The star formation history of Milky Way components through *Gaia* colour-magnitude diagram fitting

Guillem Aznar Menargues

Supervisors: Emma Fernández-Alvar
& Carme Gallart

Master's final project

Master in Astrophysics

Universidad de La Laguna

July 2023

Abstract

The formation of galaxies is still a great unknown in the world of astrophysics. Although several advances have been made in recent years, it is still not fully understood. To increase our understanding of galaxy formation, we can study the formation and evolution of our own galaxy, the Milky Way, because no other galaxy can offer us greater detail of their properties. This is in part thanks to the precise data from the *Gaia* mission.

This project provides for the first time the detailed star formation history of the Milky Way, derived in several regions that cover a range of galactocentric radius and distances from the plane contained in a 2 kpc sphere around the Sun, using the *Gaia* EDR3 catalogue and assuming symmetry respect the Galactic plane. This analysis allows us to explore the stellar populations formed in the Milky Way across time and its correlation with metallicity, and how they distribute in radial and vertical directions.

The methodology employed is based on the colour-magnitude diagram fitting technique, using a specific code called `dirSFH`. We have to take into account the effects of interstellar extinction and interstellar reddening. We have evaluated two extinction maps to compensate for these effects, choosing the Lallement et al. (2022) map for the correction. Three layers of 100 pc of height above and below the galactic plane have been selected for study, two of them dominated by the thin disc and the third by the thick disc. In each layer, rings of 500 pc width are selected, covering a galactocentric radius between 7.5 and 9 kpc for the layer closest to the Galactic plane and between 7 and 9 kpc for the other two. A brief study is also carried out to settle the maximum extinction cut-off, deciding to leave the upper value of 0.5 magnitudes.

In the two layers closest to the Galactic plane, a decreasing metallicity trend with time is observed for old populations, while young populations increase their metallicity with time. In the uppermost layer these trends are not as clear, with the majority of the population being old and splitting in a double metallicity sequence. It can be seen in all layers that the young regions increase as the regions get further radially away from the Galactic centre, as well as old-intermediate metal-rich stellar populations which can not be explained through chemical evolution models appear. Studying the cumulative function of the star formation rate shows a decrease of the age as the radius increases and as the height from the plane decreases. An analysis of the metallicity distribution function reveals no clear gradients of metallicity with radius, although a slightly decrease of metallicity with height above the Galactic plane is observed.

Our results show a great coherence with the chemical trends observed in previous works from the literature. The dominance of young stellar populations with radial increase is consistent with the decrease of the abundance of high-[α/Fe] stars, indicative as well of an inside-out scenario. The abundance of [α/Fe] can also be related to the dominance of young populations as the height above the Galactic plane decreases. As for metallicity, it can be seen from observations that as the height increases the stars become less metallic. The obtained old-intermediate metal-rich stellar populations can be explained by radial migration of metal-rich stars formed in the inner regions of the Galaxy to the solar vecinity. The results of this work open the door to future research on the SFH of the Galaxy and to explore the impact of various formation processes, such as the radial migration, in the configuration of the its current morphology.

Resumen

La formación de las galaxias es todavía una gran incógnita en el mundo de la astrofísica. Aunque se han hecho diversos avances en los últimos años, esta sigue sin conocerse en su totalidad. Para aumentar nuestro conocimiento sobre la formación galáctica, podemos estudiar la formación y evolución de nuestra propia galaxia, la Vía Láctea, la cual nos proporciona un detalle de sus poblaciones estelares como ninguna otra. La disciplina que se encarga de este estudio es conocida como arqueología galáctica, que examina los registros fósiles de procesos de formación y evolución de la Galaxia a lo largo del tiempo mediante observaciones detalladas de las estrellas, el gas y otros componentes. Una de sus principales herramientas es el estudio de la historia de formación estelar (SFH), que puede ser determinada a partir de la comparación entre diagramas color-magnitud (CMD) observacionales con CMDs sintéticos. Este estudio era muy limitado en nuestra galaxia debido a la dificultad de determinar la distancia de las estrellas de forma precisa, hasta que la misión *Gaia* proporcionó datos astrométricos y fotométricos para aproximadamente mil millones de estrellas con una gran precisión nunca antes vista.

En el presente trabajo se proporciona por primera vez la SFH detallada de diversas regiones de la Vía Láctea, suponiendo simetría respecto al plano galáctico. Los datos utilizados se corresponden al catálogo EDR3 de *Gaia*, acotándolos a una esfera de 2 kpc alrededor del Sol.

La metodología empleada se basa en el ajuste de CMDs, es decir, en la búsqueda de combinaciones lineales de poblaciones estelares simples sintéticas para reproducir los CMDs observados. Nuestra metodología se divide en tres partes. En primer lugar, en la creación del CMD sintético. Seguidamente, se le aplican a este errores observacionales, es decir, incompletitud y errores fotométricos, mediante el código `New Dispar`. Por último, el programa `dirSFH` se encarga de comparar ambos diagramas y calcular la SFH.

Al ser nuestra esfera de 2 kpc, la extinción y el enrojecimiento interestelar se deben tener en cuenta a la hora de pasar a magnitudes absolutas. Por lo tanto, se deben obtener estos valores para corregir las magnitudes absolutas. Para ello, se ha hecho uso de los mapas de extinción, una especie de representación cartográfica que proporciona información sobre el grado de la extinción en una dirección y distancia específicas. Se comparan los dos mapas de extinción más actualizados hasta la fecha, Green et al. (2019) y Lallement et al. (2022). Este último es el elegido para realizar nuestras correcciones, ya que cubre todo el cielo y los resultados de distribuciones de metalicidad obtenidos utilizando estas correcciones se parecen más a las restricciones observacionales.

Una vez se corrigen la extinción y el enrojecimiento, y se tienen las magnitudes absolutas, se realiza un estudio para seleccionar un número de regiones, representativo de las distintas zonas de nuestra galaxia, a ser estudiadas. Se aplican diversos criterios para así obtener una muestra de estrellas de alta calidad. También se calcula a partir de las coordenadas galácticas, las coordenadas cartesianas galactocéntricas, que incluyen la altura sobre el plano galáctico y la proyección radial sobre este o radio galactocéntrico. Estudiando el número de estrellas en intervalos de altura sobre el plano galáctico de 100 pc y su completitud, se seleccionan tres capas a estudiar, en dos de las cuales predomina el disco fino y en una tercera el disco grueso. En cada capa se seleccionan anillos de 500 pc de amplitud radial, cubriendo los radios galactocéntricos de 7.5 a 9 kpc para la capa más cercana al plano galáctico, y de 7 a 9.5 kpc para las otras dos. Se realiza también un pequeño estudio para elegir el corte en extinción, es decir, el valor máximo de

extinción que tiene que tener una estrella para poder ser estudiada, decidiendo un límite superior de 0.5 magnitudes. Se muestra, por último, los CMDs a ser simulados y una tabla recopilatoria del número de estrellas y la completitud.

Los resultados se exponen mediante diversas representaciones. En primer lugar, se muestra el plano edad-metalicidad de cada región, donde se pueden observar las distintas poblaciones estelares que se han calculado en la simulación. La mayoría de poblaciones resultantes aparecen en todas las regiones estudiadas. En las dos capas más cercanas al plano galáctico, se observa una tendencia decreciente de metalicidad con el tiempo para poblaciones viejas, mientras que las jóvenes aumentan su metalicidad con el tiempo. En la capa más alta estas tendencias no son tan claras, siendo la mayor parte de la población vieja y apareciendo esta duplicada a una metalicidad más baja. Se puede observar en todas las capas como las regiones jóvenes aumentan conforme estas se encuentran más lejos radialmente del centro galáctico. También aparecen en todas las regiones poblaciones estelares viejas y de mediana edad con alta metalicidad que no pueden ser explicadas mediante los modelos de evolución química.

En segundo lugar, se estudia el ritmo de formación estelar. Para ello se dibuja la fracción de masa acumulada de cada región mediante una función cumulativa. En otro gráfico, se selecciona la edad en la que la región ha acumulado el 50 % y 90 % de masa. Se observa un aumento de estrellas más viejas con radios próximos al centro galáctico, lo cual sugiere que la Vía Láctea se formó siguiendo un escenario de dentro a fuera, y a grandes alturas sobre el plano galáctico.

Por último, se estudia la función de distribución de metalicidades. No se observan gradientes claros de metalicidad con el radio, mientras que sí se observa una ligera disminución de esta con la altura sobre el plano galáctico.

Nuestros resultados son coherentes con las secuencias químicas publicadas en la literatura. El dominio de poblaciones estelares jóvenes con el aumento radial se puede relacionar con la disminución de la abundancia de elementos α sobre la abundancia de Fe con el radio, siendo estrellas viejas las que presentan mayor proporción de estos elementos, así como con el escenario *inside-out*, que propone que las galaxias espirales se forman de dentro hacia afuera. La abundancia de elementos α también se puede relacionar con el dominio de poblaciones jóvenes conforme disminuye la altura sobre el plano galáctico. En cuanto a la metalicidad, a partir de las observaciones se puede ver como a medida que aumenta la altura las estrellas son menos metálicas. Además, las poblaciones estelares viejas y de edad intermedia con alta metalicidad pueden ser explicadas mediante el fenómeno de la migración radial, que expone que ciertas estrellas se han podido formar en radios más cercanos al centro galáctico y posteriormente se trasladaron a radios más lejanos.

Las conclusiones de este trabajo abren la puerta a diversas investigaciones futuras, por lo que el estudio de la SFH de la Vía Láctea no ha hecho nada más que empezar.

Contents

1 Introduction **5**
1.1 Objectives 7

2 Methods **8**
2.1 Mother diagrams 8
2.2 Dispersed mothers 9
2.3 dirSFH 9

3 The data **13**
3.1 *Gaia* EDR3 catalogue 13
3.2 Extinction maps 15
3.3 Selection of regions 18

4 Results **26**
4.1 Age-metallicity plane 26
4.2 Cumulative stellar mass fraction 30
4.3 Metallicity distribution function 32

5 Discussion **35**

6 Conclusions **37**

Acknowledgements **39**

References **40**

Appendices **43**
A Detected problems of Lallement+18 43
B Study of shifts 44

1 Introduction

Human beings have always wondered about their existence and their past. Questions such as "Where do we come from?" and "Why do we exist?" have been asked since the dawn of humanity. It is not our intention to focus on these questions, but we will try to make some progress by revealing the history of our galaxy.

A galaxy is a system of stars, stellar remnants, interstellar gas and dust bound together by gravity, embedded in a large dark matter halo. These cosmic entities span a wide range in terms of size and stellar population. Galaxies are formed through the presence of hydrogen, helium and dark matter following the Big Bang. Quantum fluctuations during the early universe led to the emergence of regions with varying densities. The denser regions underwent gravitational collapse due to the substantial amount of dark matter present. Within these collapsed regions, the gas cooled and fragmented into distinct clouds at the potential centers of dark matter halos, eventually giving rise to star formation. Feedback processes as massive star formation, supernovae and active galactic nuclei play a role in heating the gas and regulating the star formation process (Cimatti et al., 2019). However, many aspects of galaxy formation remain poorly understood. To try to learn more about how this formation takes place, we can study the formation and evolution of our own galaxy, the Milky Way.

The field of study that explores the history of our nearby universe is commonly referred to as galactic archaeology. It aims to investigate our Galaxy's history, examining the stars, interstellar gas and other structures, to understand the formation and development of our Galaxy. Through these detailed observations, researchers can reconstruct past events within the Galaxy's lifespan and unravel its evolutionary journey. To accomplish this, we focus our attention on a region within our reach, taking advantage of our privileged position within the Milky Way. By focusing on our galaxy, we can investigate it with an unparalleled level of precision that is currently unattainable for other galaxies. This exceptional laboratory provides us with a unique opportunity 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 Galactic regions close to the Sun up to the current time.

One of the primary methodologies employed in galactic archaeology is the analysis of the star formation history (SFH). It seeks the characterisation of the processes by which stars have emerged across both temporal and spatial dimensions, encompassing both rapid episodes of formation and longer periods. The analysis of the SFH's fundamental objective provides insights into the distribution of stellar ages and metallicities within a galaxy, thus offering valuable information about its evolution. The SFH as we will define in this work comprises both the evolution of the star formation rate (SFR) and the chemical enrichment as a function of time. This function quantifies the conversion of mass into stars of a given metallicity as a function of time.

The SFH can be inferred by analysing the present-day population of stars and examining their distribution over stellar type, characterised by their colour and magnitude. Both the magnitude and colour of a star undergo changes throughout its lifespan as it transitions through different evolutionary stages. By plotting these stellar properties on a colour-magnitude diagram (CMD),

analogous to the Hertzsprung-Russell diagram, the position of a star within the diagram provides valuable insights into its age.

The distribution of stars in the CMD derived from observations can be compared with models that incorporate collections of stars with varying masses, ages and chemical compositions. This approach, known as CMD fitting, allows us to assess the consistency between the observed stellar distribution and the predicted distributions from stellar population models. However, it is important to note that the accuracy of CMD fitting is inherently limited due to uncertainties related to the influence of age and chemical composition on the observed colour and luminosity of stars. These ambiguities can introduce challenges in accurately determining the age and chemical composition of stellar populations based solely on CMD fitting techniques.

Nevertheless, the advent of the *Gaia* mission has significantly improved our ability to resolve these ambiguities. The mission's astrometric data (which includes precise measurements of stellar positions and distances) along with its photometric data (measuring stellar magnitudes at different spectral bands) and spectroscopic data (providing information on metallicity, distance and extinction) offer a wealth of information for studying the Milky Way.

Gaia is a pioneering mission aimed at mapping the Milky Way in three dimensions (Gaia Collaboration, 2016). By doing so, *Gaia* unveils valuable insights into the composition, formation and evolution of our Galaxy. With remarkable precision, *Gaia* acquires positional and radial velocity measurements, enabling the creation of an extensive catalogue encompassing approximately one billion stars within our Galaxy and the Local Group. This accounts for roughly 1% of the entire stellar population in the Milky Way.

Gaia is stationed in a heliocentric orbit at the second Lagrange point, L2. Positioned approximately 1.5 million kilometres away from Earth in the anti-Sun direction, the L2 point remains synchronised with Earth as it completes its one-year orbit around the Sun. An L2 orbit presents a significant advantage of uninterrupted observations, free from eclipses. This orbital vantage point allows *Gaia* to capture a comprehensive view of the entire celestial sphere over the course of a year.

During its sky survey, *Gaia* is poised to unveil numerous scientific breakthroughs. Over its operational span, it will meticulously observe each of its one billion sources approximately 70 times, thereby capturing an extensive chronicle of the luminosity and positional changes exhibited by each source over time. Accurate astrometry and photometry have allowed us to derive with sufficient precision the SFH of the Galaxy, as well as formation processes such as the accretion of a satellite galaxy into the halo and thick disc (Gallart et al., 2019). Apart from the boost in galactic archaeology, *Gaia* exceptional precision of astrometric measurements will facilitate the detection of exoplanets orbiting distant stars, asteroids within our own Solar System, icy bodies situated in the outer reaches of our Solar System, brown dwarfs, supernovae in far-flung locations, and quasars. Furthermore, these observations will provide unprecedented opportunities for scrutinising Einstein's general relativity theory to an unprecedented degree.

1.1 Objectives

The project has outlined several main objectives, which are as follows:

- To investigate the SFH of various regions within the Milky Way Galaxy by employing the CMD fitting technique, assuming symmetry respect the Galactic plane.
- To establish the age and metallicity distribution of formed stellar populations in the Milky Way, examining both radial and vertical dependencies.
- To enhance the efficacy of the methodology employed in this study, aiming to refine and improve the accuracy of the SFH determination.
- To determine the optimal extinction map to be utilised for correcting extinction and reddening effects in the data analysis.

This report is structured as follows. In Section 2, we explain the methodology employed to derive the SFH, including details on the CMD fitting technique. Section 3 provides an overview of the data introducing the highlights of the *Gaia* EDR3 catalogue and the extinction maps we have in our disposition. We also select the regions to perform the CMD fitting technique and the extinction cut to use. The results obtained for the selected regions are presented in Section 4, doing a qualitative description of the age-metallicity plane and studying quantitatively the variation of the SFR and the metallicity distribution function (MDF) for each region. In Section 5 we discuss with the literature the results obtained, debating their implications and significance and analysing any discrepancies or agreements between the findings and previous studies. Finally, we summarise the conclusions reached, identifying areas for further research, and suggest future directions for the project in Section 6.

2 Methods

Utilising synthesised stellar populations to fit observational CMDs in the absolute plane has been an effective approach for extracting the SFHs of stellar systems, including our own Milky Way Galaxy. In this study, we employ an enhanced CMD fitting methodology specifically designed for *Gaia* data. This methodology involves a statistical search for the optimal linear combination of simple stellar populations (SSP) that accurately reproduces the observed CMDs. The methodology can be divided into three main components: the generation of synthetic diagrams (referred to as *mother diagrams*), the simulation of errors on these diagrams and the computation of the SFH (Ruiz-Lara et al., 2022a).

