Departamento de Astrofísica
Universidad de La Laguna

M.Sc. in Astrophysics
MASTER'S THESIS

A first robust star formation history of the solar neighbourhood within 100 pc using the Gaia Catalogue of Nearby Stars

---

Supervisor: Carme Gallart & Tomás Ruiz-Lara

Author: Judith Santos Torres

July 2022, San Cristóbal de La Laguna

List of Contents

Resumen

Para descifrar la compleja historia de la Vía Láctea y comprender cómo era originalmente y cual ha sido su evolución, la arqueología galáctica examina los registros fósiles de nuestra galaxia mediante observaciones detalladas de las estrellas, el gas y otras estructuras. Uno de los componentes clave para rastrear la formación y evolución de las galaxias es la llamada Historia de Formación Estelar, que revela cómo y cuándo se formaron las estrellas a lo largo de la vida de la galaxia.

En este trabajo proporcionamos por primera vez la Historia de Formación Estelar (SFH) detallada del *Gaia* Catalogue of Nearby Stars (GCNS), una esfera de 100 pc en la vecindad solar con objetos muy bien caracterizados utilizando la alta calidad de los datos de *Gaia* eDR3.

Los Diagramas de Color-Magnitud en el plano absoluto (aCMDs) albergan información de las poblaciones estelares, lo cual los convierte en una excelente herramienta para determinar las SFHs. En el presente proyecto empleamos una técnica actualizada de ajuste del aCMD a través de la comparación entre el aCMD observado y los aCMDs sintetizados según las predicciones teóricas de librerías de evolución estelar.

Para familiarizarnos con los datos, examinamos el aCMD del GCNS realizando un primer estudio de las poblaciones estelares. En él, identificamos las estrellas con velocidades radiales válidas y las clasificamos según la componente galáctica a la que pertenecían: disco fino, grueso o halo. Además, verificamos la calidad de los datos y comprobamos que los errores de la distancia y la magnitud aparente G aumentaban a medida que los objetos eran más distantes y más débiles, respectivamente. También estudiamos los efectos de la extinción interestelar que resultaron ser despreciables. Finalmente, probamos las características generales de las estrellas contenidas en el catálogo mediante la superposición de isocronas de diferentes edades y metalicidades usando dos librerías estelares distintas: la de BaSTI y la de PARSEC.

En esta última prueba detectamos que los modelos estelares de BaSTI presentaban una apreciable discrepancia hacia el azul en relación con los datos observacionales, localizada en la baja secuencia principal del aCMD. Esta desviación resultó estar vinculada con el cálculo de las correcciones bolométricas y se resolvió implementando correcciones adicionales actualizadas. A pesar de que esta región no afecta a la determinación de la SFH que nos ocupa, cerciorarnos de este problema promovió la mejora en la calidad de la librería, lo cual es de gran importancia para futuros trabajos.

La misión *Gaia* destaca por la precisión sin precedentes de sus datos astrométricos y fotométricos. Aunque su poder de resolución es alto, *Gaia* detecta algunas binarias como estrellas individuales que son más brillantes de lo que deberían, dando lugar a un flujo que es en realidad la suma de los flujos de las dos estrellas que componen el sistema. Estos sistemas binarios trazan una banda paralela a la secuencia principal desplazada hacia magnitudes más brillantes. Una parte importante de nuestro trabajo consistió en utilizar el conjunto de datos del GCNS para estimar la población local de binarias no resueltas. Este análisis se llevó a cabo mediante el cálculo de diferentes SFHs usando distintos aCMDs sintetizados variando las fracciones de binarias y ratios de masas de las compañeras secundarias. Posteriormente se compararon las distribuciones estelares en la secuencia principal y en la secuencia de binarias no resueltas del aCMD observado y de los diferentes aCMDs solución. Esto permitió acotar las características de la población de binarias no resueltas en el GCNS, en términos de fracción de binarias y distribución de masa de las secundarias.

En nuestros cálculos tuvimos en cuenta los modelos estelares de BaSTI y PARSEC, también aCMD sintéticos con diferente número de estrellas, experimentamos con diferentes parametrizaciones de binarias no resueltas y finalmente tuvimos en cuenta distintos parámetros de entrada para el código de ajuste de la SFH. A lo largo de todo el proceso, las diferentes SFHs resultantes nos permitieron comprobar en qué medida afectan a la solución los distintos parámetros seleccionados. Observamos que las principales particularidades de las SFHs obtenidas se conservaban. Aún así, elegimos los parámetros que proporcionaron soluciones más detalladas y que, a su vez, hacían más eficiente la ejecución del código, y procedimos con la derivación de la SFH precisa de la muestra.

La SFH del GCNS es principalmente la historia de formación estelar del disco fino ya que la mayoría de las estrellas de este volumen pertenecen a esta componente. Los ingredientes que adoptamos tras el muestreo para extraer la SFH completa fueron un diagrama sintético de 100 millones de estrellas calculado con la librería BaSTI, considerando una fracción de binarias de 0.3 y una distribución de masas de las secundarias entre 0.1 y 1 veces la masa de la primaria. Y por otro lado, en cuanto a los parámetros introducidos en el código, consideramos pesos uniformes en todas las regiones del aCMD y limitamos nuestros cálculos a magnitudes G absolutas, $M_G = 4$ mag.

Las principales características de la SFH calculada resultaron ser: épocas con episodios de formación estelar (Star Formation Rates (SFRs)) situados entorno a 0.5, 2.3, 5.5 y posiblemente a 10 Gyrs, distribuciones de metalicidades con

un pico significativo entorno a metalicidades solares y, por último, estructuras diferenciadas en el plano de edad-metalicidad, entre las cuales destacamos una población vieja poco metálica que podrían ser estrellas del halo y otra sobredensidad de 3 Gyrs de edad y [Fe/H]=-0.6. Tras comparar nuestras soluciones con publicaciones de estudios realizados en este campo, pudimos verificar elementos coincidentes, reforzando así la confianza en nuestros resultados.

Abstract

We aim to determine for the first time the accurate Star Formation History (SFH) of the the Gaia Catalogue of Nearby Stars (GCNS), a clean and well-characterised bubble of 100 pc radius from the Sun, using the high quality Gaia eDR3 data. Since Colour-Magnitude Diagrams in the absolute plane (aCMDs) harbour information of stellar populations, they are an excellent tool to retrieve SFHs. Therefore, we employed an updated aCMD-fitting technique through the comparison between the observed aCMD and the aCMDs synthesised according to the predictions of stellar evolution libraries.

Special attention was given to the local unresolved binary systems residing below the old Main Sequence (MS) turn-off point, tracing a brighter and parallel band to the MS where single stars lie. We seeked for the best fit of the synthetic binary population to the data by exploring different binary fractions and mass ratios in the models. Additionally, we tested other parameters intrinsic to the code and others in relation to the synthesised diagrams to find the best configuration and derive the SFH of the data-set.

Our study suggested that the best binary recipe in concert with the observational data was a binary fraction of 0.3 and a minimal mass ratio of 0.1, together with the simulation of a 100 million star mother diagram using BaSTI. We found that the GCNS mainly comprises solar-metallicity stars and experienced several star formation epochs (around 0.5, 2.3, 5.5 and possibly 10 Gyrs) combined with quiescent periods. Other particular features regarding the Star Formation Rate (SFR) in the age-metallicity plane revealed structures like a low density of metal-poor and old stars possibly reflecting the halo population, and a stellar over-density of approximately 3 Gyrs old with [Fe/H]=-0.6. All these characterisations are in good agreement with previous studies of other data volumes of the Milky Way.

1 Introduction

Galaxies are gravitationally bound entities, typically consisting of dark matter, gas, dust and stars. They populate the Universe, mainly residing in clusters and groups. The most well-known galaxy is our own, the Milky Way, a spiral galaxy that, together with its galactic neighbours, belongs to the Local Group. The exact number of galaxies in the Local Group and its immediate surroundings is unknown as some are occluded by the Milky Way. However, about 80 galaxies of different morphological types, most of which are dwarf galaxies, are known. The group itself is a part of a larger one, the Virgo Supercluster which, on a larger scale, may be a part of the Laniakea Supercluster.

Unravelling the history of our nearby Universe is of utmost importance to understand the formation and evolution of all types of galaxies. This is the goal of galactic archaeology, which studies many local examples that can be resolved into individual stars with a clearness impossible elsewhere. Detailed observations of stellar fossil records, gas and other structures are the key to assemble past events in the galaxy's lifetime and thus, trace their formation and evolution history. For this purpose, we focus on a portion within our reach, as it is our privileged position within the Milky Way, to study it with a level of detail currently only attainable for this galaxy. We will use this exceptional laboratory to perform galactic archaeology and shed light into the enigmatic history of spiral galaxies, by opening a unique window directly onto the first epoch of star formation in the Universe, and following the evolution of the solar neighbourhood up to the current time.

The aim is to study in detail today's snapshot in time to find clues to help us understand the history of development of our Galaxy. Our knowledge and registered sources of the Galaxy, including the local region, is incomplete so far (Gaia Collaboration, 2021). However, a step forward has been made possible thanks to the successful achievements of the *Gaia* mission, a space observatory of the European Space Agency (ESA), launched in 2013, designed to gather stellar measurements of 1.7 billion stars in our galaxy with unparalleled precision. Its main ambition is to chart the most precise 3-dimensional space catalogue ever made, and producing a kinematic census of about one billion stars in our Galaxy. It is continuously scanning the sky measuring the positions, photom-

etry and spectra of stars with unprecedented precision and accuracy. With such exceptional information provided, we are able to derive magnitudes and colours, proper motions, parallaxes and thus distances, chemical composition and radial velocities, hence all this essential information to decode the formation and evolution of the Galaxy (Santos & Jordi, 2020).

Throughout this project we used the recently published results in the well-characterised *Gaia* Catalogue of Nearby Stars (GCNS). It comprises the complete and deep census of 331312 stars residing within a sphere of 100 pc radius centered on the Sun, selected from the full *Gaia* EDR3 sample. All objects with reliable and highly accurate astrometry and photometry down to the nominal G=20.7 magnitude limit of *Gaia* are entailed within the volume of data.

In addition, the GCNS is a concise, clean and close sample since (i) it counts on a high degree of completeness, (ii) the influence of extinction is practically non-existent and the photometric errors are minimal and, (iii) it is the best observed and classified set of objects with a well-known uniform composition to that of our Sun. All these features, make of this data set look like an almost theoretical sample.

Despite its incredible resolving power and astrometric accuracy, *Gaia* is not able to resolve certain binary stars in their individual component stars. This sample is an excellent test set for exploring the unresolved binary population. This is one of the main goals of this project, but to achieve it, we will retrieve even more important, and so far unclear, information for the solar neighbourhood: its detailed Star Formation History (SFH) from its early times to the present.

The SFH is usually defined as a function of both the Star Formation Rate (SFR) and the chemical enrichment in terms of their respective evolution in time (Aparicio & Gallart, 2004). It also involves other characteristics of a stellar system such as the Initial Mass Function (IMF) or the fraction and mass distribution of binary stars. Colour-Magnitude Diagrams (CMDs) that reach the oldest MS turnoffs (defined below), display stars born throughout the lifetime of a galaxy and are fossil records of its SFH. Retrieving a quantitative and accurate SFH from the observed CMD, through comparison with synthetic CMDs calculated using the predictions of stellar evolution theory, after a careful simulation of the corresponding observational errors is complicated and requires relatively sophisticated techniques (Gallart et al., 2005), that we will refer to as CMD-fitting for simplicity. These techniques have been applied to CMDs in the absolute plane, namely aCMDs, for stars located in a large volume of the

Milky Way, encompassing different Galactic structural components, resulting in excellent SFHs and stellar age and metallicity distributions (Gallart et al., 2019).

The GCNS data-set is ideal to determine the SFH by means of the CMD-fitting technique. Different ingredients are needed for the computation of the required synthetic aCMDs. One of which is a parametrisation of the unresolved binaries (i.e. considering different binary fractions and distributions of masses of the secondaries). On account of that, in the present work we intend to use the GCNS sample to empirically constrain the properties of the unresolved local binary population. In the process, we will obtain accurate SFHs and examine how different decisions about the parameters used in the SFH derivation process influence the outcome, which will be useful to learn how best to approach the SFH in other, less well-characterised volumes of our Galaxy.

The report is structured as follows: In Section 2, we provide an overview of the data introducing the highlights of the GCNS, followed by a careful study of the stellar content within the data volume: a sample quality test, a study of the impact of reddening and an analysis of the stellar content and the accuracy of the stellar models by comparing isochrones with the aCMDs. Section 3 discusses the methodology to derive the SFH and to estimate the characteristics of the unresolved binary population. The obtained Results are presented in Section 4, testing the input parameters of the SFH derivation code to select the most suitable for our case, the further study of the binaries and the computed final SFH of the GCNS. Finally, we summarize the conclusions reached and briefly mention what lies ahead in Section 5.

2 The data

We performed an analysis to become acquainted with the data in the catalogue. The foremost particularities of the sample are presented within the present section.

2.1 The *G*aia Catalogue of Nearby Stars (GCNS)

Exploring the essentials of the GCNS we found that it is a clean and well-characterised catalogue of 331312 objects in the solar neighbourhood. Comparisons made to previous catalogues as *Gaia* DR2 revealed that the level of contamination has decreased in the recent results (Gaia Collaboration, 2018).

To create this data-set, Gaia Collaboration (2021) used information from the *Gaia* EDR3 including source ID, basic astrometry, photometry, and a number of quality flags. The selection of candidates within the sphere of 100 pc involved selected training sets, machine learning procedures such as random forest classifiers and quality flag indicators. Distance probability densities were computed for all sources using the parallaxes and a simple distance prior not depending on sky position or type of star. Specifically, the prior was the distance distribution of stars in the GeDR3mock (Rybizki et al., 2020) catalogue for stars with observed parallax greater than 8 mas.

Gaia Collaboration (2021) provide a detailed examination of the open clusters contained in the volume, like the Hyades and the Coma Berenices, the population of white dwarfs and wide-binary systems (a catalogue of 16556 high-probability resolved binary candidates with members classified into the corresponding types). Moreover, 12 members of Gaia Enceladus have been identified among the detection of local streams, and halo objects. A total of 74281 stars have validated radial velocities, 95% of which are thin disc stars, 4.6% are part of the thick disc stellar population and finally, 0.4% belong to the galactic halo.

