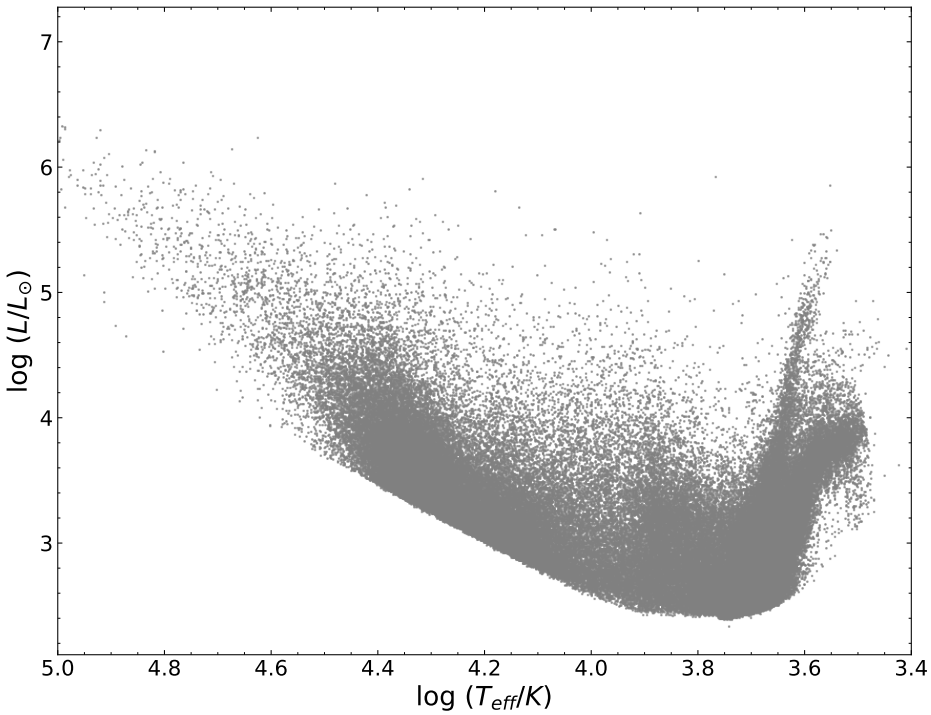


Figure 2.15: Hertzsprung-Russell diagram of stars in the Small Magellanic Cloud.

2.3.1 Errors in HRD position

As mentioned in Section 2.2.6, the uncertainty in the visual extinction is the most dominant source of error. Figure 2.17 shows how extinction affects the position of the source in the HRD. The red arrows indicate an increase in extinction from $A_v = 0.3$ to $A_v = 0.4$, and the blue arrows indicate a decrease in extinction from $A_v = 0.3$ to $A_v = 0.2$. As expected, an increase in the extinction shifts the colour of

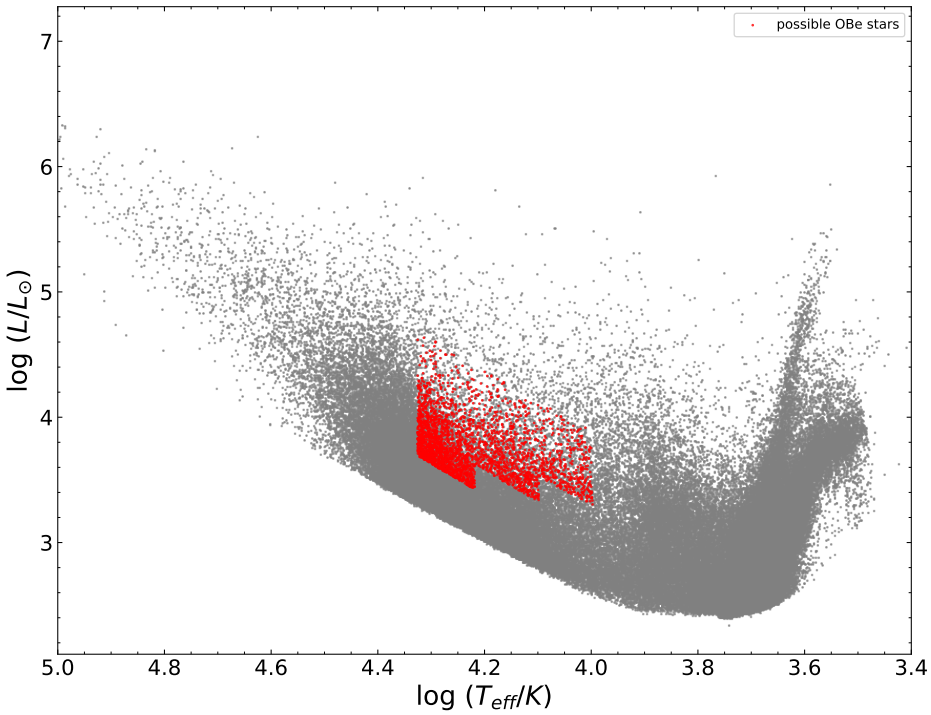


Figure 2.16: Hertzsprung-Russell diagram of stars in the Small Magellanic Cloud highlighting OBe stars.

the star towards bluer colours and hence a higher effective temperature is determined. An increase in extinction increases the determined bolometric luminosity of the star as we assume that a lower fraction of flux from this source reaches us so that the non-extinct magnitude would be brighter and, therefore, the bolometric luminosity is larger. We confirm that for bluer stars, extinction has a more significant effect on the position of the star in the HRD than for redder stars.

To quantify the errors further, Figure 2.18 and Figure 2.19 show the uncertainties $\sigma(\log(T_{\rm eff}/{\rm K}))$ vs. $\log(T_{\rm eff}/{\rm K})$ and $\sigma(\log(L/L_\odot))$ vs. $\log(T_{\rm eff}/{\rm K})$ respectively. We see an upward trend for the uncertainty in effective temperature and bolometric luminosity for higher effective temperatures. Deviations away from the dense main body of sources are likely due to anomalies in the uncertainties of the photometric data. Polynomial functions were fitted for both low and high temperatures to provide functions on how the uncertainties change with effective temperature. We will use the functions in Section 4.

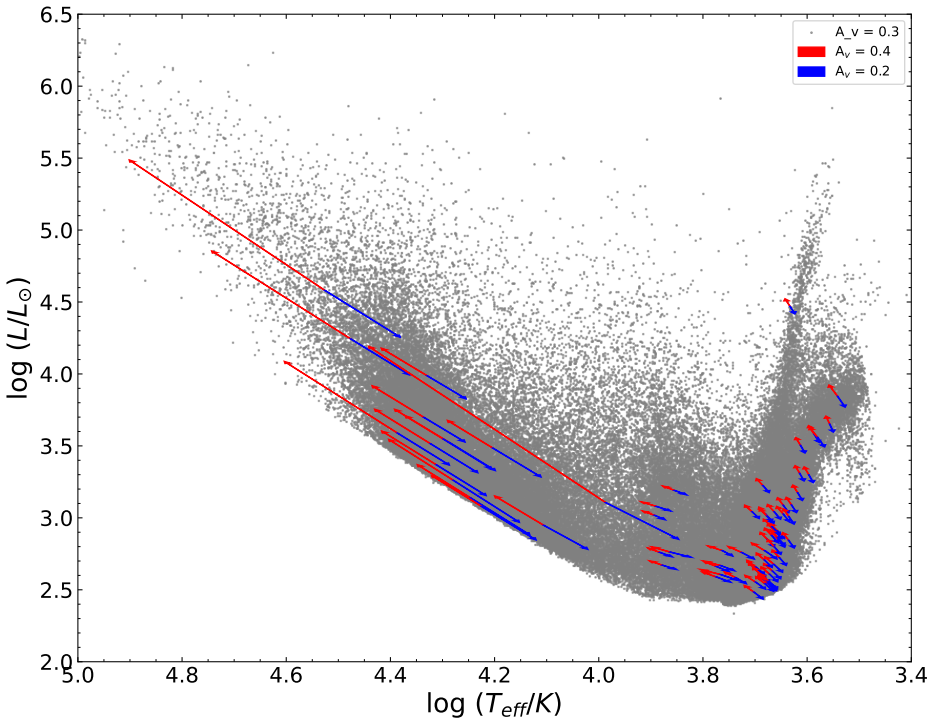


Figure 2.17: Hertzsprung-Russell diagram of stars in the Small Magellanic Cloud highlighting the effects of extinction.

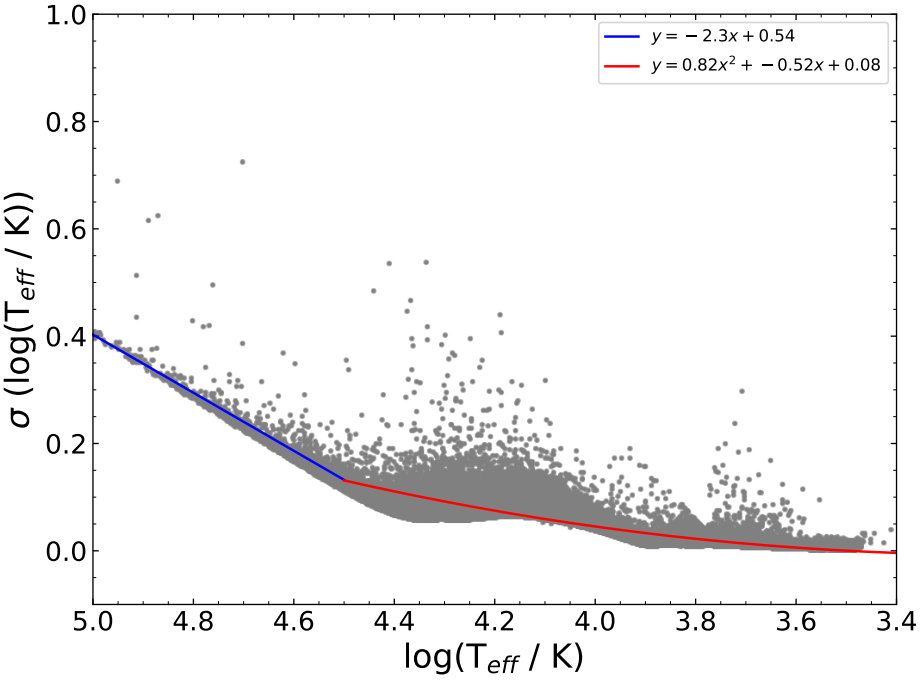


Figure 2.18: Uncertainties in the effective temperature as a function of effective temperature.

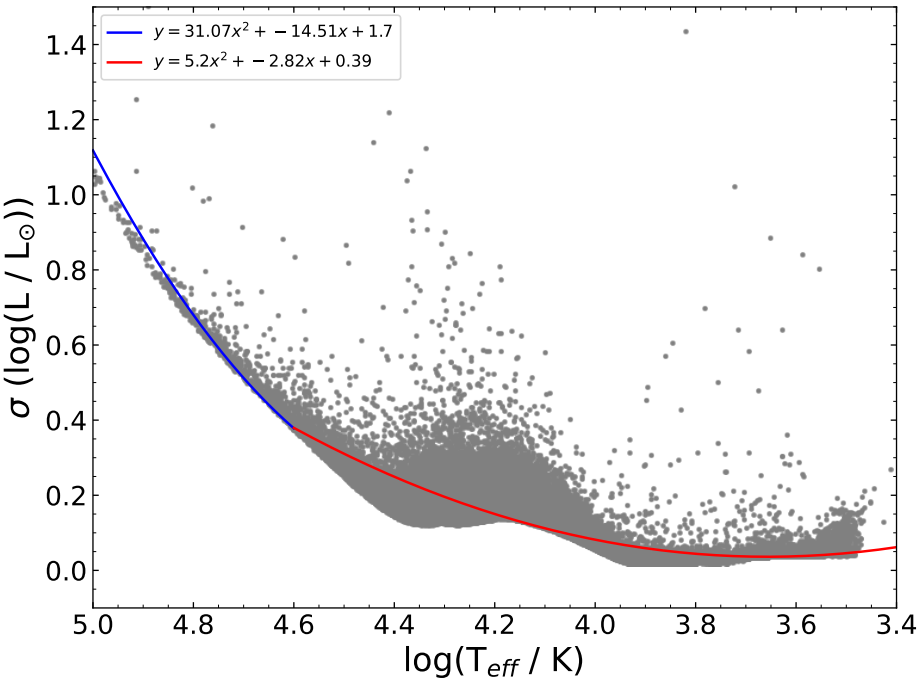


Figure 2.19: Uncertainties in the bolometric luminosity as a function of effective temperature.

2.3.2 Identifying the main sequence

To count the number of blue and red helium-burning stars, we need to separate them from the main sequence stars whose extinction is underestimated and therefore have lower effective temperatures. We must identify the main sequence and remove the main sequence stars from our data set. Effective temperature histograms of the sources in seven luminosity bins were created, shown in Figure 2.20. The expected distribution of stars should look like a sharp increase in the number of stars at temperatures around the ZAMS (Feltzing, Holmberg and Hurley 2001). The main sequence lasts for approximately 90% of the lifetime of stars, most of which is spent close to the ZAMS. As the SMC has regions of active star formation, we should see a high number of stars close to the ZAMS. Then the number of stars should decrease towards the TAMS. After the TAMS, we do not expect the number of stars to fall off entirely. Instead, we expect helium-burning stars. These stars should make up approximately 10% of the number of stars in the main sequence as the lifetime of this phase is only about 10% of the main sequence lifetime. Therefore, we expect a distribution with a sharp peak towards higher effective temperatures that drops off more gradually towards lower effective temperatures but not reaching zero counts.

Contrary to the expected sources distribution, we see that the number of sources follows a Gaussian-shaped distribution in the lower luminosity ranges. This is not representative of the theory. A plausible explanation for this Gaussian-shaped distribution is the range of visual extinctions in the SMC, as shown in Schootemeijer, Langer, D. Lennon et al. 2021. There is a peak of visual extinction at $A_v = 0.3$ with a Gaussian-like distribution. Therefore, the distribution of effective temperatures of the sources in this data set may be dominated by Gaussian errors. Furthermore, the effect of taking horizontal luminosity bins in the HRD as opposed to diagonal bins normal to the ZAMS means that we are comparing less evolved higher mass stars with more evolved lower mass stars due to the shape of the evolutionary tracks of the stars.

Figure 2.20 shows that the number of sources drops as expected from the initial mass function for the higher luminosity bins. In the upper luminosity bins, low number statistics and lack of structure in the distribution prevent us from drawing accurate Gaussian fits. Therefore stars with $\log(T_{\rm eff}/{\rm K}) \geq 4.25$ in the upper luminosity bins were removed from our data set. This temperature was chosen conservatively to avoid possible main sequence stars, but its position is somewhat arbitrary. The blue-to-red supergiant ratio in the upper luminosity bins will depend on this temperature limit, so no strict conclusions can be drawn from these two uppermost ones. To create the Gaussian fits, the stars with an effective temperature of $\log(T_{\rm eff}/{\rm K}) \geq 4.25$ were used. This temperature was chosen to avoid cutting too deep into the assumed helium-burning region of the distribution. The resulting Gaussian function was integrated within 5σ on either side of the peak of the distribution. The number of stars under this area was subtracted from the total number of blue stars (with $G_{\rm bp} - G_{\rm rp} \leq 0.872$) to determine the number of blue helium-burning stars.

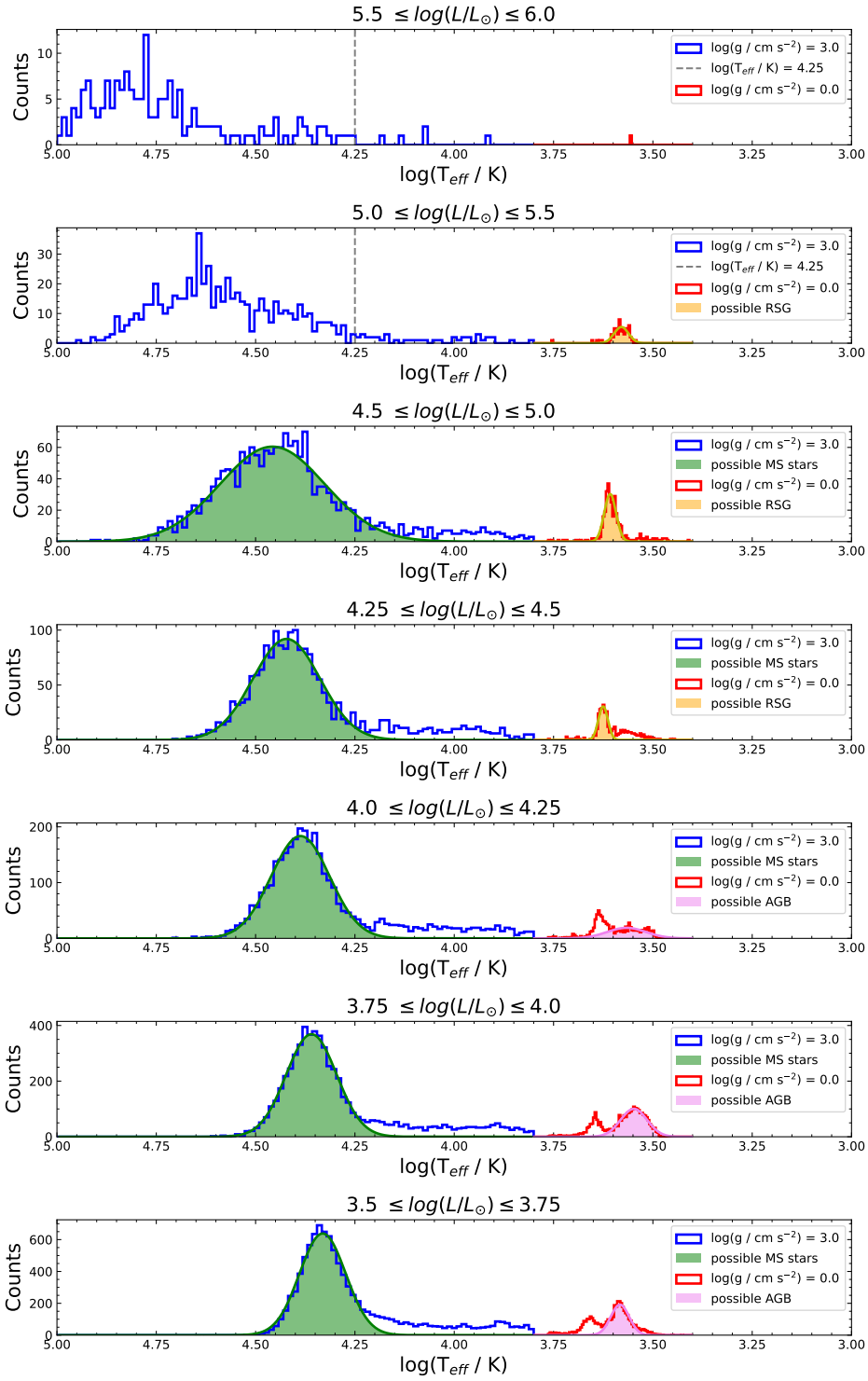


Figure 2.20: Effective temperature histograms for intervals of bolometric luminosity for $A_v = 0.3$. Blue shows the stars with log(g / cms^{-2}) = 3.0, and red shows the stars with log(g / cms^{-2}) = 0.0. Gaussian fits were used to identify the main sequence.

2.3.3 Removing the AGB stars

The AGB stars must be separated from the red helium-burning stars to count the number of red supergiants. As one can see in Figure 2.15, the AGB lies close to the RSG branch in the luminosity range $3.5 \leq \log(L/L_{\odot}) \leq 4.25$. The red supergiant branch was identified similarly to the main sequence. In the luminosity bins from $3.5 \leq \log(L/L_{\odot}) \leq 4.25$, the data set was divided into two at $\log(T_{\rm eff}/{\rm K}) = 3.625$ and Gaussian fits were made to each set. At this temperature, the number of sources is at an apparent minimum. The number of stars under the hotter Gaussian was determined to be the number of possible red supergiants. Above $\log(L/L_{\odot}) = 4.25$, only one Gaussian fit was used due to the low contamination from AGB stars and the function was again integrated over the 5σ range.