2.1 Mother diagrams

As previously mentioned, the *mother diagrams* are synthetic representations of stellar populations that are generated based on the predictions from stellar evolution libraries such as BaSTI (Hidalgo et al., 2018) and PARSEC (Bressan et al., 2012). These diagrams can be regarded as theoretical or synthetic CMDs that serve as the foundation for our analysis. By using these *mother diagrams*, we can compare them with observed CMDs to determine the best-fit stellar population models and derive meaningful information about the SFH.

The specific *mother diagram* utilised in our analysis, which has been provided by co-supervisor Dr Carme Gallart, is named `EDR3_q01_b03_120M_basti_iacSS_kroupa_MG5`. The name itself carries important information about the characteristics of the diagram. Let's break it down:

- The first word indicates the observational data which is going to be compared with, in our case, *Gaia* EDR3 catalogue. This implies a given set of bolometric corrections to go from the theoretically predicted effective temperature and luminosities, to magnitudes in the *Gaia* bands.
- q represents the minimum mass ratio for binaries and in this case, it is set to 0.1.
- β (in file's name, b) denotes the fraction of unresolved binaries, which is 30% in this case.
- The diagram consists of a population of 120 million stars with magnitude brighter than 5 in the M_G band.
- The stellar evolution library employed for generating the diagram is BaSTI, in its solar-scaled version.
- The initial mass function (IMF) used is the Kroupa IMF (Kroupa, 2001).

The stars within this *mother diagram* are generated with a constant star formation rate from 0.08 to 13.5 Gyr, and metallicities Z ranging from 0.0001 to 0.039. These characteristics enable us to compare the synthetic CMDs with observed CMDs and extract valuable information regarding the SFH of the system.

2.2 Dispersed mothers

A proper comparison of simulated CMD with observations requires the simulation of observational errors. It is also crucial to account for incompleteness, due to the stars removed for quality criteria. These observational effects introduce uncertainties in the measured colours and magnitudes of stars, deviating them from their true values. To address this, we employ a dedicated code called `New DisPar` (Ruiz-Lara et al., 2021), which is specifically designed for simulating these observational effects on the synthetic CMDs. This step is essential in order to accurately analyse and interpret the observed data and derive meaningful SFHs.

In the entry file, several parameters need to be specified to perform the error simulation. These parameters include:

- The observed file. This file contains the observed CMD data with no quality cuts or other filters applied.
- The *mother diagram*. The chosen *mother CMD* to be compared with the observed data.
- Specifications of the *mother diagram*. Additional details about the *mother diagram*, such as the mass of the *mother CMD* to disperse or the approximate maximum number of stars from the *mother CMD* to disperse in one iteration.
- The maximum allowed parallax error and extinction values that are allowed in the CMD that will be used for the SFH derivation. In this CMD, stars are filtered according to their parallax error and the amount of extinction present in the regions of interest.
- Extinction map. The extinction map to be used in the analysis, which will be discussed in subsection 3.2. This map provides information about the distribution of interstellar dust and its effects on the observed stellar properties.

To study different regions, it is necessary to have *dispersed mothers* specific to each region. These *dispersed mothers* are derived from the *mother diagram* and account for the unique characteristics and properties of each region.

2.3 dirSFH

Once we have dispersed the *mother diagram*, we can derive the SFH. This is performed using the program `dirSFH` (Gallart et al., in prep.), which is an improved version of previous tools like IAC-pop (Hidalgo et al., 2009) and TheStorm (Bernard et al., 2018). `dirSFH` is responsible for deriving the SFH from the dispersed *mother diagram* and observed CMDs. In short, `dirSFH` defines a series of SSPs from the error-convolved synthetic CMD using a Dirichlet tessellation from a grid of seed points within the available range of ages and metallicities. The code then finds the combination of SSPs that best fits the observed CMD based on the Skellam probability distribution of the difference between two statistically independent distributions (observed and simulated CMDs).

The program utilises a *.in* file as input, where various specifications are provided. These include the observed file (which can now incorporate quality cuts), the dispersed *mother diagram*, the desired names for each test, and the output path. Additionally, it is possible to designate

distinct paths or names for individual files generated during the execution, rather than assigning a single name to the entire test. The code can be segmented into four distinct processes or components:

1. *dirPrep*. Transformation of the data for subsequent processes by extracting only the essential information required for the SFH calculation. This entails reading the specified observed and *dispersed mother* files, which are in the *.hdf5* format, and converting them into the more accessible *.rds* format. This conversion facilitates efficient access to the data during the subsequent stages.

2. *dirWin*. Definition of the region within the CMD, known as a *bundle*, whose stars will be utilised for the computation of the SFH. Once the *bundle* is interactively delimited, several calculations take place. Firstly, observed CMD matrices are computed, i.e., a grid that records the number counts of stars within boxes defined by colour and magnitude. Additionally, weights are calculated across the dispersed *mother diagram* (which will be employed in *dirall.R* with the "wei" option). These weights are determined as the logarithm of the inverse variance of ages across the synthetic CMD. In essence, the weights are assigned based on the representativeness of a given age population within a specific region of the CMD. Within the input file, it is possible to specify the breakpoints for colour and magnitude in the CMD matrices, i.e., the boundaries of the two variables utilised for plotting the CMD in the pop-up window, along with their corresponding grid size.

3. *getSM*. Estimation of the shift between the observed CMD and the dispersed *mother diagram*, which arises due to the possibly imperfect colour transformations in the models to the observed photometric system. To determine this shift, the observed CMD is randomly shifted within specified limits in both colour and magnitude. For each shift, a SFH solution is computed. By considering the residuals from all these solutions, the likely shift between the observed and *mother CMD* is determined. In the input file, you can specify the lengths of the colour and magnitude axes of the initial CM-shift ellipse, the number of evaluation points (typically around 2000), and the number of iterations (usually 3). Additionally, you need to specify the age and metallicity bins as "seeds".

4. *dirall*. Calculation of a specified number of SFHs using the files generated in the previous steps. The weight matrix created by *dirWin* may be utilised, or uniform weights across the CMD can be indicated. These SFHs are computed by exploring different dissections of the age-metallicity space by semi-randomly changing the set of seeds from the original ages and metallicities. Once all the SFHs are calculated, they are combined to obtain the final SFH, with the variations between individual solutions serving as uncertainties. In our case, unless specified otherwise, we use 80 individual runs for the final results and 64 for the tests. Additionally, we utilise the *weighted* class rather than the *uniform* class. In the input file, we can specify which magnitude and colour shifts should be considered in the process and the transformation parameters between metallicity Z and iron abundance ratio [Fe/H].

Below are the parameters we considered and included in the input file of `dirSFH` to obtain the results of our study:

- A *bundle* comprising the highest concentration of stars in the CMD within approximately the magnitude range of $-3 < M_G < 4.2$ and the colour range of $-0.5 < (G_{BP} - G_{RP})_0 < 2.5$. The faint limiting magnitude of this *bundle* corresponds to the magnitude of the subgiant branch of the oldest and more metal-rich population included in the *mother CMD*.
- A customised initial age-metallicity bin configuration, featuring specific age values of (0.08, 0.126, 0.192, 0.262, 0.334, 0.404, 0.469, 0.532, 0.596, 0.656, 0.718, 0.784, 0.857, 0.938, 1.028, 1.128, 1.244, 1.391, 1.576, 1.81, 2.066, 2.337, 2.609, 2.882, 3.156, 3.427, 3.695, 3.978, 4.272, 4.581, 4.946, 5.389, 5.858, 6.351, 6.861, 7.372, 7.883, 8.393, 8.904, 9.415, 9.925, 10.436, 10.947, 11.457, 11.968, 12.479, 12.989, 13.5) Gyr, and metallicity values ranging from $Z = 0.001$ to $Z = 0.039$, with increments of 0.1 dex in the chemical abundance ratio [M/H].
- The calculated colour-magnitude shift for the *Gaia* Catalogue of Nearby Stars (GCNS) sample, (-0.036, 0.04).[1]

In Figure 1, which is an example of the kind of output figure generated by `dirSFH`, we can observe a summary of the SFH results presented in various projections. The two panels in the top left corner represent the observed CMD and the solution provided by `dirSFH`, respectively, colour-coded with the density of stars and with the defined *bundle* superimposed. The subsequent panel in the top row illustrates the resulting SFR as a function of time, while the following panel

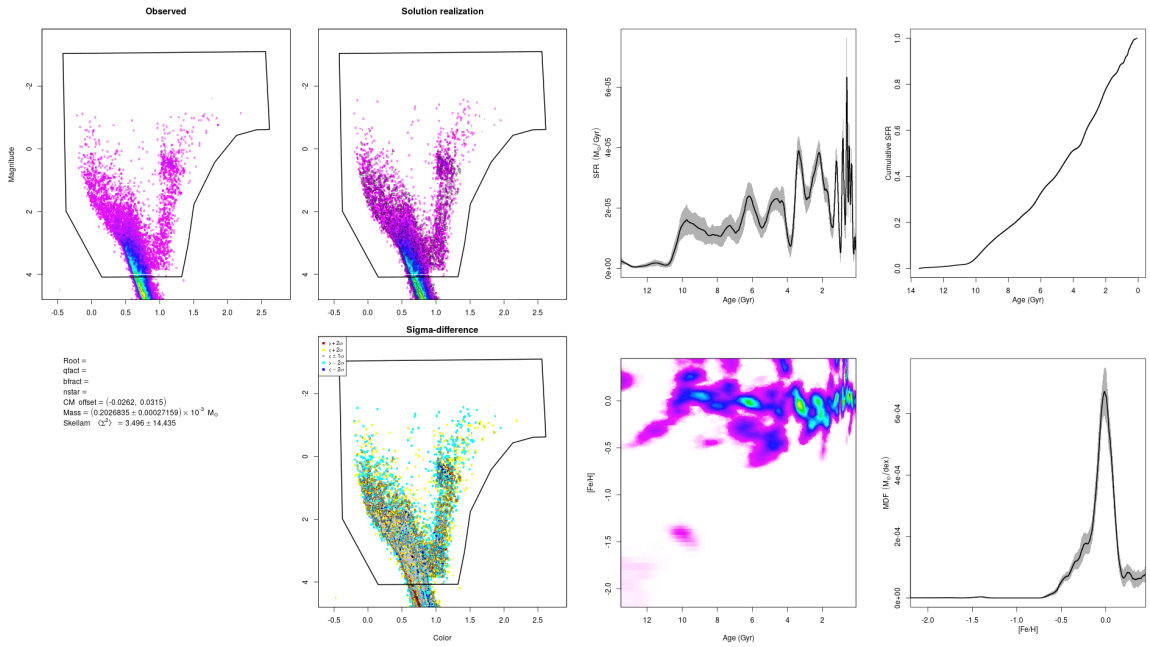

Figure 1: *Output plots of* `dirSFH`*. Top panels, from left to right: observed CMD, simulated CMD, SFR, cumulative SFR. Bottom panels: residuals, SFH in the age-metallicity plane and MDF.*

[1]For detailed information on the methodology and rationale behind the selection of this shift, please refer to appendix *B. Study of shifts*.

displays its cumulative distribution. The first panel in the bottom row shows the residuals between the observed and calculated CMD, providing insights into the discrepancies between the two. Moving on, the plot in the middle represents the SFR (colour code), in solar masses per unit of time and metallicity, in the age-metallicity plane, revealing the metallicities of stars formed at different ages. Lastly, the bottom right panel presents the MDF, which depicts the distribution of stellar mass converted into stars at each metallicity.

It is also possible to derive the MDF as the number of stars per metallicity bin from the metallicity values of each star in the CMD solution. A similar age distribution function can also be derived. This would be the distributions corresponding to the stars alive today and comparable with observations.

3 The data

This section introduces the catalogue utilised for the project and outlines the process undertaken to carefully curate the appropriate data for deriving the SFH. Additionally, we highlight the selection criteria employed to ensure the accuracy and reliability of our dataset, as well as the choice of extinction map and regions to study.

3.1 *Gaia* EDR3 catalogue

The *Gaia* Early Data Release 3 (*Gaia* EDR3), which was made publicly available on the 3rd December 2020, serves as the primary catalogue utilised in this project (Gaia Collaboration, 2021a). It encompasses data gathered over a duration of 34 months, from the 25th July 2014 (10:30 UTC) to the 28th May 2017 (08:44 UTC). The *Gaia* EDR3 data can be obtained through the *Gaia* Archive and partner data centres (CDS, ASDC, ARI, AIP, and Flatiron). Researchers can extract the desired information by executing ADQL queries, and subsequently downloading the corresponding results tables from the *Gaia* Archive.

The *Gaia* EDR3 serves as a precursor to the *Gaia* Data Release 3 (*Gaia* DR3), which was published on the 13th June 2022 (Gaia Collaboration, 2023a). *Gaia* DR3 incorporates previously published data products with additional new ones, including spectroscopic information, all derived from the same time frame and set of observations. For the purpose of our specific work, the following contents of *Gaia* EDR3, which are unchanged in DR3, are particularly noteworthy:

- $1,467,744,818$ sources with a comprehensive astrometry solution, providing precise positions in the sky, parallax, and proper motion. The limiting magnitude for these sources is approximately $G \approx 21$, with a bright limit of about $G \approx 3$.
- $1,806,254,432$ sources with available G magnitudes.
- $1,542,033,472$ sources with G_{BP}-band photometry.
- $1,554,997,939$ sources with G_{RP}-band photometry.

By utilising the parallaxes provided in the catalogue, we derive the distances d to the stars, taking into account the parallax zero point of 0.017 mas determined in Lindegren et al. (2021). We employ the distance modulus equation to convert the apparent magnitude G to the absolute magnitude M_G, leveraging the ultra-precise photometric data available. This conversion also takes into account the impact of interstellar extinction A_G caused by dust and gas along the line of sight:

$$M_G = G - 5\log d[\text{pc}] + 5 - A_G \tag{1}$$

For the colour, we subtract the reddening $E(G_{BP} - G_{RP})$ to the value of colour given by *Gaia*, $G_{BP} - G_{RP}$:

$$(G_{BP} - G_{RP})_0 = G_{BP} - G_{RP} - E(G_{BP} - G_{RP}) \tag{2}$$

Both extinction and reddening calculation will be explained in next section, *3.2. Extinction maps*.

In Figure 2, we present the CMD of the parent sample from which we will extract subsamples corresponding to different Milky Way volumes. It is worth noting that the CMD is subject to a

limiting magnitude of 5.5 and has undergone quality cuts (that will be explained latter). Within the diagram, we can discern various prominent structures, particularly in regions of higher stellar density (indicated by yellower and greener areas).

The main sequence (MS) is observed as a diagonal band extending from bright blue stars ($(G_{BP} - G_{RP})_0 \sim 0$ and $M_G \sim -4$) to faint red stars ($(G_{BP} - G_{RP})_0 \sim 1$ and $M_G \sim 5.5$), representing stars that are currently undergoing hydrogen fusion in their cores. As these stars exhaust their hydrogen fuel, they depart from the MS, marking the turn-off point (TO) ($(G_{BP} - G_{RP})_0 \sim 1$ and $M_G \sim 4$). The TO is a vital indicator for determining the age of a single stellar population: a brighter and bluer TO signifies a younger stellar population.

Following the TO, stars initiate hydrogen fusion in a shell and transition to the red giant branch (RGB), where their luminosity and size increase while their temperature decreases (until $(G_{BP} - G_{RP})_0 \sim 3$ and $M_G \sim -2$). Subsequently, stars may reach the helium ignition stage in their cores, progressing to the horizontal branch (HB) or the red clump (RC), which is clearly distinguishable in this CMD at $(G_{BP} - G_{RP})_0 \sim 1.5$ and $M_G \sim 0.5$.

After the core He-burning phase, stars enter the asymptotic giant branch (AGB) phase, which is not clearly distinguished in this CMD from the RGB. During the AGB phase, helium fusion occurs in an additional layer around the inert carbon core, while hydrogen fusion takes place in another one surrounding the helium-burning region. Eventually, a star may shed its envelope and may transform into a white dwarf (WD). As white dwarfs have $M_G > 5.5$, they are not visible in this CMD.