Using the median distances provided in the catalogue ($d_{\rm pc}$=Dist50), and the ultra-precise *Gaia* EDR3 photometric data, we converted apparent G magni-

tudes (G) into absolute via:

$$M_G = G - 5\log_{10}(d_{\rm pc}) + 5 \tag{2.1}$$

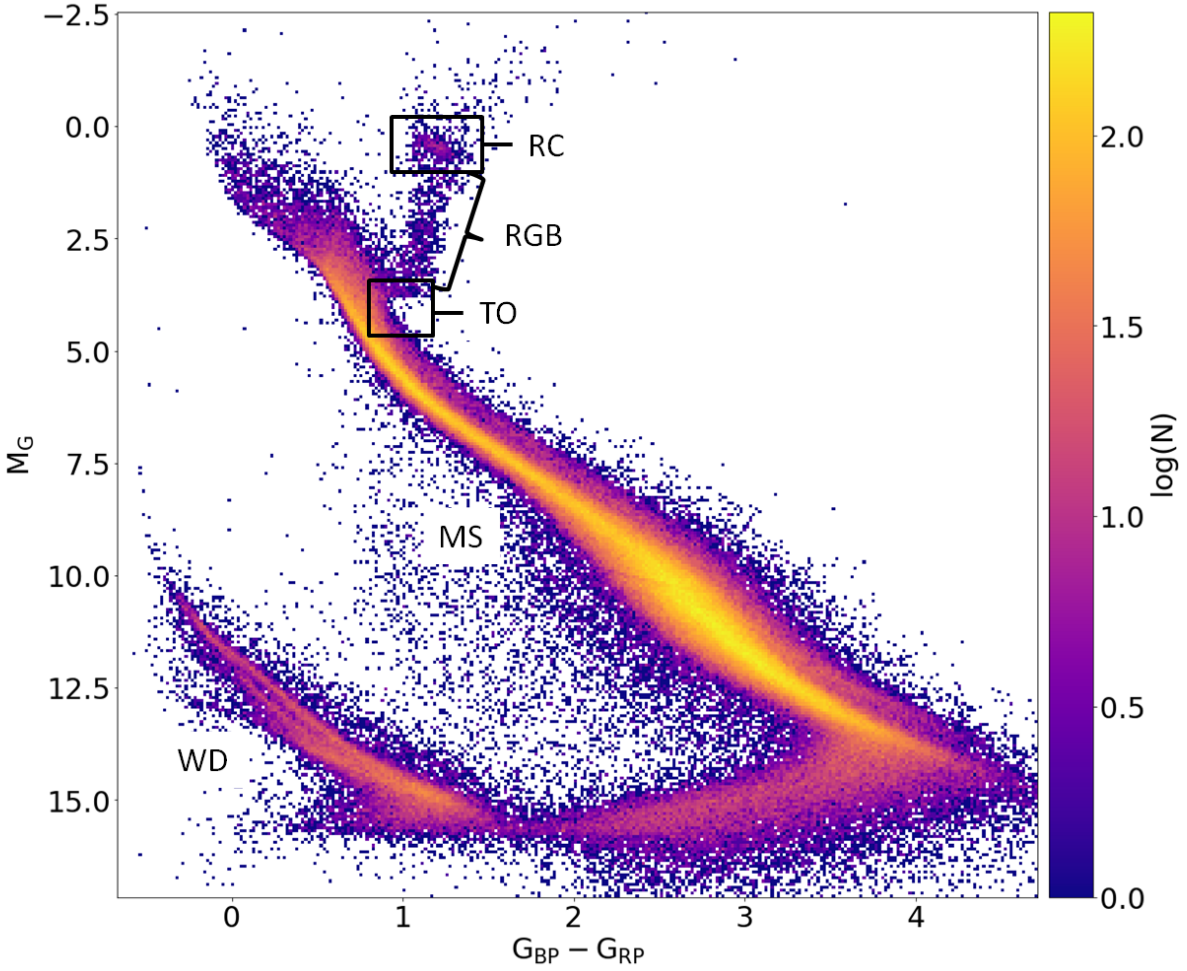

Figure 2.1: Absolute Colour Magnitude Diagram of the complete GCNS sample, colour-coded according to the logarithmic number of stars. All the components are labelled. Note that the highest density of stars is located in the Main Sequence.

In this way, in Fig. 2.1 we show the absolute Colour Magnitude Diagram (aCMD) of the GCNS. The main components of this high quality aCMD are the following: the Main Sequence (MS), a diagonal band from bright-blue ($G_{BP} - G_{RP} \sim 0$ mag and $M_G \sim 0$ mag) to faint-red ($G_{BP} - G_{RP} \sim 5$ mag and $M_G \sim 15$ mag), containing H-core burning stars. These stars will leave the MS after the exhaustion of Hydrogen tracing the Turn-Off point (TO), that continues to a "knee"-shaped region that is the most accurate means of determining the stellar ages, i.e. the lower the turnoff point, the older the population of stars. After that, the turn-off stars enter a more evolved stage fusing H in shell: they become giants and ascend the Red Giant Branch (RGB), rising their luminosity and size until they start the core-helium ignition stage. It is then when they become Red Clump (RC) stars, constituting a clustering region found at $G_{BP} - G_{RP} \sim 1.5$ mag and at $M_G = 0$ mag. Last of all, located below the MS, lies the White

Dwarf (WD) population, a density of extremely faint and usually blue stars. Finally, the strip of source with nominally good solutions connecting the main and WD sequence, found at $2 < G_{BP} - G_{RP} < 4$ mag and $M_G = 15$ mag is unexpected and due to contamination of the GCNS by faint objects at distances of 80-120 pc, according to Gaia Collaboration (2021).

Gaia Collaboration (2021) also provide a cross-match with photometry from other catalogues such as GUNN (g, i, z) and 2MASS (j, h, k). We compared aCMDs combining different photometric passbands other than *Gaia* and found that they showed a significant lack of quality when compared to the complete aCMD obtained with the *Gaia* bands. Additionally we took into account all the possible combinations of the three *Gaia*'s photometric passbands. The most compact and finest aCMD of all was that using $G_{BP} - G_{RP}$. We believe this is because these two bands are independent one from another. Therefore, we proceeded with our further analyses using this colour.

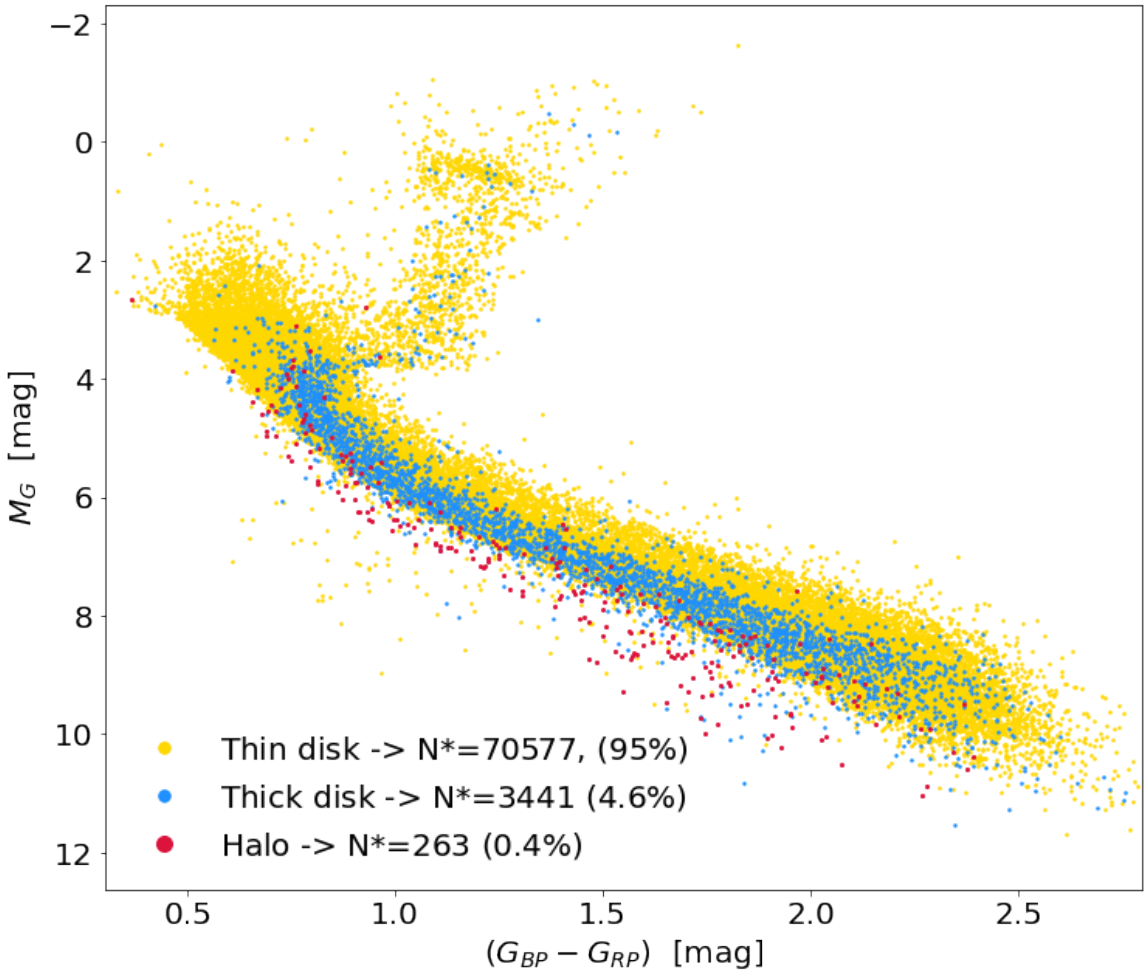

Figure 2.2: aCMD of the 74281 stars with verified radial velocities, classified according to the galactic components: the thin disc (yellow), the thick disc (blue) or the halo (red).

As mentioned above, Gaia Collaboration (2021) also found that the GCNS sample comprises stars from three Galactic components. First of all, the thin

disc, the site of ongoing star formation. It is the most characteristic structure of the Galaxy and most of its stars move on fairly circular orbits. In second place, a more diffuse and hotter component, the thick disc, populated by older stars. Last but not least, the galactic halo, the most extended spheroidal old component.

In Fig. 2.2 we represented the stars identified kinematically as belonging to the thin and thick disc and halo labeled with different colours. We also provide the numbers in each population verifying their correspondent percentage towards the total of stars with valid radial velocities. The more numerous thin disc stars trace all the CMD components, namely the MS, the RGB and the RC. The MS of thick disc stars does not reach a magnitude as bright as that of the thin disc, indicating an older age, and it is concentrated in the blue side, hinting lower metallicity. Lastly, the halo stars, the eldest of the sample, are only spotted in the MS instead, lying in a fainter, bluer locus than the disc stars, hence indicating even lower metallicities.

2.2 Preliminary study of the data-set

Subsequently, we present a prompt review of the initial study performed to the data-set, consisting in a quality check, the analysis of the influence of the extinction and the inference of the general properties of the stellar populations present in the aCMD by comparing with isochrones from two different stellar evolution libraries.

2.2.1 The quality of the data

Various errors affect the determination of the stars in the observed aCMD. For this reason, it was fundamental to perform a first analysis to the main sources of uncertainty, such as the photometric errors and the errors in the calculation of the stellar distances.

We examined the effect of the distance errors with the median stellar distance. In Fig. 2.3, we plotted the distance uncertainties, as a function of the median stellar distance (Dist50). The 1σ uncertainties have been defined as (Dist84-Dist16)/2, using the 84 and 16 percentile of the distance probability distribution for each star. As can be noticed, more distant stars have larger errors than the nearest within the sample.

In Fig. 2.4 we plotted the photometric errors as a function of the G magni-

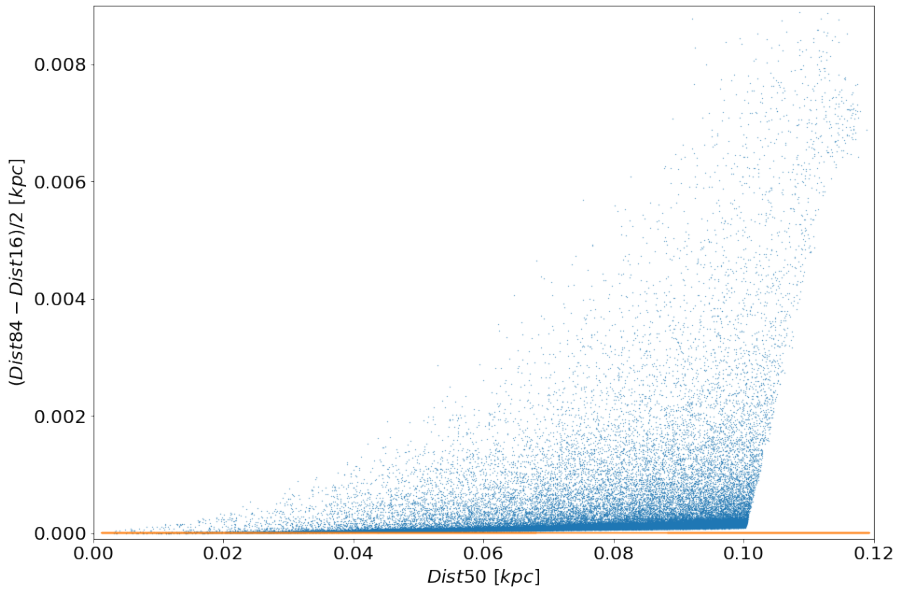

Figure 2.3: Distance uncertainties defined as (Dist84-Dist16)/2 versus the median stellar distance (Dist50). The solid orange line indicates the position of the zero on the abscissa.

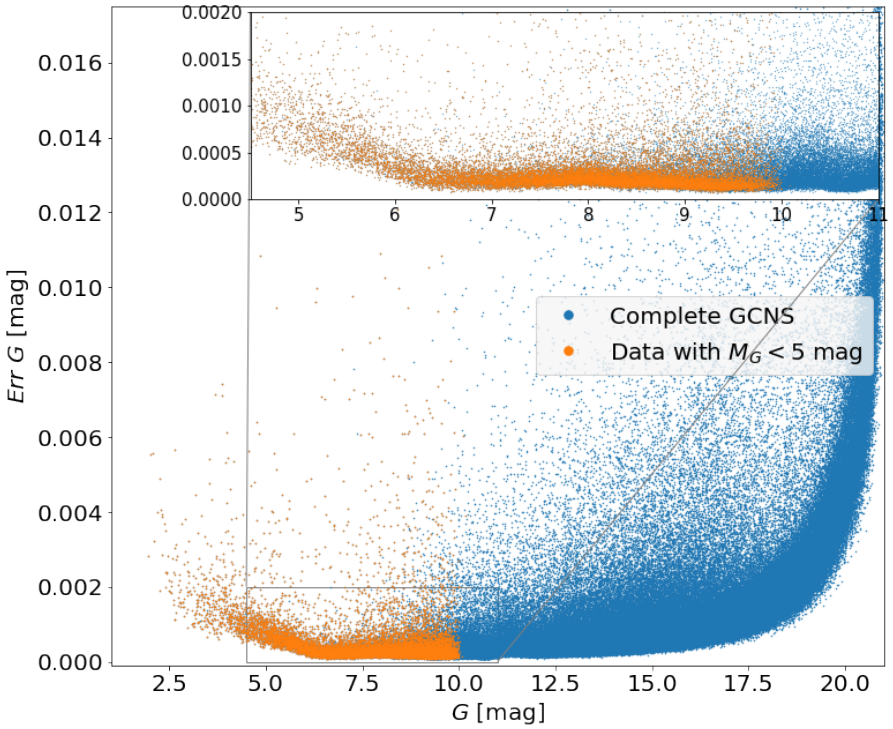

Figure 2.4: Photometric error of the apparent G magnitude vs the G magnitude. The data fulfilling the condition $M_G < 5$ mag was highlighted in orange.

tude[1]. To check how big are the underlying errors within the portion of the diagram that we will use to derive the SFH (i.e. stars brighter than 5 mag in absolute G magnitude, emphasized stars in orange). For the G band, the values of the photometric error were considered directly without taking into account the extinction[2]. We found an increasing trend of photometric errors

[1]To convert the flux error provided by *Gaia* to the photometric G errors we used Err G= $\left(\frac{2.5}{ln(10)}\right)\left(\frac{1}{\texttt{phot\_g\_mean\_flux\_over\_error}}\right)$

[2]To convert to absolute magnitudes it would be necessary to consider extinction but in the following Section 2.2.2 it is shown that its effects are negligible in our data-set.

with magnitude for stars fainter than $G \sim 6$ mag. However brighter stars than this limit have increased photometric errors. Overall, the photometric errors are in fact very small (~ 0.001 mag), specially for most of the stars within the region of interest.