2.3.4 Extracting blue and red supergiant stars

After removing OBe, AGB, and main sequence stars, the remaining stars are blue and red supergiant stars. As a validation check of the removal of OBe, AGB, and main sequence stars, we compared the number of blue and red supergiant stars to the number of main sequence stars. The expected value of this should be approximately 10% due to the nuclear timescale of the main sequence compared to the nuclear timescale of the helium burning phase. The ratio of supergiant (SG) stars to main sequence (MS) stars plus supergiant stars was $SG/(MS+SG) = 0.283^{+0.145}_{-0.048}$, where the full details can be seen in Table 2.2. For visual extinction of $A_v = 0.2$ and 0.4, see Figure A.6 and Figure A.7 and Table A.1 and Table A.2.

$log(L/L_{\odot})$	MS	BSG	RSG	AGB	OBe	$\frac{\rm SG}{\rm (MS + SG)}$
6.0 - 5.5	183	6	1	0	0	0.037
5.5 - 5.0	605	41	62	1	0	0.145
5.0 - 4.5	1541	242	335	1	13	0.272
4.5 - 4.25	1589	312	369	1	76	0.3
4.25 - 4.0	2806	618	325	538	320	0.252
4.0 - 3.75	4956	1088	652	2040	934	0.26
3.75 - 3.5	7646	2124	1445	2903	1446	0.318
Total	19326	4431	3189	5484	2789	0.283

Table 2.2: The number of main sequence (MS), blue helium supergiant stars (BSG), red supergiant stars (RSG), asymptotic giant branch (AGB), non-supergiant OB stars (OBe), and all supergiant stars (SG) to main sequence stars plus supergiant stars in varying luminosity intervals for $A_v = 0.3$.

The percentage of helium-burning stars is greater than expected from theory, which further suggests there are missing young blue stars as discussed by Schootemeijer, Langer, D. Lennon et al. 2021. However, as a star evolves, the luminosity increases; thus, when we make horizontal luminosity bins in the HRD, we compare younger high-mass stars against more evolved low-mass stars. Due to the initial mass function, we have more lower-mass stars than high-mass stars, which skews our results towards an inflated number of stars closer to the TAMS than expected. If we include the fact these stars could have their extinction under-corrected, these stars could appear in the helium-burning region of the HRD. We would therefore count them as blue supergiants. This will artificially inflate the blue-to-red supergiant ratio. Furthermore, this could account for obtaining a higher SG/(MS + SG) than what is expected. In

Section 4.3, we show that horizontal HRD cuts do increase the SG/(MS + SG) compared to the expected result.

By further increasing the visual extinction, we see that the number of supergiant stars to the number of main sequence plus supergiant stars decreases. Schootemeijer, Langer, D. Lennon et al. 2021 investigated the visual extinction in the SMC, where sources with values $A_v > 0.5$ were less probable compared to $A_v = 0.2, 0.3$ or 0.4. Therefore, increasing the visual extinction to $A_v > 0.4$ is not recommended to decrease the number of supergiant to main sequence plus supergiant stars to match theory. Consequently, we conclude that even with attempts to mitigate contamination from OBe and main sequence stars, we may still overestimate the actual number of blue supergiants.

In Section 3, theoretical models with varying internal mixing parameters are used to create a synthetic distribution of stars. This is done to test if we can match the luminosity-dependent blue-to-red supergiant ratio for a given set of internal mixing parameters to that of observations. The observational and synthetic blue-to-red supergiant ratio is quantified in Section 4.

CHAPTER 3

Stellar population synthesis

After identifying the blue-to-red supergiant stars from the observational data, we turn to investigate theoretical models. A previously run grid of stellar models with varying mass, semiconvection efficiency α_{sc}, and overshooting efficiency α_{ov} from Schootemeijer, Langer, Grin et al. 2019 was investigated. The masses of stars in this grid ranged from 3 $M_\odot$ to 100 $M_\odot$. These models allow us to study combinations of α_{sc} and α_{ov} to determine the best pair to reproduce the blue-to-red supergiant ratio in the SMC. There were 60 combinations of α_{sc} and α_{sc}, allowing a detailed study of the parameter space. In this study, we will consider the following efficiencies of semiconvection and overshooting: $\alpha_{sc} = [0.1, 1.0, 10.0, 100.0]$ and $\alpha_{ov} = [0.0, 0.12, 0.23, 0.34, 0.45, 0.56]$

Schootemeijer, Langer, Grin et al. 2019 showed that models with very inefficient semiconvection $\alpha_{sc} = 0.1$ and a semiconvection efficiency of $\alpha_{sc} = 1$ produce a significant number of RSGs but very few BSGs. This contradicts the observational findings in Section 4.1. However, we should take care in ruling out inefficient mixing as we extend our luminosity range significantly below the work of Schootemeijer, Langer, Grin et al. 2019 to $\log(L/L_\odot) \geq 3.5$.

3.1 Creating the stellar population

Synthetic populations of stars were created by using the provided grid of models. Each star within this synthetic population will have a random age and mass randomly generated according to the initial mass function.

We assume a constant star formation rate of the SMC, which is a large simplification compared to reality. The star formation history of the SMC is detailed in Harris and Zaritsky 2004. A randomly generated age of the star between 0 and 100 Myr is then selected. 100 Myr was considered as the maximum age to include 5 $M_\odot$ post-main sequence stars while conserving the number of high-mass stars to reduce the problem of low-number statistics in the upper luminosity bins. Naturally, the star's randomly derived age must be less than the star's lifetime to exist; if this is true, then we continue. If the age exceeds the star's lifetime, we discard the star as it would have ended as a supernova or black hole. We now need the mass of the star to determine the effective temperature and luminosity at this age.

To do this, we borrow a Python script[1] that utilises the Kroupa 2001 initial mass function for a cluster of mass M. The number of stars generated within this synthetic star cluster depends on the mass of the

[1] https://github.com/keflavich/imf

constituents. A cluster of mass $M = 4 \times 10^6\,\mathrm{M}_\odot$ produces approximately $\sim 30,000$ stars depending on if their age is less than the lifetime of the star. This is equivalent to the number of main sequence stars plus supergiant stars above $\log(L/L_\odot) \geq 3.5$ from our observation catalogue. As we are only interested in massive stars that evolve into supergiants, we consider only stars with mass $m \geq 5\,\mathrm{M}_\odot$. While we are confident that a $3\,\mathrm{M}_\odot$ star will not evolve into a supergiant star from synthetic stellar populations, some $5\,\mathrm{M}_\odot$ did produce RSGs with $\log(L/L_\odot) \geq 3.5$. We do not consider lower masses than $5\,\mathrm{M}_\odot$ as we note that these stars will not produce supergiant stars with luminosities $\log(L/L_\odot) \geq 3.5$. Therefore stars with masses below $5\,\mathrm{M}_\odot$ are not present in the synthetic populations.

To determine a star's effective temperature and bolometric luminosity, the model with the closest modelled mass to the randomly determined mass is selected to prevent the need to run a multitude of models of different masses. The history.data file from the MESA output of this model is read, and the effective temperature and bolometric luminosity as functions of age are stored in two arrays. We have a problem finding the star's effective temperature and bolometric luminosity at the randomly determined age. It is possible to produce more data in smaller time steps. However, this is computationally expensive. We are therefore left to use the effective temperature and bolometric luminosity of the closest modelled age. The star's effective temperature and bolometric luminosity are then saved, and the process is repeated for each star. If we plot the results directly, we will produce a HRD where stars follow the evolutionary track lines of specific masses perfectly. An example of a synthetic stellar population with a set of mixing parameter efficiencies is shown in Figure 3.1. However, as we see, this is not the case observationally. Therefore, we need to remove resolution effects to make meaningful comparisons to the observations.

3.2 Applying errors to synthetic population

We need to remove the resolution effects and account for the fact that we have to choose the closest modelled mass and age to determine the effective temperature and bolometric luminosity. We look at the errors we calculated in the HRD from Section 2.3.1 to do this. Namely, we look at Figures 2.18 and 2.19, where we show the error in effective temperature and bolometric luminosity as a function of effective temperature. We see that as the effective temperature increases, the errors increase. The stars' effective temperature and bolometric luminosities have asymmetrical errors due to non-monotonically decreasing colour-effective temperature relation. In these figures, we plot the minimum of the asymmetrical errors. If we use the mean or the maximum errors, the main sequence of the synthetic population becomes too broad, and we cannot draw any meaningful conclusions.

Polynomial functions were created to fit the errors in effective temperature and bolometric luminosities. The effective temperature of a synthetically generated star was used to determine the error in effective temperature and bolometric luminosity. Then Python's `np.random.normal()` function was used to choose a value of the effective temperature and bolometric luminosity according to a Gaussian distribution centred around the effective temperature and bolometric luminosity determined by the age and mass of the star. The standard deviation of the Gaussian was the error determined by the polynomial fits. Finally, solely for the bolometric luminosity, an additional uncertainty of $\sigma(\log(L/L_\odot)) = 0.1$ was added to stars with $\log(T_{\mathrm{eff}/K}) \leq 3.763$ as the minimum error from our bolometric luminosity determination was not enough to remove resolution effects for the RSGs. Thus, we now have an effective temperature and bolometric luminosity of the synthetic stars that no longer precisely follow the evolutionary tracks but instead are scattered around their original position.

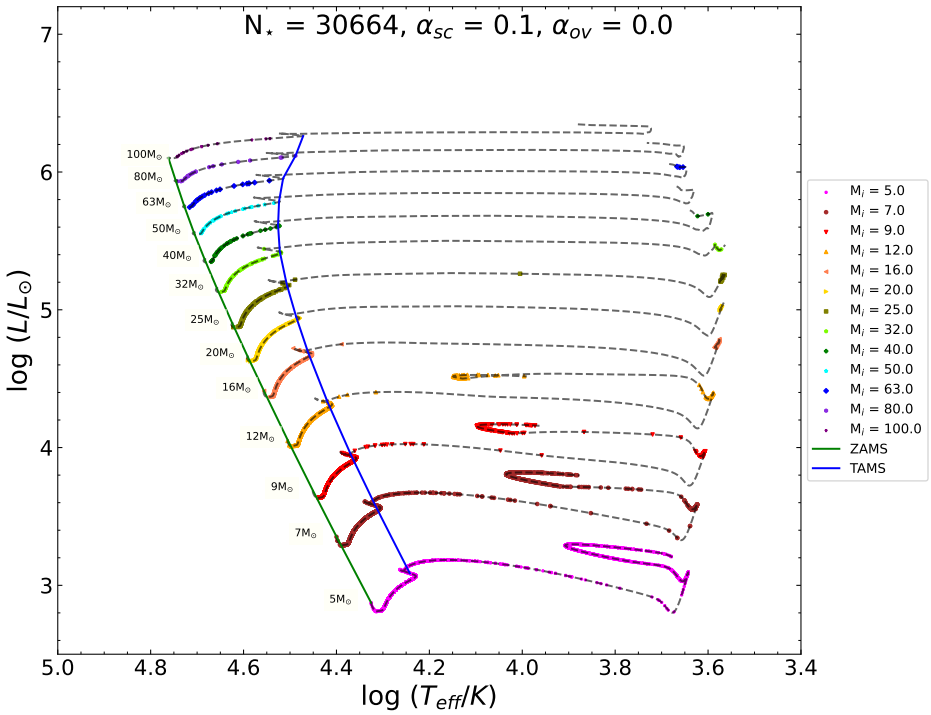


Figure 3.1: Synthetic HRD for $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$. The colour of the data points is for the corresponding initial mass in the legend.

The method above is repeated for multiple combinations of $\alpha_{\rm sc}$ and $\alpha_{\rm ov}$. An example of a synthetic HRD of approximately $\sim 30,000$ stars with $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$ is shown in Figure 3.2.

This figure shows that using the minimum errors in effective temperatures and bolometric luminosities causes some problematic features in the synthetic HRD. Most notably, we see many stars hotter than the ZAMS, which was not seen in the observational HRD (Figure 2.15). We note that in the observational data, a magnitude limit of $G \leq 18$ mag resulted in a sharp feature running diagonally in the HRD where we cannot see if there are stars hotter than the TAMS for $\log(L/L_\odot) \lesssim 3.5$. However, above this luminosity, there are still significantly more stars in the synthetic population beyond the ZAMS. Another noticeable result of this scatter by errors is that there are more stars with luminosities $\log(L/L_\odot) \geq 6.5$ than observed. Therefore, this suggests our method of determining the effective temperature and bolometric luminosity using bolometric correction tables introduces too high uncertainties. As the largest uncertainty in this work is visual extinction, the assumed uncertainty value of visual extinction could be smaller than 0.1.

In Figure 3.3, we can identify where the blue and red supergiants end up in the HRD. The supergiant stars were classified as stars whose core hydrogen mass fraction fell below 1×10^{-2}. Blue and red supergiant stars were again separated by their effective temperature, where blue supergiants have $\log(T_{\rm eff}/K) > 3.763$ and red supergiants have $\log(T_{\rm eff}/K) \leq 3.763$. Figure 3.3 we see how the blue supergiants are mixed well with the main sequence, shown in grey data points. Therefore, we show how difficult it is to accurately separate the blue helium-burning stars from the main sequence when A_v is not

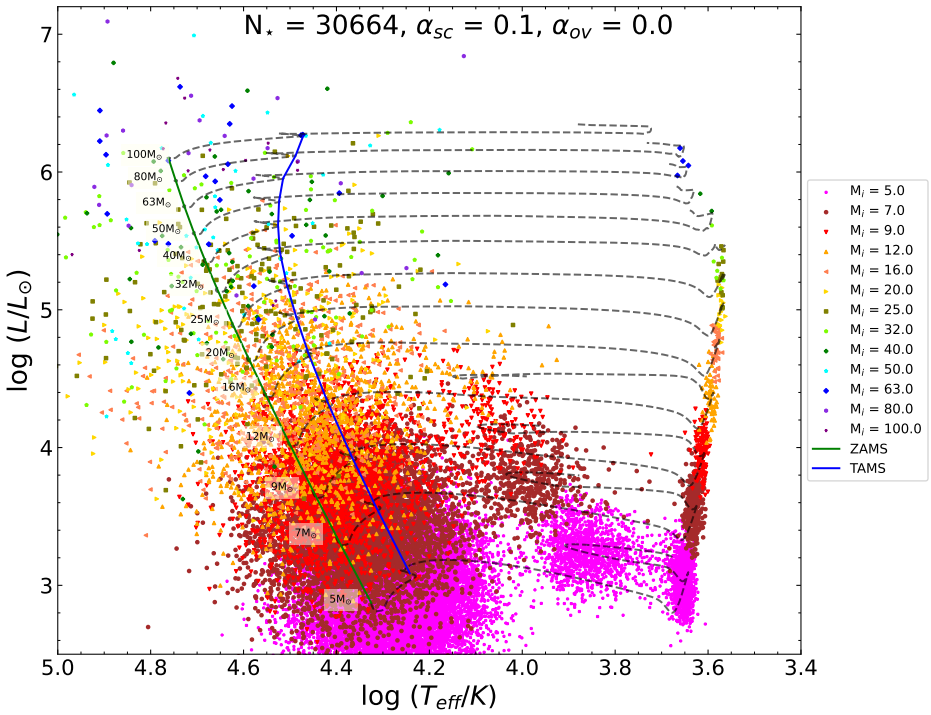


Figure 3.2: The same as Figure 3.1, but the points were scattered according to the errors in effective temperature and bolometric luminosity determined from observations.

explicitly known. Section 4.2 determines the blue-to-red supergiant ratio.

We again create effective temperature histograms of the synthetic HRD in the same luminosity bins as the observational data. Figure 3.4 shows an example for a set of internal mixing parameter efficiencies $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$. Notable is the Gaussian shape of the distribution of the hottest stars. This is likely due to using the `np.random.normal()` function, which applies a Gaussian uncertainty to each point. However, the advantage of this is that we can again use the same approach as in the observations to separate the main sequence stars and helium-burning stars. The resulting B/Rs of various synthetic stellar populations are shown in the results section.

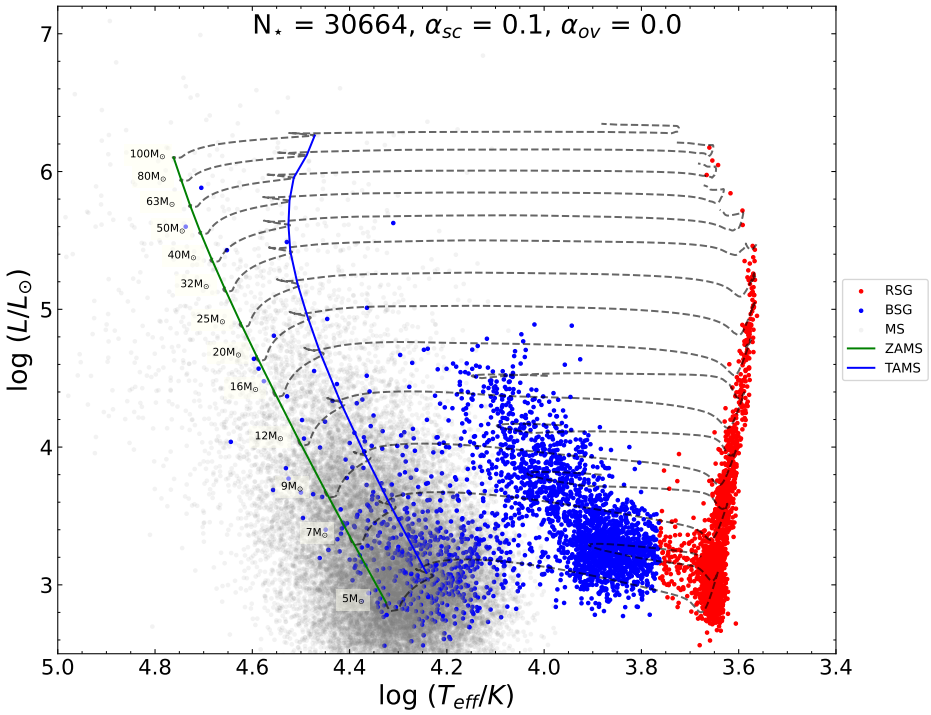


Figure 3.3: Synthetic HRD for $\alpha_{\rm sc}$ = 0.1 and $\alpha_{\rm ov}$ = 0.0. Blue points are considered blue supergiants, and red points are considered red supergiants.