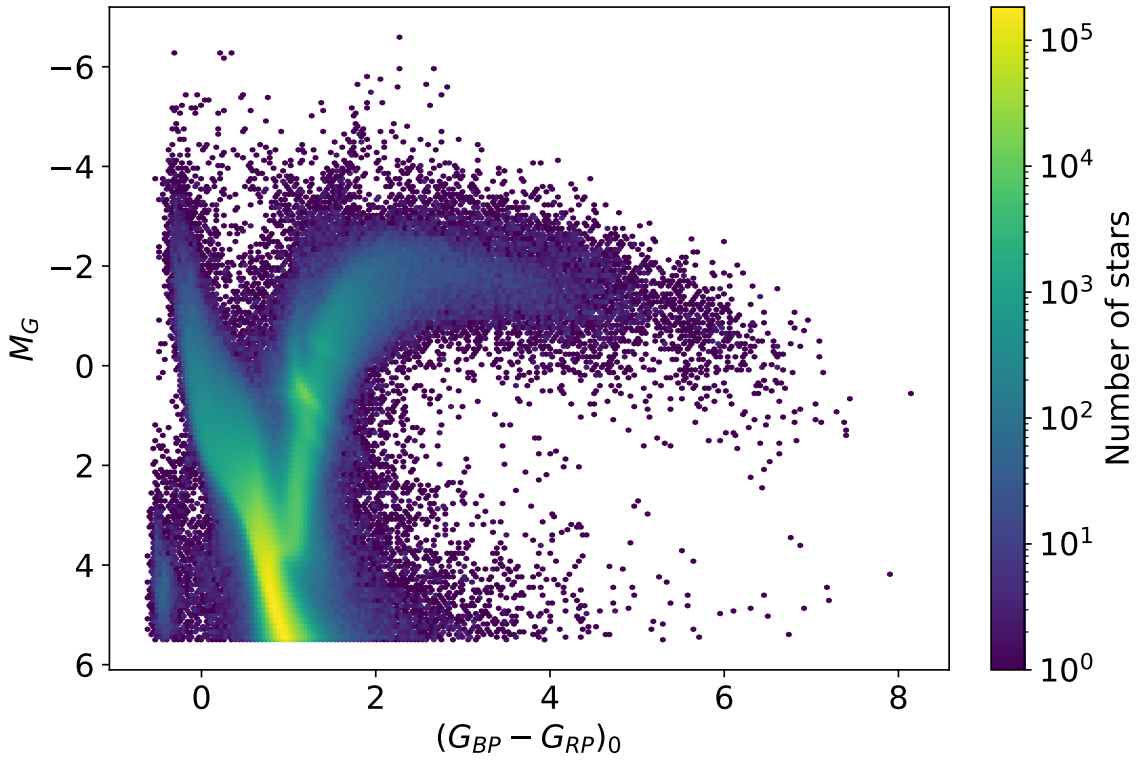

Figure 2: *CMD of the* 2 kpc *bubble around the Sun, using Lallement+22 extinction map with quality cuts and colour-coded according to the number of stars with logarithmic normalisation.*

3.2 Extinction maps

The term *interstellar matter* is commonly used by astronomers to refer to all the material present between stars. This vast collection of material is known as the interstellar medium (ISM). The ISM is primarily composed of gas, accounting for about 99% of the material between stars. This gas exists in the form of individual atoms or molecules. In contrast, the remaining 1% of the ISM is comprised of solid particles referred to as interstellar grains or interstellar dust. These dust grains are frozen conglomerates made up of numerous atoms and molecules. A typical interstellar dust grain consists of a core composed of rock-like substances, surrounded by a mantle of ices. Dusty clouds in space obstruct light, emit energy in the infrared, reflect light, and cause stars to appear redder. It is this latter effect that we are particularly interested in discussing.

Interstellar dust absorbs and scatters starlight, leading to dimming and reddening effects known as interstellar extinction and reddening. Different colours of light interact with dust in varying ways, with shorter wavelengths being scattered or absorbed more than longer wavelengths. This results in stars appearing redder and, in extreme cases, invisible at visible wavelengths.[2]

Dust that causes extinction in the Milky Way Galaxy is primarily concentrated along its spiral arms, a pattern observed in other spiral galaxies as well. Astronomers have conducted extensive studies to determine the three-dimensional distribution of extinction within our region of the Galaxy. These investigations involve analysing visible and near-infrared observations of stars, combined with a model of stellar distribution. The result of this analysis is commonly referred to as an extinction map.

An extinction map can be considered a celestial cartographic representation that provides information about the extent of extinction in a specific space direction and distance. These maps are invaluable tools for correcting the effects of extinction in astronomical data, allowing for more precise measurements of the luminosity, colour, and other properties of celestial objects.

The differences among the various 3D reconstructions of extinction primarily stem from the selection of photometric and/or spectro-photometric catalogues used to estimate individual extinctions, the methodology employed for determining distances (such as parallax, photometry, or a combination of both), and the specific inversion technique utilised. Within the scope of our project, we will specifically focus on two extinction maps: one derived by Green et al. (2019) (which we will refer to as Green) and another by Lallement et al. (2022) (Lallement+22).[3]

On one hand, the Green extinction map utilises *Gaia* DR2 parallaxes and broadband optical and near-infrared stellar photometry obtained from Pan-STARRS 1 and 2MASS. This comprehensive dataset covers the stellar population in the northern sky with a declination larger than -30° and extends up to a few kiloparsecs in distance. Stars are employed as tracers of the dust

[2]It is worth noting that the term *reddening* used to describe this process is not entirely accurate, as no red colour is added to the star's light. Instead, the effect arises from the subtraction of blues and related colours, leading to a relative increase in the prominence of red light. Strictly speaking, the process could be more accurately described as *deblueing* (Fraknoi et al., 2018).

[3]Prior to acquiring the Lallement+22 data, our study initially utilised the Lallement et al. (2018) extinction map (Lallement+18), an earlier version of the same map. However, during our analysis, we encountered certain complexities and challenges associated with this specific map. For a detailed account of these complications, please refer to the appendix *A. Detected problems of Lallement+18.*

column, and their apparent magnitudes and parallaxes are modelled as a function of their type, distance, and foreground reddening. By grouping these stars into discrete sightlines and utilising their distances and reddening values, the distribution of dust as a function of distance can be constrained. Notably, the uncertainties in the derived reddening values are approximately 30% smaller compared to those reported in the *Gaia* DR2 catalogue, which can be attributed to the inclusion of a greater number of photometric passbands in the analysis.

On the other hand, the Lallement+22 extinction map combines *Gaia* EDR3 photometric data with 2MASS measurements to determine the extinction towards stars with precise photometry and relative uncertainties of less than 20% on EDR3 parallaxes. This map employs distance-extinction pairs of stars in the vicinity of the Sun to estimate the extinction at three-dimensional coordinates. The celestial space is divided into volumes, and the average extinction within each volume is computed. Subsequently, inversions are performed, and the extinction density profiles are accumulated in a three-dimensional cube. By utilising averaged data and employing a Bayesian inversion approach, the extinction density model is iteratively refined at various scales. This iterative process continues until a desired resolution is reached, resulting in three maps with different resolutions: 10, 25, and 50 parsecs. Depending on the region of the Galaxy being studied, we utilise one of the three Lallement+22 maps to obtain accurate estimates of extinction. This approach allows us to account for variations in extinction across different regions of space and obtain reliable extinction values for our analysis.

Now, we need to determine which extinction map is better for our study. To evaluate the strengths and weaknesses of both extinction maps, we begin by plotting the number of stars in the galactocentric x-y plane of a specific layer z, such as $0.1 < |z|$ [kpc] < 0.2, using data from both maps (Figure 3).

As depicted in Figure 3, the Green data exhibits a lack of stars in a particular quadrant. This discrepancy arises from the fact that, as we commented before, the Green extinction map solely utilises observations from the northern hemisphere, thereby impeding extinction estimations

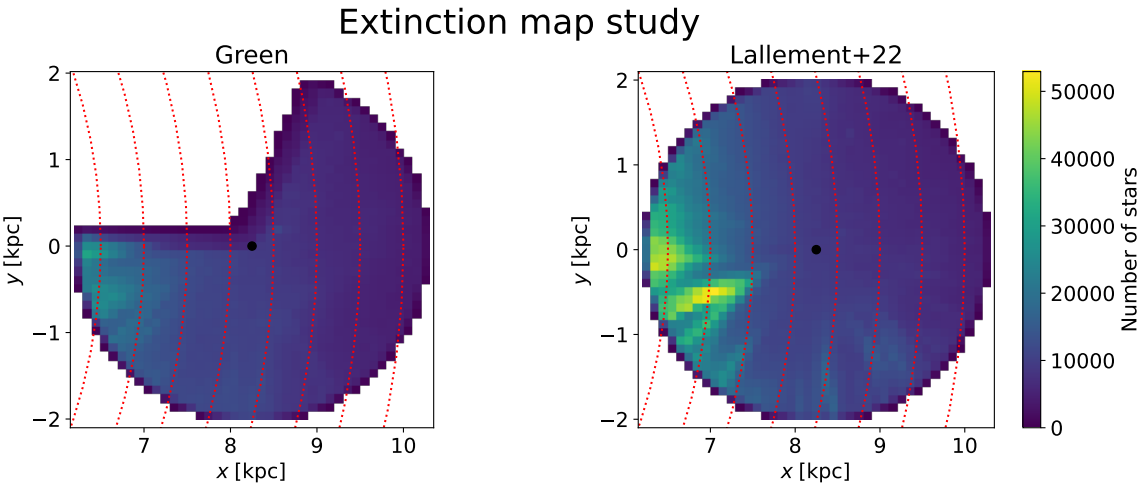

Figure 3: *Radial plane of the* $0.1 < |z|$ [kpc] < 0.2 *layer with the number of stars with* $M_G < 5.5$ *in* 0.1×0.1 kpc *bins. The black dot marks the place where is the Sun, whereas the red dotted lines represent galactocentric radius* R *with values* 6.5, 7, 7.5, 8, 8.5, 9, 9.5 *and* 10 kpc. *Left: Green extinction map. Right: Lallement+22 extinction map.*

for stars in the southern hemisphere. Consequently, at first glance, one might perceive the Lallement+22 data as a better option for our purpose due to its comprehensive coverage of the entire sky.

Next, we perform a comparison between our derived MDF of alive stars, for each extinction map, and an observational MDF. To accomplish this, we apply `dirSFH` to a cross-match between the Green map and the Lallement+22 map stars. That is to say, we select only the stars for which extinction values are available in both maps.

For the observational MDF, we consider the seventeenth data release (DR17, Abdurro'uf et al. (2022)) of the high-resolution ($R \sim 15,000$) spectroscopic survey Apache Point Observatory Galactic Evolution Experiment (APOGEE, Majewski et al. (2017)), part of the Sloan Digital Sky Survey (SDSS, Blanton et al. (2017)) program. This survey obtained near-infrared stellar spectra of mostly RGB stars in the Galactic disc. It is one of the most complete and accurate high-resolution spectroscopic surveys of disc Milky Way stars. From spectral fitting with a library of synthetic spectra the APOGEE collaboration derived stellar parameters and chemical abundances, in particular the abundance of iron, [Fe/H] (Jönsson et al., 2020).

We will focus our analysis on the same layer as in Figure 3, which corresponds to $0.1 < |z|$ [kpc] < 0.2, and on the solar radius bin, $8.0 < R$ [kpc] < 8.5. Not all stars that are present in our *Gaia* CMD have spectroscopic measurements, and the stars lacking them are not distributed evenly across the different evolutionary phases present in the CMD. To ensure a fair comparison, we divide the CMD into small boxes and calculate the fraction of stars with spectroscopic metallicity relative to the number of observed stars in our sample within each box. Subsequently, we randomly select the same fraction of stars from each simulated CMD region based on these percentages. This approach allows us to assign weights that replicate completeness, resulting in the creation of a spectroscopic simulated CMD. We calculate the MDF of this last CMD and convert the metallicity Z obtained from our code into [M/H] using the following formula[4] with the coefficients extracted from Hidalgo et al. (2018),

$$[\mathrm{M/H}] = \log\left(\frac{Z}{1.0 - (0.247 + 1.31Z) - Z}\right) + 1.67985 \tag{3}$$

We also mimic the observational errors by adding to these [M/H] solution values a normal distribution with a standard deviation of 0.15, the typical dispersion observed in APOGEE [Fe/H] measurements.

We compare the resulting MDF with the one derived from APOGEE observations. The results are shown in Figure 4. It is evident that in both cases, the results are strikingly similar. The overall shape of the spectroscopic MDF is faithfully reproduced, with the primary peak aligning in both extinction maps. Additionally, an extended tail corresponding to metal-poor stars is observed in both simulations. This indicates that the choice of extinction map has minimal impact on the metallicity results. However, it should be noted that the Green map tends to overestimate metal-poor regions, and there are instances in both solutions where metal-rich stars are present but not observed in the actual spectroscopic data.

[4]This form comes from the development of the following equations:

$$[\mathrm{M/H}] = \log(Z/X) - \log(Z/X)_\odot \qquad X = 1 - Y - Z \qquad Y = 0.247 + 1.31Z$$

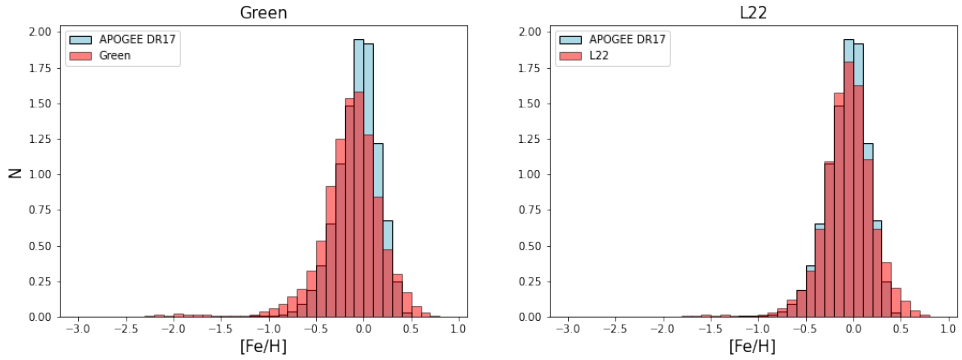

Figure 4: *Comparison of the APOGEE MDF (blue bars) with the MDF obtained from the solution CMD (red bars) for the region* $0.1 < |z|$ [kpc] < 0.2 *and* $8.0 < R$ [kpc] < 8.5. *Left: MDF obtained from the solution CMD using Green extinction map. Right: MDF obtained from the solution CMD using Lallement+22 extinction map.*

There are no dramatic distinctions between the two extinction maps. Consequently, we have opted to select the Lallement+22 map for our analysis. This decision is based on the aforementioned reasons, including its comprehensive coverage of the entire sky and the better replica of the metallicities derived considering their extinction corrections with observational constraints.

3.3 Selection of regions

In this subsection, we will select the specific layers and radii for conducting our study on the SFH spatial variation. Additionally, we will determine the appropriate extinction cut to be applied in our analysis.

In order to choose the regions to perform our analysis we need to evaluate where we have enough completeness in the CMD to be able to infer the SFH with our methodology. We have evaluated two different sets of data (provided by Dr Tomás Ruiz-Lara and co-supervisor Dr Emma Fernández-Alvar). The first set, referred to as the *full* data, includes all stars observed with *Gaia* with an absolute magnitude M_G smaller than 5.5 (bright stars) within a 2 kpc radius around the Sun. The second set, referred to as the *clean* data, is a subset of the *full* data that has passed certain quality selection criteria. These criteria include:

- A `parallax_over_error` ratio greater than 5. It indicates that the error in the parallax measurement is less than 20% of the parallax value.
- A $(G_{BP} - G_{RP})$ excess factor values, obtained from Ruiz-Lara et al. (2022b), verifying

$$0.001 + 0.039(\texttt{bp\_rp}) < \log(\texttt{phot\_bp\_rp\_excess\_factor}) < 0.12 + 0.039(\texttt{bp\_rp})$$

- Low extinction stars, with a tentative extinction value `AG` less than 0.5.

By applying these selection criteria, we ensure that the *clean* data consists of high-quality stars with reliable parallax measurements and low extinction values, and thus, well determined colours and magnitudes.

Before analysing the data, we performed a coordinate transformation from the galactic spherical coordinate system to a Cartesian system centred at the Galactic centre. The galactic coordinate system is a celestial one that uses spherical coordinates, with the Sun as its centre, the primary direction aligned with the centre of the Milky Way Galaxy, and the fundamental plane parallel to the Galactic plane but offset to its north.

To perform the transformation, we used the common spherical to Cartesian conversion, taking into account the distance between the Sun and the Galactic centre. We consider the value 8.249 kpc as in GRAVITY Collaboration (2020). Additionally, we considered that the Sun is located approximately 0.0196 kpc above the Galactic plane (Reed, 2006). This vertical offset is relatively small compared to the distance between the Sun and the Galactic centre, so it has a negligible effect on the x and y coordinates.

The transformation equations are as follows:

$$x\ [\text{kpc}] = 8.249 - d \sin b \cos l \tag{4}$$

$$y\ [\text{kpc}] = -d \cos b \sin l \tag{5}$$

$$z\ [\text{kpc}] = 0.0196 + d \sin b \tag{6}$$

Here, b, l, and d represent the galactic latitude, longitude, and distance (as the inverse of the parallax) obtained from *Gaia*, respectively. In Equation 4, the minus sign is included because $\cos l$ is positive for stars closer to the Galactic centre, and in Equation 5, the minus sign is included to ensure that y is negative for stars to the left of our reference point.

