MASARYK UNIVERSITY
FACULTY OF SCIENCE

Ph.D. Dissertation

BRNO 2022 **MARTIN PIECKA**

MASARYKOVA
UNIVERZITA

MASARYK UNIVERSITY
FACULTY OF SCIENCE
DEPARTMENT OF THEORETICAL PHYSICS AND ASTROPHYSICS

Studying Open Clusters and DIBs During the Gaia Era

Ph.D. Dissertation

Martin Piecka

Supervisor: doc. Ernst Paunzen, Dr.rer.nat
Department of Theoretical Physics and Astrophysics

Brno 2022

Bibliographic Entry

Author:	Martin Piecka Faculty of Science, Masaryk University Department of Theoretical Physics and Astrophysics
Title of Thesis:	Studying Open Clusters and DIBs During the Gaia Era
Degree Programme:	Physics
Field of Study:	Astrophysics
Supervisor:	doc. Ernst Paunzen, Dr.rer.nat Faculty of Science, Masaryk University Department of Theoretical Physics and Astrophysics
Academic Year:	2021/2022
Number of Pages:	167
Keywords:	Gaia; Astrometry; Open clusters; Galactic kinematics; Diffuse interstellar bands

Bibliografický záznam

Autor:	Martin Piecka Přírodovědecká fakulta, Masarykova univerzita Ústav teoretické fyziky a astrofyziky

Název práce:	Studying Open Clusters and DIBs During the Gaia Era
Studijní program:	Fyzika
Specializace:	Astrofyzika

Vedoucí práce:	doc. Ernst Paunzen, Dr.rer.nat Přírodovědecká fakulta, Masarykova univerzita Ústav teoretické fyziky a astrofyziky

Akademický rok:	2021/2022
Počet stran:	167
Klíčová slova:	Gaia; Astrometrie; Otevřené hvězdokupy; Kinematika Galaxie; Difúzní mezihvězdní pásy

Abstract

Open clusters represent one of the most important objects for studying stars. They can be used to study multiple aspects of star-formation, stellar evolution, kinematics and dynamics. With the information provided by Gaia in its multiple data releases, we have the ability to study clusters with much higher precision than ever before – this affects not only the physics probed by the astrometric quantities but also the derived membership probabilities.

Several problems tackled in the literature are still poorly resolved. For example, the well-known issues with the transformation of parallaxes to distances are still often ignored. We provide an analysis of the problem and propose that the inverted mean parallaxes are still reasonable measures of distances up to a few kpc. We argue that the elongation of clusters along the line of sight will remain almost unchanged regardless of the chosen parallax-distance transformation method.

Another issue can be encountered when analysing open clusters from the recently published catalogues based on the UPMASK method. None of the popular works have taken into account that clusters may significantly overlap in the phase space. We point out that this oversight results in a bad membership classification in such cases of clusters. This affects all the information we can get for these clusters – mean parallaxes, mean proper motion, ages and metallicities.

In one of our latest studies, we analysed the ability of open clusters and diffuse interstellar bands to probe the nearby Galactic spiral arms. We prepared a relatively simple approach that yields results comparable with more complicated numerical methods. It is shown that only clusters younger than about 100 Myr should be considered as suitable tracers of spiral arms in our Galaxy. Our method can be easily extended to other spiral galaxies and requires only the knowledge of the rotational curve, the distance and the pattern speed associated with the spiral arms.

Before we started the analysis of spiral arms with diffuse interstellar bands, we wanted to improve upon our older version of the map of these tracers of the interstellar medium. The revised maps provide several insights into the distribution of the carriers in the Solar Neighbourhood. However, neither these maps nor the more detailed maps based on an infrared band can be proved to trace spiral arms. More research in this topic is still required.

Abstrakt

Jedny z nejdůležitějších objektů pro studium hvězd jsou otevřené hvězdokupy. Mohou být použity ke studiu mnoha aspektů vzniku, vývoje, kinematiky a dynamiky hvězd. S informacemi, které poskytuje Gaia ve svých četných datových vydáních, jsme schopni studovat hvězdokupy s mnohem větší přesností než kdy dřív – to ovlivňuje nejen fyziku zkoumanou astrometrickými veličinami, ale také odvozené pravděpodobnosti členství.

Několik problémů studovaných v literatuře je stále špatně vyřešeno. Příkladem jsou známé problémy s transformací paralax na vzdálenosti, které jsou stále často ignorovány. Poskytujeme naši analýzu výše zmíněného problému a tvrdíme, že převrácené hodnoty střední paralaxy představují dobrou míru odhadu vzdálenosti až do několik kpc. Ukazujeme, že prodloužení hvězdokup podél směru pohledu zůstane téměř nezměněno, bez ohledu na zvolenou metodu transformace mezi paralaxami a vzdálenostmi.

Na další problém můžeme narazit při analýze otevřených hvězdokup z nedávno publikovaných katalogů založených na metodě UPMASK. Žádná z populárních prací nezohlednila, že se hvězdokupy mohou ve fázovém prostoru výrazně překrývat. Ukazujeme, že v některých případech má toto přehlédnutí za následek špatnou klasifikaci členů hvězdokupy. Toto ovlivňuje všechny údaje, které můžeme pro dané hvězdokupy získat – ne jenom průměrné paralaxy a průměrné vlastní pohyby, ale také věk a chemické složení.

V jedné z našich nejnovějších studií jsme analyzovali schopnost otevřených hvězdokup a difúzních mezihvězdných pásů (diffuse interstellar bands) sledovat blízká spirální ramena naši Galaxie. Připravili jsme poměrně jednoduchý postup, který přináší výsledky srovnatelné se složitějšími numerickými metodami. Ukazuje se, že pouze hvězdokupy mladší než zhruba 100 Myr by měly být považovány za vhodné indikátory spirálních ramen v naší Galaxii. Naše metoda může být snadno použita pro jiné spirální galaxie – vyžaduje se pouze znalost rotační křivky, vzdálenosti a rychlosti vzoru (pattern speed) spojené se spirálními rameny.

Než jsme začali s analýzou spirálních ramen pomocí difúzních mezihvězdních pásů, chtěli jsme vylepšit naši starší verzii mapy těchto indikátorů mezihvězdného prostředí. Revidované mapy poskytují několik pohledů na rozložení zodpovědných molekul (carriers) v okolí Slunce. Avšak, ani tyto mapy, ani podrobnější mapy založené na infračerveném pásu nemohou prokázat, že by difúzní mezihvězdné pásy sledovaly spirální ramena. V této oblasti výzkumu je zapotřebí ještě mnoho práce.

Acknowledgement

I would like to express my gratitude to my supervisor, Ernst Paunzen, who not only supported me throughout my whole studies, but also taught me a lot of what is required when doing a scientific research. Our conversations have been the foremost source of inspiration to me.

Special thanks goes to my beloved wife. Without her, I would not have been able to get as far in my life as I was able to.

I would like to thank our colleagues from Jena for their hospitality during my visits and for allowing me to learn from them. I also greatly appreciate all the support I got from my family, friends and colleagues.

Prohlášení

Prohlašuji, že jsem svoji dizertační práci vypracoval samostatně s využitím informačních zdrojů, které jsou v práci citovány.

Brno 2022 . . . . . .

Martin Piecka

Contents

Chapter 1

Introduction

The stars, that astronomers and astrophysicists aim to study, are born in the interstellar medium (ISM) and are made from the material present in the star-forming regions. Almost every star spends most of its life on the main sequence (MS), which represents the region of the Hertzsprung-Russell diagram where stars are burning hydrogen in their cores. The energy of the thermonuclear reactions supports stellar existence and fuels stellar luminosity for a period of time, which depends on the initial mass of the star when it enters the MS phase. Once the hydrogen concentration in the stellar core drops to a value when the radiation pressure cannot support the gravitational collapse, the star will start evolving. The specific evolutionary track depends on the stellar initial mass (and partially on the chemical composition of the stellar core, stellar rotation, and mass-loss rate).

Many of the stars will end their life in a spectacular stellar activity. For lower-mass and intermediate-mass stars ($\sim 1M_{\odot}$ up to $\sim 8M_{\odot}$, ignoring stellar rotation and other effects which could influence the stellar collapse), the evolution reaches its end at the asymptotic giant branch (AGB, Iben & Renzini 1983). AGB stars will lose a significant fraction of their mass in the form of a dust-driven stellar wind (Höfner & Olofsson 2018). At later stages of AGB evolution, the appearing UV radiation ionises the outgoing material, which can recombine and radiate energy in the form of a planetary nebula. The remnant stellar core becomes a white dwarf. For higher-mass stars (which have high enough initial masses to overcome the electron-degenerate pressure of the iron stellar core at the final stage of the evolution) and for some of the stars which are found in contact binary systems, the stellar life ends in the form of a supernova.

Whatever the end stage of the stellar life is, the stellar activity returns a considerable amount (and sometimes most) of the stellar mass back to the interstellar medium. Due to the processes that turn lighter chemical elements into more massive elements, the returning material differs in the chemical composition when compared to the chemical composition of the medium from which the star was born. The material originating from the stellar activity will, over time, mix with the interstellar medium from which new stars are to be born in the future. Therefore, we run to the conclusion that the chemical composition of a galaxy (in this text, lower case *g* is reserved for any galaxy, upper case *G* is used when talking about our Galaxy, the Milky Way) and of the universe evolves in time and is influenced by star-formation.

Our knowledge of the processes in the star-forming regions are heavily constrained

by studying the state of the medium prior and subsequent to the star-formation. In the case of the former, one wishes to study interstellar clouds (and most importantly, the giant molecular clouds and their sub-structures) in which most of the stars are born. As for the latter, the giant molecular clouds collapse to form multiple stars. They form kinematic groups which can be studied to gain additional information about the chemical composition of these stars or about some details regarding the star-formation (what is the approximate age of these stars? what is the initial mass function?).

Our understanding of the star-formation has significantly improved, from the perspective of molecular clouds, over the last two decades. It is generally believed that molecular clouds are typically bound by self-gravity and in the state of virial equilibrium – they neither collapse nor expand. However, the equilibrium may be unstable and a spontaneous collapse may occur. Furthermore, this state can be changed by a large number of events, including (Elmegreen 1998; Barton et al. 2000; Li et al. 2014; Dugan et al. 2014):

- A stellar activity from a nearby source, or sources (H II regions, stellar winds, supernovae)
- A collision with another cloud
- An interaction with a density wave travelling across the galaxy (excited by the gravitational potential of the host galaxy)
- A gravitational interaction with another (possibly colliding) galaxy
- An interaction with a relativistic jet originating from a super-massive black hole

The cloud begins to collapse if the triggering event created a gravitationally unstable region within the cloud by exceeding the Jeans density. However, we should note that not all molecular clouds will collapse into stars and are instead disrupted. Furthermore, the star-formation efficiency in our Galaxy seems to be rather small – less than 25 % (and usually much less) of the cloud's material is turned into stars (Murray 2011). In the star-forming scenario, the cloud will collapse into fragments and those fragments into smaller fragments, turning into regions from which the stars can be born. Due to the gravitational collapse and the presence of turbulence within the medium, elongated over-densities, called molecular filaments, are formed (McKee & Ostriker 2007). It is within these filaments where the stars are actually born. On the other hand, rotation of the cloud and the magnetic field within the cloud can reduce the star-formation rate in some scenarios. Getting such specific knowledge of the regions from which stars are later born, we can put new constraints on the star-formation processes.

As for the kinematic groups of stars, the situation is a little bit more confusing. One usually distinguishes between several types of stellar groups (e.g. Montes et al. 2001):

- Embedded clusters – stars are gravitationally bound together and moving in the same direction, they are surrounded by the remnants of the progenitor molecular cloud
- Open clusters – stars are gravitationally bound together and moving in the same direction, the stellar activity has already cleared the surrounding medium

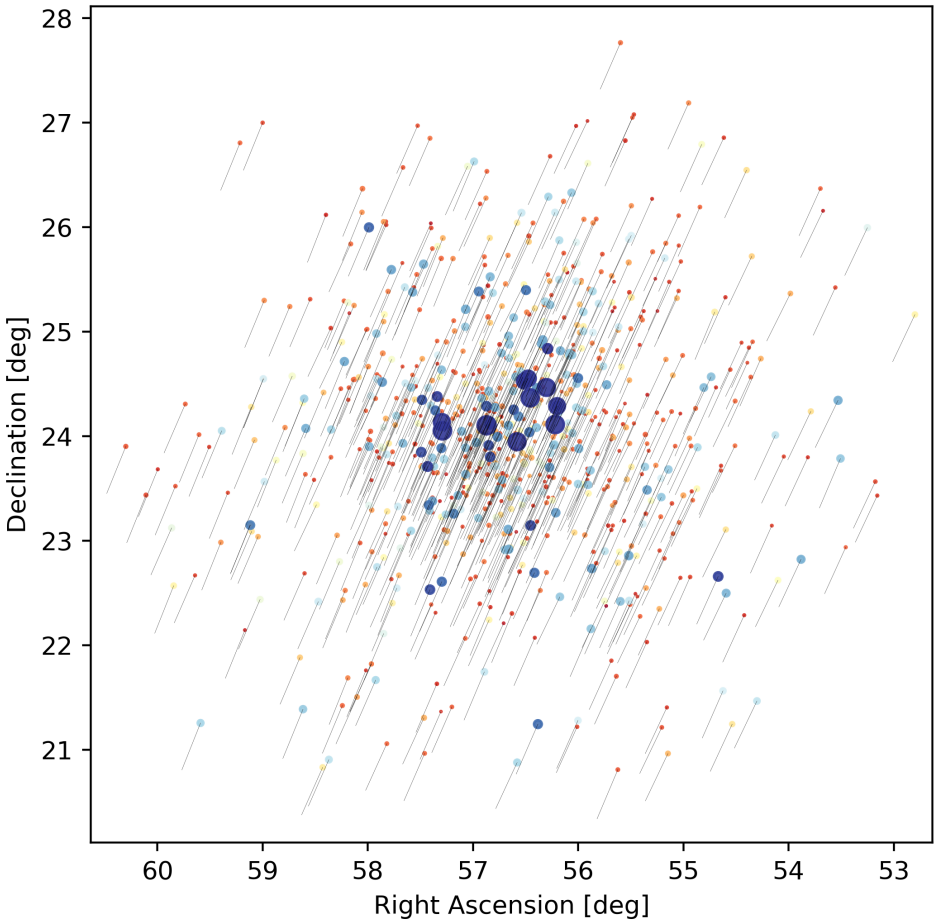


Figure 1.1: Diagram of the open cluster Pleiades (M 45, Melotte 22) based on the data from Gaia Data Release 2. The members were taken from Cantat-Gaudin et al. (2018) with the exception of the (displayed) ten brightest stars that were not included in their work. Only stars with membership probabilities higher than (or equal to) 90 % are displayed. Colour of the points represents their photometric colour (lower $B_P - R_P$ values towards blue, higher towards red) and the size of the points indicates the brightness. Lines show the direction of motion of the individual stars and clearly point out the common motion of the members.

- Stellar associations – stars are moving in the same direction (the motion is quite collimated) but they are not gravitationally bound, the stars are still relatively young
- Moving groups – stars are moving in the same direction (the motion is less collimated) but they are not gravitationally bound, the stars are older than in the case of stellar associations

It is clear that there is no proper distinction between stellar associations and moving groups. We run into additional issues when we realise that star clusters evolve a loose members (Lamers et al. 2005a; Gieles et al. 2006; Famaey et al. 2007). Stars become unbound due to the tidal forces produced by the galactic potential or due to an interaction with another large gravitationally bound object. The overall picture of a star cluster, therefore, involves stars which are bound in the core of the cluster and unbound stars which form the halo (or corona) of the cluster (Nilakshi et al. 2002). In some cases, we also observe another structure that is designated as a tidal tail (e.g. Jerabkova et al. 2021). Occasionally, a stellar association can be born together with a star cluster in the near vicinity ($\sim$ 100 pc or closer), giving almost the same age to both objects. If their motion is also very similar, it will become impossible to distinguish between the cluster's halo and the stellar association.