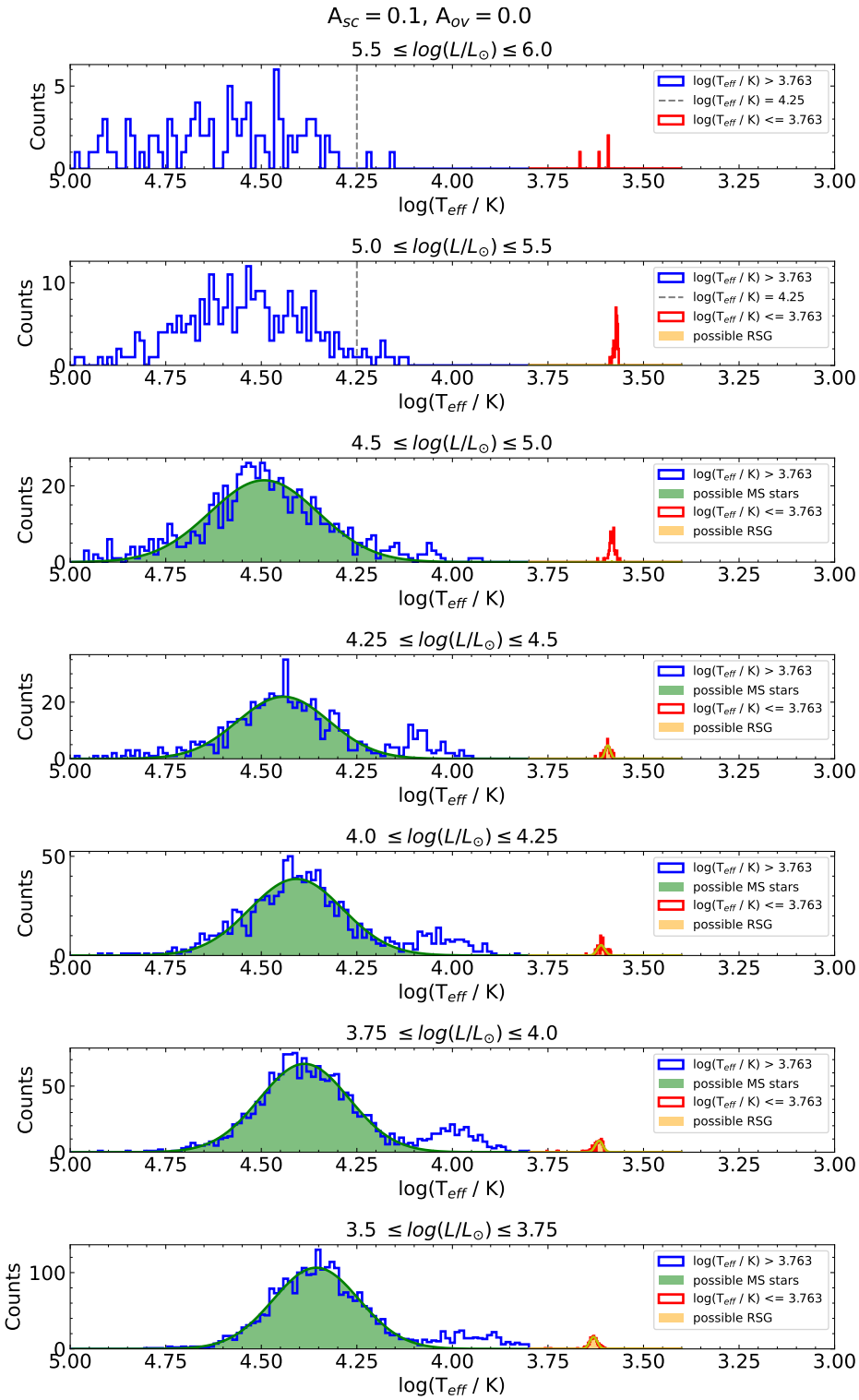


Figure 3.4: Effective temperature histogram for different bolometric luminosity intervals for the synthetic stellar population with internal mixing parameters: $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$.

CHAPTER 4

Results

4.1 Observed blue-to-red supergiant ratio

Once the observed blue and red helium-burning stars were identified through the main sequence star and AGB star separation method detailed in Sections 2.3.2 and 2.3.3, the next step was to divide the number of blue helium-burning stars by the red helium-burning stars for each luminosity bin. This provides us with a luminosity-dependent blue-to-red helium-burning star ratio. If we assume all of these stars are, in fact, supergiants due to the removal of OBe stars and AGB stars, then we can call this the blue-to-red supergiant ratio (B/R).

The B/R can be calculated for A_v = 0.3, 0.2, and 0.4 from Tables 2.2, A.1, and A.2 respectively. This is shown in Table 4.1. The total B/R of the SMC can also be easily calculated by totalling the blue and red helium-burning stars in all luminosity bins above $\log(L/L_\odot) \geq 3.5$. The total value for the SMC is B/R = $1.389^{+0.207}_{-0.452}$. Figure 4.1 graphically shows these results for the different luminosity bins.

$log(L/L_\odot)$	B/R		
	$A_v = 0.2$	$A_v = 0.3$	$A_v = 0.4$
6.0 - 5.5	7.0	6.0	1.143
5.5 - 5.0	0.818	0.661	0.56
5.0 - 4.5	0.901	0.722	0.697
4.5 - 4.25	1.031	0.846	0.741
4.25 - 4.0	2.616	1.902	1.565
4.0 - 3.75	2.32	1.669	1.354
3.75 - 3.5	1.97	1.47	1.256
Total	1.596	1.389	0.937

Table 4.1: Blue-to-red helium burning star ratio B/R for A_v = 0.2, 0.3, and 0.4.

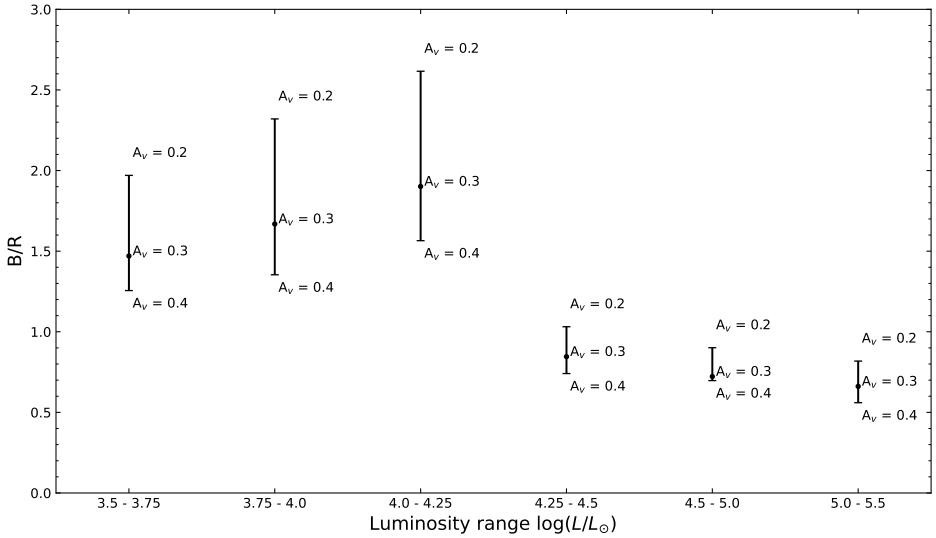


Figure 4.1: Blue-to-red supergiant ratio for the SMC, given a visual extinction $A_v = 0.3 \pm 0.1$. The total B/R = $1.389^{+0.207}_{-0.452}$. The highest luminosity bin is left out here due to the wide scatter of data points.

4.2 Synthetic blue-to-red supergiant ratio

The blue-to-red supergiant ratio was then determined from the synthetic stellar populations for different mixing parameter combinations, as described in Section 3 and compared against the observational result from Section 4.1.

We implemented the same procedure as with the observational data to do this. Effective temperature histograms were created for the same luminosity bins as before. The main sequence and AGB stars were removed using the same Gaussian fit removal method. The blue and red supergiants were then identified and counted. The luminosity-dependent blue-to-red supergiant ratios were then determined and compared with the observational findings. This will be known as method A.

An advantage of synthetic models is that we can examine the output history.data file from MESA and identify which stars are burning helium in the core and which are still burning hydrogen. More specifically, we call supergiants, stars with core hydrogen fraction $< 1 \times 10^{-2}$. We call this method B. We can accurately produce an SG/(MS+SG) ratio and compare that to the observational ratio. As it is not clear precisely at which effective temperature a supergiant is considered red or blue, we use the same transition colour as we did previously to separate blue and red stars. Unless otherwise stated, we will use method A in all future comparisons. This allowed us to make a fair comparison to observations.

The following tables show the number of supergiants and main sequence stars obtained from a synthetic population with $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$. Table 4.2 shows the number of main sequence and supergiant stars alongside the respective ratios of these stars using method B. As we can see from Table 4.2, the ratio of helium-burning supergiants compared to all stars is ~ 0.18, which is closer to the expected value according to the ratio of lifetimes of helium-burning stars to the main sequence than for the observed ratio which was ~ 0.28. As mentioned previously, in method A, we compare low-mass stars with high-mass stars when computing the B/R. This also results in a higher-than-expected SG/(MS+SG).

Table 4.3 represents the results of testing method A in regards to identifying main sequence stars and helium-burning stars. Figure 3.4 shows graphically the distribution of sources in the effective temperature bins for different bolometric luminosity intervals. The number of blue supergiants and red supergiants from method A match well with the number of blue and red supergiants from method B. Furthermore, the ratio of helium-burning stars to all stars in method A also matches the ratio in method B. Therefore, we are confident that we can separate the blue-supergiant stars from the main sequence. As the models were computed until central helium exhaustion, no AGB branch is visible in the synthetic data. We could not comment on how well method A separates RSG from AGB stars.

4.3 Comparison between observed and synthetic B/R

Now that we have luminosity-dependent blue-to-red supergiant ratios for different mixing parameter combinations, we can determine the set of parameters that best match the observed trend. Figure 4.2 shows the observed blue-to-red supergiant ratio in the SMC in black with a green shaded area covering the uncertainty region due to the uncertainty in the visual extinction. The blue line shows the best matching mixing parameters using method A are $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$. Furthermore, 4.3 shows for $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$, how similar the B/R from method B (red line) compares with method A (blue line). Other combinations can be found in the subsequent Figures 4.4 - 4.7, where the blue line always represents method A.

$log(L/L_{\odot})$	MS	BSG	RSG	$\frac{\text{SG}}{\text{(MS + SG)}}$	B/R
6.0 - 5.5	116	2	6	0.065	0.333
5.5 - 5.0	256	8	28	0.123	0.286
5.0 - 4.5	602	39	62	0.144	0.629
4.5 - 4.25	592	88	48	0.187	1.833
4.25 - 4.0	950	153	59	0.182	2.593
4.0 - 3.75	1513	299	124	0.218	2.411
3.75 - 3.5	2385	350	202	0.188	1.733
Total	6414	939	529	0.186	1.775

Table 4.2: The number of main sequence (MS), blue helium supergiant stars (BSG), red supergiant stars (RSG), the ratio of all supergiant stars (SG) to main sequence stars plus supergiant stars, and the blue-to-red supergiant ratio (B/R) in varying luminosity intervals for method B. $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$.

$log(L/L_{\odot})$	MS	BSG	RSG	$\frac{\text{SG}}{\text{(MS + SG)}}$	B/R
6.0 - 5.5	116	2	6	0.065	0.333
5.5 - 5.0	253	11	27	0.131	0.407
5.0 - 4.5	609	32	61	0.132	0.525
4.5 - 4.25	610	70	47	0.161	1.489
4.25 - 4.0	975	128	58	0.16	2.207
4.0 - 3.75	1518	294	123	0.216	2.39
3.75 - 3.5	2380	355	201	0.189	1.766
Total	6461	892	523	0.18	1.706

Table 4.3: The same table as Table 4.2 but using method A. $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$.

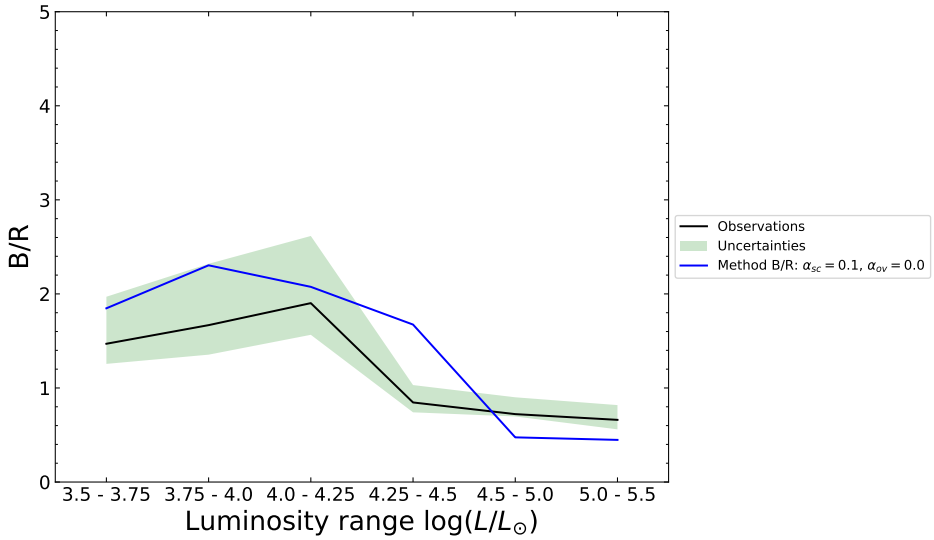


Figure 4.2: Blue-to-red supergiant ratio for the SMC and the synthetic stellar population for $\alpha_{\rm sc} = 0.1$ and $\alpha_{\rm ov} = 0.0$. The highest luminosity bin is left out here due to the wide scatter of data points.

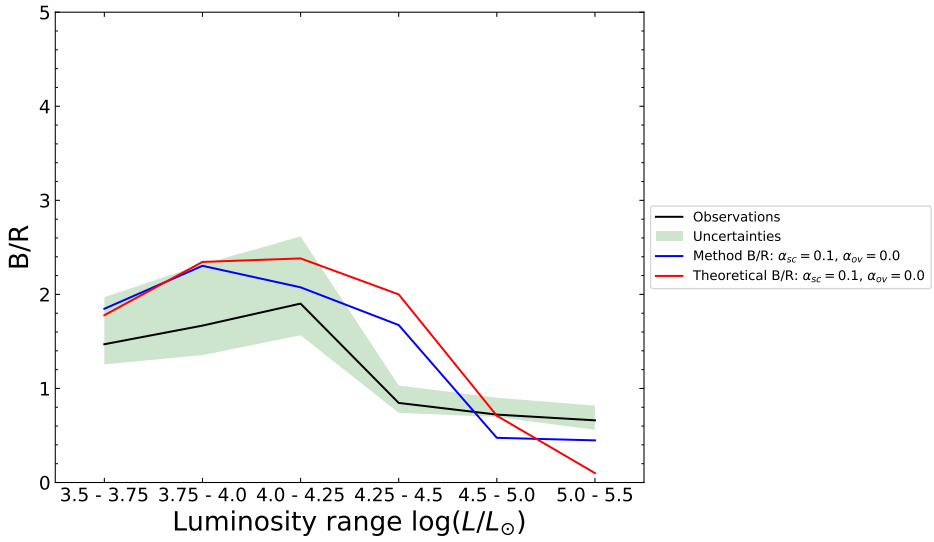
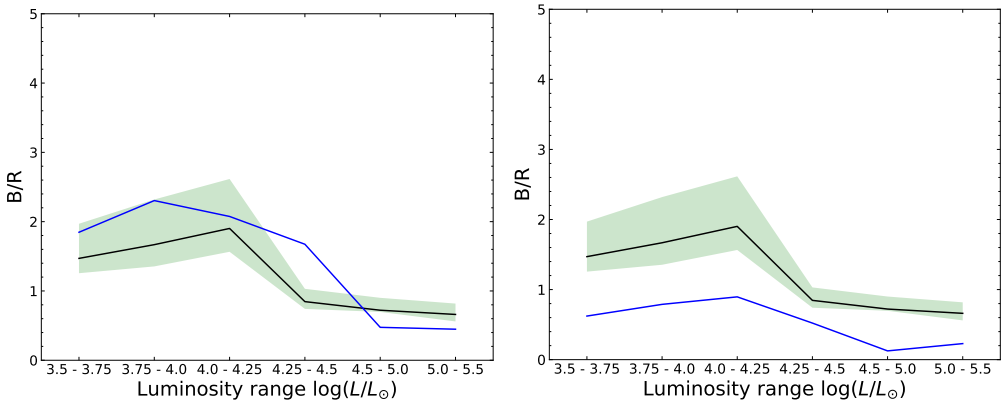
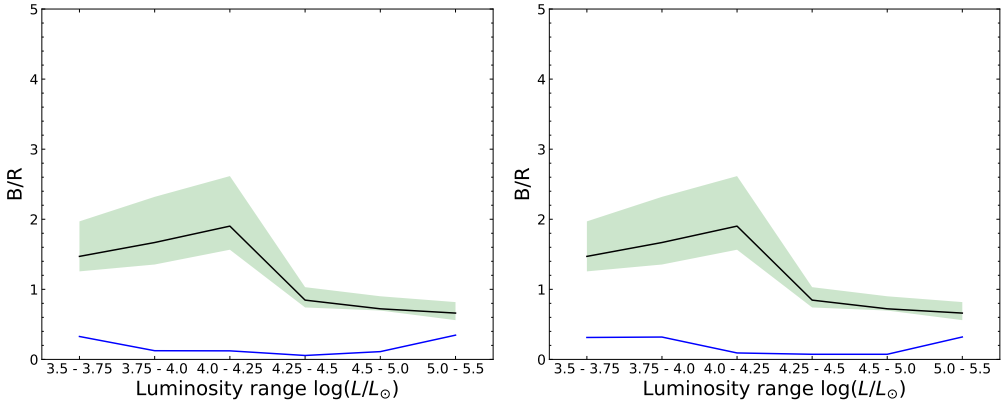


Figure 4.3: Same as Figure 4.2 but with an additional red line showing the B/R obtained using if we consider helium-burning stars as stars with core hydrogen mass fraction below 1×10^{-2}.



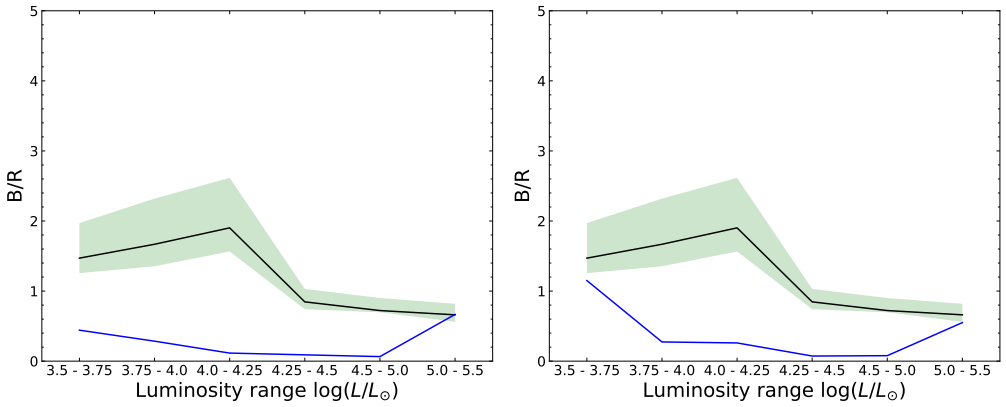
(a) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.0$

(b) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.12$



(c) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.23$

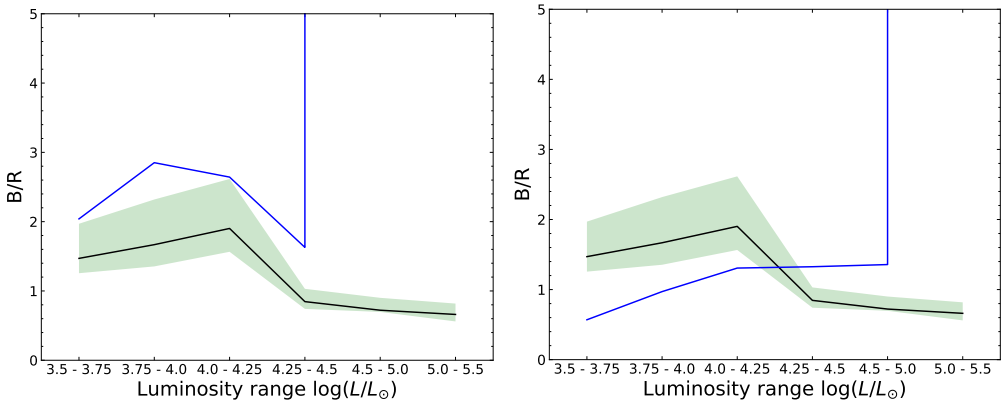
(d) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.34$



(e) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.45$

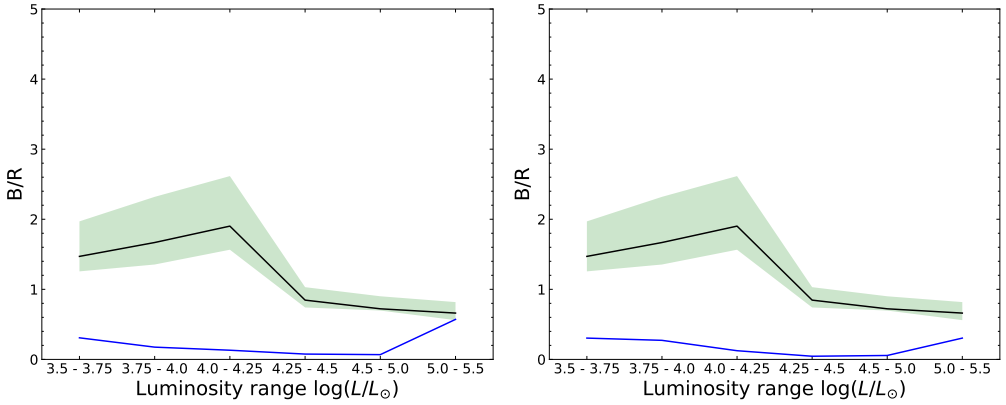
(f) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.56$

Figure 4.4: B/R of the synthetic stellar population for varying $\alpha_{sc} = 0.1$ and varying overshooting efficiencies shown in blue and the observed B/R is shown in black. The uncertainties of the B/R due to visual extinction is shown in green.