Once we have calculated the values of x and y for each star, we can determine the radius of the star projected onto the plane of the Galactic disc R, measured from the Galactic centre, also known as galactocentric radius. The radius is simply the magnitude or modulus of the two coordinates mentioned earlier, given by:

$$R\ [\text{kpc}] = \sqrt{x^2 + y^2} \tag{7}$$

First, we assess the distribution of stars within each R and z bin to determine the stellar density in different regions of the Galaxy. This analysis is crucial in identifying areas with an adequate number of stars spanning the CMD. It allows us to identify regions where our methodology can be applied reliably and yield statistically significant results. By considering the population density across various R and z bins, we ensure the robustness of our approach and the validity of our conclusions.

In the left panel of Figure 5, we present the distribution of star counts for the *clean* sample as a function of the radius R of the projection on the Galactic plane for a number of layers with specific z ranges in bins of 100 pc. The R bins were defined of 0.1 kpc width, with the two first and last bins grouped together due to the lower number of stars in those regions ($6.249 < R\ [\text{kpc}] < 10.249$ for the Galactic plane layer). Consequently, the first bin spans from 6.2 kpc to 6.4 kpc, while the last bin ranges from 10.1 kpc to 10.3 kpc. In the plot, we have included multiple lines, each corresponding to a different limiting magnitude. These lines represent the number of stars within an R bin that have a magnitude brighter than a specific threshold. It is important to note that the depicted stars fall within the CMD region defined by $M_G > -3$ and $-0.5 < (G_{BP} - G_{RP})_0 < 2.5$, ensuring that only stars within this magnitude and colour range are considered in the analysis.

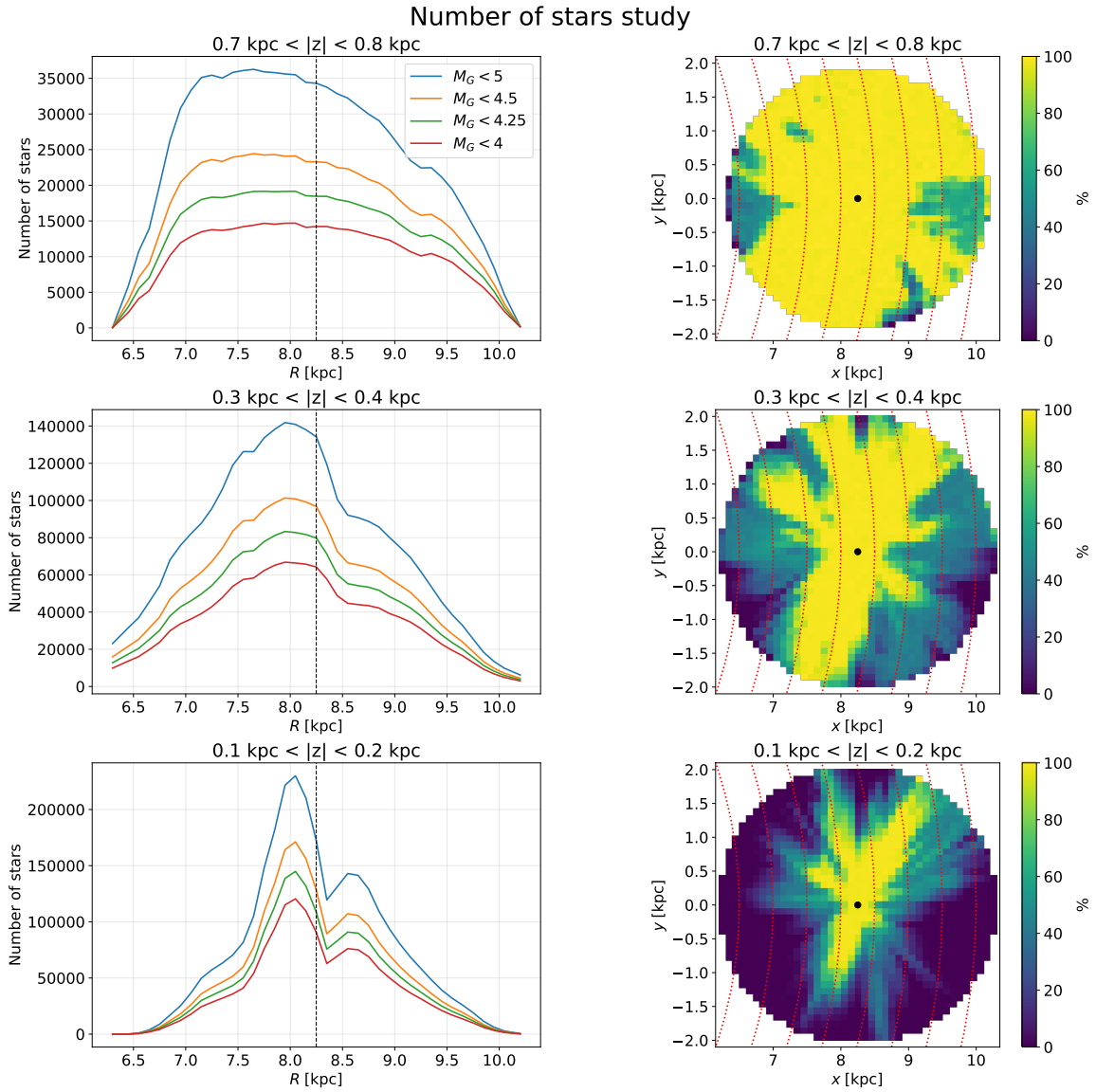

Figure 5: *Plots of some z layers obtained using the Lallement+22 extinction map. From bottom to top:* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers. Left: Number of clean stars brighter than different limiting magnitudes, and with* $M_G > -3$ *and* $-0.5 < (G_{BP} - G_{RP})_0 < 2.5$, *as a function of* R, *with bins of* 0.1 kpc. *Blue, orange, green and red solid lines indicate the limiting magnitude* 5, 4.5, 4.25 *and* 4, *respectively. The black dashed line shows the position of the Sun. Right: Percentage of stars in the bundle region that have passed our quality process in the* x-y *plane. The black dot marks the place where is the Sun, whereas the red dotted lines represent* R *with values* 6.5, 7, 7.5, 8, 8.5, 9, 9.5 *and* 10 kpc.

The second graphic we generated (right panel of Figure 5) illustrates the completeness of stars for each layer z. Specifically, it represents the percentage of stars that have passed our quality selection process (*clean* stars) compared to the total number of stars (*full* stars) for each z bin in the x-y plane. We also only choose the stars in the region of the CMD in which the bundle is contained, i.e., $-3 < M_G < 4.2$ and $-0.5 < G_{BP} - G_{RP} < 2.5$.[5] To ensure a meaningful representation, the x-y bins have a fixed size of 0.1 kpc. This allows us to assess the completeness of stars in different areas of the radial plane with a consistent spatial resolution.

Based on the analysis of the plots for each z layer, we have decided to focus our study on the SFH of the following layers:

- $0.1 < |z|$ [kpc] < 0.2.
- $0.3 < |z|$ [kpc] < 0.4.
- $0.7 < |z|$ [kpc] < 0.8.

They represent a sample with a good number of stars in a layer near the Galactic plane (dominated by thin disc stars), a little bit far from the Galactic plane (where we expect a larger mix between thin and thick disc stars) and far from it (dominated by thick disc stars). We have discarded the Galactic plane layer, $0.0 < |z|$ [kpc] < 0.1, due to the great extinction suffered by the stars in this part of the Galaxy. Also we did not choose higher layers because the number of stars and the number of eligible radial bins are small to carry out our analysis. It is worth recalling that the radius of the data circle in the radial plane decreases with height because our data is a sphere around the Sun.

For each layer, we have selected specific bins in the radial direction R based on the completeness and the number of stars available after quality cuts. We have chosen a bin size of 0.5 kpc for R. The selected bins for each layer are as follows:

- For the $0.1 < |z|$ [kpc] < 0.2 layer, we have 3 bins ranging from 7.5 kpc to 9.0 kpc.
- For the $0.3 < |z|$ [kpc] < 0.4 layer, we have 5 bins ranging from 7.0 kpc to 9.5 kpc.
- For the $0.7 < |z|$ [kpc] < 0.8 layer, we have 5 bins ranging from 7.0 kpc to 9.5 kpc.

In the first layer, only three radial bins have been chosen because the completeness is low due to the high extinction of the layers close to the Galactic plane, as can be seen in the right part of the Figure 5. By selecting these specific layers and radial bins, we aim to obtain a representative sample of stars in different regions of the Galaxy, allowing us to study the SFH as a function of galactocentric radius and distance from the plane.

Now that we have defined the layers and radius for computing the SFH, we will conduct a brief study on the distribution of stars based on their extinction values in order to explore whether we can restrict more our upper extinction limit. Additionally, we will consider the number of stars with extinction values below certain thresholds.

For a more comprehensive visualisation, we have generated histograms depicting the number of stars within each extinction bin (Figure 6). The range of extinction values considered is $0 < A_G < 1$, and the histogram bins have a size of 0.05 magnitudes. These histograms provide

[5] To save time, we will refer to this region of the CMD as the *bundle region* from now on.

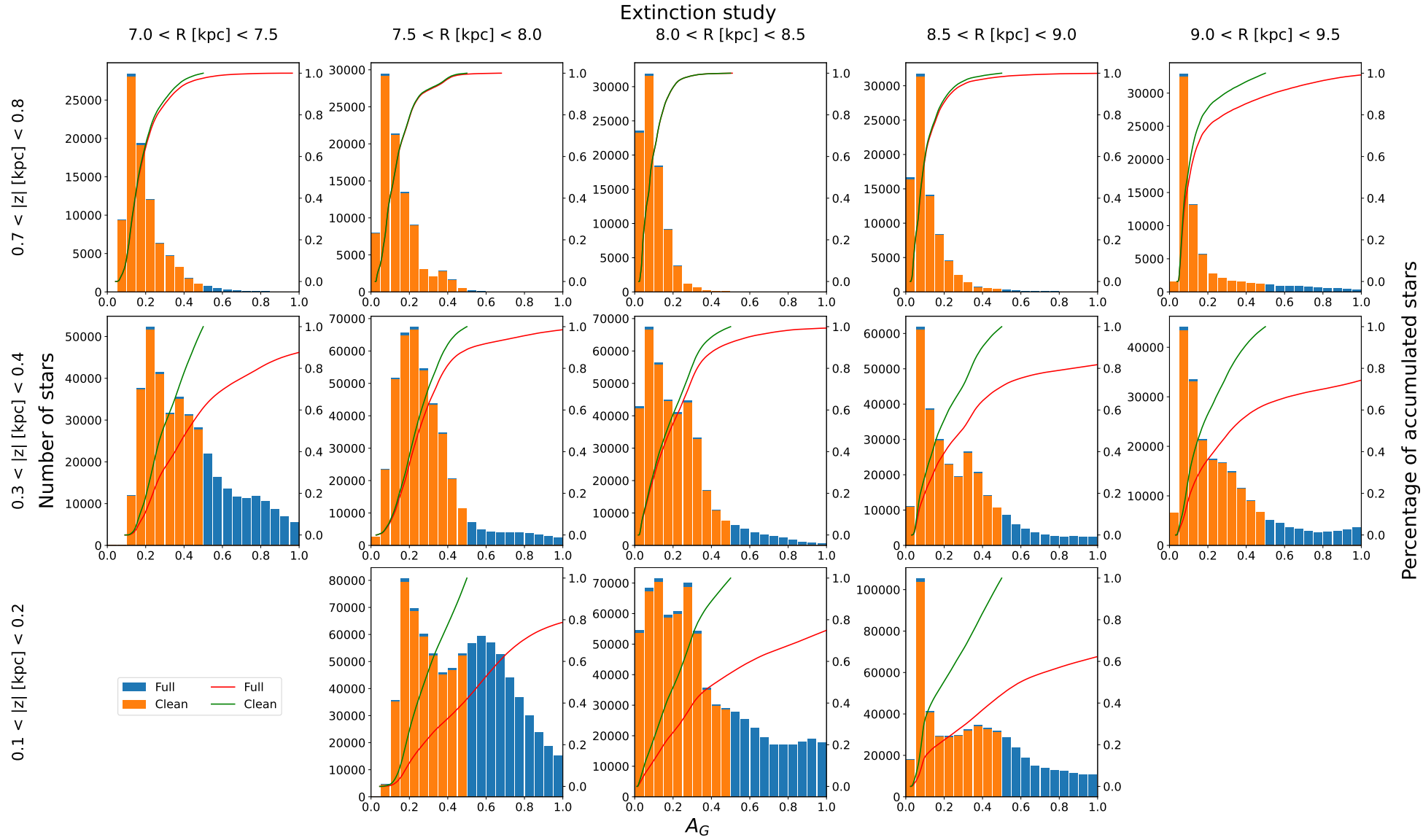

Figure 6: *Extinction histograms, with bins of* 0.05 *magnitudes in the range* $0 < A_G < 1$*, of our study regions obtained with Lallement+22 extinction, for both full (blue bars) and clean data (orange bars) in the bundle region. The plotted red and green lines indicate the cumulative function for full and clean data, respectively. From bottom to top:* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers. From left to right:* $7.0 < R$ [kpc] < 7.5, $7.5 < R$ [kpc] < 8.0, $8.0 < R$ [kpc] < 8.5, $8.5 < R$ [kpc] < 9.0 *and* $9.0 < R$ [kpc] < 9.5 *radius bins.*

insights into the distribution of stars across different levels of extinction for each region. The comparison between the *clean* and *full* data allows us to assess the impact of our quality selection criteria on the completeness of the selected samples.

Additionally, we have plotted the cumulative function, which represents the fraction of stars with extinction values below a given threshold. By examining the cumulative function for both the *clean* and *full* data, we can evaluate the completeness of our dataset concerning different extinction thresholds.

We observe a notable increase in the number of stars with $A_G > 0.5$ for lower R and $|z|$, especially for the second one. This finding aligns with theoretical expectations, as the Galactic plane is known to harbour a significant concentration of interstellar gas and dust. The comparable heights of the bars in both datasets suggest that the exclusion of stars from the analysis, due to the quality cut-off, can be attributed basically to the imposed extinction cut-off.

Indeed, it is crucial to strike a balance between excluding stars that may be affected by high extinction and ensuring an adequate sample size that is representative of the full sample for a robust statistical analysis. The selection of an appropriate extinction threshold should be guided by a thorough evaluation of the data, taking into account factors such as the completeness and reliability of the sample.

The cumulative function provides valuable information to guide the choice of the extinction limit. It represents the percentage of stars below a certain extinction value compared to the total sample, indicating the completeness of the *clean* data relative to the *full* data. For instance, in the first layer, at $A_G = 0.5$, the *full* percentage is approximately 0.5, indicating that only half of the stars in the region are selected. If we were to lower the extinction threshold to $A_G = 0.3$, we would retain only about 30% of the total stars. Therefore, after considering the cumulative distribution across different z and R values, we have decided to maintain the initial extinction limit of $A_G < 0.5$. This choice strikes a reasonable compromise, as the cumulative function

Layers [kpc]	**Radius bins [kpc]**	**Clean**	**Full**	**Completeness [%]**
$0.1 < \|z\| < 0.2$	$7.5 < R < 8.0$	442,609	1,072,464	41.27
	$8.0 < R < 8.5$	524,884	982,531	53.42
	$8.5 < R < 9.0$	378,701	873,162	43.37
$0.3 < \|z\| < 0.4$	$7.0 < R < 7.5$	266,479	442,441	60.23
	$7.5 < R < 8.0$	371,457	422,299	87.96
	$8.0 < R < 8.5$	360,631	395,901	91.09
	$8.5 < R < 9.0$	252,709	359,464	70.30
	$9.0 < R < 9.5$	179,018	290,774	61.57
$0.7 < \|z\| < 0.8$	$7.0 < R < 7.5$	85,573	88,711	96.46
	$7.5 < R < 8.0$	90,542	92,063	98.35
	$8.0 < R < 8.5$	88,126	89,251	98.74
	$8.5 < R < 9.0$	79,971	82,425	97.02
	$9.0 < R < 9.5$	63,302	72,014	87.90

Table 1: *Number of stars in the bundle region in each one of our regions with quality cuts, without them and the percentage that has passed them.*

demonstrates its effectiveness in preserving a significant fraction of stars while maintaining a suitable level of completeness across the different regions of interest.

In the final step of our analysis, we present the CMDs for the regions that will be simulated by `dirSFH` and subsequently studied. These CMDs are depicted in Figure 7. Additionally, Table 1 presents the number of stars included in each CMD, the count of *full* data within those CMD limits, and the corresponding percentage of stars that have successfully passed the quality cuts (completeness).

