MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Creating an empirical isochrone archive“

verfasst von / submitted by

Alena Kristina Rottensteiner BSc BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 861
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Astronomie
Betreut von / Supervisor:	Dr. Stefan Meingast, Bakk. MSc

Acknowledgements

> We are all in the gutter, but some of us are looking at the stars.
>
> *Oscar Wilde*

This thesis represents the conclusion of a project that has accompanied me for the better part of my Master's studies, and during which I have received a lot of advice and encouragement from many people dear to me. First and foremost, I would like to thank my supervisor Stefan Meingast for his dedicated support and counsel, as well as for his willingness and alacrity to discuss the progress of the project at every stage in its development. Thank you for every valuable suggestion and for providing guidance whenever I felt stuck with a particularly persistent problem.

Furthermore, I am extremely grateful for having been given the chance to conduct this project as a student in the research group of João Alves at the Department of Astrophysics, where I have gained firsthand experience of astronomical project work, in a motivated and friendly working environment. I want to thank the whole research group for their inspiration and the pieces of advice they have provided concerning the development of the extraction method.

On the other hand, I want to express my heartfelt gratitude to my family, not only for tirelessly supporting me in pursuing my goals, but also for always showing interest for the topics of my studies, no matter which depths of the universe they revolve around.

Lastly, I want to give special thanks to my partner Marcus, for offering valuable advice for this project, for patiently weathering my many late-night writing sessions, and for always motivating me to keep improving myself and continuing to grow.

Abstract

CONTEXT. Stellar groups, and open clusters in particular, are among the most popular tracers of stellar evolution in the Milky Way galaxy. Studying them provides valuable insights into large-scale structure and kinematics, among many other topics. However, inferring information about structure evolution or cluster formation processes requires knowledge about a crucial, yet elusive physical parameter – the age of a cluster.
AIMS. To date, age estimation of clusters and stars remains a complex process, and its results are often associated with high uncertainties and dependencies on the completeness and correctness of theoretical models. Moreover, over an age range of around $100 - 500$ Myr, age determinations for clusters prove to be extremely difficult and even a blindspot for theoretical isochrones can be observed. With the goal of circumventing the drawbacks associated with model-dependent ages, an age estimation method based exclusively on observational data, which allows the derivation of *relative* cluster ages in the form of empirical isochrones, is developed. By providing such isochrones for a comprehensive, representative, and relevant selection of nearby open clusters, a way of quantifying ages via a relative comparison between individual stellar populations is established. The isochrones are derived from high-quality observational data of *Gaia* DR2 and EDR3, for a selection of 67 clusters within 500 pc.
METHODS. The empirical isochrones are extracted from color-magnitude diagrams (CMD) of the individual clusters, which are based on member selections found in contemporary literature. In a first step, various methods stemming from either the field of image processing or data analysis are tested on selected clusters and compared with respect to their performance, universality, and computational complexity. In addition, procedures regarding the pre-processing of cluster data, which could improve the results or applicability of an isochrone extraction method, are examined. The investigations show that the best method for the purposes of the project is support vector regression (SVR), which allows a thorough tailoring to each cluster via an automated hyperparameter tuning, and its extracted isochrones reproduce the shape of the cluster distribution in the CMD in great detail. Combined with data pre-processing via principal component analysis (PCA), the method can be generalized over a large range of cluster ages.
RESULTS. Empirical isochrones are extracted for each of the 67 archive clusters by creating a large number of SVR-isochrones via bootstrapping of the cluster data and calculating the median values from all resampled curves. By further determining the values of the 5^{th} and 95^{th} percentiles, an uncertainty region around the extracted isochrone is defined. The results are compared to cluster age estimates found in literature and similarities as well as discrepancies are highlighted. Subsequently, the relevance of the archive is discussed, and two case studies on clusters of special interest are conducted.
CONCLUSION. By providing an archive of empirical isochrones for nearby clusters, an age quantification of stellar populations that is independent from stellar evolution modelling is created. The robust method allows for a relative comparison of both known and new stellar conglomerates and yields valuable information regarding the quest for a deeper understanding of the formation and evolution of mid- to large-scale structures in the galaxy.

Kurzfassung

Die Bestimmung von absoluten Sternaltern ist bis heute ein schwieriges Unterfangen, da sie sich im Regelfall nur schwer aus messbaren Parameter ableiten lassen. Aus diesem Grund wird bei der Untersuchung von Sternhaufen oder im Rahmen groß angelegter Altersstudien häufig auf theoretische Modelle zur Schätzung der ungefähren Populationsalter zurückgegriffen. Infolgedessen ist die Qualität der erhaltenen Alterswerte an die Vollständigkeit und Richtigkeit der theoretisch beschriebenen Sternentwicklungsstadien gebunden und systematische Fehler können nicht ausgeschlossen werden. Für Sternhaufen im Alter von etwa $100-500$ Millionen Jahren gestaltet sich die Altersbestimmung oft besonders schwierig, da Modelle wie theoretische Isochronen, also Indikationslinien für Sterne gleichen Alters im Hertzsprung-Russel-Diagramm (HRD) oder Farben-Helligkeits-Diagramm (FHD), über diesen Altersspanne oft sehr ähnlich verlaufen und überdies kaum eindeutige Referenzpunkte für Modelle in den Observationsdaten zu finden sind. Sofern keine weiteren globalen Parameter von Sternhaufen, wie beispielsweise Metallizitäts- oder Extinktionswerte, bekannt sind, kann ein "blinder Fleck" in der Altersbestimmung in dieser Altersregion entstehen, welcher zur Verfälschung der geschätzten Haufenalter beiträgt.

Um eine Modellabhängigkeit und ihre assoziierten Nachteile bei der Altersbestimmung von Sternhaufen zu umgehen, besteht die Zielsetzung der vorliegenden Arbeit darin, ein rein auf Beobachtungsdaten basierendes System zur Altersbestimmung zu entwickeln. Unter Verwendung von Observationsdaten des *Gaia* Satelliten, sowie publizierten Mitgliedslisten von Sternhaufen wird ein Archiv aus 67 solcher Objekte, die sich in naher Umgebung der Sonne ($\leq 500\,\mathrm{pc}$) befinden, erstellt. Für dieses werden anschließend empirische Isochronen ermittelt und so eine relative Art der Altersbestimmung geschaffen.

Alle empirischen Isochronen werden aus den Farben-Helligkeit-Diagrammen der jeweiligen Sternhaufen extrahiert. Für ihre Erzeugung aus dieser auf reinen Beobachtungsdaten basierenden Darstellungsform werden verschiedene Zugangsweisen aus den zwei unterschiedlichen Bereichen der Bildverarbeitung und der Datenanalyse untersucht und an ausgewählten Sternhaufen des Archivs getestet. Konkret handelt es sich hierbei um die drei Methoden *Percentile fitting*, *Principal curves* und *Support vector regression* aus dem Feld der Datenanalyse, sowie die beiden Methoden *Skeletonization* und *Contour evaluation* aus dem Fachbereich der Bildverarbeitung. Alle Methoden werden in Bezug auf die Qualität ihrer Ergebnisse, ihre universellen Anwendbarkeit, sowie ihren numerischen Rechenaufwand miteinander verglichen und bewertet. Zusätzlich werden die Methoden *Extreme deconvolution* und *Principal component analysis* bezüglich ihrer Eignung zur Aufbereitung der Beobachtungsdaten erforscht.

Im Allgemeinen erscheinen die Methoden aus dem Bereich der Datenanalyse besser für die Extraktion von empirischen Isochronen geeignet, als jene aus dem Feld der Bildverarbeitung. Der Algorithmus der *Support vector regression* (SVR) produziert die besten Isochronen, ist sehr gut automatisierbar und kann genau auf jeden Sternhaufen zugeschnitten werden. In Verbindung mit der Datenaufbereitung durch *Principal component analysis* (PCA) ist SVR außerdem auf eine große Altersspanne der Sternhaufen generalisierbar. Unter Verwendung der Methode werden zuerst für jeden Sternhaufen die SVR-Hyperparameter automatisch optimiert. Im Anschluss wird eine große Anzahl an Isochronen mithilfe von wiederholtem “bootstrapping” der Ausgangsdaten berechnet. Bei dieser Technik handelt es sich um ein Ziehen mit Zurücklegen von Wertepaaren aus den ursprünglichen FHD-Daten eines Sternhaufens. Dadurch entsteht eine neue Datengrundlage für die Berechnung der SVR-Isochronen, in der manche Ursprungswerte mehrmals vorkommen, während andere nicht mehr vertreten sind. Aus der Gesamtheit aller Einzelisochronen wird abschließend die endgültige empirische Isochrone durch eine lokale Ermittlung des Medianwerts generiert. Zusätzlich werden das 5. und 95. Perzentil der Isochronen-Sammlung berechnet, um so einen Schwankungsbereich der Isochronenposition zu definieren.

Unter Verwendung der SVR-Methode werden Isochronen für alle 67 Sternhaufen des Archivs generiert und bezüglich ihrer Qualität und Aussagekraft analysiert. Außerdem werden die erhaltenen Ergebnisse mit gängigen Haufenaltern aus der Literatur verglichen, um auf Übereinstimmungen sowie Diskrepanzen zu schließen. Abschließend wird die Relevanz des Archivs diskutiert und zwei Fallbeispiele von Sternhaufen von besonderem Interesse werden vorgestellt.

Das Archiv empirischer Isochronen für nahe Sternhaufen eröffnet die Möglichkeit, eine relative Altersbestimmung von Sternhaufen und Assoziationen durchzuführen, welche nicht auf theoretischen Sternentwicklungsmodellen basiert und deswegen auch nicht mit den gleichen Unsicherheiten behaftet ist. Allein der relative Vergleich zwischen bekannten und neu entdeckten Sternpopulationen könnte wertvolle Erkenntnisse bezüglich der Entstehungsgeschichte der Milchstraße liefern.

Contents

List of Tables

List of Figures

1. Introduction

Although their astonishing, prominent appearance on the night sky is gradually being diminished by the nightly glow of human civilisation, from a scientific perspective there has never been a better time to study stars in the history of humankind. Across millennia, the distant, bright dots have been pondered by people all over the globe, catering inspiration on a philosophical and a spiritual level, and later serving more practical purposes like reference points for maps and calendars, as natural philosophy slowly, but steadily, evolved into science. But nowadays, the observation of stars does not only concern applications regarding spatial and temporal orientations on Earth, but has become even more important now that we know to look beyond our own planet. With the fast-paced rhythm of technological progress, we now possess instruments that are capable of tapping into the deepest, farthest regions of the universe, space probes that have helped us witness events back to the first time that space became transparent to light, and telescopes which can map billions of stars around us, in the attempt to create a stellar census. Instead of viewing stars as unchanging, eternal pinpricks of light in the sky, we can now measure the smallest changes in their appearances and convert this information into insight about their kinematics, positions, and evolutionary states. Never before have we had such a wealth of data on our hands that we can use to understand the distant siblings of our host star.

WHY ARE WE INTERESTED IN STUDYING STARS? Already from a purely conceptional point of view, the study of stars is a fascinating subject. In spite of their vast distances to Earth, which even light can only bridge over the course of multiple years at minimum, a surprising amount of essential stellar properties can be derived directly from observations, provided that at least a part of the stellar spectra can be measured. For example, there exists a clear relationship between the luminosity, mass and radius of a star, while its color and size are correlated to the expected temperature at its surface. If a star's spectrum can be resolved to a degree that allows the distinction of individual absorption lines, one can draw a rough image of its internal chemical composition and structure. Combined with positional information, accessible for instance through the parallactic shift of a celestial object when viewed from a moving observer, larger scale structures and stellar kinematics can be made visible. However, one of the most fundamental stellar parameters eludes quantification through direct observation or simple derivation: Neither luminosity, nor position or any other stellar parameter by itself carries enough information to reliably determine the actual, precise age of a star. In fact, the only concise absolute age that is currently known – that of the sun – is based on age-dating Solar System materials, not on measurements of the star itself (see Soderblom 2010, and references therein).

BUT WHY ARE STELLAR AGES IMPORTANT IN THE FIRST PLACE? Depending on the underlying scientific objective, the answer to this question might vary greatly. From the perspective of stars and stellar populations, ages are the key to performing so-called galaxy archaeology, meaning discovering the formation process of the Milky Way and being able to extrapolate the findings towards the general theory of galaxy formation. Furthermore, the dynamics playing into galactic evolution could be accessed by studying the formation timescales and kinematics of the various stellar populations distributed throughout the galaxy. To establish a holistic view of our home galaxy, from the large-scale structure down to the formation and evolution of stellar agglomerates and individual stars, it is imperative to tie the different available tracers into a time frame. However, estimating ages currently almost exclusively works via one of two approaches: Either employing stellar evolution models, or using empirical methods, which are usually tailored to a specific age group or require extremely precise and elaborate measurements. Seeing as there exists no ideal age indicator, different stellar age studies lead to numerous, often quite contrasting, age estimates for stars and populations. To tackle this issue, this project is centered around the creation of a new, empirical age estimation method for clusters, which is not focused on assigning absolute ages to populations. Instead, it offers a concept that is designed for quantifying ages in a relative manner – an age-based ranking of cluster ages, so to speak.

Why do we look to open clusters when dealing with the ages of stars? Star clusters are natural points of interest for numerous reasons. From a scientific objective, expanding the focus from single sources to stellar populations like clusters, permits the mapping of the conditions and structures across the galaxy. Concerning age investigations, different tools have been established for clusters, which is always fortunate for comparison and error detection in methods, measurements, or underlying physics. Furthermore, clusters are commonly used as age benchmarks for empirical age estimation methods (e.g. Soderblom 2010). From the perspective of the current project, though, clusters are attractive due to three vital characteristics: Firstly, they provide numerous coeval objects that formed under identical, or at least very similar, conditions. Secondly, a large number of young open clusters is located within the Solar Neighborhood, meaning even their low mass members are bright enough to be observed. Linking the two facts, one arrives at the single most important reason for using clusters as objects of interest: A very specific tool can be used to estimate their ages – the Hertzsprung-Russel diagram (HRD).

The body of this thesis is structured as follows: The most important theoretical background on the HRD, stellar evolution, star clusters, and stellar age determination is presented in Chapter 2. Age estimation based on the use of isochrones, as well as its caveats and difficulties, such as the occurrence of an isochrone blindspot, are discussed separately in Chapter 3. The centerpiece of this project, the empirical isochrone archive, is introduced in Chapter 4, with the different prospective extraction methods tested on selected archive clusters being described in Chapter 5. The best-performing isochrone extraction procedure is presented along with the results in Chapter 6. Furthermore, the empirical isochrones are compared to age estimates found in literature, and analyzed with regard to their quality and reliability. The relevance of the isochrone archive is discussed in Chapter 7, and two case studies on the special objects of interest NGC 2516 and Melotte 22 are conducted. A final summary is provided in the concluding Chapter 8. Regarding the graphic presentation of the various findings of this work, all non-referenced plots or images are created with the `python` packages `seaborn` (Waskom 2021) and `matplotlib` (Hunter 2007).

2. Theoretical background

Finding the right start towards the project is difficult, particularly since its key topics, namely AGE DETERMINATION – STELLAR EVOLUTION – HRD – CLUSTERS – ISOCHRONES are so tightly interwoven. For instance, one cannot talk about stellar age without mentioning stellar evolution, and evolution in turn requires the introduction of the most important diagnostic tool for tracing the different phases in the lives of stars – the Hertzsprung-Russel diagram. Applying this important representation method to simple stellar populations, such as clusters, results in the observation that stars with different masses, luminosities, colors and sizes, but with the same age naturally form characteristic lines. Such lines are referred to as isochrones and they are represented among the most common age estimation methods in the form of isochrone fitting. Thus, a full circle from the last to the first of the mentioned topics can be formed.

As the HRD is the starting point for the extraction of the empirical isochrones, the journey into the theoretical foundations of this project also begins with this particular diagram. However, the diagram is really a window into the evolutionary state of a population and allows scientists to explain the reasons behind the shape that an isochrone takes. Therefore, the different regions of the HRD and the evolutionary paths that stars of different masses trace inside the HRD, need to be discussed in combination with the HRD itself. Subsequently, stellar populations and their age estimation methods are placed into focus in this chapter. However, as already emphasized, each topic connects with all the others, regardless of their order in the following paragraphs.

2.1. The HRD and stellar evolution in a nutshell

The fact that astronomers consider the age of a star an important parameter directly gives rise to the implication, that it does not behave completely static, but changes in some way. In other words, the scientific expectation is that with the passing of time, stars *evolve*. What might seem like an intuitive, trivial fact, has really been a groundbreaking insight that took decades in the making. It was not until the late 19th century, that astronomy transitioned from primarily counting and cataloguing stars to mapping actual stellar physics. With the simultaneous rise of spectroscopy, photometry, and photography in the late 1850s, stars could be identified as self-luminous spheres of gaseous materials, instead of solid, possibly habitable rocks, as proposed by William Herschel in 1795 (see Tassoul 2014, and references therein). The presence of different spectroscopic absorption lines in stellar spectra also incited the first extensive sequential classification of stars. The accumulated findings of the preceding decades culminated around 1910 in the creation of the Hertzsprung-Russel diagram. The diagram improved the fundamental understanding of stars in two major ways: Not only did it first illustrate a physical diversity of groups of stars, but astronomers were gradually able to link the observed properties of various stellar objects to different stages in their evolution (de Boer & Seggewiss 2021). Initially, two different populations, the “dwarfs” occupying the so-called main sequence and the “giants” on a separate branch region, were deduced from the diagram by its founders. In 1914, Russel composed the first evolutionary scheme between the two groups that considered subatomic processes as stellar energy source.[1] About a decade later, Robert Trumpler correctly interpreted the different source densities and shapes of the HRDs of clusters as indications of their different respective ages. Lastly, Bengt Strömgren showed in 1932, that the evolutionary path of a star in the HRD would run from the main sequence towards the giant branch, and not like previously thought the other way around (see Tassoul 2014, and references therein).

2.1.1. The Hertzsprung-Russel diagram

Generally speaking, the HRD belongs to the family of phase diagrams, which are used to characterize relationships between different physical parameters. Depending on which stellar properties are used for spanning the parameter space, three different types of the diagram can be distinguished:

- the spectroscopic HRD
- the theoretical HRD
- the color magnitude diagram (CMD)

As can be seen from the axis labels of the three representations in Fig. 2.1, all variations of the HRD only relay the surface properties of stars and do not intrinsically contain physical information about their interiors (de Boer & Seggewiss 2021).

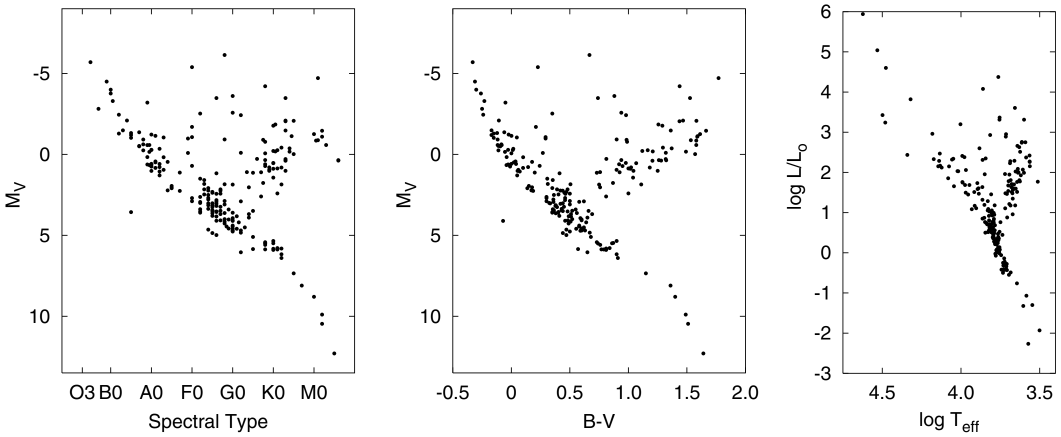

Figure 2.1.: The three types of the Hertzsprung-Russel diagram (HRD). Observational forms of the diagram are either the original spectral type diagram (*left panel*), or the color magnitude diagram (CMD) (*center panel*). The theoretical HRD (*right panel*) is a double-logarithmic plot of stellar luminosity against effective temperature. Image credit: de Boer & Seggewiss (2021).

---

[1] The scheme itself was later shown to be incorrect, as it entailed giant stars evolving into dwarf stars. However, such an evolution would be in direct contrast to the mass-luminosity relation found by Eddington (Tassoul 2014).

The spectroscopic HRD

The original configuration of the HRD is created by plotting the absolute stellar magnitude in the photometric V-band (M_V) against the stellar spectral types of the Harvard classification. In general, a star's absolute magnitude M can be calculated via the **distance modulus**

$$M - m = 5 \log_{10}(d) - 5, \tag{2.1}$$

which is used to calibrate the observed stellar flux, or apparent brightness m, according to its distance d.[2] The spectral type of a star on the other hand, is a categorical variable based on the presence and strength of certain spectral absorption lines, especially the hydrogen lines, in an observed stellar spectrum. Since the appearance of absorption lines is related to the surface temperature of a star, the established spectral sequence is approximately parallel to a temperature sequence (de Boer & Seggewiss 2021; Tassoul 2014). It should however be noted, that the connection between temperature and stellar spectral type is not uniform across all stars. The spectral absorption lines emerge due to the ionization of the gas in the stellar atmosphere. However, following the Saha-equation (Saha 1920, 1921), due to the lower electron pressure in their extended envelopes, the ionization process can begin at lower temperatures for giants and supergiants, than for other stars. Therefore, lines of equal temperature would be slightly diagonal and negatively sloped in the spectroscopic HRD (Voigt et al. 2012).

Since spectral types are discrete in nature and carry a certain freedom in possible definitions of further or refined spectral types, the ordinate axis of the spectroscopic HRD is only limited by the applicable conventions concerning the spectral classification (de Boer & Seggewiss 2021).

The theoretical HRD

The theoretical or physical Hertzsprung-Russel diagram connects stellar luminosity L, usually provided in units of the sun $L_\odot$, and effective temperature T_eff. Unlike the spectroscopic HRD, which lacks strictly defined limits on its axes, the theoretical HRD imposes hard limits for both stellar parameters, which are grounded in the physical extremes at which a star is no longer permitted to exist. Fig. 2.1 shows that the theoretical HRD is plotted narrower than its counterparts. This is because the axis ratio can be chosen in a manner that places stars with equal radii on diagonals across the diagram. The effect exploits the analytical relationships between stellar luminosity, temperature and radius: The radiation emitted from the surface of an ideal black body in thermodynamic equilibrium can be determined by integrating the Planck function over all electromagnetic frequencies ν

$$S = \int B_\nu \, d\nu = \int \frac{2h\nu^3}{c^2} \cdot \frac{1}{e^{h\nu/k_B T}} \, d\nu = \sigma T^4, \tag{2.2}$$

with the Stefan-Boltzmann constant

$$\sigma = \frac{2\pi^5 k_\mathrm{B}^4}{15 h^3 c^2} \cong 5,670 \cdot 10^{-8} \, \frac{\mathrm{W}}{\mathrm{m^2 K^4}}. \tag{2.3}$$

---

[2] The distance modulus in Eq. (2.1) usually also includes an extinction term, which has been omitted because it does not factor into the calculations performed in this work.

The integral form of the Planck equation is also known as the Stefan-Boltzmann equation, where k_B denotes the Boltzmann constant, c represents the speed of light and h indicates Planck's constant. Hence, the luminosity of a black body with radius R and surface $A = 4\pi R^2$ can be written as

$$L \equiv A \cdot S = 4\pi R^2 \cdot \sigma T^4. \tag{2.4}$$

While the spectral energy distribution a star in theory follows the Planck curve, atoms, ions and molecules in the stellar atmosphere often modify the ideal black body spectrum. Amending Eq. (2.3) for a star that does not strictly radiate according to the Planck function can be done by integrating over the distance-corrected intensity distribution $I_0(\lambda) \cdot 4\pi d^2$ of the observed star, with d again denoting its distance. The right-hand side of the resulting expression for the luminosity L

$$L = \int I_0(\lambda) \cdot 4\pi d^2 \; d\nu = 4\pi R^2 \cdot \sigma T_{\text{eff}}^4 \tag{2.5}$$

is almost identical to the result of Eq. (2.2), except for a switch from T to T_{eff}. Using the relationship in Eq. (2.5), one can infer an axis ratio for the theoretical HRD, such that stars of constant radii fall along diagonals passing the diagram at exactly an inclination of 45° (de Boer & Seggewiss 2021).

The color magnitude diagram

Often referred to as the observational counterpart to the physical HRD, the CMD is built entirely from astrometric and photometric observational data. At first glance, it might seem superfluous next to the spectroscopic HRD, since the latter includes almost the same information and usually allows for a finer differentiation between individual stars than photometry. However, to perform spectral classifications, high resolution spectra over a sufficiently large wavelength range are required. Since high-resolution spectroscopy is both more complex and much more time-consuming than photometric measurements, due to the individual refraction of the light of each source, most current astronomic surveys provide only photometric data. Therefore, the CMD has gained much popularity among the scientific community, appealing largely through its extended applicability over all stellar brightness values (e.g. Voigt et al. 2012).

In the CMD, absolute magnitude values are plotted against a color index created out of photometric data of two different passbands, here arbitrarily referred to as X and Y

$$m_X - m_Y = -2.5 \log \frac{I_X}{I_Y}. \tag{2.6}$$

The variables m_i and I_i, where $i \in X, Y$, represent the observed magnitudes and intensities within a passband, respectively. Per convention, X denotes the passband of shorter wavelengths and Y corresponds to larger wavelengths. One benefit of using a color index is, that as a relative measure between two intensities it can be calculated without knowing the distance to the observed objects. Nonetheless, the parameter is sensitive to the effects of interstellar reddening, mainly via dust absorption and scattering effects. If the amount of reddening of the sources is known or the extinction vector can be determined in another manner, the observed brightness values, and consequently the absolute magnitudes (Eq. 2.1), can be corrected. Depending on the passbands used for the color index, the CMD axis span can be restricted by the limiting slope of the Planck function towards the blue part of the spectrum, i.e. the Rayleigh-Jeans limit (de Boer & Seggewiss 2021). For star clusters, where all data points in the CMD roughly lie at the same distance, the absolute magnitudes only deviate from the observed stellar brightness measurements by a constant value (Voigt et al. 2012; Tassoul 2014).

Transformations between different HRD representations

As is evident from Fig. 2.1, there exists a tight connection between the three different representations of the HRD. The transformation between the theoretical diagram and one of the two observational diagrams, consists of two parts: The transformation between color, or spectral type, and temperature, as well as the transformation between absolute magnitudes and stellar luminosity. The former can be performed in a rather straightforward manner, since the relation between color and temperature has been globally established across the HRD. The second transformation however, is more complex and works by first introducing a new variable denoting the integrated brightness of a star in magnitudes – the bolometric magnitude

$$M_{\mathrm{bol}} = -2.5 \log \int_0^\infty I_\lambda \, d\lambda. \tag{2.7}$$

Absolute and bolometric magnitudes are connected via the spectral energy distribution (SED) of a star. Since usually only a part of the stellar SED is measured for an object, the flux not covered by the observation must be taken into account when calculating M_{bol} from an absolute magnitude measurement of a given photometric passband:

$$M_{\mathrm{bol}} \equiv M_V - B.C. \tag{2.8}$$

The bolometric correction factor $B.C.$ strongly depends on the surface temperature of a star due to its connection to the SED and can in theory be calculated for all stellar types (de Boer & Seggewiss 2021). Nonetheless, there are many stellar types for which the bolometric correction remains problematic and unreliable even today (Voigt et al. 2012). The reason behind introducing the bolometric magnitude in the first place, is that in an ideal case, where it can be calculated exactly, it is physically equivalent to the stellar luminosity, differing only by a conversion factor between magnitudes and physical units

$$M_{\mathrm{bol}} = -2.5 \log L. \tag{2.9}$$

Thus, in principle it represents the tool needed for seamless conversions between the observational and theoretical forms of the HRD. Regardless, while the theoretical framework for the transformations has already been solidified, the real-world applications are complicated by the difficulties in correctly determining the values for M_{bol} across all stellar types, and this is disregarding the fact that the latter are not even known for many stars in the first place (e.g. Voigt et al. 2012, and references therein). It should be noted, that transformations between the different representations are usually not even necessary, as their information is essentially the same and stars form typical sequences in all of them.

2.1.2. Following stellar evolution in the HRD

Despite its simple construction, the HRD provides a deep insight into stellar evolution. The HRD of a stellar population always consists of areas of preferred parameter combinations, and of voids, as displayed in Fig. 2.2. But it is not only a tool for distinguishing between sparsely populated and densely crowded areas. Instead, by connecting the different regions to different stages in stellar evolution, the HRD can be viewed as two-dimensional snapshot of the lifetime of stars, from which not only their state of evolution but also their mass can be queried exclusively from the positions they assume in the diagram. During all evolutionary processes of stars, their HRD parameters are subjugated to very distinct boundary conditions. Incidentally, the conditions are different for almost every evolutionary stage, thereby confining stars of a given mass and age to specific, rarely overlapping areas of the parameter space (e.g. Kippenhahn et al. 2012; Tassoul 2014). Due to the specific parameter choices of the HRD and their distinct variety of boundary conditions, conducting an evolutionary study of stellar populations, without explicitly adding a time dimension to the diagram, becomes possible.

By exploiting the natural segregation of the HRD into different regions, a few key characteristics and evolutionary stages can be defined. Among the earliest recognized regions were the **main sequence**, the **giant region** and the **white dwarf** region, which together hold most observed stars, as well as the **supergiant** region. With increasing knowledge about the different processes involved in stellar evolution, more regions were gradually added to the list (Tassoul 2014). In the following paragraphs, the most important stages of stellar evolution and the corresponding HRD locations are presented.

Main sequence

According to Voigt et al. (2012), the main sequence (MS) tracing a mostly diagonal line through the Hertzsprung-Russel diagram (see Fig. 2.2), can be defined as:

> THE LOCATION IN THE HRD AT WHICH STARS WITH A FIXED, HOMOGENEOUS CHEMICAL COMPOSITION ARE IN COMPLETE THERMAL AND MECHANICAL EQUILIBRIUM.

In other words, stars that are found on the main sequence are in hydrostatic equilibrium, and experience one of the longest, most stable evolutionary phases in their lifetimes. They balance their own gravitational pressure with radiation energy generated by the central burning of hydrogen to helium. The process is the same for stars of all masses, however, the absolute duration of their stay on the main sequence increases strongly towards lower masses. It can be calculated via the nuclear timescale, τ_{nuc}, which dictates stellar evolution on the MS and ranges between $\tau_{\mathrm{nuc}} \simeq 4.8 \times 10^6$ years for O5 V stars, which are among the brightest, most massive stellar objects, to $\tau_{\mathrm{nuc}} \simeq 1.2 \times 10^{13}$ years for the faint M5 dwarfs at the other end of the spectrum. Since the initial mass function dictates that low mass stars highly outnumber high-mass stars, generally around 99 % of all stars can be found on the main sequence (Voigt et al. 2012).

From a theoretical perspective, the main sequence can be divided into different subgroups, which are depicted in Fig. 2.3. If a star sustains itself via central hydrogen fusion ($\mathrm{H} \rightarrow \mathrm{He}$) and is chemically largely homogeneous, it lies on the **zero-age main sequence** (ZAMS). It

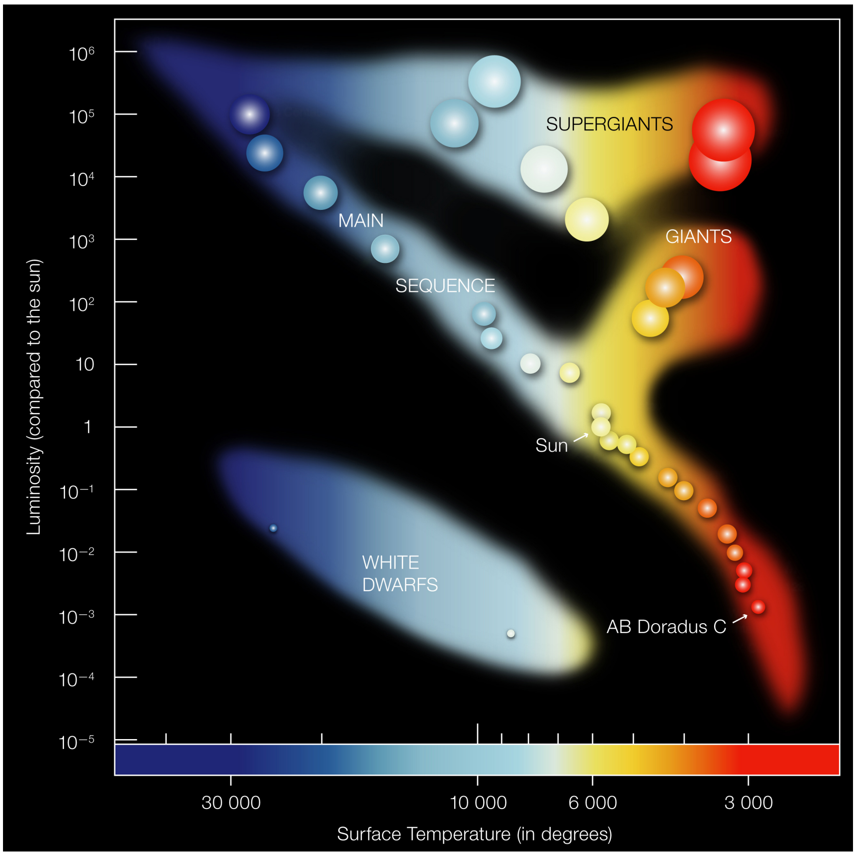

Figure 2.2.: Depiction of the most important regions in the HRD, with schematic colors and sizes of the stars populating them. Very young stars reach the main sequence from the right side of the diagram and start burning hydrogen in their cores. After exhausting the central hydrogen supply, they move through the giant or supergiant region, their precise trajectories depending on their initial masses. In the final evolutionary stages, low mass stars cool down as white dwarfs. High mass stars explode as supernovae and either become neutron stars or black holes, thereby disappearing from the HRD. Image credit: ESO.

represents the earliest, and longest core burning phase, because the object at that point consists almost entirely of the fusion material and, moreover, hydrogen fusion yields the highest energy output of all fusion reactions. Even if it is often not explicitly stated, most literature refers to the ZAMS when using the term "main sequence" (Voigt et al. 2012; Pols 2011).

The **helium main sequence** (He-MS) refers to stars with stable central helium burning. It can only be reached at a much later evolutionary stage than the ZAMS, after both the hydrogen supply in the core and the innermost shells has been depleted and an inert helium core has formed in the center of a star. Only stars massive enough to attain a core temperature of about 10^8 K can start the process of helium fusion in their cores. Even still, lower mass stars would not be found on the He-MS, as their helium burning phase is neither continuous, nor does it happen in hydrostatic equilibrium. The phase of helium core burning is much shorter than the hydrogen fusion phase, since helium only produces about 10 % of the radiative energy that hydrogen fusion provides, despite having the second highest fusion energy yield of all elements (Voigt et al. 2012; Pols 2011). The He-MS is more a theoretical construct than an observational feature, mostly because most stars are no longer homogeneous when they enter their core helium burning phase. To fulfill this criterion, they would have to be perfectly mixed or would have lost their hydrogen envelopes, which is both highly improbable. Nonetheless, the He-MS has useful applications in terms of stellar evolution modelling (Kippenhahn et al. 2012).

If a star is massive enough to ignite its carbon core and would theoretically still be both homogeneous and in an equilibrium state, it would populate the **carbon main sequence** (C-MS). Like the He-MS, the carbon main sequence lacks practical applications and is not usually observed in CMDs. However, the properties derived from modelling it can aid the understanding of more evolved stars (Kippenhahn et al. 2012).

As can be seen in Fig. 2.3, the ZAMS is located at lower stellar temperatures compared to the other two sequences and it also extends to fainter luminosity values. The dashed diagonal lines of equal stellar radii in the figure further indicate that stars on the He-MS and C-MS become increasingly more compact (Voigt et al. 2012; Kippenhahn et al. 2012).

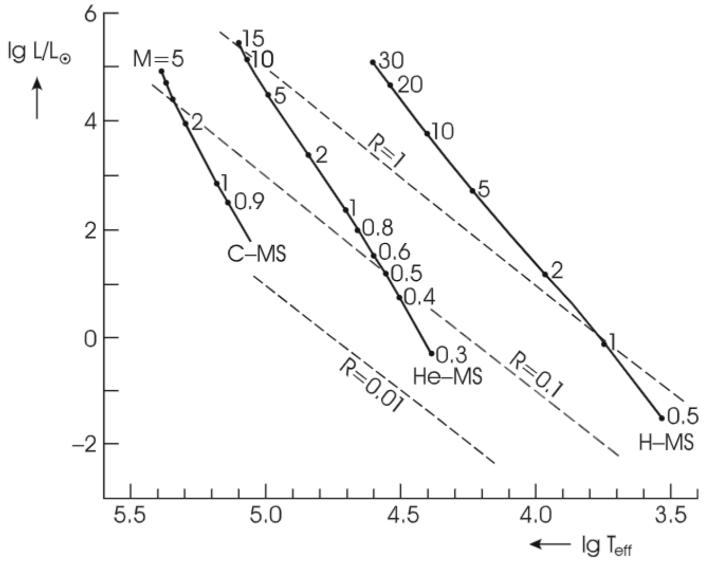

Figure 2.3.: HRD showing the schematic locations of the hydrogen main sequence (H-MS or ZAMS), the helium main sequence (He-MS) and the carbon main sequence (C-MS). The numerical markers along the sequences correspond to different stellar masses in units of $M_\odot$ and the dashed lines indicate lines of constant stellar radii. Image credit: Kippenhahn et al. (2012).

Main sequence turn-off

The main sequence turn-off (MSTO) is a characteristic feature in the HRD which possesses special importance, since it is also a popular age indicator for clusters. The term refers to the evolutionary phase, where stars begin exhausting their core hydrogen supply, having processed approximately 12 % of their initial mass (Voigt et al. 2012). Due to the lowered radiation output of the burning regions, a star in this phase struggles with maintaining an equilibrium state. As a result, gravity starts bearing down on the dense helium core that accumulated during the hydrogen burning process. However, at the same time the stellar envelope begins to expand and grow thinner in an effect that, due to the very contrary behaviour of the interior and exterior regions, is called "mirror principle". Because the star expands, it appears redder and more luminous, which means that in the HRD it moves upwards and to the right of its initial position on the ZAMS. In a cluster, where all stars of similar mass enter this phase roughly simultaneously, one can discern a knee-shaped curve away from the main sequence in the cluster distribution – the MSTO. Since the stellar lifetime on the MS strongly depends on the initial stellar mass, the position of the turn-off point can be used as a very distinct age indicator for clusters and associations, provided they are old enough for their most massive members to have used up most of their core hydrogen. Massive stars are located towards the upper end of the main sequence of a given cluster and finish their central hydrogen burning phase faster than their less massive cluster siblings. This means that for two clusters with comparable extinction values and similar initial chemical compositions but different ages, the MSTO of the older cluster is located at a lower position in the HRD (Soderblom 2010; Kippenhahn et al. 2012; Voigt et al. 2012). The phenomenological evolution of the MSTO with cluster age is shown in Fig. 2.4 on the example of calculated isochrones for eleven clusters, spanning an age range between 4 Myr and 7 Gyr (Lohmann 1957).