2.2.2 The effects of the interstellar extinction

The so-called interstellar extinction is due to the absorption and scattering of electromagnetic radiation by gas and dust existing in the interstellar medium between an emitting source and the observer. Since blue light is more strongly attenuated than red light, extinction causes objects to appear redder than they actually are. Hence, this phenomenon is also referred to as reddening.

We determined whether a reddening correction was necessary to perform a more accurate analysis. This study was not carried out in Gaia Collaboration (2021), so it is merely our own. When taking into account the interstellar extinction, the absolute magnitude and colour of any star within the data sample are described as follows:

$$\begin{aligned} M_G &= G + 5\log(d_{pc}) + 5 - A_G \\ (G_{BP} - G_{RP})_0 &= (G_{BP} - G_{RP}) - E(G_{BP} - G_{RP}) \end{aligned} \tag{2.2}$$

We used dustmaps, a unified interface for several 2D and 3D maps of interstellar dust, reddening and extinction. In particular, we used the maps from Green (Green et al., 2019), based on *Gaia* parallaxes and stellar photometry from Pan-STARRS 1 and 2MASS, covering the sky north of a $dec > -30$ deg, amounting to three-quarters of the sky.

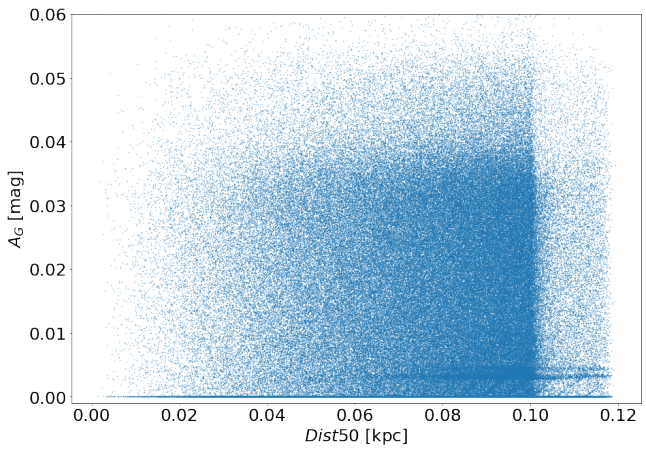

Figure 2.5: Extinction versus the median stellar distance.

We could expect to find increasing values of the interstellar extinction in the

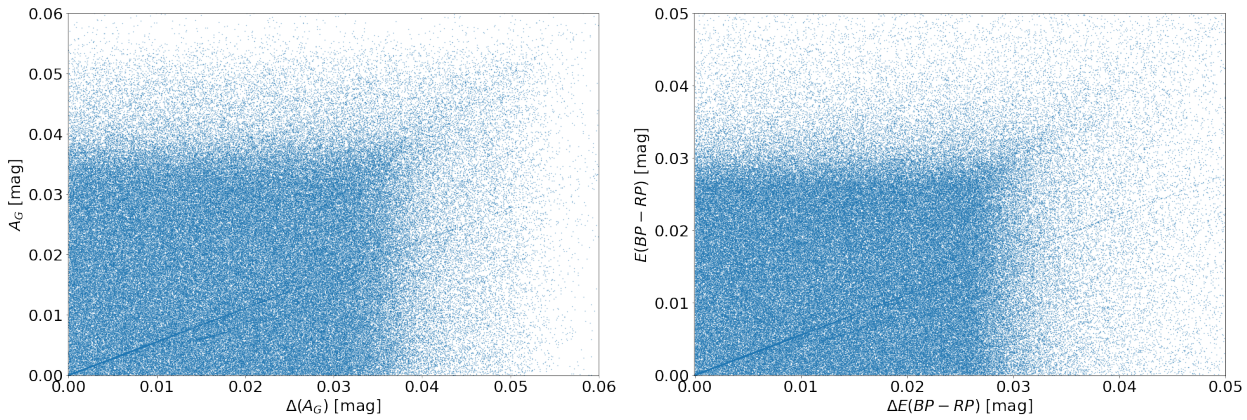

Figure 2.6: Left panel: Extinction versus its corresponding error. Right panel: Colour excess factor versus its error.

G band (A_G) with the distance but we found another behaviour instead. In Fig. 2.5, a homogeneous trend of the extinction can be noticed throughout the whole extension of the sample. An equal situation was observed in both panels of Fig. 2.6, which also shows a uniform pattern when comparing the extinction and the reddening (or colour excess, $E(G_{BP} - G_{RP})$) with their corresponding errors. The values of both variables are small and comparable to their errors. We conclude that such low values ensure that the corrections in Eq.2.2 can be therefore neglected.

2.2.3 Comparison with isochrones of different sets

A comparison between the observational data of the GCNS and the stellar evolution libraries, such as BaSTI (Hidalgo et al., 2018) and PARSEC (Bressan et al., 2012), was performed as a test for the models to see how well they reproduce *Gaia*'s excellent quality data.

In Fig. 2.7 we overplotted a set of isochrones of various ages from young to old (0.1 Gyr-12.1 Gyr) at the closest metallicity to the solar (Z=0.017[3]) for each of the models. In the left panel, it can be noticed that the BaSTI library appears systematically blueshifted from the observed data in the faintest locus of the MS, particularly in the range $1.5 < G_{BP} - G_{RP} < 4$ mag and from $M_G = 8$ mag to the fainter magnitudes. In a simultaneous study to this one with Open Clusters (Alicia Rivero's TFG), we found the same mismatch.

In the case of BaSTI, the disagreement quite probably relied in the adopted colour-T_{eff} relationship, adopted when transferring the theoretical predictions from the H-R diagram to the *Gaia* photometric system. Dr. Santi Cassisi pro-

[3]Note that the chosen solar metallicity is the initial metallicity that a solar-type star should have, so that at its current age of 4.5 Gyr, the measured photospheric metallicity due to diffusion is of Z=0.0153.

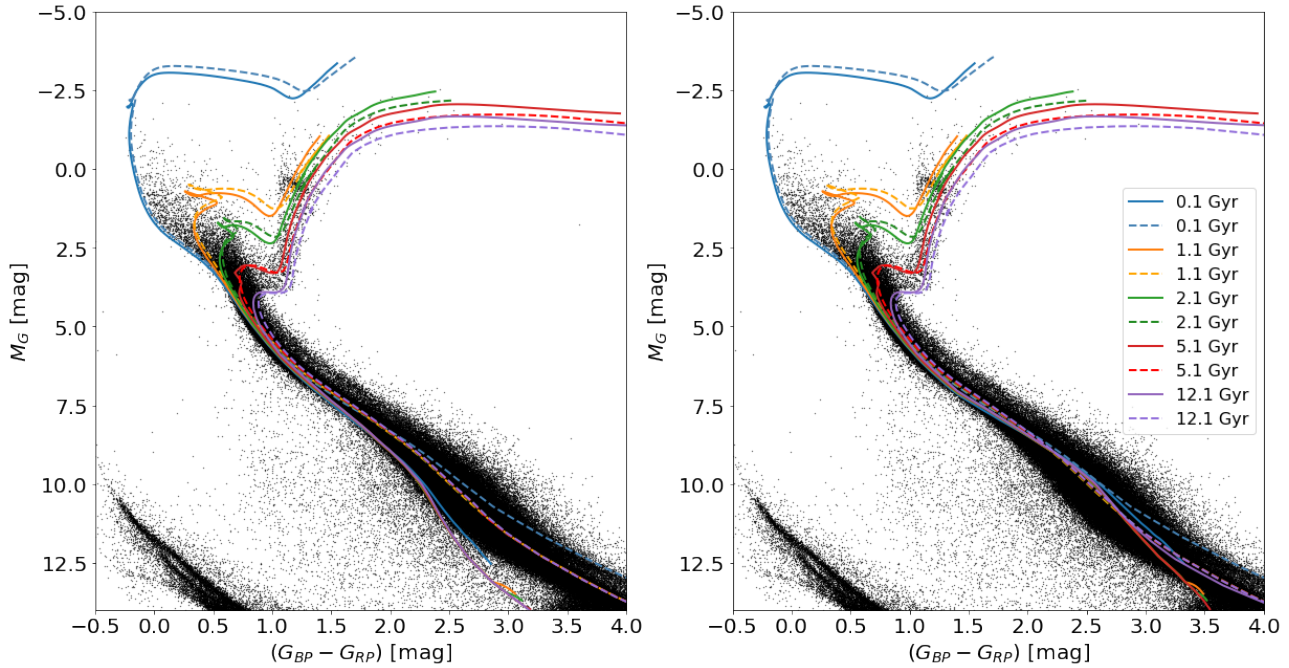

Figure 2.7: Set of different age (from 0.1 Gyr to 12.1 Gyr) isochrones of solar metallicity, Z=0.017, from the BaSTI (solid lines) and PARSEC (dashed lines) libraries, plotted with the GCNS data (in black). The left panel shows the original BaSTI isochrones, while the right panel presents the corrected BaSTI isochrones according to the Eq. 2.3.

vided updated bolometric corrections with the support of the empirical data collected by Mann et al. (2015). Hence, yielded an inversely proportional dependence of the corrections with the T_{eff}. Such corrections were only applied to the colour when fulfilling certain conditions, as described below:

$$\begin{aligned} &\texttt{if} \;\; T_{eff} \leqslant 4400K \;\; \& \;\; log(L/L_{\odot}) < -0.50 : \\ D_{colour} = -63.39 + 0.069271 \cdot T_{eff} - 2.7736 \cdot 10^{-5} \cdot T_{eff}^2 + \\ &4.8655 \cdot 10^{-9} \cdot T_{eff}^3 - 3.1723 \cdot 10^{-13} \cdot T_{eff}^4 \end{aligned} \tag{2.3}$$

Thus, the corrected *Gaia* colour is: $G_{BP} - G_{RP} = (G_{BP} - G_{RP})_{BaSTI} + D_{colour}$

Getting back to the Fig. 2.7, the right panel represents the same set of isochrones after applying the above-mentioned corrections. Notorious improvements are noticed for the BaSTI isochrones in the lower MS of the empirical data. Although these corrections won't affect our study, because we will focus on brighter magnitudes, this contribution to the improvement of the BaSTI model has provided awareness of this discrepancy in the low end of the MS and finding such an effective solution is very important for future studies.

There are slight differences between BaSTI and PARSEC, i.e the PARSEC isochrones prevail systematically redder and brighter in the red-giant branch and sub-giant branch, respectively than the BaSTI isochrones. But in general,

all the populations are well reproduced with solar metallicity isochrones within the selected age range, from the youngest to the eldest. However, bluer and fainter stars than the MS are not appropriately fit by this isochrone set.

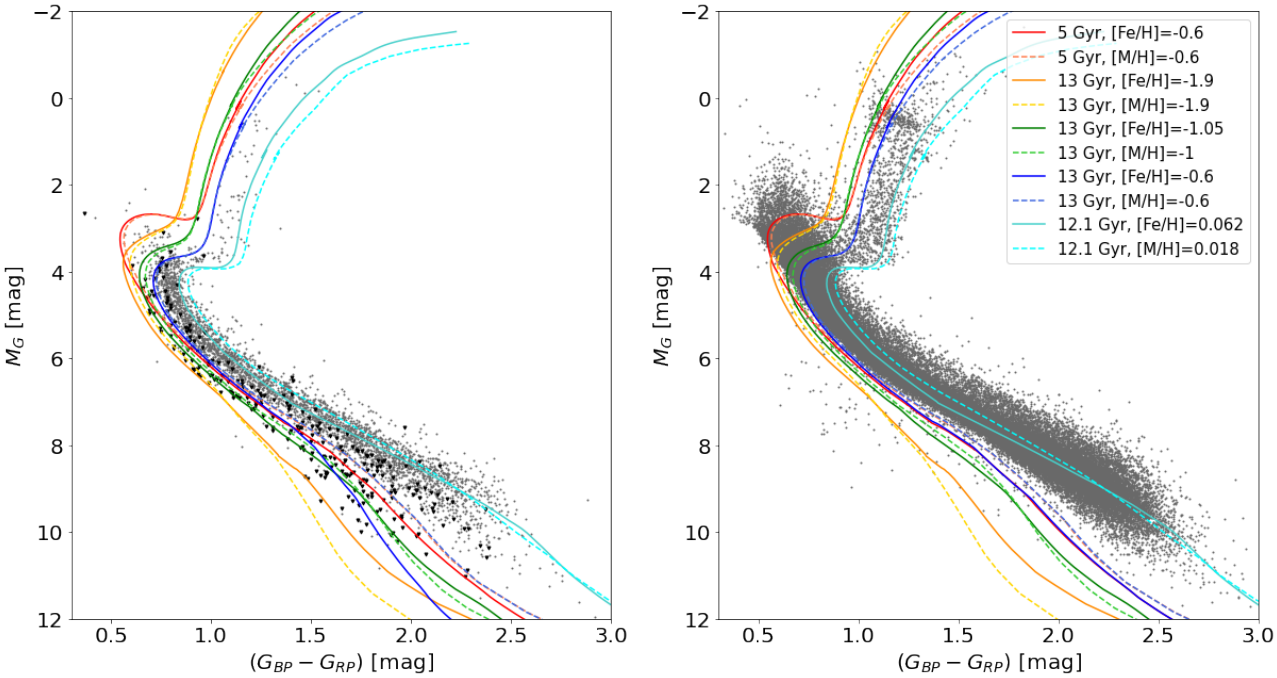

Figure 2.8: Set of isochrones of 5, 12.1 and 13 Gyrs of different metallicities ([Fe/H]=-1.9,-1,-0.6, 0.062 and 0.018) from the BaSTI (solid lines) and PARSEC (dashed lines) libraries. The GCNS data is plotted again separating the stars in the different Galactic components. In the left panel there are stars of the thick disc (grey) and the halo (black), while the thin disc stars are plotted in the right panel.

Fig. 2.8 represents corrected isochrones based on the Eq. 2.3 for both models, with a range of metallicities. We chose old isochrones (13 Gyr) of [Fe/H]=-1.9,-1 and -0.6 and a 12.1 Gyr isochrone with [Fe/H]=0.062 for BaSTI and [M/H]=0.018 for PARSEC, in addition to a mid-age isochrone of 5 Gyr with [Fe/H]=-0.6, respectively. In this plot we only presented the selected stars with valid radial velocities leading to a cleaner, but not complete sample. Therefore, there are many of the brightest stars and the bluest stars lying below the MS that do not appear in the figure.

The observational data in the background are GCNS stars from the different galactic components: the thick disc and halo on the left panel and the thin disc on the right. We checked what was mentioned above in Fig. 2.2. In general, the halo stars are the bluest, eldest and least metallic of all, and are relatively well matched by 13 Gyr and [Fe/H]=-1.9, -1 and -0.6 isochrones. Thick disc stars on the other hand, are well fitted by old and more metallic isochrones, as the 13 and 12.1 Gyr with [Fe/H]=-0.6, 0.062 and 0.018 isochrones, which constrain accordingly the metallicity range of the stars comprising the thick disk. Note that the metal poor isochrones don't match the thick disc stars in the MS. The 5 Gyr old isochrone provides a lower limit to the possible age of thick disk stars,

but note the low density of stars between the old isochrones and the 5 Gyr old one. Finally, the stars from the thin disc reach brighter magnitudes, cover a wide range of younger ages and are more metal-rich than the former. Despite being old, the 12.1 Gyr isochrone is the best option to simulate this population within this isochrone set. Providing that, we reject the other older and more metal-poor isochrones to fit this region.

To summarise, both stellar evolution models present systematic offsets, which may lead to slight age differences in the SFH that we will compute. However, we will bring special attention to the MS because it contains crucial age information, particularly we will focus in the old turn-off region upwards. Hence, it is the most representative part to consider when we derive the SFH.