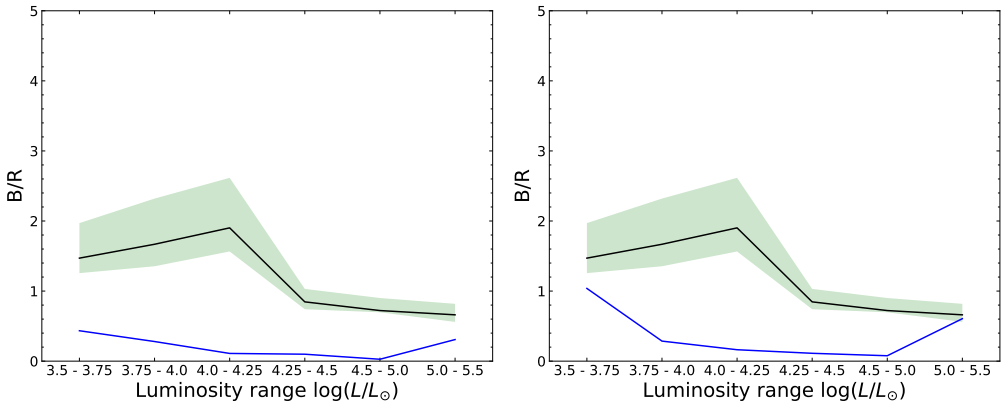
(a) $\alpha_{sc} = 1.0$, $\alpha_{ov} = 0.0$

(b) $\alpha_{sc} = 1.0$, $\alpha_{ov} = 0.12$



(c) $\alpha_{sc} = 1.0$, $\alpha_{ov} = 0.23$

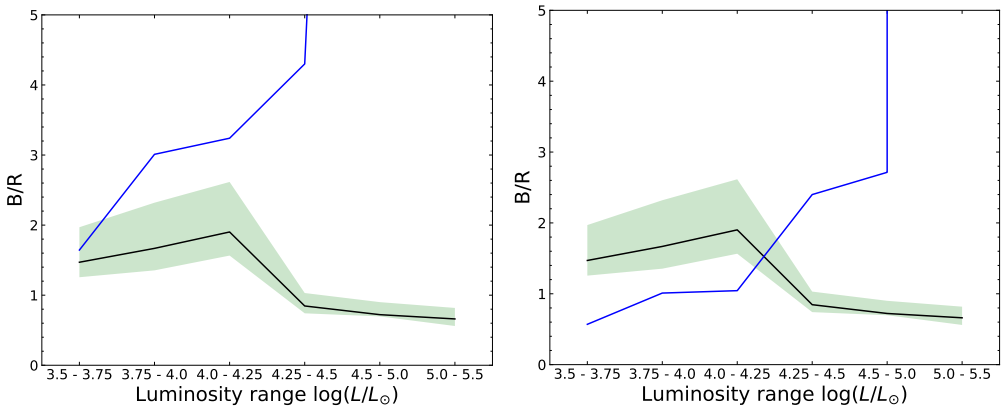
(d) $\alpha_{sc} = 1.0$, $\alpha_{ov} = 0.34$



(e) $\alpha_{sc} = 1.0$, $\alpha_{ov} = 0.45$

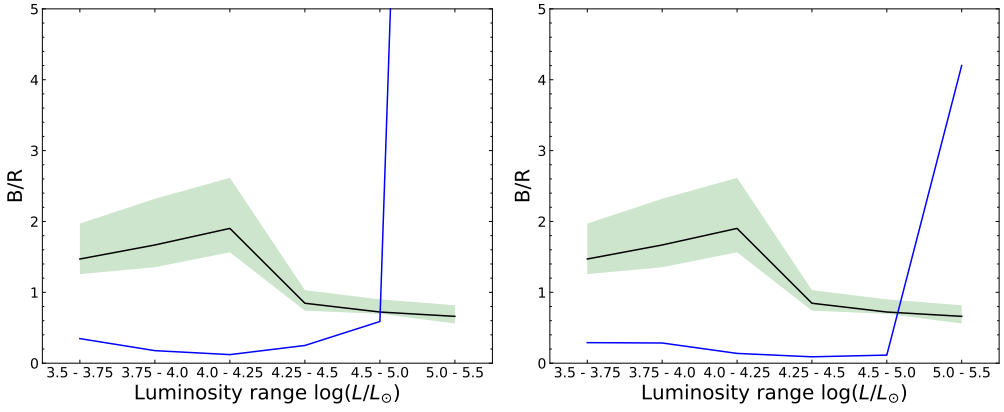
(f) $\alpha_{sc} = 1.0$, $\alpha_{ov} = 0.56$

Figure 4.5: B/R of the synthetic stellar population for varying $\alpha_{sc} = 1.0$ and varying overshooting efficiencies shown in blue and the observed B/R is shown in black. The uncertainties of the B/R due to visual extinction is shown in green.



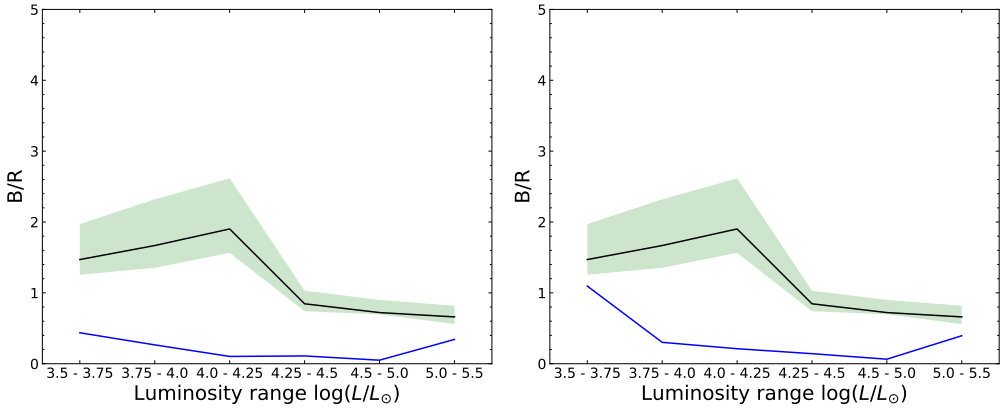
(a) $\alpha_{sc} = 10.0$, $\alpha_{ov} = 0.0$

(b) $\alpha_{sc} = 10.0$, $\alpha_{ov} = 0.12$



(c) $\alpha_{sc} = 10.0$, $\alpha_{ov} = 0.23$

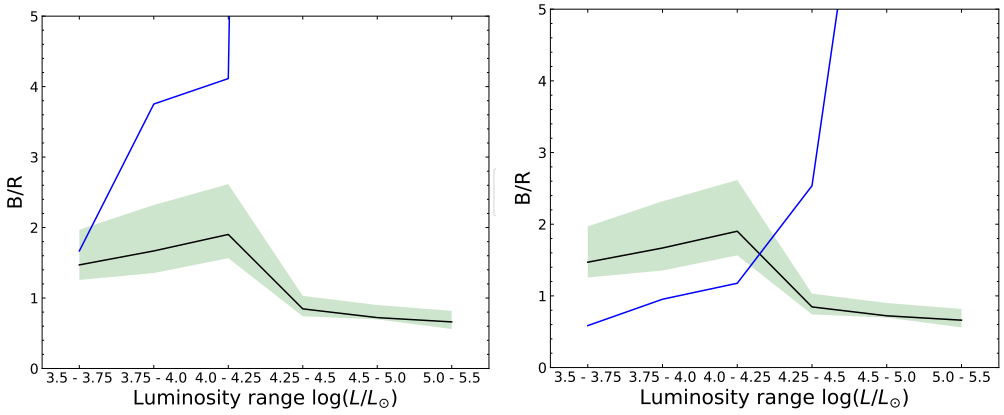
(d) $\alpha_{sc} = 10.0$, $\alpha_{ov} = 0.34$



(e) $\alpha_{sc} = 10.0$, $\alpha_{ov} = 0.45$

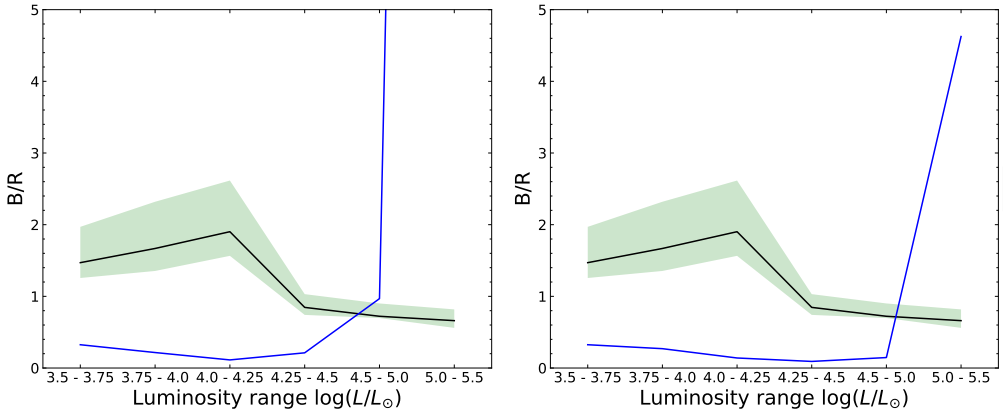
(f) $\alpha_{sc} = 10.0$, $\alpha_{ov} = 0.56$

Figure 4.6: B/R of the synthetic stellar population for varying $\alpha_{sc} = 10.0$ and varying overshooting efficiencies shown in blue and the observed B/R is shown in black. The uncertainties of the B/R due to visual extinction is shown in green.



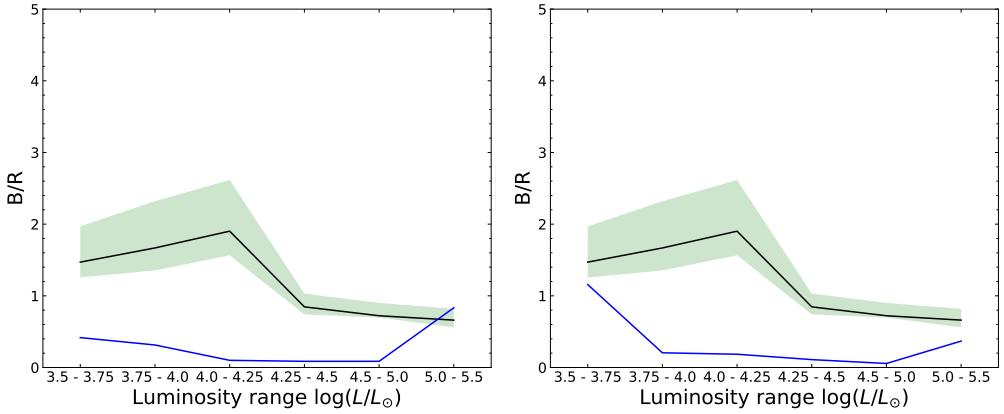
(a) $\alpha_{sc} = 100.0$, $\alpha_{ov} = 0.0$

(b) $\alpha_{sc} = 100.0$, $\alpha_{ov} = 0.12$



(c) $\alpha_{sc} = 100.0$, $\alpha_{ov} = 0.23$

(d) $\alpha_{sc} = 100.0$, $\alpha_{ov} = 0.34$



(e) $\alpha_{sc} = 100.0$, $\alpha_{ov} = 0.45$

(f) $\alpha_{sc} = 100.0$, $\alpha_{ov} = 0.56$

Figure 4.7: B/R of the synthetic stellar population for varying $\alpha_{sc} = 100.0$ and varying overshooting efficiencies shown in blue and the observed B/R is shown in black. The uncertainties of the B/R due to visual extinction is shown in green.

CHAPTER 5

Discussion

5.1 Method validation with the VSS survey

To validate this determination of effective temperatures and bolometric luminosities using bolometric correction tables, which we will call method C, we compared sources' effective temperatures and bolometric luminosities with spectroscopically determined values from other surveys. The surveys from Dufton et al. 2019, Ramachandran et al. 2019, Bouret et al. 2013, Hunter et al. 2008, Mokiem, de Koter, Evans et al. 2006, Trundle, D. J. Lennon et al. 2004 (supplemented with their later work, Trundle and D. J. Lennon 2005), and Venn 1999 were compiled together to form a various spectroscopic survey (VSS) as in Schootemeijer, Langer, D. Lennon et al. 2021. The results of which are provided in Figure 5.1 and Figure 5.2, where the dashed line represents an exact agreement between this work and the spectroscopically determined effective temperature or bolometric luminosity from the VSS.

All the following data was accessed using the Strasbourg Astronomical Data Center[1] (CDS) and cross-matched with sources in the GAIA DR3 catalogue (Gaia Collaboration, Vallenari et al. 2023) using the CDS X-Match Service [2]. The cross-match position had a radius set to $1''$; the majority of the sources were within $0.5''$, so we are assured that false matches were not common. Further care was taken here; the output CSV files were scanned for duplicate matches, the match with the lowest angular distance was kept, and the other was discarded. Additionally, OBe stars were removed from the cross-matched catalogue if the spectral type was given.

A short description of the data from the VSS studies we compared this work with is given below, but more details can be found in their papers.

Dufton et al. 2019 investigated 363 O- and B-type stars in the young open cluster NGC 346 where the reddening E(B-V) = 0.09 is given. We will use the visual extinction law $A_v = 2.72 \cdot$E(B-V) for the SMC from Bouchet et al. 1985 for all further calculations unless otherwise specified. The visual extinction is thus $A_v \sim 0.25$, which is lower than the assumption of $A_v = 0.3$ used in this work but still within our error of $\sigma(A_v) = 0.1$. Only 53 sources were cross-matched successfully and were used for the comparison. The spectroscopic effective temperatures were determined using TLUSTY stellar atmosphere models. The effective temperature we calculate is higher than the spectroscopically derived value. This can be seen in Figure 5.1. However, some notable exceptions do not fit this trend at higher

[1] https://cds.unistra.fr/

[2] http://cdsxmatch.u-strasbg.fr/

effective temperatures. The luminosity was inferred from the effective temperature, rotational velocity, and photometric magnitude. The luminosity has a scatter around the dashed line in Figure 5.2, where most points are found below the dashed line. Therefore, we more often overestimate the value of the bolometric luminosity due to our higher value of visual extinction.

Ramachandran et al. 2019 focused on O- and B-type stars in the wing of the SMC where the reddening is lower than the bulk of the SMC at E(B-V) = 0.08 corresponding to $A_v \sim 0.22$, again lower than our assumption. Out of the 320 stars they investigated, 44 were found to be OBe stars. Four stars had bluer colours than the colour-effective temperature relation could fit, so it was impossible to include them in this comparison. We were left with 272 stars used in our comparison. Their effective temperatures were determined using PoWR model atmospheres. We have overestimated the visual extinction of these objects, and therefore we determine higher effective temperatures. The bolometric luminosity is again overestimated, such as the case with Dufton et al. 2019, as the visual extinction used in this work is higher than the estimated value in Ramachandran et al. 2019.

Bouret et al. 2013 investigated 23 O-type dwarfs in the SMC in optical and UV filters. All 23 cross-matched successfully with GAIA DR3, and no OBe stars were present in the final data set. The reddening varied from $0.04 \leq$ E(B-V) ≤ 0.13 or $0.1 \leq A_v \leq 0.35$, which results in this work mostly overestimating the visual extinction but perhaps in some cases underestimating the visual extinction. This is represented in Figure 5.1, showing the orange data points scattered around the dashed line. The effective temperatures were determined using the CMFGEN model atmosphere code.

Hunter et al. 2008 had a catalogue of ~400 early-type objects from the SMC, LMC, NGC 346, and NGC 330. There were 125 matches with GAIA DR3, and one was flagged as an OBe star. The reddening varied depending on the different regions sampled, E(B-V) = 0.09 for NGC 346 and E(B-V) = 0.06 for NGC 330 or $A_v \sim 0.25$ and $A_v \sim 0.16$, respectively. Using method C, the effective temperature is overestimated for lower effective temperatures, but at higher effective temperatures, the effective temperatures are underestimated as compared to Hunter et al. 2008. The overestimated effective temperatures come from the overestimated visual extinction. Regarding bolometric luminosity, the green data points are much more confined to a line just more luminous than the dashed line, confirming that the extinction was overestimated.

Mokiem, de Koter, Evans et al. 2006 studied 31 O- and early B-type stars in the SMC, most of which resided in NGC 346. It is worth noting that Hunter et al. 2008 also contains sources from Mokiem, de Koter, Evans et al. 2006, Trundle, D. J. Lennon et al. 2004 and Trundle and D. J. Lennon 2005, but the catalogue was rechecked for duplicates. Eight sources found matches in CDS X-match, none of which were OBe stars. The visual extinctions, which range from $0.03 \leq A_v \leq 0.74$, were directly given in Table 1 of Mokiem, de Koter, Evans et al. 2006. As with Hunter et al. 2008, the surface gravity was greater than $\log(g \text{ / cms}^{-2}) \geq 3.0$ and, therefore, can account for the underestimated effective temperatures in this work. The bolometric luminosities derived in this work are slightly underestimated compared to the spectroscopically derived bolometric luminosities.

Trundle, D. J. Lennon et al. 2004 and Trundle and D. J. Lennon 2005 measured the effective temperatures of 17 B-type supergiants and 1 giant star in the SMC using the silicon and carbon resonance lines. Eighteen sources were cross-matched successfully. The reddening was given as E(B-V) ~ 0.07, which is a visual extinction of $A_v \sim 0.19$. The effective temperatures of the stars are in good agreement with the dashed line. However, the underestimated effective temperatures do not agree with the overestimated visual extinction. The surface gravities also range from $2.0 \leq \log(g \text{ /cms}^{-2}) \leq 3.6$ with an average value of $\log(g \text{ /cms}^{-2}) = 2.5$, which also does not explain the discrepancy between this work and the work of Trundle, D. J. Lennon et al. 2004 and Trundle and D. J. Lennon 2005. The

bolometric luminosities agree with the dashed line, although slightly underestimated.