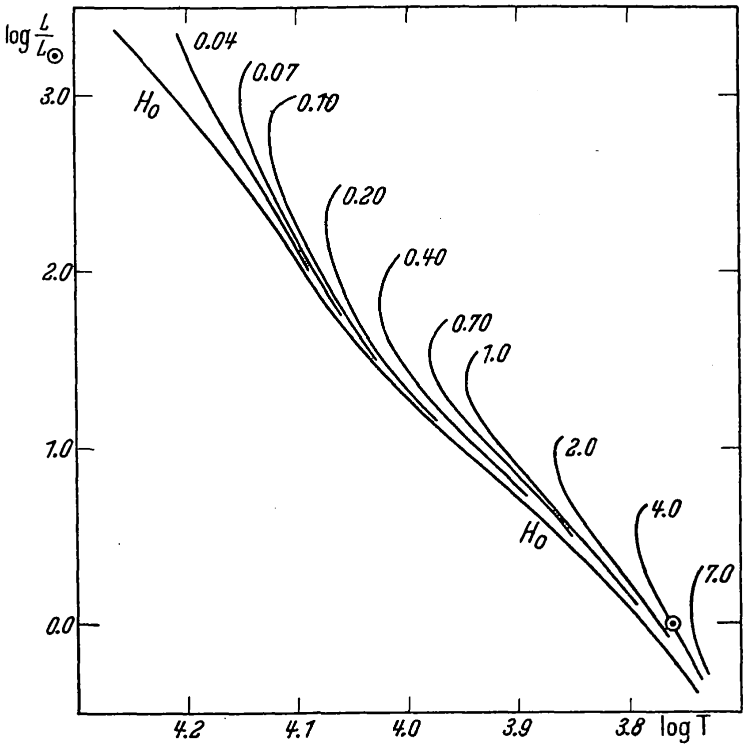

Figure 2.4.: Calculated main sequence turn-off points (MSTO) for clusters between 4 Myr and 7 Gyr. The MSTO continuously shifts downwards on the MS with increasing cluster age. Image credit: Lohmann (1957).

Red giant region

As stars travel across the bend that makes up the MSTO, they begin to fully exhaust their core hydrogen supply. As a result, gravitational forces start compressing the inner stellar regions more vigorously due to the lack of counteracting radiation pressure. The subsequent stages of stellar evolution depend strongly on the respective masses of different stars:

For almost all low mass stars, the hydrogen region directly above the helium core becomes hot and dense enough to provide the necessary conditions for hydrogen fusion – a process called shell hydrogen burning can commence in a continuous manner. For stars with masses $\leq 2\ M_\odot$, the helium core becomes degenerate either already on, or shortly after the main sequence, which means that no rapid core contraction sets in. Thus, for stars in this mass range there exists no Hertzsprung gap. The stars become redder, brighter, and possess very extended envelopes (Kippenhahn et al. 2012). While their colors and surface temperatures approach the same values as those of lower main sequence dwarfs, their radii and luminosities are several magnitudes larger than those of their fainter, smaller siblings (see Fig. 2.2). Following the original classification of the HRD by Hertzsprung, the second prominent arm of the HRD, which is formed by the large and luminous objects, is called the **red giant branch** (RGB) (de Boer & Seggewiss 2021). Low mass stars move upwards along the RGB by increasing their luminosities and radii, while their interior becomes up to more than 70 % convective. During the shell burning phase of a star, mass is added to the core, which heats up until at around 10^8 K the degenerate matter in the stellar interior ignites in an off-center, highly unstable helium flash that terminates the upward evolution along the RGB. After the flash phase, the stellar luminosity decreases significantly and stars move diagonally down and to the left side of the HRD before settling on the **horizontal branch** (Kippenhahn et al. 2012).

For a star of intermediate mass between $2.5\ M_\odot < M < 10\ M_\odot$, the evolution into the red giant region is qualitatively similar to, but differently timed than that of a low mass star, as the helium burning commences before the stellar core becomes degenerate. The hydrogen burning in the shell adds further mass onto the core until the Schönberg-Chandrasekhar limit is reached. At this critical point, the core compresses strongly, while the envelope expands and the stellar luminosity increases, resulting in a rapid transformation of the star into a giant (Fig. 2.2). The considerable speed with which stars in this mass range pass this stage in their evolution, leads to a very sparsely populated area between the main sequence and the red giant region, called the **Hertzsprung gap**. The rapid core contraction and heating leads to the start of central helium fusion, which sustains hydrostatic equilibrium for about 20 % of the duration of the hydrogen burning phase. At this evolutionary stage, most of the total energy output of a star stems from hydrogen fusion in a shell surrounding the core region, as opposed to helium burning. Depending on their respective masses, stars can perform loops or almost vertical upward and downward motions inside the giant region of the HRD during this burning phase. As it lasts rather long, the individual morphologies of the stellar loops can have a significant influence on the appearance of the red giant region of a given cluster CMD. Once the central helium burning ceases and the fusion process is shifted to a burning shell of hydrogen and of helium, respectively, the star moves steeply upward in the HRD, a phenomenon that may again be accompanied by loops. Stars with masses $M \geq 8 - 10\ M_\odot$ usually continue their evolution as **supergiants** (Kippenhahn et al. 2012).

Horizontal branch

The horizontal branch (HB) phase (see Fig. 2.5) follows the helium-flash of low mass stars. The extent of the HB in the HRD of old stellar populations, like globular clusters, provides evidence for the existence of different mass-loss mechanisms during the giant phase, with stars that loose more mass landing on the blue side of the branch (BHB), while those with little mass-loss fall on its red side. Furthermore, it is an indicator for stellar composition, as in more metal-rich clusters the majority of the stars populate the red region, whereas in metal-poor clusters the star concentration is shifted towards the blue region. The stellar concentration parameter is called the first parameter in HB morphology.[3] Despite the name of the region suggesting otherwise, the appearance of the HB in the HRD is never completely horizontal (Kippenhahn et al. 2012). Since 1974, it is known that a region adjacent to the original HB region and populated by O-type and B-type subdwarfs (sdO, sdB) exists. The two parts of the HB are separated by a gap at a surface temperature of roughly 20 000 K, as can be seen in Fig. 2.5. Inspired by its far-left location in the HRD, the extension of the HB was dubbed the **extreme horizontal branch** (Heber 2009).

Asymptotic giant branch

Upon moving away from the horizontal branch, low to intermediate mass stars evolve back towards higher luminosities and lower temperatures, i.e. to the upper right part of the HRD, while again performing loops along their general trajectories. In doing so, they again approach the RGB in an upward slanted arc, called the asymptotic giant branch (AGB). At that point in their evolution, the helium content in their cores has been exhausted as well, but they are not massive enough to reach core temperatures that would allow the ignition of carbon fusion ($M < 8\ M_{\odot}$). Therefore, an AGB star consists of an inert carbon-oxygen core, a helium burning shell and a hydrogen burning shell providing the means for energy production, and an extended hydrogen envelope. On the AGB, two different evolutionary phases can be distinguished. The first one is referred to as **early AGB** and refers to a point at which the hydrogen shell of a star extinguishes after a continued phase of outward burning. The fusion can no longer be maintained because the temperatures in the exterior shells are too low. The helium burning shell on the other hand stays active and burns outward at a much higher temperature than what is needed for hydrogen ignition. Once it approaches the bottom of the hydrogen-rich envelope, hydrogen fusion starts again in a shell around the helium burning layer. However, the two shells are no longer burning in a stable manner, leading to a thermal runaway and cyclic thermal pulses (TP) of the stellar envelope. This process constitutes to the second AGB phase known as **TP-AGB**. While the inner region of the star already resembles a white dwarf, the presence of a thin, extended, hydrogen-rich envelope makes it appear as a red giant. The AGB phase ends with the stellar envelope shrinking rapidly due to strong stellar winds, coupled with mass-loss via shell burning. As a result, an AGB star rapidly moves to the left of the main sequence and down into the white dwarf region, as soon as its shell burning ceases (Kippenhahn et al. 2012).

---

[3] It is speculated that there exists a second parameter, which has yet to be determined (Kippenhahn et al. 2012).

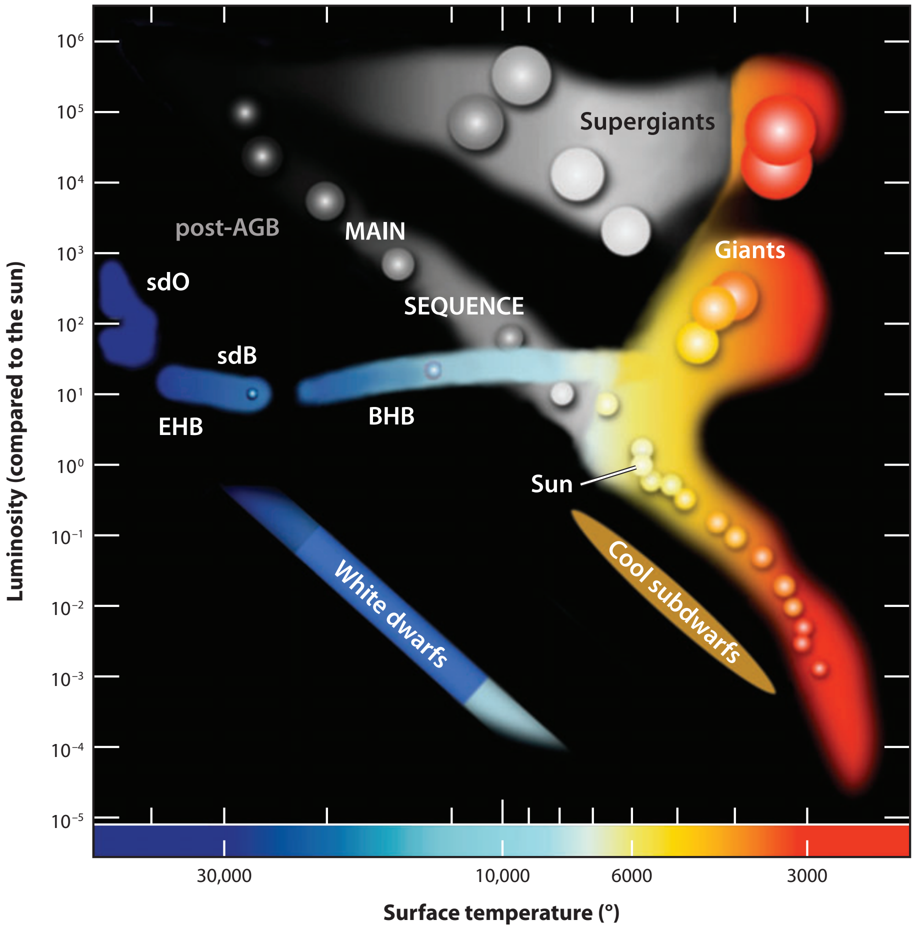

Figure 2.5.: Modified version of Fig. 2.2 with additional information about the horizontal branch and the subdwarf regions. Image credit: (Heber 2009).

Supergiants

In principle, high mass stars ($M \geq 8 - 10\ M_{\odot}$) would evolve in a similar manner to intermediate mass stars from central hydrogen fusion to central helium burning, were they not subjected to the additional effects of semi-convection, overshooting and mass-loss. Unlike a low mass star, which has a radiative core surrounded by a convective outer region, a massive star possesses a convective core. During its evolution on the main sequence, the stellar core then effectively shrinks as a result of the increasing concentration of nuclear energy production towards its very center. Hence, a diversely mixed, semi-convective layer starts to build up around, and encompass the convective star center. It is possible that the two regions connect and the core receives further hydrogen supply for its fusion process. This effect is called overshooting and, together with the effect of mass-loss through strong stellar winds, can lead to an extended MS evolution at higher luminosities. Additionally, the core helium burning can already begin during the evolution of a star towards the red supergiant region, if said effects occur. In terms of the evolution in the HRD, massive stars move almost horizontally from their location on the main sequence towards the red part of the supergiant region, which is indicated in Fig. 2.2, as well as in Fig. 2.5. It should be noted, that the stellar mass required for a star to qualify as supergiant varies between different works in literature, as well as between various stellar evolution models. The masses cited in this work should therefore only be seen as mass estimates. Furthermore, it is also possible that stars evolve upwards through the giant region, but still reach supergiant luminosities during or after the AGB phase via an off-center carbon ignition. Nevertheless, they are usually not considered to be supergiants themselves, because they evolve like AGB stars (Kippenhahn et al. 2012).

Subdwarfs

In the traditional sense, the term subwarf refers to stars that are located slightly to the left of the lower MS in the HRD, with spectral types ranging roughly between F and K. Compared to their main sequence neighbors with equal absolute magnitude values, these **cool subdwarfs** seen in Fig. 2.5 display less metal lines in their spectra, making them appear as “earlier” stellar types. Concerning their evolutionary phase, such low mass stars are also producing energy via central hydrogen fusion, and are simply less enriched in metallicity than their peers on the main sequence. In more recent years, another kind of subdwarfs, referred to as **hot subdwarfs**, has been defined. With spectral types between B and O, and they can be found between the main sequence and the white dwarf region in the HRD, as well as on the aforementioned extreme horizontal branch. Unlike the traditional subdwarfs, they represent very late stages in the evolution of low-mass stars. B-type subdwarfs (sdB) make up the blue end of the horizontal branch and produce energy through core helium burning, while O-type subdwarfs (sdO) include post-RGB, post-HB as well as post-AGB stars. In some cases, central stars of planetary nebulae are also classified as hot subdwarfs (Heber 2009).

White dwarf region

Located in the lower left corner of the HRD in Fig. 2.2 and Fig. 2.5, at very high temperatures but low luminosities, lies a region populated by stars in their final phase of evolution, which are called white dwarfs (WD). Like the MSTO, they can also be used as age indicators for stellar populations, as shown in Koester (2002); von Hippel (2005); Winget & Kepler (2008), for instance. However, age estimation using white dwarfs is based on vastly different physics than those of age indicators referring to other stellar types (Soderblom 2010). Just the appearance of white dwarfs in a cluster population is a valuable clue for cluster age estimations, as it implies that stars of low to intermediate masses, which could not ignite their carbon cores after finishing their helium burning phase, have already existed long enough to reach the last evolutionary stages of their lives. Having previously shed their thin, gaseous envelopes as planetary nebulae, only their inert carbon cores, which are the origin of their name, remain and cool down slowly during a process that lasts several Gyr (Kippenhahn et al. 2012).

2.2. Clusters

Star clusters are among the most popular objects of interest in mid- to large-scale astronomy, so much so, that they have garnered a reputation for being a go-to field laboratory among many astrophysicists. Their unique status among celestial objects is largely owned to the fact, that they conveniently reduce the astronomical parameter space. As all cluster members presumably stem from the same nascent cloud complex, clusters provide a large number of sources that are, at least to a certain extent, representative of the initial mass function (IMF), in one single, spatially and kinematically constrained observation volume. Depending on whether one looks at clusters from an observational or a theoretical point of view, sets of constrained cluster parameters can be grouped as follows:

- Observational: Small dispersions in the radial velocity and proper motion measurements, as well as similar parallaxes.
- Physical: Approximately coeval distances and initial chemical compositions, in addition to a very low age dispersion.

As the IMF dictates, not every star born from the molecular cloud of a cluster has the same mass (e.g Kroupa 2001). Since mass directly connects to stellar evolution, choosing clusters as objects of study allows scientists to study certain physical parameters and conditions, or test theoretical models on a wide variety of stars in different stages of evolution, without having to consider possible influences from the parameters that are identical for each cluster member. With this highly beneficial combination of similar and disparate variables, clusters have become useful tools to verify or reject theories and to accumulate a deeper knowledge of the processes of stellar evolution, the IMF, and stellar structure. Moreover, age, metallicity, extinction, and spatial distribution of clusters permit the inference of possible formation scenarios of galactic structure (Voigt et al. 2012).

There is a second aspect to the importance of star clusters that has become increasingly relevant in recent years and has turned out to be fundamental in the general understanding of star formation history and structure evolution: According to current scientific knowledge, an estimated 90 to 100 per cent of *all* stars are thought to have formed within clusters or associations (e.g. Voigt et al. 2012; Clarke et al. 2000; Lada & Lada 2003; Portegies Zwart et al. 2010; cf. Ward & Kruijssen 2018). This fact on its own is already a major contributor to the motivation for studying these peculiar overdensities of stars.

2.2.1. Cluster types

Historically, clusters have been divided into the two classes of open clusters (OC) and globular clusters (GC), based on characteristics like stellar mass, morphology, and location. Table 2.1 lists some of the key parameters for both cluster types. As can be seen, open clusters by far dominate the overall cluster census. Globular clusters are often very old and therefore inherently faint objects, which are mostly located at such large distances, that parallax measurements can no longer be conducted accurately for most of their members (Voigt et al. 2012; Soderblom 2010). Since the present work is focused on clusters in the Solar Neighborhood[4] ($d \leq 500\,\mathrm{pc}$), and is heavily reliant upon astrometric data, globular

---

[4] The nearest globular cluster to the sun is Messier 4 (Garner 2017), with an estimated distance of about 1720 pc (NASA 2003).

clusters do not fit its criteria and thus lie beyond the scope of this thesis. As a result, the terms "open cluster" and "cluster" are used interchangeably, and always refer to OCs, unless explicitly specified otherwise, throughout the present work.

Table 2.1.: Typical characteristics of open and globular clusters. It should be noted that the values serve only as a general description of the trends, as there are also known clusters with significantly different parameters. References: Voigt et al. (2012); Meingast et al. (2021); de Boer & Seggewiss (2021).

Parameter	Globular clusters	Open clusters
DIAMETER	5-20 pc	several hundred pc
ESTIMATED NUMBER	$\sim 180 - 200$	$> 10^3$ known, $\sim 10^4$ estimated
(MAX) STELLAR MASS	$\approx 10^6\ M_\odot$	$\sim 100\ M_\odot - 10^4\ M_\odot$
SPATIAL DISTRIBUTION	bulge and halo	thin and thick disk
MORPHOLOGY	high spherical symmetry	loose group, extended envelopes
STAR CONCENTRATION	high ($\rho_{*,\mathrm{c}} \sim 10^5\ M_\odot \mathrm{pc}^3$)	low
AGE	up to 10^{10} yr	predominantly young

2.2.2. Open clusters

Open clusters populate the thin and thick disk regions and young OCs are strongly linked to star formation regions within the spiral arms of the galaxy. This observation reinforces the argument that at least a majority of stars are born inside such a stellar cradle. Some clusters are even hidden behind dust clouds for large portions of their lives, and are called embedded clusters. Due to their special locations and multitude, OCs are popular probes of galactic disk parameters (Voigt et al. 2012).

The general cluster characteristics of Tab. 2.1 notwithstanding, the **morphology** of open clusters is not at all uniform and can be strongly influenced by the stellar density, cluster age, as well as by the stellar richness, i.e. the number of member stars. Individual OCs can have similar stellar densities to an average GC, but more frequently they have star concentrations that only slightly vary from the field star concentration, which renders defining cluster members a non-trivial task. The influence of age on cluster morphology can for example be seen in its overall luminosity, which for young clusters is likely dominated by a few bright supergiants, whereas the morphology of an old cluster is much more diffuse and redder (Voigt et al. 2012).

Concerning their **lifetime**, open cluster members are much less gravitationally bound than GC populations due to their generally low star density and total mass content ($\leq 10^4\ M_\odot$). Combined with the fact that open clusters are most commonly found within or near the galactic disk, theories suggest that galactic forces may disrupt OCs on a timescale of about ~ 200 Myr. The findings are supported by the fact that a large fraction of observed open clusters appears to be very young. Since old OCs are an astrophysical rarity and the few known ones in the galaxy are located at much larger distances than their younger counterparts, studying their populations proves rather difficult. Moreover, most age-related quantities of their members, such as rotation, activity or lithium abundance, become weaker with increasing age. Their observations therefore require a higher S/N-spectra and better resolution, the fainter the stars of interest get. Although it would generally be profitable to use the few old open clusters for the calibration of age indicators and global cluster parameters, one should also consider the possibility that they themselves might not actually be representative cluster examples, but more likely represent outliers, lest they would have dissolved on similar timescales as most observed clusters (Soderblom 2010; Voigt et al. 2012).

Most open clusters in the Milky Way have estimated **metallicity** values that are comparable to that of the sun, with only very few objects being very metal-poor (< -0.5 dex). Studies suggest, that there is also no evidence for a pronounced radial metallicity gradient or an age-metallicity relation of OCs within the galaxy. These findings imply, that the galactic disk has not experienced a metallicity enrichment within the last 10 Gyr, and that the determining factor for the metallicity value of an OC is therefore its location within the disk, rather than its age (Voigt et al. 2012).

2.2.3. CMD morphologies of open clusters

The evolution of clusters is best studied with a CMD or HRD. However, the success of creating the latter strongly depends on how complete and reliable their member selections (see Chapter 4) are. In observed cluster CMDs, the MS is often the only discernible feature in the source distribution, because there are simply too few cluster members observed and displayed in the diagram. Regardless, for a well-sampled cluster or in the case of a theoretical HRD, one can observe numerous differences between less and more evolved clusters, as their respective member stars are often at vastly different respective evolutionary stages (Voigt et al. 2012).

Young clusters

In young clusters, the evolutionary stages after a star has left the ZAMS happen very fast, meaning that at any given point in time, there are only very few stars occupying the corresponding late-stage regions in the CMD. Due to the comparatively low number of cluster members with large masses (e.g. Kroupa 2001) and the possibility that the cluster is embedded for the first few million years of its existence, it is possible that the red and blue supergiant regions of such young clusters are entirely invisible in the HRD. Less massive stars evolve on much longer timescales, going through the phases of the horizontal branch, the asymptotic branch, and the thermal pulses, while experiencing mass loss in varying degrees, according to the description in Section 2.1.2 (Voigt et al. 2012).

Old clusters

In older clusters, stars with convective cores on the main sequence move abruptly to higher effective temperatures once the burning of the hydrogen shell sets in. The older the cluster population, the less massive are the stars that are at the MSTO and thus, the more degraded their helium core becomes. The stars move upwards on the RGB and their envelopes expand significantly, until the helium flash sets in. The HB and the AGB are located very near to the RGB in the HRD of such clusters, the latter often being indiscernible from the RGB. Old open clusters also show stars populating the white dwarf region, but one needs to be aware of the risk of contamination of this region in the case of using observational HRDs, for example from quasi-stellar objects (QSOs). Incidentally, observational HRDs of older clusters generally show fewer white dwarfs than theories predict, which implies that either there are unaccounted “hidden” white dwarfs in binary systems, or that the asymmetric mass loss of stars on the AGB may lead to an expulsion of white dwarfs from the cluster (Voigt et al. 2012).

Blue stragglers

A very delicate problem for observers is the appearance of blue stragglers in the CMD of a cluster population. Such stars apparently have higher luminosities and effective temperatures than the stars that make up the MSTO, which suggests that they also have larger masses. However, as has been discussed, a higher stellar mass dictates a faster evolution, therefore the region on the MS above the MSTO should be devoid of stars. At first glance, such objects are an outright contradiction to current stellar evolution theories and point to the presence of outliers in the observational data of a cluster, yet they have been found in most clusters where observers specifically searched for them. A possible explanation for their existence could be a mass transfer between cluster members, either in binary systems or due to collisions of stars in the cluster cores (Voigt et al. 2012).

Blue stragglers are especially problematic for membership determinations or age determinations, as it is very difficult to correctly distinguish between real cluster members and field star contaminants in this situation.

2.3. Stellar age determination

The importance of stellar ages – and especially the ages of populations – has already been established in the introduction of this project. To recap, clusters hold a large part of the key to unlocking the formation timescales and kinematics of the Milky Way galaxy, as they are scattered across the galactic disk and act like probes of the galactic environment (Voigt et al. 2012). They are also popular targets for age determination as, regardless of the estimation method, the age can be based on much more data, compared to the age determination for an individual star. On the one hand, the availability of more data makes the development of methods tailored specifically to stellar populations attractive. On the other hand, methods that apply to single sources can also profit from cluster data, because the calculated individual ages can be verified by comparing the results between different cluster members (Soderblom 2010). Furthermore, with regard to empirical estimation methods, clusters are not only utilized for verification purposes, but also used as calibration benchmarks (e.g. David & Hillenbrand 2015).

But which tools for age estimation are at the disposal of the scientific community? In the subsequent paragraphs, the makings of a good age indicator, as well as the definition of an age-zeropoint for stars, are discussed, before some of the most popular age estimation methods are presented. The statements and information about the various methods are based on the review article of Soderblom (2010), unless referenced otherwise. Due to its special relevance for this project, the popular model-dependent **isochrone fitting** method is described separately in Chapter 3.

2.3.1. Zeropoint definitions

A vital decision with respect to ages, which must already be made before developing an age estimation strategy, is where to define the start of the stellar lifetime. Since the birth point of a star is still ill-defined in literature, every age estimate for a single source or a population is inherently uncertain to some degree. According to stellar evolution theories, a sensible definition of a zeropoint would be the first appearance of hydrostatic equilibrium in a star. However, since it is not yet possible to observe and measure the conditions beneath a star's surface, the time at which this purely physical balance is established is difficult to pinpoint from an observational point of view. Another possibility of a zeropoint definition would be the first emergence of a star onto the ZAMS, since it is well-defined, with respect to theory as well as to observations. Still, it remains a problematic approach because a star does not land on the ZAMS immediately after its formation within a molecular cloud. The time it takes to reach the ZAMS would therefore have to be denoted with negative ages, which are physically meaningless.

As a way to circumvent this impassé, Stahler (1988) proposed the point in time at which a pre-MS star, that supports itself through deuterium burning as a predecessor to hydrogen fusion, is no longer hidden by accretion material, and instead reveals itself to direct observations, as zeropoint. Combining evolutionary track calculations for stars of different masses creates a theoretical visualization of this point in the HRD, which is referred to as "birthline". A second zeropoint proposal was made by Wuchterl & Tscharnuter (2003), who suggest the first formation of a stellar photosphere as $t = 0$. The photosphere condition is defined by a Rosseland mean opacity of $\kappa = 2/3$, and is met about 10^4 years earlier than the event of crossing the birthline. However, for the age estimation methods discussed in the context of this work, both the zeropoint proposals of Stahler (1988) and Wuchterl & Tscharnuter (2003) are valid. As the project is centered around clusters that are estimated to be at least a few Myr old (e.g. Bossini et al. 2019; Dias et al. 2021), the 10^4 Myr zeropoint difference becomes negligible compared to the absolute ages of the populations.

2.3.2. Attributes of useful age indicators

Perhaps the most important point to emphasize when talking about age indicators is the following: According to current scientific knowledge, there is no single, ideal age indicator, that is applicable to star ensembles, individuals, and the entire possible age range all at once. Each of the many physical and chemical processes occurring during stellar evolution, contributes to the complexity of the system or object under observation, and it is highly unlikely that the key to solve it manifests in a single parameter. Moreover, observational data represent the observational perception of the internal stellar processes, which does not always mirror the physical reality. Considering the limited amount of directly accessible stellar properties, Barnes (2007) lists some of the attributes an (observable) age indicator should ideally possess:

- Measurability for single stars
- Sensitivity to age
- Insensitivity to other parameters
- Possibility for calibration, inversion, verification, and error analysis

In other words, an age indicator should present a single, well-defined, easily measurable quantity, that is preferably observable in single stars and insensitive to all stellar variables but the age (i.e. no age-metallicity degeneration). Ideally, it should be susceptible to calibration to eliminate systematic errors, and the relationship between age and the indicator should furthermore work in both directions. In the best case, the physics behind such a relation should also be understood. In the real world however, scaling down these expectations for an age indicator is unavoidable. For instance, the complexity of the involved physics might not allow for the indicator to be separated from other influences, meaning that degeneracies will always play into the equation. Furthermore, the possibility of systematic errors factoring into the values and uncertainties of observed quantities needs to be recognized.

The demands imposed on an age indicator are also strongly tied to the individual objectives one wants to achieve. Sometimes it might be enough to place an upper or lower boundary on the age of an object, while in other instances numerical ages and durations are needed to connect fundamentally different processes. According to the definition by Soderblom (2010), age indicators can be used for:

1. **Classification**, e.g. determining if an object is a T Tauri star from an Li feature.
2. **Limitation**, by defining upper or lower age bounds.
3. **Sorting**, by ordering objects to establish a chronological sequence of events.
4. **Quantification**, by assigning numeric ages or age ranges to objects.

2.3.3. Popular age estimation methods

Age dating methods can be divided into five categories, as displayed in Tab. 2.2, according to their level of quality.

Table 2.2.: Summary of popular age-dating methods, as discussed in Soderblom (2010). The applicability of each method for ensembles ([E], good fit [$\boldsymbol{E}$]), and individuals ([I], good fit [$\boldsymbol{I}$], only at times [i]) is denoted in the right column.

Type	Method	Usefulness
semi-fundamental	nucleocosmochronometry	E, I
	kinematic ages	$\boldsymbol{E}$
Model-dependent	isochrone fitting	$\boldsymbol{E}$, i
	lithium depletion boundary	$\boldsymbol{E}$
	asteroseismology	$\boldsymbol{I}$
Empirical	gyrochronology	$\boldsymbol{E}$, $\boldsymbol{I}$
	decay of activity	E, i
	lithium decline	E
Statistical	age-metallicity relation	i
	galactic disk heating	E

The first category, semi-fundamental ages, refers to methods for which the underlying physics are largely understood and only very few assumptions are necessary. It would only be preceded by fundamental ages, where only observations and no assumptions are required, and the physics of the processes are understood in their entirety. Since no methods in this category have been established yet,[5] it does not appear in Tab. 2.2. A large portion of age estimation methods belong to the class of MODEL-DEPENDENT ages. They have predictive power, but can be biased by their built-in dependencies. EMPIRICAL ages on the other hand, are established by observing relations between different properties, have no predictive power because the underlying physics are not fully understood, and need to be calibrated against age benchmarks, such as clusters. The calibrations are most often based on ages derived by model-dependent methods, which means empirical ages inherit the uncertainties associated with model-dependent age estimations. The last category, STATISTICAL ages, contains two statistical correlations that have been discussed in literature. While each of the methods listed in Tab. 2.2 is a valuable contribution to the science of age determination, the focus of the current work lies with star clusters and not with individual sources. For this reason, only the methods that are relevant with respect to star ensembles (***E***) are described in the following paragraphs. For further details regarding the other techniques, readers are referred to the work of Soderblom (2010); Soderblom et al. (2014).

Lithium depletion boundary

Using lithium depletion in clusters to determine their ages is a model-dependent approach, but the underlying physics of the depletion process are fairly well understood. Originally proposed as a division between brown dwarfs (BD) and very low mass stars, the lithium depletion boundary (LDB) was first proposed by Rebolo et al. (1992) and is based on the core temperature of stellar objects, which can be calculated in a straightforward manner. Objects with $M > 0.06\ M_\odot$ eventually reach temperatures larger than 2.5×10^6 K, above which the element lithium (^{7}Li) starts to be fused in proton capture reactions (e.g. Basri et al. 1996). Pre-MS stars arrive at this temperature on a timescale that depends on their initial mass and, once the lithium burning has started, the element is rapidly depleted throughout the entire star. As a result, a clear, age-dependent cut – the lithium depletion boundary – between stars retaining their initial Li and those that have already depleted it, can be observed. The benefits of the method include the well understood physics, the weak dependence of the ages on stellar composition and measurement uncertainties, as well as the minimal amount of analysis involved in the LDB observations: The Li feature in stellar spectra is either clearly present, or absent. However, the method is limited to an age range between $\sim\ 20 - 200$ Myr due to both physical and observational constraints. The upper limit arises as the LDB is measured in stars of very low mass, which are inherently very faint objects. The observation of the Li I feature ($\lambda = 6708$ Å) requires a high spectral resolution $R \geq 3000$ and signal-to-noise ratio (S/N), respectively, and the member selection of the observed cluster stars needs to be robust against contamination by old and faint field stars. The faintness of the LDB, with $L_\mathrm{bol}/L_\odot \simeq 5 \times 10^{-4}$ at $t \simeq 200$ Myr, ultimately imposes an upper age limit on the applicability of the method. It also restricts the use of the LDB method to very near clusters, for which faint stars can be detected. The lower age limit stems from a combination of model predictions about when lithium depletion sets in, and large age differences between various LDB calculations for models younger than $\simeq 20$ Myr (Soderblom et al. 2014).

---

[5] The only fundamental age Soderblom (2010) lists is that of the sun, determined via age-dating Solar System material.

Kinematic ages

Kinematic ages are an umbrella term for a few semi-fundamental methods, which are all based on studying the motions of very young, unbound associations or groups of stars. With “expansion ages” one uses the observed stellar motions to calculate the expansion rate of stellar groups, while with “traceback ages” the orbits of stellar ensembles are simulated back towards a common physical origin. A third method, called “flyby ages” works by determining the time of the smallest separation between different stellar groups, or between groups and individual sources. Kinematic ages have the advantage that they do not depend on stellar evolution models and the difficulties associated with them. Additionally, they provide age estimates for groups and clusters younger than 20 Myr, which are not covered by other estimation methods like LDB (Soderblom et al. 2014). Moreover, the kinematic study of stars only requires a single assumption, which concerns the shape of the galactic potential, and is valid for all stellar masses. However, a drawback exists with respect to the practical application of the method, as high-quality kinematic data about the three-dimensional space motion, radial velocities, and parallaxes of the stars are required for the calculations to succeed. Due to the dissolution and orbital changes of clusters on a timescale of a few hundred Myr (e.g. Voigt et al. 2012), the methods are only applicable to very young groups of stars. In practise, the upper age limit of associations that can be estimated using kinematic ages lies at roughly 30 Myr, as beyond that age the uncertainties in the calculated kinematic evolution become too large. Furthermore, after a cluster or association lifetime of around 50 Myr tidal disruptions would need to be considered as well.

Gyrochronology

Gyrochronology is an empirical approach to estimating stellar ages that involves measuring stellar spin-down. The method is based on observing rotation periods of stars and using the well-founded relation that older solar-type stars lose part of their angular momentum over time. Gyrochronology benefits from the fact, that rotation periods are generally easy to measure and not apt to large variability. On the downside, the method is limited to sun-like stars and is only applicable to objects as old as, or younger than the sun as well. What is more, rotational period distributions measured for different clusters with similar ages have been shown to not be identical, implying that the calibration of the method might still be incomplete, or that further, as of yet unknown parameters influence stellar rotation periods.

3. Isochrones and their importance for age determination

Throughout the previous chapter, the thread connecting the most important aspects of stellar evolution, open clusters, and the field of stellar age determination has begun to unravel. During this process, the focus of this work has been placed on putting the roles of these topics in uncovering the formation and evolution of larger structures into perspective. By building on the established theoretical foundations, it now becomes possible to discuss the age determination method of isochrone fitting. Besides being among the most popular and widely used tools for cluster age estimation, the procedure is most closely related to the topic of this project – the empirical isochrone archive. After a brief definition of the general concept behind an isochrone, model isochrones and their applications for cluster age estimations are presented. At the same time, some dependencies and problems that can occur when relying on them are highlighted. The chapter concludes with the realization, that over a significant age range spanning multiple hundred Myr, astronomers are potentially age-blind over most of the dynamical range of a cluster CMD, if only model isochrones are consulted as a means of age determination.

3.1. Definition

The term "isochrone" is derived from a compound of two Greek words that, in a literal translation, mean "same" (gr. *isos*) and "time" (gr. *khronos*). In the context of stellar astrophysics, an isochrone may be defined as

> ... THE LINE CONNECTING POINTS OF EQUAL AGES OF THE STELLAR EVOLUTIONARY PATHS IN THE HRD OR THE CMD. – Voigt et al. (2012)

> ... THE LOCUS OF STELLAR MODELS OF IDENTICAL AGE, BUT DIFFERENT MASS IN THE HERTZSPRUNG-RUSSEL DIAGRAM. – Kippenhahn et al. (2012)

This means that for clusters, their distribution of sources in the CMD would already *be* their isochrones, if they were not influenced by the effects of measurement errors, interstellar extinction, as well as by uncertainties or incompleteness in their member selections (Voigt et al. 2012).

3.1.1. Difference to evolutionary tracks

Isochrones have to be distinguished from a very similar tool in stellar evolution modelling, which is known as evolutionary tracks. Both terms refer to curves that originate in, and apply to the HRD, are profoundly shaped by stellar evolution, and are connected to stellar ages. However, in contrast to an isochrone, an evolutionary track describes only the trajectory of an *individual* star, with a given initial mass, through the HRD. By computing

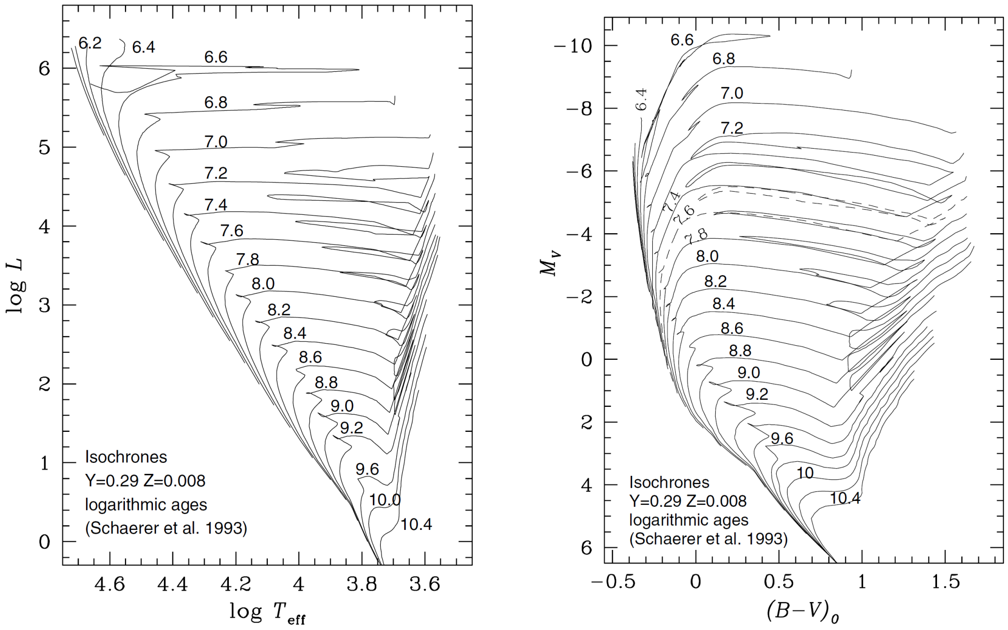

Figure 3.1.: Model isochrones for different ages (2.5 Myr to 25 Gyr), shown in the theoretical HRD (*left*), and the CMD (*right*). The isochrones are calculated for fixed stellar compositions (X, Y, Z), and the values next to the curves correspond to the logarithmic cluster ages. Image credit: Schaerer et al. (1993).

evolutionary tracks for many stars with different masses and extracting their positions at the same point in time t, a **theoretical isochrone** corresponding to the age t is created. A compilation of such calculated curves for different ages is shown in Fig. 3.1 for a theoretical HRD and a CMD representation. Due to their close connection, the two curves can often look very similar (de Boer & Seggewiss 2021).

3.1.2. Characteristics of an isochrone

The shape of an isochrone can be influenced by two main factors: age and metallicity. Regarding the former, the entire curve is basically dictated by the evolutionary processes described in Section 2.1.2, which occur simultaneously for increasing star masses. This means for instance, that the MSTO continually shifts down along the MS, the older a population grows. In addition, later evolutionary stages, like the RGB or the AGB, become more densely populated and distinct in the CMDs of older clusters. The effects of the metallicity parameter on the shape of an isochrone can be seen for instance in the regions of the MS and the giant branch, which appear redder in a more metal-poor cluster CMD. Additionally, the horizontal branch becomes less distinct, and its stars are more densely concentrated towards its red part, as compared to more metal-rich clusters (Voigt et al. 2012). Since metallicity affects the opacity of the stellar gas, metal-poor massive stars are observed to evolve with blue loops and reach higher effective temperatures than metal rich stars (de Boer & Seggewiss 2021).