3 Methods

3.1 Deriving the SFH

Within this section we determined the SFH of the GCNS sample by using the CMD fitting technique. The method is based on the comparison between the distribution of stars in the observed CMD and that in synthesised stellar populations computed from the predictions of stellar evolution libraries such as BaSTI and PARSEC. We will call these theoretical CMDs *mother diagrams*. We have computed several of them under a number of assumptions, including different stellar models and unresolved binary recipes.

Synthetic Mother Diagrams						
Name	q	β	Stellar library	Lim. M_G	Calc. N*	N* ($M_G \leq 5$)
q01b03_100Mbasti_MG11	0.1	0.3	BaSTI	11	100M	7360717
q01b03_100Mparsec_MG10	0.1	0.3	PARSEC	10	100M	4062354
q04b03_30Mbasti_MG11	0.4	0.3	BaSTI	11	30M	2210156
q01b07_30Mbasti_MG11	0.1	0.7	BaSTI	11	30M	2205296
q06b01_30Mbasti_MG10	0.6	0.1	BaSTI	10	30M	3536055

Table 3.1: Features of the synthetic mother diagrams used throughout the analysis to derive the SFH of the GCNS. From left to right: the name by which we identify them, the minimum binary mass ratio (q), the binary fraction (β), the stellar evolution library, the limiting M_G magnitude, the number of stars computed ('M' means millions of stars) and, finally, the number of stars with a $M_G \leq 5$ mag.

Table 3.1 displays the main features of the synthesised mother diagrams that we used in our study. Note that the cut of the absolute G magnitude was 10-11 mag but for the calculation of the SFH, only brighter stars are used ($M_G \leq 5$ mag), thus, reducing significantly the number of stars of the samples. This number of stars is given in the table for reference. The mother CMDs contained stars with a flat distribution of age, ranging from 0.08 to 13.5 Gyr, and metallicity, $0.0001 < Z < 0.039$ for BaSTI and $0.0001 < Z < 0.027$ for PARSEC, respectively. In both cases, the models we used had solar-scale elemental abundances. We assumed for all of them a Kroupa IMF (Kroupa, 2001) and varied the stellar evolution model involved in the calculation and the parameters of the unresolved binaries for each test.

Previous to the comparison between the observed and synthetic star distribution across the aCMD, it is necessary to simulate the observational errors in

the synthetic CMD. For that, we used and algorithm provided by co-advisor T. Ruiz-Lara. We considered that the main sources of uncertainty affecting the position of an observed star in the aCMD were solely the photometric errors and the error in the determination of the distance. In order to implement these observational effects on the synthetic diagrams, we first assigned to each synthetic star a distance following the global distribution of stellar distances in the GCNS ('Dist50'). This preliminary step allowed us to, applying the relation between apparent and absolute magnitudes, move our synthetic CMD to the apparent plane (extinction was not considered). As shown in Riello, M. et al. (2021) and in Fig.2.4, there is a clear trend of photometric errors with magnitude, with the brigther tail ($G \gtrsim 6$) presenting larger uncertainties, and a smooth trend to larger uncertainties for fainter stars as well. We found a similar trend with distance as well (see Fig. 2.3), more distant stars displaying larger uncertainties, which we defined as (Dist84-Dist16)/2.0, being 'Dist84' and 'Dist16' the 1σ upper and lower bound of the distance probability distribution for each star. In both cases, we fitted a 5th degree polynomial (P_i) to the run of the error in the parameter i as a function of the value of such parameter, with i corresponding to apparent BP, RP, G or distance. Also, we characterised the running standard deviation of each parameter (σ_i). Thus, to each star s with a given set of parameters i (G, BP, RP, distance), we assigned random errors (ϵ_i) following a Gaussian centred at $P_i(i_s)$ and with sigma $\sigma_i(i_s)$, with i_s referring to the value of the parameter i for the star s.

Once we have photometric and distance errors for each synthetic star, we perform a quadratic propagation of uncertainties to translate these observational errors into an error in the absolute magnitude and colour. Finally, we proceed to simulate the errors by modifying the colour and magnitude of each star by adding to its colour and magnitude a correction following a Gaussian distribution centred at zero and, with sigma the propagated error in colour and magnitude. Regarding completeness, we have considered that the sample is complete and thus have not applied any completeness simulation in the mother aCMD.

The SFH derivation has been carried out using *dirSFH* [Gallart et al. in prep.] for the first time, which is an improvement and extension of the well-known tools IACpop (Hidalgo et al., 2009) and TheStorm (Gallart et al., 2021). The most important difference lies in the way the individual Simple Stellar Populations (SSPs) in the synthetic aCMD are defined. The SSPs are stars in a small range of age and metallicity. The age and metallicity bins are treated as "seed points" to perform a Dirichlet tessellation (Green & Sibson, 1978) with which the SSPs have not 'square' shapes anymore, to better sample both parameters. Another

new implementation to the code is that there are two weighting strategies as a function of colour and magnitude: "uniform", meaning flat all across the CMD, or "weighted", as the logarithm of the inverse of the variance of the ages across the synthetic aCMD. In other words, weights have been implemented based on how precisely a given region in the CMD is representative of a population of a given age.

The code is based in 4 different processes:

- *dirPrep*: prepares data for processing, reads the mother and the observed CMDs specified in the input file, and converts them in a new format for easier access.
- *dirWin*: a window pops up and allows the user to define the region of the aCMD to be used as the main delimiter of the fitting space for the SFH computation (bundle). Once outlined, it calculates the weights across the CMD and the main CMD matrices.
- *getSM*: runs several SFH estimations shifting the observed aCMD within a specified maximum range of colour and magnitude to sample the SFH with the purpose of detecting possible offsets between the observational data and the stellar evolution models.
- *dirall*: a number (specified by the user) of SFHs are calculated with different dissections of the age-metalliciy space. These are obtained by randomly changing the age and metallicity seed points, within the boundaries of their total ranges, and keeping their overall distribution as a function of age and metallicity. From these seed points, different Dirichlet tesselations of the age-metallicity space are produced, which result in the SSPs to be used for each SFH run. Then, all the solutions are combined to produce the final SFH and uncertainties are derived from the variations between the individual solutions.

In Fig. 3.1 we show the output figure from executing the code and which summarizes the SFH results in different projections. The two panels on the upper left part of the figure are the observed (GCNS) and the solution CMDs, respectively, both of them bounded by the defined bundle. The following top plot represents the SFR as a function of time, whereas the last top panel displays the Cumulative distribution of SFR over the years. In addition, the first bottom panel includes the residuals between the observed and solution CMDs, which allows us to seize how good is the fit. The next panel displays the SFR in the age-metallicity plane, which provides an insight into the epochs of star formation at different ages and their metallicities. Finally, the last plot reveals

the Metallicity Distribution Function (MDF).

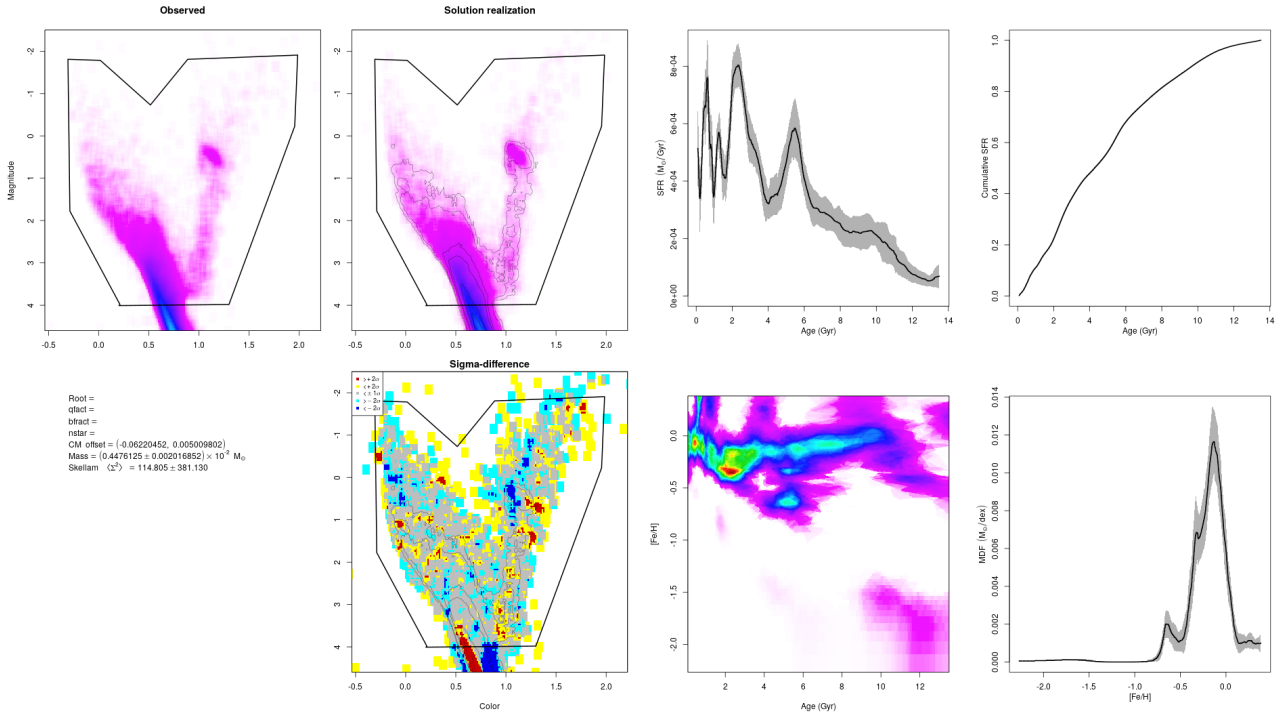

Figure 3.1: Output plots of the code *dirSFH*. From left to right, top panels: Observed CMD, solution CMD, SFR, cumulative SFR distribution. Id., bottom panels: residuals between the empirical data and the resulting solution CMDs, SFH in the age-metallicity plane and MDF.

3.2 Assessing the binary recipe

From the study of the MS structures, it is possible to obtain information about the unresolved binary stellar systems. Binaries with a MS primary and a giant companion would lie in the RGB region, on the other hand, binaries with a late-type main-sequence primary and a WD companion are usually found between the WD and the MS (Penoyre et al., 2022). When both stars are MS stars, they appear offset upward ~ 0.5 mag from the MS ridge line. We will concentrate in the latter, which are the ones most easily identifiable.

To derive the SFH we focused on the portion above the oldest MS turn-off, brighter than approximately M_G= 5 mag, which is the one containing most age information (see Section 2.1). However, the clearest clues of the unresolved binary population can be observed in a range of magnitudes below the old MS turn-off, in the form of a parallel, less populated sequence, above and to the red of the main sequence of single stars. We will derive SFHs with several mother diagrams calculated adopting different binary fractions and secondary mass ratios, and compare the distribution of stars in the main and binary sequence in this area, in order to find the best fit with the observational data. Hereafter,

we display into a little more detail the followed procedure.

Specifically, to study the unresolved binary fraction we will limit to the colour range, $0.95 < G_{BP} - G_{RP} < 1.4$ mag, and to the absolute G magnitude, $4 < M_G < 7.5$ mag.

As can be seen in Fig. 3.2, we compared the MS of the observational data (in blue) with that of the BaSTI aCMD solution of 100M stars (in orange). Note that the model is systematically offset towards fainter magnitudes within our boundaries. Below the MS, there is a trace of stars that are more numerous in the BaSTI's solution CMD than in the observed data. In this region we could expect the presence of field stars known as 'sub-dwarfs' because they are low-mass and metal-poor stars. Furthermore, there is a brighter tail of stars lying above the MS, differentiated from the single star MS, which we identify as the binary sequence.

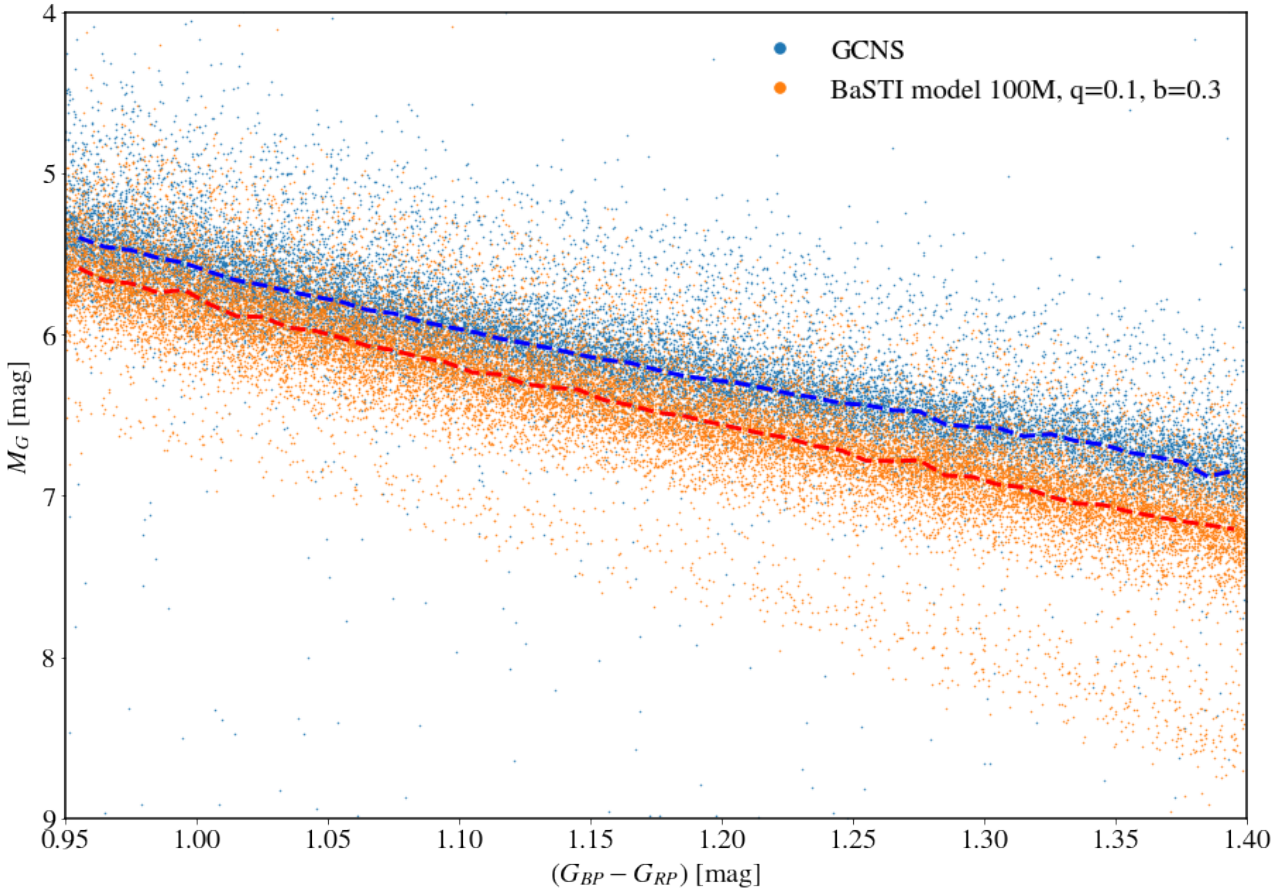

Figure 3.2: Zoomed region of the CMD of the GCNS data (blue) and the BaSTI model (orange). The blue and red dashed lines are the polynomial fits applied to the GCNS data and the BaSTI solution, respectively.

The stellar distribution as a function of M_G for each colour will reveal the binary distribution in relation to the single stars. As there is a shift in the position of the MS between the observations and the models, the M_G distributions

will not be directly comparable. Thus, we will perform this comparison on a relative basis with respect to the maximum of each of the distributions.