Venn 1999 investigated 10 A-type supergiants in the Magellanic Bridge, all cross-matched successfully. There was no comment on the reddening, but the surface gravities of these stars are approximately two orders of magnitude less than for an O- or a B-type star. Thus we have grossly overestimated the surface gravity. However, as seen in Figure 2.9, the colour-effective temperature relation for $\log(g/\mathrm{cms}^{-2}) = 0.0$ and 3.0, the same colour gives a very similar effective temperature. This is the plausible reason why the purple data points still fit well to this dashed line. There was no bolometric luminosity data on these sources.

A summary of the VSS visual extinctions is given in Table 5.1.

Author	A_v
Dufton et al. 2019	~ 0.25
Ramachandran et al. 2019	~ 0.22
Bouret et al. 2013	$0.1 \leq A_v \leq 0.35$
Hunter et al. 2008	~ 0.25 (NGC 346) and ~ 0.16 (NGC 330)
Mokiem, de Koter, Evans et al. 2006	$0.03 \leq A_v \leq 0.74$
Trundle, D. J. Lennon et al. 2004 and Trundle and D. J. Lennon 2005	~ 0.19
Venn 1999	undefined

Table 5.1: Visual extinction values within the VSS.

We conclude that the majority of effective temperatures and bolometric luminosities determined using method C align with the spectroscopically determined values within the given error bars. The average value and standard deviation of the difference between the effective temperatures determined from bolometric correction tables versus the VSS, $\Delta\log T = \log(T_{\mathrm{eff}}/\mathrm{K}) - \log(T_{\mathrm{eff}}^{\mathrm{VSS}}/\mathrm{K})$, are given as $\Delta\bar{\log}T = 0.048$ and $\sigma(\Delta\log T) = 0.130$. The average value and standard deviation of the difference between the bolometric luminosity determined from bolometric correction tables versus the VSS, $\Delta\log L = \log(L/\mathrm{L}_\odot) - \log(L^{\mathrm{VSS}}/\mathrm{L}_\odot)$, are given as $\Delta\bar{\log}L = 0.362$ and $\sigma(\Delta\log L) = 0.738$.

We can exclude the Ramachandran et al. 2019 survey as they focus on stars in the wing of the SMC, which does not contain as many stars as the rest of the SMC and has a lower visual extinction of $A_v = 0.2$. The average values and standard deviations are then, $\Delta\bar{\log}T = -0.0011$, $\sigma(\Delta\log T) = 0.131$, $\Delta\bar{\log}L = 0.359$, and $\sigma(\Delta\log L) = 1.024$. We see that the average values decrease, indicating a closer correlation between the derived value using method C and the VSS values. However, the standard deviation increases due to reducing the number of data points.

Any significant deviation in effective temperature and bolometric luminosity arises from under- or overestimating the visual extinction outwith our given errors. We recognise a diagonal shift in the bolometric luminosities for $\log(L/L_\odot) \leq 5.0$, where we consistently determine a higher bolometric luminosity than the spectroscopically derived value. If individual distances to sources were used rather than a single value, the bolometric luminosities determined through method C may be closer to the VSS values. A value of $A_v = 0.2$ could have resulted in more satisfactory results when compared to Ramachandran et al. 2019 as a systematic shift of method C's effective temperature and bolometric luminosities towards lower values appears to fit the dashed line better. $A_v = 0.2$ is still within our range of uncertainty. Therefore we are confident that we can still determine the bolometric luminosity of a source

within a reasonable uncertainty. Furthermore, the other VSS studies do show a better agreement with $A_v = 0.3$. With this discussion on the validity of our effective temperature and bolometric luminosity determinations, we are confident that we produce meaningful results when looking at the B/R.

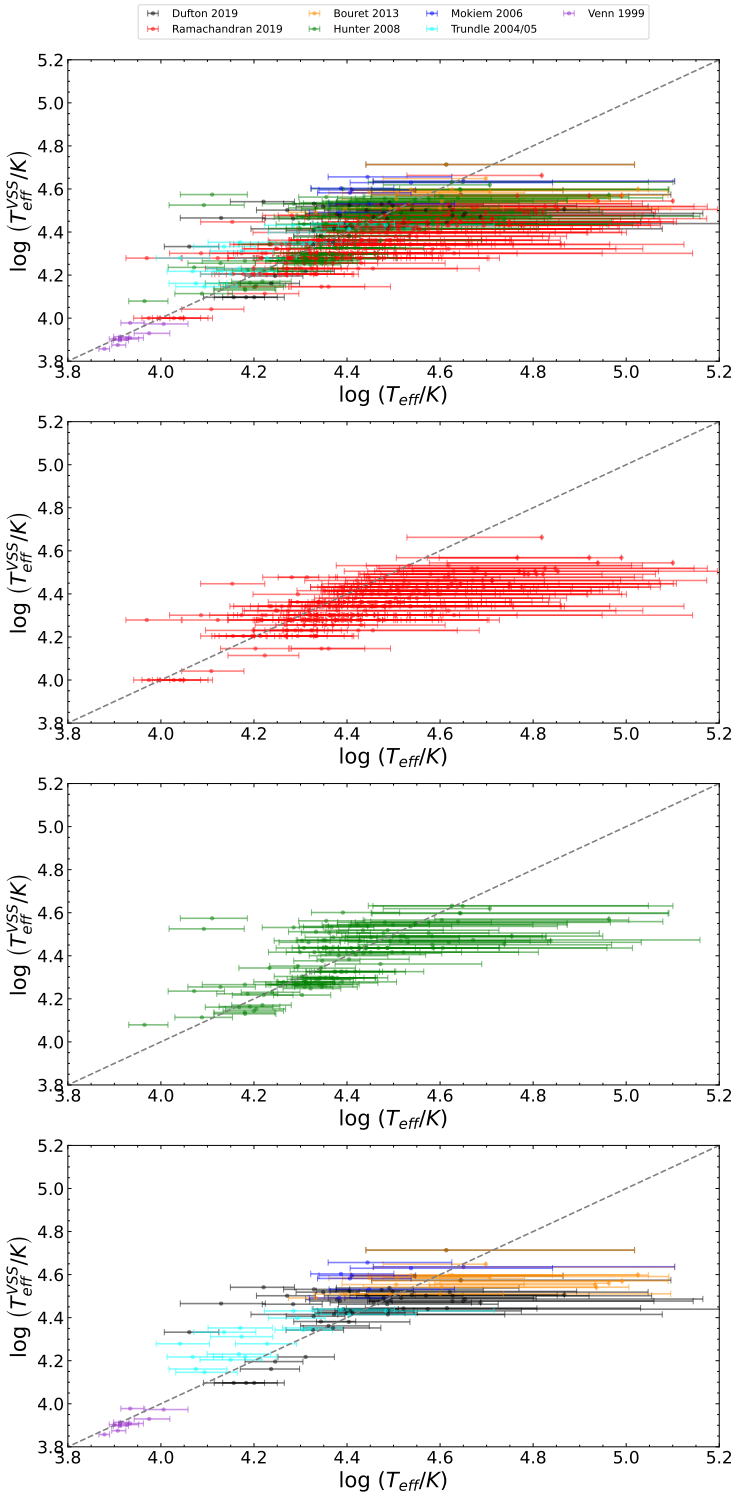


Figure 5.1: Effective temperatures of sources using bolometric correction tables compared to spectroscopically determined values from the VSS. The upper plot shows all VSS data. The subsequent plots underneath show only some VSS for clarity.

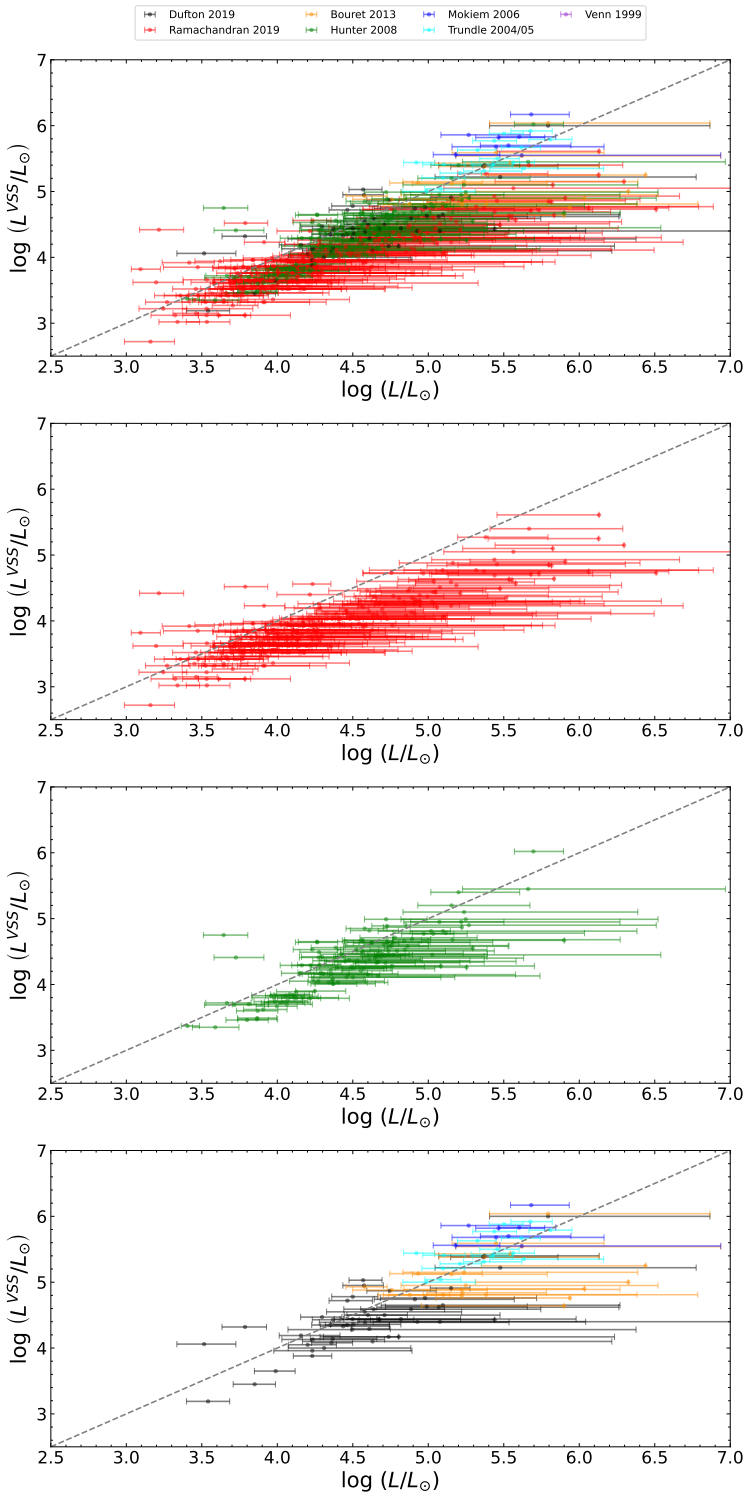


Figure 5.2: Bolometric luminosities of sources using bolometric correction tables compared to spectroscopically determined values from the VSS. The upper plot shows all VSS data. The subsequent plots underneath show only some VSS for clarity.

5.2 Caveats

Here we summarise all of the assumptions we made that led us to our results and how these could negatively impact the validity of our results.

5.2.1 Visual extinction

Firstly, visual extinction was selected based on a study of massive stars in the SMC. The visual extinction could be biased towards higher values as more massive stars form from massive molecular clouds, and in the SMC, where the metallicity is low, the weaker stellar winds may not have been able to clear out the surrounding molecular clouds. This would result in a higher visual extinction than for low-mass stars where the molecular cloud the star formed from is not as massive, and hence the star is not as extinct. Perhaps a location-based visual extinction determined from extinction maps of the SMC, such as Dobashi et al. 2009, could allow for a more precise determination of the SMC stars' effective temperature and bolometric luminosity. However, extinction maps are created using low-mass stars, which may not accurately describe the extent of the visual extinction of high-mass stars. There is still an open question on the depth of the SMC; therefore, the possibility of still over or underestimating the visual extinction is still plausible. As a first approach, we have shown that a nominal value of $A_v = 0.3$ with an uncertainty of ± 0.1, results in large uncertainties in the determined effective temperatures and bolometric luminosities of sources in the SMC. When we applied these uncertainties to the synthetic stellar population, the width of the MS was much larger than calculated for the observational data. Therefore, a smaller uncertainty on the visual extinction could be applied to the effective temperature and bolometric luminosity determination. However, the goal of this project was not to accurately determine the effective temperature and bolometric luminosity of individual stars in the SMC but rather to determine the blue-to-red supergiant ratio. Changing the uncertainty on the visual extinction would reduce the width of the Gaussian-shaped distribution of MS and AGB stars in the effective temperature histograms for the synthetic stellar population. A tighter distribution of stars in the effective temperature histogram could be better modelled, and more accurate numbers of MS, AGB, BSG, and RSG stars could be determined. If we can more accurately identify the main sequence, we would remove contamination in B, reducing the B/R. Additionally, more accurately identifying the AGB branch would reduce the number of contaminants in R, increasing the B/R. The overall effect of the B/R would either increase or decrease based on the ratio of high-mass main sequence stars to evolved AGB stars. Therefore, for younger dwarf galaxies and clusters, the B/R would decrease because there is not enough time for AGB stars to form. Conversely, in older dwarf galaxies and clusters with low star formation rates, the number of AGB stars could outweigh the number of high-mass main sequence stars and result in an increase in the B/R. Reducing the errors in the effective temperature and bolometric luminosity is something to strive for, and future work should look into constraining the visual extinction used to determine the effective temperatures and bolometric luminosities.

5.2.2 Surface gravity

The second assumption we made was the surface gravity of blue and red stars. Blue stars were assumed to have a high surface gravity of $\log(g/\mathrm{cm\,s^{-2}}) = 3.0$, which is appropriate for a high-mass star during the MS or shortly after but not during the envelope expansion phase after hydrogen shell ignition. This surface gravity is also not appropriate for low-mass stars that only reach this surface gravity late

into the helium-burning phase (see Figure 2.8). Red stars were assumed to have a surface gravity of $\log(g/\mathrm{cm\,s^{-2}}) = 0.0$, which is appropriate for stars that have expanded greatly during the late evolutionary phases. According to Figure 2.8, stars under $12\,\mathrm{M}_\odot$ struggle to reach $\log(g/\mathrm{cm\,s^{-2}}) = 0.0$ during their core helium-burning phase. The surface gravity assumption affects the values of bolometric corrections we use to create the colour-effective temperature and effective temperature-bolometric correction functions. Therefore, it is recognised that different surface gravities would result in different determined effective temperatures and bolometric luminosities of the sources in the SMC. As mentioned before, the project's goal was not to accurately determine these parameters for individual stars. It is worth noting that the surface gravity assumption has more of an effect on cooler stars, as the higher the temperature, the more often a blackbody assumption is used to determine the bolometric corrections. Therefore, it is recognised that the chosen value of surface gravity will affect the determined number of RSGs more than BSGs. Future work could improve on this by taking more surface gravity assumptions between $\log(g/\mathrm{cm\,s^{-2}}) = 0.0$ and 3.0.

5.2.3 Colour separation

Additionally, blue and red stars were separated by a colour of $G_{\mathrm{bp}} - G_{\mathrm{rp}} = 0.872$. Looking at Figure 2.4, this corresponded to an effective temperature of $\log(T_{\mathrm{eff}/K}) = 3.673$ or $\sim 5\,800\,\mathrm{K}$, which is higher than others consider as the upper-temperature limit of RSGs (Massey and Olsen 2003, Massey, Neugent et al. 2021). If we take a redder colour to separate the blue and red stars, we arrive at a higher blue-to-red supergiant ratio as we consider cooler stars as blue. There is no strict definition of the exact temperature to consider a supergiant blue or red; therefore, it is difficult to improve on this choice. One could choose to keep in line with other work for easy comparisons. However, $G_{\mathrm{bp}} - G_{\mathrm{rp}} = 0.872$ was chosen to avoid using bolometric corrections determined from blackbody assumptions in lower temperatures where this approximation is not the best representation of reality.

5.2.4 Distance

Moreover, the distance to each source in the SMC is kept the same. This affects the determined bolometric luminosity of each source. As there is no consensus on the full extent of the SMC, it was safer to use a single value of distance than to consider an array of different distances. As the distance to each star is unknown, it would be impossible to use individual distances. If we assume the SMC is spherical with a diameter of 5.78 kpc according to NASA/IPAC Extragalactic Database NED[3], the uncertainty on distance is less than 10%, which is also less significant than the uncertainty on the visual extinction. As this affects the vertical position of sources in the HRD, there is a possibility that changing the distance to the source will shift the distribution of blue and red supergiants into different luminosity bins. There is no reason to suggest that more stars would be moved into upper luminosity bins than lower, and therefore, we expect the effect on the ratio of blue-to-red supergiant stars to be minimal.

5.2.5 Effective temperature histograms

In addition, horizontal luminosity bins were taken when building the effective temperature histograms. Another suggested method is to build effective temperature histograms perpendicular to the ZAMS. This way, we minimise the effect of comparing younger high-mass stars with more evolved low-mass stars.

[3] http://ned.ipac.caltech.edu/

The distribution of stars in the effective temperature plane would no longer be Gaussian shaped, and we would not be able to use the Gaussian-fit method to determine the number of main sequence and AGB stars. It is uncertain how this could affect the blue-to-red supergiant ratio as there would need to be new assumptions to be made on where to start counting helium-burning stars. An advantage to using horizontal luminosity bins is that we can implement the Gaussian-fit method to determine the number of blue and red supergiants. Future work could investigate the results of creating histograms perpendicular to the ZAMS and compare them against the horizontal luminosity bin case.

5.2.6 Synthetic stellar populations

Further assumptions we made affect the synthetic stellar populations. It is extremely computationally expensive to create synthetic stellar populations equivalent to the number of stars in the SMC, and therefore, less massive synthetic populations were tested. We produced a similar number of stars above $\log(L/\mathrm{L}_{\odot}) = 3.5$ for as many synthetic populations as possible. As as seen from the effect of changing the efficiency of semiconvection and overshooting, the lifetime of the stars changes and therefore, the number of stars that survive up to 100 Myrs increases. Using a random number generator to determine the ages and masses of stars is safer on larger data sets as low-number statistics more afflict smaller data sets. Larger stellar populations should be tested in the future to see if the resulting best-fit parameters are the same.

A constant star formation rate was assumed, and the ages of stars were assumed to be between 0 and 100 Myr. In reality, the SMC has active star-forming regions. In the future, the SMC's star formation history (SFH) could be used to obtain a more realistic distribution of the ages of stars in the synthetic stellar populations. The ages of the stars determine their position in the HRD. If the star formation rate is high and there are more young high-mass stars, we will determine a larger number of blue supergiants through method A, increasing the B/R. If the star formation rate is low, then we should have fewer high-mass main sequence stars as many will have died out due to their short lifetimes, and additionally, we will have more AGB stars. This would inflate the number of RSGs we determine in method A and lower the B/R.

By increasing the age range in the synthetic stellar population, we can postulate that by assuming stars can be older than we have previously accounted for, more low-mass SGs are expected. Such stars will be $5\,\mathrm{M}_{\odot}$ stars which we have shown in the synthetic stellar populations, can be bright enough to be considered in our luminosity range. Whether these stars are blue or red depends on the internal mixing efficiencies. Active star-forming regions could increase the number of massive stars in the synthetic populations, inflating the total number of SGs. Depending on the internal mixing efficiencies, this would result in more RSGs or BSGs.