3.1.3. Transformation of isochrones

Similar to the transformation between different representations of the HRD (Section 2.1.1), isochrones can be transformed from the (L, T_{eff})-space to the (M_X, $m_x - m_Y$)-space and vice versa. The process does however require a calibration with data from well-known stars with reliable parameters such as distance, radius, metallicity as well as broad SED coverage. Still, transformations often remain imperfect to a degree because the important role of stellar atmospheres cannot be comprehensively reproduced in the transformation process in many cases (de Boer & Seggewiss 2021).

3.2. Isochrone fitting with model isochrones

Since their first digitized mentions (e.g. Barbaro et al. 1969, as *isochronous curves*), theoretical or model isochrones have gained widespread popularity as a method of age determination for stellar populations. The workflow of the process starts with the calculation of initial theoretical models, before varying properties like metallicity, extinction and age, oftentimes outlining a parameter grid. Subsequently, the calculated collection of model isochrones is compared to the observational data, using either a form of visual evaluation or regression tools. The chosen candidate from the isochrone grid can either serve as an initial point for a finer parameter tuning, or it is chosen as final result and its corresponding age is assigned to the investigated cluster.

Compared to the age determination methods discussed in Section 2.3, the pre-eminent advantage of the method is its use of the whole cluster HRD, including key features like the MSTO, which depend on very specific stellar processes. On behalf of this fact, isochrone fitting also does not discriminate against any stellar mass range and is applicable to all age ranges, provided the cluster members are visible in a CMD. A direct consequence of using the HRD as age estimation tool, is that it simultaneously works as a test for the quality of the employed stellar evolution models and assumptions that ultimately produce a model isochrone. A good, reliable fit of the model to the observational data can only be achieved if the evolutionary processes are well represented across all stellar mass ranges (Soderblom 2010). Another compelling benefit of model isochrones that has aided their widespread use within the scientific community is grounded in a user's perspective: Their calculation is facilitated by a large number of open-source codes, with PARSEC[6] (Bressan et al. 2012; Marigo et al. 2017), MIST[7] (Choi et al. 2016; Dotter 2016), and BaSTI[8] (Hidalgo et al. 2018; Pietrinferni et al. 2021; Salaris et al. 2021), being some of the most popular implementations. These readily available, user-friendly stellar evolution codes offer quick access to model isochrones, which are highly amendable, but also yield good results with the provided default parameters.

As isochrone fitting is a model-dependent method (see Tab. 2.2), important quality considerations need to be made to ensure their reliability. A non-exhaustive list includes assumptions about: chemical composition (X, Y, Z), opacity functions, nuclear fusion reactions, energy transport processes (e.g. convection), stellar atmosphere physics, and

---

[6] http://stev.oapd.inaf.it/cgi-bin/cmd

[7] http://waps.cfa.harvard.edu/MIST/index.html

[8] http://basti.oa-teramo.inaf.it/index.html

stellar structure (e.g. connection between rotation and color). All these parameters factor into the calculation and the success of the method hinges on how accurate they can be portrayed in the stellar evolution models. If the completeness or accuracy of the models is lacking, isochrone fitting will produce systematic errors in the age estimates. From the observational side, further error sources contribute to the estimation in the form of accuracy and completeness of the photometry, field star contamination, as well as influences of metallicity, distance, and extinction (Soderblom 2010; de Boer & Seggewiss 2021).

3.3. Prospective difficulties with model isochrones

Apart from performance or quality concerns regarding stellar evolution models or observational data, isochrone fitting also carries the potential for pitfalls in terms of practical applicability. The following paragraphs address two very specific sources of difficulties and illustrate each on a scientific example.

3.3.1. Self-dependency of ages based on isochrone fitting

Table 3.1.: Input parameters for the PARSEC model isochrones displayed in Fig. 3.2.

Name	$\log_{10}$(age)	Extinction A_V	[Fe/H]	Ref.
Orion group 1	6.954	0.2	-	Swiggum et al. (2021)
IC 2391	7.561	0.093	0.000	Bossini et al. (2019)
Hyades	8.922	0.034	0.149	Dias et al. (2021)

When estimating cluster ages via isochrone fitting, the derived ages are somewhat contingent on how old the clusters themselves are. More precisely, it often depends on the cluster age, which CMD features are used as principal orientation points for the fitting routine. For instance, very young clusters usually have ages determined from their massive stars at the MSTO (Soderblom 2010). However, the initial mass function for clusters is strongly biased against the number of high-mass stars (e.g. Kroupa 2001), meaning there are generally very few stars in young clusters that could define the MSTO in the first place. Moreover, instruments like *Gaia* do not observe very bright stars (Gaia Collaboration et al. 2016), which means they are often not included in Cluster CMDs and can therefore not be used as guide points for the fitting process. For intermediate-age clusters, the positions of intermediate-mass stars, which have different interior physics than massive bright stars, are often a decisive factor in isochrone fitting, whereas ages of old clusters often depend on the locations of evolved stars in the HRD (Soderblom 2010).

To illustrate the age self-dependency on an example, member selections of three stellar groups are shown in the left panel of Fig. 3.2. The color gradient in the selections stems from performing a Gaussian kernel density estimation for the CMD data and using a power-norm colormap to indicate the density of the various regions. A set of PARSEC isochrones, corresponding to the cluster parameters listed in Tab. 3.1, are displayed in the right panel of the figure. The oldest population, drawn in violet, is known as the approximately $\sim 625-650$ Myr old Hyades cluster (Perryman et al. 1998; Lodieu 2020; Dias et al. 2021), and the green data points denote the $\sim 36-49$ Myr old IC 2391 cluster (Bossini

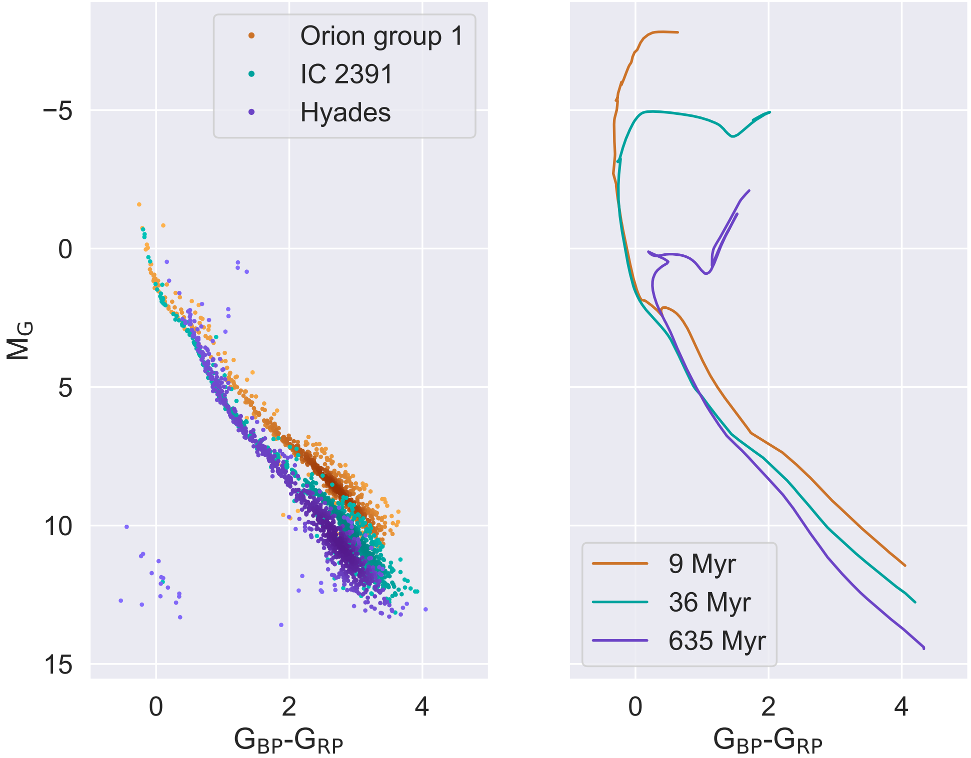

Figure 3.2.: *Left*: Cluster CMDs of the Hyades cluster, IC 2391 and Orion group 1, based on EDR3 data of the member selections by Meingast (priv. comm.), and Chen et al. (2020); Meingast et al. (2021). *Right*: PARSEC isochrones corresponding to the approximate cluster ages found in literature (Bossini et al. 2019; Perryman et al. 1998), assuming a solar-like metallicity Z.

et al. 2019; Dias et al. 2021). The orange-colored population is a subgroup of the Orion OB association, which is estimated to be about $\sim 7-10$ Myr old (Swiggum et al. 2021; Briceño et al. 2007). The member selections for the two clusters are from Meingast et al. (2021) and Meingast (priv. comm.), while the Orion group 1 selection is based on Chen et al. (2020).

As is visible in the cluster distributions in the left panel of the figure, the MSTO cannot be used to fit an isochrone to IC 2391 or Orion group 1 since it is not sampled in the member selections of the two populations. The plot further shows, that the upper main sequence region around 2 to -1 mag is almost identical for the two younger populations. However, the youngest group has a flatter MS shape around 2 mag, which is mirrored by a bend in the corresponding isochrone in the right panel of the figure. The initially prominent shift and bend of the 9 Myr isochrone becomes less and less noticeable with rising cluster age and already for the 36 Myr isochrone of IC 2391, it is no longer visible. In younger populations, the region around the bend and the lower MS would probably be used as a reference for judging the quality of the fit of a model isochrone. In contrast, the Hyades cluster has a very different upper main sequence region compared to the other two populations, which implies that its age can already be estimated by looking at the positions of the MSTO and evolved intermediate mass stars.

The fact that with isochrone fitting, the age of a cluster is often determined by a specific, age-dependent indicator in the CMD on its own might not be a source for errors. However, it can become a problem for clusters in which no "classical" reference points are readily available. Here, the choice of which features of a cluster population are used for the fitting process, can add to the uncertainty and ambiguity of the age estimate, as can be seen in Fig. 3.3.

3.3.2. Ambiguity of cluster ages

Another problem of isochrone fitting is that age estimates for identical populations are not always consistent throughout the literature, even though the various works essentially employ the same method for their age determinations. Deviations between the results can trace back as far as regarding the choice of the stellar evolution code for the calculation of isochrones, as for instance shown in Bouy et al. (2015), but even when using the same code, the derived cluster ages are often ambiguous.

Fig. 3.3 illustrates such an example of age deviations on the account of estimates found in literature for the nearby, well-studied Pleiades cluster, also referred to as Melotte 22. Specifically, the age estimate of the publication Bossini et al. (2019), who place the age of the cluster at around ~ 86 Myr, is compared to the estimate of Dias et al. (2021), who propose a considerably higher age of ~ 131 Myr. It should be especially emphasized that in both works the isochrone grids used for fitting the cluster data are calculated with the PARSEC (Bressan et al. 2012; Marigo et al. 2017) stellar evolution code. Moreover, the age estimates are both based on *Gaia* DR2 observations of the cluster data. However, the former calculations are performed with the *Gaia* passband correction proposed in Evans et al. (2018), while the latter work uses the correction published by Maíz Apellániz & Weiler (2018). Regarding the respective member selections for the cluster, it can be assumed that there does exist an overlap or even an equality between the two publications, as the membership catalogue of Cantat-Gaudin & Anders (2020) is listed as a source for both works. However, Dias et al. (2021) includes multiple membership catalogues as input for their overall cluster selection (for details see references in Dias et al. 2021), therefore it cannot simply be assumed that both results are based on identical cluster data.

In Fig. 3.3, the member selection ($p \geq 0.5$) of the Cantat-Gaudin & Anders (2020) catalogue for the Melotte 22 cluster is displayed. In contrast to most other plots shown in this work, the CMD in this figure shows *Gaia* DR2 data, allowing for proper comparisons between the results of the two age estimates and the cluster data. The model isochrones shown in the plots are reproduced in accordance with the published cluster parameters $\log_{10}(\text{age})$, extinction A_V, and metallicity in units of [Fe/H], which are listed in Table 3.1. The isochrone calculation also takes into account the passbands and corrections of the respective publications, and the PARSEC code (Bressan et al. 2012; Marigo et al. 2017) is used as well. For the isochrone calculated with the corrections of Maíz Apellániz & Weiler (2018), the $\mathrm{G_{BP}}$ band for the faint magnitude range is plotted, although it should be noted that the differences to the bright $\mathrm{G_{BP}}$ band are almost imperceptible.

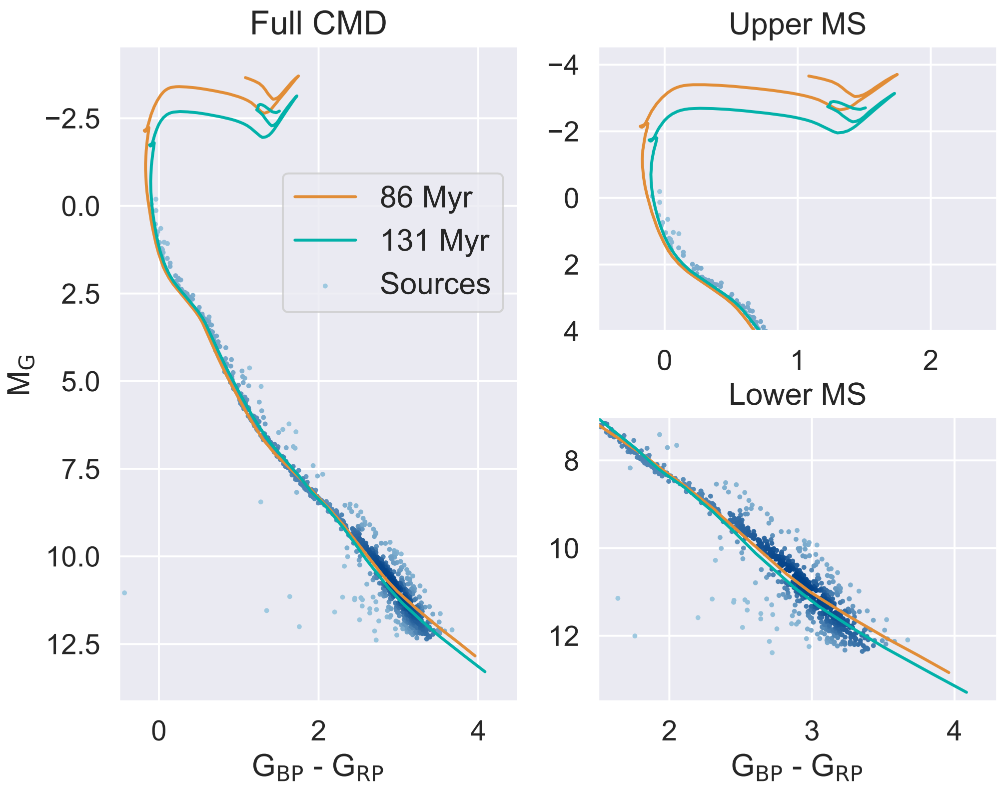

Figure 3.3.: Comparison of two different age estimates for the Melotte 22 cluster, based on isochrone fitting. The 86 Myr isochrone (*orange line*) corresponds to the age estimate of Bossini et al. (2019), whereas the 131 Myr isochrone (*green line*) corresponds to the parameters published in Dias et al. (2021). The data points represent the published Cantat-Gaudin & Anders (2020) selection ($p \geq 0.5$) for the cluster and the isochrones are produced with the PARSEC code (Bressan et al. 2012; Marigo et al. 2017).

As can be seen in the figure, both isochrones provide a very good fit to the cluster data, despite the large difference in their associated cluster ages. For the most part of the main sequences, the curves are almost identical, in spite of their different parameters (see Tab. 3.1). A magnification of the upper and lower ends of the main sequence, as depicted in the right column of the figure, only indicates slight differences between the two curves. In the plots, the isochrone of Dias et al. (2021) seems to be a bit closer to the depicted sources of the lower MS, than the one of Bossini et al. (2019), and vice versa in the upper MS part. Significant differences in the theoretical isochrones themselves become apparent only for magnitudes brighter than about -1 mag, where no sources are located within the shown selection and thus no information for isochrone fitting is available.

There are many factors that play into the deviating age values, the possible use of different member selections and deviating assumptions concerning the *Gaia* passband corrections

Table 3.2.: Published cluster ages, extinction, and metallicity values for Melotte 22 for the two isochrones shown in Fig. 3.3.

Publication	$\log_{10}(\text{age})$	Extinction A_V	[Fe/H]
Bossini et al. (2019)	7.937	0.140	0.000
Dias et al. (2021)	8.116	0.168	0.032

having already been mentioned. Additionally, Dias et al. (2021) vary the metallicity during their fitting procedure as well, while Bossini et al. (2019) instead rely on literature values and assume [Fe/H] = 0 for clusters with no prior metallicity information available. Lastly, it appears as if the two isochrones may have been primarily judged at different reference points concerning the best fit, which would explain why one is a closer fit to the upper MS and vice versa for the lower MS. The interplay between all these factors leads to the result, that although using the *same* age determination method, the independently proposed ages for the cluster disagree with one another by almost 50 Myr. With respect to the absolute cluster age, this difference makes up an estimated 33 – 51 per cent. The fact that age estimates deviate this much for one of the nearest, best studied open clusters, which is moreover a popular benchmark cluster for calibrating other age estimation methods and relevant for the cutting-edge scientific objective of searching for exoplanets (e.g. David & Hillenbrand 2015), illustrates just how much uncertainty of the results can remain when using isochrone fitting. The procedure is so very sensitive to its input parameters and the quality of the observational data, that even quantifying the age of a popular, nearby object is only possible with deviations up to 50 % of the proposed cluster age itself.

3.4. Isochrone blindspot

As has been shown in Section 3.3, isochrone fitting is neither a trivial, nor a straightforward process, and it hinges significantly on the different features that divide young and old clusters. But while the differences are readily visible for comparisons between very young and old clusters (see Fig. 3.2), the situation changes drastically for isochrones that denote ages somewhere in between the two extremes. The plot in Fig. 3.4 again shows a range of PARSEC isochrones, calculated with default metallicity and extinction values, this time encompassing an age range between 50 and 700 Myr. The strong differences between very young and very old isochrones (*dotted lines*) disappear almost completely over a dynamical range of 12 magnitudes (*dashed horizontal lines*) and an age range between $\sim 100-500$ Myr. This means, that from the point at which the low mass stars of a cluster are no longer separated from populations of older clusters by their shifted, flatter MS, until the time at which the MSTO has shifted to lower stellar masses and has become more densely populated, there exist virtually no decisive features in the shapes of theoretical isochrones over a large dynamical range – which results in the appearance of an “isochrone blindspot”. Additional complications arise from the issue, that the lower MS of a cluster cannot be sampled as easily or accurately than the brighter upper MS regions, as lower mass stars are much fainter and thus harder to observe. For clusters falling within the age range of the isochrone blindspot, bright member stars are often scarce, and possibly cannot even be captured by *Gaia*, whose brightest observed star to date is reported at G = 1.73 mag (Gaia Collaboration et al. 2021). At the same time, the intermediate and low mass cluster

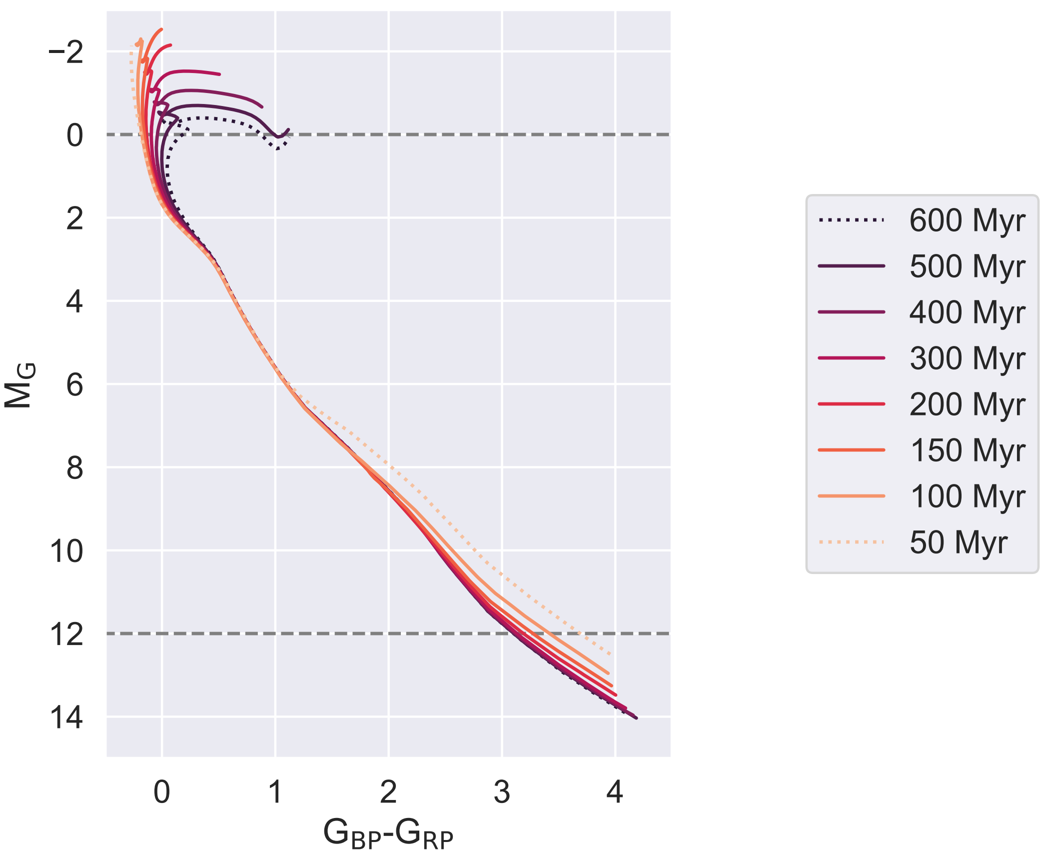

Figure 3.4.: The isochrone blindspot refers to a largely featureless region of model isochrones, covering an age span of approximately 100 – 500 Myr and about 12 – 0 absolute magnitudes of the *Gaia* G band. The plot features PARSEC isochrones (Bressan et al. 2012; Marigo et al. 2017) with default parameters.

stars have not yet reached the MSTO. The differences between the model isochrones in Fig. 3.4 in the MS region are vanishingly small, even without considering the effects of a possible age-metallicity degeneracy, making age determinations from this region virtually impossible without further information. What is more, over this particular age range, a partial blindspot exists not only with respect to theoretical isochrones, but also considering other age determination methods for stellar populations. As has been established in Section 2.3.3, gyrochronology only works for single, solar-type stars, whereas LDB or kinematic modelling apply only to young groups of ≤ 200 Myr and ≤ 30 Myr, respectively.

4. The empirical isochrone archive for open clusters

The theoretical background, as well as the discussion regarding the difficulties and limitations of age determination tools presented in the previous section serve to stress a single fact: To this day, absolute age determinations, no matter how sophisticated the method, remain a serious challenge. Even for well-studied, local clusters for which high-quality observational data are available, the process is often arduous and characterized by the need for assumptions and numerical models. Moreover, uncertainties or gaps in current stellar evolution theories, as well as measurement errors in observational data, which are still prevalent for the numerically most abundant low mass, faint dwarfs, preclude the determination of absolute ages with universal, unbiased and academically undisputed methods (see e.g. Soderblom 2010, and references therein). But while the present state of technology and the level of physical knowledge might not quite be at a stage that permits *quantifying* cluster ages en masse, one can refer to the different approaches to age determination (see Section 2.3) and try the next best option – *sorting* them. The goal of the empirical isochrone archive for open clusters is the development of a tool that allows a (physically) model-independent and relative age determination for stellar populations. Instead of designing models to try and capture the complexity of stellar astrophysics, and thereby unavoidably having to rely on many assumptions or simplifications (see e.g. Section 3.2), the directive of the empirical isochrone archive is to only utilize information that is available as a measurable quantity. The approach is furthermore completely independent of any a priori estimations concerning the global extinction vector or metallicity of a cluster and only relies on high-quality stellar parallaxes and photometric data, from which cluster CMDs are created and empirical isochrones are extracted. Even if in using this approach, no numerical ages can be assigned to the isochrones, the possibility of making a qualitative comparison between new and/or known clusters can yield valuable insights into cluster evolution, as well as into the properties of large-scale structures in the galaxy. Moreover, if the observational data are very precise, extracting and comparing empirical isochrones could lead to the discovery of new features – be they common or distinct – in their shapes. As a final benefit, empirical isochrones may provide calibration data for various stellar evolution models, possibly improving the quality of model isochrones and in the best case also broadening the general understanding of stellar evolution.

In the following sections, the main expectations for the archive content are outlined along with the limitations that need to be taken into account in this undertaking. Next, the choice of photometric system for the archive and the foundations for the cluster member selection are presented, before introducing the final cluster selection that makes up the basis for the archive.

4.1. Scientific objectives

The first step towards stacking data into the empirical isochrone archive is defining the requirements and scientific objectives that the archive as a whole should fulfill, and discussing their implications regarding the properties of individual clusters that should be contained within it. Therefore, two science questions are formulated as follows:

1. Which archive properties and contents will maximize the information gain of the observational data, as well as the usefulness for age comparisons or model calibrations?
2. Where do the limitations in realizing the desired properties and contents of the archive lie and what are their specifications?

Concerning the first question, three key requirements for the contents of an ideal archive, or any kind of scientific collection for that matter, are defined as:

Relevant: Providing a useful tool for relative age determination of stellar populations requires that, first and foremost, the archive ought to include well-studied and popular clusters. Such groups may have already been used as benchmarks for astrophysical purposes, including but not limited to age determination. As a result, they have been repeatedly observed and their internal processes and environments are likely better understood than those of newly found clusters. Using well-studied examples permits in-depth analyses of apparent disparities in the empirical isochrones between clusters, which in other aspects might be characterized by very similar parameters. Furthermore, it also enables insightful comparisons between benchmark clusters and recently discovered stellar conglomerates.

Comprehensive: Including many open clusters in the archive enables the access to a large number of isochrones for comparing, sorting and calibrating ages. Possessing data for many clusters is also important with respect to the statistical relevance of comparisons between isochrone characteristics, especially when discussing global similarities or discrepancies between isochrone shapes.

Representative: With respect to the isochrone archive, being representative means that the represented clusters should ideally cover a broad age range, starting from very young clusters and spanning to even the very oldest open clusters that inhabit the Milky Way.

4.2. Limitations

As the archive is based solely on observational data, or more precisely parallaxes and photometric measurements of stars, three main limitations to the empirical isochrone archive can be defined: Firstly, the quality of the extracted isochrones is strongly connected to both the quality and quantity of the measurements of the cluster members. Hence, there exist limits to certain observational cluster parameters, beyond which no reliable isochrones cannot be extracted anymore. Additionally, the observational input data neither contain information about the metallicity Z, nor about the extinction value A_V of a cluster, which are both parameters that can considerably influence the shape of a theoretical isochrone. The following paragraphs are dedicated to a discussion of the specifics of these limitations.

4.2.1. Observational constraints of cluster parameters

Reliable isochrones can only be derived, if the CMD distribution of cluster stars forms a sharp and well-defined sequence. Consequently, there also needs to be enough data available to trace a representative cluster shape over a large part of the dynamical range of the diagram. The quality and usefulness of a given cluster CMD is governed by the three parameters distance, cluster age, and number of member stars.

Distance

From an observational perspective, sources become fainter with increasing distance, possibly leading to an under-representation of the low mass stars in clusters, if they are located at a large distance to the sun. Furthermore, as parallaxes decrease with increasing distance to a source, measurement errors grow comparatively large for more distant sources. For instance, Cantat-Gaudin & Anders (2020) describe that for clusters located beyond $\sim 500\,\mathrm{pc}$, the uncertainty in the proper motion measurements of the second *Gaia* data release contributes significantly to the scattering of the cluster members. At roughly $1\,\mathrm{kpc}$, measurement uncertainties become the dominating factors in the observational data. Another point to consider with respect to cluster distance is the cluster environment: Parameters like extinction A_V or metallicity Z of clusters can change significantly throughout the galaxy. Since neither effect can be considered in the creation of empirical isochrones (see Sections 4.2.2 and 4.2.3), the best available option for mitigating their influence is to pick samples with similar environmental conditions – in the Solar Neighborhood.

Age bias

Extinction also affects the span of the age range that can be considered in the archive: Very young clusters ($\lesssim 5$ Myr) may at least partly obscured by their parental dust cloud, leading to significant reddening in the observational data (Lada & Lada 2003, and references therein). Hence, very young or embedded clusters are excluded from the archive. The upper limit to the possible cluster age range is dictated by the comparatively short lifetimes of most open clusters (see Section 2.2.2). Due to their advanced evolution, older clusters are also generally fainter and thus harder to observe, and also much less abundant (e.g. Voigt et al. 2012; Lada & Lada 2003).

Number of cluster members

Since isochrones essentially trace the shape of a whole population in the CMD, obtaining measurements that preferably represent a large portion of the cluster IMF is a vital quality concern. Likewise, empirical isochrones are extracted with statistical methods of varying degrees of complexity, which require a sufficiently large set of data points for each cluster to yield useful results.

4.2.2. Influence of stellar metallictiy

The influence of the metal content of a cluster on the shape of its CMD has been well established (e.g. Voigt et al. 2012), and yet empirical isochrones completely neglect this essential parameter, to which theoretical isochrones are very sensitive (e.g. Voigt et al. 2012; de Boer & Seggewiss 2021). However, metallicity is a parameter that is not only problematic for empirical isochrones, but also for theoretical ones. While it is generally an observable cluster parameter, its measurement is a more difficult undertaking than flux or parallax observations, as it again requires high-resolution spectroscopy measurements. To give an illustrative example, at the time of its publication, Voigt et al. (2012) referred to modern catalogues containing about 1800 known open clusters, an estimated 1100 of which included distance estimates, but only 140 OCs also had metallicity measurements.

Curiously, 950 of the mentioned clusters had age estimates, despite the fact that no information about their metallicity was available (Voigt et al. 2012). This indicates, that the age estimations were largely done without any knowledge about the cluster metallicities. It is actually common astronomic practise to treat the metallicity value of a cluster like a free parameter in the process of isochrone fitting (e.g. Dias et al. 2021), when no a priori information is available. Other methods for example only use a specific Z value in the fitting process when measurements are available for a given cluster, and to otherwise ignore it by assuming Solar metallicity for clusters (e.g. Bossini et al. 2019). Objectively, there exists no clear answer, which of the two approaches is more correct than the other. It can only be seen, that the end results do not necessarily agree, as for example in the case of Melotte 22 in Fig. 3.3. The empirical isochrones cannot yield insight into cluster metallicity, but at the current level of technology, neither can model isochrones, without prior knowledge.

However, one can make educated assumptions about cluster metallicity, based on other parameters like the region of the galaxy that a cluster is located in. Extrapolating from the negligible metallicity gradient within the galactic disk (Voigt et al. 2012), it can be assumed that the metallicity values do not vary extremely between the archive clusters, as their distance is already constrained to the Solar Neighborhood according to data quality and quantity reasons. Especially for young clusters, the assumption of Solar metallicity is generally a fair one (Voigt et al. 2012).

4.2.3. Influence of interstellar extinction

Extinction influences a cluster CMD in a more straightforward way than metallicity does. It usually enters the color index formulation in Eq. (2.6) as a constant term and can often be described or approximated by a two-dimensional extinction vector. In the CMD, high extinction causes stars to appear redder and fainter, thereby moving the population towards the bottom right corner of the diagram. While the empirical isochrones cannot account for this parameter either, they can be easily adapted in case information about a cluster's extinction parameter exists, by incorporating the extinction term when calculating the color index (Voigt et al. 2012). As has already been discussed, extinction can also limit the performance and applicability of empirical isochrones, due to the high scatter it can cause in cluster CMDs. Its effects are more prevalent for younger clusters as they are more likely to be embedded (e.g. Lada & Lada 2003), and can also mitigated up to an extent by restricting the spatial volume of possible archive clusters to a region with a similar, and preferable solar-like environment.

4.3. Choice of photometric system

Having discussed the data quality considerations, the question which photometric systems to include in the archive arises. Not only do different surveys or instruments rarely share a set of passbands, but they also often contain many filters, indicating a multitude of possible combinations for the creation of cluster CMDs. To illustrate the scope of the problem, Tab. 4.1 lists the passbands of the popular photometric systems *2MASS* (Skrutskie et al. 2006), *Gaia* (Gaia Collaboration et al. 2016), *panSTARRs* (Chambers et al. 2016), *SDSS* (Fukugita et al. 1996), and *WISE* (Wright et al. 2010). The number of possible filter combinations for creating a cluster CMD, N_c, within a photometric system with n different filters is calculated with $N_c = n \cdot \sum_{i=1}^{n}(n-i)$. The equation is valid for the twin assumptions that the absolute magnitude shown in the CMD can be represented with every filter, and that the color index is calculated according to the convention of Eq. (2.6). Purposely exaggerating by summing over all combinations, disregarding their physical merit, would equate to 142 CMD representations for a *single cluster*, which would all require empirical isochrones.[9] What is more, even passbands for the same telescope can be adapted over time, as is for example the case for the *Gaia* data releases DR2 and EDR3. Due to a different internal calibration process and more efficient error suppression, new passbands have been published for the latter and astronomers are further advised to treat the DR2 and EDR3 passbands as altogether different photometric systems (Gaia Collaboration et al. 2021).

Table 4.1.: List of different popular photometric systems and the sum of possible CMD diagram representations, not including passband adaptions within surveys.

Survey	Photometric system	N_c
2MASS	J, H, K	9
Gaia	G, $\mathrm{G_{BP}}$, $\mathrm{G_{RP}}$	9
panSTARRs	g_{P1}, r_{P1}, i_{P1}, z_{P1}, y_{P1}	50
SDSS	u' ,g', r', i', z'	50
WISE	$W1$, $W2$, $W3$, $W4$	24

4.4. Gaia data

Including multiple photometric systems in the archive would be beyond the limits of the present project, which mainly revolves around the development stage of the archive. For this reason, only one of the systems listed in Tab. 4.1 is chosen to serve as a pilot for the extraction of empirical isochrone. As has been established, the observational data need to be abundant and should fulfil high quality standards. Thus, the choice falls on the data of the *Gaia* mission, which best suits the requirements of this project. The list of observable parameters of the survey explicitly includes the input data for cluster CMDs, namely stellar parallaxes ϖ, and stellar photometry. The large source catalogue of the survey renders cross-matching data with other optical surveys mostly obsolete (Cantat-Gaudin & Anders 2020). In other words, the *Gaia* source list comes as close to meeting

---

[9] This is a very generous estimate, as in reality some filter combinations are far more commonly used, and much more physically relevant than others.

the requirements of the empirical isochrone archive, as currently possible. Nonetheless, it should be emphasized that the archive can always be extended towards other photometric systems, provided that enough members have been observed and the shape of their distribution in the CMD is sharp enough to permit a reliable isochrone extraction for a given cluster.

Due to the importance of the mission, both to the wider field of astronomy and to this project in particular, a general overview of the scientific mission and the characteristics that ensure the high quality and quantity of data products is given in the subsequent paragraphs. Additionally, the latest data releases, DR2 and EDR3, which are the basis for the archive data, are introduced.

4.4.1. Scientific objectives

The abundance of possible applications for *Gaia* data is undisputed, reaching from small scales, like the search for exoplanets or the cataloguing of near-earth objects (NEOs), to mapping large-scale structures or conducting fundamental physics experiments. Nevertheless, at its core, *Gaia* remains a star census, with the aim of measuring about 1 % of the stars in the galaxy. Objects are preferably measured with a five parameter astrometric solution of right-ascension, declination, proper motions and parallaxes (α, δ, $\mu_{\alpha*}$, μ_{δ}, ϖ), as well as with photometry in the three passbands G, $\mathrm{G_{BP}}$, and $\mathrm{G_{RP}}$ (Gaia Collaboration et al. 2021).

4.4.2. Mission overview

Launched in 2013 into an Lissajous-orbit around the second Lagrangian point (L2, nominal distances $\sim 1.5 \times 10^6$ km) the *Gaia* spacecraft started collecting data for its nominal 60 month mission in 2014 (Gaia Collaboration et al. 2016). The original mission concluded on July 16, 2019 but has since been extended until the end of 2022 (Gaia Collaboration et al. 2021). Profiting from continued instrumental health, and with the extent of the radiation damage still well below anticipated levels, the only factor dictating the telescope's end of life at this point is the limited amount of propellant for the micro-propulsion system. Current projections postulate that the spin rate of the spacecraft should be maintained until the beginning of 2025 at the least. Should the survey runtime ultimately match the whole ten years of the instrument‘s lifetime, the precision of all observation data products could be boosted by up to 40 %, while the proper motion measurements would improve by a factor of 3, compared to a five year mission (Gaia Collaboration et al. 2021). The extrapolated absolute errors in astrometry and photometry are estimated to be in the range of micro-arcseconds and milli-magnitudes, respectively (Gaia Collaboration et al. 2016).

4.4.3. Instrument and measurement principle

The satellite has two fields of view á $0.725\,\mathrm{m}^2$ each, at a basic angle of 106.5 degrees, and an overall focal length of 35 m. Infalling light from both fields is guided onto a shared focal plane (938 million pixels in 7×17 CCDs), thereby enabling a trigonometric calculation of absolute parallaxes, in lieu of providing only relative measurements. As the satellite spins around its axis, the crossing times of targets over the focal plane are measured and their positions relative to the instrument axes are determined. The observational data are then used to reconstruct the instrument pointing and geometric calibration, making

Gaia a self-calibrating mission, with improved systematic errors for superseding observation periods. The basic angle is monitored using focal plane wavefront sensors, which are sensitive to misalignment on a picometer scale. The source catalogue is built up along with the detection on the focal plane, and the SED of the observed objects is determined via prism dispersion. The photometric system consists of a white-light broad-band filter G (330−1050 nm), a blue G_{BP} (330−680 nm) and a red G_{RP} (680−1050 nm) photometer (Gaia Collaboration et al. 2021). The overall sensitivity of the telescope ranges down to 21 G band magnitudes.[10] A field spectrograph ($R \approx 11700$) centered on the Calcium triplet (845−872 nm) is used to determine the radial velocity of some of the sources on a dedicated array on the focal plane (3×4 CCDs). The combined space motion of the satellite orbit around L2, the Earth's motion around the sun, and the uniform scanning mode of the spacecraft is mirrored in the **scanning law** of *Gaia*: Certain areas of the sky are scanned more often than others and at large angles different scanning strips overlap, enabling the derivation of 2D stellar positions (Gaia Collaboration et al. 2016, 2021).

4.4.4. DR 2

The second data release (DR2) (Gaia Collaboration et al. 2018b) includes parallaxes, proper motions, and photometry for about 1.7 billion sources, with a five parameter astrometric solution (α, δ, $\mu_{\alpha *}$, μ_δ, ϖ), and magnitudes in all three passbands provided for about 1.3 billion stars. Radial velocities are measured for about 10 per cent of the sources, and for an added few million stars parameters like surface temperature, variability, radius or luminosity are provided as well.[11] However, significant differences in precision of DR2 can be observed: For instance, between sources brighter than $G \geq 14$ and those with $G \approx 21$, nominal uncertainties of DR2 parallaxes increase from 0.02 mas to 2 mas, and proper motion uncertainties rise from 0.05 mas/yr to 5 mas/yr (Cantat-Gaudin et al. 2018a).

4.4.5. EDR 3

The early third data release (EDR3) (Gaia Collaboration et al. 2021) incorporates raw data from 34 months of observations, resulting in astrometric and photometric measurements for 1.8 billion sources with $G > 21$ mag. For 1.5 billion stars, parallaxes, proper motions, and photometric data in all three passbands are available. No new radial velocity measurements are added with respect to DR2, but the astrometric and photometric measurements reportedly improved significantly in terms of precision, accuracy, and homogeneity. Parallax precision has increased by 30 %, while proper motion accuracy has been boosted by a factor of 2. Even more, systematic errors in the astrometric measurements have been cut by 30-40 % for parallaxes, and by a factor of ~ 2.5 for proper motions. Concerning photometry, in addition to an increased precision of the data, the homogeneity across position, color, and magnitude has improved and no systematic effects above the 1 per cent level ought to remain within a passband. Due to the changes in magnitude and color precision between DR2 and EDR3, a new set of passbands has been published for the latter Riello et al. (2021).