Still, we can learn a lot from the stellar groups. It is presently believed that most (if not all) of the stars are born in such groups. Studying star clusters gives important information about all of the well-established members. Most of the information is encoded in the four main clusters parameters: the distance, the reddening, the age and the metallicity. Considering that the cores of clusters cannot have radii larger than a certain critical value (typically around 5-10 pc in our Galaxy, Kharchenko et al. 2013), the sizes of clusters are generally much smaller than their distances from the Solar System. For this reason, we can approximate that all of the members have the same distance. If we have independent measurements of the distances of the members, we can determine the overall distance much more accurately (in this approximation, each star represents an independent measurement of the distance). Furthermore, the members are typically born within several Myr, and we usually conclude that they have almost the same ages. If we, again, assume that the progenitor cloud was chemically well-mixed, the initial chemical composition of each member had to be the same as well. Finally, we can often use the same reddening (or identically, interstellar extinction) value for each of the stars. However, the error due to this final approximation can become quite large if the cluster displays what we call differential reddening. This effect can usually be detected using Hertzsprung-Russell diagrams or colour-magnitude diagrams (CMDs).

When we want to study the detailed chemical composition of the universe, we are often interested in studying the atomic and molecular populations. The information about different species is contained in the observed stellar spectra in the form of absorption (and sometimes also emission) spectral lines and bands. Occasionally, our search for previously undiscovered interstellar molecules is prompted by laboratory research. However, most of the time we take the inverse approach – unidentified observed spectral lines are matched with the quantum mechanical models of atoms and molecules, and with the laboratory spectra. Curiously, the carriers of hundreds of lines in the optical, near infrared (IR), and near ultraviolet (UV) part of the spectrum remain unknown to this date. These absorption features are called the diffuse interstellar bands (DIBs, see for example Hobbs et al. 2008;

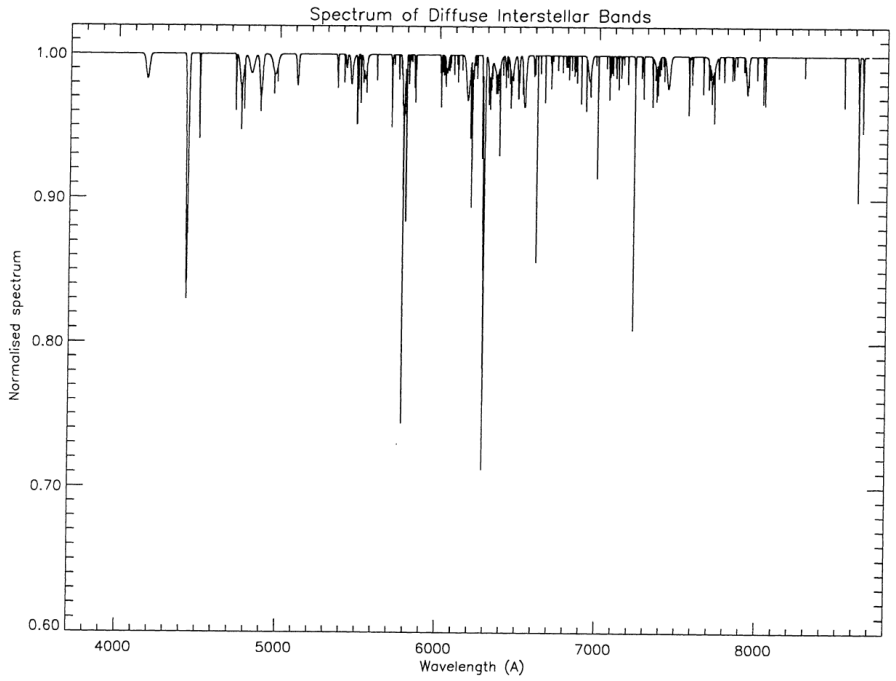


Figure 1.2: Spectrum of the DIBs in the wavelength range $3800 - 8680$ Å. Taken from Jenniskens & Desert (1994).

Krelowski 2018). Their name originates from the fact that the lines are mostly quite narrow (< 0.1 nm, although broader DIBs are also observed) but their profiles are not sharp, and they were proven to be formed in the interstellar medium (and not in stellar atmospheres or in the circumstellar medium). Up to this date, only one molecule has been successfully proven to be responsible for at least some of these features (Campbell et al. 2015). It is believed that the carriers are large organic or carbonaceous molecules (such as polycyclic aromatic hydrocarbons or fullerenes). But how are these molecules formed? How are they able to survive in the interstellar medium? The answers to these questions would improve our knowledge of the chemical composition (and evolution) of the interstellar medium. Furthermore, DIBs might serve as a useful tool for probing the interstellar environment.

Chapter 2

Gaia Overview

Before we can start to investigate the kinematic groups or the ISM, we have to get the information about the observed stars. To study the ISM which obscures the lines of sight toward the stars, we will generally need stellar positions, distances (often computed with the use of parallaxes), photometric data and (or) spectra. On the other hand, we also require the knowledge of the stellar motion for kinematic groups. This motion is observed in the form of the astrometric proper motion (the perpendicular component of the velocity vector) and the radial velocity (for which we need a spectrum).

The launch of the Hipparcos space satellite (Perryman et al. 1997) lead to the observation of a large number of stars. It provided us with photometric and astrometric solutions (position in the sky, parallax, proper motion) for over 100 000 objects in the Solar Neighbourhood. This catalogue was later revisited and corrected, providing more precise astrometric solutions for the brighter stars (van Leeuwen 2007). Together with the Hipparcos catalogue, Tycho catalogue was released (Hoeg et al. 1997; Høg et al. 2000), giving the information about positions and proper motions for over 2 million stars (together with photometric data).

We use the stellar positions, proper motion vectors, parallaxes and radial velocities to compute the full 6D kinematic properties (galactic position, dimensions, mean galactic motion, rotation, expansion) of the kinematic groups (but also of the field stars). Next, we can construct the CMDs with the help of the photometric data, which are used to analyse ages and, potentially, metallicities of these groups using stellar evolutionary tracks called the isochrones (Jørgensen & Lindegren 2005). Furthermore, we can use the CMDs to study the individual stars as well. This helps us to better understand stellar evolution (Tripicco et al. 1993) and put constraints on the stellar evolution models.

Based on astrometry, photometry and spectroscopy, we can study the motion of stars in our Galaxy (Dehnen & Binney 1998), the spiral structure of the Galaxy (Fernández et al. 2001), kinematic groups (de Zeeuw et al. 1999; Pinsonneault et al. 2000; Castellani et al. 2002), mechanisms that remove stars from binary systems (e.g. Hoogerwerf et al. 2000, 2001), and more.

Since the Hipparcos mission, new massive catalogues containing up to almost 0.5 billion objects have been released. Some examples are:

- The Two Micron All Sky Survey (2MASS, Skrutskie et al. 2006)
- The Sloan Digital Sky Survey (SDSS, York et al. 2000; Ahn et al. 2012)

- USNO CCD Astrograph Catalog (UCAC, Zacharias et al. 2000, 2013)
- PPMXL Catalog (Roeser et al. 2010)

While providing good results, Hipparcos needed to be superseded by a new mission that would give us precise astrometric data (especially parallaxes) for a larger number of objects (many of which could be found in the huge all-sky catalogues). For this reason, the ESA Gaia mission was proposed and accepted in the early 2000s (Perryman et al. 2001).

2.1 Mission Description

The aim of the Gaia mission proposal was to construct a space telescope capable of observing up to a billion different objects, giving a 5D (without radial velocities) or full 6D astrometric information about them. Included would be not only the stars from our Galaxy, but also from other galaxies in the Local Group. The main aim was to study the Galactic structure and dynamics, star-formation history (based on absolute luminosities and metallicities), and stellar properties and evolution. Secondary objectives include, for example, studies of binary systems and exoplanets, and also providing a new and updated frame of reference ICRF3 (International Celestial Reference Frame, third realisation).

Gaia satellite (Gaia Collaboration et al. 2016b) makes use of two identical telescopes (aperture 1.45 m $\times$ 0.50 m) separated by a constant angle and mounted on a support composed of 17 segments (called the optical bench). The beams from the telescopes are combined onto a common focal plane. The focal plane consists of 106 CCD detectors of three different types, the broad-band, the blue-enhanced and the red-enhanced, with a total of 938 million pixels. The detectors are separated into five groups:

- Two of the CCDs are used as the wave-front sensors – these serve to monitor the optical performance of the telescopes. Furthermore, two CCDs are reserved for basic angle monitoring.
- The sky mapper. Each telescope uses seven CCDs for detecting the objects which enter the field of view.
- The astrometric field. Uses 62 CCDs are is used to extract the 5-parametric (astrometric) solutions. Together with the sky mapper, they are used for the Gaia (G) broad-band photometry.
- Seven CCDs are allocated for the blue (B_P) and seven more CCDs for the red (R_P) photometers.
- The radial velocity spectrometer. Provides high-resolution spectroscopic data in the red band. The main purpose is to get the radial velocities of stars.

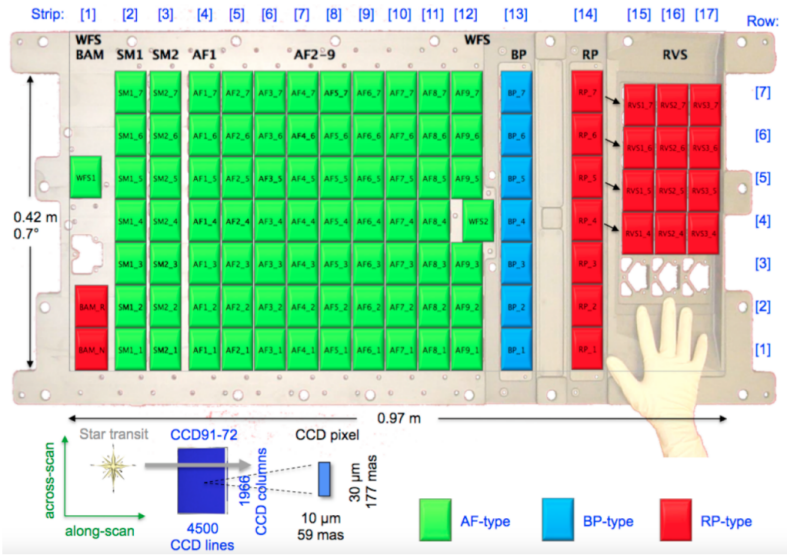


Figure 2.1: Schematic overview of the focal plane assembly. Taken from Gaia Collaboration et al. (2016b).

2.1.1 Data Gathering and Processing

The satellite was launched on 19 December 2013 and is operating up to this date. It is located on a Lissajous-type orbit around the L_2 point of the Sun-(Earth-Moon) system.

The observations are based on the principle of scanning space astrometry (Lindegren & Bastian 2010), in which the satellite slowly spins at a constant rate (≈ 59.9605 arcsec per second) and measures the focal plane crossing time of the observed objects. Specifically, this time is measured at the point when the centre of the object has a well-defined position on the CCD detector. The so-called observation time then represents the one-dimensional measurement of the object's position relative to the orientation of the instrument. The astrometric solution (described in Lindegren et al. 2012) of the observed objects is built using several such observation times.

The satellite makes use of two fields of view which are separated by a basic angle of about 106.1° – the plane given by these two directions is perpendicular to the spinning axis of the satellite. As the satellite spins, the two fields of view scan all stars along the great circle over the period of six hours. To scan across the whole sky, the orientation of the spinning axis has to change as a function of time, and is given by the "scanning law". The introduced precession period (around the Sun) of the spinning axis is about 63 days.

Due to the overall large number of pixels on the CCD cameras, it is impossible to store all of the information and then transmit it from the satellite to the Earth. To deal with the data challenge, Gaia DPAC (Data Processing and Analysis Consortium) was made in charge of the data pre-processing, the details of which can be found in Fabricius et al. (2016). The final output of the data processing is the release of data to the publicly available archives. Up to this date, there have been three data releases (with the next one to be released mid-2022) – these are described below.

2.1.2 Initial Issues

Before Gaia started gathering useful data, three main issues were discovered that affected the observations beyond the expected errors: water ice contamination, straylight and periodic basic angle variations (Gaia Collaboration et al. 2016b).

Water Ice Contamination

The water ice contamination of the focal plane on some of the mirrors was discovered short after the launch. It is believed that this ice is formed from air molecules that were initially trapped in different parts of the satellite (insulation blankets, fibre-reinforced polymer structural parts) and later released from the material, turning into solid state on the surface.

The problem was resolved by repeatedly heating up the surfaces of the contaminated mirrors and releasing the contaminant through the apertures.

Straylight

The problem with the straylight had two origins: the scattered light from the Sun and the integrated brightness from our Galaxy. Overall, the straylight contamination was about two orders of magnitude above the expected limits. If untreated, this would heavily affect the observations of fainter objects – the largest effect would be seen in the radial velocity measurements.

The scattered sunlight can enter the aperture via fibres placed between sunshield frames. To deal with this, a strategy for the radial velocity measurement needs to be followed, were the across-scan size of the windows (two-dimensional images) is adjusted to decrease the noise.

Periodic Basic Angle Variations

As mentioned above, it is assumed that the angular distance between the two lines of sight is constant. Of course, this is not absolutely true – it is impossible to completely remove vibrations from the structure of the telescope. This was well-known prior to the launch. However, once switched on, the basic angle monitor showed that the angle variations were a couple of magnitudes higher than predicted. Furthermore, the variations in the direction of the preceding telescope turned out to be about five times higher than for the following telescope. The amplitudes of these oscillations are between 0.1 and 1.0 mas.

The fluctuations show three types of modulations: 6-hour (satellite rotation period) modulation, 24-hour modulation and longer-scale modulations (due to the changes in the radiation pressure from the Sun). Based on the analysis of the problem, it was deduced that the monitoring system was performing correctly. Although the candidates for the source of the individual modulations were identified, some of the specifics of the problem still remain unknown.

Additionally, lower-magnitude jumps in the value of the basic angle (reaching a few 0.01 mas) were also discovered. It is believed that they originate from the mechanical structure of the payload.

This problem has been mitigated with the second Gaia data release. This was possible thanks to the improvements in the data processing (Gaia Collaboration et al. 2018b).

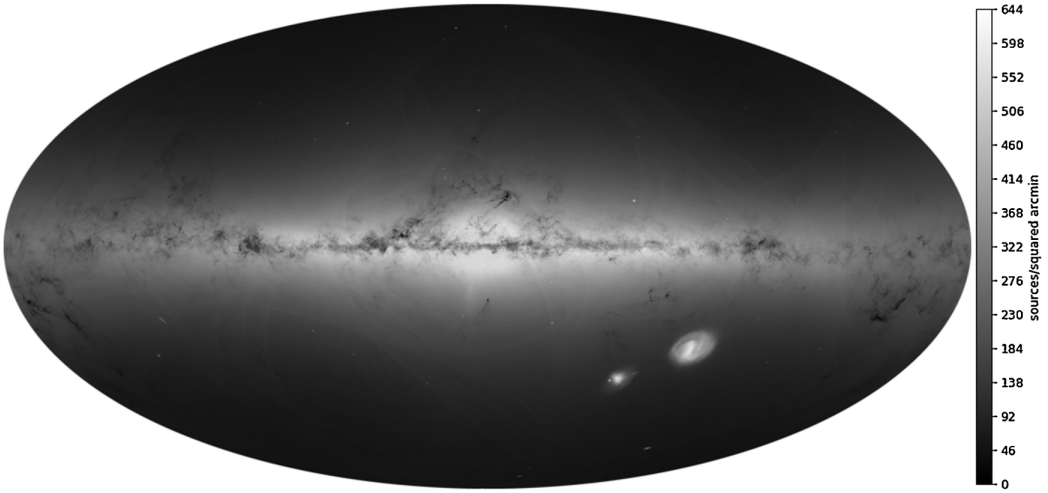


Figure 2.2: Map of the Gaia DR2 sources. Clearly visible are, for example, the Magellanic Clouds south of the Galactic disk and the globular clusters 47 Tuc (bright spot just left of the Small Magellanic Cloud) and ω Cen (just above the Galactic disk at a longitude between the Magellanic Clouds). Taken from Gaia Collaboration et al. (2018b).

2.1.3 Gaia Data Release 1

The Gaia Data Release 1 (Gaia DR1, Gaia Collaboration et al. 2016a) was published in 2016, after around three years of observations. The catalogue consists of over 1 billion objects. For most of those, astrometric and photometric data with unprecedented precision were available. Still, some of the errors were somewhat larger than expected (see the list of issues above), which gives a room for improvement in the next data release.