Rather than computing more models of different masses, we used the minimum uncertainty in the determined effective temperatures and bolometric luminosities of the observed stars to scatter the synthetic population in the HRD. More modelled masses and a lower uncertainty on the visual extinction would reduce the spread of points in the HRD. Firstly, more modelled masses would mean stars can occupy different positions in the HRD before being scattered by the minimum uncertainty approach we took. Therefore, a lower uncertainty in the visual extinction could be taken and still remove the resolution effects of the synthetic population. This would only be possible if we were more certain about the range of visual extinction on high-mass stars in the SMC. Therefore, we would reduce the number of stars that are hotter than the ZAMS, making the effective temperature and bolometric luminosity values more trustworthy. Reducing the spread of data points would tighten the Gaussian-like distribution of stars in

the effective temperature histogram, which could better determine the number of MS and AGB stars under the Gaussian curve. This would result in a better determination of the number of BSGs and RSGs and, therefore, the B/R. As we can see from Figure 4.3, there is still a good agreement between method A and method B in determining the B/R, and both methods are able to reproduce the observational trend of the B/R in the SMC.

5.3 Best-fit parameters

By comparing observations and the grid of synthetic stellar populations, the best-fit parameters were $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$. This implies that semiconvective mixing is inefficient and convective overshooting is non-existent. If we compare our results to Schootemeijer, Langer, Grin et al. 2019, who studied the B/R in the SMC in the luminosity range $4.8 \leq \log(L/L_{\odot}) \leq 6.0$, we reach different internal mixing efficiencies. Schootemeijer, Langer, Grin et al. 2019 quote that inefficient semiconvection never produces the observed number of blue supergiants regardless of the efficiency of overshooting. We confirm the same in the upper luminosity bins $\log(L/L_{\odot}) \geq 4.5$ for our best-fit parameters. However, we find good agreement between the best-fitting mixing efficiencies and the observed blue-to-red supergiants ratio for the lower luminosity bins.

The synthetic stellar population for the best-fit parameters show more red supergiants in the luminosity range $\log(L/L_{\odot}) \geq 5.5$ than is observed in the SMC. In the observed HRD only one RSG has a $\log(L/L_{\odot}) \geq 5.5$, whereas the synthetic population has 7. The number of blue supergiants in the luminosity range $\log(L/L_{\odot}) \geq 5.0$ is roughly the same between the observed and synthetic population as seen by Figure 2.20 and Figure 3.4. However, as mentioned in the methods section, we conservatively selected helium-burning stars as stars with $\log(T_{\rm eff}/K) \leq 4.25$. Therefore, the B/R depends on the exact temperature we choose here, and conclusions drawn from the highest luminosity bins were considered difficult. Therefore, internal mixing efficiencies that matched well with observations in the lower luminosity bins carry more significance in this work. This is because the B/R is not driven by a strict effective temperature limit on the supergiants, but the B/R is determined from the Gaussian method.

If we consider more efficient semiconvection as predicted by Schootemeijer, Langer, Grin et al. 2019, such as $\alpha_{sc} = 10.0$ and $\alpha_{ov} = 0.23$, we find that there is a good agreement between the synthetic B/R and observational B/R in the luminosity bin $4.5 \leq \log(L/L_{\odot}) \leq 5.0$. However, there are no RSGs within $5.0 \leq \log(L/L_{\odot}) \leq 5.5$ to confirm these efficiencies reproduce the observed trend at higher luminosities. These mixing parameter efficiencies do not agree with observational B/R at lower luminosities. Therefore, we conclude the mixing parameter efficiencies, as suggested by Schootemeijer, Langer, Grin et al. 2019, do not produce the observed blue-to-red supergiant ratio in the SMC.

In this work, there does not exist a set of internal mixing parameter efficiencies that consistently reproduce the observed B/R in the upper and lower luminosity bins better than $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$. This leads us to the conclusion that we are missing some physics implemented in the models, which can include but is not limited to rotation and binarity effects. The results of the best-fit parameters show that semiconvective mixing does not appear to be efficient in stars, and convective overshooting either does not exist or the efficiency is between 0.0 and 0.12.

5.4 Comparison of B/R with other work

The total observed B/R was determined to be B/R = $1.389^{+0.207}_{-0.452}$ and the best-fit parameters resulted in B/R = 1.706 using method A. The B/R determined through method A is slightly larger than the upper limit of the observed B/R, which is 1.596 for $A_v = 0.2$. This is expected as the B/R of the synthetic stellar population is more often larger than the observed trend for $A_v = 0.2$. An exception to this is in for luminosities $\log(L/L_\odot) \geq 4.5$, where there are fewer stars to greatly impact the total B/R.

Schootemeijer, Langer, Grin et al. 2019 quotes a B/R = 0.852 in the luminosity range $4.8 \leq \log(L/L_\odot) \leq 6.0$, which falls short of our lower uncertainty limits for $A_v = 0.4$ but can be expected from our wider luminosity range. Eggenberger, Meynet and Maeder 2002 shows that the B/R = 0.6 when only considering B-type supergiants in B and K-type and M-type supergiants in R for the cluster NGC 330 in the SMC. This B/R is expected to be larger when considering A-type supergiants in B and further possible contamination of O-stars in B. The observed B/R is larger than in both of these studies. This can be accounted for by using a more complete sample of photometric data from GAIA DR3. The increase in the B/R is suggestive of an overestimation of the number of blue supergiants due to the contamination of O-type main sequence stars with an underestimated visual extinction. Extinction maps of the SMC could be utilised to more accurately account for the visual extinction of the sources in the SMC to study the impact on the observed B/R. However, the contamination of the expected helium-burning region in the HRD by main sequence stars is a more complex problem that needs to be solved in future work.

The B/R for the galactic centre of the Milky Way is B/R = 4.1 when considering the same as above for NGC 330 (Eggenberger, Meynet and Maeder 2002), showing that the B/R is, in fact, a rising function of metallicity as predicted by Eggenberger, Meynet and Maeder 2002. However, if we use the SMC metallicity in the function provided by Eggenberger, Meynet and Maeder 2002 (Equation 5.1) and include O, B, and A supergiant stars in B, and $(B/R)_\odot = 3.0$, we predict a B/R = 0.273 for the SMC. This contrasts with the value we determine and therefore shows that the B/R is a rising function of metallicity, but it is not as steep as suggested by Eggenberger, Meynet and Maeder 2002. The mass range of supergiants in their work is limited between $8\,M_\odot$ and $30\,M_\odot$, whereas our work includes stars from $5\,M_\odot$ to $100\,M_\odot$. As it is not possible for us to determine the exact number of O, B and A supergiant stars from the photometric data we present here, the B/R is expected to be higher than the equation predicts.

$$\frac{B/R}{(B/R)_\odot} = 0.05 \cdot e^{3 \cdot \frac{Z}{Z_\odot}} \tag{5.1}$$

CHAPTER 6

Conclusion

A Hertzsprung-Russell diagram of the Small Magellanic Cloud was created by use of GAIA DR3 photometric data and bolometric correction tables from a grid of 1D stellar atmosphere models. The determined effective temperatures and bolometric luminosities of sources in the SMC are within good agreement with the sources in various spectroscopic surveys (VSS), especially when we exclude the SMC wing sources. The total observed blue-to-red supergiant ratio is determined to be B/R = $1.389^{+0.207}_{-0.452}$ and is comparable with other work as discussed in Section 5.4.

Various synthetic stellar populations were created using a previously run grid of stellar evolutionary models courtesy of Schootemeijer, Langer, Grin et al. 2019 with SMC metallicity, varying masses, and varying internal mixing parameter efficiencies α_{sc} and α_{ov}. A constant star formation rate and Kroupa 2001's initial mass function (IMF) were used to draw synthetic HRDs. The effective temperature and bolometric luminosity of each star were determined by the model with the closest mass as drawn by the IMF and then by finding $\log(T_{\rm eff}/{\rm K})$ and $\log(L/{\rm L}_\odot)$ for the closest modelled age. To remove resolution effects, the uncertainties in effective temperatures and bolometric luminosities were used to create so-called dirty data. We use Python's `np.random.normal()` function to randomly select a value of $\log(T_{\rm eff}/{\rm K})$ and $\log(L/{\rm L}_\odot)$ based on the uncertainties of those parameters for a given $\log(T_{\rm eff}/{\rm K})$. After removing the resolution effects, we could directly compare the blue-to-red supergiant ratio with the observed trend.

After removing the resolution effects in the data, the width of the main sequence in the synthetic stellar population arises from the uncertainties in $\log(T_{\rm eff}/{\rm K})$ and $\log(L/{\rm L}_\odot)$. These uncertainties are determined mainly by the uncertainty on visual extinction $A_v = 0.3 \pm 0.1$. The width of the main sequence is larger than the theoretical width of the main sequence in all of the synthetic stellar populations. This indicates that the uncertainty value on visual extinction could be reduced to ± 0.1 in future work. This highlights that the main source of uncertainty in this thesis was overestimated, and the reduction of this uncertainty could make way for more precise determinations of the B/R.

A blue-to-red supergiant ratio was determined for each combination of internal mixing efficiencies. The best-fit parameters that matched with the observed ratio were $\alpha_{sc} = 0.1$ and $\alpha_{ov} = 0.0$. These efficiencies fail to accurately recreate the observed distribution of blue and red supergiants at higher luminosities, but these parameters best match observations at lower luminosities. More efficient semiconvection, $\alpha_{sc} > 0.1$, produced too few red supergiants at luminosities $\log(L/{\rm L}_\odot) \geq 4.25$ than observed. More effective overshooting $\alpha_{ov} \geq 0.11$, extended the main sequence lifetime of stars, reduced the total number of supergiants, and decreased the B/R. Therefore, even though a set of best-fit parameters were

determined, other mechanisms, such as rotation and binarity, are likely to influence these results.

There is further possible room for improvement in future work by reducing the uncertainty on A_v, considering stars with different surface gravities other than $\log(g/\mathrm{cm\,s^{-2}}) = 0.0$ and 3.0, as well as other assumptions covered in Section 5.

To summarise, appropriate steps were taken to make justified decisions at every obstacle in determining the best-fit internal mixing parameter efficiencies that closely reproduced the observed blue-to-red supergiant ratio in the SMC without considering binarity or rotation effects. However, the effects of binarity on the ratio are not yet well understood, and work should be done to deepen our understanding on this front.

APPENDIX A

Appendix

A.1 Colour-effective temperature relations

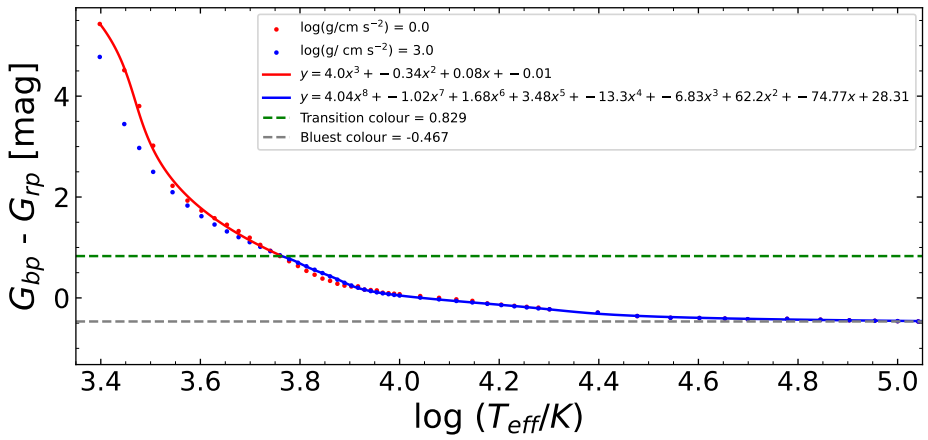


Figure A.1: Colour-effective temperature relation for $A_v = 0.2$ using bolometric corrections and effective temperatures from MIST. The red points are for a log(g / cms^{-2}) = 0.0 and the blue points for log(g / cms^{-2}) = 3.0. Red stars are those with colours above the green line, and blue stars are those with colours below the green line.

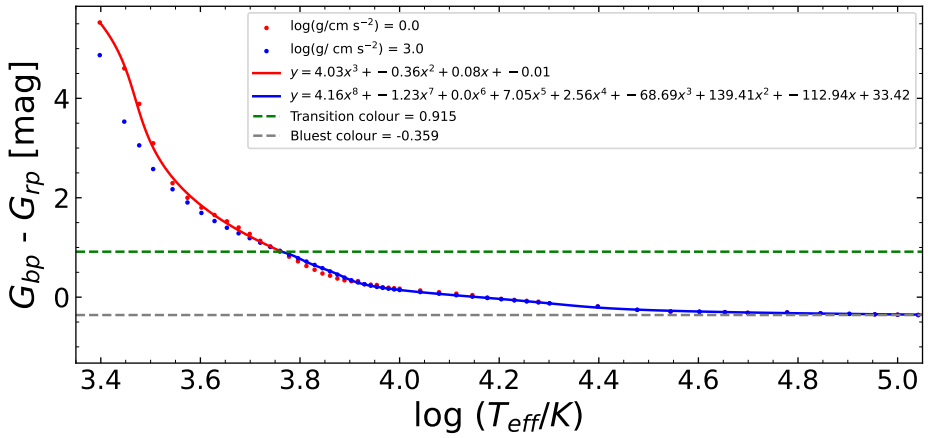


Figure A.2: Colour-effective temperature relation for $A_v = 0.4$ using bolometric corrections and effective temperatures from MIST. The red points are for a log(g / cms^{-2}) = 0.0 and the blue points for log(g / cms^{-2}) = 3.0. Red stars are those with colours above the green line, and blue stars are those with colours below the green line.

A.2 Effective temperature-bolometric correction relations

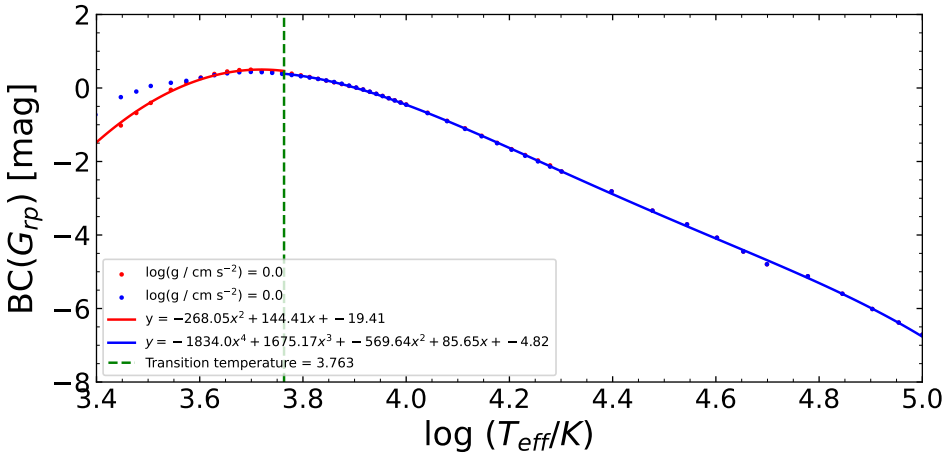


Figure A.3: Temperature-bolometric correction relation for $A_v = 0.2$ using bolometric corrections and effective temperatures from MIST. The red points are for a $\log(g / cms^{-2}) = 0.0$ and the blue points for $\log(g / cms^{-2}) = 3.0$. Red stars are those with colours below the green line, and blue stars are those with colours above the green line.

brightest blue star

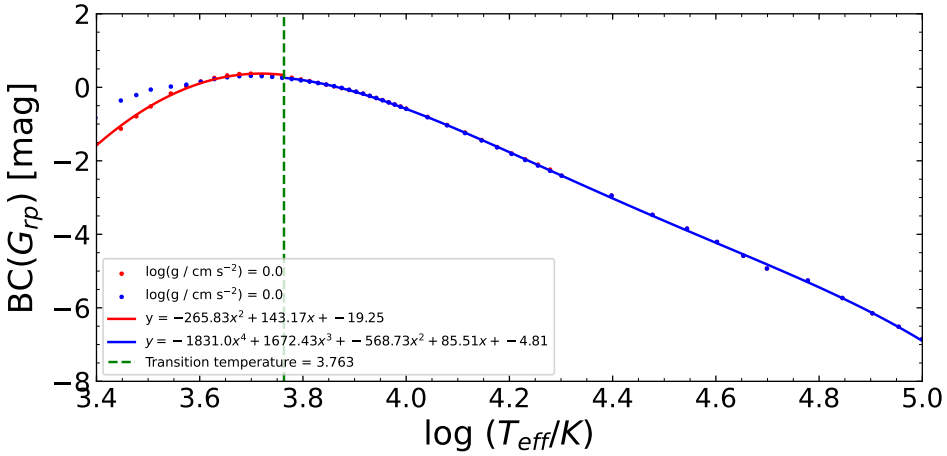


Figure A.4: Temperature-bolometric correction relation for $A_v = 0.4$ using bolometric corrections and effective temperatures from MIST. The red points are for a $\log(g / cms^{-2}) = 0.0$ and the blue points for $\log(g / cms^{-2}) = 3.0$. Red stars are those with colours below the green line, and blue stars are those with colours above the green line.

A.3 OBe stars

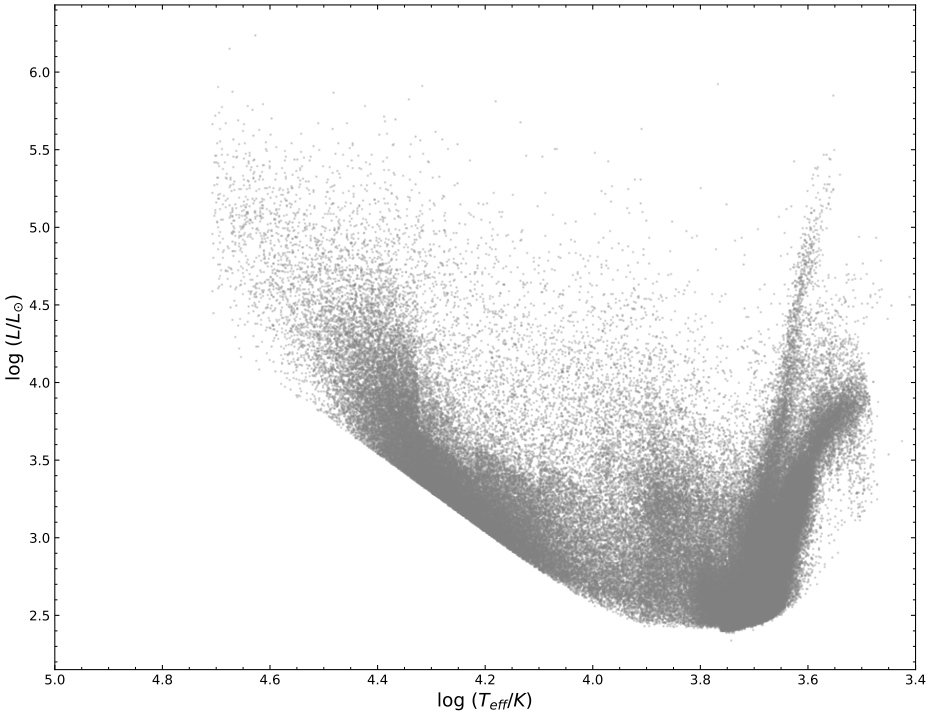


Figure A.5: Hertzsprung-Russell diagram of stars in the Small Magellanic Cloud without OBe stars.

A.4 Helium burning stars

$log(L/L_{\odot})$	MS	BSG	RSG	AGB	OBe	$\frac{SG}{(MS + SG)}$
6.0 - 5.5	9	5	1	0	0	0.4
5.5 - 5.0	217	45	55	1	0	0.315
5.0 - 4.5	1012	264	293	1	0	0.355
4.5 - 4.25	1028	329	319	1	30	0.387
4.25 - 4.0	1865	620	237	529	125	0.315
4.0 - 3.75	3111	1195	515	1640	497	0.355
3.75 - 3.5	4757	2435	1236	3097	1224	0.436
Total	11999	4893	2656	5269	1876	0.386

Table A.1: The number of main sequence (MS), blue helium supergiant stars (BSG), red supergiant stars (RSG), asymptotic giant branch (AGB), non-supergiant OB stars (OBe), and all supergiant stars (SG) to main sequence stars plus supergiant stars in varying luminosity intervals for $A_v = 0.2$.