To study the binary region, we will use Gaussian Mixture Models (GMMs) for histograms defined in two different ways (see below). A Gaussian Mixture is a function that is comprised of several Gaussians, each identified by 3 parameters: the mean defining its center, the covariance referring to its widths and the mixing probability setting how big or small the Gaussian function will be.

First, we used GMMs to trace the position of the MS peak as a function of colour, in colour bins of 0.01 mag. For each colour bin, we calculated the maximum of the M_G distribution, as the highest M_G mean value of the dominant Gaussian component. We also calculated the colour average for each colour bin. From this process we obtained two output arrays, one for the colour and the other for the absolute G magnitude. Thereafter, we defined a polynomial to fit the MS band using these arrays. In Fig. 3.2 we present the resulting fit, in blue for the observed and in red for the solution aCMD. Notice the obvious shift between both data-sets.

In the next step we defined the $\Delta G = M_{G,pol} - M_{G,*}$, such that $M_{G,*}$ refers to the absolute G magnitude of each star in the aCMD and, $M_{G,pol}$ is the G magnitude interpolated in the polynomial fit for the colour of the star. In summary, we used the line representative of the general trend of the MS in the aCMD to subtract the absolute G magnitude of the fit to that of each star located at a vertical distance and finally get the overall distribution of ΔG for the whole colour interval.

From the calculations of the ΔG values, we computed the histogram and the corresponding GMM. In the GMMs, the idea is to fit individual Gaussians (black lines) whose sum gives the wrapping Gaussian (red line) of the studied distribution. We only fitted 5 Gaussian components in all of our tests so that they were comparable with each other and therefore, ease their visualisation.

Two examples of the final results are shown in Fig. 3.3, where the distribution of ΔG for the GCNS aCMD is presented on the left and that of the 100M star BaSTI solution on the right. Both GMMs are characterised by a small bump in the left side of the main component which correspond to the unresolved binary population. Besides, only the BaSTI solution shows a few stars towards positive values of ΔG, which we believe are associate to the presence of synthetic metal-poor stars lying under the MS (see Fig. 3.2). The fact that this feature is missing in the observed aCMD indicates an underlying subtle mismatch between

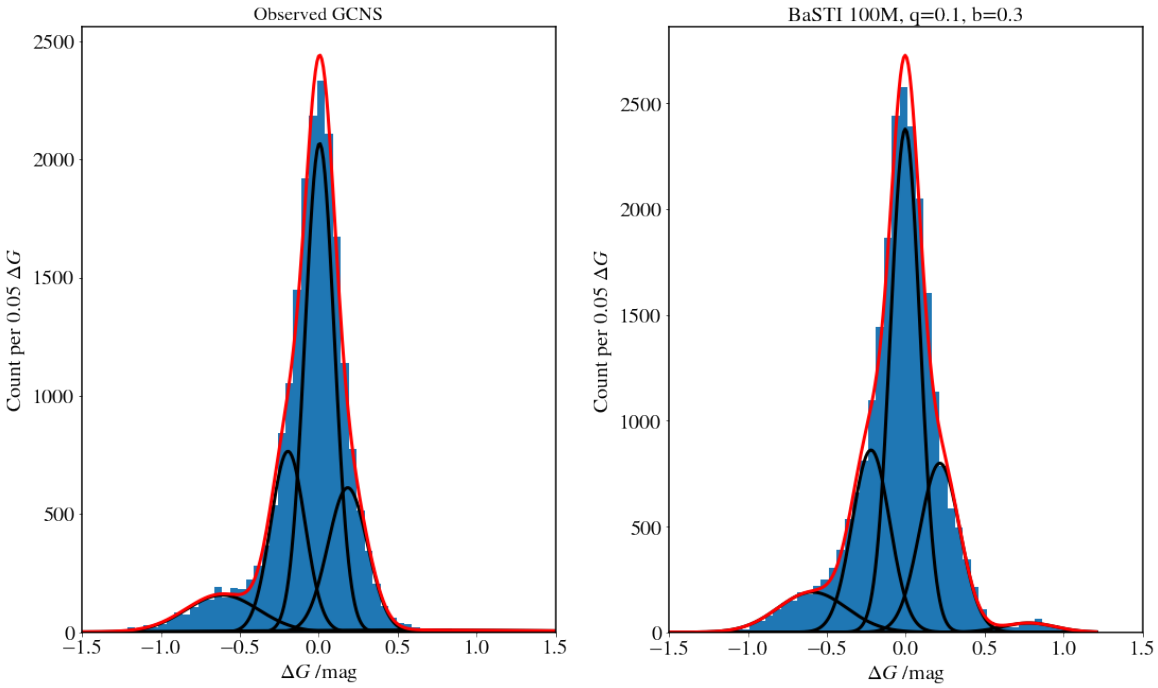

Figure 3.3: Gaussian Mixture Models of the distribution of star counts per 0.05 ΔG bin in the range $0.95 < G_{BP} - G_{RP} < 1.4$ after the slope of the main sequence has been subtracted in the aCMD. The GCNS data is shown in the left panel and in the right one, the BaSTI model solution considering 100M of stars in the synthetic mother diagram.

the empirical data and the theoretical models. In any case, this only affects a minority of stars, so it won't have a substantial effect on the conclusions.

4 Results and Discussion

Extensive testing using different synthetic aCMDs (stellar models, unresolved binary recipes, etc.) and various internal parameters of *dirSFH* (age-metallicity grids, weighting strategies, etc.) reveal that the solutions are robust. Hereby we present the results from our study by discussing the different sets of input parameters and the resulting SFHs.

We will present the results with a simplified visualisation using a subset of the panels in Fig. 3.1, showing only the evolution of the SFR over time (top) and the SFR in the age-metallicity plane (bottom), as these figures carry the most valuable information on the SFH.

There are common features in the derived SFHs, i.e. the SFH is quite bursty, the SFR in the age-metallicity plane reveals an old population of stars with very low metallicity and a densest area of stars around 3 Gyr at solar metallicities and, finally the metallicity distribution uncovers that most of the stars have solar metallicity, as expected.

For all the calculated solutions the faint magnitude limit of the bundle was set at M_G=4 mag, unless otherwise specified.

4.1 Testing parameters

4.1.1 Age-Metallicity binning size

The effects on the SFH of adopting different sizes of the bins of the age and the metallicity were tested by using the mother diagram "q01b03_100Mbasti_MG11" from Table 3.1.

First of all, we explored the solution by considering what we named "type 1" age and metallicity bins. In terms of the age, we considered the range from 0.03-13.5 Gyr, in particular (0.03, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1.2, 1.6, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10.0, 12.0, 13.5). Regarding metallicity, we used an interval within $0.0001 < Z < 0.039$ with increasing bin size towards higher Z values, i.e. (0.0001, 0.0007, 0.0016, 0.0029, 0.0046, 0.0067, 0.009, 0.016, 0.027, 0.039).

The second test used what we will call "type 2" bins of both parameters. On the one hand, the age was constrained between 0.08 and 13.5 Gyr with smaller bins: (0.08, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0, 13.0, 13.5). On the other hand, the metallicity range taken under consideration in this case counted on more intermediate values than in the former case, resulting also in smaller metallicity bins: (0.0001, 0.0004, 0.0007, 0.0011, 0.0016, 0.0022, 0.0029, 0.0037, 0.0046, 0.0056, 0.0067, 0.0079, 0.009, 0.012, 0.016, 0.021, 0.027, 0.039).

The third attempt was a combination of both cases above, a "mixed" approach in which we considered "type 2" (smaller) age bins and "type 1" (larger) metallicity bins.

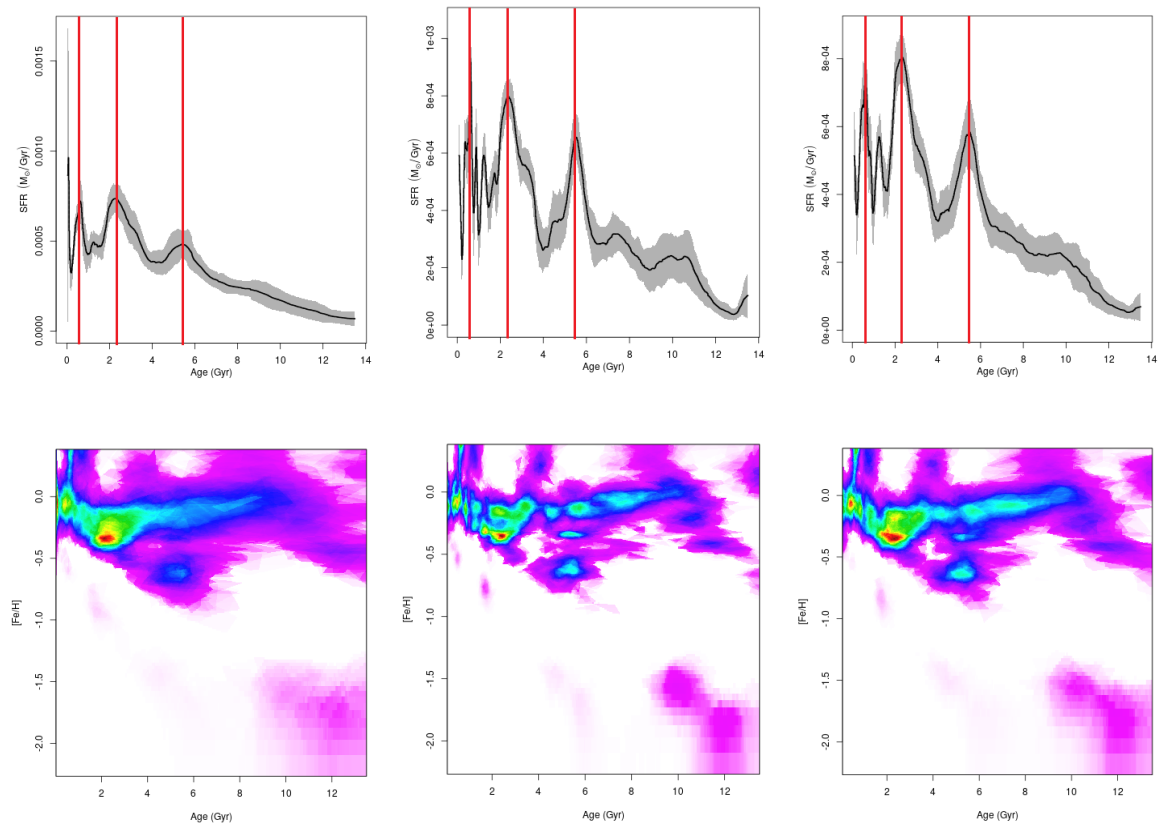

Figure 4.1: SFR over time (top) and SFR age-metallicity plane (bottom) of the three tests performed by changing the age and metallicity binning sizes. The positions of the star-forming bursts in the top panel are indicated with vertical red lines, taking as reference the location of the bursts in the solution shown in the rightmost panel. From left to right: "type 1" bins, "type 2" bins and the combination of "type 1" age bins and "type 2" metallicity bins (see the text for more details).

In Fig. 4.1 we show the "type 1" bins test, followed by the "type 2" bins test and the "mixed" test. The leftmost panel displays a rather blurred solution in which not many structures can be shaped in the SFR plane of age-[Fe/H]. In addition, given the width of the bins, the SFR over time shows very weak and wider peaks of star formation. In the second binning size test, the top panel

revealed differentiated star formation bursts and many structures were more clearly outlined in the SFR age-metallicity plane. Finally, the combination of both "type 1" age bins and "type 2" metallicity bins resulted in a figure fairly similar to the central.

In short, we saw that the structures that appear are the same in all the tests, however, the smaller bins reveal more defined structures. Whether the solution obtained in the second test, which is the sharpest, is the best, is yet unknown. Therefore, in order to save computational time and gain efficiency, in general we chose to move on using the third test 'mixed' parameterisation, combining narrower bins for age and larger bins for metallicity, thus ensuring that the essence of the results is preserved.

4.1.2 Weighting strategies

We performed attempts using two weighting strategies: the "weighted" and the "uniform". The synthetic mother diagram used was "q01b03_100Mbasti_MG11" and in this case we used the 'type 1' age-metallicity bins.

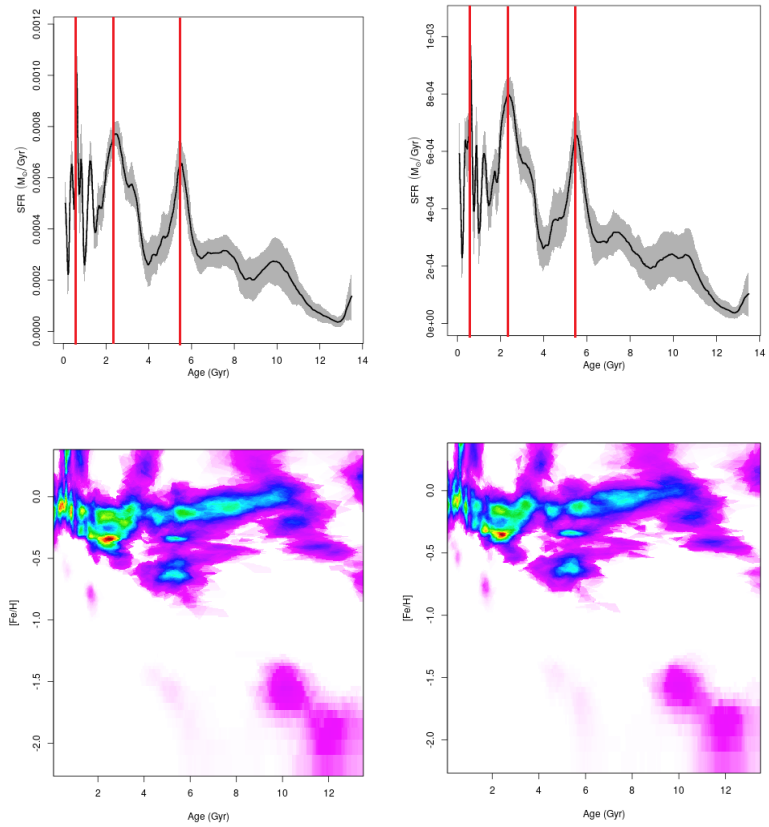

Figure 4.2: SFR over time (top) and SFR age-metallicity plane (bottom) of the two tests performed by changing the weighting strategies throughout the computing process. The positions of the star-forming bursts in the top panel are indicated with vertical red lines, with the same location in both left and right panels. In the left two panels we show the "weighted" method, while in the right ones, the "uniform" plan is considered.

The left two panels of Fig. 4.2 show the solution obtained considering more weight in the regions of the CMD with more uniform age distribution. In the right hand pair of plots, we tested the case of giving uniform weights across the aCMD. Both solutions are very similar, revealing almost identical features in the SFR age-metallicity plane and showing resembling SFRs over time, except for a minor difference in the widths of the peaks. The "uniform" test on the right presents slightly narrower peaks than the "weighted" case. Yet, the dissimilarities are almost negligible according to our present interest. Since running the "weighted" test took more execution time, we preferred to use the "uniform" approach in the future tests.

4.1.3 Stellar models

As it relates to the theoretical stellar evolution models, we used BaSTI and PARSEC, both solar scaled, with the synthetic mother diagrams "q01b03_100M basti_MG11" and "q01b03_100Mparsec_MG10", respectively.