Despite the large number of objects in the catalogue, most of these were given only 2-parametric solution (position in the sky) together with the photometric data in the G band. The 5-parametric solution (position in the sky, proper motion, parallax) was available only for around 2 million stars (Lindegren et al. 2016). The results in the B_P and the R_P bands were missing, preventing the construction of CMDs based on the Gaia photometry, only. Still, Gaia DR1 results superseded those given by Hipparcos, especially in terms of the precisely determined parallaxes.

Based on the Gaia DR1, several very important works were published. As specified by the main aim of the mission, Galactic kinematics and dynamics were studied, giving us updated information about the Galactic rotation and velocity ellipsoid (Bovy 2017a; Anguiano et al. 2018). Next, it was shown data Gaia data supported the idea of spiral arms (in our Galaxy) being transient arms rather than static density-wave arms (Baba et al. 2018). In different studies, Gaia data were used to analyse the stellar density distribution in the Solar Neighbourhood (Bovy 2017b) and to study the properties of white dwarfs (Tremblay et al. 2017). Finally, it should be mentioned that Gaia DR1 showed stellar tidal tails between the Magellanic Clouds (Belokurov et al. 2017), proving that the Gaia data could be used outside of our Galaxy.

2.1.4 Gaia Data Release 2

The second data release has been out since 2018. Gaia DR2 (Gaia Collaboration et al. 2018b) provided a much richer catalogue of objects, adding 0.5 billion stars on top of what was given by Gaia DR1. More importantly, 5-parametric solutions were provided for more that 75 % of all objects – this covers almost 1 % of all stars in our Galaxy, for which we have the 2D motion in the sky and the distances.

One of the reasons behind the success of Gaia DR2 were the additional two years of observations – as was mentioned in the description of the mission, the astrometric solution is built on having multiple observations of the same object. Secondly, the Gaia team managed to improve the data processing and partially mitigate the problems caused by the oscillations in the basic angle value. Next, modelling of the pointing (direction of the spinning axis of the satellite) was changed, further improving the astrometric precision.

This data release also provided new information about the observed objects. This includes the radial velocities for over 7 million stars and the photometric measurements in B_P and R_P in over 1.3 billion lines of sight. Furthermore, a new reference frame (ICRF3) was built based on the DR2. Finally, a data set of variable sources was given, together with the light curves for many of those objects.

One of the problems with Gaia is that each release is supposed to be treated independently (Gaia Collaboration et al. 2018b). This is due to the changes in the way how the data are handled in the pre-processing and processing stages. One of the main points is that the photometric passbands change a little with each data release. Secondly, the identification of the objects may change – the comparison of the same object between two data released should be treated with care.

An incredible number of papers were published based on the data from Gaia DR2. These, of course, include new studies of the Galactic kinematics (e.g. Gaia Collaboration et al. 2018c) and of the structure of the Solar Neighbourhood (Bennett & Bovy 2019). Stars with very high velocities compared to the local field of stars were analysed (Boubert et al. 2018). New catalogues of open clusters were published (Cantat-Gaudin et al. 2018; Cantat-Gaudin & Anders 2020; Dias et al. 2021). Individual and global kinematics of open clusters and of stellar associations were investigated (Soubiran et al. 2018; Kuhn et al. 2019; Cantat-Gaudin et al. 2019b), together with the cluster ages (Bossini et al. 2019; Dias et al. 2021). One should also note the discovery of tidal tails around star clusters, such as Hyades (Röser et al. 2019; Meingast & Alves 2019). Gaia DR2 was also used to probe the internal kinematics of the Large Magellanic Cloud (Vasiliev 2018).

As for the interstellar medium, star-formation was studied in the local region (Dzib et al. 2018), as well as across the whole Galaxy (Kounkel & Covey 2019). Distances toward molecular clouds were studied with the use of Gaia DR2 (Yan et al. 2019). By combining Gaia DR2 and 2MASS photometry, interstellar dust maps were constructed up to the distance of a few kpc (Lallement et al. 2019).

All of these works represent only a small part of the whole set of published works. Other works also include studies of the origin of young open clusters, of several specific star-forming regions, studies combining asteroseismology with the Gaia data, analyses of abundancies and metallicities of individual stars and of the whole Galaxy, and much more.

2.1.5 Gaia Early Data Release 3

Due to some issues with data processing together with the effect of the global pandemic situation, the third data release needed to be postponed. Instead, the early version of the third Gaia data release (EDR3, Gaia Collaboration et al. 2021a) was published at the end of 2020.

Overall, this data release did not contribute to our knowledge as much as the DR2. Still, EDR3 gives more precise astrometric solutions and slightly extends the total number of observed objects (up to around 1.8 billion). Sources for the Gaia reference frame have been extended from around 0.5 million up to 1.5 million quasars. However, we still have to wait for the new catalogue of the Gaia radial velocity measurements, as EDR3 only contains the revised values from the DR2 (Seabroke et al. 2021).

With EDR3, the spiral structure of our Galaxy was revisited (Xu et al. 2021; Castro-Ginard et al. 2021). The structure of the Magellanic Clouds was re-analysed (Gaia Collaboration et al. 2021b), showing improvements in the precision of the new data release. Hundreds of other works were also published, generally revisiting the issues studied with the use of the DR2.

2.1.6 Gaia Data Release 3

At the time of writing this thesis, Gaia DR3 (`https://www.cosmos.esa.int/web/gaia/dr3`) is scheduled to be released in mid-2022. It will be the final data release of the main Gaia mission, although additional data releases based on the extended mission (post-2022) are also planned.

Gaia DR3 is supposed to deliver additional information compared to the previous data releases (in addition to the improvements in the astrometric solutions over the last one and a half year). Included should be: B_P and R_P spectra for around 100 million sources, radial velocity spectra for around 1 million sources, radial velocity measurements for over 30 million sources, and more.

2.2 Astrometry

Astrometric measurements are possibly the most important result of the Gaia mission. In this section, the process of deriving the astrometric solution in Gaia DR2 is reviewed. Afterwards, the handling of the errors is described. Finally, the proper motion and the parallax measurements are analysed in detail.

2.2.1 Constructing Solution

Details regarding this topic are described in Lindegren et al. (2018). The position on the CCD of the observed target serves as the main input in constructing the solution. Each source is given a specific ID, and all measurements for a given source are combined together. If available, the photometric colour $B_P - R_P$ is used in the chromacity calibration which improves the astrometric solution. In the solution itself, it is assumed that the motion of the observed source is uniform with respect to the barycentre of the Solar System. Clearly,

this assumption is incorrect in the case of binary stars (such sources will be treated in the Gaia DR3).

In order to construct the solution, we are required to build the source model (described in Gaia Collaboration et al. 2016b). Next, the knowledge of the orientation of the optical instrument in ICRS (International Celestial Reference System, specified by the positions of the observed quasars) as a function of time is required – this is described by the attitude model. Finally, a calibration model is constructed, based on a combination of basis functions (see Lindegren et al. 2018, Table 2). The calibration model included in the Gaia DR2 ignores small-scale spatial effects (which are to be accounted for in Gaia DR3). Part of the calibrations also deals with the variations in the basic angle.

The complete astrometric solution consists of the primary solution and of the secondary solutions. The primary solution makes use of only a small number of sources. For these, the attitude model parameters, the calibration model parameters and the astrometric parameters are adjusted at the same time in an iterative process. Quasars are part of the primary sources, and hence this approach is also used to adjust the reference frame. As can be seen, the astrometric solution is self-adjusting. As for the secondary solutions, the attitude and the calibration parameters are kept fixed while searching for the astrometric solutions of every source on the list. During the calculation of the secondary solutions, astrometric parameters of the primary sources are calculated as well, as a part of the self-consistency check.

The astrometric solution for the Gaia DR2 was derived in two steps. First, a provisional solution was calculated which consisted of the parameters for about 1.6 billion sources. It served for the identification of possible issues and for the wavelength calibration of the instruments used for deriving photometric data and radial velocities. Afterwards, the final solution took into account the chromacity calibration and updates to the model parameters based on the provisional solution.

The construction of the solution can be summarised in the following steps:

1. Astrometric global iterative solution (AGIS) pre-processing. Serves for collecting the input data for all sources.

2. Preparation of the preliminary secondary solutions (based on the provisional solution). The purpose of this step was the collection of the source statistics, which is to be used in the next step.

3. Selection of the primary ($\sim$ 16 billion) sources.

4. Deriving the primary solution.

5. Preparation of the final secondary solutions. At this stage, sources are separated into 5- and 2-parametric solutions.

6. Regeneration of attitude and calibration. Necessary for making use of the observations during time intervals when the observations were problematic or non-ideal (moments of problematic observations, the first month of nominal operations).

7. AGIS post-processing. In this final step, the data are converted into the final format and stored in the main database.

2.2.2 Covariances and Error Propagation

As a measurement value, each of the astrometric parameters (the position in ICRS given by α and δ, the proper motion components μ_α and μ_δ, the parallax ϖ) has an associated measurement error. Furthermore, it is obvious from the process in which the solution is constructed that the parameters are generally not independent but are, instead, correlated.

This becomes important when (for example) calculating quantities which depend on multiple astrometric parameters, when analysing the proper motion diagram, or when transforming from ICRS to a different coordinate system. In such a case, one is required to work with a covariance matrix. While the terms on the main diagonal (corresponding to the associated errors) are typically the most important, the off-diagonal terms should not be neglected (or at least considered).

The mathematical approach to this problem is discussed in Holl & Lindegren (2012).

2.2.3 Proper Motion

The proper motion plays a key role in the kinematic studies of stars, kinematic groups, or galaxies (either on its own or together with the radial velocity). It represents the tangential motion vector of the stars of the sky. As such, it has two components, which are typically taken to correspond to the motion in right ascension α and in declination δ. However, the angular motion μ_α must be treated with care. If we were to work with α and δ, we would be assuming an equirectangular map projection of the stellar positions on the sphere. In such a map projection, the meridians are at a constant distance from each other. However, they meet at the poles on a sphere. To account for the distortion near the poles, one can multiply the $x-$coordinate by $\cos\delta$. Indeed, it is this corrected value of μ_α which is given in the Gaia data (where it is labelled as $\mu_{\alpha*}$).

There is a dependence of the absolute error of a proper motion component on the brightness of the source due to the photometric noise. The brighter the object is, the lower the error (typically, but the brightest sources also represent a problem). In Gaia Collaboration et al. (2018b) (Table 3), this is summed up as follows:

- $G < 15$ mag, typical error of about 0.07 mas yr^{-1}
- $G = 17$ mag, typical error of about 0.2 mas yr^{-1}
- $G = 20$ mag, typical error of about 1.2 mas yr^{-1}
- $G = 21$ mag, typical error of about 3 mas yr^{-1}

It should be noted that systematic errors are present in the astrometric results of the Gaia observations (Luri et al. 2018). Usually, one only considers the systematics for parallaxes. However, it is important to be aware of the fact that they also affect the proper motion and are of the order of several 0.01 mas yr^{-1}. Accounting for systematics may be important for studies that aim to search for exoplanets. For open clusters and stellar associations, this effect is usually negligible.

2.2.4 Parallax

Treatment of the parallax measurement is much more complicated than dealing with the proper motion. In this section, all of the aspects of how to deal with the parallax are described (especially when trying to get the distance of the source).

First of all, the parallax was shown to be affected by systematic errors of the order of < 0.1 mas (Gaia Collaboration et al. 2018b). In later works, the systematics were studied by analysing the parallaxes of eclipsing binaries. Based on the studies published by Stassun & Torres (2018, 2021), the Gaia parallaxes seem to be somewhat smaller than they should be. In their first study, it was concluded that the systematics are in the range of -0.082 ± 0.033 mas for stars with $5 \leq G \leq 12$ mag and at distances starting from 30 pc and going up to 3 kpc. This is in agreement with the predicted systematic errors. Improvements in the astrometric solution for Gaia EDR3 (Lindegren et al. 2021b) lead to the reduction of the systematics down to -0.037 ± 0.020 mas. Furthermore, the application of the parallax correction recommended in Lindegren et al. (2021a) resulted in the final value of systematics -0.015 ± 0.018 mas, almost eliminating the problem.

It can be seen in the Gaia data that some of the parallaxes have negative values. This is the result of how the astrometric solution is being constructed – for an illustration of such a problem, see Section 3.1 in Luri et al. (2018). Such an issue usually occurs when the measurement noise becomes of the order of the "true" parallax value (this is the actual value of the parallax which we cannot really know for a real observation; usually useful for simulations where the value can be known). Traditionally, such a situation is resolved by removing the measurement from the scientific analysis. However, this may lead to biases. An example of such a case would be the study of a cluster (open or globular) at a large distance of about 6 kpc (or $\varpi = 0.17$ mas). Assuming a parallax uncertainty of 0.10 mas, this would lead to a very high probability of a small fraction of the observed cluster members having negative parallaxes. Still, the overall distribution of values would be centred around the "true" value of the cluster's mean parallax. However, if we remove the negative parallaxes, the mean of the distribution shifts toward larger (more positive) values, away from the "true" value.

Another problem is encountered when one wants to get the actual distance d of the observed source. First of all, negative distances have no physical meaning. Secondly, one will find that the parallax with its measurement error may not transform correctly to the distance (and its uncertainty) when simply using $d = 1/\varpi$. If we started with a parallax modelled by a normal distribution, such a distribution would be transformed by this relationship, resulting in an extended tail towards larger distances (when compared with the one towards shorter distances). Moreover, the maximum of the distribution would be shifted towards lower values (when compared with the "true" distance). For this reason, it is recommended to avoid the traditional approach, unless the parallax is a positive value and its relative error is very small (ideally much smaller than 10 %).

An example of a proper way of dealing with such a problem has been proposed in Bailer-Jones (2015) and further explained in Luri et al. (2018). It makes use of a Bayesian approach to the problem. The main task then becomes the choice of the prior. Different choices are detailed in the former mentioned work, while the latter work focuses on the use of the exponentially decreasing space density prior. This option was used in the work by Bailer-Jones et al. (2018), who determined distances toward most of the stars from Gaia

DR2. Regarding the uncertainties associated with the distances, it is recommended (and was used in the Bailer-Jones works) to use quantiles (such as 5 % and 95 %) instead of a standard deviation. In a later work, Bailer-Jones et al. (2021) extracted distances from Gaia EDR3, using a more sophisticated prior.

It should be pointed out that the mentioned choice of the prior is mostly useful only for the field stars. For stars located in clusters, a different choice would be more practical. An example can be seen in the work by Carrera et al. (2019), who worked with a combination of the exponentially decreasing space density function and of a Gaussian model of the cluster's profile.

Similar to what was shown for the proper motion, the parallax errors are magnitude-dependent (DR2):

- $G < 15$ mag, typical error of about 0.03 mas
- $G = 17$ mag, typical error of about 0.1 mas
- $G = 20$ mag, typical error of about 0.7 mas
- $G = 21$ mag, typical error of about 2 mas

Finally, it should be pointed out that the parallaxes (not only) from Gaia have certain limitations. A very good example is the star Betelgeuse (α Ori). This star is a nearby ($\sim$ 100 pc) evolved red supergiant. It has a very large angular diameter of above 40 mas (Montargès et al. 2016), which makes it much larger than the value of the parallax. Hence, determination of the parallax of Betelgeuse (Harper et al. 2017) is very difficult and will always yield significantly larger errors than for most of the other stars at similar distances. It should be noted that the parallax of Betelgeuse was excluded from Gaia DR1, DR2 and EDR3.

2.3 Photometry

The introduction to the Gaia broad-band photometry was first presented in Jordi et al. (2010). The spectral energy distribution of the sources is represented by the brightness in three broad filters. The exact calibration of the passbands changes with each data release, but they can be described in a more general way:

- G-band covering $\lambda \in (330, 1050)$ nm, centered near 673 nm, optimised for astrometric measurements.
- B_P-band covering $\lambda \in (330, 680)$ nm, centered near 532 nm, calculated by integrating the low-resolution spectra from the blue photometer.
- R_P-band covering $\lambda \in (630, 1050)$ nm, centered near 797 nm, calculated by integrating the low-resolution spectra from the red photometer.