---

[10] Although *Gaia* observes in the optical wavelength range, its lower G band limit surpasses the *J*-band completeness of *2MASS* by 5 mag. Therefore, even infrared clusters discovered with *2MASS* should be visible with *Gaia* (Cantat-Gaudin et al. 2018a).

[11] To put these numbers into perspective, the second catalogue from the *Hipparcos* mission, which represents a pioneer in the field of astrometry, included only about 2 million measurements (Gaia Collaboration et al. 2016).

4.5. Input catalogues and member selection

Once the decision concerning the photometric systems has been made, the next step is to acquire cluster data to serve as input for the isochrone extraction. The prominence and visibility of many of the nearby open clusters have promoted a long and thorough history of observations, culminating in multiple catalogues listing their positions and characteristics. For instance, Charles Messier published his celestial object catalogue even before the turn of the 18th century (Messier 1781), and roughly 100 years later nearly half of the currently known open clusters were already included in the New General Catalogue (NGC) (Dreyer 1888). However, the progress away from calculating mean cluster parameters on the basis of a few prominent cluster stars and towards attaining complete membership lists of open clusters is still ongoing. There exist numerous approaches to the determination of cluster memberships, from older methods like consulting cluster CMDs, or measuring distances between stars and the estimated cluster cores (see references in Voigt et al. 2012), to sophisticated studies of stellar kinematics (Meingast et al. 2021), or the application of machine learning algorithms (Cantat-Gaudin & Anders 2020). Following the onset of large astrometric surveys like *Hipparcos* (ESA 1997) or *Tycho-2* (Høg et al. 2000), selections of cluster members are often based on scanning the higher-dimensional velocity-position space (five parameter solution) of the combined data for overdensities indicating possible clusters. A non-exhaustive list of popular studies based on astrometric measurements includes Dias et al. (2001); Alessi et al. (2003), who use *Tycho-2* data, Kharchenko et al. (2012), using PPMXL data from Roeser et al. (2010), or Cantat-Gaudin et al. (2018b), as well as Cantat-Gaudin et al. (2018a); Cantat-Gaudin & Anders (2020); Meingast et al. (2021), who use *Gaia* DR1 and DR2, respectively. Since membership determinations pose a challenging subject on their own and developing an individual selection method would exceed the bounds of this project, two already existing member compilations are utilized for the present work.

4.5.1. Catalogue I

As has been established, the most accurate and comprehensive data are currently supplied by *Gaia*. Thus, the DR2-based membership lists for 1481 clusters by Cantat-Gaudin & Anders (2020) marks a good starting point for the empirical isochrone archive. In short, they compile a list of known clusters from various sources in literature, and use the astrometric solution of DR2 sources in a volume surrounding the cluster center to determine membership probabilities. Mitigating the effect of the lessened precision for fainter sources in DR2 (see Section 4.4), they conduct their selection exclusively on sources brighter than $G = 18$, corresponding to uncertainties of 0.3 mas/yr in proper motion, and 0.15 mas in parallaxes. They perform a cone-search around the approximate cluster positions listed in literature and initially apply a broad parallax cut (within 0.5 mas of the expected parallax) and no proper motion cut. Depending on the success of the first five iterations of their clustering algorithm, parallax and proper motion values may be further restricted to 0.3 mas and 2 mas/yr, respectively, with the exception of a few nearby clusters with naturally high proper motion values. By applying the clustering algorithm 10 times to each selected space volume, membership probabilities between 0 and 100 % with an increment of 10 % are obtained. The photometric data of DR2 are then used to verify the membership lists via the resulting cluster CMDs (Cantat-Gaudin et al. 2018a).

UPMASK algorithm

The membership determination is conducted by applying a procedure called UPMASK (Unsupervised Photometric Membership Assignment in Stellar Clusters) by Krone-Martins & Moitinho (2014). The working principle of the algorithm relies on two main steps:

1. Identification of stellar agglomerates in the 3D parallax and proper motion space via k-means clustering.
2. Veto step: Verification that the on-sky distribution of the cluster is more concentrated than a random distribution.

Repeating the process with different values for (ϖ, $\mu_{\alpha*}$, μ_δ), drawn from their distributions given their listed uncertainty values and the parameter correlations, transforms the binary veto step value into a membership probability for each star (Cantat-Gaudin et al. 2018a).

4.5.2. Catalogue II

For a subset of the ten open clusters α Per (here: Melotte 20), Blanco 1, IC 2602, IC 2391, Messier 39 (here: NGC 7092), NGC 2451A, NGC 2516, NGC 2547, Platais 9, and the Pleiades (here: Melotte 22), the member catalogue provided by Meingast et al. (2021) is consulted in addition to Catalogue I. The authors determine cluster memberships by inferring the spatial distribution of the clusters from their bulk velocity. The mean cluster velocities are calculated via radial velocity and proper motion measurements for a subset of well-established cluster members and the spatial distribution is then deconvolved with a Gaussian Mixture Model. With their method, the field star component can be eliminated effectively and the line-of-sight errors are taken into consideration as well. In addition to reconstructing the core regions of the investigated clusters, so-called coronae extending $\sim$100 pc from the central regions can be resolved. In terms of the cluster member selections, the recovery of the low density coronae results in comparatively more detected members and thus sharp and well-populated CMDs, down to the low mass end of the MS. Due to the smaller sample size and abundant source data, Catalogue II is often used as a preliminary test sample for the various methods presented in Chapter 5.

4.6. Cluster selection

To get the cluster data for Catalogue I, the corresponding `ViZier` catalogue[12] is accessed and hard cuts in three observational parameters are applied. Firstly, only stars with a membership probability greater than 50 % are accepted as cluster members. Following the discussion on the archive limitations from Section 4.1, the space volume of the catalogue is limited to objects within 500 pc and clusters are further required to have at least 100 accepted members. No cuts need to be applied regarding Catalogue II[13], since all clusters of its selection are located within the defined distance limit, consist of more than 100 sources, and the cluster membership is not based on a probability parameter. Applying the cuts to Catalogue I results in a selection of 67 clusters, which represent an age range between 3 Myr (L1641-S, Da Rio et al. 2017) and 3 Gyr (Ruprecht 147, Curtis et al. 2013).[14]

---

[12] https://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=J/A%2bA/633/A99/members

[13] https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=J/A+A/645/A84

[14] *Note:* The original archive catalogue was Cantat-Gaudin et al. (2018a), which yielded 59 clusters after the quality cuts. Catalogue I retains all clusters of the former publication and adds some of the recently

Table 4.2.: Cut-off values and global parameter ranges for the archive clusters. The listed ages are based on Da Rio et al. (2017); Curtis et al. (2013). The final member count of a cluster may be lower than the initial cut, if sources are removed in a later processing step due to missing parameters in the input catalogue.

Cuts for Catalogue I	
distance (pc)	≤ 500
membership probability	≥ 0.5
number of member stars	≥ 100
Global archive characteristics	
sample size	67
age range (Myr)	$\approx 3 - 3000$
N_*	$99 - 1856$

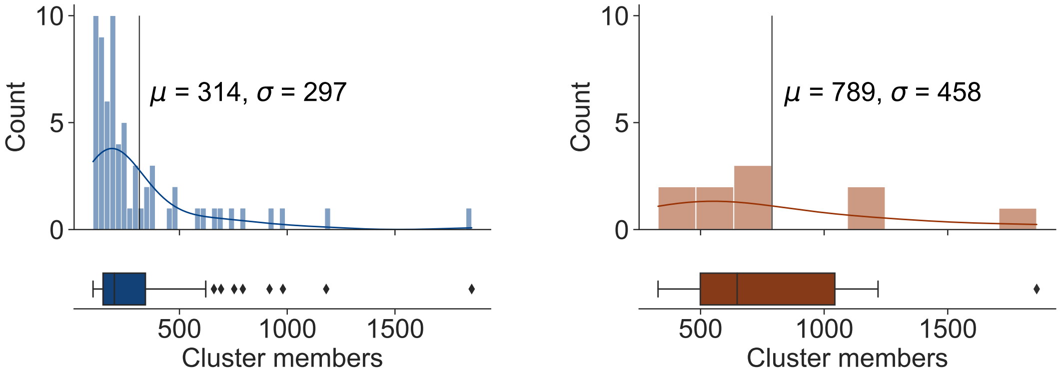

Figure 4.1.: Histogram and boxplot depicting the distribution of the member counts N_* of the archive clusters, after crossmatching the respective member selections of Catalogue I (*left*) and Catalogue II (*right*) with *Gaia* EDR3.

Figure 4.1 shows a histogram and a boxplot of the member counts N_* of the archive clusters, according to their member selections after performing a crossmatch between EDR3 and Catalogue I and II, respectively.[15] As can be seen, many clusters of the archive have only a few hundred members, whereas very abundant groups are rare. It is also apparent, that the clusters of Catalogue II have comparatively larger member selections over whole sample, from a statistical point of view.

In Fig. 4.2, the spatial distribution of the clusters is shown in a cartesian representation of the galactic reference frame. The cluster positions are listed in the table of mean cluster parameters[16] by Cantat-Gaudin & Anders (2020), and the transformation into the galactic frame is performed with the `python` package `astropy` (Astropy Collaboration et al. 2013, 2018). A CMD representation for each cluster can be seen in Appendix A.

---

discovered UPK (Sim et al. 2019) and UBC clusters (Castro-Ginard et al. 2018, 2019) to the archive.

[15] The count N_* of a cluster may vary slightly (absolute deviation usually ≤ 5) when using different color indices or data releases, for a given member star the photometry may not be provided in all passbands.

[16] https://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=J/A%2bA/633/A99/table1

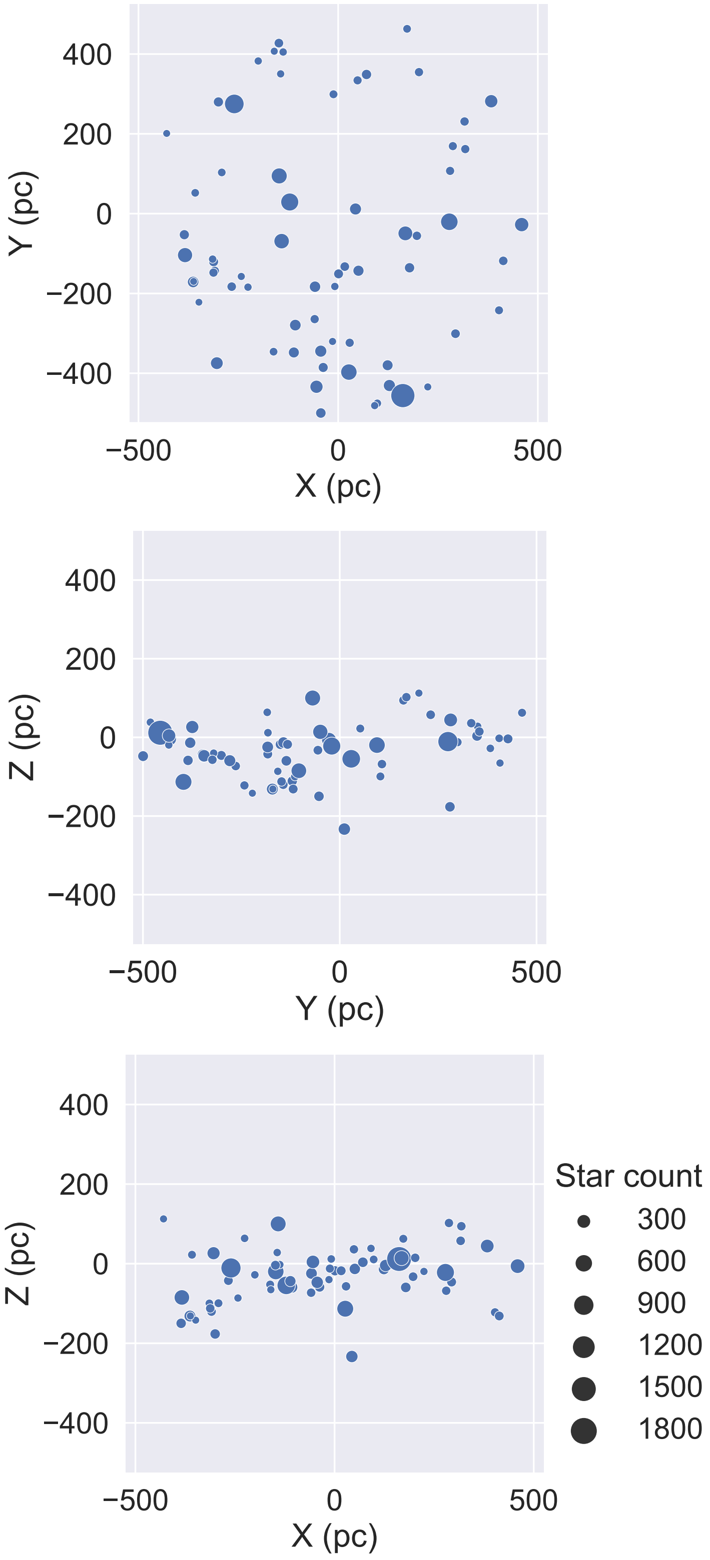

Figure 4.2.: Cartesian representation of the galactic positions and relative member counts of the 67 archive clusters. The size of each data point is scaled according to the cluster member count.

5. Methods

The primary property that distinguishes the empirical isochrone archive from other stellar age determination techniques that also employ isochrones, is its purely empirical nature. The goal of the archive is to avoid the caveats and drawbacks associated with the model-dependent isochrone fitting method, while simultaneously taking advantage of its large applicability over stellar age and mass ranges, as compared to other age determination methods. In creating the archive, one can profit from the direct origin of isochrones in the HRD or CMD, because its construction requires only three observational parameters. As a result, an empirical isochrone can be directly extracted from brightness measurements in two passbands and a stellar parallax for each cluster member. While in the plots shown throughout this work usually all three *Gaia* passbands are used, with the G band magnitude depicted on the abscissa and the color index generated from the e G_{BP} and G_{RP} bands, including the G band in the color index and omitting one of the other passbands would work just as well for the isochrone generation. However, when using cluster CMDs as input data, special observational artefacts, called unresolved binaries, need to be taken into account. The location and meaning of these objects is discussed in a subsequent section.

When confronted with the development of an extraction method for the isochrones, two fundamentally different approaches are taken to provide multiple perspectives on the methodology of possible extraction algorithms:

1. **Data analysis**: Only the data of an individual cluster is considered, and the isochrone extraction happens through exploration of the underlying relationships between the different input variables. The two-dimensional CMD shape of the cluster does not explicitly factor into the extraction process, however, the diagram is used to evaluate the quality of the result and thereby assess the performance of the extraction method.
2. **Image processing**: A cluster CMD can also be treated like an actual 2D image and isochrones can be extracted on the basis of the apparent two dimensional features.

To discover the approach that is better suited for the extraction of reliable and robust isochrones, a total of five different methods stemming from one of the two different realms are presented and tested on selected sample clusters in this chapter. An additional two methods are investigated with respect to the possibility of data pre-processing, which could improve the performance of various extraction methods. The goal is to find a method, or a combination of methods, that ideally performs equally well for different clusters in the sample, without requiring extensive manual parameter-tuning. It should be emphasized, that neither the two different approaches to the problem, nor the selected methods discussed in this chapter explore all possibilities of dealing with the extraction of empirical isochrones. Rather, in minding the scope of the project as a whole, the methods were chosen based on their working principles and on their applicability to the process of isochrone extraction.

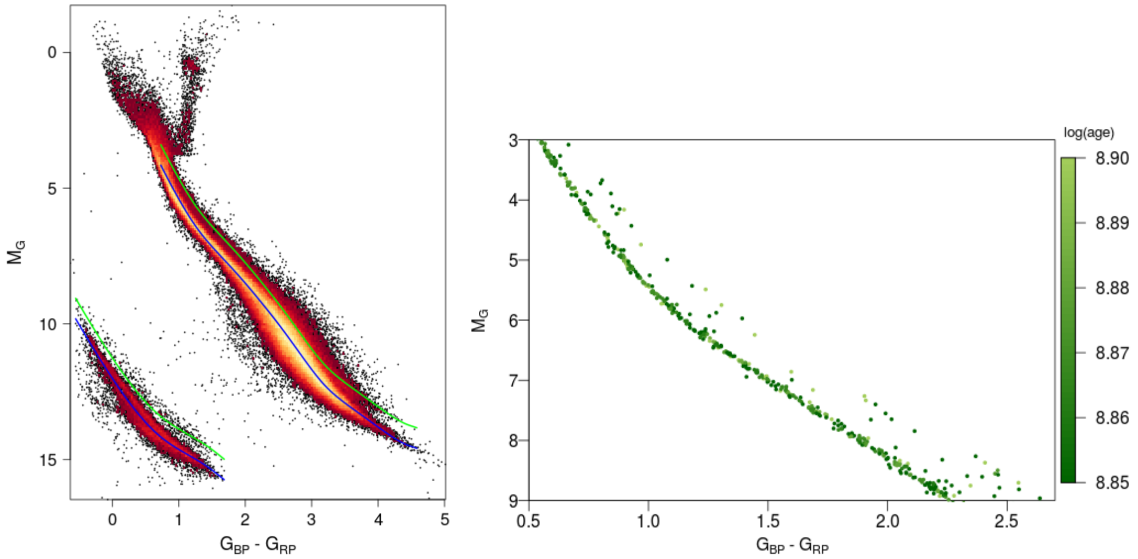

Figure 5.1.: *Left:* CMD including about 2×10^5 sources within 100 pc of the sun. The median fiducial (*blue line*), and its -0.753 mag shift (*green line*) are also displayed. *Right:* Scattering of unresolved binaries above the main sequence of the CMDs of the Hyades and Praesepe clusters. Image credit: Gaia Collaboration et al. (2018a).

Each method is introduced by outlining its working principle, before the reasons for choosing it are stated. Subsequently, the method is applied to a test sample and the results are visualized. Lastly, its performance is evaluated and the decision whether it is retained for further development or discarded in favor of a more promising method is declared. The tests are performed on EDR3 data, because it has the smaller measurement errors of the two *Gaia* data releases and is therefore more likely to produce a high-quality empirical isochrone in the development phase of the extraction tool. The application of all methods is conducted with with `python`, using the packages referenced in the introduction of the respective method, if applicable. The source code for all presented methods, as well as for the final extraction algorithm (see Chapter 6), is available on the `GitHub` page of the empirical isochrone archive project: `https://github.com/rottenstea/Isochrone_Archive`.

5.1. Unresolved binaries

Before delving into the specifics of the various extraction and pre-processing methods, a peculiarity of the CMD representation of stellar data needs to be addressed: The CMD representation of a cluster (using *Gaia* data) is often accompanied by a feature along the main sequence, that is especially clearly visible for open clusters and not connected to stellar evolution, but instead results from the resolution limit of telescopes (Gaia Collaboration et al. 2018a). The incoming light of a celestial object is usually fitted with a point-spread-function (psf) to determine the electromagnetic flux. However, there exists a limit of how close to each other two sources can appear in the field of view of a telescope, to still be detected as separate sources. Usually, the separation limit is defined as the point, at which the first minima of the respective point-spread-functions of two sources are just shy of overlapping. If the limit is not fulfilled, the light of the two sources, which are most likely double stars or binary systems, is detected as a single data point, which leads to a scattering of unresolved binaries throughout the cluster MS in the CMD. The location of such a scattered data point can be determined theoretically: An unresolved binary system consisting of two stars with equal masses is measured at the same color, however, the apparent magnitude difference to a single star is (Voigt et al. 2012)

$$\Delta m = 2.5 \log(2) \simeq -0.753 \text{ mag}. \tag{5.1}$$

As a result, unresolved, equal-mass binaries experience a vertical displacement of -0.753 mag in the cluster CMD, which is highlighted by the green fiducial curve in the left panel of Fig. 5.1. The distinct but much more sparsely populated sequence parallel to the MS formed by these observational artifacts is conventionally referred to as **binary sequence**, and can be seen for example in the CMD on the right-hand side of the figure for the Praesepe and Hyades open clusters. The prominence of the binary sequence is not necessarily caused by a bias of binary systems towards a mass ratio $q = M_2/M_1 = 1$. Rather, as described by Hurley & Tout (1998), systems of two stars with masses $M_1 > M_2, 0 \leq q \leq 1$ can also occur on the same-mass binary sequence, because for unresolved binaries of different masses a change in the observed color of the unresolved data point occurs. As M_2 decreases with respect to M_1, the binary data point appears fainter and redder for q greater than a critical value q_{crit}. Therefore, for a given binary ratio q, the data point can align with the equal-mass binary sequence of a cluster, if $q_{\text{crit}} < q \leq 1$ is fulfilled. The critical ratio varies in dependence of the binary masses, for example, the authors find $q_{\text{crit}} \simeq 0.5$ for $M_1 \geq 3\ M_\odot$, or $q_{\text{crit}} \simeq 0.9$ for $M_1 \leq 2\ M_\odot$. The location of unresolved binary systems that consist of stars in different evolutionary stages can either be above the binary sequence, if a MS star is paired with a red giant, or between the MS and the white dwarf region, in the case of a MS star and a white dwarf binary (Gaia Collaboration et al. 2018a). Both regions are visible in the left panel of Fig. 5.1 and they are clearly populated.

Although many unresolved binaries have already been purged in the finished *Gaia* data releases (e.g. Gaia Collaboration et al. 2018a), some remain in the cluster CMDs of the archive clusters, as can be seen for many of the clusters in Appendix A. Consequently, one has to be cautious to avoid biasing the empirical isochrones with their presence, especially with respect to methods from the approach of data analysis, which have been discovered to be more susceptible to these kind of outliers in the data, than the image processing approach.

5.2. Pre-processing methods

When faced with the task of developing an isochrone extraction routine, it can be beneficial to not merely take the raw observational data as fixed input. Instead, one can also explore different possibilities regarding data pre-processing, which might ultimately yield better results in combination with an extraction method. To this end, the methods extreme deconvolution (XD) and principal component analysis (PCA), which do not produce empirical isochrones when applied on their own, are investigated.

5.2.1. Extreme deconvolution

The first pre-processing method under consideration is the "Extreme Deconvolution of Gaussian Mixture Models" (XDGMM) described in Željko Ivezić et al. (2019), who also developed an implementation of the method in the `astroML` package[17] and whose work is the basis for the following description of the method. Generally, XD can either be used for density estimation, or for clustering of point distributions in multiple dimensions. The aim of the procedure is to infer the underlying probability density function (pdf) of a distribution by smoothing the original data points. Mixture models, such as the GMM employed in the present method, are parametric estimation tools, in which the number of kernels, as well as their locations and widths are varied to obtain a good fit to the true pdf of the data. They can be seen as a bridge between non-parametric density estimation, such as kernel density estimation, where each data point is represented by a kernel function and the aim is to capture every aspect of the underlying pdf, and clustering, where the locations and number of kernels take on physical importance.

Principle

In a multivariate Gaussian Mixture Model (GMM), the density distribution of a dataset is represented by a sum of M Gaussians $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, with given means $\boldsymbol{\mu}_i$, variances $\boldsymbol{\Sigma}_i$, as well as assigned weights α_i, with the condition $\sum_{i=1}^{M} \alpha_i = 1$. Every point in the distribution is therefore sampled from one of the Gaussian functions and the pdf of $\boldsymbol{x}$ can be formulated as follows

$$p(\boldsymbol{x}) = \sum_{i=1}^{M} \alpha_i \, \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \tag{5.2}$$

where the multivariate form of a Gaussian function in D dimensions can be written as

$$\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{\sqrt{(2\pi)^D \det(\boldsymbol{\Sigma}_i)}} \exp\left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right). \tag{5.3}$$

However, in observational astronomy one generally has to account for measurement errors of the data, as only a noisy version of the true pdf can be observed. Extreme deconvolution incorporates such errors into a GMM via a Bayesian approach. It is based on the assumption that the true values $\boldsymbol{v}_i$ of the dataset are related to the measurements $\boldsymbol{x}_i$ via

$$\boldsymbol{x}_i = \boldsymbol{R}_i \boldsymbol{v}_i + \boldsymbol{\varepsilon}_i. \tag{5.4}$$

The variable $\boldsymbol{R}_i$ denotes a projection matrix and $\boldsymbol{\varepsilon}_i$ symbolizes the noise component, which is assumed to be drawn from a Gaussian distribution with mean zero and variance $\boldsymbol{S}_i$. The method of extreme deconvolution works by using $\boldsymbol{R}_i$ and $\boldsymbol{S}_i$ as priors to find values for $\boldsymbol{\mu}_i$, $\boldsymbol{\Sigma}_i$, and the weights α_i of the individual Gaussians, which maximize the likelihood of the observed data. The GMM is called a generative model, which means that it fully describes the data.

It should be emphasized that the amount of Gaussians M does not necessarily correspond to the number of clusters found in a dataset – on the contrary, it almost always deviates

---

[17] https://www.astroml.org/index.html

from the number of clusters, for instance because their shapes are usually not Gaussian or because the point distribution includes a strong noise background. In the case of density estimation, it can even be useful to set M to a much larger value than the suspected cluster number to map the pdf more accurately. The number of Gaussian components can then for example be divided into sub-clusters and background distributions via a density threshold.

Aims

The motivation for employing XDGMM as pre-processing procedure is three-fold: Firstly, the expression of the data through multiple Gaussians could make it possible to precisely locate the ones representing the less densely populated binary sequence and omit them. Without the presence of the biasing effects of the binary sequence, extracting the isochrone with statistical tools, for example by locally averaging over the data, would become more reliable. Secondly, the GMM representation could be utilized to clone the original dataset, meaning that one could simulate additional cluster members by drawing from the pdf that describes the observations. Boosting the amount of data in this fashion would be especially useful for clusters with few members and sparsely populated CMDs, which do not provide as good an input for isochrone extractions than clusters with a larger N_* value. Lastly, the method already incorporates the measurement errors associated with the observational data, which have to be considered separately for other methods.

Application

In Fig. 5.2 the XDGMM method implemented in the `astroML` package is applied to the Melotte 22 cluster of Catalogue II. The true source distribution is shown in the plot in the left corner (*blue data points*), and the four smaller plots correspond to illustrations of the locations and shapes of the M Gaussian components (*left panels*), where $M \in \{5, 10, 50, 100\}$, and simulated data points that were resampled from them (*orange data points, right panels*). The maximum number of iterations for the calculations in the figure is fixed to 1000, however convergence is always reached beforehand for the example in the figure. As can be seen, a low number of components ($\approx 5 - 10$) is not enough to differentiate between the binary and the main sequence, whereas increasing the components by a factor of a magnitude leads to an unwieldy amount of clusters that also do not separate naturally into a MS and a binary sequence branch. Furthermore, the computation time increases significantly with the increase of M. It can also be observed that some of the key features of the cluster, like the shape of the upper main sequence, are partly lost or washed out in the resampled distribution.

Conclusion

XDGMM is a powerful data analysis tool that can be used for both density estimation and clustering in astronomy. Unlike most other methods presented in this chapter, it is already designed to expect and incorporate measurement errors and can, in theory, be employed to boost the CMD population for clusters with few members. Furthermore, the `astroML` implementation presents a very convenient and user-friendly way to apply the tool to data, even for astronomers who are unfamiliar with the concept of extreme deconvolution.

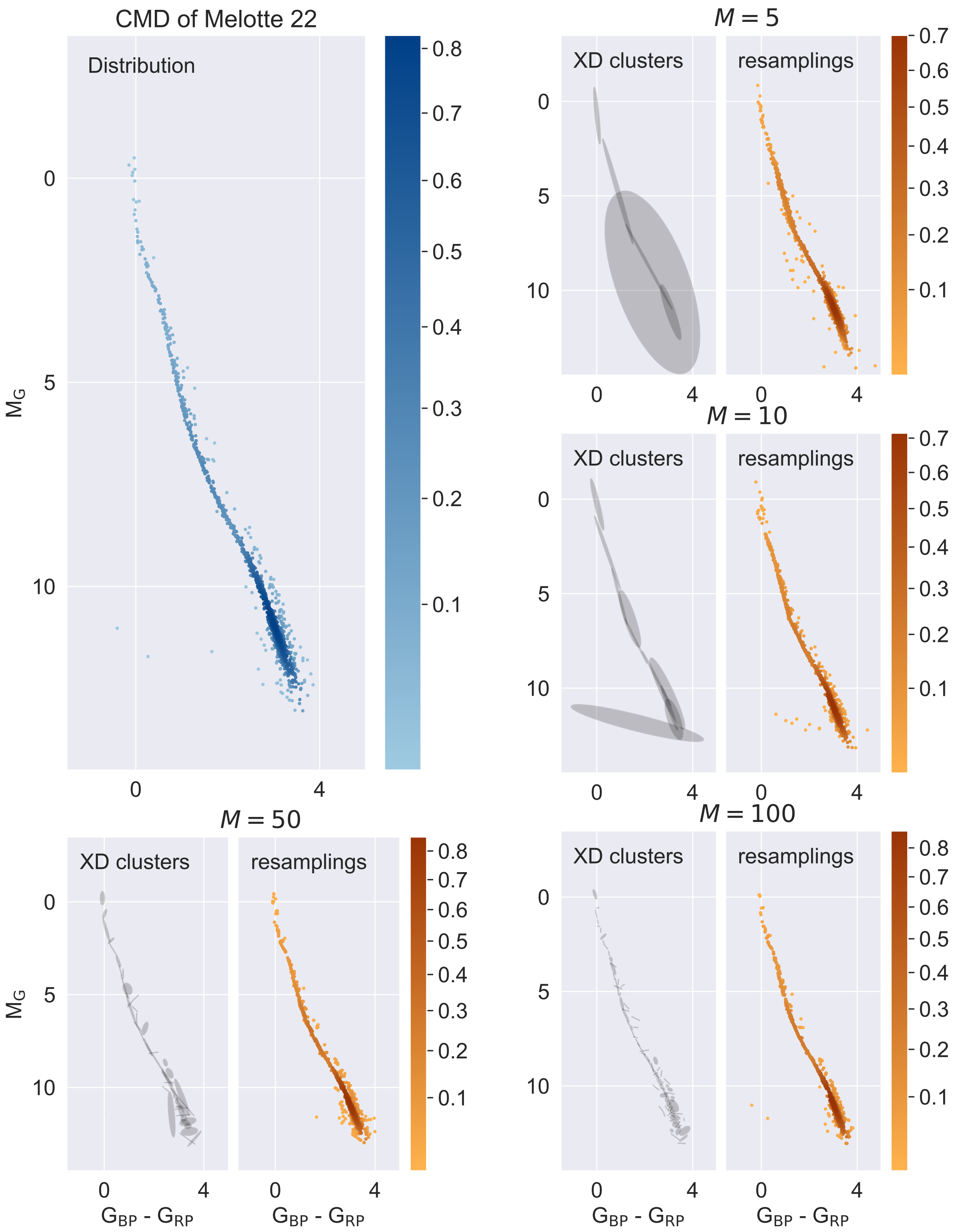

Figure 5.2.: *Upper left:* CMD of Melotte 22. *Other panels:* Application of the XDGMM method to the cluster under variation of the number of Gaussian components M. The subplots on the left-hand side depict the locations and shapes of the GMM, right-hand side panels depict the resampled cluster CMDs.

On the downside, like it is often the case when clustering is concerned in astronomy, the definition of the "correct" number of Gaussian components M is not straightforward and greatly affects the outcome of the calculation. As can be seen for the test cluster in Fig. 5.2, the number of components could not be optimized in a way that allows for the exclusion of the binary sequence, while simultaneously preserving all features of the CMD. While such a sweet spot might still exist for a value of M, it would likely be different for every cluster, depending on the individual number of observations and the shape of each individual CMD. As the calculation of XDGMM is also computationally expensive, varying M without prior knowledge or certainty that there will be a decisive result for all clusters of the archive seems ill-advised.

Concerning the simulation of additional cluster data, the figure shows that certain features of the CMD are washed out or altered, which negates the aim of preserving the cluster shape in the empirical isochrones as completely as possible.

Ultimately, the method does not substantiate the hopes of eliminating the binary sequence and bolstering the empirical cluster data, which is why it is not pursued further in this work.

5.2.2. Principal component analysis

The second pre-processing method under investigation is a technique called principal component analysis (PCA). It is based on finding linear combinations of variables that are characterized by special variances, meaning that the first principal component is the normalized linear combination of the data that has the maximum variance and so forth. Due to its special principle, which produces a matrix of variables with descending variance values, PCA is very popular for the purpose of dimensionality reduction, with special applications in areas like image analysis, visualization, pattern recognition, as well as regression. For the purpose of this project, the PCA function implemented in the `scikit-learn` package (Pedregosa et al. 2011) is utilized. It is created from a probabilistic formulation based on a maximum likelihood framework, which broadens the original PCA formulation towards non-linear applications. For further details, readers are referred to the publication of Tipping & Bishop (1999).

Principle

For a basic definition of principal components, a p-dimensional random vector $\boldsymbol{x}$ with expectation value $E(\boldsymbol{x}) = \boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma} = E\left[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\right]$ is defined. The vector entries x_i should then be projected by using a $p \times p$ matrix Γ with orthogonal columns and unitary vector entries $\boldsymbol{\gamma}_i$

$$z_i = \boldsymbol{\gamma}_i^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu}), \tag{5.5}$$

$$\mathrm{Var}(z_i) = \boldsymbol{\gamma}_i^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\gamma}_i. \tag{5.6}$$

The goal is to choose the transformation matrix $\boldsymbol{\Gamma}$ in a way that maximizes the variance of the components of $\boldsymbol{z}$, which is expressed in Eq. (5.6). The solution to such a maximization problem can be formulated with the help of Lagrangian multipliers (see e.g. Anderson 2003, for a detailed derivation) and results in an eigenvector/eigenvalue problem. Iteratively solving the equations is done in such a manner, that the first principal component z_1, for which the variance $\mathrm{Var}(z_1)$ has been maximized, is the one that points in the direction of the eigenvector $\boldsymbol{\gamma}_1$ corresponding to the largest eigenvalue, the second component z_2 corresponds to the second largest eigenvalue, and so on. Collecting all p eigenvectors in $\boldsymbol{\Gamma}$ and storing the corresponding eigenvalues in descending order of value in a diagonal matrix $\boldsymbol{A}$, permits the expression of the principal component solution in matrix form as a **spectral decomposition**

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{A}\boldsymbol{\Gamma}^{\mathrm{T}}. \tag{5.7}$$

The linear transformation from Eq. (5.5) is called **principal component transformation**, the elements of the ordered vector $\boldsymbol{z}$ are called **principal components**, and $\boldsymbol{\Gamma}$ is known as **loading matrix**. The elements γ_{ij} of the latter reflect the influence the initial entries x_i have on the entries z_j.

When applying the PCA concept to observational data in form of a matrix $\boldsymbol{X}$, one estimates the expectation vector and covariance matrix of the data via the sample mean $\hat{\boldsymbol{\mu}}$ and sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ and defines the principal components $\boldsymbol{Z}$ and the spectral decomposition of the data analogous to Eq. (5.5) and Eq. (5.7)

$$\boldsymbol{Z} = (\boldsymbol{X} - \mathbf{1}\hat{\boldsymbol{\mu}}^{\mathrm{T}})\hat{\boldsymbol{\Gamma}} \tag{5.8}$$

$$\boldsymbol{S} = \hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{A}}\hat{\boldsymbol{\Gamma}}^{\mathrm{T}}. \tag{5.9}$$

The matrix $\boldsymbol{S}$ is the eigenvector matrix and $\hat{\boldsymbol{A}}$ holds the eigenvalues (Anderson 2003). The observational data can either be directly transformed into PCA space, or the dimensionality of the dataset can be reduced first by choosing to only include the parameters with the highest variance values. This way, a transformation equals a transformation into the **principal subspace** (Tipping & Bishop 1999, and references therein).

In its original form, PCA only describes linear transformations of data, thereby limiting its applicability. For this reason, various amendments to the method have been proposed to add a more sophisticated, nonlinear character to it (see e.g. Tipping & Bishop 1999).

Aims

The first motivator for performing PCA on the input data is the resulting rotation of the coordinate system, such that extraction methods may become more efficient. The point distribution of a cluster CMD using Gaia passbands generally spans a large dynamical range on the y-axis, and a comparatively narrow range on the ordinate axis. As a result, the variance in the two dimensional data can be divided into a clearly dominant first principal component and a much smaller second principal component, which are illustrated on the example of Melotte 22 in the left panel of Fig. 5.3. Hence, projecting the data into PCA space, as seen in the right panel of the figure, significantly reduces the variance of the data points, which in turn can be beneficial for isochrone extraction tools. The second reason for transforming cluster data into PCA space regards the handling of evolved objects: Clusters that display a MSTO in their CMD distributions are often not handled well by isochrone extraction methods from the field of data analysis, as most of them depend on some form

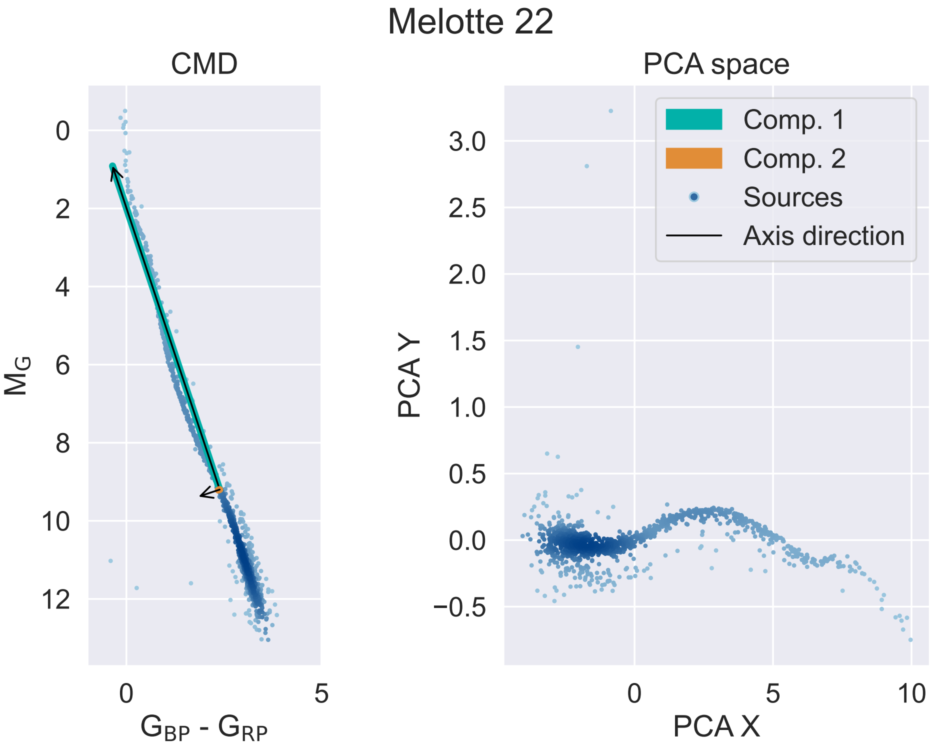

Figure 5.3.: *Left:* Extent and orientation of the principal components of the cluster data of the use-case Melotte 22. The first component (*green*) is much larger than the second one (*orange*). The directions of the PCA space axes are denoted by black arrows, where the second one has been magnified by a factor of 15 for better visibility. *Right:* Transformed cluster data in PCA space, with the new axes pointing in the direction of the two PCA components.

of functional dependency between the input variables. While this dependency ceases to exist in the CMD space of older clusters with a MSTO, a transformation of the data into PCA space re-establishes it.