Fig. 4.3 shows the BaSTI model test on the left and the PARSEC one the right. Again, vertical red lines have been drawn as reference coinciding with the three most prominent burst in the BaSTI solution. In this case, some differences in the SFHs are apparent. The burst at ~ 5.5 Gyr in the BaSTI solution may be identified with that happening at 4.5 Gyr in PARSEC. In the case of the bursts more recent than 4 Gyr, the ones that can be observed in the PARSEC solution are systematically older than the BaSTI ones, which is consistent with the fact of PARSEC presenting brighter isochrones (see Fig. 2.7). Furthermore, note that there is a tail of a burst at 13 Gyr and high metallicity in PARSEC while it is missing in BaSTI, which instead presents trace of a population old and metal poor. Overall, the PARSEC solution has larger error bars. These differences in PARSEC can also be observed to the age-metallicity plane, where a less reasonable structure is observed than in the BaSTI model, with narrow stripes at specific metalicities. In particular, PARSEC shows a flattened distribution of stars at close to solar metallicities.

The features in the age-metallicity plane in the case of the PARSEC solution possibly point to some, yet unidentified problem in the implementation of the models in the code used to compute the mother synthetic CMDs, and therefore we won't use these models further in this work.

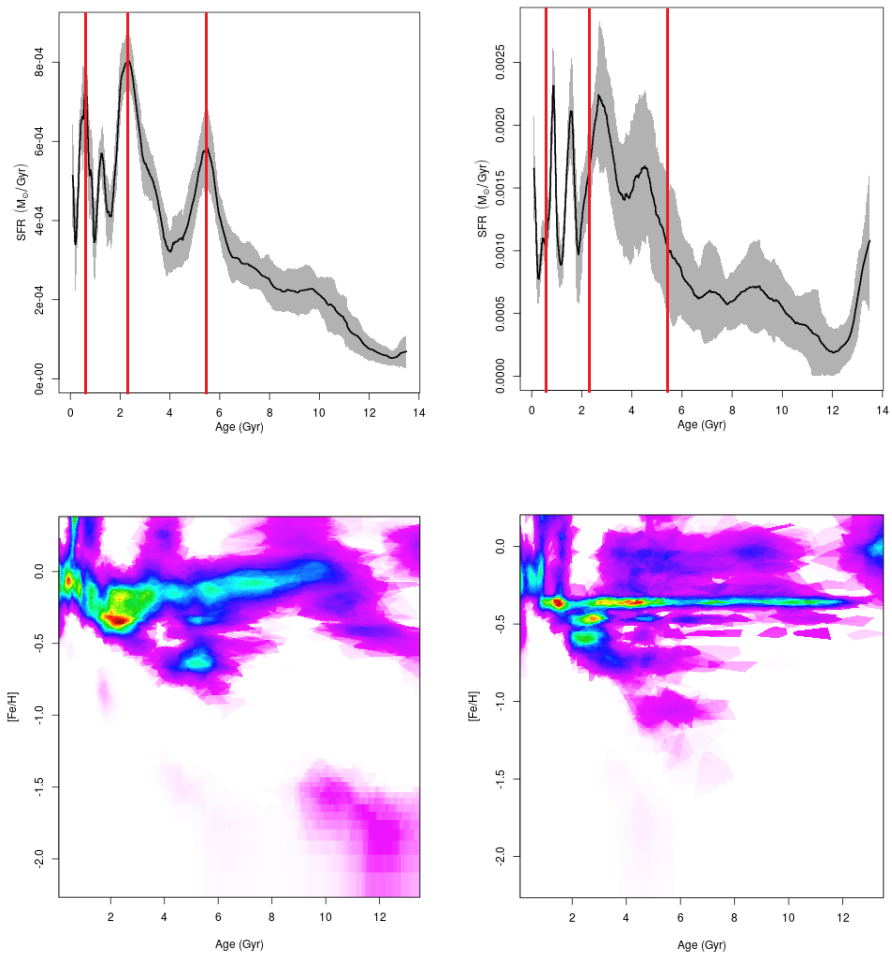

Figure 4.3: SFR over time (top) and SFR age-metallicity plane (bottom) of the two tests performed by changing the stellar evolution models. The positions of the star-forming bursts in the top panel, corresponding to the BaSTI solutions, are indicated with vertical red lines. The solution obtained with the BaSTI models is shown on the left, whereas the one obtained with the PARSEC library is presented in the right two panels. Note that the latter shows rather peculiar, unexpected structures in the lower panel, thus, not a satisfactory solution.

4.1.4 Limiting magnitude

Another test of special interest was the analysis of the resulting SFH using different limiting M_G when defining the bundle. We carried on this study by using our reference mother diagram, "q01b03_100Mbasti_MG11".

The solutions are shown in Fig. 4.4. From left to right, the panels correspond to increasing limiting M_G of 4, 4.5 and 5 mag. Let's concentrate first in the upper panels, where vertical lines indicate the position of the main star formation bursts in the solution depicted in the left panel. The solutions limiting the bundle at M_G of 4 and 4.5 (left and central panels), show star formation burst at the same age, but in the solution shown in the central panel, a burst occurring 10 Gyr ago, only hinted in the solution in the left panel, appears prominently. The SFR as a function of time shown in the right panel is very

different from the other two. The youngest bursts are very similar, but the one at $\sim$ 2.3 Gyr ago is wider than in the other two solutions and the one at 5.5 Gyr almost totally disappears. Two bumps $\sim$ 8 and 11 Gyr ago are visible as opposed to the other two solutions. In the lower panels displaying the SFR as a function of age and metallicity, the sharp features that can be observed in the solution of the left panel (like the strip at solar metallicity and ages between 4 and 10 Gyr ago, and the sharp $\sim$ 5.5 Gyr old population with [Fe/H]=-0.5) disappear progressively in the other two solutions, as the old and metal poor feature does.

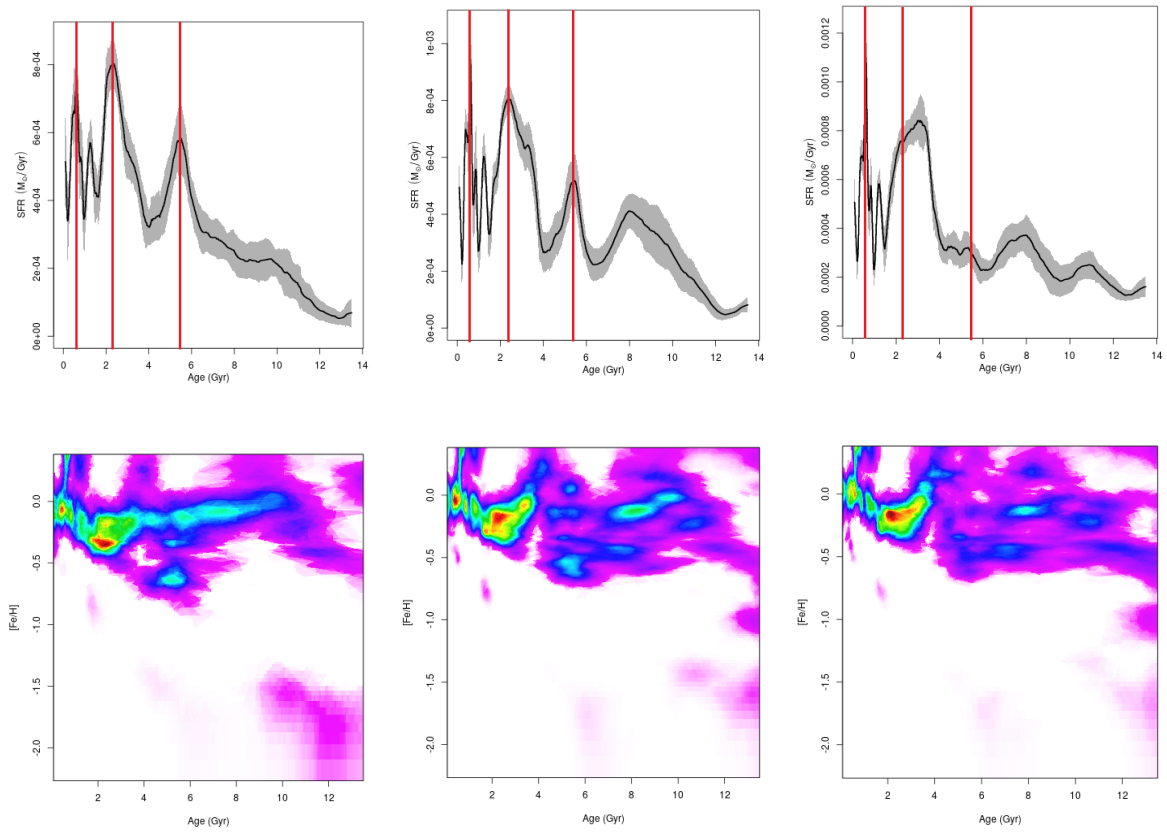

Figure 4.4: SFR over time (top) and SFR age-metallicity plane (bottom) of the three tests performed by changing the faint limiting magnitude of the bundle. The positions of the star-forming bursts in the top panel are indicated with vertical red lines, taking as reference the location of the bursts in the solution shown in the lefmost panel. From left to right, the threshold G magnitude that have been considered for each case are: $M_G = 4, 4.5$ and 5 mag.

We don't have a clear criteria to decide which of the magnitude limit cuts above lead to a more reliable SFH, as the true SFH of the solar neighbourhood is not known by any independent method. To try to understand the differences, we will examine whether we are actually providing useful age information when adding fainter stars to the fit. To this end, we plotted Fig. 4.5, which is a zoomed version of Fig. 2.7 indicating the limiting M_G considered cuts (solid black lines) at 4, 4.5 and 5 mag. As can be seen, the fainter the magnitude, the closer together the isochrones of different ages are to each other, thus,

the distinction of the stellar ages is more complex. This idea is consistent with the above-mentioned about the progressive increase in the width and poor definition of the star formation episodes. For this reason, we decided to place our absolute magnitude G cut-off at 4 mag because increasing the limit half or 1 extra magnitude did not seem to result in more age information.

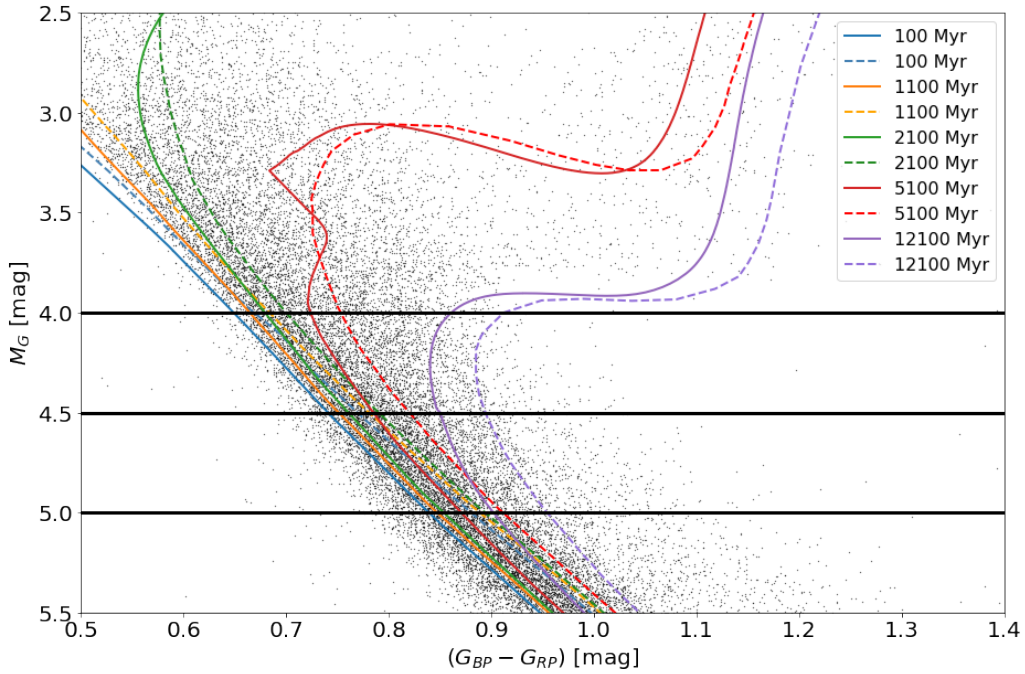

Figure 4.5: Zoomed region of Fig. 2.7 with the different limiting magnitudes considered in the bundles (solid black horizontal lines) at: $M_G = 4$, 4.5 and 5 mag. Note that at $M_G = 4$ mag, the isochrones are more separated from each other, making it easier to identify the age of the stars from this threshold cut towards brighter magnitudes.

4.1.5 Number of stars in the mother CMD

The following test was focused in the number of stars encompassed in the synthetic mother CMD used in the calculation to determine the SFH. The solutions were all computed by using sub-samples of the "q01b03_100Mbasti_MG11" mother diagram. In each case we selected the amount of stars of interest in the mother diagram with a limiting M_G at 11 mag. The solutions are presented in Fig. 4.6.

In first place we included 1M of stars in the mother diagram (corresponding to $\sim$ 70000 at magnitudes brighter than $M_G \leq 5$ mag, useful for SFH determination)[1]. The result was as expected (see top left panels). With so few stars in each SSP, starbursts are hardly differentiated in the SFR over time panel and also the solution has bigger error bars. The SFR age-metallicity plane on the other hand shows a diffuse stellar density distribution.

[1]For comparison, the aCMD of the GCNS has about 20000 stars above this magnitude limit.

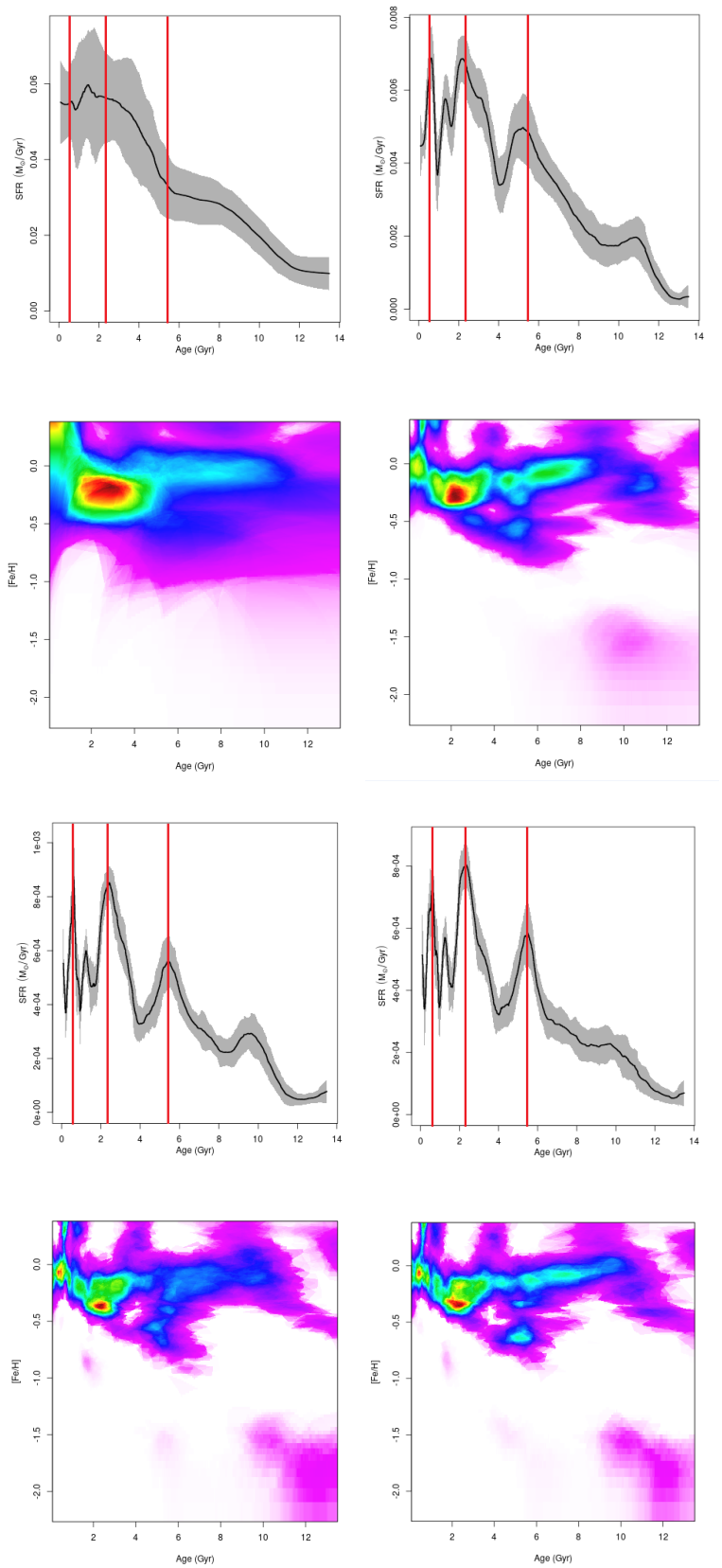

Figure 4.6: SFR over time (top) and SFR age-metallicity plane (bottom) of the tests performed by changing the number of stars of the mother diagrams. The positions of the star-forming bursts in the top panel are indicated with vertical red lines, taking as reference the location of the bursts in the solution shown in the bottom right panel. From left to right and top to bottom, the panels show tests with increasing number of stars: 1M, 10M, 60M and 100M.