Generally, the photometric calibration does not deal well with the saturation effects that occur in the case of brighter sources. To deal with this problem, the magnitudes need

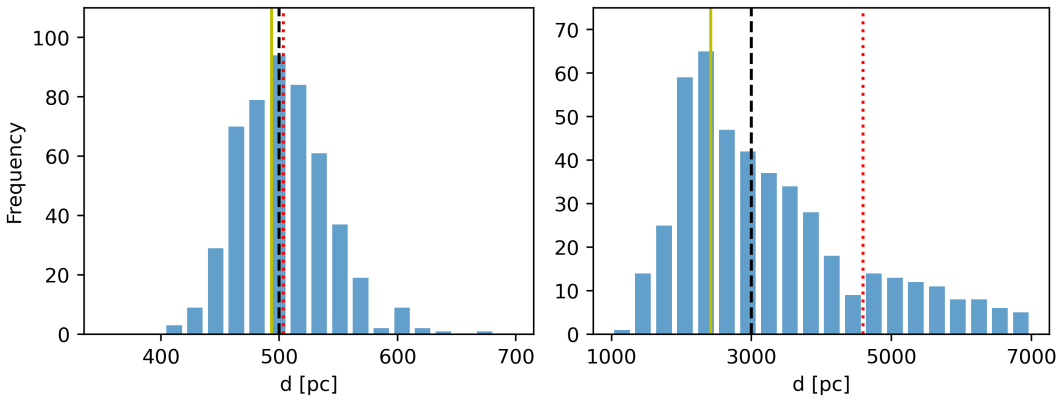


Figure 2.3: Examples of the transformation issue that occurs when the distances of the cluster members are calculated using the inverse-parallax approach. Left: a cluster at the "true" distance of 500 pc. Right: a cluster at the "true" distance of 3000 pc. The absolute parallax error is the same for both simulated cases ($\sigma_{\varpi} = 0.15$ mas) and both clusters have the same number of members $N = 500$. The errors are assumed to follow a Gaussian distribution. The black dashed vertical lines display the "true" distances of the clusters, red dotted lines point out the means of distances and yellow solid lines give the modes of distances.

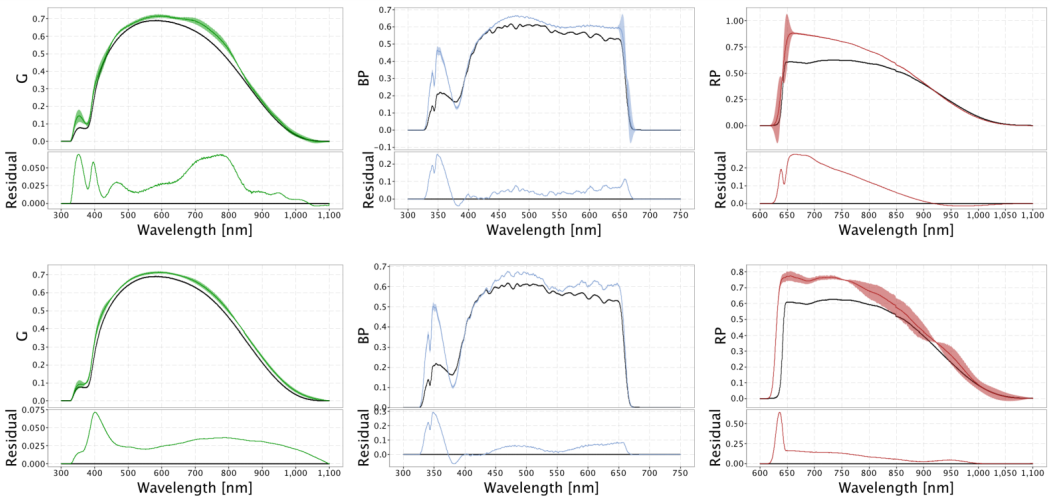


Figure 2.4: The two sets of Gaia passbands shown in Evans et al. (2018). The upper set represents the original Gaia passbands, the lower set is designated as the revised set of passbands which has been recommended for Gaia DR2.

to be corrected by using the relationships specified in Evans et al. (2018, DR2) and Riello et al. (2021, EDR3).

Additional issues and inconsistencies were also detected (usually on the mmag level). For example, studies of open clusters may be affected due to the stronger flux excess in B_P than in R_P in the lower parts of the CMDs (Gaia Collaboration et al. 2018b; Evans et al. 2018). Even more affected are the more crowded regions (e.g. near cores of globular clusters) and the regions around brighter stars. It is recommended to use $G-R_P$ instead of B_P-R_P when studying faint red sources. The situation has slightly improved with EDR3 (Riello et al. 2021). Another important effect was described after EDR3, where systematics depending on B_P-R_P were introduced in the determination of the flux in G. Polynomial relations were provided to deal with this issue (for stars $G > 13$ mag).

The information about the B_P and the R_P magnitudes has been provided only since Gaia DR2. This gives us the ability for studying the stellar effective temperatures (e.g. Mucciarelli & Bellazzini 2020) and to produce the CMDs. This is most useful when studying star clusters. However, the data can be used to obtain effective temperatures for all stars (Evans et al. 2018) together with the interstellar extinction, although with various degrees of success.

At some point, a researcher will likely want to get the absolute magnitudes in various Gaia passbands. To get the absolute magnitudes, one must know the distance from the source and the interstellar extinction in a given photometric band. Usually, we have good knowledge of the extinction A_V or of the reddening $E(B-V)$. One-parameter transformations are often available to get the extinction (or reddening) in any other filter (or colour). However, the situation for Gaia is more complicated due to the non-negligible broadness of the passbands. Instead of a one-parameter transformation, one should use colour-dependent polynomial relations (see Gaia Collaboration et al. 2018a).

2.4 Spectroscopy

The details about Gaia's onboard spectrometer, the Radial Velocity Spectrometer (RVS), can be found in Cropper et al. (2018). To process the Gaia DR2 data with the aim to extract radial velocities (wavelength calibration, dealing with scattered light, processing spectra, and so on), a special spectroscopic pipeline needed to be introduced (Sartoretti et al. 2018). It was applied to stars brighter than $G_{\rm RVS} = 12$ mag (brightness of the star in the passband used for the radial-velocity-spectra), with a possibility to extend down to $G_{\rm RVS} = 14$ mag for Gaia DR3. It must be emphasised that in the Gaia DR2 catalogue, radial velocities are included only for stars with $3500\ {\rm K} < T_{\rm eff} < 7000\ {\rm K}$.

As for the precision, the median error in the radial velocity is below $0.5\ {\rm km\,s^{-1}}$ for stars brighter than $G_{\rm RVS} = 8$ mag. The median error rises to about $1\ {\rm km\,s^{-1}}$ at $G_{\rm RVS} = 10.5$ mag and even further up to $2\ {\rm km\,s^{-1}}$ at $G_{\rm RVS} = 11.5$ mag. It should also be noted that the radial velocity measurements are also affected by systematic offsets of about $+0.3\ {\rm km\,s^{-1}}$, which is within the range of values predicted prior to the launch of the satellite. When compared with the radial velocities obtained from APOGEE (Katz et al. 2019), it was shown that Gaia radial velocities exhibit an increasing trend in the residuals for stars with $9\ {\rm mag} < G_{\rm RVS} < 12\ {\rm mag}$.

To understand how precisely the radial velocities of open clusters are determined, it is

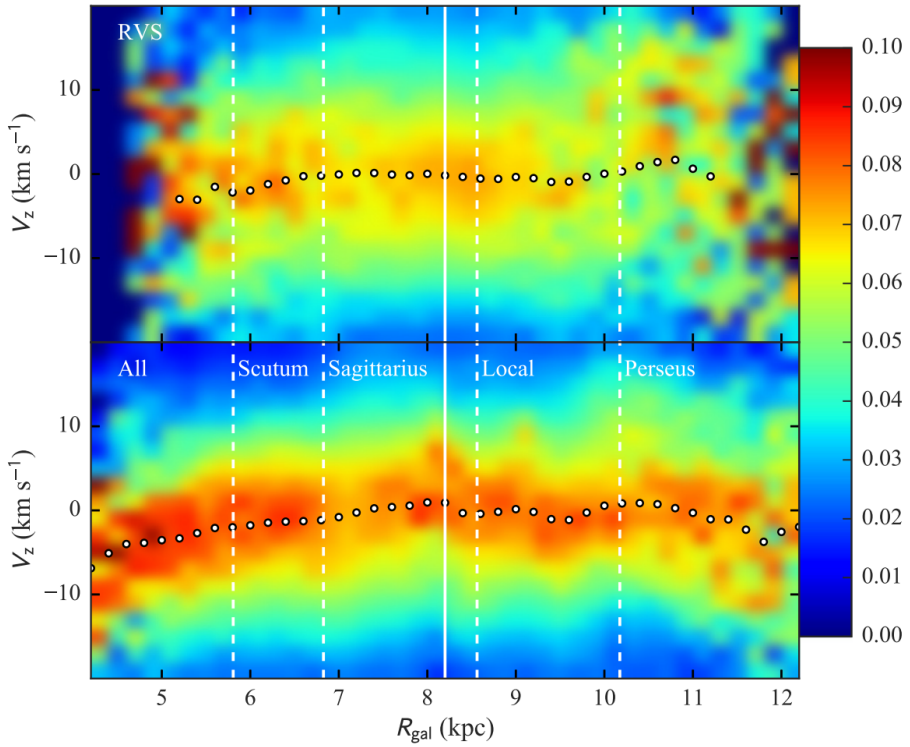


Figure 2.5: The wave-like structure seen in the vertical component of the Galactic velocity as a function of the Galacto-centric distance (Kawata et al. 2018).

best to look at the work by Soubiran et al. (2018). We can notice that for clusters where more than five members were used for the analysis, a typical error will be more than 0.3 km s^{-1}. If we were to select the clusters for which the number of members with radial velocity measurements is smaller, we would find a typical errors about three times larger. Such precision is more than adequate for both, statistical studies (motion in the Galactic disk) and studies of the individual clusters.

It is important to also mention the velocity maps produced by Gaia Collaboration et al. (2018c) and Kawata et al. (2018). Focusing on the latter work, these maps were created for the region within 3.5 kpc from the Sun using Gaia DR2 radial velocities and distances. Such maps give information about the kinematic structure of our Galaxy and shine some light on the origin of the Galactic spiral arms. The results support the notion that the spiral arms are transient (created by a superposition of unstable density waves) rather than long-lived features (see Sellwood 2011). Furthermore, a wave-like structure can be seen in the relation between the vertical velocity component and the Galacto-centric distance.

Chapter 3

Studies of Open Clusters

In this section, the summary of our knowledge prior to the 1990s (from before the Hipparcos era) is presented. Afterwards, the results of the more recent works are compiled. Included are the works that had either a huge impact on the advances in this field of research (star clusters) or have a great potential in future research. Furthermore, works based on the data from other projects than Hipparcos and Gaia are also mentioned.

The research during the Gaia era is still ongoing (2022 and onward), but the presented compilation of papers is based on the works published before 2022. Moreover, while Gaia has already superseded Hipparcos in almost every way, the data from Hipparcos can still be very useful.

3.1 Situation Before 1990s

It should be mentioned that prior to the Hipparcos data release, precise astrometric results were available only for several clusters (and stellar associations) located within 500 pc from the Sun. Instead, the research of star clusters was mostly focused on photometric studies (mainly Hertzsprung-Russell diagrams and CMDs). From the photometric point of view, current age has significantly changed from the mentioned time period – while our current detectors are almost always based on the CCD (charge-coupled device, invented in the 1960s, first tested in the early 1970s, applied in astronomy since late 1970s – see Tompsett et al. 1970; Young et al. 1978; Christian et al. 1985) technology, past measurements were typically made either with photographic plates (Trumpler 1921; Hanson 1975) or with photoelectric photometers (Kron & Mayall 1960; Lindenmann & Hauck 1973).

Already in the early decades of the 20th century, the existence of star clusters and their various types was well known (for example Plummer 1911; Shapley 1930). It should be noticed that star clusters were divided into two main sub-groups. The first group is represented by globular clusters, which have a very clear spherically symmetric shape on the images. One of the other main characteristics is their brightness – globular clusters appear to be quite bright (especially near the centre of the cluster) due to the relatively larger number ($10^3 - 10^6$) of cluster members. Moreover, the 3D positions of the individual globular clusters suggest that they form a spherically symmetrical distribution (belonging to the Galactic halo or to the Galactic bulge) around a point in our Galaxy, which is believed to coincide with the centre of the Galaxy (Shapley 1918; Racine & Harris 1989).

The second group of clusters is formed by, so-called, open (or galactic, Trumpler 1925) clusters. Unlike their globular counterparts, they consist of a lower number of members (typically $\sim 10^2$, up to $\sim 10^3$) and they do not have a clear spherical shape. Moreover, they are usually found in the disks of spiral galaxies, suggesting possible differences in the formation of the two sub-groups of star clusters (Miller & Scalo 1978; Fall & Rees 1985). While we know with certainty that open clusters are a by-product of the star-formation from the progenitor molecular cloud, the origin of the globular clusters remains unknown to this date.

While kinematic studies of their stellar members were rare (an example of such work would be Mayall 1946), early attempts to quantify distances and sizes of clusters do exist (Trumpler 1930). These works represented an important step towards (what we presently call) modern times. Other important insights were provided by studies of the distribution of cluster-member stars. Even today, King's profile and radius (King 1962) remain an important tool for investigating both sub-groups of star clusters.

A very interesting mystery regarding the stars present in some of the star clusters remains unsolved. These stars do not belong to the cluster's MS (although they are MS stars) or to the red-giant branch, and are instead shifted blue-ward from the turnoff point of the MS (Sandage 1953). For this reason, these stars were labelled as blue stragglers. While theories about the formation of such stars do exist (Wheeler 1979; Leonard 1989), none of them manage to answer all questions about their existence.

Open clusters were used to study the structure of our Galaxy even before the large astrometric surveys. For example, Janes & Adler (1982) made use of the available open cluster data (distances, reddening values, ages) for over 400 objects, a good and representative sample of then-known open clusters. One of the main presented findings shows that clusters do not seem to follow spiral arms very tightly. Instead, they form irregularly distributed clumps and segments of Galactic arms. The authors argued that these results may be somewhat biased by ignoring the faint objects (especially those located behind dense interstellar clouds).

Various biases tend to affect most of the astronomical observations. Malmquist bias (Malmquist 1922) represents a selection bias based on a preferential choice of brighter objects. This can happen when the observed sample of stars is magnitude-limited. For example, it was shown that this kind of bias affects distances calculated using the Tully-Fisher relation for extra-Galactic sources (Bottinelli et al. 1986, 1988). While a correction has been proposed by Malmquist, the correction of a bias may be a non-trivial problem and it will depend on the specific situation. To avoid the problem, one usually restricts the sample of observed objects in the distance (or the parallax) in order to work within a volume of space in which the sample is unbiased. However, this approach leads to a different kind of bias (Lutz & Kelker 1973) which systematically increases the value of the parallax (when compared with the "true" parallax). Lutz & Kelker proposed a way to correct the bias, which itself is, unfortunately, problematic (e.g. Francis 2014).

Finally, the state of our knowledge about the open clusters by the end of this era can be summed up in four main points:

- **Membership probabilities:** The probability of whether or not a star is an actual member of an open cluster was often ignored. The main way of deriving membership probabilities was to use the CMDs (for example Janes & Adler 1982). Often, the

procedure did not yield actual probability measures but only a flag indicating whether the star is a member (or not). A database of open clusters was developed (Mermilliod 1995, currently hosted at Masaryk University, `https://webda.physics.muni.cz/`).

- **Ages and metallicities:** The oldest known open cluster was Melotte 66 ($\log \mathrm{Age} = 9.8$, Demarque 1980). An isochrone fitting approach was the most advanced method (Patenaude 1978) and is still one of the most popular methods today. The precision with which the age (and other cluster parameters) can be extracted depends heavily on the quality of the photometric data, the way how the distance is derived and, of course, on the quality of the evolutionary tracks. The photometrically derived metallicities (or chemical abundancies) were not very reliable (Demarque 1980) – instead, the spectroscopic approach began to be more popular by the end of the era (Boesgaard 1989). One of the most interesting topics in global studies of the Galactic chemical composition is the observed metallicity gradient, which was suggested based on radio observations in Churchwell & Walmsley (1975). Shaver et al. (1983) further analysed the problem and arrived at relatively precise gradient relationships for several elements. It should be pointed out that these works were based on studying H II regions, the distances of which cannot be precisely established, especially when compared with star clusters. Still, such research had higher statistical significance than the research done with open clusters (such as Janes 1979).

- **Spatial and kinematic maps:** As was mentioned above, the most advanced spatial studies could not unambiguously show that open clusters in our Galaxy are well-correlated with the positions of the spiral arms. Kinematic maps were generally quite poor-quality and often useful only for studies of several specific clusters (e.g. Hyades, Eggen 1960).