The morphology of the CMDs already gives clues of the underlying stellar populations. The position of the old TO point, practically the same in all regions, indicates that the oldest stars formed at approximately the same time. In addition, we can see that as we approach the Galactic plane, more stars appear in the bright part of the MS. This feature indicates that these layers contain a larger number of young stars. However, these are only qualitative observations, and to study the formation and evolution of the stellar populations in a quantitative way, we must obtain their SFHs.

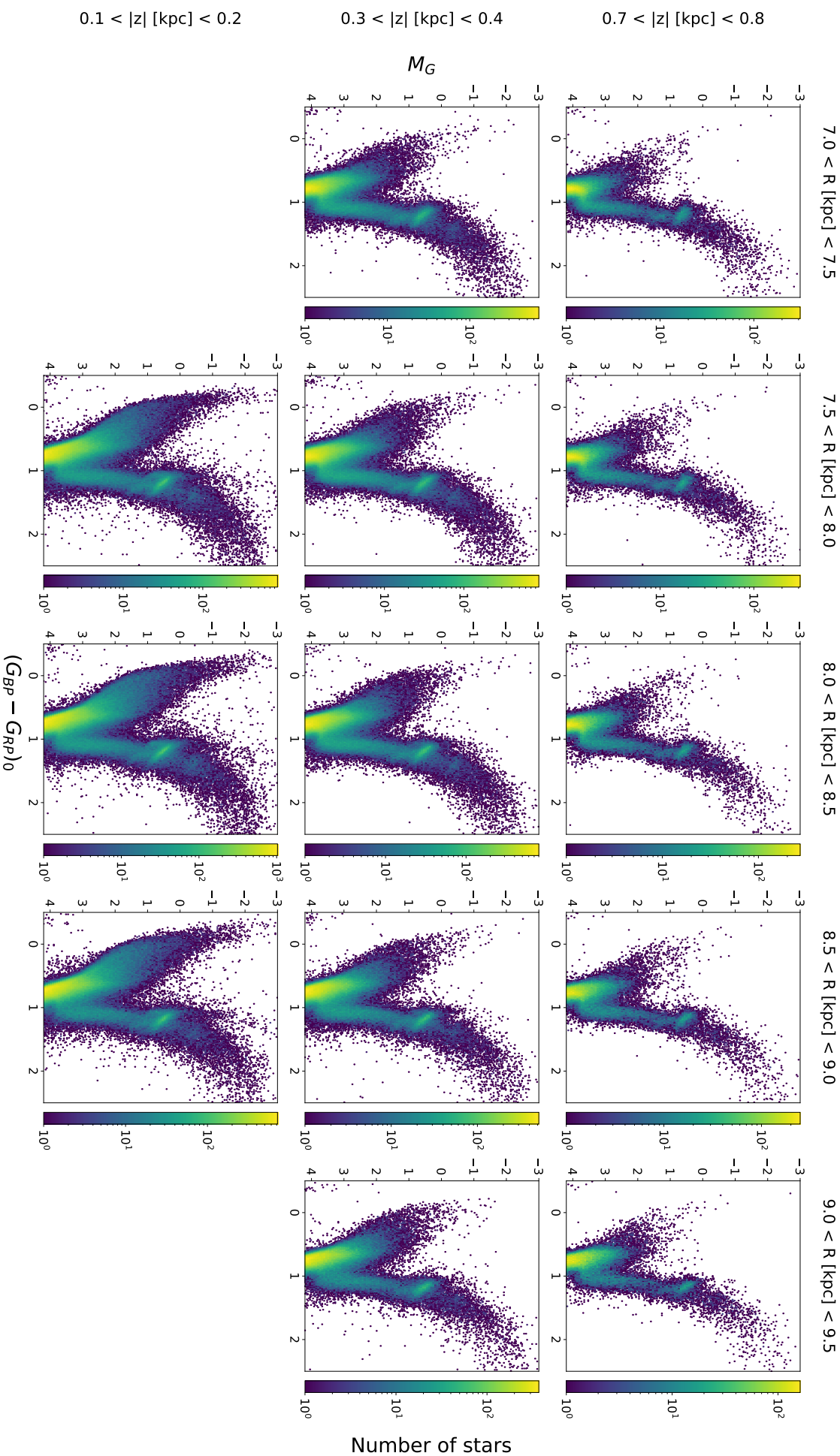

Figure 7: *CMDs of our study regions obtained with Lallement+22 extinction, limited by the bundle region and colour-coded according to the number of stars with logarithmic normalisation. From bottom to top:* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers. From left to right:* $7.0 < R$ [kpc] < 7.5, $7.5 < R$ [kpc] < 8.0, $8.0 < R$ [kpc] < 8.5, $8.5 < R$ [kpc] < 9.0 *and* $9.0 < R$ [kpc] < 9.5 *radius bins.*

4 Results

In this section, we present the results of our study, i.e., the derived SFHs for the thirteen regions that were defined in subsection 3.3.

4.1 Age-metallicity plane

To initiate our analysis, we will conduct a qualitative examination of the resulting SFHs obtained. To facilitate this analysis, we utilise the age-metallicity plane generated by the `dirSFH` program.

Our results revealed that there are some stellar populations in the age-metallicity plane that are present in most of the regions, with more or less intensity. To facilitate the description of these results, we present these populations in Figure 8, which displays the layer $0.1 < |z|$ [kpc] < 0.2 and the radius bin $8.0 < R$ [kpc] < 8.5. This figure illustrates the various stellar populations that will be discussed in the subsequent analysis. The main characteristics of the stellar population defined in Figure 8 are described in Table 2. For a comprehensive view of the age-metallicity planes of all thirteen regions, please refer to Figure 9.

The SFR is the mass of gas turned into stars per unit of time, as a function of time and metallicity, in order to reproduce the stars present in the CMD. Therefore, the SFR values appear smaller in regions with fewer stars (as indicated in Table 1). Furthermore, the colour scale is normalised based on the maximum relative SFR value within each region.

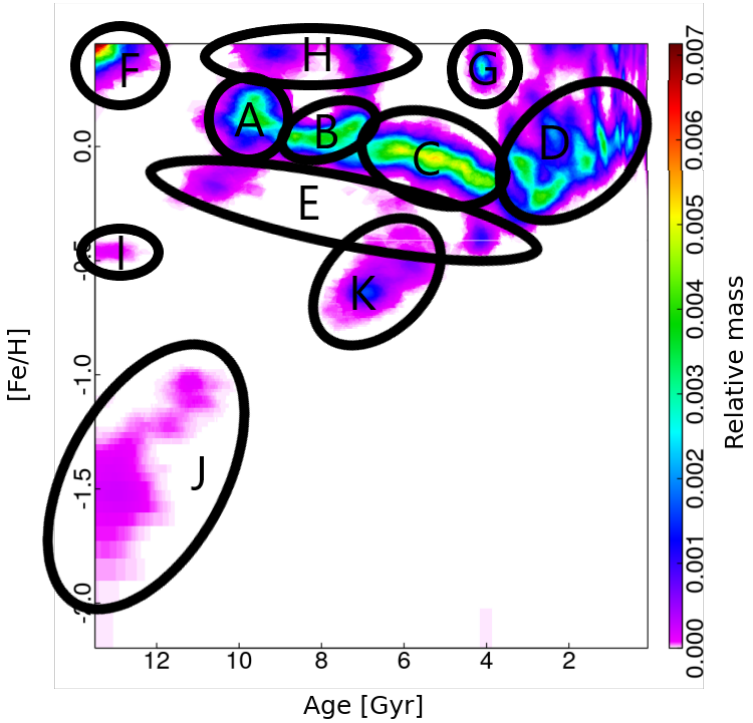

Figure 8: *Age - metallicity plane obtained by* `dirSFH` *for* $0.1 < |z|$ [kpc] < 0.2 *and* $8.0 < R$ [kpc] < 8.5 *region, with the main stellar populations encircled and named alphabetically. The coloured bar indicates the SFR, which is the relative mass turned into stars at a given age and metallicity.*

Population	Type	Age [Gyr]	[Fe/H]
A	Old and solar metallicity	~ 10	~ 0.1
B	Middle-aged and solar metallicity	[9,7]	~ 0
C	Middle-aged and solar metallicity	[7,4]	[-0.2,0]
D	Young and solar metallicity	[4,0]	[-0.2,0.2]
E	Middle-aged and slightly below solar metallicity	[11,4]	[-0.5,0.1]
F	Old and metal-rich	~ 13	~ 0.5
G	Young and metal-rich	~ 4	~ 0.3
H	Middle-aged and metal-rich	[10,6]	~ 0.5
I	Old and metal-poor	[13,11]	~ -0.5
J	Old and metal-poor	[13,10]	[-2,-1]
K	Middle-aged and metal-poor	[8,5]	[-0.75,-0.5]

Table 2: *Characteristics of the stellar populations that appear in our age-metallicity planes, grouped according to their metallicity characteristics.*

First of all, we will focus on the $0.1 < |z|$ [kpc] < 0.2 layer, the one closest to the Galactic plane. We observe a notable similarity in the populations across the three radius bins. The A to C populations are prominent in this layer. They span a range of ages between $10 \gtrsim$ Age [Gyr] $\gtrsim 4$ and metallicities ranging between $-0.25 \lesssim$ [Fe/H] $\lesssim 0.25$. It is worth noting that the metallicities decrease as the age decreases, indicating a trend of decreasing metallicity over time.

We can also observe various younger stellar populations with a similar range of metallicities, the denoted population D. Interestingly, in this case, the metallicities increase over time, suggesting an opposite trend compared to the aforementioned populations.

Populations A, B, C and D are consistent over the three radii, but towards the outside, a decrease in population A in favour of population C is observed. By examining the evolution of the SFR (represented by the range of colours), we can observe that the mass concentration shifts with galactocentric radius R towards younger ages. This indicates that stars are younger as R increases, suggesting a radial age gradient within the studied region.

We also observe the presence of secondary populations. For example, populations F, G and H, which are characterised by very high metallicities. These populations are consistently present in all the evaluated regions. Their presence is intriguing, because they are difficult to explain from what we know of the chemical evolution of the Galaxy. Populations with such high metallicity should not appear in this part of the Galaxy.

Population F, which corresponds to the locus of the age-metallicity plane with the largest SFR in the first layer, exhibits the oldest age and the highest metallicity. That behaviour is not what we expect, as theory tells us that the oldest stars are the least metallic because the medium has not been enriched with heavy elements expelled by supernovae. Population G, on the other hand, represents a younger population with slightly lower metallicity. In addition, population H represents a middle-aged population with relatively high metallicity.

It is possible that these metal-rich populations are result of an error or discrepancy in the SFH determination. The age and metallicity values of the three populations may be influenced

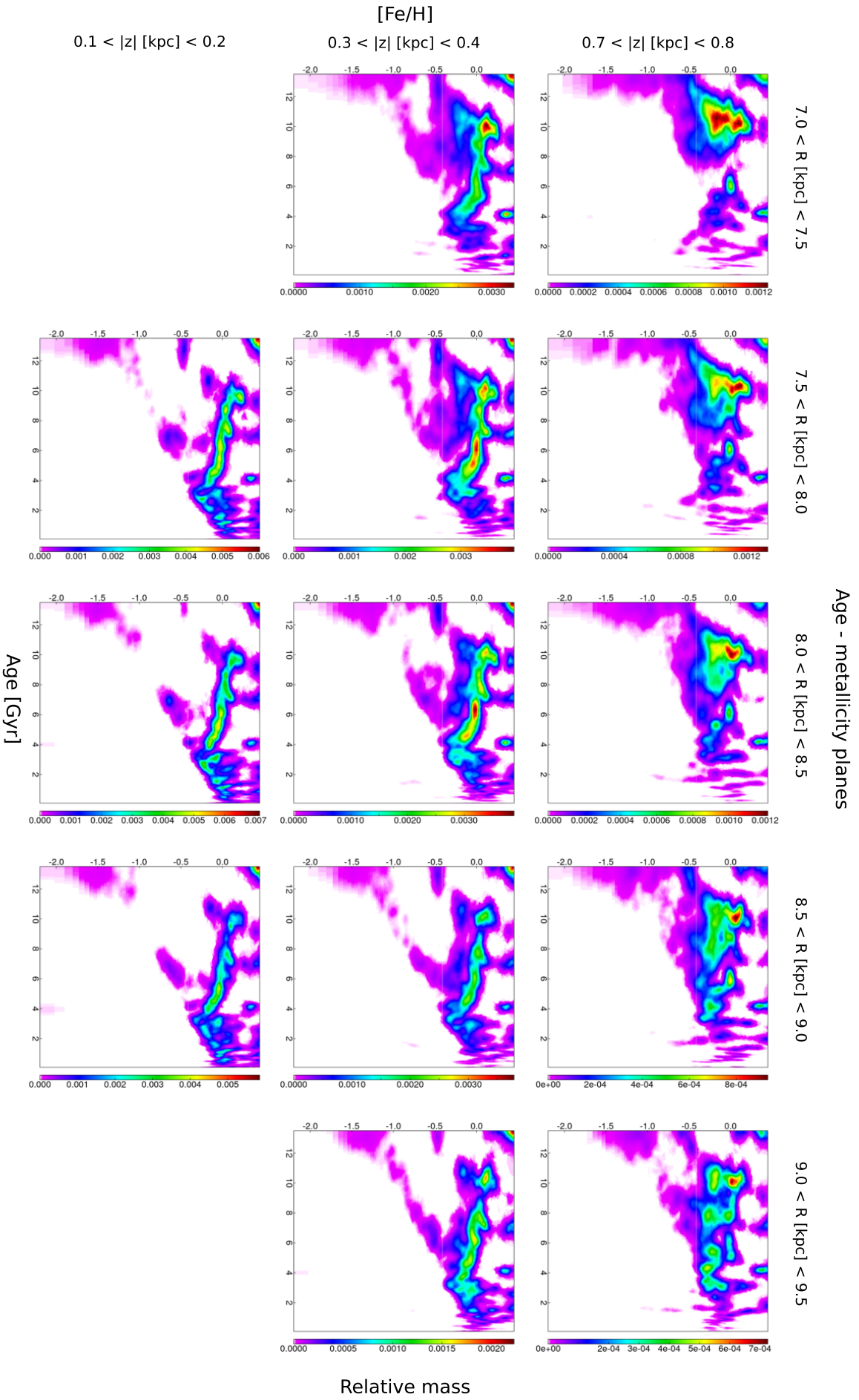

Figure 9: *Age - metallicity planes obtained by* `dirSFH` *of our thirteen regions. The coloured bar indicates the relative mass turned into stars at a given age and metallicity. From bottom to top:* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers. From left to right:* $7.0 < R$ [kpc] < 7.5, $7.5 < R$ [kpc] < 8.0, $8.0 < R$ [kpc] < 8.5, $8.5 < R$ [kpc] < 9.0 *and* $9.0 < R$ [kpc] < 9.5 *radius bins.*

by the presence of stars with underestimated extinction.[6] The fact that these populations show a relatively small decrease in size as R increases within the first layer, may support this hypothesis because extinction affects less the outer disc. Alternatively, these high metallicity stars could be stars formed in more internal regions of the Milky Way, known to be more metal-rich, and "migrated" to outer radius—a phenomenon commonly known as radial migration. We will discuss these hypothesis in more detail in section *5. Discussion*. Further investigation is required to understand the origins and characteristics of populations F to H.

There are several other secondary metal-poor populations that exhibit distinct characteristics. Population I is an old population with a metallicity around [Fe/H] ~ -0.5. It is notable that the intensity of this population decreases with increasing R, eventually disappearing at the outermost radius. Population J represents an old population with the lowest metallicity. Lastly, population K is a middle-aged population with relatively poor metallicity. In the first layer, population K does not appear to vary significantly with R. These secondary populations provide additional insights into the age and metallicity distribution within the studied regions, highlighting the diversity and complexity of stellar populations in the Galactic disc.

We now direct our attention to the second layer, characterised by a height range of $0.3 < |z|$ [kpc] < 0.4. In these age-metallicity planes, we continue to observe a clear continuity of populations A to C with decreasing metallicity, as well as an increasing metallicity population D. Once again, the evolution of the SFR indicates that stars have more recently formed at greater radii, as evidenced by the rightward shift in the concentration of stars. A notable distinction from the previous layer is the presence of a faint secondary population below the main ones, denoted as population E, which exhibits lower metallicity values. Once again, the evolution of the SFR indicates that stars have more recently formed at greater radii, as evidenced by the rightward shift in the concentration of stars. In the first radius bin, population A exhibits the largest SFR value, whereas in the third radius bin, it is population C. Additionally, it is worth noting that at distant radii, the stellar distribution breaks at approximately 9 Gyr, giving rise to a defined stellar population with an age of approximately 10 Gyr and a metallicity of [Fe/H] ~ 0.1, which corresponds to population A.

Populations F to H also appear here and the same considerations apply as for the previous layer. Populations I, J, and K are interconnected, along with population E in the innermost layer, but they gradually separate as R increases. However, population K remains in contact with population E and shows some remnants of association with population J.