Application

Figure 5.3 illustrates the two principal components for the cluster data of Melotte 22, which have been scaled by their corresponding variance values. As can be seen, the first component (*green*) is extremely dominant compared to the second one (*orange*), with the extent of the second component not even being visible in the plot. The black arrows in the figure display the directions of the components, or rather their corresponding eigenvectors, and therefore the PCA space axes. The arrow corresponding to the smaller component has been magnified by a factor of 15 to enhance its visibility in the plot. The PCA-transformed cluster data can be seen in the right panel of Fig. 5.3. A look at the y-axis of the PCA space confirms that the scatter in the data is much smaller compared to the cluster CMD in the left panel of the figure.

Conclusion

Principal component analysis is a computationally efficient, easily implemented method for pre-processing cluster data. While the method is not used for dimensionality reduction in the present work, as the input data are already only two-dimensional, it can be employed to the benefit of reducing its variance. The effects of the transformation are for example apparent in a smoother empirical isochrone in the lower main sequence of a cluster CMD, when used in combination with percentile fitting (see Fig. 5.5). Another important benefit of the procedure is that it preserves the functional relationship between the two input variables for evolved clusters with a MSTO, which is beneficial for many methods based on data analysis. The good results of the preliminary tests of combinations between PCA and isochrone extraction methods motivate the decision of retaining the method and combining it with the best-performing extraction algorithm in Chapter 6.

5.3. Data analysis methods

Having explored different ways of improving the input data, the subsequent paragraphs are focused on finding extraction methods based solely on exploiting the relationships between the different input variables of the data. Overall, three different methods are discussed, starting from a straightforward approach percentile fitting, towards the increasing computational and theoretical complexity of the two methods principal curves, and support vector regression.

5.3.1. Percentile fitting

Percentile fitting represents an intuitive, straightforward approach to the extraction of empirical isochrones. It is based on the idea of locally approximating the point distribution of a cluster CMD, by compiling the cluster data into a sorted array and then calculating a statistically relevant parameter, such as the median, average, or an arbitrary percentile value, at discrete points along the parameter range of the explanatory variable. The empirical isochrone is then obtained by interpolating between the discretely determined values. As the method is very simple, only the `numpy` package[18] (Harris et al. 2020) is required for its implementation.

Principle

The basic idea of percentile fitting is to determine an explanatory and a response variable and sort the input data along the axis of the former. Contrary to intuitive belief, it can be helpful to determine the absolute magnitude variable of a cluster CMD as explanatory parameter. In doing so, the CMD axis are practically flipped during the percentile fitting procedure, and the functional relationship between the CMD variables is maintained, even for clusters with a MSTO. Figure 5.4 serves as an illustration of how the percentile fitting method performs for the old cluster Ruprecht 147, depending on which variable is chosen as the explanatory and the response variable. Empirically, it can be established that using the absolute magnitudes as explanatory variables of the CMD yields smoother empirical isochrones for all clusters when the CMD aces are flipped. After defining the dependent and independent variables, the procedure iterates through the sorted data and calculate

---

[18] https://numpy.org/

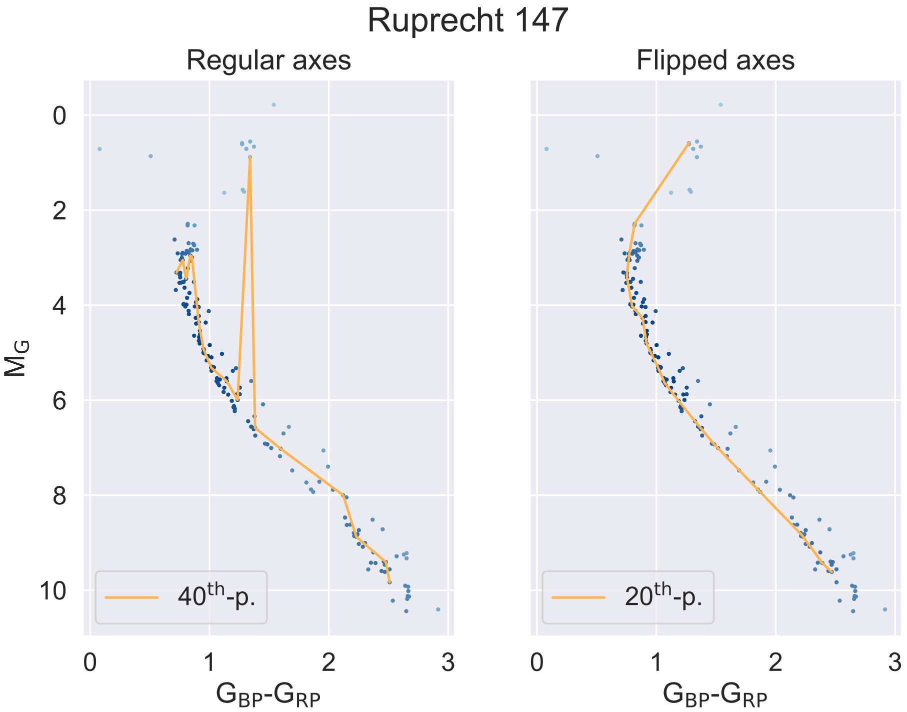

Figure 5.4.: Use-case example of applying the percentile fitting method on the evolved cluster Ruprecht 147. Choosing the color index as independent variable is problematic for the method due to the populated MSTO and giant region. Flipping the axes, however, solves the problem. The percentiles of the isochrones are chosen to represent the best results.

the chosen statistical value for a fixed number of data points. Subsequently, the nearest corresponding value to the statistical value in the sample is extracted via interpolation, along with the corresponding explanatory variable at the same index. The coordinates can either be directly connected to form an empirical isochrone, or be post-processed via Gaussian deconvolution with a one-dimensional Gaussian kernel to produce a smoother curve.

Aims

The percentile fitting method appeals through its idea of simplicity, as well as through its direct implementation, which leaves it completely adaptable to the specific requirements of the present project. The straightforward approach limits the possibility of overfitting the data and producing unwanted artefacts. By using arbitrarily chosen percentiles instead of the average value for determining the shape of an empirical isochrone, biasing effects of the binary sequence can be avoided. Additionally, the method can easily be expanded by adding PCA as a pre-processing method, and could be adapted to produce uncertainty regions by repeating the extraction on bootstrapped cluster data and then calculating different percentile or quantile values from the collection of all resampled isochrones.

Application

The right panel of Fig. 5.5 shows an example of applying percentile fitting straight to the CMD data of the cluster Platais 9 (*red line*). To generate the isochrone, a fixed stepsize $n = 10$ was used, indicating that a discrete point of the empirical isochrone was extracted from every ten sources in the ordered cluster data. By varying the percentile values and through visual verification, the best shape for the empirical isochrone can be determined. It should not be biased by binaries and should pass through the most densely populated regions. The left panel of the figure displays an example of an isochrone that is extracted from pre-processed data in PCA space and then projected back into the CMD space (*orange line*). As can be seen, pre-processing the data before extracting the isochrone via percentile fitting yields a smoother curve and additionally eliminates the need to flip the cluster axes for evolved clusters with a MSTO. However, since the method works by interpolating between points that are already part of the cluster distribution, the extracted isochrones can assume unnaturally straight shapes and edges. Furthermore, the results are very sensitive to the choice of the stepsize n for the percentile sampling and the edge regions cannot be fully reproduced with the extraction process.

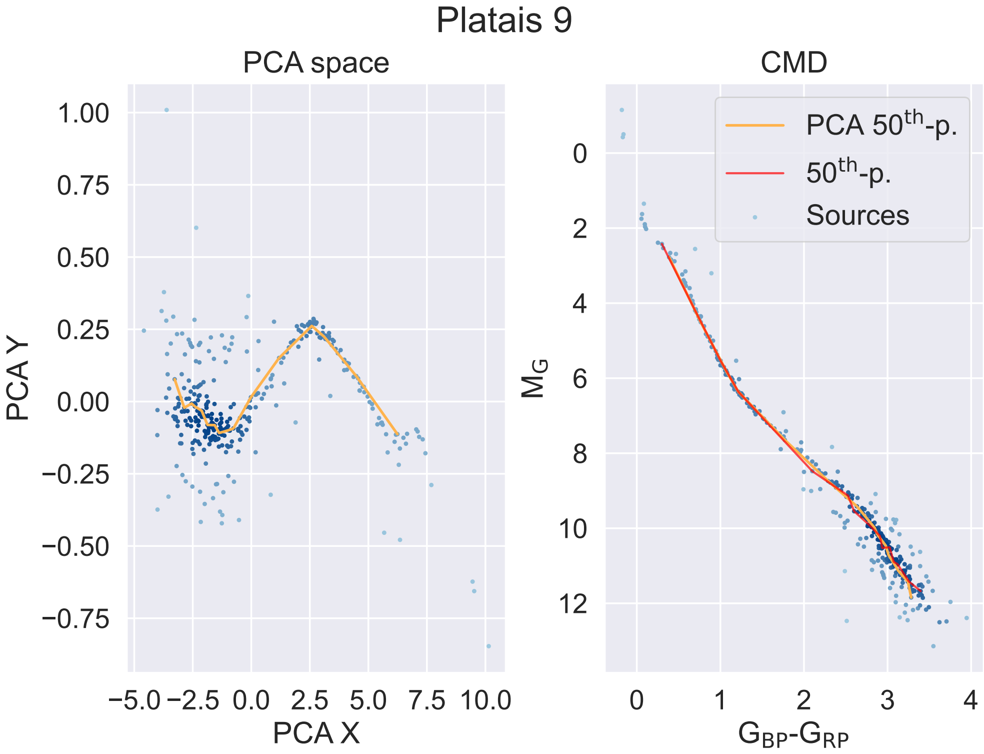

Figure 5.5.: Observational data for the cluster Platais 9, shown in PCA space (*left panel*) and the CMD space (*right panel*). The orange line corresponds to an empirical isochrone extracted via percentile fitting in the PCA space, whereas the red line represents an isochrone extracted directly in the CMD space using the same method and parameters.

Conclusion

Percentile fitting is a promising, straightforward procedure for extracting empirical isochrones, which is not only highly adjustable to individual clusters, but also not discriminatory against cluster ages when used in combination with PCA. Furthermore, pairing the method with bootstrapping would allow for an easy extraction of uncertainty boundaries, represented by quantile or percentile values that could be extracted alongside the actual isochrone values. On the downside, owning to the working principle of the method, the isochrones can display unnaturally straight stretches or sudden edges, which are unwanted features that do not reproduce the shape of the cluster distribution in the CMD. The quality of the results strongly depends on the choice of the stepsize n, which in turn relies heavily on the cluster member number N_*. If the stepsize is chosen too small, the lower part of the isochrone displays wiggles and edges, whereas if the stepsize is too large, the upper MS region of the cluster cannot be represented accurately anymore. The situation is further complicated by the fact, that the stepsize concerning the lower part of the cluster CMD would need to act as an opposite to the stepsize regarding the upper part: For the former, large bins holding many members are needed to avoid wiggles in the empirical isochrone. In the upper region, bins that only hold very few members would be needed to accurately reproduce the characteristic upper MS sequence of each cluster.

All in all, the method shows good success considering its simplicity on the test sample, and can be quickly improved by adapting the input parameters according to each cluster. Regardless, in-depth investigations of ways to optimize the stepsize n for each cluster, preferably in an automated procedure, turned out to be unsuccessful. Compared to other methods presented in this chapter, the inflexibility of percentile fitting with respect to the stepsize is a significant drawback, which can also heavily impact the quality of the results. Therefore, the method is not further developed.

5.3.2. Principal curves

The concept of principal curves (PC) was first proposed by Hastie & Stuetzle (1989), as a way of a smooth and non-linear summary of the pattern drawn by observational data, for example within a scatter-plot consisting of two variables. The term principal curve refers to a non-parametric 1D curve that marks the middle of a p-dimensional dataset. As the name of the method suggests, there exists a connection between principal component analysis (Section 5.2.2) and principal curves, in the sense that the first principal component of the input data is often used as a starting point for the PC algorithm. By iteratively determining the local average of the p-dimensional points in a smooth manner until convergence is reached, a principal curve of the dataset is created. The concepts of the algorithm described in Hastie & Stuetzle (1989) are realized in the `python` package `pcurvepy`, which is available at `Github`.[19]

Principle

As PC is one of the most sophisticated methods in the method selection, capturing the theoretical basis for its algorithm in a strictly mathematical fashion lies beyond the scope of this thesis. Thus, in the short description of working principles, the focus is placed on

---

[19] https://github.com/zsteve/pcurvepy/blob/master/README.md

the application aspect of the method and the algorithm for finding principal curves, rather than on formal definitions. For an exhaustive discussion of principal curves, readers are referred to the original paper of Hastie & Stuetzle (1989).

The main motivation of the principal curves procedure in two dimensions, is to treat both variables symmetrically and find a smooth, self-consistent line that passes exactly through the middle of the data – called a principal curve. The connection between principal curves and PCA stems from the fact, that for a set of observations, a principal curve is found in an iterative manner and corresponding algorithm that is applied to the data requires an initial estimate for a principal curve. For this reason, the first principal component of the input data, which is a straight line that minimizes the squared deviations in both variables (see Section 5.2.2), is calculated and used as a starting point for the first iteration. In a subsequent step, the curve estimate is inspected with regard to its self-consistency within the dataset. The analysis is conducted by looking at each point on the curve and collecting all data that projects onto it. When taking the average of these data points, it should coincide with the point on the curve. In the algorithm, the projection and averaging of the data points are performed locally along the curve, and are called the **projection step** and **expectation step**, respectively. If the curve is not self-consistent, the averaged values are used as a new starting point for the next iteration, until convergence, usually defined by a tolerance criterion for the changes between expectation and averaging, is reached (Hastie & Stuetzle 1989).

Aims

The concept of using principal curves as empirical isochrones has already been employed by Bouy et al. (2015) and Olivares et al. (2019), for example. The primary benefit of the method is, that it has been specifically designed with the intent of representing shapes and patterns in the input data – exactly like an empirical isochrone should. Furthermore, the procedure guarantees a smooth curve without wiggles or sharp edges in the lower main sequence region, which can appear when using other methods, like percentile fitting. Moreover, the concept seems similar to using splines, albeit circumventing the disadvantages of choosing where to define spline nodes and how to handle the boundary regions.

Application

The performance of the principal curves procedure is visualized for three different clusters in Fig. 5.6. It was decided to use samples from Catalogue I in this use-case example, as they work well in illustrating the advantages and drawbacks of the method. As is apparent from the right panel of the figure, the method has no problems in reproducing the curved shape of the MSTO of the old Ruprecht 147 cluster, which is not surprising as it treats both input variables symmetrically and does not depend on a functional relation between them. However, the curve does not directly pass through the stars forming the MS, but slightly above them, as it is influenced by the binary sequence. For clusters with larger scatter and less distinctive features in their CMDs, such as Collinder 69 in the center panel of the figure, the procedure also produces good results, although the isochrone deviates from the data points in the upper region of the MS. For a sharp, well-defined upper MS, as seen for IC 2602 in the left panel of the figure, the produced curve is too smooth and does not capture all specific characteristics of the cluster distribution in the CMD. Investigations of the source code of the PC implementation employed in this work show, that the curve representation is realized via univariate spline interpolation and that the shape of the principal curve can be tuned via a parameter $k \in [2, 5]$. As the goal of the isochrone extraction is to preserve as many cluster distribution characteristics as possible, k is set to the maximum value for all PC calculations.

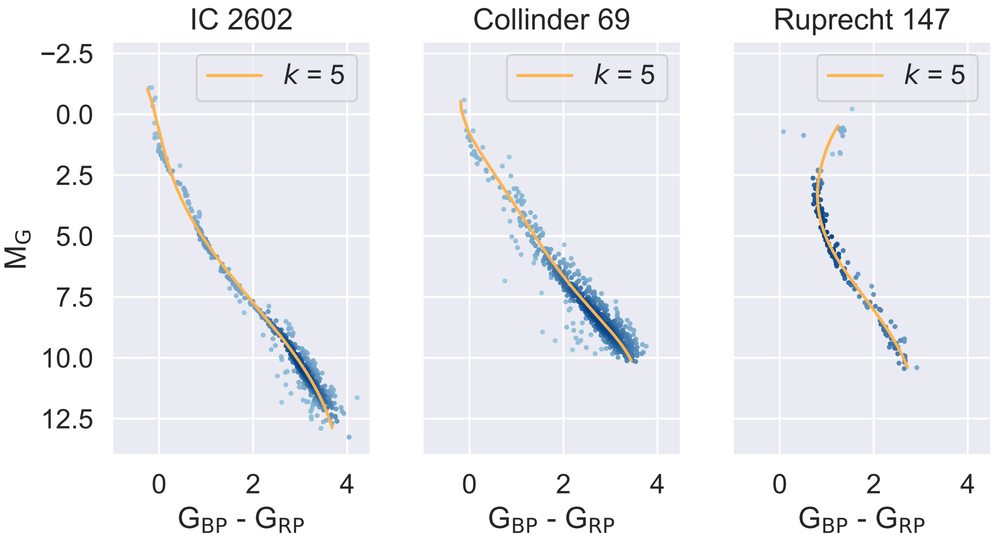

Figure 5.6.: Principal curves calculated with the `pycurve` implementation for the three different clusters IC 2602, Collinder 69 and Ruprecht 147. While the MSTO can be reproduced without problems (*right panel*), and the curve also looks good for less sharp cluster CMDs (*center panel*), sharper characteristics, like the upper-MS shape of IC 2602 are not reproduced as well (*left panel*).

Conclusion

As expected, employing the principal curves method yields smooth, well-defined curves for the cluster data. It can handle older clusters as it correctly maps the MSTO, and also works well for clusters with less concentrated CMD distributions. However, at least with the implementation of the concept that is employed in this project, the smoothing is sometimes too extensive for the desired purposes. The goal of retaining as much of the individual shape and features of a cluster distribution in the CMD contrasts with the results of the method shown in Fig. 5.6, especially for the IC 2602 cluster (*left panel*). Moreover, since the curve is defined as the *middle* of the two-dimensional data, clusters with few members and a binary sequence, or faint field star contaminants in their CMD distributions, the lower MS region of a PC isochrone will become biased. Lastly, the implementation represents the isochrones by applying splines, which opens the questions about how many nodes to choose. Using only the limited k parameter to regulate the isochrone shape is not sufficient if the isochrone should be particularly detailed for sharp cluster CMD distributions. The sum of the disadvantages of the method when applied in this work warrants the decision to not pursue it further for as a prospective isochrone extraction method.

5.3.3. Support vector regression

The theoretical concept of support vector regression (SVR) originates in statistical learning theory (Smola & Schölkopf 2004) and is based on providing a training dataset for a model, which then generates a regression curve for the whole dataset. Its benefits include a high effectiveness in high dimensional space, memory efficacy and versatility, the latter resulting from the fact that the approach is suitable for different kernel functions (see. Eq. (5.12)). The working principles of the method are described in accordance with its implementation in `scikit-learn` (Pedregosa et al. 2011). The theoretical descriptions are further based on the work of Bishop (2006), to whom interested readers are referred for a more detailed description of SVR or machine learning methods in general.

Principle

SVR differs from other regression methods primarily due to its property of sparseness, which is achieved by implementing an ε-insensitive loss function L_ε with the properties

$$L_\varepsilon = \begin{cases} 0 \quad \text{for} \quad |y_i - \hat{y}_i| \leq \varepsilon \\ |y_i - \hat{y}_i| - \varepsilon \quad \text{otherwise,} \end{cases} \tag{5.10}$$

into a regularized regression problem with target values y_i and prediction values $\hat{y}_i$. It should be noted that L_ε does not have to be linear, as long as its special characteristic of equating to zero for data points located within the so-called ε-tube around the target values is preserved. Additionally, one can explicitly define the upper and lower margins of the function via the **Slack variables** $\zeta_i^+ \geq 0$ and $\zeta_i^- \geq 0$ for all $i = 1, ...n$, which describe data points above and below the predicted regression curve, respectively. The regularized loss function that needs to be minimized for the regression can be written as

$$L = C \sum_{i=1}^{n} \left(\zeta_i^+ + \zeta_i^-\right) + \frac{1}{2}\|\boldsymbol{w}\|^2, \tag{5.11}$$

with the constant C denoting the **Slack penalty**. The weights $\boldsymbol{w}$ are included in Eq. (5.10) in the form of a norm, as they should be kept as small as possible. Using slack variables has the effect of allowing data points that lie beyond the ε-tube to affect the loss function and places the additional constraints

$$y_i - \hat{y}_i \leq \varepsilon + \zeta_i^+$$
$$\hat{y}_i - y_i \leq \varepsilon + \zeta_i^-$$

on the minimization of the loss function. A solution for the minimization problem of Eq. (5.11) can be obtained by introducing Lagrange multipliers and optimizing the resulting Lagrangian, which, after a few steps of calculation that can for example be found in Bishop (2006), leads to the formulation of an equation for predicted values

$$\hat{y}(\boldsymbol{x}) = \sum_{i=1}^{n} \left(\alpha_i^+ - \alpha_i^-\right) k(\boldsymbol{x}, \boldsymbol{x}_i) + b. \tag{5.12}$$

The equation now includes the Lagrangian multipliers $\alpha_i^+ \geq 0$ and $\alpha_i^- \geq 0$, which describe the two conditions of the previously used Slack variables, as well as a kernel function k.

As given data point can never be both above and below the regression curve, one of the α variables always equates to zero for the same index i. Support vectors are points that contribute to predictions generated by Eq. (5.12), meaning those for which either α_i^+ or α_i^- is greater than zero. The coefficient b can be either calculated using the Karush-Kuhn-Tucker conditions (for details see e.g. Bishop 2006, and references therein) or approximated in a numerical fashion.

Aims

With support vector regression, the question of how a rather sophisticated machine learning procedure performs on the archive dataset is investigated. The method produces isochrones in a straightforward manner and its implementation is comparatively easy to understand and only relies on a few essential parameters. While those parameters need to be optimized for each cluster to generate a smooth, well-placed SVR-isochrone, the search can be automated by employing a parameter gridsearch for each archive object. The SVR method is computationally very efficient, although the gridsearch may take a while to complete, depending on the extent of the investigated parameter grid. When using SVR, on can also capitalize on the fact, that for the purpose of this project, the training of the SVR model can be performed on the whole dataset of a cluster, to reproduce the shape of the CMD distribution as accurately as possible. Finally, the procedure is both flexible enough to allow a combination with PCA for handling evolved clusters, as well as computationally efficient enough to consider resampling a large number of isochrones from bootstrapped cluster data and using them to define a fluctuation region and a robust final isochrone.

Application

Due to the high computational efficiency and straightforward implementation of the method, it is applied to the entire test sample of Catalogue II. The first step is to perform a gridsearch for the best SVR hyperparameters for each individual cluster, using a 5-fold cross-validation. Next, an SVR model with the tuned parameters is fitted to the entire cluster data and an empirical isochrone is generated by predicting the response variable for each datapoint in the cluster distribution. The results of applying the SVR method to the ten clusters of Catalogue II are displayed in Fig. 5.7.

The first observation that can be made from the results, is that the SVR curves generally trace the CMD distribution of the clusters very well. Even the specific details in the upper MS regions can be preserved, despite the fact that there are generally fewer sources in those areas that can be used as training data, than in the mid and lower MS regions. In the lower MS region, the isochrones pass through the most densely populated areas of the distribution, as desired. However, while the isochrones in principle span the same dynamical range than as the corresponding cluster distributions in the CMD, it can be observed that the edge regions of the SVR-isochrones do not always match the data points in the CMD. Even more, they sometimes describe a right turn at the lower end of the MS, which does not make sense from a physical point of view, but can be explained by the increased scatter of the cluster data in the lower MS region, which renders precise isochrone extractions more difficult.

For the case of the NGC 7092 cluster (*second row, fourth panel*) however, the SVR method does not produce a good fit to the data. The bad performance of the SVR model in that particular case is likely the result of poorly chosen hyperparameters, which could have been influenced by the presence of evolved white dwarfs in the cluster data. Preliminary testing of the method on older clusters further shows that the SVR procedure breaks down when a MSTO in the cluster distribution in the CMD is encountered. This issue could be rectified by using PCA as a data pre-processing method.

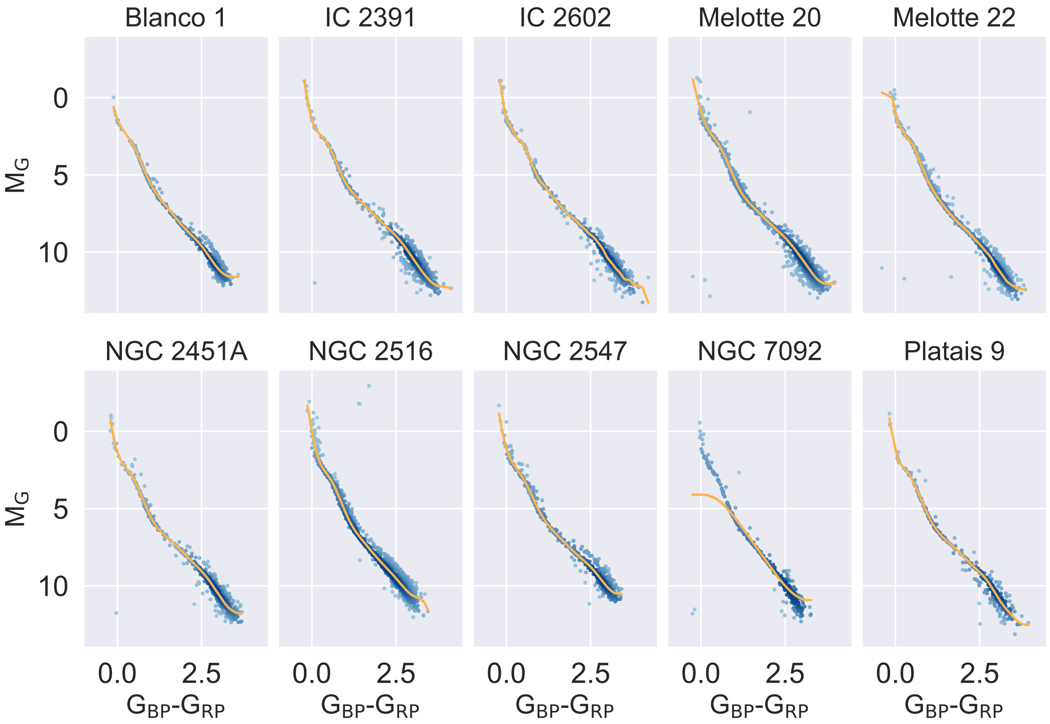

Figure 5.7.: SVR-based empirical isochrones for the clusters from Catalogue II. For most cases, the isochrone fits the distributions very well, except for the very edge regions. However, at times the method does not perform as well, as seen for example in the case of NGC 7092 (*second row, fourth panel*).

Conclusion

The implementation of SVR for the extraction of empirical isochrones works well for most clusters of the preliminary sample, as displayed in Fig. 5.7. Although for one cluster the hyperparameters were not determined correctly, it can be seen that the lines extracted for the other clusters indeed pass through the most densely populated regions in the CMDs. Overall, the method performs very well and offers the prospect of a combination with PCA for handling older clusters. Furthermore, SVR-isochrones could be resampled from bootstrapped cluster data and used to calculate robust final values, as well as uncertainty bounds. Due to the promising performance of the method and the possibility of further refinements, it is retained and developed further.

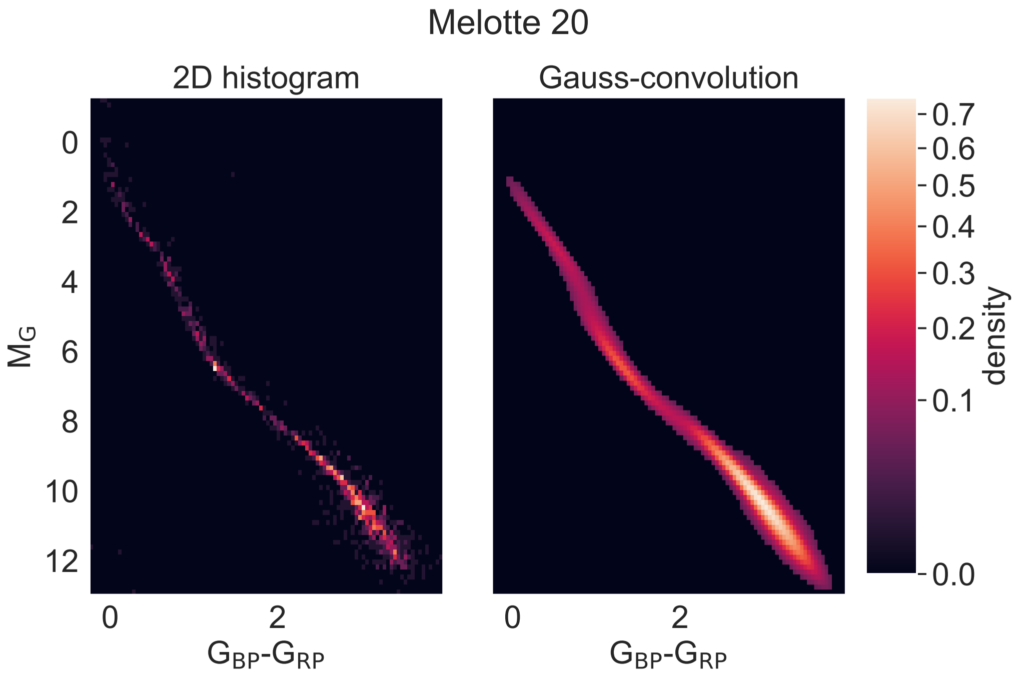

Figure 5.8.: *Left panel:* 2D histogram consisting of 100 bins of the cluster CMD of Melotte 20. *Right panel*: Gaussian convolution of the 2D histogram, after applying a density threshold value of 0.05 to the cluster density map, which causes the “cut-out” optics of the CMD distribution.

5.4. Image processing methods

The second way of creating an isochrone extraction procedure is to refrain from viewing the observations for a cluster as a data analysis problem, and to instead regard the cluster distribution in the CMD as a two-dimensional image and to extract the isochrone from it that way. The starting point for all image processing methods is a 2D histogram of the cluster data, like the one seen in the left panel of Fig. 5.8 for the cluster Melotte 20. The binned cluster data are then smoothed further via a Gaussian kernel density evaluation with a threshold value of 0.05, to generate the image displayed in the plot on the right-hand side of the figure. Both image processing methods tested in this section are available as pre-defined functions in the `scikit-image` package[20] (van der Walt et al. 2014).

It should be noted that more image processing tools were tested, than are presented in this section, the initial speculations were that the image processing methods would perform better than the data analysis approach. However, the methods ridge regression or edge detection were not applicable at all to the present project and have therefore been omitted from a more detailed method discussion.

---

[20] https://scikit-image.org/

5.4.1. Skeletonization

The skeletonization method for two-dimensional images implemented in the `scikit-image` library is based on the publication of Zhang & Suen (1984). It revolves around the concept of thinning a pattern of an image until only a "skeleton" of unitary thickness remains. Originally developed for the purpose of pattern recognition and feature extraction of alphanumeric characters and ideographs, the principle can be applied to any input image with a connected pattern. The authors also report that the distortion produced by artificial thinning of patterns is kept as low as possible.

Principle

Skeletonization employs a parallel thinning algorithm, which works iteratively and consists of two main steps. The algorithm accepts a binary matrix B of image pixels $B(i,j)$ as input. If a pixel contributes to the pattern of the image, its binary value is one, otherwise it is zero. Furthermore, each stroke of the original image pattern must exceed a thickness of one pixel. The algorithm processes different areas of the image in a parallel manner, using 3×3 pixel windows, denoted by $P_1, ..., P_9$ and removing all contour points of the binary image that do not make up the skeleton. Since the connectivity of the whole skeleton should be preserved, the thinning process takes place in two stages. First, a given contour point P_1 is removed if the conditions

$$\begin{aligned}
&2 \leq B(P_1) \leq 6 \\
&A(P_1) = 1 \\
&P_2 \cdot P_4 \cdot P_6 = 0 \\
&P_4 \cdot P_6 \cdot P_8 = 0,
\end{aligned}$$

are fulfilled. The term $A(P_1)$ denotes the number of 01-patterns in the ordered list of matrix elements $P_i,\ i = 1, ..., 9$ whereas $B(P_1)$ stands for the number of non-zero neighbors of P_1. For the second iteration step, the latter conditions are amended as follows

$$\begin{aligned}
&P_2 \cdot P_4 \cdot P_8 = 0 \\
&P_2 \cdot P_6 \cdot P_8 = 0.
\end{aligned}$$

As can be shown (see Zhang & Suen 1984, for details and illustrations), the first step removes the south-east boundary points and the north-west corner points of the pattern, provided that they do not belong to the skeleton. In a similar fashion, the second step removes the north-west boundaries and south-east corner points. The first condition preserves the endpoints of the skeleton, while the second one ensures its global connectivity (Zhang & Suen 1984).

Aims

At first glance, the method seems promising, as it provides a straightforward way to reduce a cluster CMD down to a single line at the approximate location of an isochrone. Moreover, it is computationally inexpensive and does not require extensive pre-processing of the data.

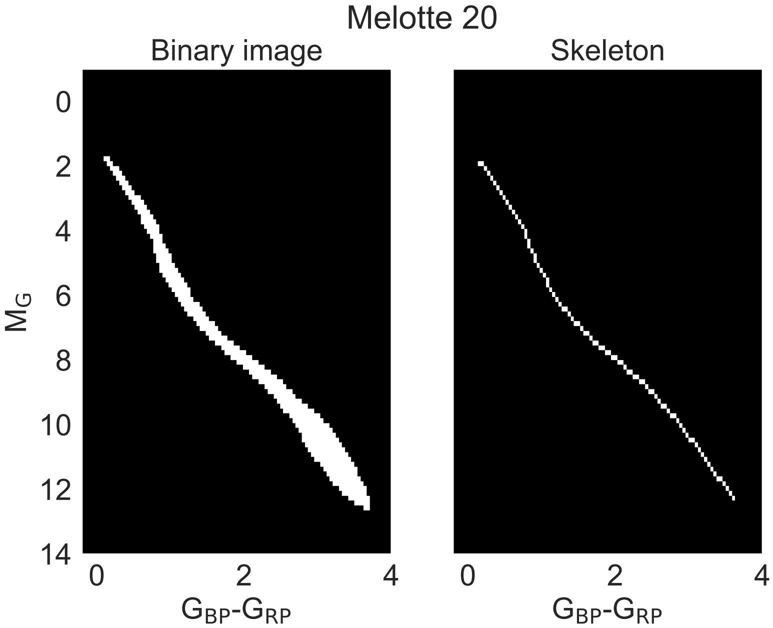

Figure 5.9.: Use-case example of applying the skeletonization method to the cluster Melotte 20. The left panel depicts a binary version of the Gauss-convolved 2D histogram of the cluster CMD, whereas the right panel shows the skeleton.

Application

To perform the skeletonization on an archive cluster, the two dimensional Gaussian convolution of its CMD (see Fig. 5.8), is transformed into a binary matrix B by assigning each image pixel $B(i, j)$ the value 1 if it is greater than zero. Figure 5.9 displays the binary version of the input image (*left panel*), as well as the resulting skeleton (*right panel*) for the cluster Melotte 20. The figure shows, that indeed a single line is produced, however, the result is merely a mask of ones and zeros that spans the input image, which makes it difficult to transform the result back into the CMD space and create an actual isochrone. Moreover, due to the working principle of the method, the isochrone is not guaranteed to pass through the densest CMD regions, but instead runs symmetrically through the middle of the data.

Conclusion

Although the method appeals through its simplicity in both idea and execution, it is too inflexible to suit the demands of the project: A big disadvantage of the skeletonization procedure is the lack of possibilities to include a weight or density map in the calculations. Since an empirical isochrone should pass through the areas of highest member density in cluster CMD, neglecting the density parameter negatively impacts the quality of the results. Furthermore, the CMD of a cluster is not globally symmetrical, which means that the extracted isochrone suffers from distortion effects mentioned in Zhang & Suen (1984). As further research into skeletonization methods that include weight maps or similar concepts turned out unsuccessful, the skeletonization algorithm is not pursued further.

5.4.2. Contour evaluation

Finding image contours is based on determining lines that run along a constant value in a two-dimensional image. The search for contours is conducted with the **marching squares** algorithm, which represents a special case of the 3D marching cubes algorithm Lorensen & Cline (1987) developed for constructing 3D iso-surfaces for medical imaging.

Principle

The marching cube algorithm works approximates the structure of 3D data by creating a grid of eight-pixel cubes over small regions of the input data. The intersection points between the straight lines of the grid and the 3D data structure create triangular shapes, which are then used to represent the structure at the intersections. For each intersection point, the algorithm determines whether it belongs to a structure in the data, by comparing the pixel value of the intersection point to a pre-defined threshold value. In this manner, an iso-surface representing the three-dimensional structure of the whole image is created (Lorensen & Cline 1987; Mantz et al. 2008). Likewise, in the 2D marching square algorithm, squares are used to determine a one-dimensional boundary of a given structure. Depending on a threshold value, for example the density of a given point in a 2D histogram, a square covering a part of the image is split into partitions that are above or below the threshold value. Interpolating between the four corners and four midpoints of the square edges yields 16 unique configurations of such partitions. Additionally, using weighted side lengths or corners allows for a more accurate form of interpolation. By connecting all the partitioned squares of the input image, one obtains a boundary line encompassing a surface with values larger than the threshold (Mantz et al. 2008). If the input structure is globally connected and the threshold value is chosen correctly, this line represents an unbroken contour.

Aims

Similarly to the skeletonization method, the contour determination is a computationally cheap way of extracting an isochrone-like shape from a given cluster CMD. Moreover, one could exploit the fact that the contour should encompass the whole structure in an unbroken line to determine not only an isochrone, but also a confidence interval surrounding the densest CMD regions. Lastly, other works using empirical isochrones, like Bouy et al. (2015), purposely place their isochrones on the lower left side of the CMD data via a manual offset. While it was ultimately decided that the empirical isochrones of the present work should pass through the densest regions of the cluster distributions in the CMDs, the contour method was originally chosen in part because its lower part could mimic an isochrone shifted to the lower left of a cluster distribution, without applying manual offsets.

Application

The input image for the contour method is a Gaussian-convolved 2D histogram, as displayed in Fig. 5.8. The contours are then determined along a user-defined threshold value r. Figure 5.10 depicts the results for four different values $r \in \{0.05, 0.1, 0.15, 0.2\}$ for the cluster Melotte 20. As can be seen, a very low threshold value yields a single contour (*first row*), which can encompass the cluster quite well but can also be very generous in the lower main sequence region, whereas a higher threshold induces the creation of multiple, unconnected contours (*second row*). Moreover, with increasing threshold values, more and more of the upper MS region, which is generally less densely populated than its fainter counterpart, is excluded from the region within the contours.