- **Evolution of open clusters:** A successful theory of open clusters should be able to explain how they are formed, how they evolve and how they disperse in a galaxy. Up to this point, it was shown that clusters are born from giant molecular clouds (Bash et al. 1977). The comparison of the lifetimes of clusters was compared with the predictions by N-body simulations (which agreed with the observational results, Terlevich 1987). Encounters with typical interstellar clouds were found to barely affect the existence of clusters. The effect of the mass segregation on the velocity distribution (and on stellar mass distribution) was known.

- **Beyond our Galaxy:** Studies about open clusters outside of our Galaxy were published. The most studied were the Andromeda Galaxy (Hodge 1979) and the Magellanic Clouds (van den Bergh 1984). In the former study, it was shown that the distribution of clusters follows a spiral pattern and coincides with the galactic disk. However, several questions remained unanswered, for example, the observed star-formation enhancement at the intermediate galacto-centric distances when compared to the halted star-formation in the regions with older stars at larger distances. On the other hand, studies of Magellanic Clouds showed that the differences between globular and open clusters are not as apparent as in our Galaxy.

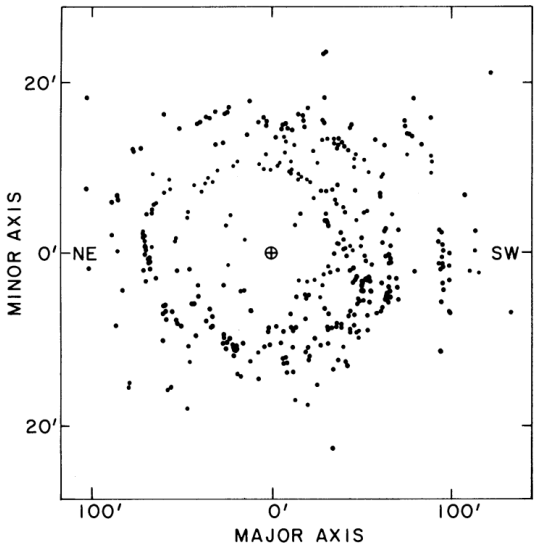


Figure 3.1: Distribution of open clusters in the plane of the Andromeda Galaxy (Hodge 1979).

- **Distance of Pleiades:** The consensus regarding the distance of Pleiades based on various methods (parallaxes from Hipparcos and the distance moduli from CMDs) could not be reached for a long time. This represented a big problem – how can we study clusters beyond 1 kpc when we cannot agree on the distance of one of the closest clusters? At this point, we note that the distance toward Pleiades is determined fairly consistently, yielding $d \approx 130$ pc based on the photometric approach (Crawford & Perry 1976; O'dell et al. 1994).

3.2 Hipparcos Era (1997-2016)

As was mentioned previously, Hipparcos satellite brought us astrometric measurements for millions of stars. This gave the option to study open clusters in different ways than just (mostly) with photometry. Indeed, parallaxes of clusters were studied only about two years after the Hipparcos data release (van Leeuwen 1999; Robichon et al. 1999).

Open clusters provided a good option for studying different stellar objects. For example, open clusters were used as a tool for analysing theories that aimed to explain the existence of Be stars (McSwain & Gies 2005) – these are fast-rotating B-type stars surrounded by a circumstellar disk (Slettebak 1949), which can drastically affect the shape of stellar spectral lines. Clusters were used to determine the age of the studied Be stars and it was shown that the fraction of such stars increases with the age of the host cluster. Furthermore, it was suggested that most of the (studied) Be stars may have gained their high rotation frequency from a binary companion via mass transfer.

Another example of how open clusters can be used is studying the properties of binary stars (e.g. Southworth et al. 2004). In such a case, the distance and the age (and potentially the metallicity) of the host cluster can be used to infer these properties of the binary star.

This leads to very precisely derived parameters (masses, radii, effective temperatures) of the components of the studied binary star.

During this era, we can notice a large increase in the empirical relations used for deriving effective temperatures and bolometric corrections based on stellar photometric colours and metallicities (Flower 1996; Alonso et al. 1996, 1999; VandenBerg & Clem 2003; Masana et al. 2006; González Hernández & Bonifacio 2009; Pecaut & Mamajek 2013). Such calibrations not only serve for estimating stellar parameters (Torres & Ribas 2002; Torres et al. 2010), they are also very useful when comparing observed data with stellar and evolutionary models (especially when studying star clusters, Pöhnl & Paunzen 2010).

Besides Hipparcos, advances in the research of star clusters were made also with the use of other instruments. Reports of multiple stellar populations observed in globular clusters (e.g. Piotto et al. 2007) represented a new challenge for the theories of these objects. Several clusters showed up to three distinct populations, pointing to the fact that globular clusters tend to host several generations of stars. To reproduce such an effect (for the majority of globular clusters), star-formation is believed to have happened in a series of events. A review of these new findings was presented by Gratton et al. (2012). The effect of including multiple populations was also studied in relation to the evolution of the binary stars in globular clusters (Hong et al. 2015). Based on two populations and with the use of N-body simulations, it was found that first-generation binaries tend to survive encounters in the host cluster more often than the second-generation binaries. However, it was also suggested that the central regions of the cluster can be dominated by the second generation of binary stars.

3.2.1 Membership Probabilities

Hipparcos astrometric data allowed for more sophisticated methods for determining cluster members. Besides making use of the CMDs, astrometric parameters were used to identify stars that form a cluster. For example, Robichon et al. (1999) search for in the following way:

- Find the vector $x_i = (\varpi_i, \mu_{\alpha*i}, \mu_{\delta i})$ and the corresponding covariance matrix Σ_i for each investigated star.
- Find the vector $x_0 = (\varpi_0, \mu_{\alpha*0}, \mu_{\delta 0})$ and the corresponding covariance matrix Σ_0 for the cluster's centre.
- Assume that the depth of the cluster is zero (the whole cluster can be described by a single value ϖ_0). At the position of each star i, determine the vector $x_{0i} = (\varpi_0, \mu_{\alpha*0i}, \mu_{\delta 0i})$, the components of which can be calculated as

$$\begin{aligned} \mu_{\alpha*0i} = & \cos(\alpha_i - \alpha_0)\mu_{\alpha*0} \\ & + \sin\delta_0 \sin(\alpha_i - \alpha_0)\mu_{\delta 0} \\ & - \cos\delta_0 \sin(\alpha_i - \alpha_0)(V_{R0}\varpi_0/4.74)\,, \end{aligned} \tag{3.1}$$

$$\begin{aligned}\mu_{\delta 0i} = &- \sin\delta_i \sin(\alpha_i - \alpha_0)\mu_{\alpha*0} \\ &+ [\cos\delta_i \cos\delta_0 + \sin\delta_i \sin\delta_0 \cos(\alpha_i - \alpha_0)]\,\mu_{\delta 0} \\ &+ [\cos\delta_i \sin\delta_0 - \sin\delta_i \cos\delta_0 \cos(\alpha_i - \alpha_0)]\,(V_{R0}\varpi_0/4.74)\ ,\end{aligned} \tag{3.2}$$

where V_{R0} is the mean radial velocity of the cluster. We may notice that the equations correspond to a rotation transformation of x_i onto the model of the cluster described by x_0. Within the brackets at the end of the last term of each equation is the "radial proper motion" component, which corresponds to the transverse motion.

- Let us label the associated covariance matrix as Σ_{0i} and assume that the individual components of the matrix follow a multivariate Gaussian distribution. Star i is considered as a member of the cluster if

$$\chi^2 = (x_i - x_{0i})^T (\Sigma_i + \Sigma_{0i})^{-1} (x_i - x_{0i}) < 14.16\ , \tag{3.3}$$

  where χ^2 follows a Chi-squared distribution with three degrees of freedom.

- Check the standard error of the parallax and of proper motion for each star. Remove from the analysis all stars for which $\sigma_\varpi > 3$ mas or $\sigma_\mu > 3$ mas yr^{-1}.

- Check the distance of each star perpendicular to the line of sight. Remove from the analysis all stars for which $d_\perp > 10$ pc.

- Analyse the remainder of the stars in the CMD.

The results clearly depend on the knowledge of the radial velocities, which needed to be taken from other sources. For this reason, only 18 clusters were analysed in the study by Robichon et al. (1999). Still, the study provided very precise mean values of cluster parallaxes, proper motion and radial velocities, showing the strength of the precise determination of cluster members.

In the work by Dias et al. (2001), the membership probabilities are estimated based on the maximum likelihood procedure described in Sanders (1971). It is assumed that the distribution of stars in the examined field has the form of a bivariate Gaussian distribution, consisting of the members of the cluster (circular model, but can be generalised to an elliptical one) and of the field stars (elliptical model). The problem can be separated into a set of eight independent non-linear equations that need to be solved with the use of an iterative procedure. The solution to this problem (in the work by Dias et al. 2001) resulted in a catalogue of 112 open clusters, for which mean proper motion and mean parallax values were published. Dias et al. (2002) further extended this list of open clusters (ending up with over 1500 entries) and added additional information about their ages, angular diameters, colour excess and distances.

Next, Kharchenko et al. (2005) published an independent catalogue of clusters. They searched the All-Sky Compiled Catalogue of 2.5 Million Stars (ASCC-2.5, based on Hipparcos data release) and identified over 500 (previously known) open clusters. The membership probabilities (for detailed information regarding the procedure, see Kharchenko et al. 2004) were computed with the use of the distribution in the sky, proper motion and photometric measurements. The procedure works as follows. First, field stars are separated from the cluster members. This is done by first identifying over-densities in the sky and

defining the cluster in the proper motion diagram. The member candidates are then checked in the CMD, where outliers are removed from the list. The procedure is then repeated until the list of members at the beginning and at the end of the iteration is the same (typically achieved in the second iteration). In each iteration (with the obvious exception of the first one), the cluster is identified with the help of the projected stellar (surface) density from the previous iteration, giving an estimate of the position of the cluster's centre and of its projected radius. The membership probability of a star is calculated as a measure of the distance from the position of the cluster in the proper motion diagram or in the CMD.

Kharchenko et al. (2013) published one of the final catalogues of open clusters in this era. The catalogue contains information about 3006 objects, most of them being open clusters (and others representing stellar associations and globular clusters). Included are the distances from the Sun, reddening values $E(B-V)$, $E(J-K_S)$ and $E(J-H)$, estimates of ages and literature metallicities with the corresponding errors, and the parameters of the King's profile. It should be noted that not all of the clusters have values included for all of these quantities. The applied membership identification procedure is described in Kharchenko et al. (2012) and is similar to the one from Kharchenko et al. (2004).

Three very advanced procedures for deriving membership probabilities for kinematic groups were developed, which are being used up to this date.

- BANYAN (Malo et al. 2013): This tool is based on a Bayesian approach to the determination of the membership probabilities. It requires a kinematic model of the studied kinematic group, which consists of the vector containing means for the 3D spatial positions and velocity components, and of the corresponding covariance matrix. Furthermore, the observable quantities are analysed using a Bayesian approach. The main advantage of BANYAN is its application to almost any type of kinematic group. The authors demonstrated the use of this tool on a couple of moving groups and on several stellar associations, to which new low-mass stars were assigned.

- UPMASK (Krone-Martins & Moitinho 2014): An unsupervised tool for membership assignment in stellar clusters. It makes use of the photometric data and of the stellar positions in the sky. The only assumption used in this analysis is that cluster members are grouped together in various diagrams (most important is the positional space) and that they can be distinguished from the field population of stars, which does not cluster. An area in the sky is then analysed by an iterative process that focuses on finding the stellar over-density (taking measurement errors into account). As demonstrated by the authors on synthetic (but also on observed) clusters, the recovery rate for the members with 90 % membership probabilities is very high for various cluster parameters.

- ASteCA (Perren et al. 2015): an open-source tool designed for the analysis of star clusters. To determine the membership probabilities, stars are removed at random from the original sample. Assuming that the removed stars are mostly part of the field population, the remaining stars should resemble a non-contaminated region of the cluster. This process is repeated multiple times (1000 by default) and the final membership probability of a given star is calculated as an arithmetic mean of the individual probabilities (determined using a Bayesian approach) from each iteration.

- CLUSTERIX (Sezima et al. 2015): An online tool (a new version available at `http://clusterix.cab.inta-csic.es/clusterix/`) used to determine membership probabilities for open clusters. It utilises a non-parametric method (classical statistics) described in Balaguer-Núñez et al. (2004) and makes use of online stellar catalogues.

3.2.2 Ages and Metallicities

As mentioned above, new catalogues of clusters published in this era also often included the information about cluster ages. Piskunov et al. (2006) investigated the ages for clusters up to about 800 pc from the Sun. As was pointed out by the authors, such investigation was very much required. While there were many works focused on the cluster ages (from which formation, evolution, and lifetimes of clusters can be deduced) prior to the Hipparcos era, none of the studies had at their disposal the amount of high precision data as was available at this point. Piskunov et al. made the assumption that the cluster formation rate is time-independent in our Galaxy. The formation rate they calculated (0.23 ± 0.03 kpc^{-2} Myr^{-1}) is based on a sub-sample of 195 open clusters from their list. The resulting mean lifetime of a cluster was found to be about $\log \mathrm{Age} = 8.5$. Furthermore, it was shown that the sample of clusters is biased, showing more younger than older clusters. This was interpreted as the result of two factors. First, clusters are disrupted over time due to the dynamical evolution and the evaporation of cluster members. Second, younger clusters contain more massive and more luminous stars than their older counterparts. For this reason, such clusters can be observed at larger distances.

The advances in the studies of metallicities were much larger. This was possible due to new photometric methods being prepared and because of the spectral studies with much higher spectral resolution (and much better signal-to-noise ratio). First, let us look at the problem from the photometric perspective. Paunzen et al. (2010) derived metallicity values for 188 (88 with multiple measurements) open clusters within 4 kpc from the Sun. This was done with the use of "standard stars" (with precisely known metallicities) together with a chosen set of colour indices that are sensitive to metallicity effects. These indices were chosen for six different photometric systems. Moreover, this study took into account the extinction in the lines of sight toward individual cluster members, rather than using a mean value that can be affected by differential reddening. The final metallicity maps show that the Sun is surrounded by a "patchy structure" which may be connected with the Local Bubble (Lallement et al. 2003).

In their follow-up work, Heiter et al. (2014) showed the spectroscopic approach to deriving metallicities of open clusters. They searched the literature (over 80 publications!) for metallicity values obtained by low- and high-resolution observations, finding results for 78 open clusters. First of all, it was pointed out that metallicities obtained by different authors do not precisely agree with each other and the authors discussed the sources of the systematics (e.g. possible issues with error determination, uncertainties in stellar parameters). When compared with their photometric study, it turned out that the spectroscopic metallicities are systematically higher by about half of the value of the scatter in the data (although the spectroscopic results do agree with a much more recent study by Netopil & Paunzen 2013). They also made a preliminary analysis of the Galactic metallicity gradient, in which they highlighted the fact that the observed gradient based on open clusters does

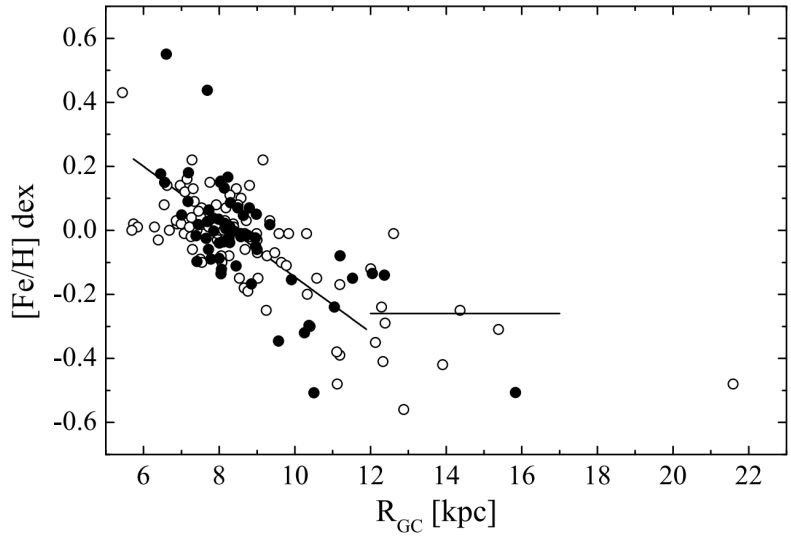


Figure 3.2: Metallicity gradient observed in our Galaxy. Taken from Netopil et al. (2016).

not agree with theoretical models.