The third layer exhibits distinct characteristics compared to the previous two layers. Clear population continuities are no longer observed. In all five R bins, two populations are consistently present: population A+E, characterised by a range of metallicities ($-0.4 \lesssim$ [Fe/H] $\lesssim 0.25$) and representing the highest SFR across all radii, and population C, with an age of ~ 6 Gyr and a near-solar metallicity ([Fe/H] ~ 0). The former population decreases in size as the distance from the Galactic centre increases, giving way to the two populations, with different metallicities, that conform it (population A with [Fe/H] ~ 0 and population E with [Fe/H] ~ -0.25). As R increases, stellar populations B and D emerge, and the concentration of star formation shifts towards younger ages, mirroring the trends observed in the previous layers. It is important to remark that population D is much less important than in the previous layers. In the outermost

[6] See appendix *A. Detected problems of Lallement+18*.

radius, it becomes evident that certain stellar populations were formed at the same age but with two different metallicities, corresponding to populations A to C ([Fe/H] ~ 0) and population E ([Fe/H] ~ -0.25).

The metal-rich populations continue to appear. Population I remains connected to population J, being both consistently connected to population E in all radii and playing a more significant role than in the previous layer. However, population K does not appear in this layer or can be considered as being older.

Finally, let's take a closer look at how the main populations vary in the age-metallicity plane with height z. If we compare the second layer with the first one, we observe that the concentration of star formation in population D is significantly lower, indicating an older population. Also the population E starts to gain prominence. In the third layer, population D almost disappears and population E is clearly visible. It is also worth mentioning that there appears to be a decrease in metallicity with increasing height z.

4.2 Cumulative stellar mass fraction

In this section, we will analyse the SFR. The code calculates the SFR values for each time interval, representing the rate at which stars were formed during that period, with their associated errors.

To plot the cumulative mass fraction, we need to convert the SFR values into mass units by multiplying them by their respective time intervals. Then, we compute a cumulative function by summing the mass of each time interval to all the previous masses. To normalise the cumulative function, we divide it by the maximum value, which corresponds to the total stellar mass. To incorporate the error into the plot, we calculate it by adding the SFR error values to each point of the cumulative function, such as is done in Gallart et al. (2021).

In Figure 10, we present the cumulative mass fraction for the thirteen regions of our study, categorised by different z layers. Within each layer, we observe that the radial bins closer to the Galactic centre exhibit a higher cumulative mass fraction at older ages. This indicates that a significant number of stars in these regions were formed during the early stages of the Galaxy. On the other hand, the radial bins located farther away from the Galactic centre show a slower increase in their cumulative mass fraction, suggesting that stars in these regions have formed more recently and require more time to accumulate, for instance, the 50% of the total mass formed.

When considering the three plots together, we can observe that, for a given radius bin, a higher cumulative mass fraction is achieved in a shorter period of time in the uppermost layers of the Galactic plane, particularly in the $0.7 < |z|$ [kpc] < 0.8 layer. This indicates that in regions with higher z values (further from the Galactic plane), the majority of stars are older, while in regions closer to the Galactic plane, the stars are relatively younger. This suggests a spatial variation in the age distribution of stars within the Galaxy, with older stars being more prevalent at larger vertical distances from the Galactic plane.

In order to gain further insights, we present Figure 11, which illustrates the ages at which the cumulative mass fraction reaches values of 50% and 90% for each of the thirteen regions, plotted as a function of galactocentric radius R. The error bars represent the uncertainties derived from the error curves obtained in the previous figure. Similar to the observations made

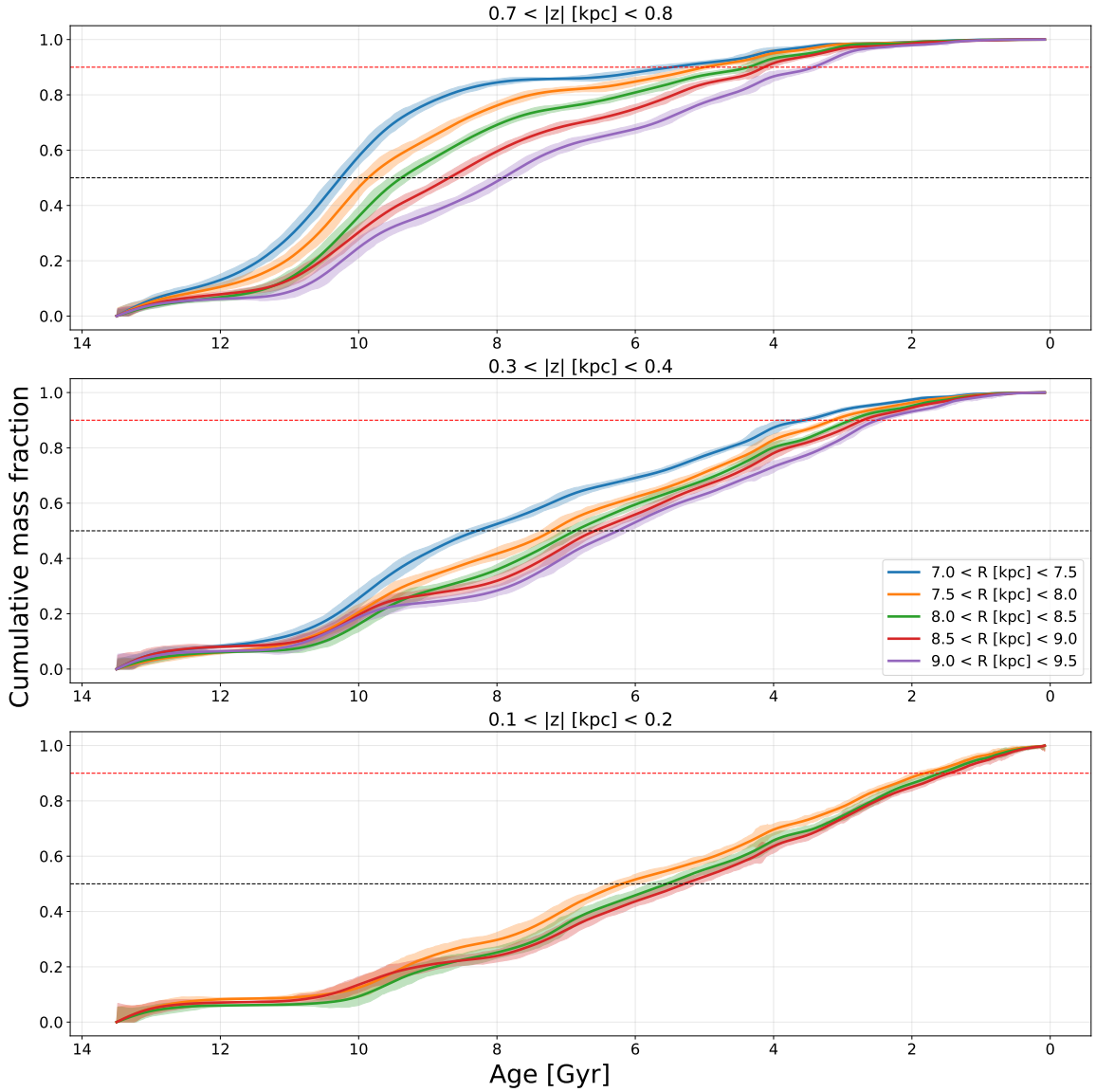

Figure 10: *Cumulative mass fraction as function of time, with its error. Blue, orange, green, red and purple solid lines correspond to* $7.0 < R$ [kpc] < 7.5, $7.5 < R$ [kpc] < 8.0, $8.0 < R$ [kpc] < 8.5, $8.5 < R$ [kpc] < 9.0 *and* $9.0 < R$ [kpc] < 9.5 *radius bins, respectively. The black and red dashed lines mark where the accumulated mass is 50% and 90% of the total, respectively. From bottom to top:* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers.*

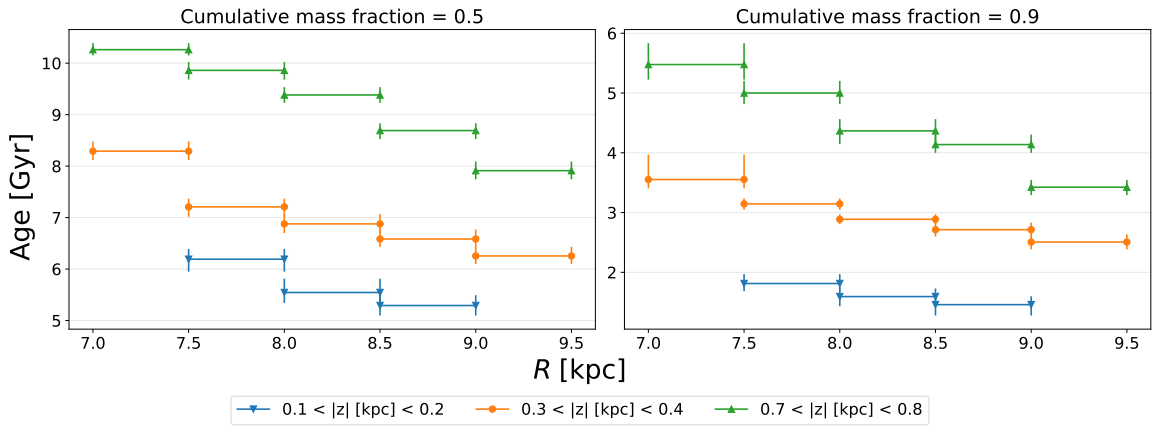

Figure 11: *Age where a given value of the cumulative mass fraction is accomplished as function of galactocentric radius for each of our thirteen regions, with its error. Blue, orange and green solid lines, with triangles pointing down, circles and triangles pointing up respectively, correspond to* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers, respectively. Left: Cumulative mass fraction equals 0.5. Right: Cumulative mass fraction equals 0.9.*

in Figure 10, we can draw the same conclusions from both percentile values. The youthfulness of stars increases as we move towards larger radii from the Galactic centre and shorter distances from the Galactic plane.

4.3 Metallicity distribution function

On the other hand, we will now examine the MDF. We evaluate the distribution function of mass formed per metallicity value. We prefer to evaluate the metallicity distribution instead of its cumulative function because it is more directly comparable with the observational metallicity distribution function derived in literature works. We normalise each region's curve to its maximum, so that we can better compare them. The error curves are calculated by error propagation.

In Figure 12, we present the MDF of the regions divided by z layers based on the MDF. One notable observation is the lack of mass converted into stars for metallicities below [Fe/H] ~ -0.5.

The first characteristic that strikes us is the appearance of three peaks or stellar populations of different metallicity in the three layers, located at [Fe/H] ~ -0.03, [Fe/H] ~ -0.19 and [Fe/H] ~ -0.45 (indicated in the figure with the black dashed lines). In the two layers closest to the galactic plane, we see how the peak of [Fe/H] ~ -0.03 dominates, while the peak of [Fe/H] ~ -0.19 does in the more distant layer. This indicates that in the latter layer the stellar population is relatively less metallic than in the two nearest layers. Between the two layers closest to the plane, we can see that in $0.3 < |z|$ [kpc] < 0.4 the population with [Fe/H] ~ -0.19 appears lighter, indicating that in this layer the fraction of metal-poor stars is higher than in $0.1 < |z|$ [kpc] < 0.2. Moreover, we can clearly notice that in the higher layers, the upward trend of the curve begins at lower metallicities, indicating that some stars formed at lower metallicity, i.e., the MDF is shifted to low metallicity values as the height above the Galactic plane increases.

If we examine a specific layer, there is minimal distinction between the lines representing different radii. However, if we take a closer look, we can see how subtly the curves of the radii

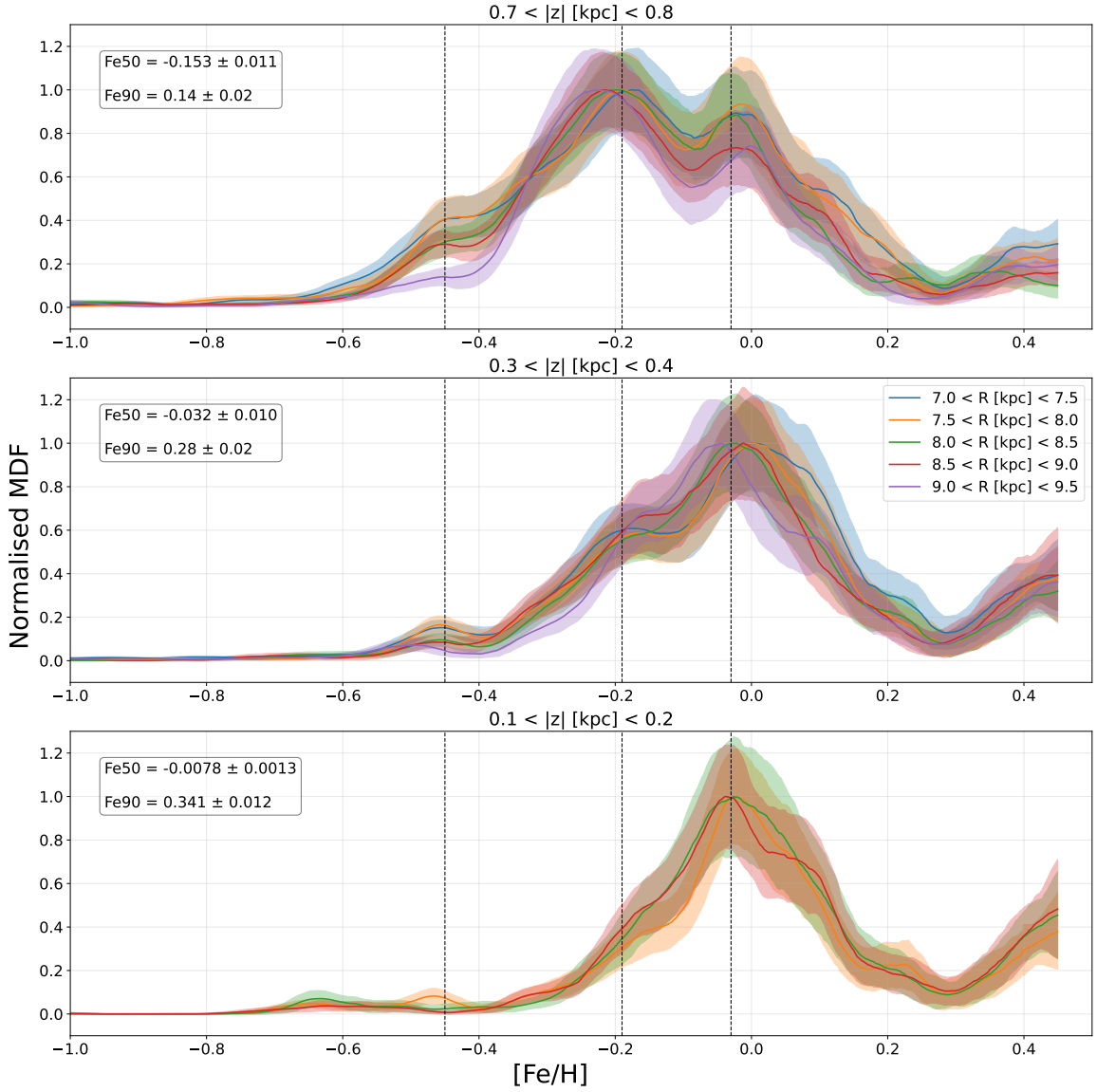

Figure 12: *Normalised MDF, with its error. Metallicity values below* [Fe/H] $= -1$ *have been cut. Blue, orange, green, red and purple solid lines correspond to* $7.0 < R$ [kpc] < 7.5, $7.5 < R$ [kpc] < 8.0, $8.0 < R$ [kpc] < 8.5, $8.5 < R$ [kpc] < 9.0 *and* $9.0 < R$ [kpc] < 9.5 *radius bins, respectively. The black dashed lines mark the metallicity peaks. The Fe50 and Fe90 values are the mean of each layer of the accumulated mass at 50% and 90% of the total, respectively. From bottom to top:* $0.1 < |z|$ [kpc] < 0.2, $0.3 < |z|$ [kpc] < 0.4 *and* $0.7 < |z|$ [kpc] < 0.8 *layers.*

furthest from the galactic centre begin the mass rises at lower metallicities, which would indicate that these regions have a higher proportion of old stars. But as we have already said, this difference is very subtle and within the error range, so we can say that there is practically no radial variation in metallicity.

Figure 12 also shows the average metallicity of the regions in each layer where the cumulative mass fraction reaches a value of 50% (Fe50) and 90% (Fe90). We see how the metallicity obtained for the two layers closest to the plane have quite similar values, while the value of the farthest layer is further away from the previous values in both percentiles. We can clearly observe a correlation between metallicity and the height above the Galactic plane z: the further away from the Galactic plane, the higher the fraction of metal-poor stars. This is in agreement with what we have observed in the MDF curves. The standard deviation of the averages of each layer also tells us that there is not much difference in metallicity with radius. Therefore, if there is any gradient at all, it is very weak.