$log(L/L_{\odot})$	MS	BSG	RSG	AGB	OBe	$\frac{SG}{(MS + SG)}$
6.0 - 5.5	538	8	7	0	0	0.027
5.5 - 5.0	1297	42	75	1	2	0.083
5.0 - 4.5	3015	251	360	1	54	0.169
4.5 - 4.25	2813	323	436	1	209	0.212
4.25 - 4.0	4505	626	400	619	710	0.185
4.0 - 3.75	6776	1049	775	2408	1405	0.212
3.75 - 3.5	9434	2245	1788	2671	695	0.299
Total	28378	4544	3841	5701	3075	0.228

Table A.2: The number of main sequence (MS), blue helium supergiant stars (BSG), red supergiant stars (RSG), asymptotic giant branch (AGB), non-supergiant OB stars (OBe), and all supergiant stars (SG) to main sequence stars plus supergiant stars in varying luminosity intervals for $A_v = 0.4$.

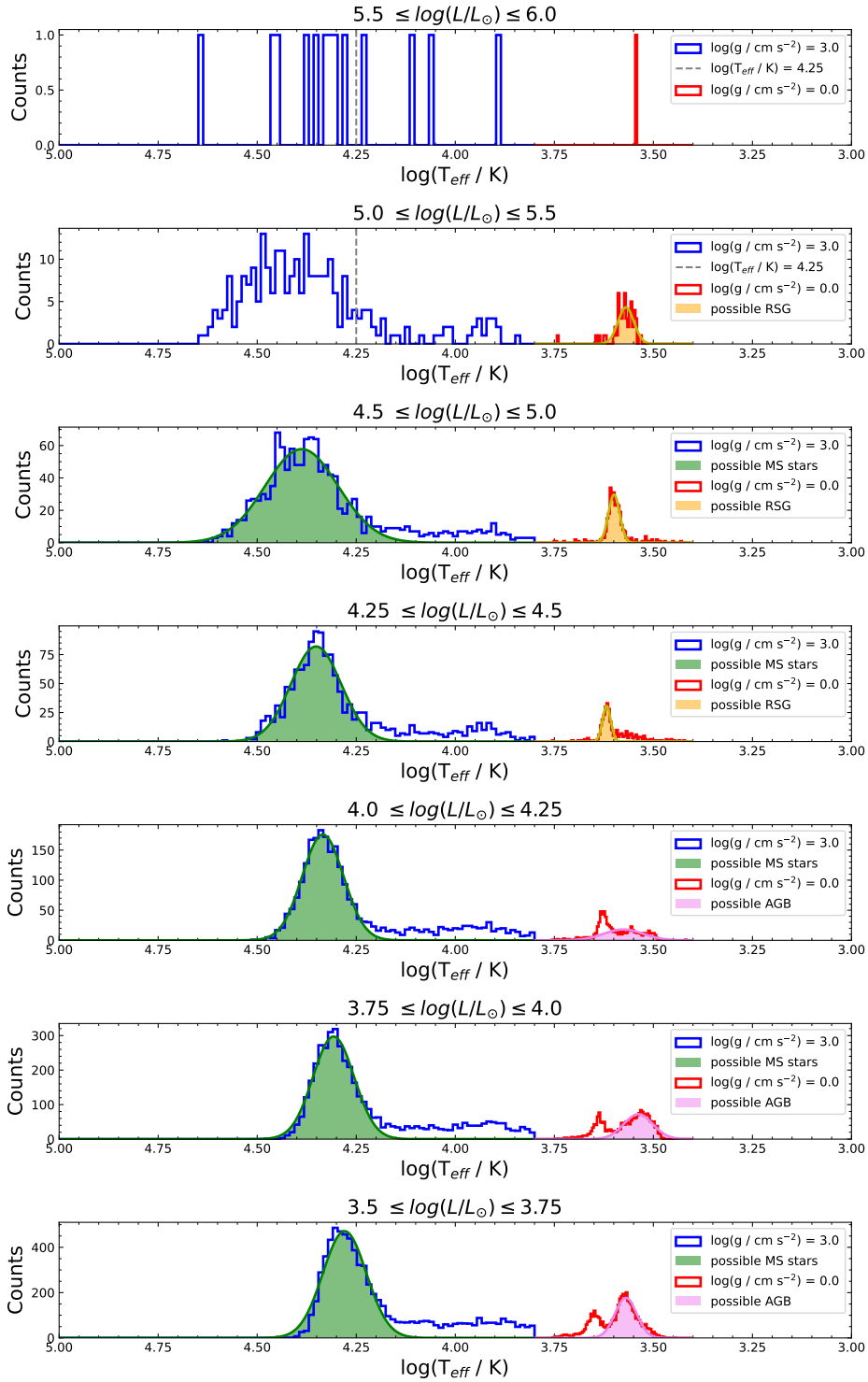


Figure A.6: Effective temperature histograms for intervals of bolometric luminosity for $A_v = 0.2$. Blue shows the stars with log(g / cms^{-2}) = 3.0 and red shows the stars with log(g / cms^{-2}) = 0.0. Gaussian fits were used to identify the main sequence.

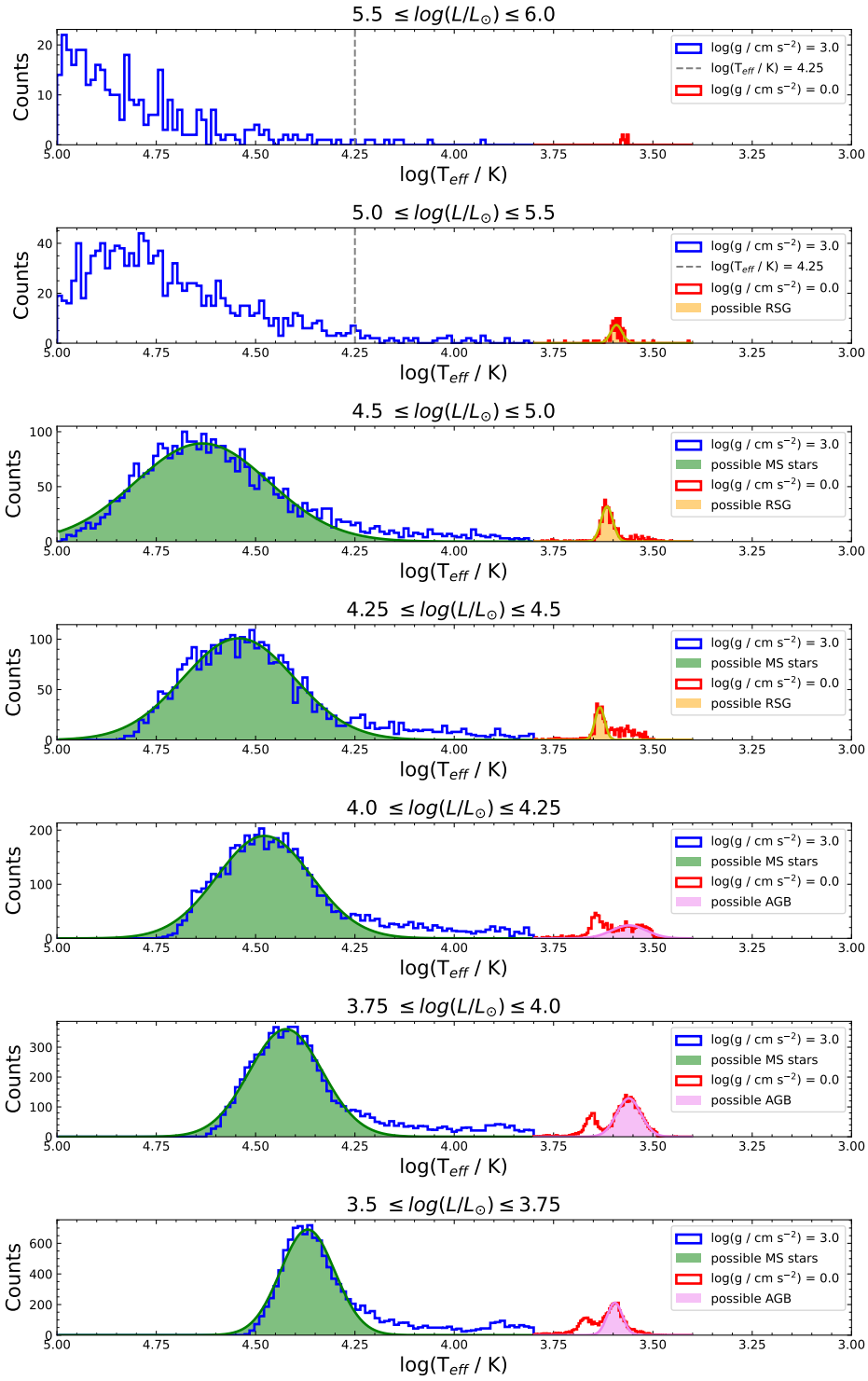
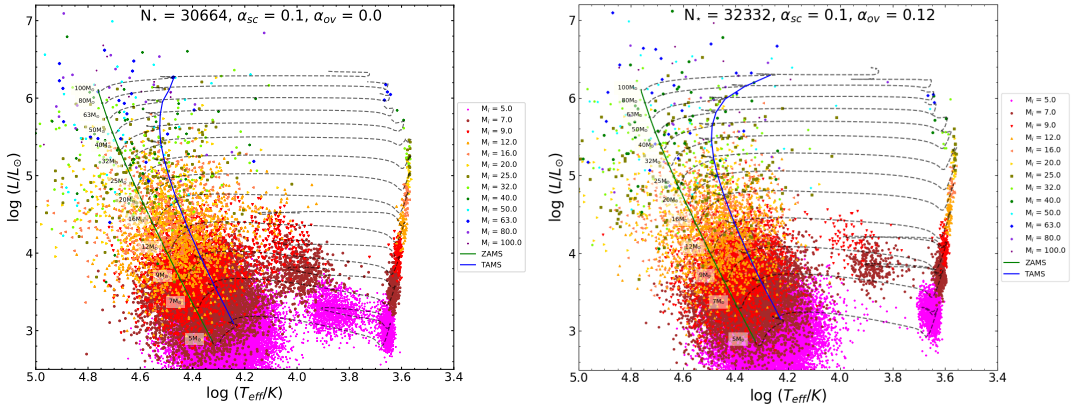
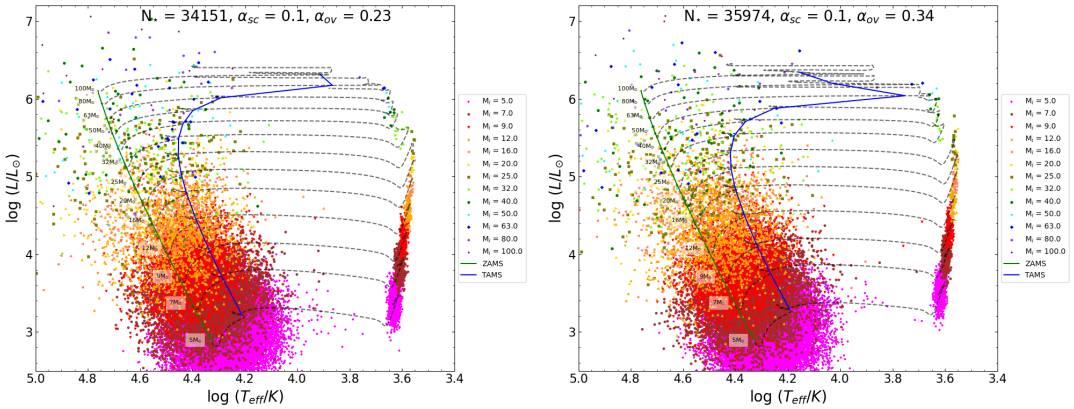


Figure A.7: Effective temperature histograms for intervals of bolometric luminosity for $A_v = 0.4$. Blue shows the stars with log(g / cms^{-2}) = 3.0 and red shows the stars with log(g / cms^{-2}) = 0.0. Gaussian fits were used to identify the main sequence.

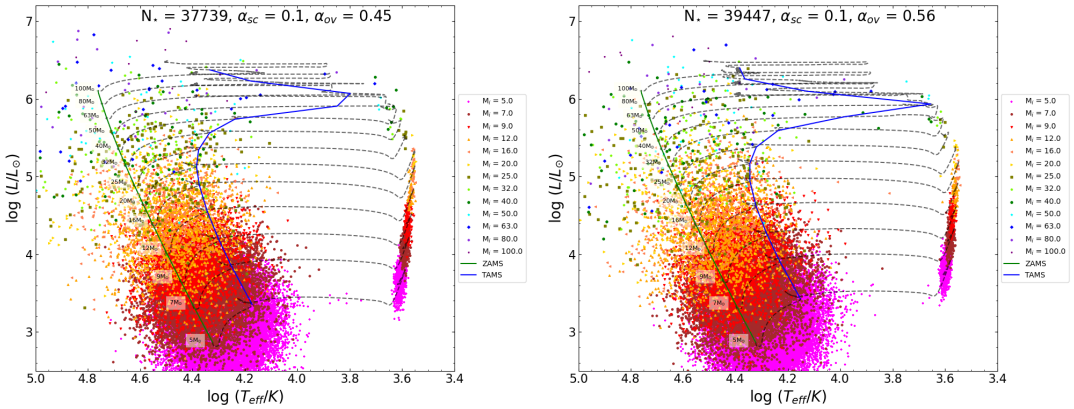
A.5 HRDs of the synthetic stellar populations



(a) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.0$ (b) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.12$

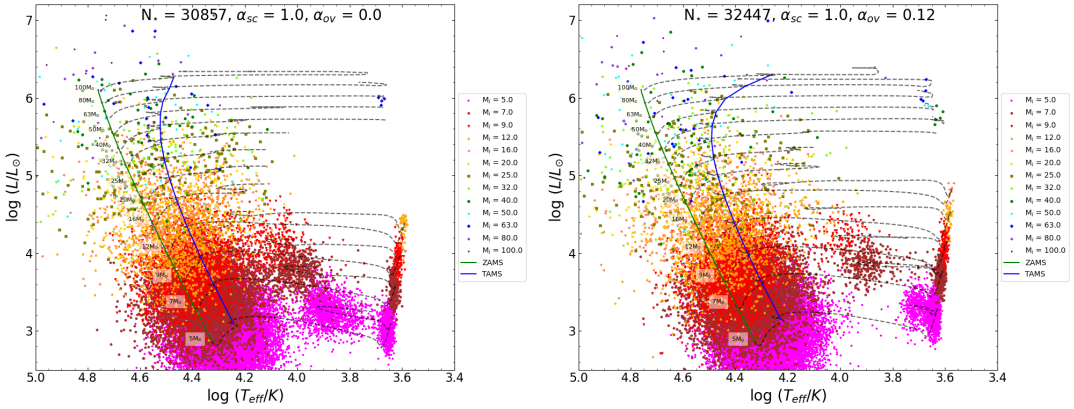


(c) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.23$ (d) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.34$



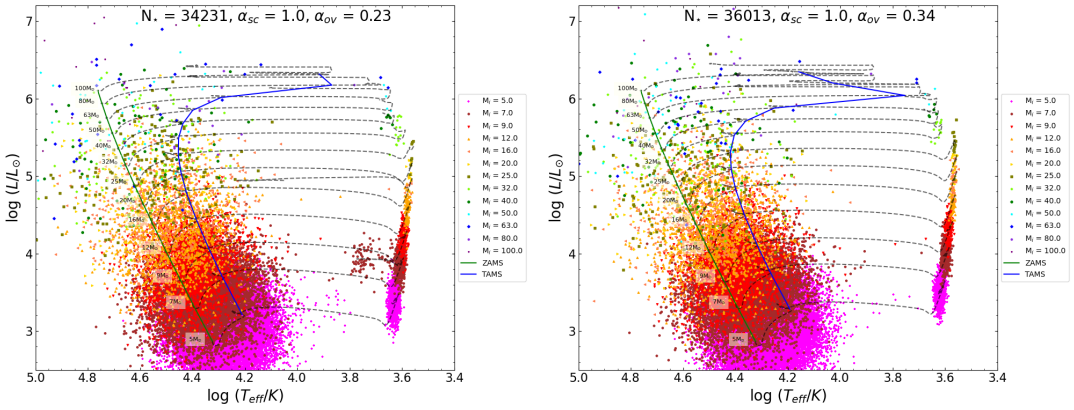
(e) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.45$ (f) $\alpha_{sc} = 0.1$, $\alpha_{ov} = 0.56$

Figure A.8: HRD of the synthetic stellar population for varying $\alpha_{sc} = 0.1$ and varying overshooting efficiencies.



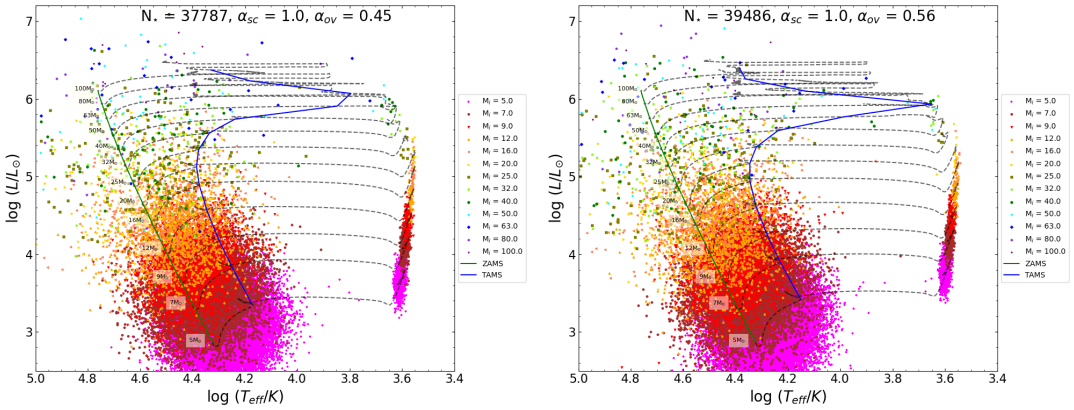
(a) $\alpha_{\rm sc} = 1.0$, $\alpha_{\rm ov} = 0.0$

(b) $\alpha_{\rm sc} = 1.0$, $\alpha_{\rm ov} = 0.12$



(c) $\alpha_{\rm sc} = 1.0$, $\alpha_{\rm ov} = 0.23$

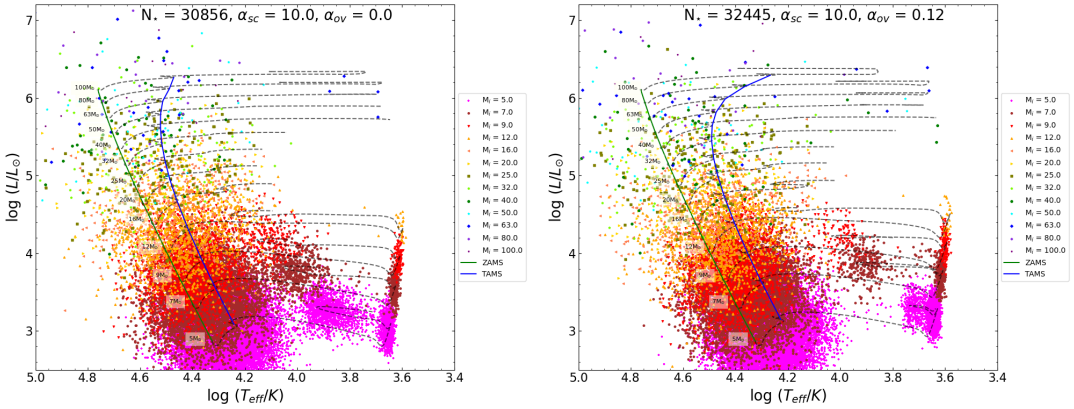
(d) $\alpha_{\rm sc} = 1.0$, $\alpha_{\rm ov} = 0.34$



(e) $\alpha_{\rm sc} = 1.0$, $\alpha_{\rm ov} = 0.45$

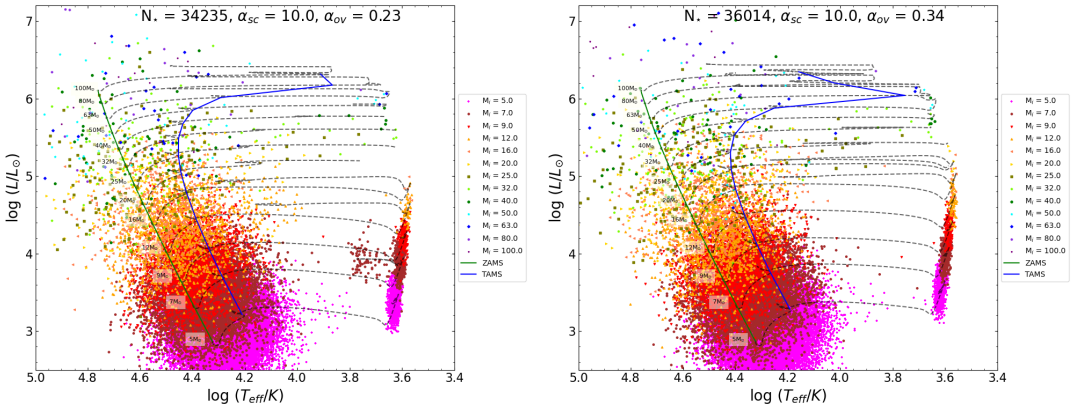
(f) $\alpha_{\rm sc} = 1.0$, $\alpha_{\rm ov} = 0.56$

Figure A.9: HRD of the synthetic stellar population for varying $\alpha_{\rm sc} = 1.0$ and varying overshooting efficiencies.