Netopil et al. (2016) further analysed the metallicity gradient with the use of a more homogeneous set of open clusters. They found that the results agree with the more recent models relatively well. However, it is also confirmed that the metallicity gradient has a different slope in the outer parts of the Galactic disk than in the inner regions – models of the Galactic gradient appeared to disagree with the behaviour in the outer parts. Furthermore, the relation between cluster ages and metallicities displayed a positive trend. Both of these results suggested a possible effect of the radial migration of stars (from the inner towards the outer regions).

It is important to know about the oldest open clusters in our Galaxy. While the ages of the oldest globular clusters agree well with the age of the universe, open clusters are believed to be shorter-lived. However, the existence of a very old open cluster could present a challenge to the theoretical understanding of open clusters. As was mentioned above, the expected lifetime of an open cluster is of the order of ~ 100 Myr. Salaris et al. (2004) studied the ages of the oldest known clusters and found a maximum age of around $\log \mathrm{Age} = 10.0 \pm 0.1$. Taking into consideration that the disk and the halo of our Galaxy may have formed at a different time, this would suggest a delay of about 2.0 Gyr, with the halo being formed first. In this work, no significant relationship between age and metallicity of clusters could be found. Friel et al. (2002) found the same lack in the metallicity-age relation for clusters in the near vicinity of the Sun but also noticed a possible dependence on age in the outer parts of the Galactic disk (similar to what was later shown by Netopil et al. 2016).

3.2.3 Spatial and Kinematic Maps

The map of open clusters from Kharchenko et al. (2013), projected onto the Galactic plane, shows the sample of around 3000 objects. This sample is (almost) completed up to the distance of about 1.8 kpc, with a possible incompleteness present in the sub-sample of

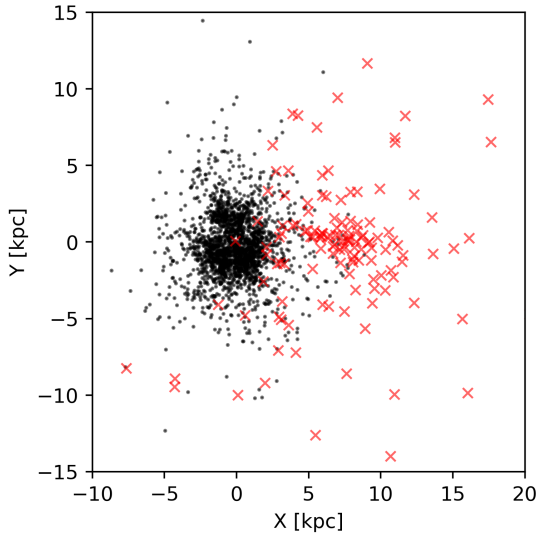


Figure 3.3: Map of the Galactic plane created using the positions of open (black dots) and globular (red crosses) clusters in our Galaxy. Based on the data from (Kharchenko et al. 2013). Sun is located at [0.0 , 0.0] and the Galactic centre can be found at the centre of the distribution of the globular clusters (approximately [8.0 , 0.0]).

the older open clusters ($\log \mathrm{Age} > 9.0$). However, if we wanted to study the details of the distribution of these clusters, the map clearly shows that the "complete" sample of clusters covers only a small part of the neighbouring inner spiral arm and barely touches the neighbouring outer spiral arm. While the authors comment that the completeness of the sub-sample allows for the studies of these arms, no analysis is presented and the map does not suggest (in an obvious way) that the distribution of clusters follows the spiral arms. The study by Bobylev & Bajkova (2014) combined the observations of young open clusters (which may be more useful for studying spiral arms) with radio observations of masers originating from young stellar objects. They were able to show that all studied objects seem to follow spiral arms very well and were able to establish the values of the pitch angle (one of the parameters used in modelling spiral arms). However, the number of included clusters (only 12) is fairly small.

Kinematic studies of open clusters allowed for new approaches. This was possible thanks to the newly prepared catalogues of clusters, measurements of radial velocities and precisely determined proper motion. For example, Wu et al. (2009) studied the motion of clusters from the Dias catalogue (see above) and were able to show that the velocity dispersion likely increases with the cluster age (although the results may be biased for the oldest clusters). Researching the velocity dispersion of various objects in the Galactic disk helps us to learn much about the disk kinematics. Moreover, Vande Putte et al. (2010) analysed orbits of the same sample of clusters and calculated the motion backwards in time. They showed that most of the orbits have small altitudes – they are concentrated close to the Galactic plane, as expected from the theory of the formation of open clusters. However, it is also discussed that four clusters (including the old cluster Melotte 66) have orbital parameters suggesting a possible extra-Galactic origin. If confirmed, such finding would

be an indication of a small population of clusters being born from a high-velocity-cloud passing through the Galactic disk. Nevertheless, it was later argued (Martinez-Medina et al. 2016) that the orbits of these clusters can also be easily explained by a formation in our Galaxy.

3.2.4 Evolution of Open Clusters

Most of the star-formation happens in giant molecular clouds, which results in the creation of star clusters. However, not all of the material is turned into stars, and the stars are not born at the exact same time (but rather within ~ 1 Myr in the case of the Orion Nebula Cluster, Palla & Stahler 1999). For these reasons, a newborn cluster is usually surrounded by the progenitor cloud, and it is labelled as an embedded cluster. Lada & Lada (2003) presented a catalogue of such clusters and researched the formation of such clusters (in the region within 2 kpc from the Sun). They concluded that ≤ 5 % of the clusters can survive this stage of evolution (with most of the remainder possibly turning into an unbound kinematic group, a stellar association). Furthermore, they show that most of the members of an embedded cluster are formed in regions with a large number of high-mass stars. For this reason, an understanding of the formation of a dense core in a giant molecular cloud is required. It is also important to understand the process of how the material inside this core is turned into high-mass stars, which appears to be different from the main process that is responsible for the formation of low-mass stars and requires a significant amount of fragmentation (and, therefore, a loss of turbulence). The authors also pointed out the large uncertainty in the knowledge of the gas-removal timescale $\tau_{\mathrm{gr}} < 5$ Myr (in which the embedded cluster sheds most of the surrounding material and disperses it into the surrounding ISM). Regarding the details surrounding the topic of star-formation, more information can be found in the works Bonnell et al. (1998) and McKee & Ostriker (2007).

Gutermuth et al. (2005) discussed how the stellar activity (e.g. stellar winds, radiation pressure) fuels further dynamical evolution of the embedded clusters, within the first few years since the stars began to form. The authors showed that some of the embedded clusters display a significant amount of asymmetry of stellar distribution when compared with the structure of the surrounding cloud. Instead, the stars appeared to be correlated with the filamentary structures detected in the clouds. Still, some of the embedded clusters (e.g. the presented NGC 7129) appear to be quite spherically symmetric and are most likely the result of a dynamical evolution.

The star-formation process is thought to be influenced by the presence of a magnetic field in the star-forming cloud. Price & Bate (2008) investigated the influence of the magnetic field on the formation of star clusters. Based on numerical simulations, they found that even relatively weak magnetic fields can severely hinder the star-formation processes. The authors discussed that the observed magnetic fields are comparable with those included in their simulations and conclude that the structures seen in their simulations are likely to be observed in molecular clouds. One of the most striking predicted features is a void of material supported by the magnetic pressure – such a void was also reported in observations of the Taurus Molecular Cloud (Goldsmith et al. 2008).

There have also been advances in the N-body simulations during the Hipparcos era. Kroupa et al. (2001) used the most realistic (at this point in history) model of the dynamical

evolution from an embedded cluster to an open cluster. They started with objects similar to Orion Nebula Cluster and assumed that the OB stars successfully dispersed 2/3 of the mass of the cloud (and the rest is contained in the member stars). The resulting cluster was very similar to Pleiades, suggesting that Orion Nebula Cluster may evolve into such a cluster. The expansion velocity of Orion Nebula Cluster is also predicted and suggested for future studies.

Remarkable improvements were made in the theory of stellar evolution, leading to the production of more realistic evolutionary tracks (isochrones). Of interest are especially:

- Geneva tracks (Lejeune & Schaerer 2001)
- Y^2 tracks (Yi et al. 2003)
- Dartmouth tracks (Dotter et al. 2008)
- Padova tracks (Bressan et al. 2012)

Many of the well-known evolutionary models were first published around the mid-1990s – the included references point to the most advanced version in the Hipparcos era. As was shown in Martins & Palacios (2013), different evolutionary tracks agree to a high degree for stars on the MS, although Geneva tracks seem to produce a MS which is too narrow. Variations appear when comparing the evolved stars, with the most significant differences seen for the cooler giants. If applied to star clusters, different tracks would likely result in different clusters parameters, affecting mostly the determination of the age and the metallicity.

It must be mentioned that many evolutionary codes were also produced. An example would be the relatively easy-to-use Modules for Experiments in Stellar Astrophysics (MESA, Paxton et al. 2011) code which is associated with Dartmouth tracks. Moreover, Padova and Trieste Stellar Evolution Code (PARSEC, Bressan et al. 2012) was published and became one of the most often applied procedures for stellar evolution associated with open clusters.

3.2.5 Beyond Our Galaxy

Andromeda Galaxy (or M 31), being a nearby large spiral galaxy, remained one of the most studied regions when searching for extra-Galactic clusters. Williams & Hodge (2001) presented an automatic method for detecting young kinematic groups in M 31 (but also useful for other nearby galaxies). One of the main limitations of this procedure is that it cannot be used for studying compact clusters. The authors were able to find 79 new candidates, for which they also derived approximate values for ages and reddening. Caldwell et al. (2009) further extended the catalogue of candidate open clusters up to 670 (with 140 young clusters). Moreover, they found that the distribution of the young kinematic groups correlates with the star-forming regions, a fact which is also supported by the analysis of the kinematic data.

Lifetimes of clusters in various galaxies have been explored. Lamers et al. (2005b) compared observations and N-body simulations of clusters in our Galaxy with the Small Magellanic Cloud, M 33 and M 51. They found that the lifetimes derived from observations

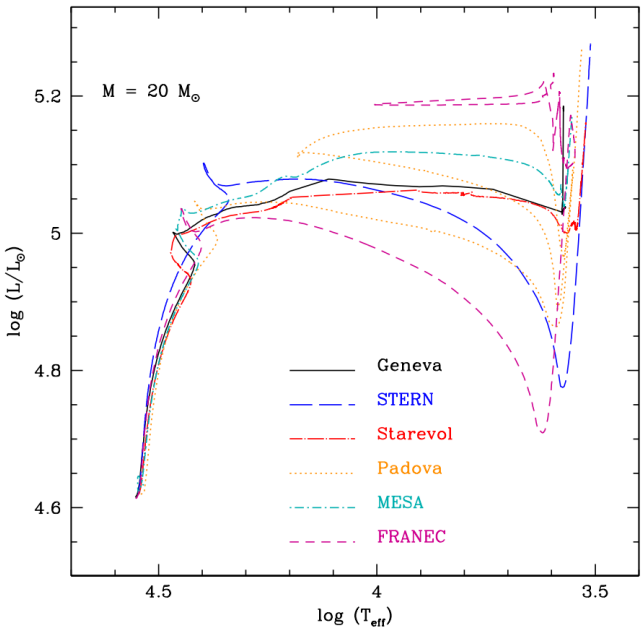


Figure 3.4: A comparison of evolutionary tracks for a single star with the mass of 20 $M_{\odot}$ (Martins & Palacios 2013). Stellar rotation was not assumed in the models.

depend on the initial masses of given clusters as $\tau_{\rm dis} \propto M_{\rm ini}^{0.60}$, agreeing with the theoretical predictions for masses between $10^3 M_{\odot}$ and $10^6 M_{\odot}$. Furthermore, it was shown that the lifetimes also depend on the mass density surrounding the cluster. The theoretical and observational predictions disagree only in the case of M 51. The authors speculate that this could be explained by tidal field variation within this galaxy.

3.2.6 Distance of Pleiades

While the problem with Hipparcos parallaxes concerns many studied clusters, the largest problem is seen with Pleiades. Soderblom et al. (1998) note that the distance derived from parallaxes lies around 120 pc, a value which is lower than in the case of the photometric measurements. The problem remains even when the errors are taken into account. In comparison, the measurements from Multichannel Astrometric Photometer and Thaw Refractor (Thaw/MAP) give a parallax consistent with the photometric measurements (Gatewood et al. 2000). A technique based on eclipsing binaries and empirical relations was used by Southworth et al. (2005) to provide a different way of determining Pleiades' distance. Their results also agree with photometry, at least within the measurement errors. Soderblom et al. (2005) also confirmed the distance of about 133 pc to be the correct value, based on the trigonometric parallaxes measured by Hubble Space Telescope's Fine Guidance Sensor.

3.3 Gaia Era (2016-2021)

For the most part, it is currently believed that the rotational (or spin) axes of stars in a cluster do not align. Theoretically, the turbulent motion within the star-forming cloud should prevent the rotational alignment. However, several studies suggest that some kind of alignment is observable. Asteroseismology shows that the alignment may survive up to several Gyr since the formation of the cluster (Corsaro et al. 2017). A possible intermediate spin-alignment was also suggested for the ≈ 0.1 Gyr old cluster NGC 2516 (Healy & McCullough 2020). It was shown by Rey-Raposo & Read (2018) that numerical simulations can predict a strong rotational alignment of cluster members. This can happen in star-forming clouds located in spiral arms but also in the inter-arm regions.

This era is also a stepping stone in gravitational wave astrophysics, following the observation of the gravitational waves produced by a binary black hole merger (Abbott et al. 2016) and from the merger of neutron stars (Abbott et al. 2017). Rastello et al. (2019) studied the evolution of a binary black hole within an open cluster consisting of 1500 stars (using N-body simulations). They suggested that clusters could be important environments for measuring gravitational waves.

Usually, open clusters are believed to host a single population of stars. Marino et al. (2018) found that CMDs of several Galactic clusters ($8.0 < \log \mathrm{Age} < 9.0$) display a broadened region near the turnoff point. The same feature was also observed in the Magellanic Clouds (e.g. Mackey & Broby Nielsen 2007). The most likely explanation for this effect is the difference in the stellar rotation. The topic was revisited by de Juan Ovelar et al. (2020), who took a sample of seven clusters, from which only one did not show the extended MS turnoff. It was concluded that this is not an effect of differential extinction. Moreover, the authors analysed the theoretical models and concluded that the inclusion of stellar rotation can explain the extended turnoff more readily than the models where stellar rotation is not taken into account. While the extended MS turnoff is believed to be a common feature among open clusters (Cordoni et al. 2018), further research may be required. For example, the study of Stock 2 by Alonso-Santiago et al. (2021) clearly shows that the rotational velocity cannot explain the extended turnoff. Instead, differential reddening seems to be more important for this cluster.

3.3.1 Membership Probabilities

With the Gaia DR2, new catalogues of open clusters were produced. Cantat-Gaudin et al. (2018) provided one of the most important catalogues in the Gaia era. In their approach, they have limited their search for clusters to $G \approx 18$ mag because of the difficulties that occur when trying to distinguish the cluster from the field stars. Next, they took the previously published cluster catalogues and prepared a list of over 3000 candidate open clusters (excluding Ursa Major moving group, Hyades and Coma Star Cluster). They searched for these clusters in the Gaia DR2 data using the UPMASK tool, which yielded a set of 1229 clusters (60 of which were newly detected objects). Furthermore, they calculated the mean proper motion and mean parallaxes of clusters using stars with membership probabilities > 50 % (excluding outliers), the projected radii of clusters (in degrees) within which 50 % of all members are located. As a secondary output, they have also calculated

cluster distances with the use of a maximum likelihood procedure (assuming Gaussian distribution). This catalogue was later expanded (Cantat-Gaudin & Anders 2020), yielding 1481 clusters. Curiously, the authors argue that many of the undetected clusters (when compared with the list of candidates) are most likely not true clusters.

At the same time, a new approach to the detection of open clusters was devised. Castro-Ginard et al. (2018) prepared a procedure that consists of the density-based clustering algorithm (DBSCAN, used for the identification of clusterings and originally proposed in Ester et al. 1996) and of an artificial neural network. With this procedure, the authors were able to find hundreds of new open clusters (Castro-Ginard et al. 2020).