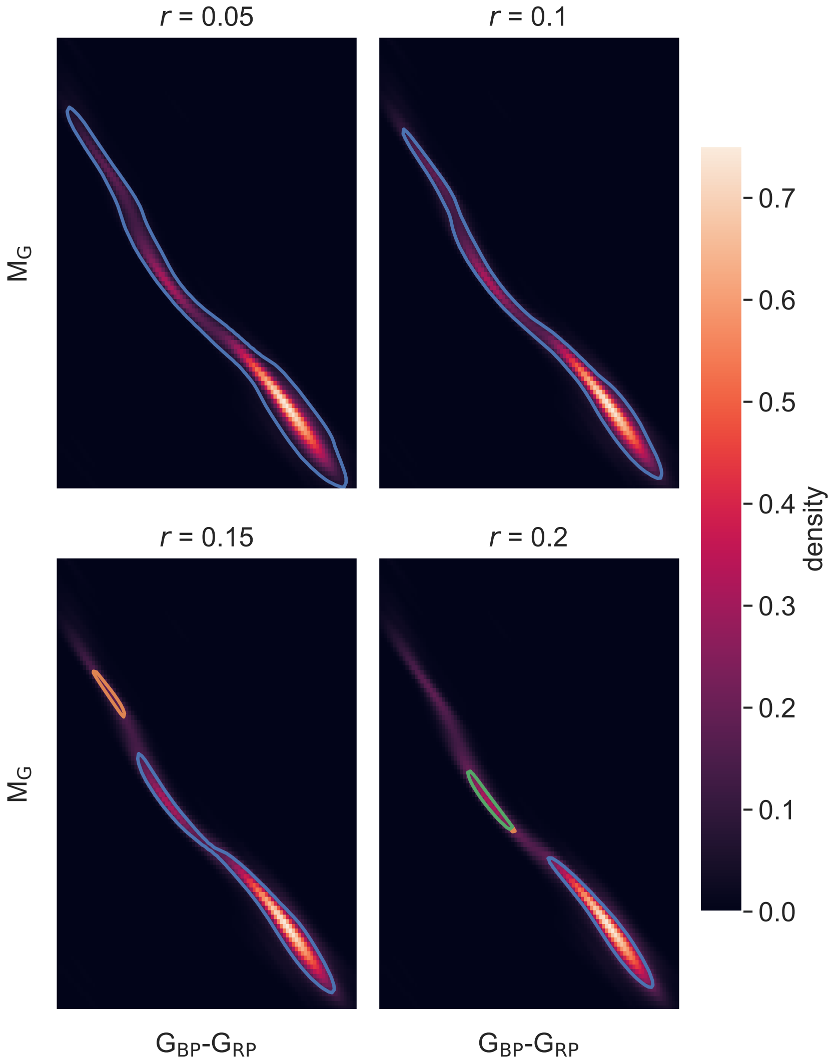

Figure 5.10.: Contours extracted for the Melotte 20 cluster for varying threshold values $r \in \{0.05, 0.1, 0.15, 0.2\}$. Increasing the threshold value leads to the appearance of multiple, disconnected contours instead of a globally connected curve.

Conclusions

As can be seen, the contour method yields mixed results. For very low threshold values r, an uneven, jagged shape around the cluster distribution in the CMD is returned, while for values larger than 0.1 in the use-case example of Fig. 5.10, the algorithm returns multiple contours with gaps in between them. Due to the stark contrasts in source density between the upper, brighter part of the CMD and the lower, fainter part, using only one contour threshold does not produce results that match the purpose of this work, which is why the method is not pursued further.

6. Results

Trying to be as open-minded as possible in the approach to both the best preparation of the cluster data, as well as in the extraction of empirical isochrones, preliminary tests have been performed with a large variety of prospective extraction procedures. Both pre-implemented algorithms, as well as newly developed ones have been taken into consideration, either on their own or coupled with data pre-processing. The preliminary results regarding the performance of each method, have been compiled throughout Chapter 5. Based on the gathered insights, it is now possible to conclude, which of the two main approaches works better for this project, before the finalized form of the isochrone extraction algorithm, and its results, are presented. The results are provided in the form of CMD plots for each archive cluster, and via summary plots where the empirical cluster isochrones are compared to age estimates found in literature, and to theoretical isochrones. Furthermore, important quality considerations with respect to the extracted empirical isochrones are mentioned.

6.1. Image processing vs. data analysis

Summing up the results of the previous Chapter 5, it can be concluded, that for the purpose of extracting empirical isochrones from open cluster CMDs, the approach of data analysis yields more favourable results. While the presented procedures originating in the field of image processing have their own merit, one could say that they are generally hard to tailor towards the specific extraction problem presented in this work. The main drawback of the skeletonization procedure is the fact, that the information about source density is lost in the binary procedure, which results in the problem that a generated skeleton isochrone merely represents the innermost points of the cluster distribution inside the CMD, in lieu of passing through the densest regions. Using contours as confidence intervals for the shape of cluster CMD distributions or as isochrones is hampered by the strongly varying densities of the different cluster CMD regions, which forces an arbitrary choice of an adequate threshold value for each individual cluster. What is more, if the density variations within a cluster CMD are high, one would either have to choose a very low threshold and include regions where no cluster members are located, as seen in the top left panel of Fig. 5.10, to retain a single contour line or one would loose the overall connectivity of the contour lines, as is the case for the plots in the second row of the figure.

Within the field of data analysis, all three presented methods yield good preliminary results. However, the percentile fitting method is ailed by the individual choice of the stepsize n for each cluster. If too few sources are represented in a bin, the empirical isochrone may smooth over individual CMD distribution characteristics of a cluster, whereas too many sources per bin result in edges and wiggles in the more densely populated CMD regions. Furthermore, the process of calculating percentiles for each bin results in a shortening of the isochrones compared to the dynamical range that the cluster members cover in the CMD. In the upper main sequence, where the cluster members are often fewer and spaced further apart, the shortening of the empirical isochrones is much more prominent than in the lower MS region (see Fig. 5.5). Even combining PCA and percentile fitting, or employing a more sophisticated source binning strategy based on kernel density estimates of the individual cluster members does not solve the issue. Principal curves can reproduce the shapes of cluster with a MSTO and are globally smooth, but are not always sensitive enough in their implementation to authentically relay the detailed characteristics of the CMD distributions of clusters with a well-sampled (upper) main sequence (Fig. 5.6). Lastly, SVR-isochrones also yield promising preliminary results, with the twin advantages of both mapping the entire dynamical range of a cluster population, and reproducing its CMD distribution in extensive detail. While the hyperparameters of SVR do need to be tailored to each cluster, the process can be automated, for example by employing a parameter gridsearch. However, the method on its own cannot handle clusters with a vertical section in the CMD, e.g. for the Melotte 22 cluster (Fig. 5.7, *first row, fifth panel*). It is also sensitive to the presence of evolved stars like white dwarfs in cluster CMDs, which can in rare cases interfere with the hyperparameter tuning, as seen for the NGC 7092 cluster (Fig. 5.7, *second row, fourth panel*). It is also not equipped to handle evolved stellar populations containing a MSTO, without further adjustments to the preliminary concept. Furthermore, it does not provide a measure of uncertainty in its original implementation. Nonetheless, of all tested methods, SVR produces the results that best fit the requirements of the project and comes with the fewest complications. For this reason, this method is chosen to be expanded, with the aims of broadening its applicability and thereby obtaining a finalized version of an isochrone extraction algorithm.

6.2. Isochrone extraction algorithm

The theoretical foundations of support vector regression have already been outlined in Section 5.3.3. Now the focus is placed on explaining the workflow of the finalized extraction procedure from a numerical, rather than a theoretical, point of view. In short, the algorithm consists of five main steps:

1. Case separation between clusters with MSTO and less evolved ones.
2. Hyperparameter tuning via a gridsearch and cross-validation.
3. Generation of an isochrone by predicting the response variable of the input data with a trained and tuned SVR model.
4. Repeating (3.) for n_{boot} observations with bootstrapped cluster data, thereby creating n_{boot} resampled isochrones.
5. Calculating the median, as well as the 5^{th} and 95^{th} percentile values from the collection of resampled isochrones.

In the following paragraphs, each step of the procedure is discussed in more detail.

6.2.1. Case separation: PCA space and CMD space

The principal drawback of the SVR method is, that similarly to the percentile fitting procedure, tracing the MSTO and giant population of older clusters in the CMD is poorly handled by the algorithm. This is because the shape of the cluster data distribution cannot be described with a functional dependency $f(x)$ in the CMD space for evolved clusters. An interconnected problem surfaces for clusters with virtually vertical member distributions for the upper MS region of the CMD, e.g. the Melotte 22 cluster in Fig. 5.7. The SVR algorithm usually recognizes the data points corresponding to evolved stars in the giant or white dwarf regions as outliers and therefore ignores them during the regression. In some cases, like for the NGC 7092 cluster in Fig. 5.7, the presence of evolved stars can even disturb the hyperparameter tuning process. For this particular example, investigations showed that the bad fit between the isochrone and the cluster data was caused by the presence of multiple white dwarfs in the cluster CMD. To generalize the extraction algorithm over a large range of cluster ages and morphologies, the handling of such exceptions needs to be incorporated into the workflow.

To this end, a principal component analysis (PCA), as presented in Section 5.2.2, is performed on clusters for which the nominal SVR routine yields poor fits. By rotating the data into the direction of the largest variance contribution, the two input variables retain a functional relationship in PCA space. An isochrone can still be extracted, by performing the training and fitting of the SVR model using the PCA transformed data, and afterwards performing an inverse transformation back into the CMD space. The PCA approach produces good results and captures the general curve of the isochrone into the direction of the giant region for evolved clusters. However, it should be emphasized that the uncertainties associated with the extracted isochrone in the region beyond the MSTO are generally much larger than those corresponding to the rest of the isochrone, since they are generally very sparsely populated. Regarding the numerical implementation of the PCA and CMD case distinction, the algorithm works identical for both cases. Only the contents of the input arrays for the gridsearch and SVR model generation, which for the PCA case consist of the already transformed data, vary, and for PCA cases, the SVR regression requires an inverse transform to yield an isochrone.

Why is the PCA transformation not employed for all clusters? Separating between the two possible extraction methods based on the morphology of a cluster's source distribution in the CMD inadvertently subtracts from the generality of the extraction method and makes it more subjective. Clusters that need to be pre-processed with PCA are also required to be flagged for the extraction algorithm. An arguably better option would be to perform PCA for all clusters of the archive, disregarding their evolutionary state or CMD distribution morphology, and to extract the isochrones for all of them in a unified fashion. In practice however, this approach includes caveats that impact the quality of individual extracted isochrones, especially concerning the edge regions. The reason behind this observation can be found by looking at the cluster shapes in PCA space. One can see that the isochrones can be roughly described as sinusoidal shapes (see e.g. Fig. 5.3). While for evolved clusters, the data points from the giant region usually align with one of the sinusoidal oscillation periods, this is not generally the case for clusters without a MSTO or giant region. Here, the sources describing the upper end of the MS can have a much larger variance compared to the rest of the data, than they have in the CMD space. As a result, they may not be recognized as supposed extension of the SVR regression curve in individual cases. As a result, the final isochrones can point away to the left of the upper MS, instead of fully tracing the member star distribution for a number of young OCs, when transforming them into PCA space and extracting the isochrones there. The central parts of the CMD are still a good fit, but more information about the CMD shape is reliably retained and represented in isochrones derived straight from the CMD space for most young clusters in the selection. As a result, the case separation between clusters is determined to be a more sensitive choice with respect to the quality and reliability of the results. It should be noted, that for evolved clusters it is still possible to calculate isochrones directly in the CMD space, provided the outliers do not interfere with the hyperparameter tuning, else they need to be excluded from the gridsearch data. In those cases the empirical isochrones will simply be shorter and incomplete in the giant region.

6.2.2. Hyperparameter tuning

The hyperparameter tuning is conducted by first splitting the two-dimensional input array (X, Y) of the CMD or PCA data of a given cluster into a training $(X_{\text{train}}, Y_{\text{train}})$ and a test set $(X_{\text{test}}, Y_{\text{test}})$, respectively, and subsequently scaling the X_{train} variable. Next, a five-fold cross-validation is defined, in which the training data is randomly split into five parts and the SVR model is trained on four different folds and validated on the fifth for one full circuit of possible fold permutations. To determine the best set of SVR hyperparameters for each cluster, a gridsearch is performed using the `scikit-learn` (Pedregosa et al. 2011) function `GridsearchCV`[21], which takes the SVR method, the CMD or PCA data, a parameter grid and the chosen cross-validation method as input parameters. Concerning the `SVR` function itself, there are three parameters relevant for the radial basis function (rbf), which is employed as kernel function for the SVR algorithm in this project. They are described in detail in the `scikit-learn` documentation of the function[22], and can be summed up as follows:

- `gamma`: The parameter that is used as a kernel coefficient for the rbf function[23]. On the example of a dataset, the value of this parameter sets the influence radius of an individual data point in an inversely proportional way. A low value of `gamma` results in a far-reaching influence radius, meaning that points that are spread further apart are recognized as one group, whereas large values of the parameter indicate a low radius of influence. Choosing the parameter too large encourages overfitting of the data, while a too small value results in a loss of complexity of the shape of the regression curve, as it becomes increasingly smooth.

- `C`: The penalty parameter represents the regularization of the algorithm, which weighs correct sample classifications against the size of the margin of the decision function. A large value indicates a high penalty for data misclassifications, whereas a low value allows for more missclassified data points. The `gamma` and `C` variables influence one another, for example, smoother models with small `gamma` values can be increased in their complexity by placing more importance on classifying data points correctly by increasing `C`.

- `epsilon`: The parameter defining the size of a tube around an actual sample value, in which a predicted value can fall without accruing a penalty in the training loss function.

---

[21] https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

[22] See https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html, and https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html

[23] The parameter is also applicable for the kernel types “sigmoid” and “polynomial”, see https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html

With respect to the cluster data investigated in this work, the parameter ranges for the gridsearch are chosen from a logarithmic scale ranging from `C` $\in [0.1, 10000]$, `gamma` $\in [0.00001, 100]$, and `epsilon` $\in [0.00001, 0.1]$. The final hyperparameters for each cluster are determined by querying the best score on the test data, which in the case of the five-fold cross-validation represents an average of all five scores on the validation data. The standard scoring method for regression models in the `GridsearchCV` function is the R^2 score[24]. Once the hyperparameters have been calculated, they can be stored for each cluster and do not have to be determined again for the creation of resampled isochrones based on bootstrapped data.

6.2.3. Generation of robust SVR-isochrones and fluctuation regions

Once a set of tuned hyperparameters has been established for a cluster, a `SVR` model is generated with the specified parameters and fitted to the entire cluster CMD or PCA data, thereby training it on the whole dataset. Subsequently, the trained model is used to predict the response variable, i.e. the absolute magnitudes or PCA Y data, yielding a regression curve that represents a single empirical SVR-isochrone of the cluster, as already shown in Fig. 5.7. In the case of fitting the SVR model to the PCA data of a cluster, it is necessary to transform the regression curve back into the real space to obtain the empirical isochrone.

However, while a single isochrone with well-tuned hyperparameters already presents a good result, it would be beneficial to create a more robust measure, as well as taking into account possible fluctuations of the different isochrone shapes, given varying cluster member selections. For this reason, instead of calculating a single SVR-isochrone, the original cluster data is bootstrapped, meaning drawing and replacing data points, to create a slightly different input array for training a new SVR model. The bootstrapped arrays have the same number of entries as the original ones, however, individual cluster members may not be present in the sample, while others appear more than once. The bootstrapped data is then used to train new SVR models, before the response variable is once again predicted to create a resampled isochrone. By always predicting values for the original array of explanatory variables, it can be ensured that all resampled isochrones share the same X positions in PCA or CMD space. After generating isochrones in this manner for a pre-defined number of observations, n_{boot}, typically either 100 or 1000, the median, as well as the 5^{th} and 9^{th} percentile values are calculated from the collection of resampled isochrones for each X position of the cluster data. In doing so, a robust final isochrone is derived, together with a region encompassing 90 per cent of the fluctuation in the resampled isochrones for each cluster. A depiction of 100 bootstrapped isochrones (*thin orange lines*) together with a reference isochrone derived from the whole cluster data (*bold orange line*) can be seen in the left panels of the plots in Figure 6.1, for the NGC 7092 and the Ruprecht 147 clusters. The first of the two clusters was not subjected to data pre-processing via PCA, whereas for the second cluster it was necessary, as it displaces a MSTO in its CMD distribution. In the right panels of the plots, the empirical isochrones derived from the median values of the 100 resampled isochrones are depicted in red, and the 90 % fluctuation region by the grey 5^{th} and 9^{th} percentile boundaries.

---

[24] See https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html

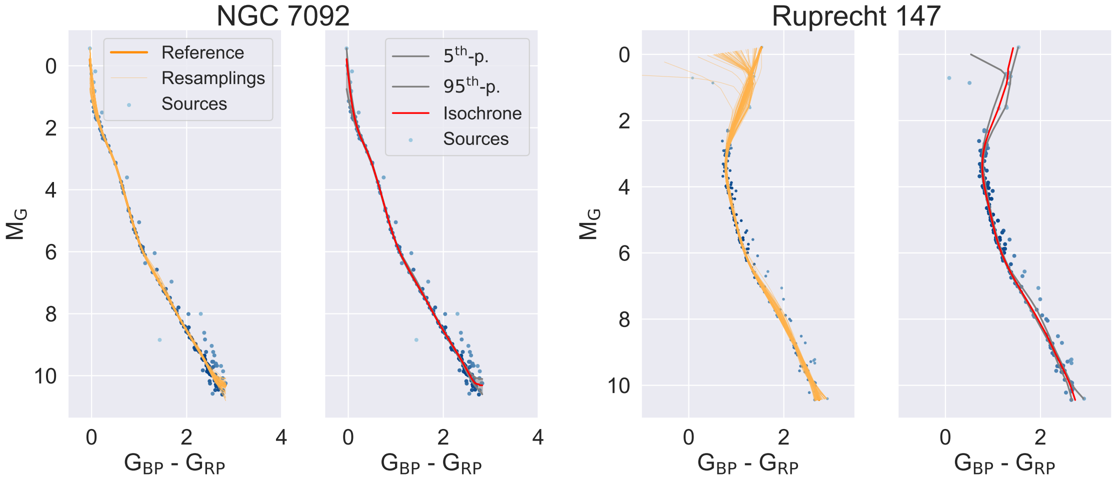

Figure 6.1.: Clusters with empirical isochrones derived via SVR and bootstrapping of the cluster data to generate 100 resampled isochrones. The isochrones for Ruprecht 147 (*right panel*) are calculated in PCA space, whereas the ones corresponding to NGC 7092 (*left panel*) are directly extracted from the CMD space.

6.2.4. Influence of measurement errors

It should be noted, that the procedure at this stage in its development does not incorporate the measurement errors associated with the observational data, which are commonly in the mmag and milli-arcsecond range for *Gaia* EDR3. However, to get a better understanding of the possible impact of measurement errors on the results of the isochrone extraction algorithm, one can for example study the absolute errors in the CMD data on a set of various archive clusters. As has been established in Section 2.1.1, the CMD variables are both compounds of two mutually independent observational values per star, and are defined via Eq. (2.6) and Eq. (2.1), respectively. This means, that they are subjected to error propagation. As the measurement errors in the *Gaia* denote uncertainty bounds of the measured values, rather than absolute numerical errors that are known including their sign, their influence on the absolute magnitude and color index can for example be estimated by using the Gaussian formula of error propagation, which for a general expression $y(x_1, x_2)$ with the measured uncertainties Δx_1 and Δx_2 can be written as follows

$$\Delta y = \sqrt{\left(\frac{\partial y}{\partial x_1} \cdot \Delta x_1\right)^2 + \left(\frac{\partial y}{\partial x_2} \cdot \Delta x_2\right)^2}. \tag{6.1}$$

The overall uncertainty of the variable y is denoted as Δy. As can be seen from the equation, the error calculation requires the partial derivatives of Eq. (2.6) and Eq. (2.1). Performing the derivations of both equations with regard to the magnitude values equates to 1, however, the derivation $\frac{\partial M}{\partial d}$ is a bit more complicated. To calculate it, the distance variable d needs to be expressed in terms of parallaxes ϖ, as those are the actual measurements provided by *Gaia* and also the parameters for which the errors $\Delta\varpi$ are provided. In the case of small measurement errors, and for the purpose of this work, the assumption that the distances correspond to the inverted parallax values, is reasonable. As parallaxes are further provided in units of milli-arcseconds, they have to be multiplied by a factor of 10^{-3} to arrive at distance measurements in units of pc. Applying the changes to Eq. (2.1) results in the following expression, which can then be derived with respect to the parallax variable

$$
\begin{aligned}
M &= m - 5\log_{10}(d) + 5 && (6.2)\\
&= m - 5\log_{10}\left(\frac{1}{\varpi}\right) + 5 && (6.3)\\
\frac{\partial M}{\partial \varpi} &= \frac{5}{\log(10)\cdot\varpi}. && (6.4)
\end{aligned}
$$

Plugging the expressions for the partial derivatives into Eq. (6.1) for both CMD variables yields

$$
\begin{aligned}
\Delta(m_X - m_Y) &= \sqrt{(\Delta m_X)^2 + (\Delta m_Y)^2} && (6.5)\\
\Delta M &= \sqrt{\left(\frac{5}{\log(10)\cdot\varpi}\cdot\Delta\varpi\right)^2 + (\Delta m)^2} && (6.6)
\end{aligned}
$$

In the case of EDR3 data, the errors in the measured magnitudes of the three passbands are already provided in the source catalogue, while for DR2 data, the error for an arbitrary magnitude m_i needs to be calculated from the corresponding measured flux F_i and its corresponding uncertainty ΔF_i via

$$\Delta m_i = \frac{2.5}{\log(10)}\cdot\frac{\Delta F_i}{F_i}.$$

Figure 6.2 displays the visualization of the calculated measurement errors for the Blanco 1 cluster of Catalogue II. Similar plots for the remainder of the Catalogue II selection can be found in Appendix B in figures B.1 to B.5. The left panel of the figure depicts errorbars created by dividing the cluster members into 12 equidistant bins, defined between the minimum and maximum absolute magnitudes of the cluster distribution, and calculating the mean errors in both variables for each bin. The errorbars have been shifted to the left side of the member distribution in the plot. In the right panel of the figure, the errors have been multiplied by a factor of 5 to enhance their visibility and to qualitatively highlight the trend of increasing measurement uncertainty towards the fainter end of the CMD distribution of the cluster. As can be seen across all clusters of Catalogue II, the errors are very small compared to the scatter in the cluster data. As a result, it can reasonably be assumed, that measurement errors in the *Gaia* data of the archive clusters would hardly influence the shapes of the empirical isochrones, especially in the upper part of the cluster MS. However, as the `SVR` implementation of `scikit-learn` (Pedregosa et al. 2011) could in principle take a weight parameter for each data point as input parameter, a future refinement of the extraction algorithm could entail the transformation of measurement errors into weights that can be directly be passed to the SVR method.

Figure 6.2.: Binned error data for the cluster CMD of Blanco 1. The center and right panel of the figure displays a five-fold magnification of the error bars.

6.3. CMD and summary plots

Figure 6.3 shows the final results of the extraction procedure on the sample clusters with selections of Catalogue II, calculated with $n_{\text{boot}} = 1000$. Of all the clusters in the depicted selection, the Melotte 22, NGC 2516 and the NGC 7092 clusters are calculated by applying a PCA transformation, as it yields better results for the specific CMD distributions of these clusters. The empirical isochrones and uncertainty boundaries ($n_{\text{boot}} = 1000$) for Catalogue I clusters are depicted in the cluster CMDs in Appendix A, throughout figures A.1 to A.8.

In a subsequent analysis, the empirically derived cluster isochrones ($n_{\text{boot}} = 1000$) are compared to estimated cluster ages found in literature. The publications employed for this purpose are the Bossini et al. (2019) catalogue, which provides age values for 41 of the 67 archive clusters, as well as the Dias et al. (2021) catalogue, in which 64 clusters of the archive selection have listed ages. The respective summary plots depicting the empirical isochrones, color coded according to the estimated ages in logarithmic values, are shown in Figure 6.4. It should be noted that the colorbars next to the respective figures display different age ranges.

6.3.1. General observations of the age comparisons

Generally, a good agreement between the extracted isochrones and the ages found in literature can be seen. The supposedly youngest clusters in the archive do indeed display isochrones that are shifted towards the upper right in the CMD, compared to the isochrones

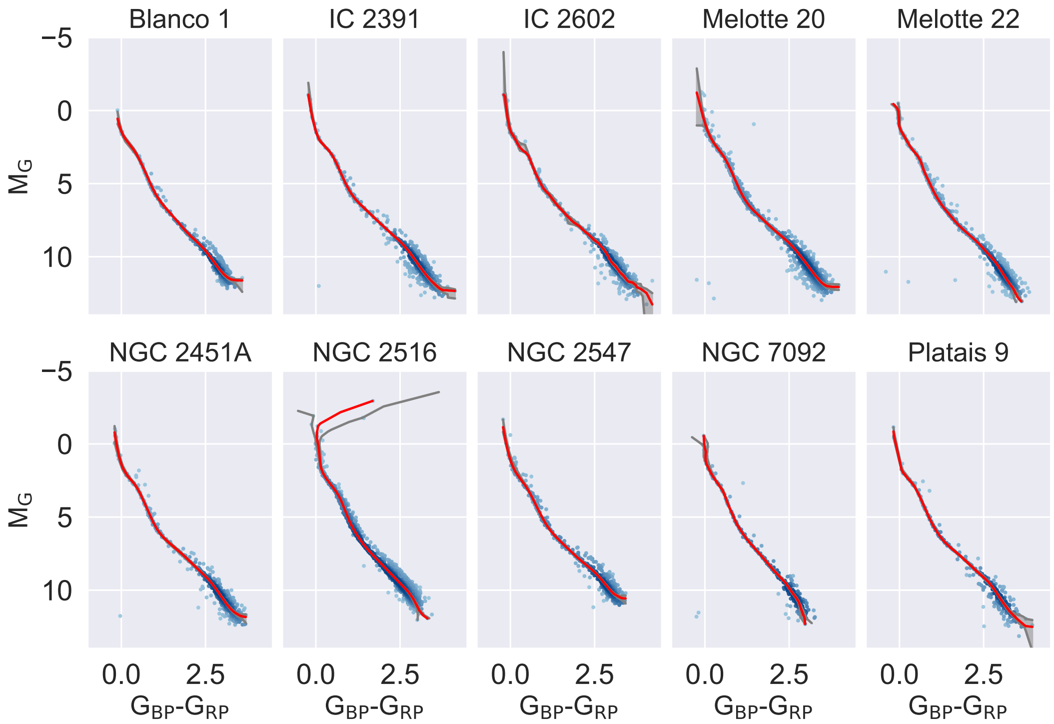

Figure 6.3.: Empirical SVR-isochrones based on 1000 bootstrapped isochrones for each of the clusters of Catalogue II. The isochrones are denoted by the red lines, whereas the grey lines and shaded areas indicate the uncertainty regions ($5^{\text{th}}/95^{\text{th}}$ percentiles).

of other clusters, and even the knee-like bend in the upper main sequence is visible for some examples. It can best be observed for the cluster IC 348, which is displayed in Fig. A.3 (*right panel*), although according to literature, L 1641S is the youngest sample in the archive selection (Da Rio et al. 2017). The cases of diffuse cluster CMDs like L 1641S are discussed in more detail in Section 6.4.

Fig. 6.4 also illustrates the downward shift of the MSTO position as clusters grow older. Notably, in the right panel of the figure, a single empirical isochrone with a MSTO can be distinguished, which is connected to a much younger estimated age than the others in its vicinity. Closer investigation reveals that the cluster in question is the Platais 8 cluster, which in the Catalogue II selection has one member star at the approximate location of the red clump region. The difference to the estimated age and expected isochrone shape for that age could suggest, that this source might be a field star contaminant instead of a member star. Overall, the agreement between the estimated ages and the empirical isochrones seems to be slightly worse for the clusters depicted in the right panel of the figure, with more empirical isochrones deviating from the expected color gradient than for the clusters depicted in the left panel of the figure. In the case of the comparison to the Bossini et al. (2019) catalogue, there seem to be three distinct regions of cluster ages indicated by the lower MS of the isochrones in the figure. The gap between the very young clusters and

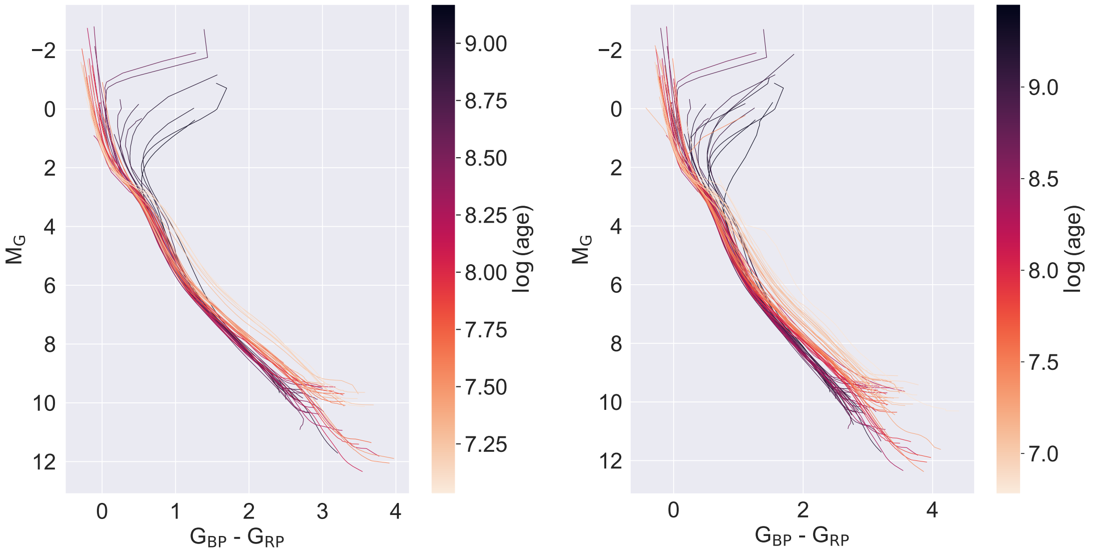

Figure 6.4.: Summary isochrone plots, colored according to cluster ages found in Bossini et al. (2019) (*left panel*) and Dias et al. (2001) (*right panel*). The colorbars represent the logarithmic ages of the clusters.

the intermediate age clusters is more pronounced than the one between the intermediate and old clusters. This agrees well with the theoretical expectations, as the lower MS is quite distinct for young clusters. The plots in Fig. 6.4 depict a mutual verification between cluster ages and empirical isochrones. The strong color gradient on the one hand shows, that apart from possible systematic errors, model-dependent cluster ages from isochrone fitting can indeed represent the stellar evolution of clusters very well. On the other hand, it proves that the archive clusters are really a collection of isochrones that can be sorted and used in terms of a ranking, when applied to other clusters or groups, simply based on the positions of the isochrones in the CMD. What the plot does not show, however, is the actual agreement between the empirical isochrones and their theoretical counterparts for individual clusters. Such a comparison is presented in Chapter 7.

6.3.2. Isochrone blindspot and empirical isochrones

Another interesting comparison that can be made using the extracted empirical isochrones, concerns the blindspot that was defined over a large dynamical range for theoretical isochrones between $100 - 500$ Myr in Section 3.4. For reference, the color-coded theoretical isochrones within this age range from Fig. 3.4 are again displayed in the left panel of Fig. 6.5. In the two panels next to it, a selection of empirical cluster isochrones is placed over the grey outlines of the theoretical isochrones. The center panel corresponds to clusters with an estimated age between $100 - 500$ Myr in the Bossini et al. (2019) catalogue, and the right panel shows the same for the catalogue of Dias et al. (2021).

The plots allow the observation, that the region where the theoretical isochrones with default extinction and metallicity values display the isochrone blindspot is not as narrow

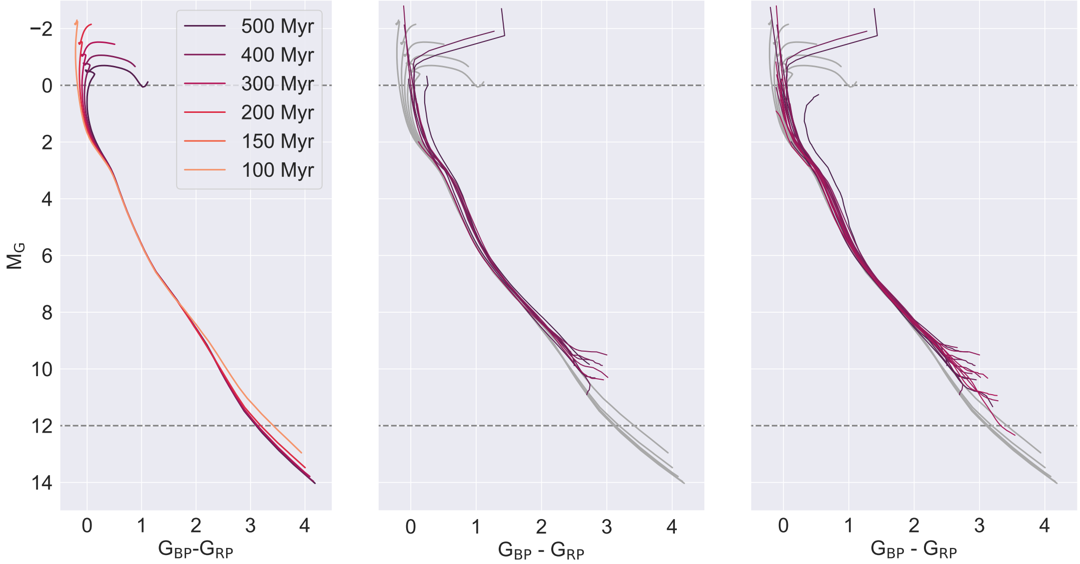

Figure 6.5.: Comparisons between empirical isochrones corresponding to estimated cluster ages between 100 − 500 Myr in the publications of Bossini et al. (2019) (*center panel*) and Dias et al. (2021) (*right panel*), respectively, and the isochrone blindspot (*left panel and grey lines*).

and more individual when traced by empirical isochrones. A possible explanation for this phenomenon could be, that the broadening is caused by the influence of extinction and metallicity values for each individual cluster, thereby again highlighting the extremely high dependence of the model-dependent isochrone fitting method on these two input parameters.

6.4. Quality of empirical isochrones

As can be seen from the cluster CMDs in Appendix A, the SVR extraction method produces very accurate reproductions of the shapes of the cluster distributions in their respective CMDs. In general, there exists little fluctuation between the various resampled isochrones and therefore the indicated uncertainty regions are very narrow. The results are an indication of a successful hyperparameter tuning, and the general good applicability of the machine learning procedure to the extraction task. Especially the central part of the cluster main sequences is exceptionally narrow and well-defined almost all archive clusters. By training each new SVR model on bootstrapped cluster data, but creating the empirical isochrones by predicting the entire original CMD data range, the uncertainty boundaries appear largely smooth and without edges or sharp angles.

Nonetheless, there are a few quality aspects that need to be addressed with respect to the empirical SVR-isochrones. Firstly, the extraction method is very sensitive to the “sharpness” of the CMD distribution of a cluster population. If the main sequence is well populated and there are no stark density contrasts between different regions, the method returns almost imperceptible uncertainties in the results. This is the case for most clusters without a MSTO. It is only when the number of cluster members in a certain CMD region drops dramatically compared to the rest of the distribution, that problems may arise. Cases where the isochrone quality and significance may be reduced include edge regions or gaps in the MS, or very diffuse clusters.

6.4.1. Edge regions

If a significant drop in member density can be observed for one or more regions in the CMD of a cluster, or if a region is particularly diffuse, the fluctuations in the shapes of the resampled isochrones can become quite large. The former is most likely to happen at the upper end of the MS, or in the giant region and MSTO for older clusters, while the latter is more relevant for the faint end of the population, and can cause the isochrone to describe a much flatter turn than expected from theoretical isochrones. As a result, the edge regions of the empirical isochrones may be subjected to larger uncertainties than the rest of the isochrone and should be treated with caution, if need be. Two examples of clusters with diverging isochrone edge regions are for example the younger cluster Collinder 140, and the more evolved cluster NGC 2632, which are displayed in Fig. 6.6. For the former (*left panel*), large uncertainty bounds in the area of the lower MS were calculated, whereas for the latter (*right panel*) the uncertainty bounds of the upper end of the isochrone strongly diverge from the empirical isochrone for the upper MS.

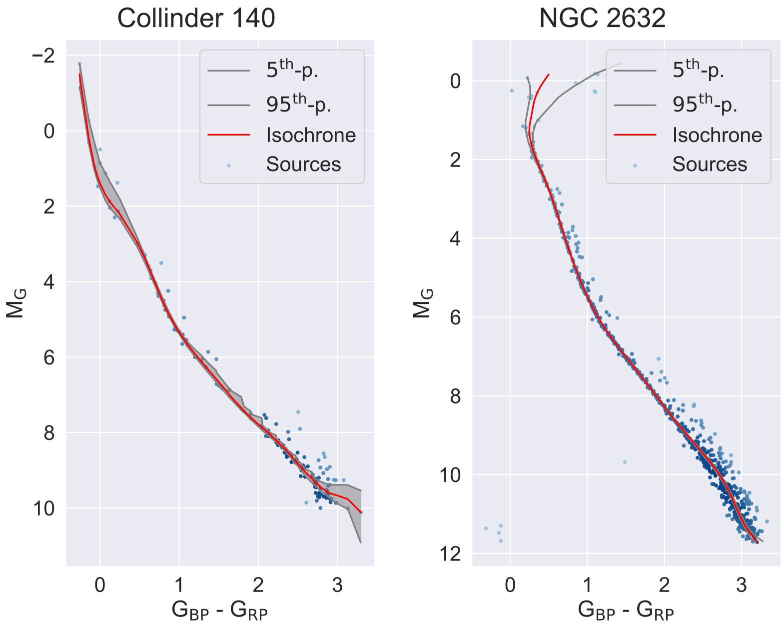

Figure 6.6.: Examples of uncertainties and non-physical shapes in the edge regions of empirical isochrones, using the CMDs of the clusters Collinder 140 (*left panel*) and NGC 2632 (*right panel*).

The edge regions diverge more strongly for evolved clusters with a MSTO than for their younger counterparts. The members populating the giant region are often very few and are at a large separation to the bulk of the cluster data, which means they are less likely sampled in the bootstrapping process, or can even be determined as outliers to the regression line of the SVR model. Therefore it is important to underline, that in the case of evolved clusters, the upper part of the empirical isochrones should be treated more like a general trend, than like an accurate reproduction of the true cluster CMD. Not only are the regions too poorly represented by cluster members to provide a robust, reliable isochrone with the extraction method developed in this project, but the regions are also vulnerable towards field star contaminants for the same reason. What is more, no empirical isochrone can reproduce the intricate shapes or curves of a theoretical one in the area of the MSTO and beyond, as the quality and quantity of observational data in these regions simply do not allow it as of today. The diverging edge regions are also visible in the form of sharp spikes or turns in the summary plot of Fig. 6.4.

6.4.2. Contaminants and binary sequence

Overall, the extraction method is very robust against outliers, such as field contaminants that fall above or below the MS, or binaries. As long as the outliers do not outnumber the actual cluster members, they do not negatively impact the isochrone shape. For example, clusters like NGC 3532 or RSG 8, shown in Fig. 6.7 display a sequence of contaminating field stars and a large binary population, respectively, and yet the calculated isochrones are not influenced by their presence. There are two exceptions in which the robustness of the method can be compromised however: Clusters with a MSTO and very diffuse clusters.