5 Discussion

After analysing the SFR and the MDF of our thirteen regions of study, it is important to discuss the obtained results and consider the findings published in the literature up to this point.

Firstly, the dominance of younger stellar populations as the galactocentric radius increases, as observed in our study, is consistent with the results obtained in the paper by Queiroz et al. (2020). In their study, they compare the standard relative-to-iron abundance diagrams, being us specifically interested in the [α/Fe] versus [Fe/H] diagram. The α elements are primarily produced by core-collapse type II supernovae, which also produce a small amount of Fe. The stars with high-[α/Fe] content are considered to be old because they are the products of SN-II only. They formed in an environment that had not been enriched yet with iron from low-mass binary stars exploding as type Ia supernovae, thus reducing the [α/Fe] ratio.

In Figure 6 of Queiroz et al. (2020), it is evident that the fraction of stars with high-[α/Fe] diminishes with increasing radius. This indicates a lower population of older stars in the outer regions of the Galaxy. This finding aligns well with our results, where we also noted younger ages when the 50% cumulative mass fraction is accomplished as the radius increases.

The agreement between our results and those presented in Queiroz et al. (2020) provides further support to the idea that younger stars dominate the outer regions of galaxies, while older stars are more prevalent in the inner regions. This consistency strengthens the evidence for a radial dependence of SFH and the role of different stellar populations in shaping galactic structures.

These results can be associated with the inside-out galaxy formation scenario, which explains the sequential development of the bulge and disc components within galaxies. According to this scenario, the bulge is formed in the early stages through the rapid collapse of gas in the innermost regions, while the disc forms later from gas possessing higher angular momentum (Kepner, 1999). In essence, this scenario suggests that stellar formation occurs radially from the inner regions towards the outer regions of galaxies. Notably, similar patterns of galaxy formation have been observed in neighbouring spiral galaxies (Muñoz-Mateos et al., 2007).

Figure 11, depicting the age at which 50% of the mass turned into stars as a function of R, indicate that the inner radii exhibit older stellar populations compared to the outer radii. Similarly, at the 90% cumulative mass fraction, the consistent relationship implies that star formation in the regions closer to the Galactic centre ceased earlier than in the regions farther away. Furthermore, the analysis of the age-metallicity plane (Figure 9) reveals that all regions show the presence of a stellar population born during an episode of star formation around 10 billion years ago. This finding suggests that the time difference in the initiation of star formation between different radii was either relatively short, unless the stellar population formed initially at radii closer to the Galactic centre have subsequently migrated to outer radii.

The metal-rich populations appearing in the age-metallicity planes (Figure 9, F to H populations) can also be explained by the radial migration phenomenon. These stars would have been born in the thin disc at radii close to the Galactic centre or even in the bulge, where the metallicity is very high, and later migrating to outer radii. As they have intermediate-old ages, they have had enough time to travel these distances. In Miglio et al. (2021), the presence of these old super-solar metallicity stars at relative near distances to the Sun is discussed. It is

debated that these stars are too metal-rich to be a result of the SFH of the solar vicinity, and cannot be explained by chemical evolution models. Their conclusion is that the population of intermediate-old, super-solar metallicity stars in very local samples has to be interpreted as clear evidence of radial migration, as these stars do not share the common chemical evolution of most of the nearest thin disc stars. However, it should be noted that the SFRs obtained for our populations have very high values, so it seems that undercorrected extinction stars also affect the population. Therefore, we could say that these populations may be truthful but we must be overestimating them.

If we revisit Figure 6 of Queiroz et al. (2020), but consider the variation of the [α/Fe] ratio with the height above the Galactic plane, we observe an increase in the population with high [α/Fe] as the height increases, accompanied by a decrease in the population with low [α/Fe]. This observation indicates that there is a larger proportion of old stars as we move away from the Galactic plane, aligning with the findings of our analysis. Our results are also in concordance with the commonly accepted view of Galactic astronomy, as it states that the thin disc (our two lowest layers) has more young stars while the thick disc ($0.7 < |z|$ [kpc] < 0.8 layer) has more aged stars.

In terms of the MDF, we found that there is a significant lack of low-metallicity stars (with metallicities lower than [Fe/H] ~ -0.5) across all regions. However, regarding the variation of metallicity with radius, our analysis did reveal non-gradient or smooth gradient.

Gaia Collaboration (2023b) presents a study of this relationship. In the top panel of Figure 13 of their study, metallicity decreases with increasing radius in the range of $7.0 < R$ [kpc] < 9.5 and $-0.5 < z$ [kpc] < 0.5 layer. This trend has been weakly observed in Figure 12, where at the farthest radii there are more stars at lower metallicity values. In the same figure of the same study, we can see that for the $-1.0 < z$ [kpc] < -0.5 and $0.5 < z$ [kpc] < 1.0 layers, there is almost no radial variation of metallicity in our study range. This is in line with our results.

Furthermore, the top panel of Figure 14 in the same study reveals the relationship between metallicity and height above the Galactic plane. Remarkably, this finding aligns with our results, as it demonstrates a consistent decrease in metallicity as we move further away from the Galactic plane.

In conclusion, our analysis of the SFR and MDF provides valuable insights into the SFH in our studied regions. It is important to continue exploring these relationships and conducting further investigations to deepen our understanding of the formation and evolution of galaxies.

6 Conclusions

In this project, we have conducted a comprehensive analysis of the SFH in thirteen nearby regions of the Milky Way. Specifically, our study focuses on regions located within a radial distance of two kiloparsecs from the Sun, utilising the data provided by the *Gaia* EDR3 catalogue. We have carried out a radial and height analysis.

Our primary objective in this project was to determine the SFH of the studied regions using a specific code that had been developed by the hosting group for this purpose. This code employs the CMD fitting technique, which involves comparing the observed CMDs of stars with simulated stellar populations generated from stellar evolution models. By comparing these distributions, we can infer the temporal and chemical distribution of stars formed within the regions under investigation.

Given the large volume of our data, it is crucial to apply extinction and reddening corrections. To accomplish this, we have access to two extinction maps: Green and Lallement+22. After conducting a preliminary analysis, we have determined that the Lallement+22 extinction map is a bit better suited for correcting our data. Therefore, we have chosen to utilise this extinction map for the purpose of correcting our dataset.

The findings of this study reveal that the majority of stellar populations formed through time are present across all regions under investigation. These populations exhibit a significant concentration within an age range spanning from 10 Gyr to 4 Gyr and possess metallicity levels spanning a mildly narrow metallicity range centred around the solar value ([Fe/H] ~ 0). Notably, in the two studied layers closest to the Galactic plane, a noticeable trend emerges where older populations exhibit a declining metallicity pattern over time, while younger populations display an increase in metallicity over time. However, in the uppermost layer, these trends become less distinct, with the predominant population consisting of older stars and displaying a second comparatively lower metallicity sequence.

The old-intermediate high-metallicity stars observed in the age-metallicity plane could be explained by the phenomenon of radial migration, i.e. born at radii closer to the Galactic centre and displaced from there to reach their current location. This must be explained in this way because these stars do not share the chemical evolution of the other stars around them. However, their obtained SFR values are possibly overestimated because of failures related to undercorrection of extinction.

We have observed a clear relationship between the galactocentric radius and the age distribution of stars in our study. Specifically, we have found that as the radius increases, the proportion of young stars also increases. This finding is in line with the patterns in the [α/Fe] versus [Fe/H] observational diagrams pointed out in the literature and supports the inside-out scenario of galaxy formation. According to this scenario, stellar formation takes place in a radial fashion, starting from the inner regions of the galaxy and progressing outward towards the outer regions.

We have also discovered a significant correlation between the age of stars and their height above the Galactic plane. Our analysis has revealed that as we move further away from the plane, the stellar population tends to be older. This finding aligns with the patterns observed

in the [α/Fe] versus [Fe/H] diagrams from literature, further reinforcing the connection between stellar age and height above the Galactic plane.

In terms of metallicity, remarkable radial gradients were not found, so we can consider they to be plain at least in the radii where we carried out our analysis. Nonetheless, we did observe a clear relationship between metallicity and the distance from the Galactic plane. Specifically, as the distance from the plane increases, the metallicity tends to decrease. This finding is consistent with observational data and provides further evidence of the connection between metallicity and height above the Galactic plane.

However, it is important to acknowledge that our results are not without limitations, as they are subject to various sources of error. While we made efforts to correct for extinction and reddening using extinction maps, it is worth noting that these maps have some uncertainties. Each extinction map employs its own methodology for calculating extinction, leading to discrepancies in the extinction values obtained. Furthermore, we must consider the inherent uncertainties associated with the SFH derivation. Each simulation run may yield slightly different results, contributing to the overall variability in our findings.

Given that the field of study is continually evolving, the code is still being updated and new data releases are anticipated from the *Gaia* mission, several avenues for future research can build upon and enhance our current findings. Here are some suggestions for future work:

- To extend the analysis to other regions of the Milky Way, such as the stellar halo, to obtain a more comprehensive understanding of the SFH across different Galactic components.
- To explore the use of alternative extinction maps, like StarHorse extinction estimations (Anders et al., 2022), to investigate potential differences and assess their impact on the derived SFH results. Comparing results obtained using different extinction maps can provide valuable insights into the robustness and reliability of the findings.
- To enhance the capabilities of the `dirSFH` code to incorporate a real estimation of the total mass converted into stars. This improvement would allow for a quantitative comparison of the age-metallicity diagrams.
- To investigate the nature of the metal-rich populations F, G and H, to determine their authenticity and assess whether they represent a real stellar population or are a result of observational biases or systematic effects.
- To research the impact of radial migration in the configuration of the current Milky Way morphology and the distribution of the stellar populations across the Galaxy.

By pursuing these future research directions, we can further refine our understanding of the SFH in the Milky Way and gain deeper insights into the formation and evolution of stellar populations in different Galactic environments.

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). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss4.org.

Personalmente, quería agradecer en primer lugar a mis tutoras, Emma y Carme. Me ayudaron desde el primer momento con cualquier duda, y han estado siempre a mi disposición, haciendo mucho más ameno el tiempo empleado en el trabajo presentado. También quería agradecer a Tomás, que ayudó en la resolución de varias dudas y siempre podíamos contar con su certera opinión, a Carlos Quintero y Carlos Allende, por dejarme usar sus ordenadores del IAC para la ejecución del código, y a Carlos Luque, que ayudó con problemas que teníamos al ejecutar el código remotamente.

Durante los dos años cursando el máster, he tenido la suerte de recibir docencia de profesores que, aparte de transmitir su pasión por la astrofísica, han sabido comprender la situación de estrés en la que mis compañeros y yo nos encontrábamos y nos han dado ánimos en todo momento. Entre ellos destaco a Ismael, Sébastien, Artemio, Sergio, Pepa, Ramón, Juan Antonio, José Alberto, Sara, Jaume y María Jesús, entre otros.

No me puedo olvidar de mis compañeros, que habiendo estado dos años en mi vida, ya van a formar siempre parte de ella. Las horas encerrados en el CCA y las excursiones a los observatorios son recuerdos que no se me van a olvidar fácilmente. Quiero agradecer a Samu, María, Pau, Elena, Amaia, Javi, Andy, José, Patri, Sara, Carlos y Sergio. Mención especial para Paula, *my Sosa canary queen*, que desde el primer momento y a pesar de ser godito me dejaste formar parte de tu vida, vida que últimamente pasamos casi siempre juntos por las largas jornadas de trabajo, casi siempre contándonos las penurias del TFM (aunque no se me ocurre mejor compañía para aliviarlas); Alberto, la persona con la que más rápido cogí confianza cuando llegué a la isla y compañero de piso durante este curso a pesar de conocer de antemano mis excentricidades; Mario, también compañero de piso y de movidas varias, que me da ánimos cuando los tengo bajos y me ha cocinado la friolera de dos creps durante todo el curso; y Yéssica, *a.k.a.* Yessikona, que desde el grado forma parte de mi vida, me invitó a la isla cuando estaba realizando el SICUE y me acogió en su casa cuando todavía no tenía: gracias a ti me enamoré de la isla y si no fuese por ti seguramente no estaría aquí.

Volia donar les gràcies a la meua família, però sobretot a la meua cosina Hermi, per sempre estar al meu costat i haver-me ajudat en qualsevol problema de lingüística, i als meus pares, que sense la seua ajuda no estaria escrivint aquest text ni en Tenerife. També a tota la gent de Crevillent i Alacant: la terreta sempre es troba a faltar.

Por último, agradecer a la isla, Tenerife, desde Roque Bermejo hasta Los Gigantes: gracias por ser como eres.

References

Abdurro'uf, K. Accetta, C. Aerts, V. Silva Aguirre, R. Ahumada, et al. The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data. *The Astrophysical Journal Supplement Series*, 259(2):35, Apr. 2022. doi: 10.3847/1538-4365/ac4414.

F. Anders, A. Khalatyan, A. B. A. Queiroz, C. Chiappini, J. Ardèvol, et al. Photo-astrometric distances, extinctions, and astrophysical parameters for Gaia EDR3 stars brighter than G = 18.5. *Astronomy & Astrophysics*, 658:A91, Feb. 2022. doi: 10.1051/0004-6361/202142369.

E. J. Bernard, M. Schultheis, P. Di Matteo, V. Hill, M. Haywood, and A. Calamida. Star formation history of the Galactic bulge from deep HST imaging of low reddening windows. *Monthly Notices of the Royal Astronomical Society*, 477(3):3507–3519, July 2018. doi: 10.1093/mnras/sty902.

M. R. Blanton, M. A. Bershady, B. Abolfathi, F. D. Albareti, C. Allende Prieto, et al. Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. *The Astronomical Journal*, 154(1):28, July 2017. doi: 10.3847/1538-3881/aa7567.

A. Bressan, P. Marigo, L. Girardi, B. Salasnich, C. Dal Cero, S. Rubele, and A. Nanni. PARSEC: stellar tracks and isochrones with the PAdova and TRieste Stellar Evolution Code. *Monthly Notices of the Royal Astronomical Society*, 427(1):127–145, Nov. 2012. doi: 10.1111/j.1365-2966.2012.21948.x.

A. Cimatti, F. Fraternali, and C. Nipoti. *Introduction to Galaxy Formation and Evolution: From Primordial Gas to Present-Day Galaxies*. Cambridge University Press, 2019.

A. Fraknoi, D. Morrison, and S. C. Wolff. *Astronomy (OpenStax)*. OpenStax, 2018.

Gaia Collaboration. The Gaia mission. *Astronomy & Astrophysics*, 595:A1, Nov. 2016. doi: 10.1051/0004-6361/201629272.

Gaia Collaboration. Gaia Early Data Release 3. Summary of the contents and survey properties. *Astronomy & Astrophysics*, 649:A1, May 2021a. doi: 10.1051/0004-6361/202039657.

Gaia Collaboration. Gaia Early Data Release 3. The Gaia Catalogue of Nearby Stars. *Astronomy & Astrophysics*, 649:A6, May 2021b. doi: 10.1051/0004-6361/202039498.

Gaia Collaboration. Gaia Data Release 3. Summary of the content and survey properties. *Astronomy & Astrophysics*, 674:A1, June 2023a. doi: 10.1051/0004-6361/202243940.

Gaia Collaboration. Gaia Data Release 3. Chemical cartography of the Milky Way. *Astronomy & Astrophysics*, 674:A38, June 2023b. doi: 10.1051/0004-6361/202243511.

C. Gallart, E. J. Bernard, C. B. Brook, T. Ruiz-Lara, S. Cassisi, V. Hill, and M. Monelli. Uncovering the birth of the Milky Way through accurate stellar ages with Gaia. *Nature Astronomy*, 3:932–939, July 2019. doi: 10.1038/s41550-019-0829-5.

C. Gallart, M. Monelli, T. Ruiz-Lara, A. Calamida, S. Cassisi, et al. The Star Formation History of Eridanus II: On the Role of Supernova Feedback in the Quenching of Ultrafaint Dwarf Galaxies. *The Astrophysical Journal*, 909(2):192, Mar. 2021. doi: 10.3847/1538-4357/abddbe.

GRAVITY Collaboration. Detection of the Schwarzschild precession in the orbit of the star S2 near the Galactic centre massive black hole. *Astronomy & Astrophysics*, 636:L5, Apr. 2020. doi: 10.1051/0004-6361/202037813.

G. M. Green, E. Schlafly, C. Zucker, J. S. Speagle, and D. Finkbeiner. A 3D Dust Map Based on Gaia, Pan-STARRS 1, and 2MASS. *The Astrophysical Journal*, 887(1):93, Dec. 2019. doi: 10.3847/1538-4357/ab5362.