It is also worth mentioning the publication by Dias et al. (2021), which is built mostly on the lists of cluster members published by Cantat-Gaudin. This catalogue includes (among other measured or calculated quantities) distances, ages, metallicities and extinction values in the *V* filter, together with the corresponding errors. The included main clusters parameters are based on the isochrone fitting technique.

The BANYAN tool received much-needed attention during this era. The procedure was updated several times, with BANYAN-Σ being the latest product (Gagné et al. 2018). This newest version is more accurate and significantly faster than the previous versions. BANYAN-Σ is freely available and was written in `IDL` and `Python 3`. Moreover, CLUSTERIX 2.0 was released after Gaia DR2 (Balaguer-Núñez et al. 2020).

3.3.2 Ages and Metallicities

As was mentioned above, Dias et al. (2021) provided a catalogue of cluster ages (and metallicities). However, it was not the only work that aimed to provide these values. Bossini et al. (2019) applied the automatic Bayesian tool BASE-9 which served to fit isochrones to Gaia CMDs. The result is a list of 269 open clusters. The authors excluded young clusters and heavily reddened clusters from their analysis. In their discussion, the cluster NGC 6793 is highlighted due to the difference in the derived age when compared with another work. They pointed out that this difference is caused by the membership probability determination – for this specific cluster, a bright star may (or may not) be included as a member, which causes a shift in the cluster age equivalent to about 150 Myr. This example clearly shows how the inclusion/exclusion of a star near the turnoff point (or in the giant branch) can have a serious impact on the age determination of a cluster.

Monteiro & Dias (2019) calculated cluster ages with the use of an advanced technique called the cross-entropy global optimization algorithm for fitting isochrones (see Monteiro et al. 2010, 2017). Their aim was to derive astrometric parameters (proper motion, parallaxes) and membership probabilities for clusters with previously unpublished values (in the old Dias catalogue) and for the newly discovered clusters by Cantat-Gaudin. About half of the candidates from the studied sample were found not to be a cluster.

It was already mentioned that photometric and spectroscopic metallicities may not always agree. This can be clearly seen in the case of NGC 6067, one of the most massive known open clusters. While spectroscopic measurements show the metallicity value of $[\mathrm{Fe/H}] = +0.19 \pm 0.05$ from Alonso-Santiago et al. (2017), photometric measurements point to a lower value. Dias et al. (2021) provide $[\mathrm{Fe/H}] = +0.089 \pm 0.042$ which does not agree with the spectroscopic measurements. Apparently, the discrepancy in cluster

metallicities derived by different approaches has not yet been resolved.

An important step forward regarding the classification of cluster ages can be made by using star clusters as age-calibrators. Based on the combination of the Gaia data with (ground-based) spectroscopic observations, Gaia-ESO Survey aimed at exploring such calibrations (Randich et al. 2018). The cluster ages determined from Gaia data (isochrone fitting) were compared with the lithium depletion boundary ages and were found to agree with each other (except for one cluster). Assuming that a precise enough relation can be established between the two quantities, the lithium depletion boundary could be (eventually) used to derive the ages of field stars.

3.3.3 Spatial and Kinematic Maps

It was mentioned in the previous era that maps of open clusters appear to be complete up to about 1.8 kpc, with the exception of the older clusters. However, Cantat-Gaudin et al. (2019a) found that this may not be completely true. The authors were able to identify new cluster candidates in the direction of the constellation Perseus, which increased the number of known clusters by about 20 % in the given region of our Galaxy. If confirmed as open clusters, this would clearly show that the sample of observed clusters is incomplete even in the near vicinity of the Sun.

Kuhn et al. (2019) studied the kinematic properties of young ($\log \mathrm{Age} < 6.7$) open clusters and stellar associations. Only on the basis of the proper motion, it was found that most of these objects seem to be expanding (consistent with the idea that clusters evolve and dissolve), with a typical expansion velocity being ~ 0.5 km s^{-1}. Rotation was detected only for one of the studied objects. It is argued that the expansion is less noticeable for these objects because most of them are still embedded in the progenitor molecular clouds.

Kinematic maps have also confirmed the presence of a large-scale warp of the disk of our Galaxy. Poggio et al. (2018) have shown this for the upper MS stars and giants. The gradient in the vertical component of the velocity vector was found to be about 5 km s^{-1} at the Galacto-centric radii from 8 kpc to 14 kpc. This coincides with the earlier results published by Kawata et al. (2018).

Cantat-Gaudin et al. (2018) used their catalogue of open clusters (Gaia DR2) to study their distribution in the Galactic disk. When compared with the models of spiral arms, their map shows that the structure of the spiral arms can be traced by the open clusters. Still, this structure was very difficult to identify on the map. The number of clusters has increased in Cantat-Gaudin et al. (2020), who showed that the local arm can be clearly traced by relatively younger clusters. However, the presence of the outer arm could still not be confirmed with such objects. The situation has not improved much with Gaia EDR3. The maps in Castro-Ginard et al. (2021) are still rather unconvincing in regions outside of the local arm. Still, important knowledge was gained from the results of this study. Most of the previous studies of open clusters supported the idea that spiral arms are long-lived structures. In contrast, Castro-Ginard et al. (2021) provide evidence that this may not be the case – instead, the transient nature of the spiral arms is suggested.

It should be pointed out that 3D analysis of a cluster is still very difficult even with the Gaia data. Lodieu et al. (2019) studied three nearby clusters (α Persei, Pleiades, Praesepe). While they were able to reproduce 3D maps of these objects, some problems accompany

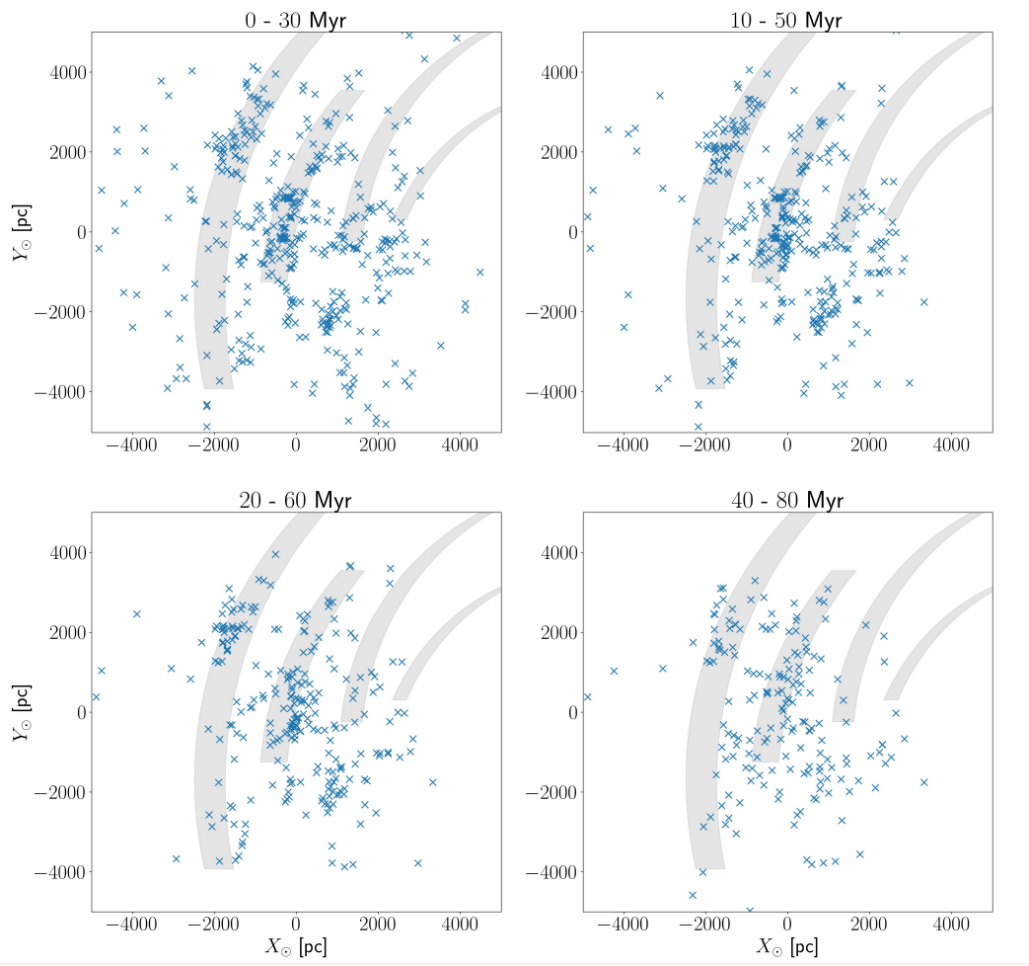


Figure 3.5: Comparison of the distribution of open clusters (grouped into ranges of their age) with the positions of nearby spiral arms. Taken from Castro-Ginard et al. (2021).

their analysis. For example, the authors note the existence of a stream of stars in the Pleiades cluster and discuss its origin. However, they failed to notice that this streaming of stars is happening along the line of sight. The same can be seen in the case of Praesepe. Clearly, the origin of this effect must be in a bias that has not been accounted for.

3.3.4 Evolution of Open Clusters

The tidal tails have been of great interest during the Gaia era. Meingast & Alves (2019) and Röser et al. (2019) have simultaneously published papers in which the tidal tails of the Hyades cluster are revealed. These (leading and trailing) tails are expected to be extended up to hundreds of parsecs away from the clusters, although the current sizes are measured to be around 100 pc. Asymmetry in the two tails was also detected, which was later confirmed by Jerabkova et al. (2021) with the use of Gaia EDR3. All of this information can provide new constraints for models of cluster evolution.

It was mentioned above that Gaia measurements should provide reliable information about the expansion of the Orion Nebula Cluster. However, the situation regarding this

topic has not yet been resolved. While Kuhn et al. (2019) report a small expansion rate for the cluster, Kim et al. (2019) conclude that there is no observable expansion of the cluster. The discrepancy in these results can be explained by an underestimated error in the expansion rate derived in the former work.

PARSEC code (and the evolutionary tracks) has received several updates since the launch of the Gaia satellite. The evolutionary tracks for various user-picked parameters can be generated using an online tool via `http://stev.oapd.inaf.it/cgi-bin/cmd` for almost any photometric system. This greatly simplifies the task of fitting isochrones, since the results are usually returned very quickly. One of the disadvantages of using this tool is the inability to alter stellar rotation in the models.

Darma et al. (2021) studied the formation of binary clusters. Based on N-body simulations, they found that such objects can originate from the (early) evolution of the star-forming region that contains a higher degree of initial sub-structure. The number of binary clusters decreases with the age of the system – it is much less likely to find old binary clusters due to the dynamical evolution of the system (mergers, dissolution).

There were also attempts to study the evolution of different observed cluster parameters. In the works by Angelo et al. (2020, 2021), the relation between the cluster's age and the concentration of stars in the cluster was analysed. While their plot of the cluster concentration radius versus the dynamical time suggests that the concentration increases with time, there appears to be no reasonable correlation between the cluster's age and the concentration measure. However, it is clearly shown that the tidal radius increases with the Galacto-centric distance (if we ignore the more concentrated clusters).

3.3.5 Beyond Our Galaxy

With the use of the Gaia DR2 data, Price-Whelan et al. (2019) discovered a young kinematic group in the Galactic halo. The authors discuss that the location of the group is fairly close to the gas stream originating from Magellanic Clouds. It is also speculated that the group may have been born from a star-forming event that was triggered by the passage of the stream through the disk of our Galaxy.

3.3.6 Distance of Pleiades

The problem with the distance of Pleiades has been finally settled by Gaia. Galli et al. (2017) reported a distance of 134 pc, which is consistent with most observations of Pleiades. The inconsistency in the distance of this cluster was clearly caused only by unknown issues with the determination of the trigonometric parallaxes by the Hipparcos satellite.

3.4 Concluding Remarks

We can infer from the cited works that spiral arms are still difficult to trace with open clusters. This represents a small problem, because we assume that most of the star-formation is happening in spiral arms and that most of the stars are born as a part of a kinematic group. One of the main causes is the inability to observe clusters in (and behind) dense regions. Secondly, the orbits of open clusters should evolve similar to the stellar orbits.

This motion is, in general, different from the "motion" of the spiral arms, which can be represented by travelling density waves. For this reason, open clusters should not stay in or near the arm of origin.

Next, it is obvious (and was pointed out in multiple works) that distances of cluster obtained from Gaia measurements are accompanied by several problems. The Lutz-Kelker bias is one of those issues. Most of these problems originate from having different measurements with different relative errors. For example, stellar positions are usually determined much more precisely when compared with parallaxes. This will result in a good interpretation of the cluster's projected shape, but it may also lead to a problematic conclusion about its shape in the line of sight. The outcome is that many researchers will be able to find a stream of stars in the radial direction of an open cluster (assuming that the typical relative parallax uncertainty for the members is high enough), although these streams are (usually) not of physical origin. Furthermore, this will affect the ability to search for tidal tails in the 3D space. A question arises – since the issue is mainly in the determination of the distance from the parallax, can we fix the problem by approaching the problem from the Bayesian perspective?

Finally, binary clusters represent a very hot topic in the study of the formation of clusters. However, it is possible to detect two clusters grouped together in the sky, although both objects do not interact with each other. This will be the case for clusters located at different distances. Moreover, if we detected clusters which are located, for example, 20 pc from each other, can they be classified as a binary system? Can we get more information about such objects using proper motion data and CMDs? Are there any unidentified binary clusters in the catalogues based on the Gaia data?

Chapter 4

Interstellar Medium and DIBs

ISM is an important part of most galaxies. Usually, one will consider a simple picture of the ISM in the form of the McKee-Ostriker model (McKee & Ostriker 1977). Most of the volume of a galaxy is dominated by a medium in which almost all of the hydrogen is in an ionised state, due to the interaction with stellar activity in the host galaxy (mostly dominated by supernovae). Pockets of over-densities can form during the recombination cooling in this very hot and ionised medium, which represent the neutral parts of the medium, often in the form of clouds. Furthermore, the surrounding hot medium is a source of high-energy particles which ionise the outer layers of these pockets. Hence, we should observe a continuous transition from the neutral to the fully ionised medium. In the clouds, the density increases and the temperature decreases towards the centre. Molecules can start forming if the density in the cloud is high enough – we label these objects as molecular clouds.

Of course, this image of the universe is oversimplified (e.g., Heiles 2001). First of all, a single supernova will not shock the whole volume of the host galaxy. Instead, only the surrounding medium will be affected, creating an object we call a superbubble (Mac Low & McCray 1988). Moreover, the ionised state of a galaxy is also affected by the inbound high-energy particles from the nearby galaxies. Furthermore, even the denser parts of the ISM seem to contain hotter and ionised sub-structures in the form of H II regions or "worms/chimneys" (Heiles 1984; Li & Ikeuchi 1990; Rand et al. 1990).

So far, only the gas-phase component of the ISM has been examined. To get a better understanding of a galaxy, the interstellar dust must also be studied. Not only does the dust affect our visual observations, but it is also believed to be one of the main catalysts for molecular formation (e.g. Goodman 1978; Katz et al. 1999). The structure of the dust particles (called grains) seems to tie together the observed polarisation of the light with the global and local magnetic fields (Davis & Greenstein 1951). This is important due to the fact that magnetic fields have an impact on the star-formation (Mestel & Spitzer 1956; Nakano 1998; Hennebelle & Inutsuka 2019).

As was highlighted above, other particles (such as cosmic rays, photons, magnetic fields, gravitational waves, ...) are also of interest. However, these are not reviewed here.