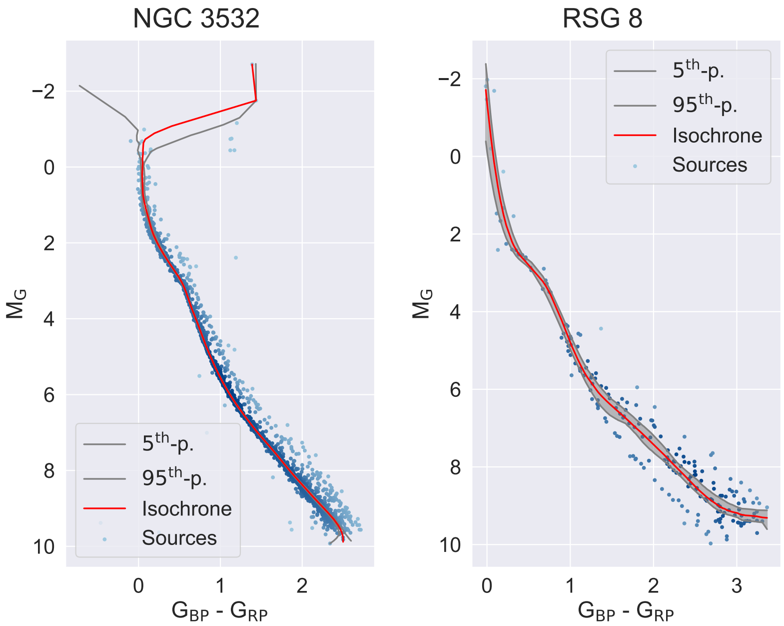

Figure 6.7.: Illustration of how extracted isochrones are influenced by data contamination like unresolved binaries (*left panel*) or unwanted field star populations (*right panel*) in the cluster data. As can be seen, neither effect influences the quality of the empirical isochrones for NGC 3532 and RSG 8.

6.4.3. Very diffuse clusters

If a cluster population appears strongly scattered in the CMD, the uncertainty regions begin to grow along all parts of the isochrone, indicating that the general quality of the fit to the data decreases as the variance in the data increases. Fig. 6.8 shows three extreme cases of very diffuse clusters, namely IC 348 (*left panel*), L 1641S (*center panel*), and RSG 7 (*right panel*).

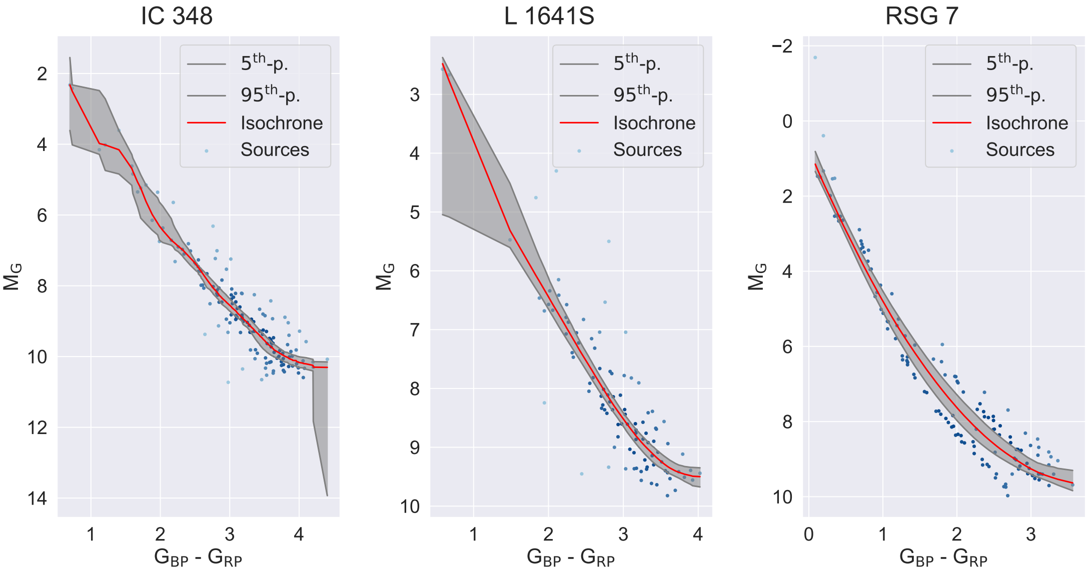

Figure 6.8.: CMDs and empirical isochrones for the very diffuse clusters IC 348, L 1614S, and RSG 7. For neither cluster a good fit between the data and the SVR-isochrone can be produced. The former two clusters are both very young, while the latter has probably a highly contaminated member selection.

In all three cases, the quality of the extracted isochrones is seriously hampered by the strong scattering of the sources in the CMD. For the first of these clusters, a very large uncertainty area and a numerically induced spike at its lower edge can be observed. While the isochrone shows a knee-like feature that is common for very young clusters like IC 348 ($\sim$ 6 Myr according to Dias et al. 2021), the isochrone should still be regarded with caution. In the cases of the two other clusters, the isochrones display almost no characteristic features and do not pass through the bulk of the member stars. Incidentally, IC 348 and L 1641S represent the youngest clusters in the isochrone archive (Dias et al. 2021; Da Rio et al. 2017), from which one can draw the conclusion, that the scattering is probably caused by high extinction values, as young clusters can often be shrouded by dust clouds (e.g. Lada & Lada 2003). As already expected in Chapter 4, the scatter of the CMD data of young clusters does indeed place a lower age boundary on the reliability of the extracted empirical isochrones. The cluster RSG 7 on the other hand, is estimated to be $\sim$ 84 Myr old (Dias et al. 2021), so it can be assumed that the bad fit may be the result of a strong field star contamination below the MS of its member distribution. While this is the same effect that has already been observed in the CMD of RSG 8 (Fig. 6.7), for RSG 7 the presence of such a large number of contaminants, combined with the low absolute member number and the comparatively high scatter in the data, results in a poorly placed, featureless empirical isochrone.

7. Discussion

With the archive of empirical isochrones now in existence, there are two important points of discussion that still have not been addressed. First the practical relevance of the archive is discussed. Afterwards, two case studies of clusters of special scientific interest, and with largely diverging age estimates in literature, are analyzed with the help of empirical isochrones.

7.1. Relevance of empirical isochrones

The value of empirical isochrones has already been recognized within the scientific community over the past few years. For example, Bouy et al. (2015); Olivares et al. (2019); Miret-Roig et al. (2019) provide a selection of empirical isochrones for the three clusters Melotte 22, Ruprecht 147 and IC 4665 in the u, g, r, i, Y, J, H, K_s passbands of the DANCe survey. Similarly to the present work, they acknowledge the added value of using empirical isochrones based on solid member selections and high-quality observational data as reference points for other groups or clusters. As has already been noted, their extraction method is based on principal curves, albeit using a different implementation than the one presented in this work. They also state, that empirical isochrones can be used to refine stellar evolution models, when they are viewed as calibration data. All three works also compare the empirically derived isochrones to model isochrones with values found in contemporary literature to gauge their quality and relevance.

Other authors employing empirical isochrones, albeit in a different context, are for example Mayne et al. (2007), who provide relative ages for young pre-MS stars, using a spline-based isochrone determination method. Additionally, the importance of a self-consistent, model-independent approach to age determination is strongly emphasized and referred to as highly relevant in Soderblom (2010).

7.2. Case studies

With the empirical isochrones of the archive now readily available for use, this section is dedicated to two case studies of two clusters that are especially interesting objects. The aims of the cluster studies, are not only to compare the empirically derived isochrones and uncertainty regions to isochrones based on various published values from the literature, but to also provide a more qualitative discussion on the general limits of the cluster ages and the difficulties of using isochrone fitting. The clusters to be investigated are NGC 2516 and Melotte 22, which both have large variations within their estimated ages in literature (see Meingast et al. 2021, and references therein). It should be mentioned however, that many of the cluster ages are not derived from *Gaia* DR2 or EDR3 data, and extinction values are often given in units of different passbands, both of which makes it more difficult to compare to the empirical isochrones of the archive to published cluster parameters.

Therefore, the cluster isochrones depicted in the following section once again correspond to the catalogues of Bossini et al. (2019) and Dias et al. (2021), as the two already show large disagreements between the cluster age estimates, and it is possible to make significant comparisons between them. To also ensure a fair comparison between the two theoretical isochrones and the empirical ones, the latter ones are also derived from DR 2 data in the context of the case studies.

7.2.1. Case study I: NGC 2516

Age estimates for the NGC 2516 cluster can be found across the broad range of ∼6 Myr Tadross et al. (2002) to ∼299 Myr (Kharchenko et al. 2013). The isochrones displayed in Fig. 7.1 correspond to the values 123 Myr and 277 Myr of the Bossini et al. (2019); Dias et al. (2021) catalogues, which disagree by an absolute value of 154 Myr, or a staggering $56 - 125$ per cent. The empirical isochrone and uncertainty bounds shown in the figure are based on the data of $n_{\text{boot}} = 100$ resampled cluster isochrones. In general, the isochrone extraction was performed in PCA space.

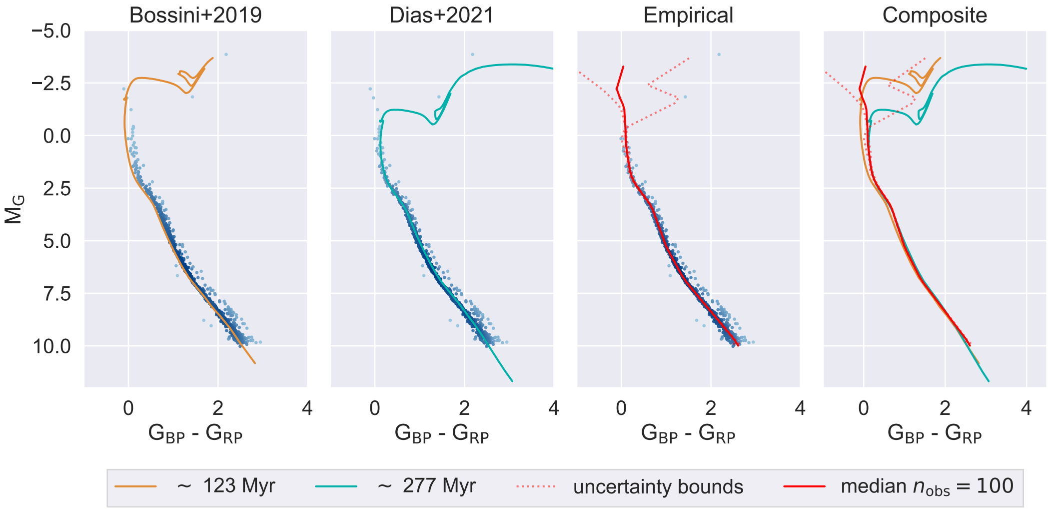

Figure 7.1.: Comparison between theoretical and empirical isochrones regarding the NGC 2516 cluster. The first three plots depict the cluster population according to Cantat-Gaudin & Anders (2020) ($p \geq 0.5$) along with either a model or an empirical isochrone, the latter including the calculated 5th and 95th percentile bounds. The panel on the far right of the figure displays a composite plot containing all three isochrones.

As can be seen from the point distribution of cluster members, there are two sources populating the red giant region, which are fitted by both theoretical isochrones. The isochrone in the second panel furthermore neglects a few sources that seem to be a part of the MS, and instead traces an earlier MSTO and an associated older age in the CMD. Neither of the two curves provides a good fit to the entire member selection of the cluster. The empirical isochrone aligns with both theoretical curves along the MS, up until ∼ 4 mag. Here, the differences between the two model isochrones also begin. Interestingly, the empirical isochrone agrees excellently with the isochrone of Dias et al. (2021) between 4 and 0 mag, but afterwards continues upwards in the indication of a much younger cluster without an MSTO. However, one needs to consider that in this region, the empirical isochrone already has very large associated uncertainty bounds. While the empirical isochrone in this case does not allow for a decisive argument for one of the two isochrones, it *does* imply that the cluster age is more likely to be younger than the postulated $277 - 299$ Myr, and that is likely even younger than the 123 Myr estimate, although the second assumption would require better knowledge about whether the stars extending the upper MS for values brighter than 0 mag, or the stars in the giant region are more likely to be field star contaminants.

There is actually a possible way to gain a bit more insight into which age should be expected for the cluster, when looking at the problem from another perspective. Inspecting the cluster in the `Aladin Sky Atlas`[25] (Bonnarel et al. 2000; Boch & Fernique 2014), shows that the very bright star HD 66194 is located right in the center of the cluster. It has been classified as a B3 or B2 spectral type according in the Morgan-Keenan classification scheme in all of the publications listed on the `Simbad` page[26] of the object (e.g. Hiltner et al. 1969; Cannon & Pickering 1993; Abt & Morgan 1969; Buscombe 1962; Evans et al. 1959), and has also been previously included as a member of the NGC 2516 cluster (e.g. Abt & Morgan 1969). Furthermore, B-type stars are very massive objects between $\geq 6\ M_\odot$ (B5) and 20 $M_\odot$ (B0) (Voigt et al. 2012), which evolve according to the processes described in Section 2.1.2. According to literature, they reach the 12 % limit of core hydrogen exhaustion after $4.6 - 46$ Myr, with the specific values again corresponding to B0 and B5 stars. To give another example, for a star with $M \approx 7\ M_\odot$, the hydrogen burning process on the MS ends after 25.5 Myr, and the RGB evolution starts already at 26.6 Myr in stellar evolution simulations (e.g. Voigt et al. 2012, and references therein). What these findings regarding the lifetime of massive O-stars are meant to convey, is that if the star is indeed a member of the NGC 2516 cluster, its physically possible age would be greatly restricted. In an effort to determine, whether HD 66194 is included in the cluster selection of this work, its EDR3 measurements are compared to the cluster member list of the crossmatch between the EDR3 and Catalogue I data: There is indeed an identical source included in the member selection of this work, with the *Gaia* EDR3 parameters displayed in Tab. 7.1. A quick comparison between the stellar parameters and the calculated CMD values reveals, that the object HD 66194 is represented by the highest extension of the MS, i.e. the data point with the coordinates (-0.104, -2.218) in the CMD of Fig. 7.1. As can be seen from the first and second plot of the figure, neither of the theoretical isochrones passes through this point, and the Dias et al. (2021) isochrone outright ignores it. However, the empirically derived median isochrone of 100 resampled isochrones actually *does* include it.

Table 7.1.: EDR3 and DR2 parameters for the star HD 66194, which can be found in the membership list of NGC 2516 by Cantat-Gaudin & Anders (2020).

Catalogue	RA	DEC	Parallax	PMRA	PMDEC	G band magnitude
EDR3	119.711	-60.8244	2.4707	-4.661	11.226	5.805
DR2	119.711	-60.8244	2.4919	-4.636	10.889	5.799

---

[25] https://aladin.u-strasbg.fr/

[26] http://simbad.u-strasbg.fr/simbad/sim-id?Ident=HIP+38994

Considering the object location with respect to the cluster from `Aladin` and `Simbad`, its presence in the cluster selection of the archive, and the fact that it is a B-type star, it would be extremely unlikely that it is not a cluster member. The additional information about the data point in the CMD finally supports the statement, that the NGC 2516 cannot possibly be as old as the proposed 277 Myr (Dias et al. 2021) if there is still a B-type star at the tip of its MS. Even the age estimate by Bossini et al. (2019) seems to be too high in light of the new findings. It is also interesting that the method treating the metallicity as free fitting parameter performs worse on the cluster data, than the one using a fixed Solar metallicity value. It again underlines how difficult and susceptible to age ambiguity the isochrone fitting process is, if no information about metallicity or extinction are available.

7.2.2. Case study II: Melotte 22

The Melotte 22 cluster, which has already been featured in Chapter 3 (see Fig. 3.3), is the topic of the second case study regarding empirical isochrones. The cluster is estimated to be as young as 86 Myr (Bossini et al. 2019), while others calculate ages of up to 176 Myr (Zhang et al. 2019). The theoretical isochrones shown in Fig. 7.2 are identical to the ones in Fig. 3.3 and the corresponding values in Tab. 3.2. The empirical isochrone including its bounds is again derived from $n_{\mathrm{obs}} = 100$ resampled isochrones in PCA space.

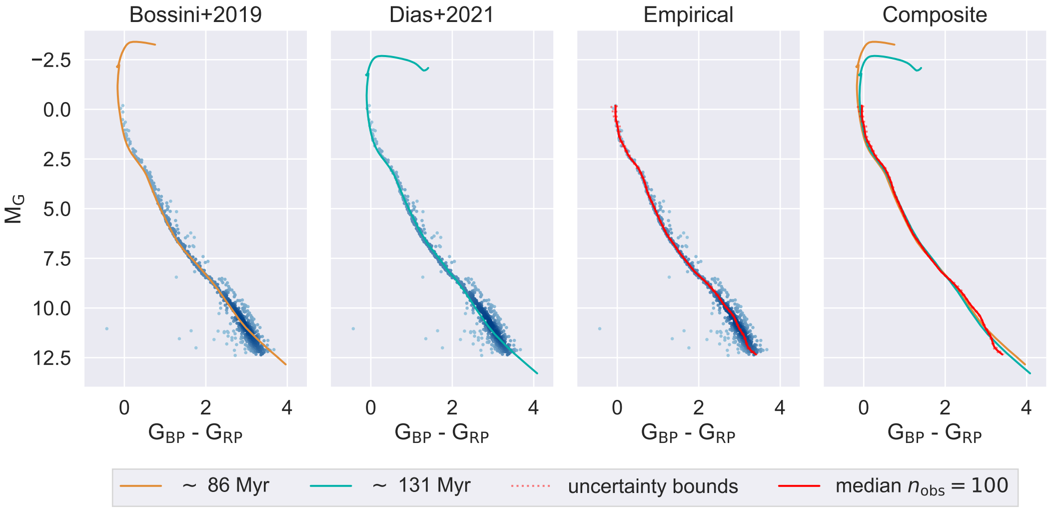

Figure 7.2.: Comparison between theoretical and empirical isochrones regarding the Melotte 22 cluster. The first three plots depict the cluster population according to Cantat-Gaudin & Anders (2020) ($p \geq 0.5$) along with either a model or an empirical isochrone, the latter including the calculated 5th and 95th percentile bounds. The panel on the far right of the figure displays a composite plot containing all three isochrones.

In contrast to the previous case study, an excellent agreement between the theoretical isochrones and the CMD distribution of the cluster can be observed. Furthermore, the empirical isochrone in this case only has a very narrow uncertainty region, indicating a very good fit as well. The problem with determining the ages in the case of this cluster, is that the upper MS of the cluster selection depicted in the plot cannot possibly be used to determine a decisive cluster age, even though the depicted theoretical isochrones are 50 Myr apart in their associated ages. The region of the cluster MSTO is not sampled by member stars and the dynamical range between ~ 0 and 12 mag falls exactly into the domain of the isochrone blindspot.

Nonetheless, a fact that can be gathered from the fits is that neither of the two theoretical isochrones fits the lower MS perfectly. The curves are generally flatter than the distribution traced by the empirical isochrone in that part of the cluster distribution. The conclusion that can be drawn from this observation, is that current stellar evolution models likely still struggle to accurately and completely incorporate the entire lower MS region, in which stars evolve under different, and arguably more complex, physical conditions than more massive stars. To reliably estimate the age of a cluster like Melotte 22 from a selection like the one shown in Fig. 7.2, it would however be crucial to create better fits to the lower MS, since it is the only region of the cluster CMD that is not as affected by the isochrone blindspot. A possible solution to the problem could be to use empricial isochrones as calibration data for stellar models, and thereby helping to improve them.

8. Summary and conclusion

A little more than a decade has passed, since David Soderblom wrote the concluding remarks of his review paper *The Ages of Stars* (Soderblom 2010). Among his predictions and hopes for the future of stellar age determination, he made the following statement:

> "THE BEST AVAILABLE DATA FOR THE NEAREST OCS WILL BE ANALYZED IN A SELF-CONSISTENT MANNER TO PRODUCE AT LEAST A MORE RELIABLE AGE ORDERING FOR THOSE CLUSTERS. THIS WILL LEAD TO SOME INSIGHTS ON WHAT FACTORS PRODUCE DIFFERENCES BETWEEN CLUSTERS WE DO NOT NOW UNDERSTAND (...)"

Incidentally, there is almost no better way of summing up the aim and purpose of the present project, than this quote. Drawing on the solid foundation of stellar evolution theory and its connection to the HRD, this work started by highlighting the importance of open clusters in unveiling the structure and history of the Milky Way. It was established that stellar ages play a crucial part in exploring the formation and evolution of larger structures, but that age determination remains a difficult task, associated with large uncertainties and model dependencies, to this day. The popular method of isochrone fitting can provide age estimates for stellar populations of almost all ages, but for an age range between $50 - 500$ Myr – named the isochrone blindspot – model isochrones show almost no discernible difference over a large dynamical range, if no additional cluster information is available.

For the purpose of providing an alternative to model-dependent cluster ages, an archive consisting of empirical isochrones for 67 open clusters in the Solar Neighborhood, based on *Gaia* DR2 and EDR3 data, was created. To extract the isochrones, various methods were tested on the data and a final extraction algorithm, consisting of a combination of the machine-learning method support vector regression and the pre-processing method principal component analysis, was developed. By resampling large numbers of SVR-based empirical isochrones from bootstrapped cluster data, uncertainty regions and robust final isochrones were provided for each cluster. Comparing the empirical isochrones for the archive clusters to age estimates found in literature, showed a generally good agreement and proves, that the archive can indeed be viewed as a relative age determination tool based on sorting ages.

Finally, the investigation of two cluster case-studies showed that empirical isochrones can help in a decision between two deviating cluster ages, and that they could also provide calibration data for stellar evolution models. At the same time, the case study of the cluster NGC 2516 yielded the result, that the cluster is likely younger than literary estimates around $123 - 277$ Myr, as it contains a bright B-type star, that is still located on the MS.

In creating the isochrone archive, a new tool has been added to the collection of modern age estimation techniques, and the added value it can provide will hopefully turn out to be a stepping stone on our path towards unlocking the physical mechanisms of our home galaxy, across both space and time.

Bibliography

Abt, H. A. & Morgan, W. W. 1969, Astronomical Journal, 74, 813

Alessi, B. S., Moitinho, A., & Dias, W. S. 2003, Astronomy and Astrophysics, 410, 565

Anderson, T. 2003, An Introduction to Multivariate Statistical Analysis, Wiley Series in Probability and Statistics (Wiley)

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, Astronomical Journal, 156, 123

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, Astronomy and Astrophysics, 558, A33

Barbaro, G., Dallaporta, N., & Fabris, G. 1969, Astrophysics and Space Science, 3, 123

Barnes, S. A. 2007, Astrophysical Journal, 669, 1167

Basri, G., Marcy, G. W., & Graham, J. R. 1996, Astrophysical Journal, 458, 600

Bishop, C. M. 2006, Pattern Recognition and Machine Learning (Information Science and Statistics) (Berlin, Heidelberg: Springer-Verlag)

Boch, T. & Fernique, P. 2014, in Astronomical Society of the Pacific Conference Series, Vol. 485, Astronomical Data Analysis Software and Systems XXIII, ed. N. Manset & P. Forshay, 277

Bonnarel, F., Fernique, P., Bienaymé, O., et al. 2000, Astronomy and Astrophysics, Supplement, 143, 33

Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, Astronomy and Astrophysics, 623, A108

Bouy, H., Bertin, E., Sarro, L. M., et al. 2015, Astronomy and Astrophysics, 577, A148

Bressan, A., Marigo, P., Girardi, L., et al. 2012, Monthly Notices of the Royal Astronomical Society, 427, 127

Briceño, C., Hartmann, L., Hernández, J., et al. 2007, Astrophysical Journal, 661, 1119

Buscombe, W. 1962, Mount Stromlo Observatory Mimeographs, 4, 1

Cannon, A. J. & Pickering, E. C. 1993, VizieR Online Data Catalog, III/135A

Cantat-Gaudin, T. & Anders, F. 2020, Astronomy and Astrophysics, 633, A99

Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018a, Astronomy and Astrophysics, 618, A93

Cantat-Gaudin, T., Vallenari, A., Sordo, R., et al. 2018b, Astronomy and Astrophysics, 615, A49

Castro-Ginard, A., Jordi, C., Luri, X., Cantat-Gaudin, T., & Balaguer-Núñez, L. 2019, Astronomy and Astrophysics, 627, A35

Castro-Ginard, A., Jordi, C., Luri, X., et al. 2018, Astronomy and Astrophysics, 618, A59

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560

Chen, B., D'Onghia, E., Alves, J., & Adamo, A. 2020, Astronomy and Astrophysics, 643, A114

Choi, J., Dotter, A., Conroy, C., et al. 2016, The Astrophysical Journal, 823, 102

Clarke, C. J., Bonnell, I. A., & Hillenbrand, L. A. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell, 151

Curtis, J. L., Wolfgang, A., Wright, J. T., Brewer, J. M., & Johnson, J. A. 2013, Astronomical Journal, 145, 134

Da Rio, N., Tan, J. C., Covey, K. R., et al. 2017, Astrophysical Journal, 845, 105

David, T. J. & Hillenbrand, L. A. 2015, Astrophysical Journal, 804, 146

de Boer, K. & Seggewiss, W. 2021, Stars and Stellar Evolution (EDP Sciences)

Dias, W. S., Lépine, J. R. D., & Alessi, B. S. 2001, Astronomy and Astrophysics, 376, 441

Dias, W. S., Monteiro, H., Moitinho, A., et al. 2021, Monthly Notices of the Royal Astronomical Society, 504, 356

Dotter, A. 2016, Astrophysical Journal, Supplement, 222, 8

Dreyer, J. L. E. 1888, Memoirs of the RAS, 49, 1

ESA. 1997, The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission

Evans, D. S., Menzies, A., & Stoy, R. H. 1959, Monthly Notices of the Royal Astronomical Society, 119, 638

Evans, D. W., Riello, M., De Angeli, F., et al. 2018, Astronomy and Astrophysics, 616, A4

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, Astronomical Journal, 111, 1748

Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, Astronomy and Astrophysics, 616, A10

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, Astronomy and Astrophysics, 616, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, Astronomy and Astrophysics, 649, A1

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, Astronomy and Astrophysics, 595, A1

Garner, R. 2017, Hubble's Messier Catalog: Messier 4, `https://www.nasa.gov/feature/goddard/2017/messier-4`, [Online; accessed 09-12-2021]

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357

Hastie, T. & Stuetzle, W. 1989, Journal of the American Statistical Association, 84, 502

Heber, U. 2009, Annual Reviews of Astronomy and Astrophysics, 47, 211

Hidalgo, S. L., Pietrinferni, A., Cassisi, S., et al. 2018, Astrophysical Journal, 856, 125

Hiltner, W. A., Garrison, R. F., & Schild, R. E. 1969, Astrophysical Journal, 157, 313

Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, Astronomy and Astrophysics, 355, L27

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90

Hurley, J. & Tout, C. A. 1998, Monthly Notices of the Royal Astronomical Society, 300, 977

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2012, Astronomy and Astrophysics, 543, A156

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2013, Astronomy and Astrophysics, 558, A53

Kippenhahn, R., Weigert, A., & Weiss, A. 2012, Stellar Structure and Evolution, Astronomy and Astrophysics Library (Springer Berlin Heidelberg)

Koester, D. 2002, Astronomy and Astrophysics Reviews, 11, 33

Krone-Martins, A. & Moitinho, A. 2014, Astronomy and Astrophysics, 561, A57

Kroupa, P. 2001, Monthly Notices of the Royal Astronomical Society, 322, 231

Lada, C. J. & Lada, E. A. 2003, Annual Reviews of Astronomy and Astrophysics, 41, 57

Lodieu, N. 2020, Mem. Societa Astronomica Italiana, 91, 84

Lohmann, W. 1957, Zeitschrift für Astrophysik, 42, 114

Lorensen, W. E. & Cline, H. E. 1987, in Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '87 (New York, NY, USA: Association for Computing Machinery), 163–169

Maíz Apellániz, J. & Weiler, M. 2018, Astronomy and Astrophysics, 619, A180

Mantz, H., Jacobs, K., & Mecke, K. 2008, Journal of Statistical Mechanics: Theory and Experiment, 2008, P12015

Marigo, P., Girardi, L., Bressan, A., et al. 2017, Astrophysical Journal, 835, 77

Mayne, N. J., Naylor, T., Littlefair, S. P., Saunders, E. S., & Jeffries, R. D. 2007, Monthly Notices of the Royal Astronomical Society, 375, 1220

Meingast, S., Alves, J., & Rottensteiner, A. 2021, Astronomy and Astrophysics, 645, A84

Messier, C. 1781, Catalogue des Nébuleuses et des Amas d'Étoiles (Catalog of Nebulae and Star Clusters), Connoissance des Temps ou des Mouvements Célestes

Miret-Roig, N., Bouy, H., Olivares, J., et al. 2019, Astronomy and Astrophysics, 631, A57

NASA. 2003, Globular Cluster M4, `https://hubblesite.org/contents/media/images/2003/19/1377-Image.html`, [Online; accessed 09-12-2021]

Olivares, J., Bouy, H., Sarro, L. M., et al. 2019, Astronomy and Astrophysics, 625, A115

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Perryman, M. A. C., Brown, A. G. A., Lebreton, Y., et al. 1998, Astronomy and Astrophysics, 331, 81

Pietrinferni, A., Hidalgo, S., Cassisi, S., et al. 2021, Astrophysical Journal, 908, 102

Pols, O. 2011, Stellar Structure and Evolution (Astronomical Institute Utrecht)

Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, Annual Reviews of Astronomy and Astrophysics, 48, 431

Rebolo, R., Martin, E. L., & Magazzu, A. 1992, Astrophysical Journal, Letters, 389, L83

Riello, M., De Angeli, F., Evans, D. W., et al. 2021, Astronomy and Astrophysics, 649, A3

Roeser, S., Demleitner, M., & Schilbach, E. 2010, Astronomical Journal, 139, 2440

Saha, M. N. 1920, LIII. Ionization in the solar chromosphere

Saha, M. N. 1921, Proceedings of the Royal Society of London Series A, 99, 135

Salaris, M., Cassisi, S., Pietrinferni, A., & Hidalgo, S. 2021, Monthly Notices of the Royal Astronomical Society, 509, 5197–5208

Schaerer, D., Meynet, G., Maeder, A., & Schaller, G. 1993, Astronomy and Astrophysics, Supplement, 98, 523

Sim, G., Lee, S. H., Ann, H. B., & Kim, S. 2019, Journal of Korean Astronomical Society, 52, 145

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, Astronomical Journal, 131, 1163

Smola, A. J. & Schölkopf, B. 2004, Statistics and Computing, 14, 199–222

Soderblom, D. R. 2010, Annual Reviews of Astronomy and Astrophysics, 48, 581

Soderblom, D. R., Hillenbrand, L. A., Jeffries, R. D., Mamajek, E. E., & Naylor, T. 2014, in Protostars and Planets VI, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 219

Stahler, S. W. 1988, Astrophysical Journal, 332, 804

Swiggum, C., D'Onghia, E., Alves, J., et al. 2021, Astrophysical Journal, 917, 21

Tadross, A., Werner, P., Osman, A., & Marie, M. 2002, New Astronomy, 7, 553

Tassoul, J.-L. 2014, A Concise History of Solar and Stellar Physics (Princeton, NJ :: Princeton University Press,)

Tipping, M. E. & Bishop, C. M. 1999, Neural Computation, 11, 443

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, PeerJ, 2, e453

Voigt, H., Röser, H., & Tscharnuter, W. 2012, Abriss der Astronomie (Wiley-VCH)

von Hippel, T. 2005, Astrophysical Journal, 622, 565

Ward, J. L. & Kruijssen, J. M. D. 2018, Monthly Notices of the Royal Astronomical Society, 475, 5659

Waskom, M. L. 2021, Journal of Open Source Software, 6, 3021

Winget, D. E. & Kepler, S. O. 2008, Annual Reviews of Astronomy and Astrophysics, 46, 157

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, Astronomical Journal, 140, 1868

Wuchterl, G. & Tscharnuter, W. M. 2003, Astronomy and Astrophysics, 398, 1081

Zhang, J., Zhao, J., Oswalt, T. D., et al. 2019, The Astrophysical Journal, 887, 84

Zhang, T. Y. & Suen, C. Y. 1984, Commun. ACM, 27, 236–239

Željko Ivezić, Connolly, A. J., VanderPlas, J. T., & Gray, A. 2019, Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, Updated Edition (Princeton University Press)

A. Cluster CMDs

The plots in this section depict the calculated SVR-isochrones for all 67 archive clusters from the selection of Catalogue 1. A collection of $n_{\text{boot}} = 1000$ bootstrapped isochrones per cluster was used to calculate the median-based, final isochrones (*red lines*), as well as the uncertainty bounds (*grey lines*), defined by the 5^{th} and 95^{th} percentiles.

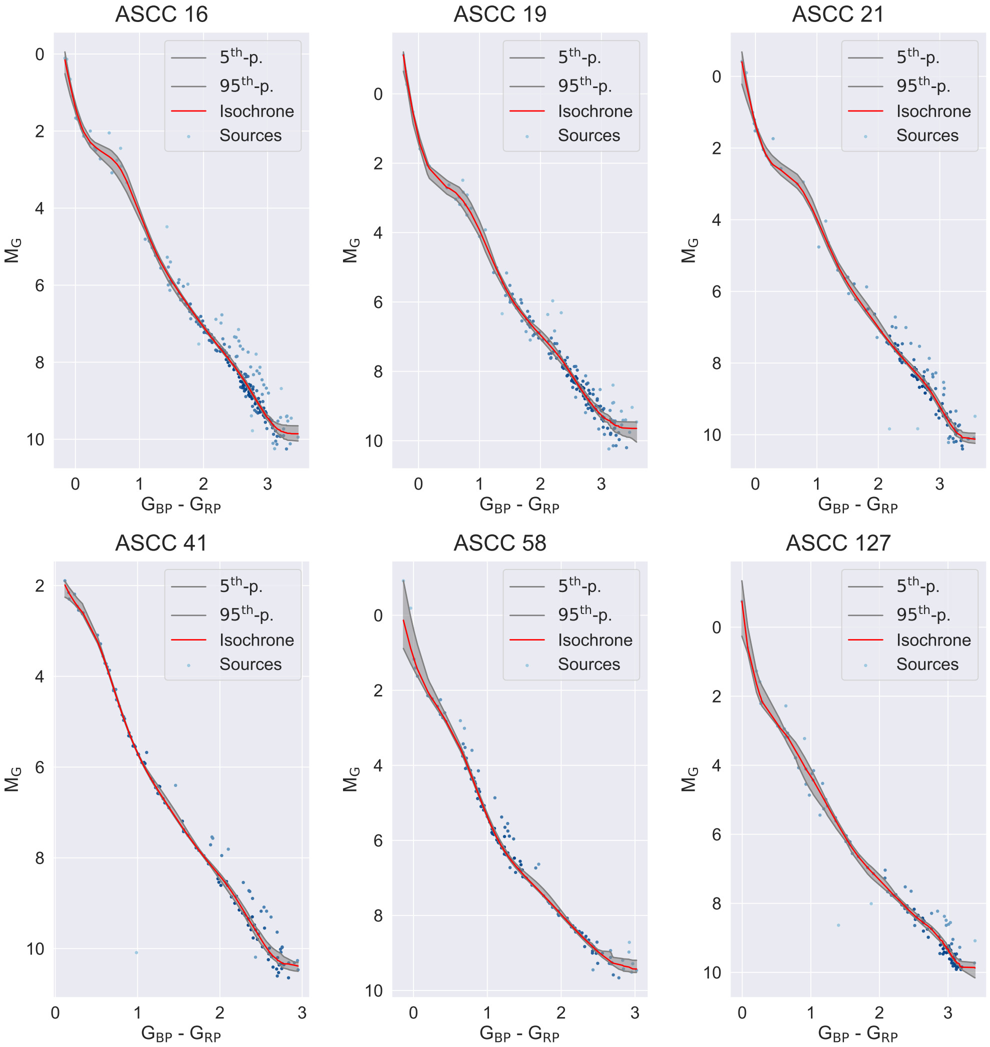

Figure A.1.: CMDs and empirical isochrones for clusters ASCC 16 to ASCC 127 (1-6).

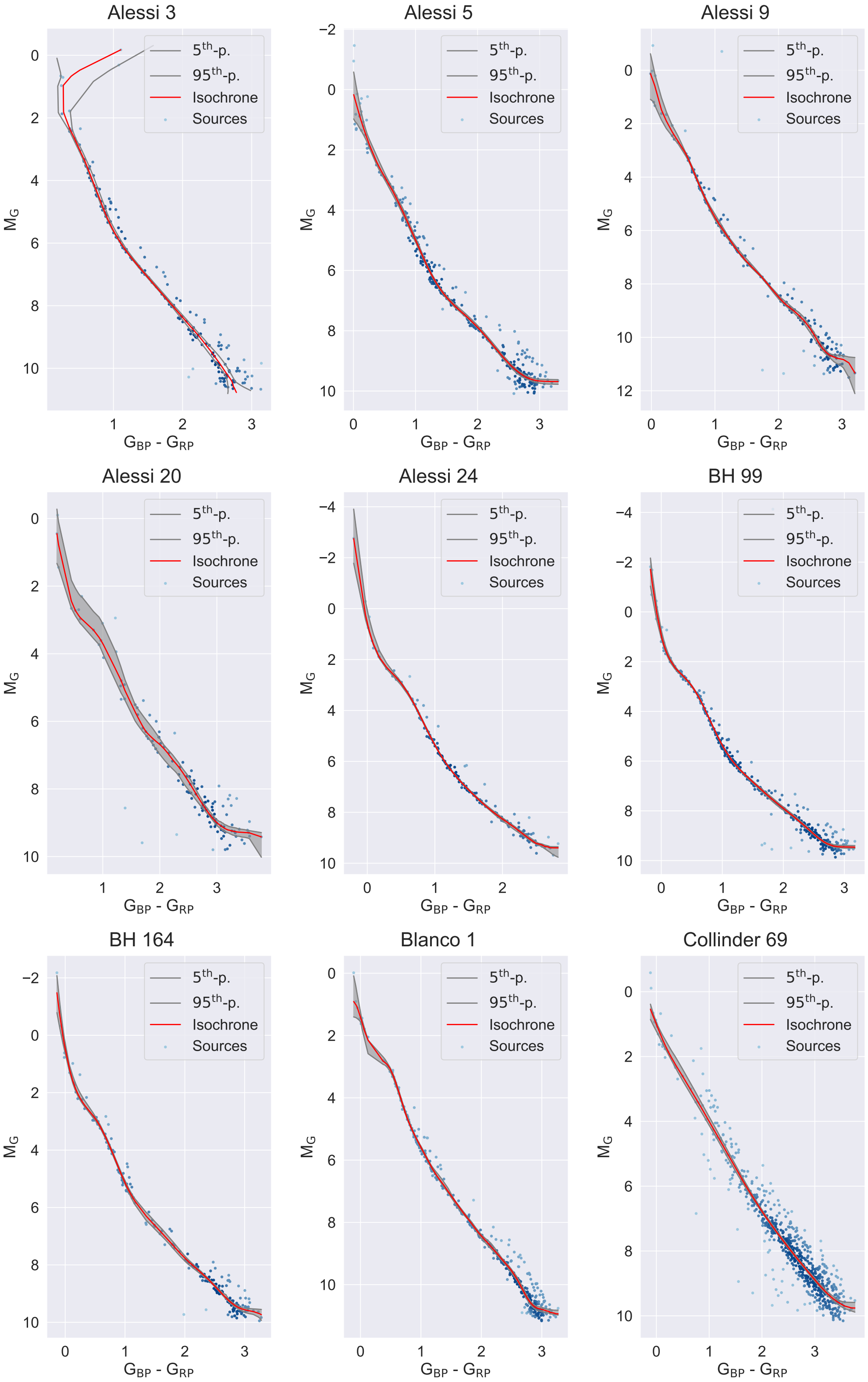

Figure A.2.: Same as Fig. A.1, for clusters Alessi 3 to Collinder 69 (7-15).