Bearing the above in mind, we increased the number of stars to 10M (corresponding to $\sim$ 700000 stars with $M_G \leq 5$ mag, top right panels). The bursts of star formation could be spotted in their expected positions overall but both the SFR as a function of time and in the age-metallicity plane are less sharp than in the SFH calculated with a larger synthetic aCMD and shown in the two solutions in the bottom panels.

The bottom left panels show the 60M solution ($\sim$ 4000000 stars brighter than $M_G = 5$ mag). This test is very similar to the previously considered of 100M stars (see bottom right panels). Either the rate of star forming bursts in time and the discerned components on the SFR age-metallicity distribution are almost equally well-resolved in these two tests.

In summary, all the panels show the same features but, as proven, the more stars are used in the calculation, the sharper the solution provided by the code. We were therefore able to set a limit in the quantity of stars in the mother aCMD from which our solutions would be acceptable and thus, reduce the computation time by using fewer stars. We adopted 30M stars in the following tests designed to study the characteristics of the unresolved binary star population.

4.2 Unresolved binaries in the GCNS

4.2.1 SFH solutions

For the binary population tests we chose the mother aCMDs of 30M stars with different q and β listed in Table 3.1 and the 60M star subset in the case of 'q01b03', since considering this amount of stars led to equally good results as the 100M set.

Binaries with a mass ratio between q=0.1-0.4 have a small secondary companion and thus the light coming from the secondaries scarcely affects the combined light of the unresolved binary. Thus, they are hardly discernible from the single stars MS. On the other hand, rising the binary mass ratio between 0.5 and 0.7, the brightness of the secondary is more significant. The latter are the binaries that constitute the band parallel to the MS, which is well differentiated in Fig. 3.2. We therefore considered different secondary mass ratios meeting these situations combined with a variation of the binary fraction within the samples.

In Fig. 4.7, we show the tests performed using different binary scenarios. These mother diagrams have been used in the different solutions, from top to

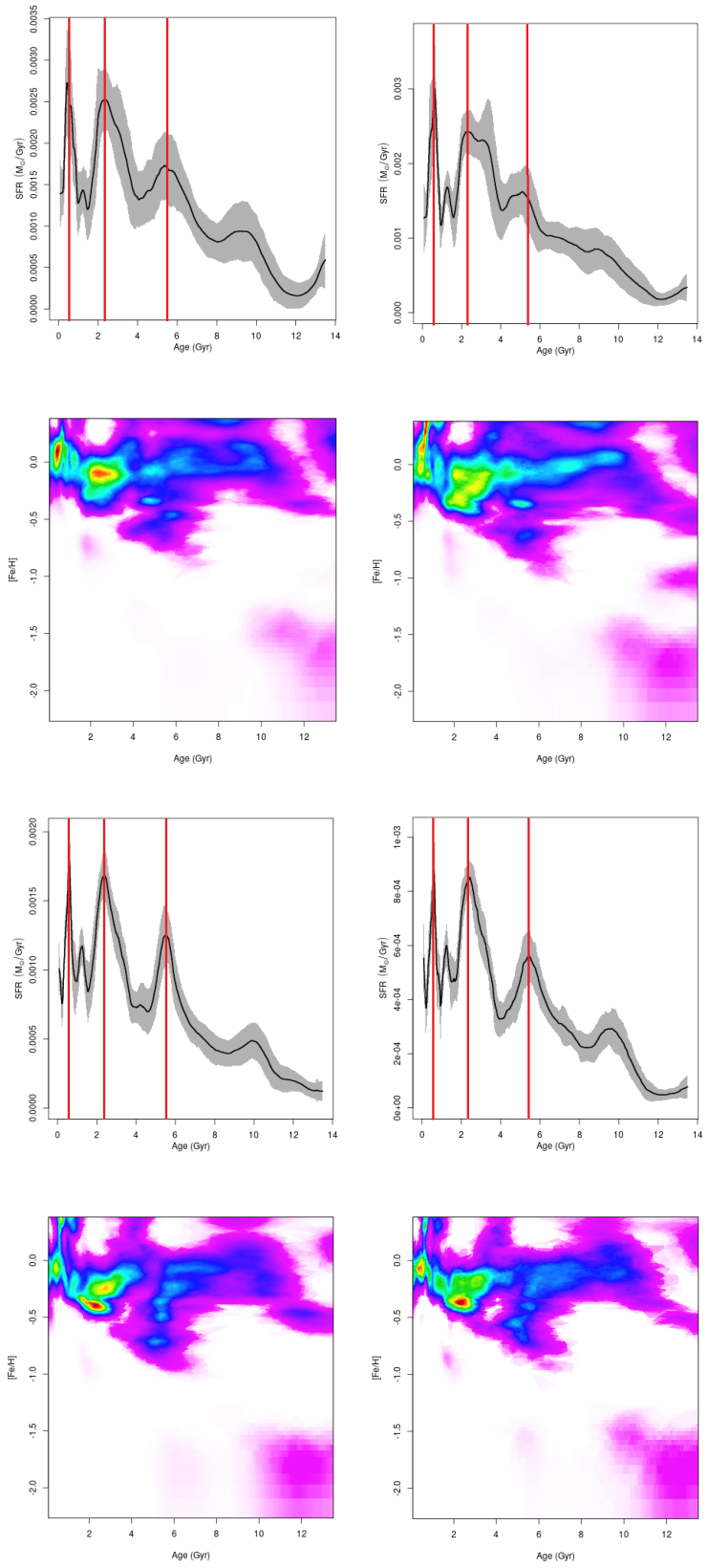

Figure 4.7: SFR over time (top) and SFR age-metallicity plane (bottom) of the four tests performed by changing the binary mass ratio and the binary fraction. The positions of the star-forming bursts in the top panel are indicated with vertical red lines, taking as a reference the location of the bursts shown in the bottom right panel. In the first three we considered mother diagrams of 30M stars, whereas the last one is a 60M star subset from our main 100M mother CMD. From left to right and from top to bottom: q=0.1 and β=0.7, q=0.4 and β=0.3, q=0.6 and β=0.1 and lastly, q=0.1 and β=0.3.

bottom and left to right: "q01b07_30Mbasti_MG11", "q04b03_30Mbasti_MG11", "q06b01_30Mbasti_MG10" and "q01b03_60Mbasti_MG11", respectively.

The four top panels show the epochs of star formation occurring at the same time. However, in the top panels 'q01b07' and 'q04b03', the bursts at 2.3 Gyr and 5.5 Gyr are wider than the bottom cases. Moreover, note that the SFR over time are noisier than in the other two cases. Regarding the SFR in the age-metallicity plane, all the tests show quite resembling features, e.g. the young (3 Gyr old) high stellar density at approximately $[Fe/H] = -0.4$, with slight stellar distribution differences towards older ages. Nevertheless, the bottom two panels are almost identical.

The latter two tests ('q06b01' and 'q01b03') have apparently very different parameterisations of binaries but their effects are practically counterbalanced. This is because in the test with q=0.6 and β=0.1, the few binaries included in the model have all bright secondaries, while in the other case with q=0.1 and β=0.3, many more binary systems have been included but at least 50% of them have a very small mass ratio for the secondaries, and thus these binaries are practically indistinguishable from a single star. Such balance between both tests explains the resemblance between the two figures. However, we could see that the solution on the right, considering more binaries of equal mass, is somewhat better (see below).

Despite the fact that some binary recipes fit the observational data better than others, the SFH main features remain mostly unchanged. Based on these results, in order to make a more in-depth comparison between the different binary parameterisations, we proceeded to calculate their ΔG distributions and compare them with the observed one.

4.2.2 GMM analysis of the ΔG distributions

Applying all the procedures to build the GMMs described in Section 3.2, we analyse the considered binary parameterisations. We adopted a unifying criteria to represent all the GMMs with 5 Gaussian components for ease of comparison.

In Fig. 4.8 we plotted the ΔG histograms and the wrapped Gaussians from the GMM fit for the 4 SFH solutions discussed in the previous section, plus the 100M star test using the PARSEC library. All of them show the binary bump in the left of the main distribution and almost all present a small bump at positive ΔG values, which are theoretical populations of stars lying between the MS and the WDs.

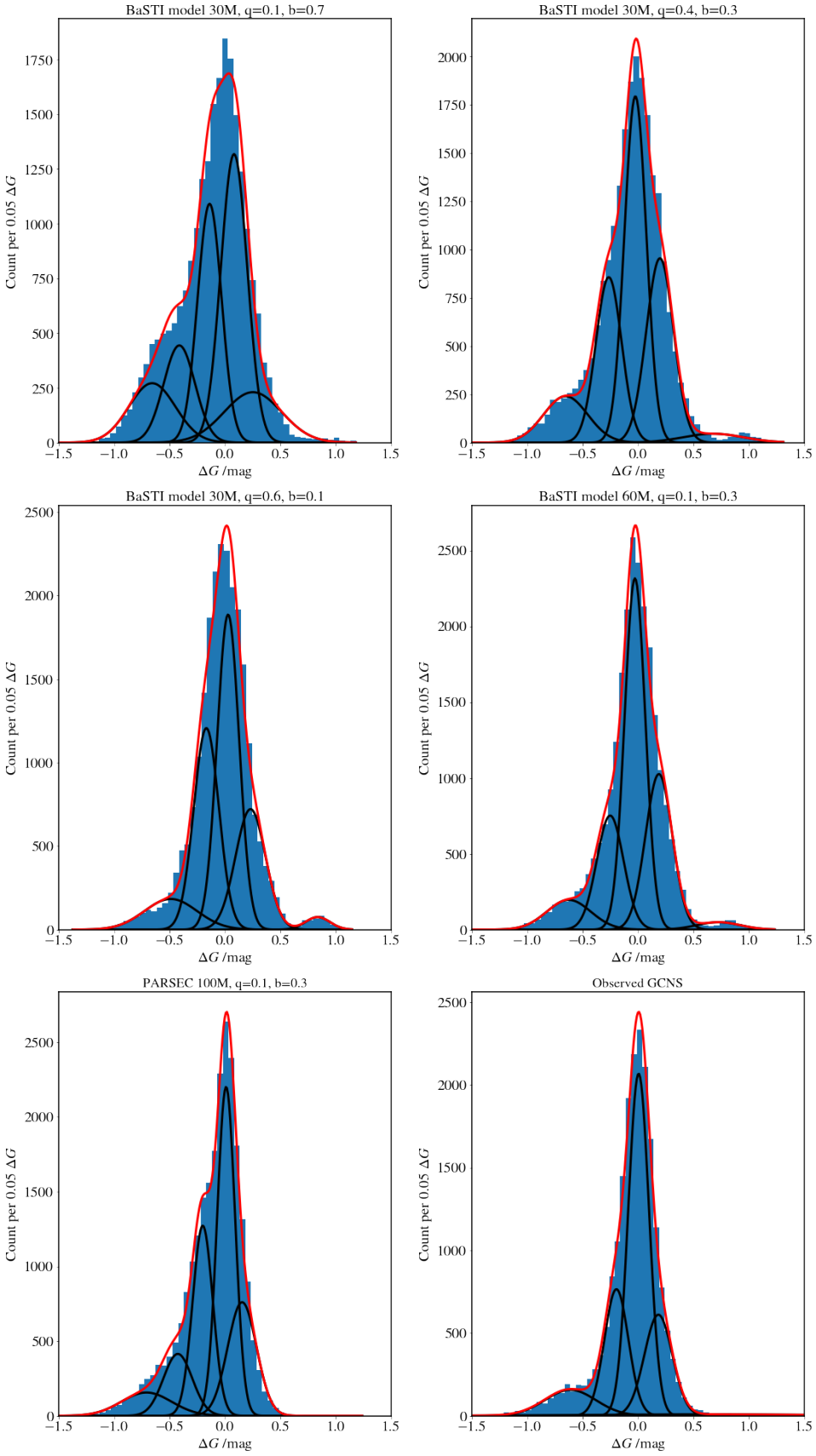

Figure 4.8: Histograms and corresponding GMMs of the distribution of star counts per 0.05 ΔG bin in the range $0.95 < G_{BP} - G_{RP} < 1.4$ after the slope of the main sequence has been subtracted in the aCMD. The mother aCMDs used in the computation are labelled. The bottom right panel corresponds to the observed distribution.

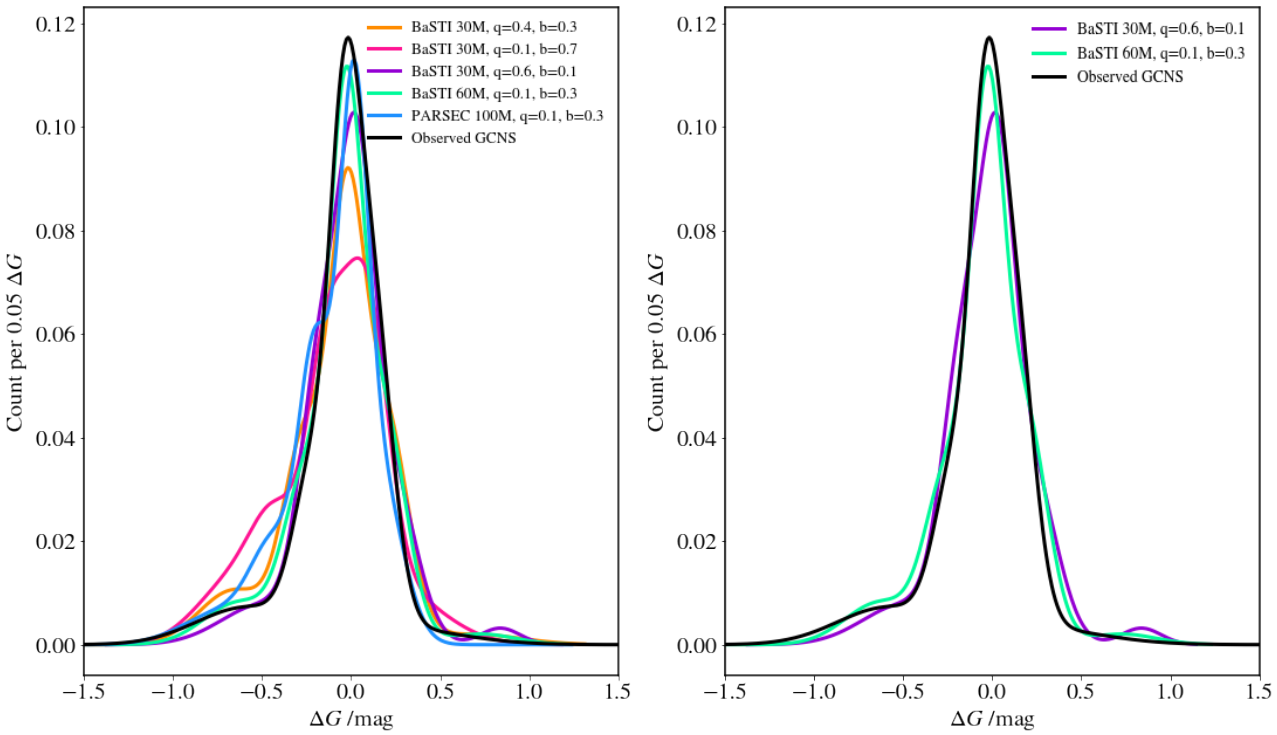

Figure 4.9: Left: Normalised composition of wrapped Gaussians of the considered distributions, namely the GCNS, the PARSEC (100M) and the BaSTI (60M, 30M), all labeled according to their binary parameterisations. Right: id., only showing the 2 distributions that best fit the observational data.