S. L. Hidalgo, A. Aparicio, and C. Gallart. Recovering the ages and metallicities of stars of a complex stellar population system. In E. E. Mamajek, D. R. Soderblom, and R. F. G. Wyse, editors, *The Ages of Stars*, volume 258, pages 245–252, June 2009. doi: 10.1017/S1743921309031895.

S. L. Hidalgo, A. Pietrinferni, S. Cassisi, M. Salaris, A. Mucciarelli, et al. The Updated BaSTI Stellar Evolution Models and Isochrones. I. Solar-scaled Calculations. *The Astrophysical Journal*, 856(2):125, Apr. 2018. doi: 10.3847/1538-4357/aab158.

H. Jönsson, J. A. Holtzman, C. Allende Prieto, K. Cunha, D. A. García-Hernández, et al. APOGEE Data and Spectral Analysis from SDSS Data Release 16: Seven Years of Observations Including First Results from APOGEE-South. *The Astronomical Journal*, 160(3):120, Sept. 2020. doi: 10.3847/1538-3881/aba592.

J. V. Kepner. Inside-out Galaxy Formation. *The Astrophysical Journal*, 520(1):59–66, July 1999. doi: 10.1086/307419.

P. Kroupa. On the variation of the initial mass function. *Monthly Notices of the Royal Astronomical Society*, 322(2):231–246, Apr. 2001. doi: 10.1046/j.1365-8711.2001.04022.x.

R. Lallement, L. Capitanio, L. Ruiz-Dern, C. Danielski, C. Babusiaux, et al. Three-dimensional maps of interstellar dust in the Local Arm: using Gaia, 2MASS, and APOGEE-DR14. *Astronomy & Astrophysics*, 616:A132, Aug. 2018. doi: 10.1051/0004-6361/201832832.

R. Lallement, J. L. Vergely, C. Babusiaux, and N. L. J. Cox. Updated Gaia-2MASS 3D maps of Galactic interstellar dust. *Astronomy & Astrophysics*, 661:A147, May 2022. doi: 10.1051/0004-6361/202142846.

L. Lindegren, S. A. Klioner, J. Hernández, A. Bombrun, M. Ramos-Lerate, et al. Gaia Early Data Release 3. The astrometric solution. *Astronomy & Astrophysics*, 649:A2, May 2021. doi: 10.1051/0004-6361/202039709.

S. R. Majewski, R. P. Schiavon, P. M. Frinchaboy, C. Allende Prieto, R. Barkhouser, et al. The Apache Point Observatory Galactic Evolution Experiment (APOGEE). *The Astronomical Journal*, 154(3):94, Sept. 2017. doi: 10.3847/1538-3881/aa784d.

A. Miglio, C. Chiappini, J. T. Mackereth, G. R. Davies, K. Brogaard, et al. Age dissection of the Milky Way discs: Red giants in the Kepler field. *Astronomy & Astrophysics*, 645:A85, Jan. 2021. doi: 10.1051/0004-6361/202038307.

J. C. Muñoz-Mateos, A. Gil de Paz, S. Boissier, J. Zamorano, T. Jarrett, et al. Specific Star Formation Rate Profiles in Nearby Spiral Galaxies: Quantifying the Inside-Out Formation of Disks. *The Astrophysical Journal*, 658(2):1006–1026, Apr. 2007. doi: 10.1086/511812.

A. B. A. Queiroz, F. Anders, C. Chiappini, A. Khalatyan, B. X. Santiago, et al. From the bulge to the outer disc: StarHorse stellar parameters, distances, and extinctions for stars in APOGEE DR16 and other spectroscopic surveys. *Astronomy & Astrophysics*, 638:A76, June 2020. doi: 10.1051/0004-6361/201937364.

B. C. Reed. The Sun's Displacement from the Galactic Plane from Spectroscopic Parallaxes of 2500 OB Stars. *Journal of The Royal Astronomical Society of Canada*, 100:146, Aug. 2006.

T. Ruiz-Lara, C. Gallart, M. Monelli, T. K. Fritz, G. Battaglia, et al. Dissecting the stellar content of Leo I: a dwarf irregular caught in transition. *Monthly Notices of the Royal Astronomical Society*, 501(3):3962–3980, Mar. 2021. doi: 10.1093/mnras/staa3871.

T. Ruiz-Lara, A. Helmi, C. Gallart, F. Surot, and S. Cassisi. Unveiling the past evolution of the progenitor of the Helmi streams. *Astronomy & Astrophysics*, 668:L10, Dec. 2022a. doi: 10.1051/0004-6361/202244127.

T. Ruiz-Lara, T. Matsuno, S. S. Lövdal, A. Helmi, E. Dodd, and H. H. Koppelman. Substructure in the stellar halo near the Sun. II. Characterisation of independent structures. *Astronomy & Astrophysics*, 665:A58, Sept. 2022b. doi: 10.1051/0004-6361/202243061.

Appendices

A Detected problems of Lallement+18

At the start of this Master's final project, only the Lallement+18 and Green extinction maps were implemented in the software. Upon examining these two extinction maps an initial observation suggested that Lallement+18 may be more favourable because it covers the whole sky, contrary to Green that lacks a quadrant (Figure 3). However, when we conducted the `dirSFH` analysis using Lallement+18 data, we noticed large differences between the observed CMD and the synthetic solution CMD, as reflected in the residuals plot, pointing to a bad fit (Figure 13).

The presence of red data points in the residuals plot points to a significant discrepancy between the observed and calculated CMDs, with the red regions indicating more observed stars. Furthermore, upon examining the MDF plot, we can observe the emergence of a tail around a metallicity of [Fe/H] ~ 0.5, indicating a substantial number of high-metallicity stars. This phenomenon can be attributed to a considerable population of stars with significant reddening that has not been adequately corrected. Heavily reddened stars manifest themselves as points located in the red region of the diagram (indicating a larger colour), but they ought to be present across the whole CMD. As more metallic stars are redder, these undercorrected, heavily reddened stars mimic a high metallicity population. The large discrepancy in the residuals plot towards the red colour is a consequence of the absence of such heavily reddened stars in the *mother diagram*, and also no extremely high metallicity stars that could eventually mimic them.

In our attempts to address the issue of the high metallicity tail, we explored the possibility of implementing more stringent extinction cuts. Despite attempting to lower the extinction limit A_G from 0.5 to 0.3, a considerable number of red points persisted in the residuals plots, and the persistent presence of the high metallicity tail persisted even though less prominent (Figure 14).

As a last resort, we computed the SFH exclusively for stars with $A_G < 0.1$ (Figure 15). In this particular case, the sigma-difference plot showed no red points, and the high metallicity tail almost disappeared. However, it is important to note that this extinction restriction significantly reduced our sample size, as we were only able to study 1% of the stars present in the *clean* data within that region (10,951 stars). Consequently, this particular case must be disregarded since

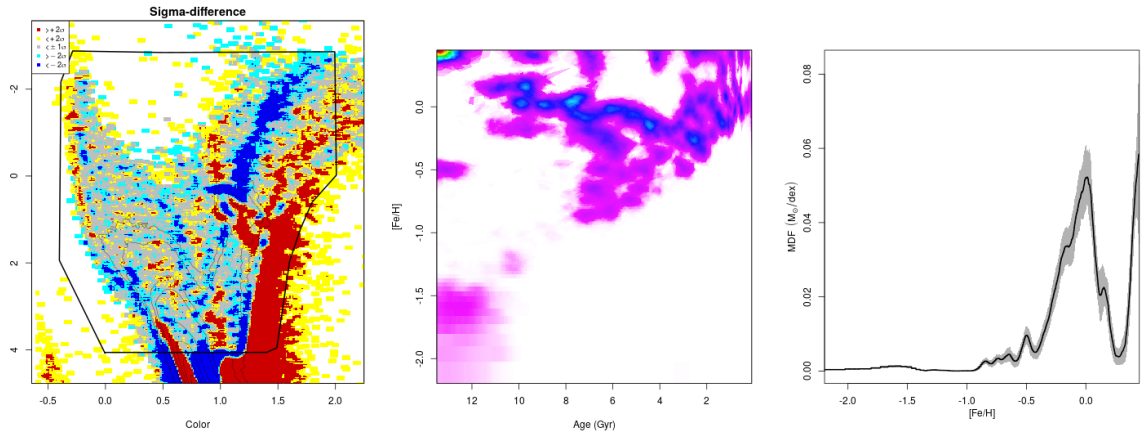

Figure 13: *Residuals, SFH in the age-metallicity plane and MDF obtained from* `dirSFH` *for the Lallement+18 data in* $0.1 < |z|$ [kpc] < 0.2 *layer with* $7.5 < R$ [kpc] < 8.0.

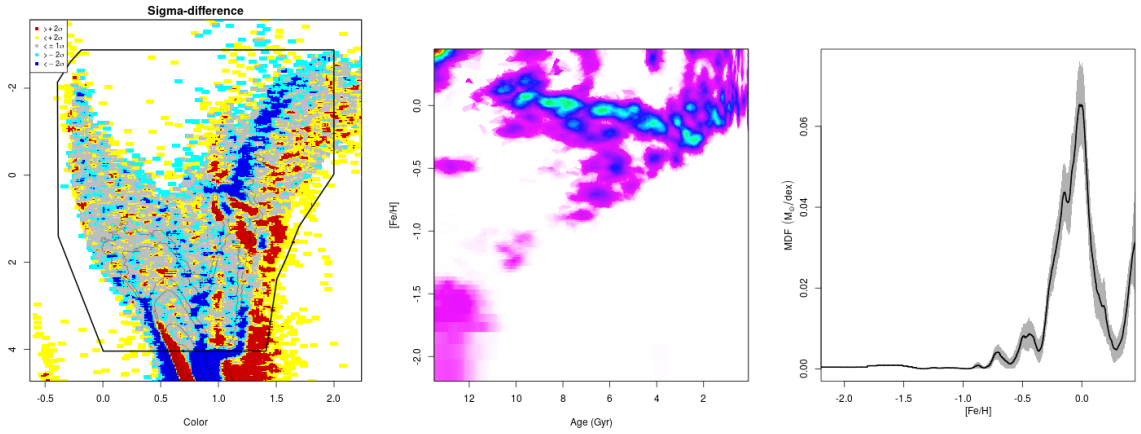

Figure 14: *Residuals, SFH in the age-metallicity plane and MDF obtained from* `dirSFH` *for the Lallement+18 data in* $0.1 < |z|$ [kpc] < 0.2 *layer with* $7.5 < R$ [kpc] < 8.0 *and* $A_G < 0.3$.

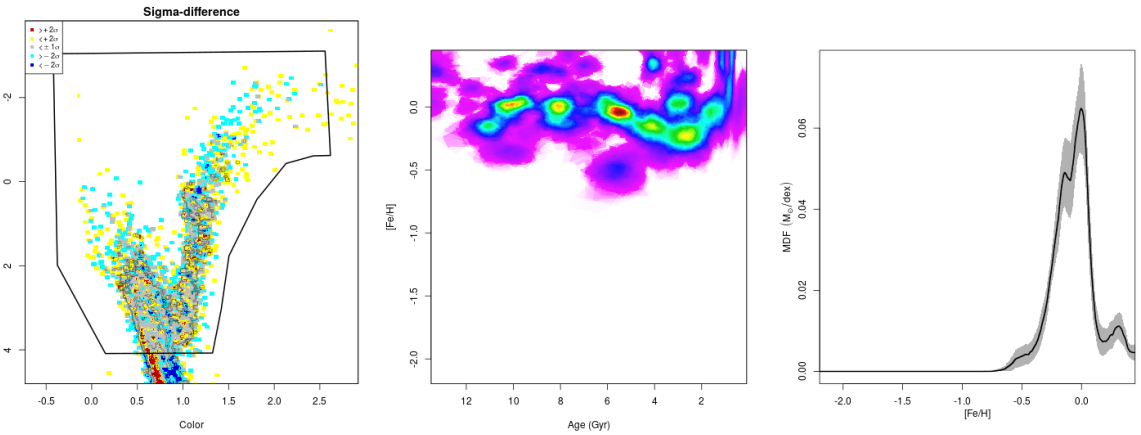

Figure 15: *Residuals, SFH in the age-metallicity plane and MDF obtained from* `dirSFH` *for the Lallement+18 data in* $0.1 < |z|$ [kpc] < 0.2 *layer with* $7.5 < R$ [kpc] < 8.0 *and* $A_G < 0.1$.

such a small percentage does not adequately represent the stellar population within the region of interest. Therefore, we concluded that implementing extinction cuts to eliminate the high metallicity tail is not a viable solution in this context.

In summary, our analysis led us to the conclusion that the Lallement+18 extinction map underestimates the amount of extinction for a significant number of stars. The presence of a high reddening tail in the CMD residuals plot and the deviation from the expected metallicity distribution in the MDF plot indicated that a considerable number of highly reddened stars were not accurately corrected in this extinction map. This discrepancy suggests that Lallement+18 may not provide reliable extinction estimates for our study, and prompted the team hosting this Master's final project to implement the newer Lallement+22 extinction map.

B Study of shifts

Throughout our project, we encountered a recurring challenge regarding the determination of the optimal colour and magnitude shift between the observed and the *mother CMD* to use in the `dirSFH` analysis. This shift is expected to be very small and to arise from slight mismatches between the *Gaia* photometric system and that of the models, after transforming the theoretical

data to the *Gaia* observational plane. However, this shift parameter plays a crucial role in aligning the observed CMD with the computed one and has a noticeable effect in the derived SFH. Finding the appropriate shift value was a topic of ongoing consideration and evaluation for us.

Two ways of calculating the shifts were implemented in *getSM*: the *shifts85* and the *Gaussian*. Initially, we believed that utilising the *shifts85* value calculated by *getSM* for the region within a sphere of 200 pc around the Sun would be the most suitable choice. This shift value corresponds to the 85-percentile of best solutions from the evaluation points computed by the *getSM* algorithm. We chose this particular 200 pc sample because it is minimally affected by reddening, as it is located in close proximity to us and the accumulated dust in any line of sight is small, and thus, it is the amount of extinction and reddening. However, upon executing the `dirSFH` analysis, we observed a significant number of stars exhibiting high metallicity, which became apparent in the MDF plot. This outcome mirrored the effect observed when using Lallement+18 data, indicating a discrepancy between the observed and calculated CMDs.

In an attempt to address this issue, we conducted an analysis of the SFH within the 200 pc sphere using the Green extinction map, which, based in our test performance, lead to a CMD less affected by highly reddened stars. We employed three different shifts calculated by *getSM* specifically for this sample:

- The *shifts85* value obtained using Lallement+18 data, which is the initial shift used.
- The *shifts85* value obtained using Green data.
- The *Gaussian* value obtained using Green data.

The *Gaussian* shift is determined by fitting a Gaussian to an indicator of the goodness of fit of the many SFH calculated by *getSM* from shifted observed CMDs and finding the most likely position of the shift. This shift is thus computed by "filling in the gaps" and taking into account the goodness of fit for the SFH at each evaluation point.

Upon analysing the results, we found that the *Gaussian* shift for Green data exhibited the least prominent high metallicity tail in the MDF plot. Consequently, it provided us with improved outcomes and offered insights into the most appropriate shift to utilise in our analysis.

To validate our earlier findings, we performed additional SFH calculations for two different regions: one closer to the Galactic centre ($7.5 < R$ [kpc] < 8.0) and one farther away ($8.5 < R$ [kpc] < 9.0). These calculations focused on the closest layer to the Galactic plane ($0.1 < |z|$ [kpc] < 0.2). We utilised both extinction maps (Green and Lallement+18, although the latter is known to be problematic) and employed the *Gaussian* and *shifts85* shifts, which were calculated for each region and extinction map using the *getSM* approach.

Upon analysing the results of this study, we arrived at a consistent conclusion: the *Gaussian* shift consistently produced superior outcomes compared to the *shifts85* shift, exhibiting a lower high metallicity tail in the resulting MDF. This further confirms that the *Gaussian* shift is a more suitable choice for our analysis.

In the final stage of our analysis, we used the *Gaussian* shift for the Green 200 pc sphere. This shift was then employed to calculate the SFH for the three regions of the first layer. The results were highly promising, showing a significant reduction in the high metallicity tail. This

confirms the effectiveness of the *Gaussian* shift in producing reliable SFH estimations for our study.

In summary, after careful evaluation and comparison of different shift values, we determined that the *Gaussian* shift is the most suitable choice for our analysis. As this shift arises from differences between photometric systems, not should be dependent on specific characteristics of the stellar population under study. However, when calculating the shift for the same region with different extinction maps, the values were not equal. This is because the extinction values are not the same in each map. Therefore, finally we decided to derive it from a region with minimal extinction. We obtained the *Gaussian* shift from the *Gaia* Catalogue of Nearby Stars (GCNS), which is a well-characterised dataset comprising 331,312 stars within a 100 pc radius sphere centred on the Sun (Gaia Collaboration, 2021b). The GCNS provides SFH solutions that closely match the spectroscopic MDF, making it a reliable source for obtaining the shift value.