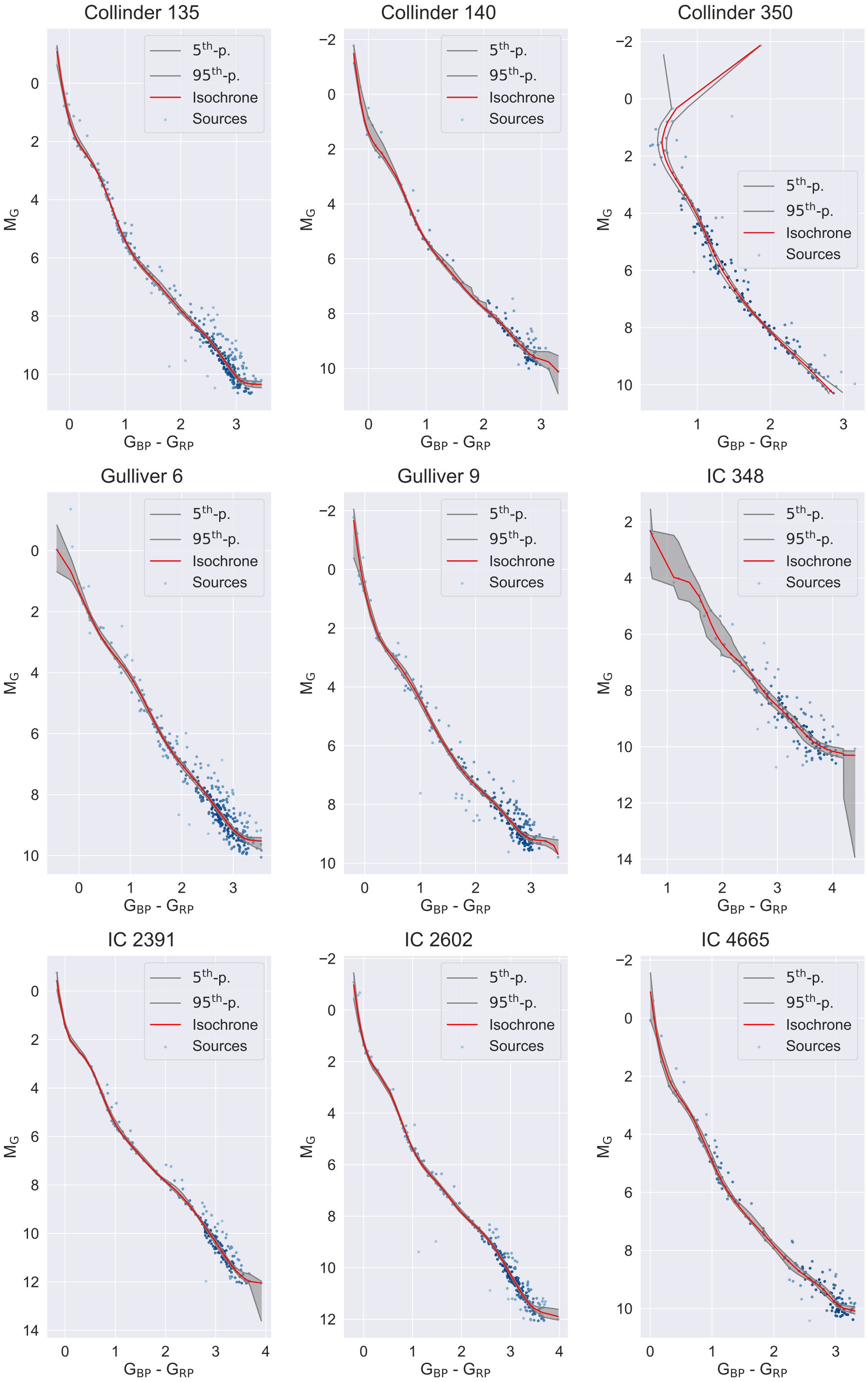

Figure A.3.: Same as Fig. A.1, for clusters Collinder 135 to IC 4665 (16-24).

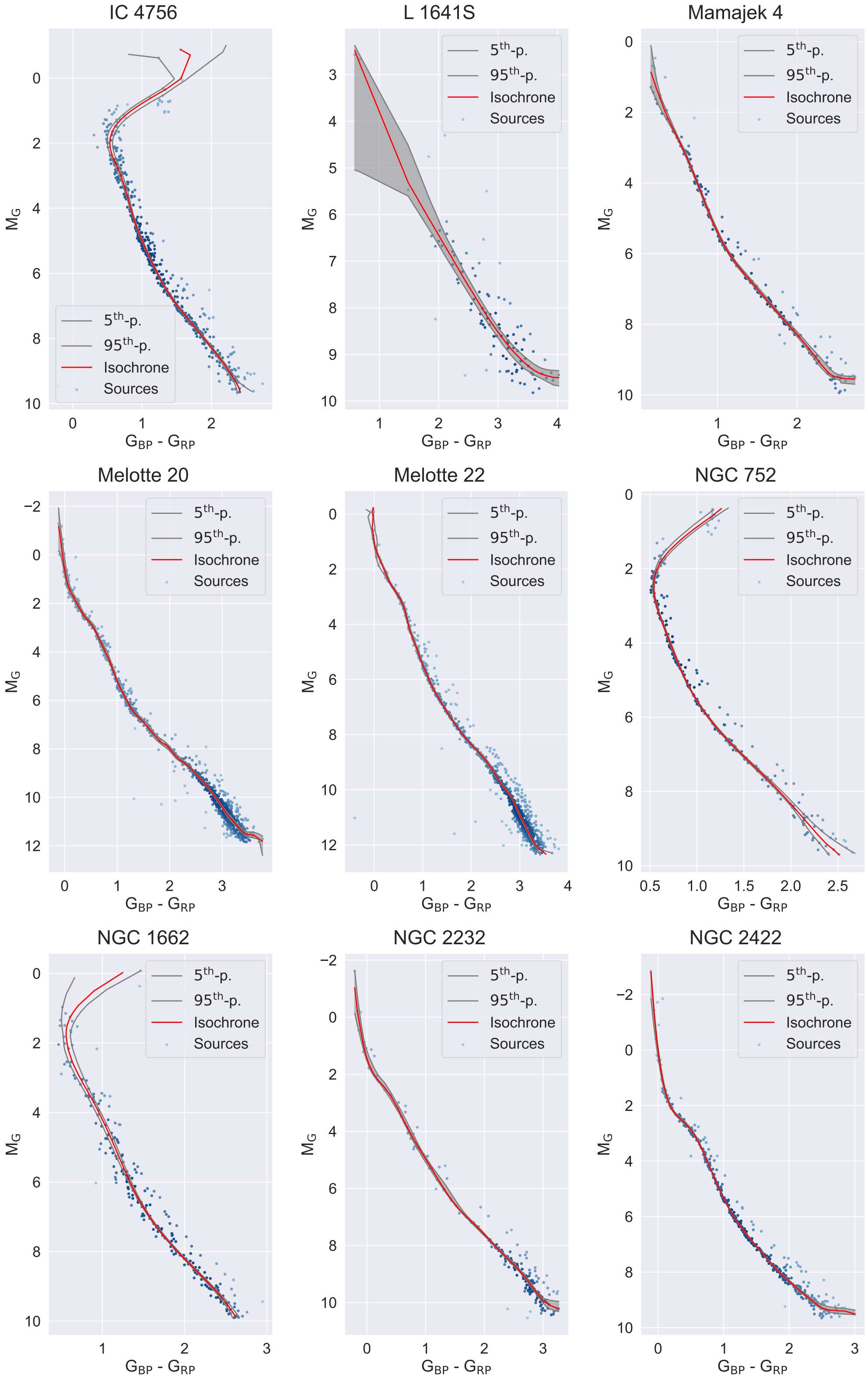

Figure A.4.: Same as Fig. A.1, for clusters IC 4756 to NGC 2422 (25-33).

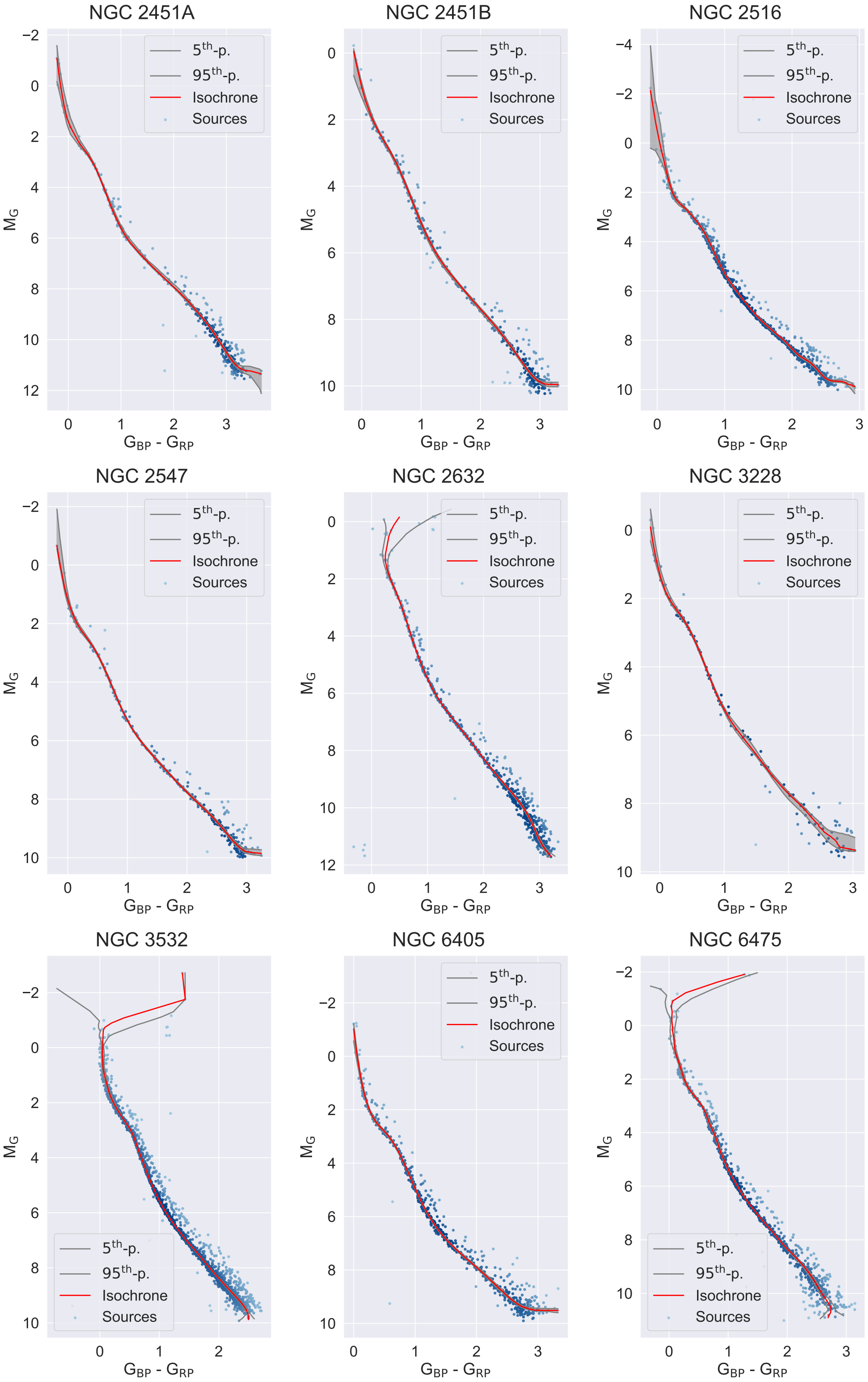

Figure A.5.: Same as Fig. A.1, for clusters NGC 2451A to NGC 6475 (34-42).

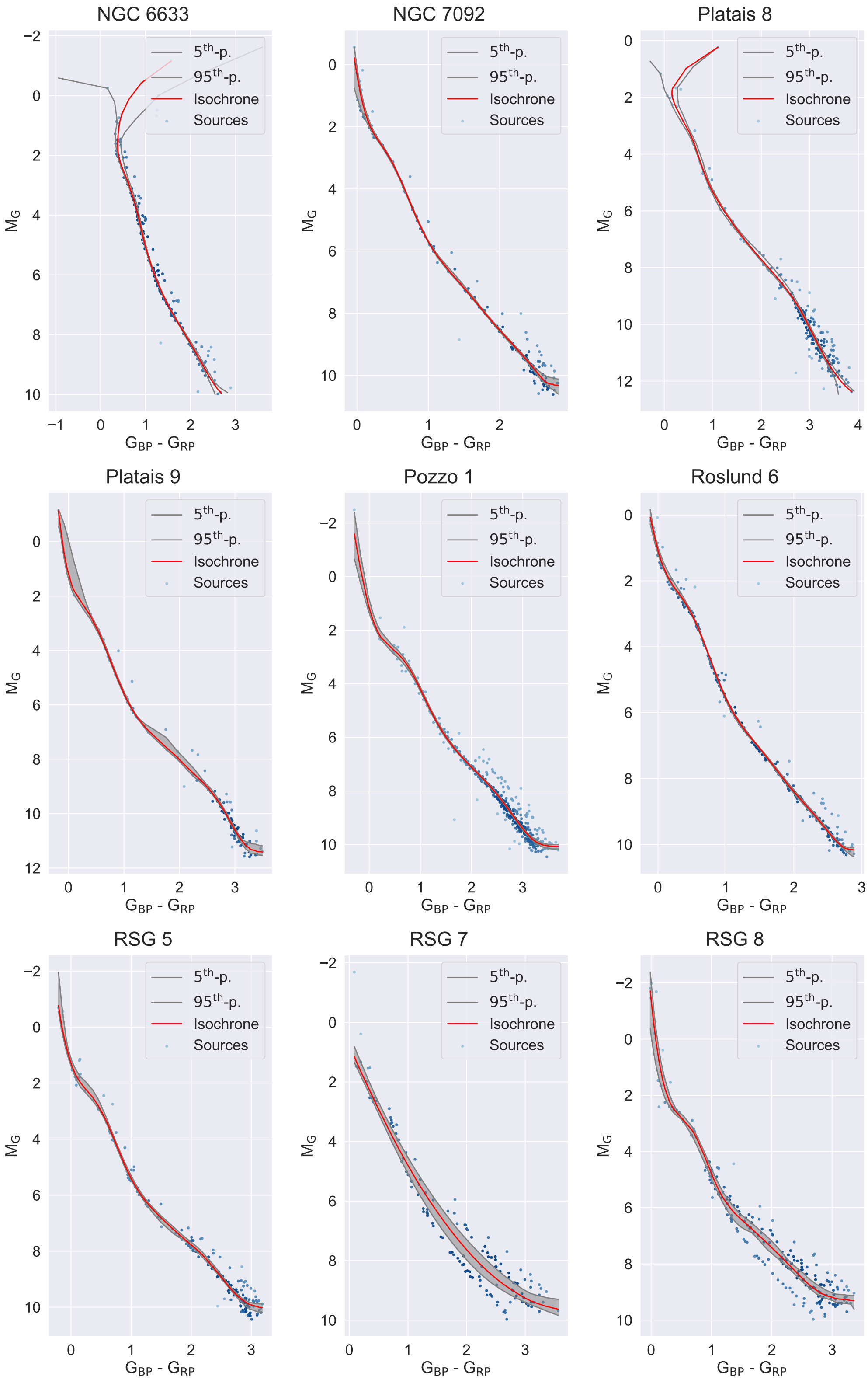

Figure A.6.: Same as Fig. A.1, for clusters NGC 6633 to RSG 8 (43-51).

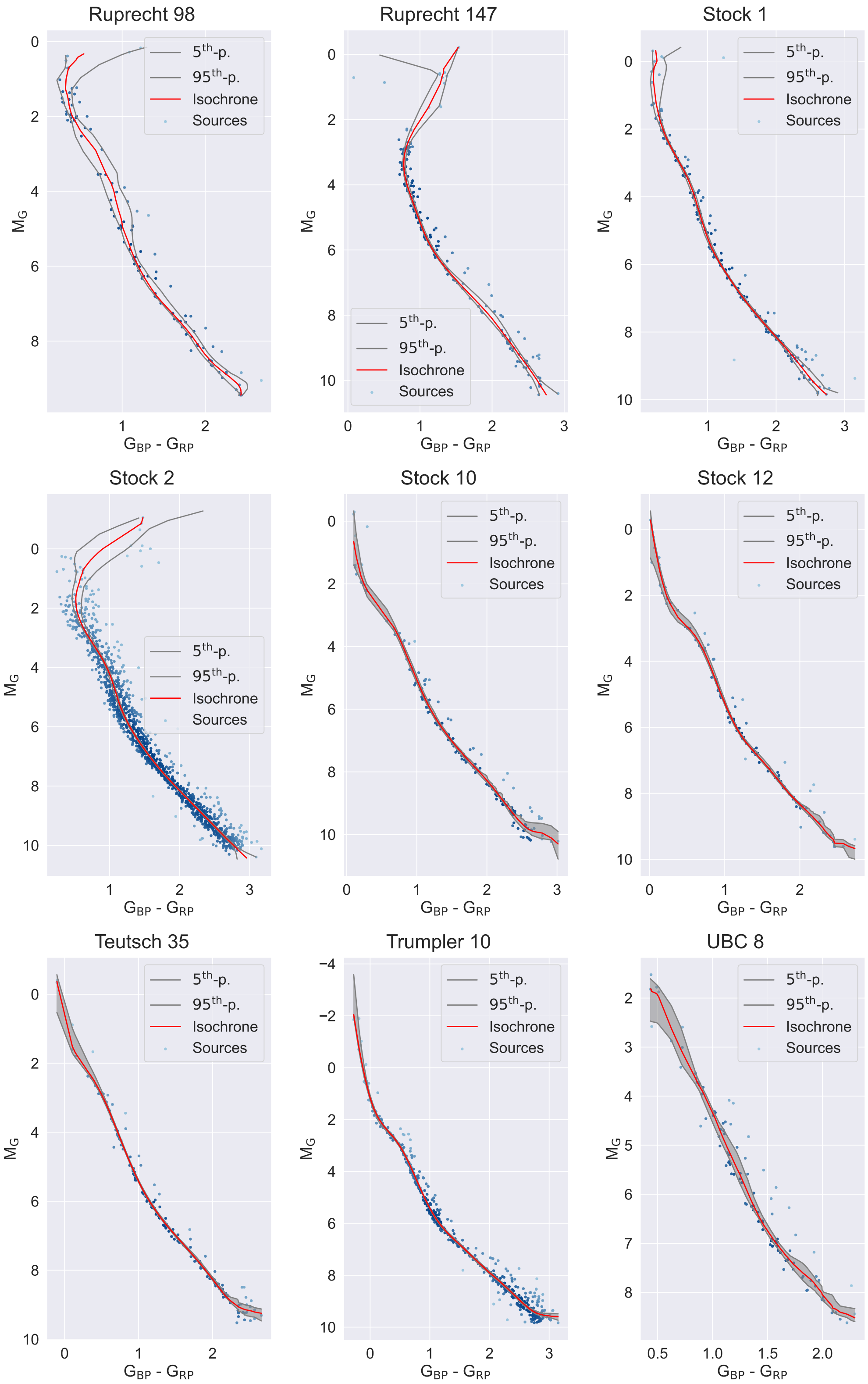

Figure A.7.: Same as Fig. A.1, for clusters Ruprecht 98 to UBC 8 (52-60).

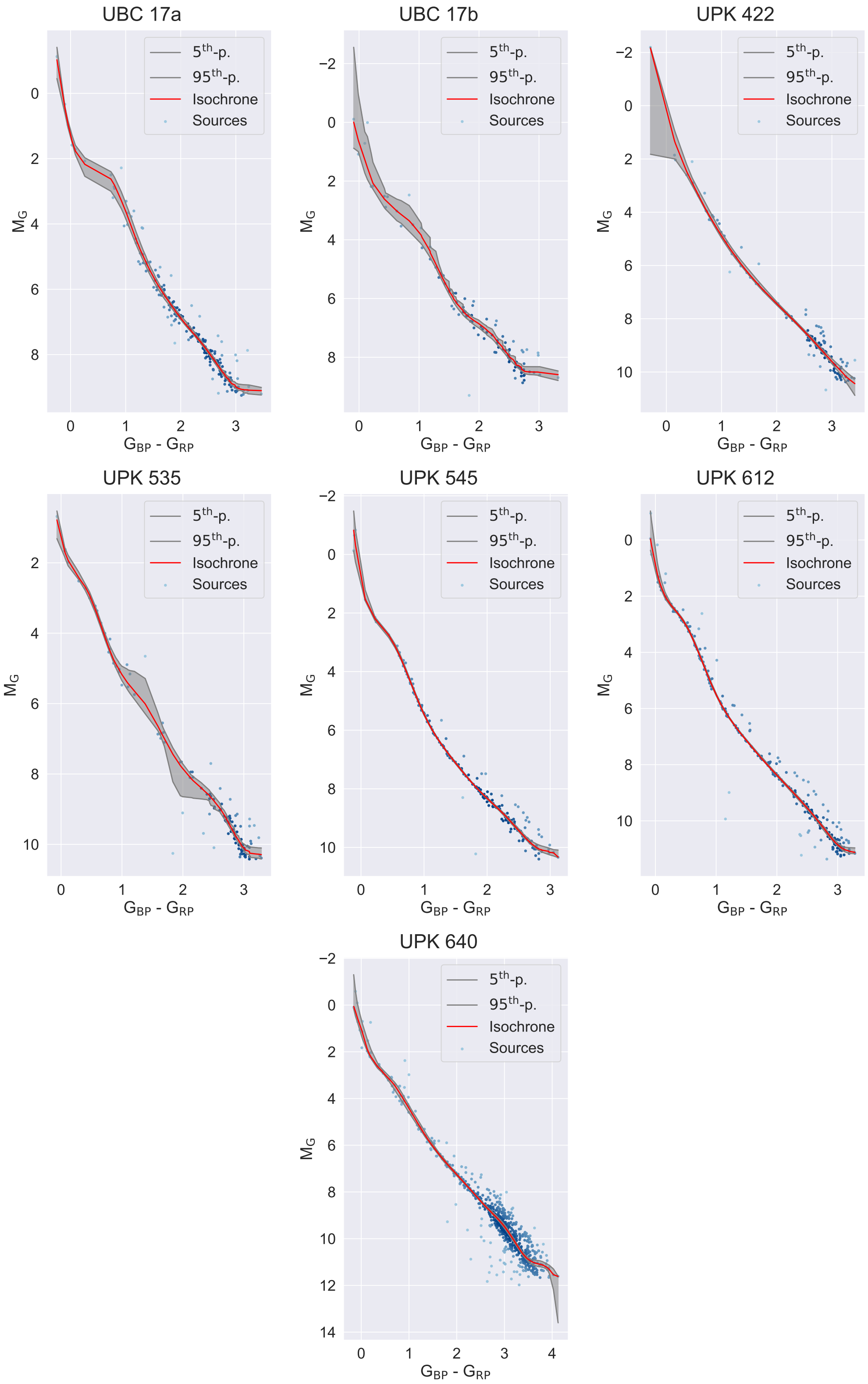

Figure A.8.: Same as Fig. A.1, for clusters UBC 17a to UPK 640 (61-67).

B. Errorplots for Catalogue II clusters

The plots shown in this section depicted the binned errors of the cluster CMD data of the selection included in Catalogue II. The errors are calculated under consideration of the effects of error propagation in both the response and the explanatory variables of the CMD, which are defined by Eq. (2.1) and Eq. (2.6), respectively. The population of each cluster is split into 20 partitions, and the mean error values in both variables are calculated for each of them. The results are displayed as errorbars beneath the cluster distributions in the CMDs, with the center panel displaying a 5× magnification of both error values, whereas the right panel displays only the errors in absolute magnitude, with a 50× magnification.

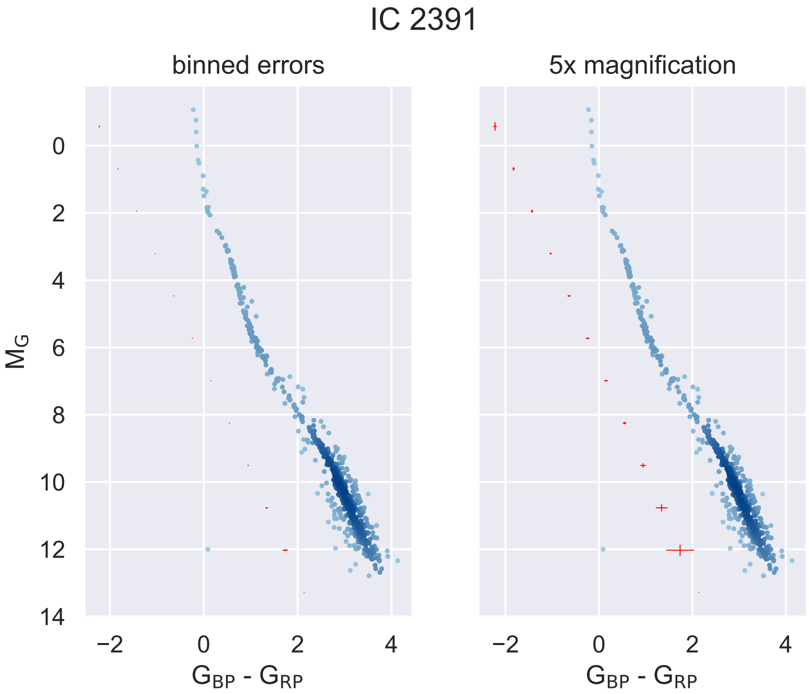

Figure B.1.: Binned CMD data errors for the cluster CMD of IC 2391. The right panel of the figure displays a five-fold magnification of the error bars.

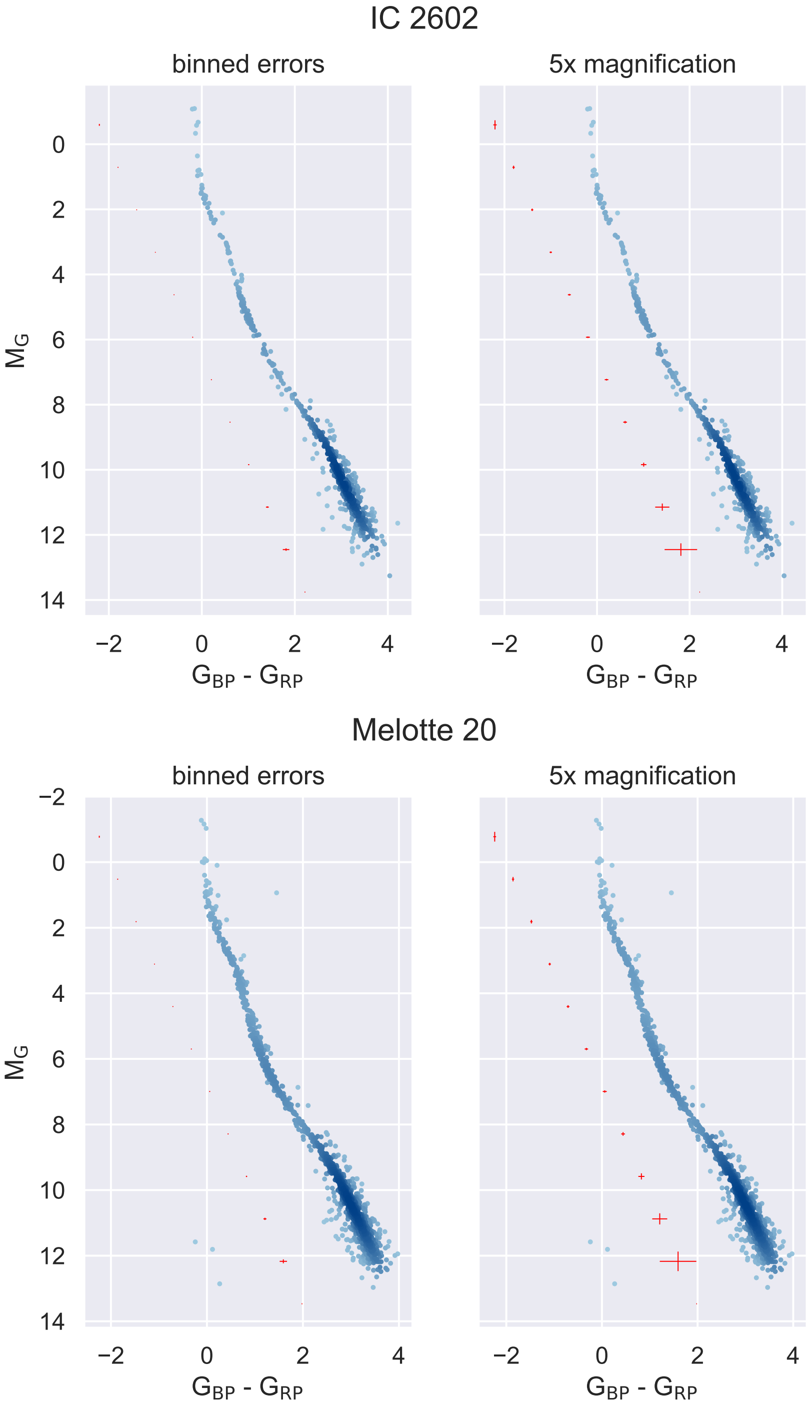

Figure B.2.: Same as Fig. B.1, for clusters IC 2602 and Melotte 20.

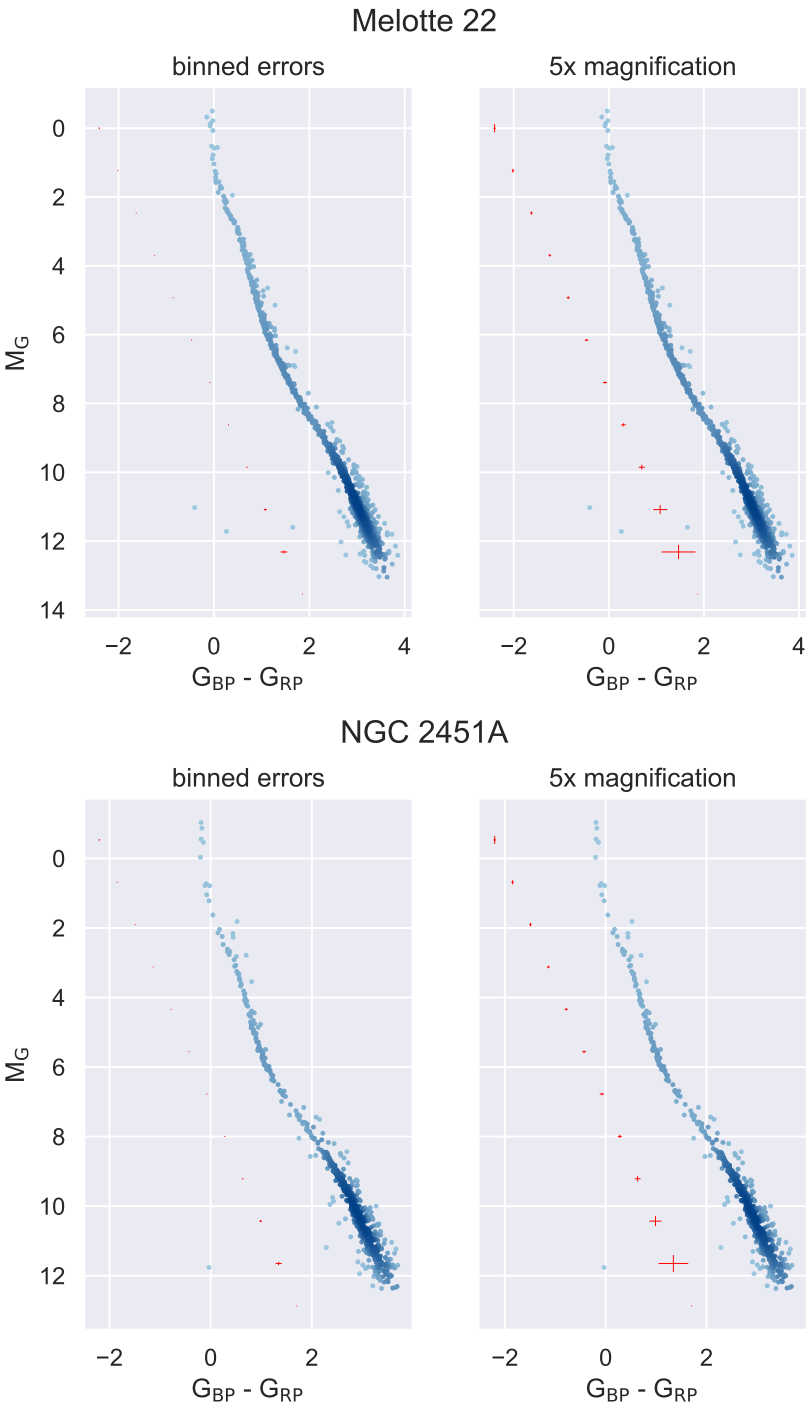

Figure B.3.: Same as Fig. B.1, for clusters Melotte 22 and NGC 2451A.

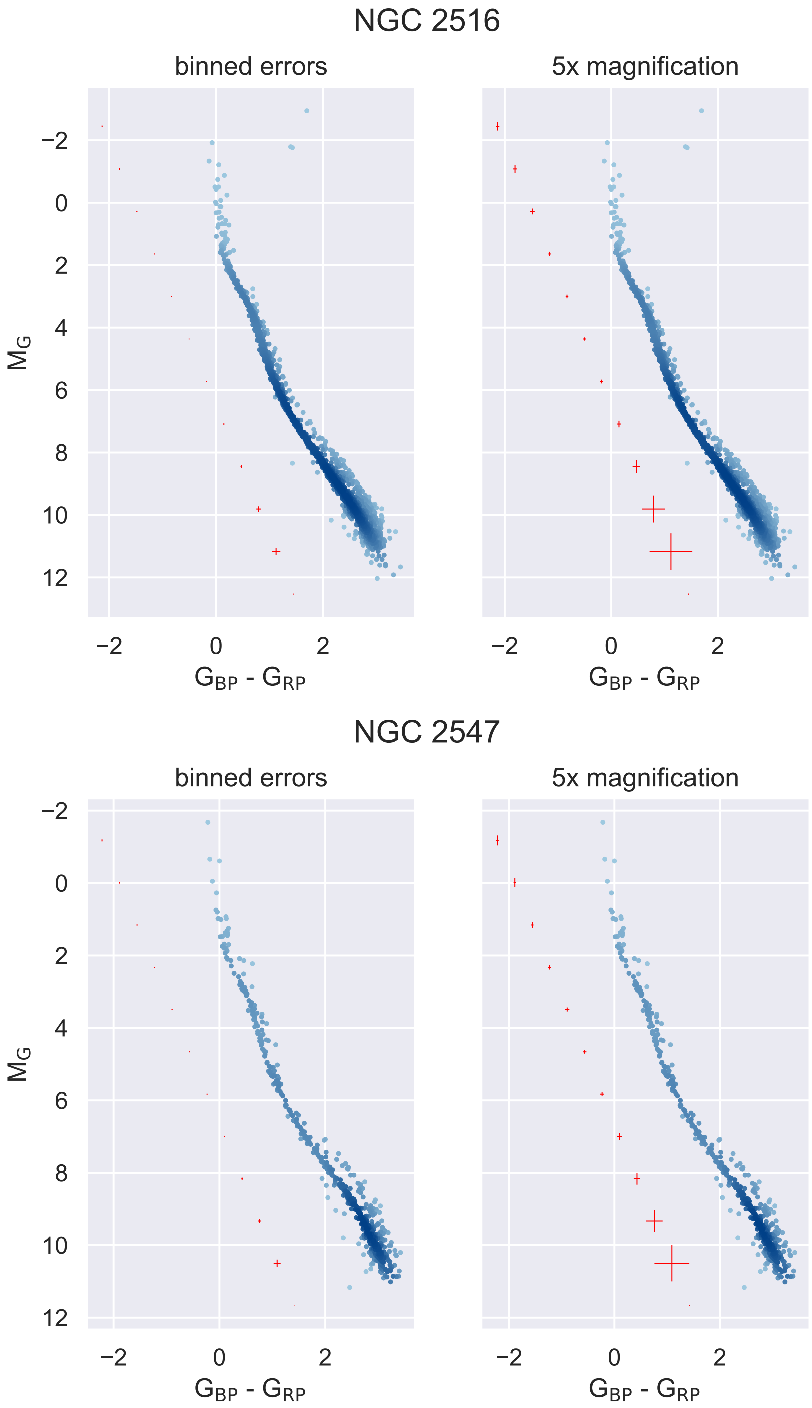

Figure B.4.: Same as Fig. B.1, for clusters NGC 2516 and NGC 2547.

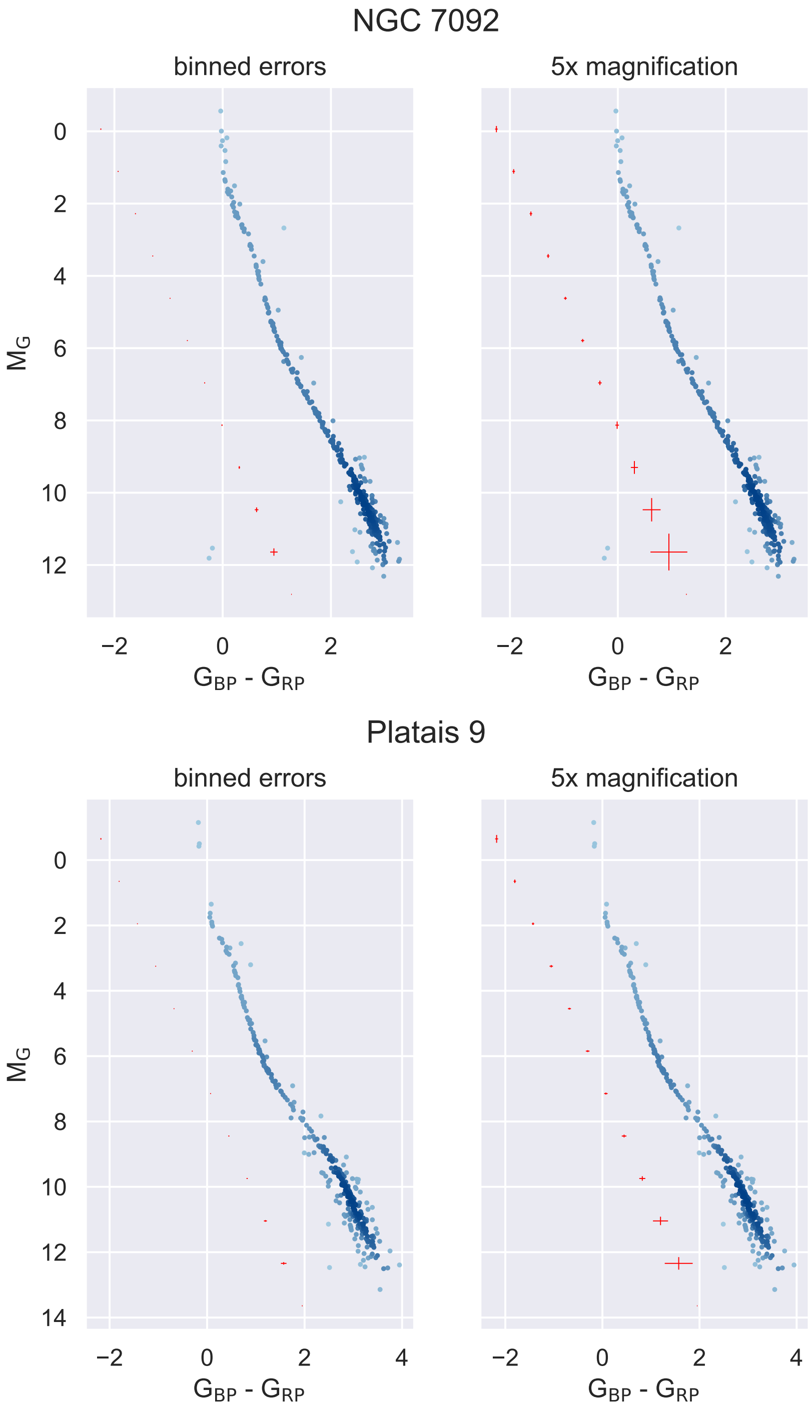

Figure B.5.: Same as Fig. B.1, for clusters NGC 7092 and Platais 9.