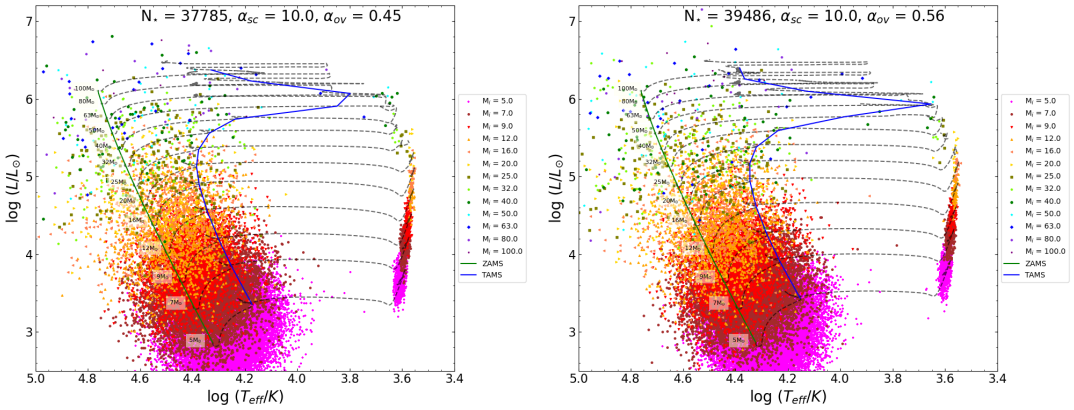
(a) $\alpha_{\rm sc} = 10.0$, $\alpha_{\rm ov} = 0.0$

(b) $\alpha_{\rm sc} = 10.0$, $\alpha_{\rm ov} = 0.12$



(c) $\alpha_{\rm sc} = 10.0$, $\alpha_{\rm ov} = 0.23$

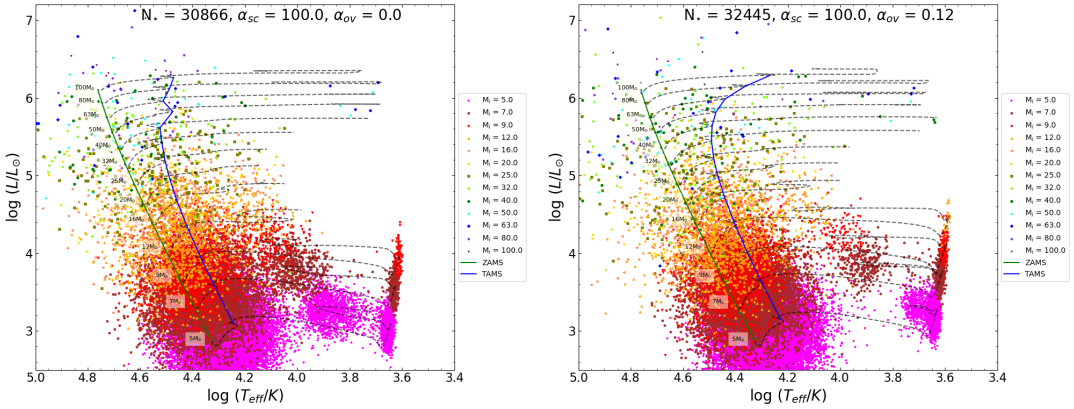
(d) $\alpha_{\rm sc} = 10.0$, $\alpha_{\rm ov} = 0.34$



(e) $\alpha_{\rm sc} = 10.0$, $\alpha_{\rm ov} = 0.45$

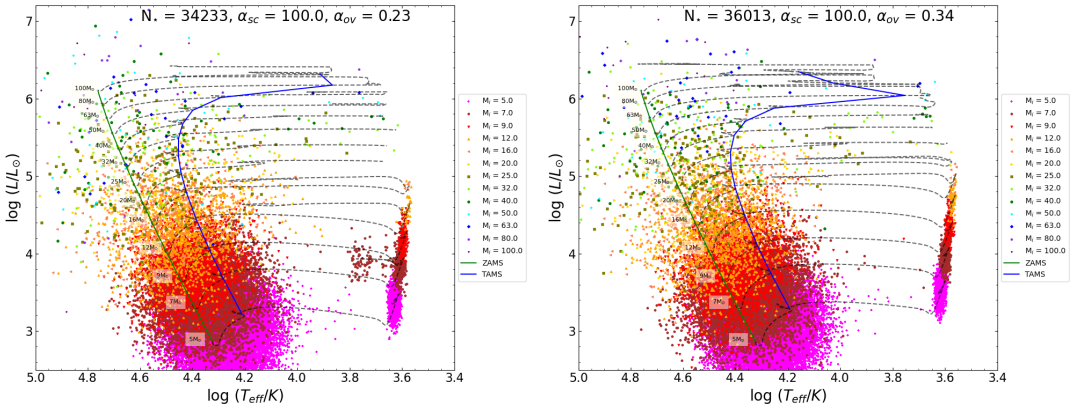
(f) $\alpha_{\rm sc} = 10.0$, $\alpha_{\rm ov} = 0.56$

Figure A.10: HRD of the synthetic stellar population for varying $\alpha_{\rm sc} = 10.0$ and varying overshooting efficiencies.



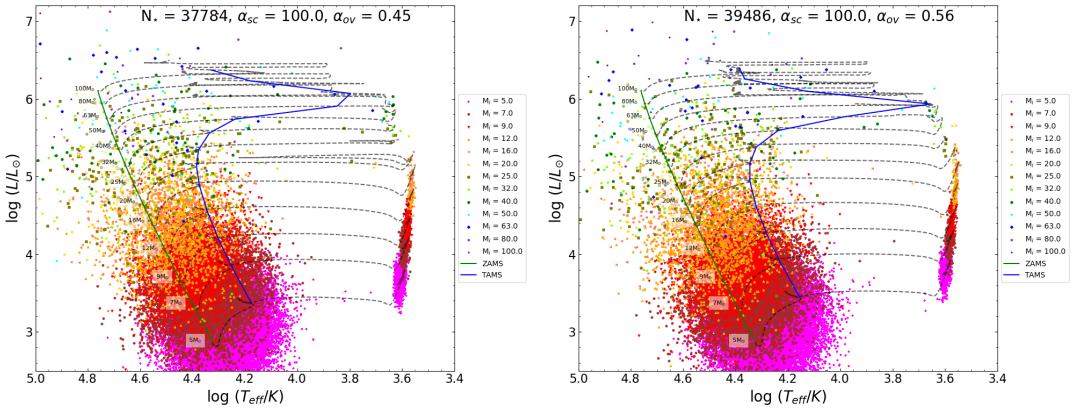
(a) $\alpha_{\mathrm{sc}} = 100.0$, $\alpha_{\mathrm{ov}} = 0.0$

(b) $\alpha_{\mathrm{sc}} = 100.0$, $\alpha_{\mathrm{ov}} = 0.12$



(c) $\alpha_{\mathrm{sc}} = 100.0$, $\alpha_{\mathrm{ov}} = 0.23$

(d) $\alpha_{\mathrm{sc}} = 100.0$, $\alpha_{\mathrm{ov}} = 0.34$



(e) $\alpha_{\mathrm{sc}} = 100.0$, $\alpha_{\mathrm{ov}} = 0.45$

(f) $\alpha_{\mathrm{sc}} = 100.0$, $\alpha_{\mathrm{ov}} = 0.56$

Figure A.11: HRD of the synthetic stellar population for varying $\alpha_{\mathrm{sc}} = 100.0$ and varying overshooting efficiencies.

Bibliography

Bailer-Jones, C. A. L., J. Rybizki, M. Fouesneau, M. Demleitner and R. Andrae (2021), *Estimating Distances from Parallaxes. V. Geometric and Photogeometric Distances to 1.47 Billion Stars in Gaia Early Data Release 3*, **161**, 147 147 (cit. on p. 9).

Bouchet, P., J. Lequeux, E. Maurice, L. Prevot and M. L. Prevot-Burnichon (1985), *The visible and infrared extinction law and the gas-to-dust ratio in the Small Magellanic Cloud.*, **149** 330 (cit. on p. 44).

Bouret, J. .-. et al. (2013), *Massive stars at low metallicity. Evolution and surface abundances of O dwarfs in the SMC*, **555**, A1 A1, arXiv: 1304.6923 [astro-ph.SR] (cit. on pp. 4, 44–46).

Chiosi, C. and A. Maeder (1986), *The evolution of massive stars with mass loss.*, **24** 329 (cit. on p. 2).

Choi, J. et al. (2016), *Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models*, **823**, 102 102, arXiv: 1604.08592 [astro-ph.SR] (cit. on pp. 13, 15).

Creevey, O. L. et al. (2022), *Gaia Data Release 3: Astrophysical parameters inference system (Apsis) I – methods and content overview*, arXiv e-prints, arXiv:2206.05864 arXiv:2206.05864, arXiv: 2206.05864 [astro-ph.GA] (cit. on p. 16).

Davies, B., P. A. Crowther and E. R. Beasor (2018), *VizieR Online Data Catalog: Luminous cool supergiants in Magellanic Clouds (Davies+, 2018)*, VizieR Online Data Catalog, J/MNRAS/478/3138 J/MNRAS/478/3138 (cit. on p. 4).

Dobashi, K. et al. (2009), *Extinction Map of the Small Magellanic Cloud Based on the SIRIUS and 6X 2MASS Point Source Catalogs*, **137** 5099 (cit. on pp. 13, 50).

Dotter, A. (2016), *MESA Isochrones and Stellar Tracks (MIST) 0: Methods for the Construction of Stellar Isochrones*, **222**, 8 8, arXiv: 1601.05144 [astro-ph.SR] (cit. on p. 13).

Dufton, P. L., C. J. Evans, I. Hunter, D. J. Lennon and F. R. N. Schneider (2019), *A census of massive stars in NGC 346. Stellar parameters and rotational velocities*, **626**, A50 A50, arXiv: 1905.03359 [astro-ph.SR] (cit. on pp. 4, 44–46).

Eggenberger, P., G. Meynet and A. Maeder (2002), *The blue to red supergiant ratio in young clusters at various metallicities*, **386** 576, arXiv: astro-ph/0202478 [astro-ph] (cit. on p. 54).

Feltzing, S., J. Holmberg and J. R. Hurley (2001), *The solar neighbourhood age-metallicity relation - Does it exist?*, **377** 911, arXiv: astro-ph/0108191 [astro-ph] (cit. on p. 25).

Gaia Collaboration, A. G. A. Brown et al. (2021), *Gaia Early Data Release 3. Summary of the contents and survey properties*, **649**, A1 A1 (cit. on p. 9).

Gaia Collaboration, A. Vallenari et al. (2023),
Gaia Data Release 3. Summary of the content and survey properties, **674**, A1 A1 (cit. on pp. 5, 9, 44).
Gordon, K. D., G. C. Clayton, K. A. Misselt, A. U. Landolt and M. J. Wolff (2003),
A Quantitative Comparison of the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way Ultraviolet to Near-Infrared Extinction Curves, **594** 279 (cit. on pp. 13, 14).
Graczyk, D. et al. (2020), *A Distance Determination to the Small Magellanic Cloud with an Accuracy of Better than Two Percent Based on Late-type Eclipsing Binary Stars*, **904**, 13 13, arXiv: 2010.08754 [astro-ph.GA] (cit. on pp. 1, 9).
Hainich, R. et al. (2015),
Wolf-Rayet stars in the Small Magellanic Cloud. I. Analysis of the single WN stars, **581**, A21 A21, arXiv: 1507.04000 [astro-ph.SR] (cit. on p. 14).
Harris, J. and D. Zaritsky (2004), *The Star Formation History of the Small Magellanic Cloud*, **127** 1531, arXiv: astro-ph/0312100 [astro-ph] (cit. on p. 29).
Hunter, I. et al. (2008), *The VLT-FLAMES survey of massive stars: atmospheric parameters and rotational velocity distributions for B-type stars in the Magellanic Clouds*, **479** 541, arXiv: 0711.2264 [astro-ph] (cit. on pp. 4, 44–46).
Korn, A. J., S. R. Becker, C. A. Gummersbach and B. Wolf (2000),
Chemical abundances from Magellanic cloud B stars, **353** 655 (cit. on p. 1).
Kroupa, P. (2001), *On the variation of the initial mass function*, **322** 231, arXiv: astro-ph/0009005 [astro-ph] (cit. on pp. 29, 55).
Langer, N. (2012), *Presupernova Evolution of Massive Single and Binary Stars*, **50** 107, arXiv: 1206.5443 [astro-ph.SR] (cit. on p. 1).
Langer, N. and A. Maeder (1995), *The problem of the blue-to-red supergiant ratio in galaxies.*, **295** 685 (cit. on pp. 1, 2).
Luri, X. et al. (2018), *Gaia Data Release 2. Using Gaia parallaxes*, **616**, A9 A9 (cit. on p. 9).
Mamajek, E. E. et al. (2015), *IAU 2015 Resolution B2 on Recommended Zero Points for the Absolute and Apparent Bolometric Magnitude Scales*, arXiv e-prints, arXiv:1510.06262 arXiv:1510.06262, arXiv: 1510.06262 [astro-ph.SR] (cit. on p. 18).
Massey, P., K. F. Neugent et al. (2021),
Testing Evolutionary Models with Red Supergiant and Wolf-Rayet Populations, **922**, 177 177, arXiv: 2107.08304 [astro-ph.SR] (cit. on p. 51).
Massey, P. and K. A. G. Olsen (2003),
The Evolution of Massive Stars. I. Red Supergiants in the Magellanic Clouds, **126** 2867, arXiv: astro-ph/0309272 [astro-ph] (cit. on p. 51).
McLaughlin, D. E. et al. (2006), *Hubble Space Telescope Proper Motions and Stellar Dynamics in the Core of the Globular Cluster 47 Tucanae*, **166** 249, arXiv: astro-ph/0607597 [astro-ph] (cit. on p. 6).
Mokiem, M. R., A. de Koter, C. J. Evans et al. (2006),
The VLT-FLAMES survey of massive stars: mass loss and rotation of early-type stars in the SMC, **456** 1131, arXiv: astro-ph/0606403 [astro-ph] (cit. on pp. 4, 44–46).
Mokiem, M. R., A. de Koter, J. S. Vink et al. (2007),
The empirical metallicity dependence of the mass-loss rate of O- and early B-type stars, **473** 603, arXiv: 0708.2042 [astro-ph] (cit. on pp. 1, 2).

Ramachandran, V. et al. (2019), *Testing massive star evolution, star formation history, and feedback at low metallicity. Spectroscopic analysis of OB stars in the SMC Wing*, **625**, A104 A104, arXiv: 1903.01762 [astro-ph.SR] (cit. on pp. 4, 44–46).

Rodrigo, C. and E. Solano (2020), "The SVO Filter Profile Service", *XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society* 182 182 (cit. on pp. 5, 14).

Rodrigo, C., E. Solano and A. Bayo (2012), *SVO Filter Profile Service Version 1.0*, IVOA Working Draft 15 October 2012 (cit. on pp. 5, 14).

Sana, H. et al. (2012), *Binary Interaction Dominates the Evolution of Massive Stars*, Science **337** 444, arXiv: 1207.6397 [astro-ph.SR] (cit. on p. 2).

Schootemeijer, A., N. Langer, N. J. Grin and C. Wang (2019), *Constraining mixing in massive stars in the Small Magellanic Cloud*, **625**, A132 A132, arXiv: 1903.10423 [astro-ph.SR] (cit. on pp. 1–3, 14, 15, 29, 53–55).

Schootemeijer, A., N. Langer, D. Lennon et al. (2021), *A dearth of young and bright massive stars in the Small Magellanic Cloud*, **646**, A106 A106, arXiv: 2012.05913 [astro-ph.GA] (cit. on pp. 13, 14, 17, 25, 27, 28, 44).

Schootemeijer, A., D. J. Lennon et al. (2022), *A census of OBe stars in nearby metal-poor dwarf galaxies reveals a high fraction of extreme rotators*, **667**, A100 A100, arXiv: 2209.04943 [astro-ph.GA] (cit. on pp. 10, 12).

Sukhbold, T. and S. E. Woosley (2014), *The Compactness of Presupernova Stellar Cores*, **783**, 10 10, arXiv: 1311.6546 [astro-ph.SR] (cit. on p. 4).

Trundle, C. and D. J. Lennon (2005), *Understanding B-type supergiants in the low metallicity environment of the SMC II*, **434** 677, arXiv: astro-ph/0501228 [astro-ph] (cit. on pp. 44–46).

Trundle, C., D. J. Lennon, J. Puls and P. L. Dufton (2004), *Understanding B-type supergiants in the low metallicity environment of the SMC*, **417** 217, arXiv: astro-ph/0312233 [astro-ph] (cit. on pp. 4, 44–46).

Venn, K. A. (1999), "Massive Stars Spectroscopy in the Magellanic Clouds", *New Views of the Magellanic Clouds*, ed. by Y. .-. Chu, N. Suntzeff, J. Hesser and D. Bohlender, vol. 190 200 (cit. on pp. 4, 44, 46).

Vink, J. S. and A. A. C. Sander (2021), *Metallicity-dependent wind parameter predictions for OB stars*, **504** 2051, arXiv: 2103.12736 [astro-ph.SR] (cit. on p. 14).

Yang, M. et al. (2019), *Evolved massive stars at low-metallicity. I. A source catalog for the Small Magellanic Cloud*, **629**, A91 A91, arXiv: 1907.06717 [astro-ph.SR] (cit. on pp. 5, 6).

List of Figures

List of Tables