For an easier comparison with the observed distribution, Fig. 4.9 depicts the normalised composition of wrapped Gaussians of the formerly studied distributions (in different colours) and that of the GCNS (in black). The left panel shows all of them, whereas the selected best fits are both presented in the right panel together with the data-set of the GCNS.

Based on the previous figures (Fig. 4.9 and Fig. 4.8), we noted that the distribution that differs most from the rest is that of the 30M model with q=0.1, β=0.7. In this case we considered the largest fraction of binaries among our tests combined with the smallest mass ratio. As a result of this, the binary bump is more noticeable, comprising a higher density of binary systems than in the other distributions. Moreover, the main peak of this distribution is wider and significantly diverging from our data.

In the second panel of Fig. 4.8, corresponding to the models with 30M stars and q=0.3, β=0.4, we also found a slightly larger binary bump than the GCNS reference. Finally, the PARSEC model, which does not present the bump in the right side of the distribution, reveals a second Gaussian component, centered at $\Delta G = -0.3$, that stands out from the rest, leaving this model also rejected as the best fit to our data.

Finally, the two tests that best reproduce the observational data are the 60M

with q=0.1, β=0.3, and the 30M with q=0.6, β=0.1 from BaSTI, as shown in Fig. 4.9. The latter shows a small bump of theoretical stars in the right hand side of the main Gaussian, whereas in the former distribution, this very same bump is more extended and less bulky.

These conclusions can be visualized in an easier way in Fig. 4.9. The binary bump differs quite a lot from the observed one for the PARSEC distribution (blue) and for the two BaSTI solutions of 30M stars with q=0.1, 0.4 and β=0.7, 0.3 (pink and orange, respectively), thus lowering their main components in relation to the GCNS distribution. Note the good agreement between the distributions corresponding to the solutions with q=0.1, 0.6 and β=0.3, 0.1 (violet and green, respectively) as shown in the right panel.

Other analyses considering a data volume of 200 pc radius around the Sun in Dal Tio et al. (2021), show that, according to their fit, the fraction of binaries is well represented in the lower MS by values of 0.4, while values twice as large fit best the upper part. In addition to that, they only simulate binaries with a mass ratio above a given threshold, which they set at 0.7, since the large majority of binaries with small mass ratios have secondaries too faint to compete with the light of the primary. In contrast to our selected unresolved binary parameterisation, in their best fit they require a much higher binary fraction and mass ratio than we do.

4.3 The SFH of the GCNS

We obtained the detailed SFH of the GCNS. In the previous sections we found that most of the parameters involved in the SFH calculation that we tested did not affect the main features of the SFH, even though some preferences could be set according to the results discussed. Based on the discussion in the previous two sections, we adopted the solution implementing the following summarised elements (see Section 4.1 for details):

- Age-metallicity bins size: age bins of "type 2" and "type 1" metallicity bins
- Weighting strategy: "uniform"
- Stellar model: BaSTI
- Limiting magnitude: bundle set at $M_G = 4$ mag
- Number of stars: 100M stars
- Binary recipe: q=0.1, β=0.3

In Fig. 4.10, we show three panels of the output SFH solution. On the left the SFR over time, the MDF in the central panel and the SFR in the age-metallicity plane on the right. It is mainly the SFH of thin disc stars since most of the stars within the sample belong to this component.

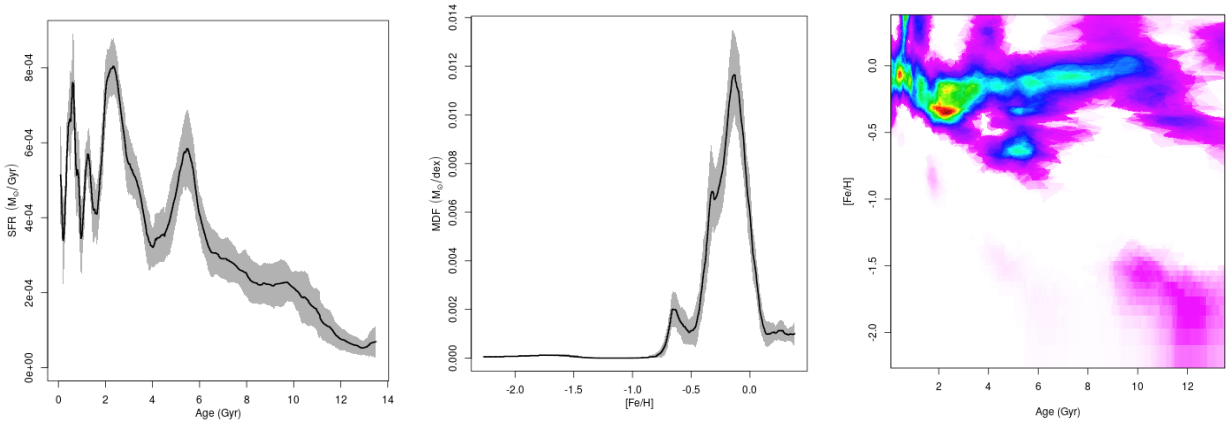

Figure 4.10: SFH of the GCNS using the mother diagram with 100M stars of BaSTI model. The resulting panels from left to right are: SFR over time, MDF and SFR in the age-metallicity plane.

The most robust result is that of the SFR over time. As can be seen, few stars formed before 12 Gyrs but from that age on, the average of newborn stars rised, turning into a highly bursty star formation. There is a hint of star formation episodes around 10 and 8 Gyrs but they are both within the error bar, so we cannot confirm the presence of bursts within this epoch. However, at 5.5 Gyr there is a striking star formation burst followed by a more quiescent period of time until 2.5 Gyr, where another significant burst happens and, finally, from 2 Gyr to the present time there are more frequent small bursts over time. Despite this, the local environment has had periods of much more star formation than it does currently.

The Metallicity Distribution Function shows the amount of solar masses that turned into stars for each metallicity. In this panel we observe that there are few stars with metallicity lower than [Fe/H]=-1 and there is a small peak around [Fe/H]=-0.6. Besides, the highest peak is found around solar metallicity, as can be expected, since our sample is constrained in the solar vicinity. In addition, we saw a small spike at [Fe/H]=-0.4 which we do not know if it is significant. Finally, the metallicity distribution decreases sharply towards higher metallicities.

In the last panel, the age-metallicity plane reveals highly resolved structures. Note that there is a low density of metal poor ([Fe/H] $<$ -1.5) old stars which are possibly the halo stars, which we have already identified in our data. Also, we identified overdensities of stars at [Fe/H]=-0.6 of 4.5 Gyr old and at solar

metallicities and around the age of 3 Gyr.

Previous studies of the SFH using larger volumes of data, such as a 2 kpc bubble around the Sun (Ruiz-Lara et al., 2020), identified that star formation bursts occurred at the age of 1, 1.9 and 5.7 Gyrs, which might be related to pericentric passages of the Sagittarius dwarf galaxy. Additionally, Gallart et al. (2019) postulated that the star formation episode at 10 Gyr could be related to the accretion of Gaia-Enceladus. Thus, our derived SFH is in good agreement with these findings, which strengthens the value and consistency to our study.

In the study carried out in a similar Milky Way volume by Dal Tio et al. (2021), they derived the SFR using a logarithmic age spacing, thus it is difficult to set a reliable comparison with our solution. However, we highlight that they obtain higher metallicities than the solar, while our metallicity limit is lower on average. Also, they obtain the most prominent star formation peak around 3 Gyr, which is in agreement with our results.

5 Summary, conclusions and future work

We have provided the precise SFH of the GCNS sample, a complete nearby star catalogue of well-characterised objects within 100pc of the Sun composed with data from the *Gaia* EDR3.

Our goal was to calculate the SFH using the new *dirSFH* code, which relies on the CMD fitting technique. The method is based on the comparison of stellar distributions of observed CMDs with simulated stellar populations computed from the predictions of stellar evolution models. We tested the effects on the solution of some input parameters from the code (i.e. the binning size of the age and metallicity, the weighting strategy, the limiting absolute G magnitude), and from the synthetised aCMDs (as the number of stars computed, the stellar library and the binary stars parameterisation).

We noticed that none of the changes applied to the input parameters significantly affected the conclusions on the main features in the local neighbourhood SFH, despite slight variations in the width and height of the peaks of the star formation bursts, the small off-set of the star formation episodes and the characteristics of the more or less diffuse structures of the SFR in the age-metallicity plane. We therefore conclude that our solutions are robust to reasonable changes in the parameters of the fit.

More specifically about the unresolved binaries, we tested various binary recipes and compared the distribution of stars in the MS below the old MS turnoff of the observed aCMD with that of the different models, using GMMs, as they did in Gaia Collaboration (2021). We found that the best fit of unresolved binaries to the observational data was that involving a binary fraction of 0.3 and a secondary lower mass ratio of 0.1, although other binary scenarios did not drastically affect the result. Thus, the 60M solution with 'q01b03' was slightly better than the 30M with 'q06b01' (see Fig. 4.9).

The computed SFH, shown in Fig. 4.10, is that of disc stars because most of the stars in the data-set reside in this part of the Galaxy, particularly in the thin disc. We identified repeated main features in the SFH derived with all of our tests such as the trace of old poor-metal stars probably belonging to the galactic halo, the high density of solar metallicity stars at about 3 Gyr and the

same epochs of star formation by highlighting the bursts at approximately 0.5, 2.3, 5.5 and possibly at 10 Gyr. These caracteristics are robust and are in good agreement with other publications (Ruiz-Lara et al. (2020) and Gallart et al. (2019)).

Additional findings have been the corrections of the bolometric corrections applied to the BaSTI models in the lower MS as a result of comparing with isochrones to determine how well the BaSTI and PARSEC models fitted the GCNS data. Changes are continuously being implemented to the stellar models according to tests like ours, which keep improving their quality. Another analysis solely ours was the effect of interstellar extinction to the data. The reddening corrections turned out to be irrelevant, thus we concluded to safely neglect its contributions.

Given the novelty of the *dirSFH* programme and and its continuous updates, as well as the recent publication of the *Gaia* DR3, there is a lot of work ahead that could enhance and sharpen all these results. In turn, we hereby suggest some of the further work, beyond the scope of this master project:

- Check if the IMF is in agreement with the data, by using the MS luminosity function of the observed and the solution CMDs. If it were not in good agreement, different values of this parameter could be tested in the synthetic CMD models.
- Refine the comparisons of the binary parameters by performing a more quantitative goodness-of-fit analysis, perhaps using the Kolmogorov-Smirnov test.
- Recalculate the ΔG used in the GMMs, this time considering the perpendicular fiducial distance between each star and the MS polynomial fit.
- Derive the SFH in other larger volumes of data.

Acknowledgements

This work has made use of data from the European Space Agency (ESA) mission Gaia (`https://www.cosmos.esa.int/gaia`), processed by the Gaia Data Processing and Analysis Consortium (DPAC,`https://www.cosmos.esa.int/web/gaia/dpac/consortium`).

I would like to express my deep and sincere gratitude towards my advisor Dr. Carme Gallart for her continued and patient guidance throughout the supervision of this project. Most of all, her persistent and enthusiastic encouragement was fundamental to fulfill and make this work possible. She shared her catching passion and knowledge with me and I cannot be more thankful.

Also, I would like to extent my appreciations to Dr. Tomás Ruiz-Lara for his exceptional help and availability throughout every step in the project, symplifying everything despite working from the distance. He gave us tips and loads of support when we needed it most. He had always a kind word and a huge smile whenever we needed a hand, which I appreciated so much.

Moreover, I acknowledge Dr. Francisco Surot for providing us the outstanding and groundbreaking code, which was the key to our success in this project and will be in the next to come.

Dr. Santi Cassissi, who nicely provided the correction for the BaSTI library, and Dr. Matteo Monelli offered their opinions and advice on numerous occasions, which we gladly listened to and followed. I am also grateful for the participation of Alicia Rivero, we helped each other with our respective projects numerous times and felt somehow accompanied. Dr. Emma Fernández also had some helpful contributions to the work and I owe her a great deal of gratitude for her kindness throughout the process of carrying out the work.

Finally, I want to thank my friends from all around (Andorra, Barcelona, Salou and Tenerife) for their persistent encouragement an also I would like to grant a special thanks to my whole family, particularly to my mum and my brother, for their unconditional support and motivation. And, above all, I am indebted to Pau, for both his endless love and attentiveness. I couldn't have made it this far without you!

Bibliography

Aparicio A., Gallart C., 2004, The Astronomical Journal, 128, 1465

Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, , 427, 127

Dal Tio P., et al., 2021, Monthly Notices of the Royal Astronomical Society, 506

Gaia Collaboration 2018, A&A, 616, A10

Gaia Collaboration 2021, A&A, 649, A6

Gallart C., Zoccali M., Aparicio A., 2005, , 43, 387

Gallart C., Bernard E., Brook C., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, Nature Astronomy, 3, 1

Gallart C., et al., 2021, apj, 909, 192

Green P. J., Sibson R., 1978, Comput. J., 21, 168

Green G. M., Schlafly E., Zucker C., Speagle J. S., Finkbeiner D., 2019, The Astrophysical Journal, 887, 93

Hidalgo S. L., Aparicio A., Martínez-Delgado D., Gallart C., 2009, The Astrophysical Journal, 705, 704

Hidalgo S. L., et al., 2018, apj, 856, 125

Kroupa P., 2001, , 322, 231

Mann A. W., Feiden G. A., Gaidos E., Boyajian T., von Braun K., 2015, The Astrophysical Journal, 804, 64

Penoyre Z., Belokurov V., Evans N., 2022, Astrometric identification of nearby binary stars II – Astrometric binaries in the Gaia Catalogue of Nearby Stars

Riello, M. et al., 2021, A&A, 649, A3

Ruiz-Lara T., Gallart C., Bernard E. J., Cassisi S., 2020, The recurrent impact of the Sagittarius dwarf on the Milky Way star formation history, doi:10.48550/ARXIV.2003.12577, `https://arxiv.org/abs/2003.12577`

Rybizki J., et al., 2020, Publications of the Astronomical Society of the Pacific, 132, 074501

Santos J., Jordi C., 2020, University of Barcelona, Faculty of Physics database