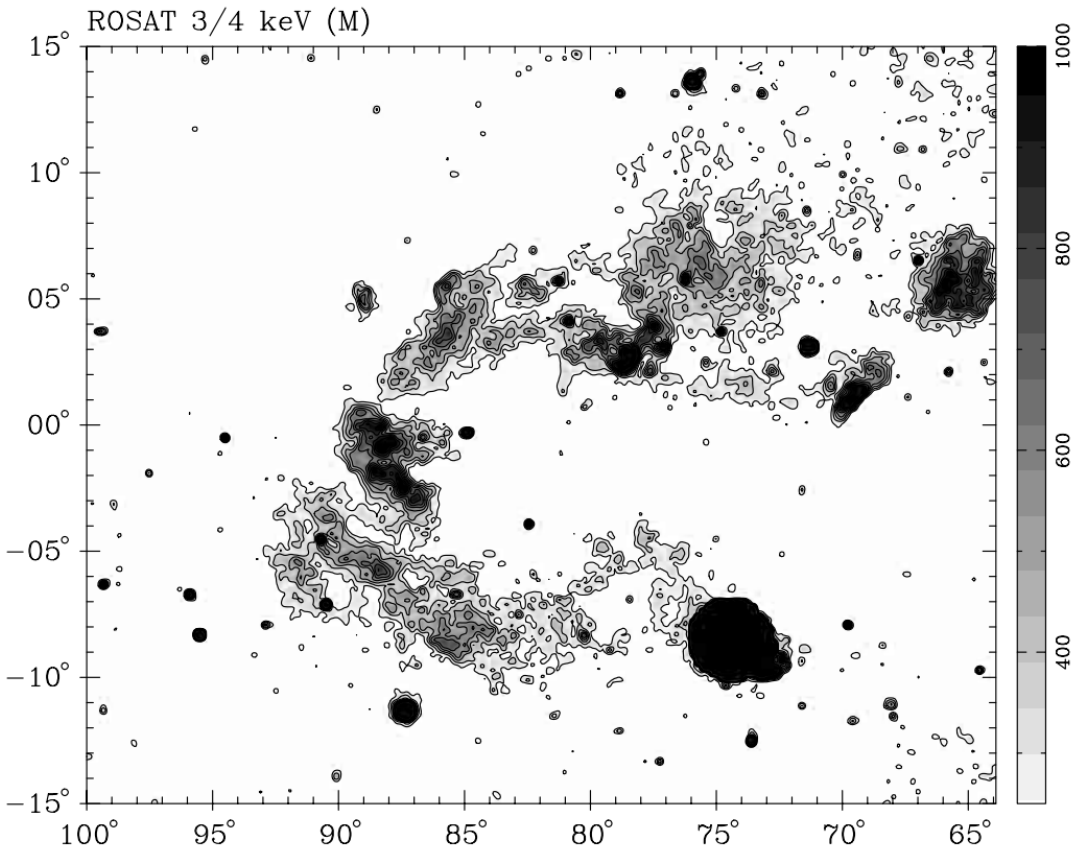


Figure 4.1: The Cygnus superbubble observed in X-rays (ROSAT 3/4 keV, Uyanıker et al. 2001) contains medium at the temperature of about 10^6 K.

4.1 Phases and Tracers of the ISM

It is obvious from the observations that the ISM consists of multiple phases. Here, the ISM phases in our Galaxy are distinguished in the same way as in Draine (2011) and the Galactic structure is followed from the outer parts of the Galactic disk towards the star-forming regions. The importance of various ISM phases in the astrophysical context and the observational tracers are discussed. It should be noted that the discussed structures are not static but evolve in time.

4.1.1 Hot Ionised Medium

The structure of the fully ionised medium depends on the location in the Galaxy. In the outer regions of our Galaxy (Galactic halo), the temperatures reach $\sim 10^5 - 10^6$ K and the densities are very low, with only $\sim 10^{-3}$ hydrogen atoms per cubic centimetre (Chevalier & Oegerle 1979). These outer regions appear to be often connected via tunnels through the neutral regions in the disk with the superbubbles, in which the density and the temperature of the medium is comparable with the Galactic halo.

Typically, the hot ionised medium is believed to be in a quasi-equilibrium state. However, superbubbles tend to expand (at some point adiabatically) which leads to cooling.

The supernovae are the source of energy and heat the medium to the cited temperatures. Hot ionised medium is a good source of X-ray radiation (Tanaka & Bleeker 1977) which also has an impact on the cooling of the medium.

Collisional interactions with the shock wave on the edge of a supernova ionise the hydrogen and other atoms. Molecules, such as oxygen, can lose up to several electrons and become good tracers of this type of ISM. For example, Jenkins (1978) studied hot regions of our Galaxy with O VI absorption lines in the UV. Of course, the mentioned X-ray radiation can also be observed and used for studying these parts of our Galaxy (but also of other galaxies, see Tüllmann et al. 2006). Finally, relativistic electrons present in the hot ionised medium can be the source of synchrotron emission at radio wavelengths (Vlasov et al. 2016).

4.1.2 Warm Ionised Medium

Some of the fully ionised regions of galaxies cool over time. When the temperature reaches about 10^4 K with the density of $n_{\mathrm{H}} \sim 0.1$ cm^{-3} (Haffner et al. 2009), we label the medium as a warm ionised medium. In a classical model of the structure of a galaxy, this type of medium represents a layer between the neutral medium in the disk and the hot ionised medium in the galactic halo. Its existence can be probed in several ways. For example, if the interstellar medium consists largely (in terms of volume) of an ionised medium, a significant number of free electrons should be present. These electrons will interact with a radio wave travelling through the medium. If a radio pulse is sent from a source to an observer, the shape of the observed pulse will be broadened due to the scattering on the free electron. Indeed, this effect was observed (e.g. Gomez-Gonzalez & Guelin 1974) and serves as one of the proves that warm ionised medium fills a large volume of our Galaxy.

At present, one of the main tracers of this part of the ISM is the emission in H_α resulting from the recombination of hydrogen with the free electrons. Fine-structure lines in the far-IR (Petuchowski & Bennett 1993) can also serve as tracers. H_α has been used to study our but also other galaxies (Dettmar 1990). This and forbidden emission lines from other atoms (see Sembach et al. 2000) serve as one of the primary cooling processes. Another important cooling process is the free-free emission (Dong & Draine 2011).

To balance the energy lost via cooling, the warm ionised medium is heated by photo-ionisation/recombination processes (Haffner et al. 2009). Reynolds & Cox (1992) studied the possibility of dust grains being another heating component in the medium. While the photo-electrons liberated from the surface of a dust grain play an important role, they serve only as a secondary source of heating.

Unlike in the case of the hot ionised medium, the warm ionised medium can be found in two different states characterised by the density and the source of ionisation. Medium ionised from the outer layers of the Galactic disk (X-ray radiation, see above) tends to have densities of about $\sim 10^{-1}$ cm^{-3}. However, hot O- and B-type stars can also be very effective in ionising the surrounding medium. These so-called H II regions tend to be much denser, reaching $n_{\mathrm{H}} \sim 10^4$ cm^{-3} (or even higher values, Habing & Israel 1979).

Figure 4.2: Composite image of the Orion Nebula (M 42) captured by the Hubble Space Telescope. We can see several phases of the ISM in the image. The red colour corresponds to the emission from H II regions (H_α), while the blue colour of the nebula is explained by the reflection of starlight caused by dust grains (small dust particles scatter more light at shorter wavelengths). The dark regions are the result of the absorption of light (by dust) in these lines of sight – this traces the very dense molecular regions where new stars are being formed. Image credit: NASA, ESA, M. Robberto (Space Telescope Science Institute/ESA) and the Hubble Space Telescope Orion Treasury Project Team (`https://esahubble.org/news/heic0601/`).

4.1.3 Cold and Warm Neutral Medium

Most of the neutral medium is observed using the 21-cm hydrogen line (e.g. Kanekar et al. 2003). The heating is dominated by the photo-electrons from dust grains (see above). Small ionisation fraction is being kept by the interaction with the interstellar radiation field and with the cosmic rays (Wolfire et al. 2003).

The density and the temperature of the neutral medium can vary considerably. The regions with $T \sim 10^3$ K have usually smaller densities ($n_{\mathrm{H}} \sim 1$ cm^{-3}) and are labelled as the warm neutral medium. They can also be observed in UV absorption lines from neutral atoms (Jenkins 2013). On the other hand, the neutral medium closer to $T \sim 100$ K tends to be much denser ($n_{\mathrm{H}} \sim 10^1$ cm^{-3}) – we call this state the cold neutral medium.

Both types of neutral medium can be found in the form of clouds and layers (warm around the cold, cold around molecular clouds, see for example Vázquez-Semadeni et al. 2006). The clouds are of great interest since they represent larger volume of the Galaxy than the molecular clouds and they do not tend to affect our observations as much. Effectively, one can observe stars through several clouds. This can be done, for example, by studying the interstellar sodium absorption lines near 5890 Å (Ferlet et al. 1985). These can be used to study distances toward stars (Munari & Zwitter 1997). Furthermore, we have also spectral evidence of multiple clouds in a line of sight (Hobbs 1969). The apparent splitting (or overlap) of the same line is caused by the presence of two or more clouds (with different radial velocities) in the direction towards the observed object.

4.1.4 Molecular Clouds

Enveloped in the cold neutral medium layers, molecular clouds tend to be shielded from the interstellar UV radiation field. Due to the high density of $n_{\mathrm{H}} > 100$ cm^{-3} ($T \sim 10^1$ K), the molecules can start forming and survive in this environment. An overview of various molecular species observed in the ISM can be found in Tielens (2013).

One of the most important tracers (but an efficient cooling source) of molecular clouds is the emission in the rotational line ($J = 1 \rightarrow 0$) of CO at 2.6 mm. It is often used as a detector of the denser parts of the ISM and was shown to be able to map the spiral structure of galaxies (Ryden & Stark 1986; Planesas et al. 1991). Another important indicator of the molecular clouds is the dust – a characteristic feature of the dust is the emission in the IR part of the spectrum (Telesco & Harper 1980; Li & Draine 2001). Furthermore, H_2O masers are often used to probe the star-forming regions in molecular clouds (Genzel & Downes 1977; Wouterloot & Walmsley 1986).

The importance of molecular clouds in star-formation was mentioned already in the text above. For this reason, further discussion is omitted in this section.

4.2 Interstellar Dust and Extinction Curves

Some of the effects of interstellar dust on our observations and its importance in the formation of interstellar molecules has been already mentioned. Besides the observational reasons, dust is thought to be a key ingredient in the formation of exoplanets (Williams &

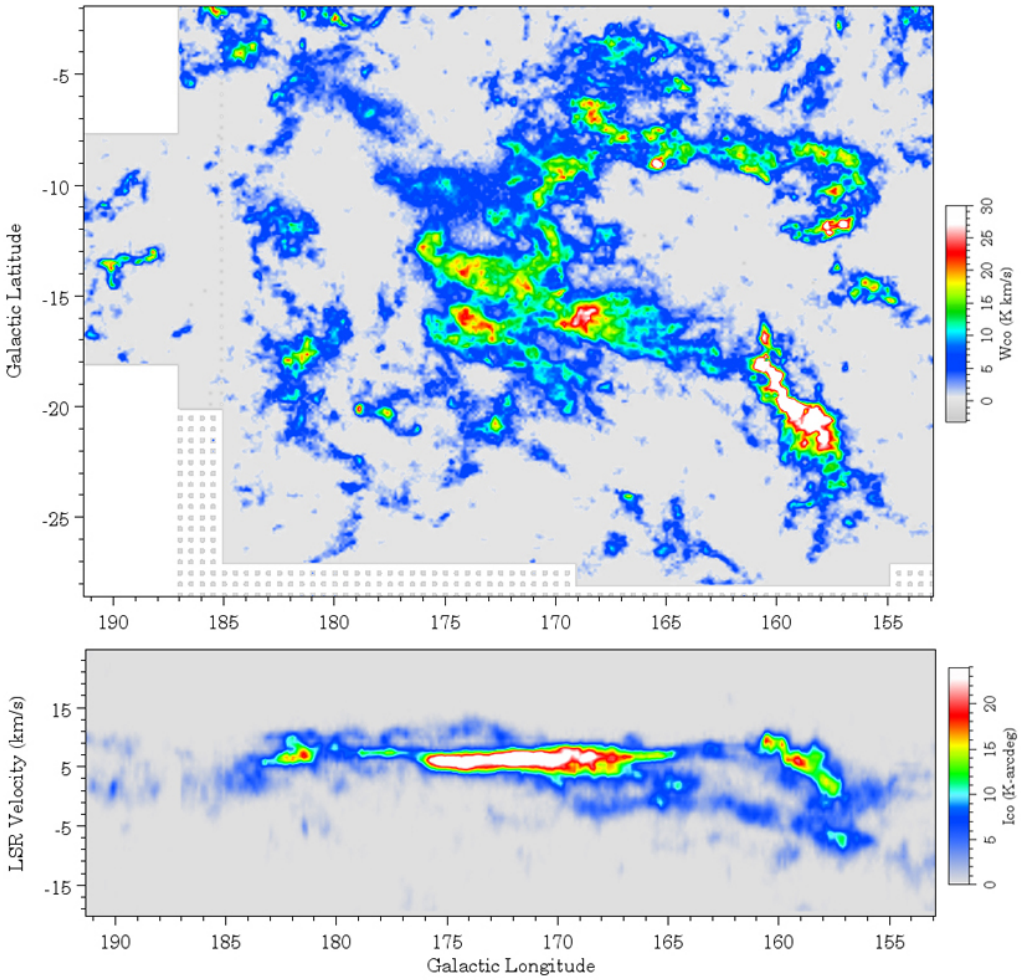


Figure 4.3: The ^{12}CO map of the molecular clouds in the Taurus region. We can see several prominent clouds. 1) The structure in the upper-right corner of the map (155 deg $< l <$ 170 deg, -14 deg $< b < -5$ deg) represents the California Molecular Cloud. 2) The bright concentration in the lower-right corner is the nearby Perseus Molecular Cloud, which contains two nearby open clusters IC 348 and NGC 1333. Finally, 3) the central parts of the image show the Taurus Molecular Cloud, which is believed to be the closest observed molecular cloud ($d \sim 140$ pc). Originally published by Dame et al. (2001), the image was taken from Dame et al. (2011).

Cieza 2011). Clearly, understanding the properties of the dust is very important for multiple fields of research.

It is believed that the dust originates from the supernovae (Todini & Ferrara 2001; Nozawa et al. 2003; Bianchi & Schneider 2007). However, the dust will not survive for longer than ~ 100 Myr in the ISM (Jones et al. 1994). According to Jones & Nuth (2011), at least some amount of dust needs to be recycled within the ISM, probably on the surface of the surviving grains.

4.2.1 Observational Evidence

Extinction

Due to the presence of small (typically sub-micron) solid particles in the ISM, the light from the stars that we observe tends to get dimmer. This will affect our ability to observe objects through thick layers of dense ISM. In the optical part of the spectrum, two quantities are usually used as measures of the attenuation of light. The first one is the interstellar reddening (or colour excess) $E(B-V)$ which tells us how many of the higher-energy photons are absorbed and re-emitted at longer wavelengths. Secondly, we can directly describe how much of the light (at a given wavelength or in a given photometric filter) has been absorbed – we call this measure the interstellar extinction A_λ. Both quantities have units of magnitudes and are related via relations $E(B-V) = A_B - A_V$ and $A_V = R_V\, E(B-V)$. In a typical line of sight, $R_V \approx 3.1$ (Schultz & Wiemer 1975), although this value can get higher near molecular clouds (where also the grain sizes can get larger).

Of course, a single observation tells us nothing about the stellar spectrum and the spectral effect of the dust. Instead, we see both components combined together. To measure any of the attenuation measures, one has to disentangle the two from each other. This can be done in two ways – the spectrum of the observed star is compared either with a model spectrum of such star (extinction-without-standards method, see Fitzpatrick & Massa 2005; Fitzpatrick et al. 2019) or with another (almost unreddened) star of the same spectral type (pair method, see Aiello et al. 1988; Fitzpatrick & Massa 1990). In both cases, the spectral type of the observed star has to be known in advance. The final result is the extinction curve, which describes the interstellar extinction as a function of the wavelength.

The knowledge of the extinction curves is very important for spectral and photometric studies. Let us focus on the studies of open clusters. Let us assume a transformation such that $A_\lambda = k_\lambda A_V$, where k_λ is a constant factor which depends on the wavelength (up to a scaling factor equivalent to the relationship between the extinction and the reddening mentioned above). Such transformation is often used in the Johnson or even 2MASS photometric systems (see McCall 2004). However, the bandwidth of a filter may represent a problem if the wavelength coverage is too large. An example of such a case is Gaia photometry. The passbands are broad enough to capture the information about the change in the extinction (as a function of wavelength). For this reason, a simple constant factor R_λ or k_λ will not be sufficient. It is preferred to use the empirical relations between k_λ, A_V and the Gaia photometric colour $B_P - R_P$ (Gaia Collaboration et al. 2018a):

$$\begin{aligned} k_\lambda = {} & c_1 + c_2\,(B_P - R_P) + c_3\,(B_P - R_P)^2 + c_4\,(B_P - R_P)^3 \\ & + c_5 A_V + c_6 A_V^2 + c_7 A_V\,(B_P - R_P)\,, \end{aligned} \tag{4.1}$$

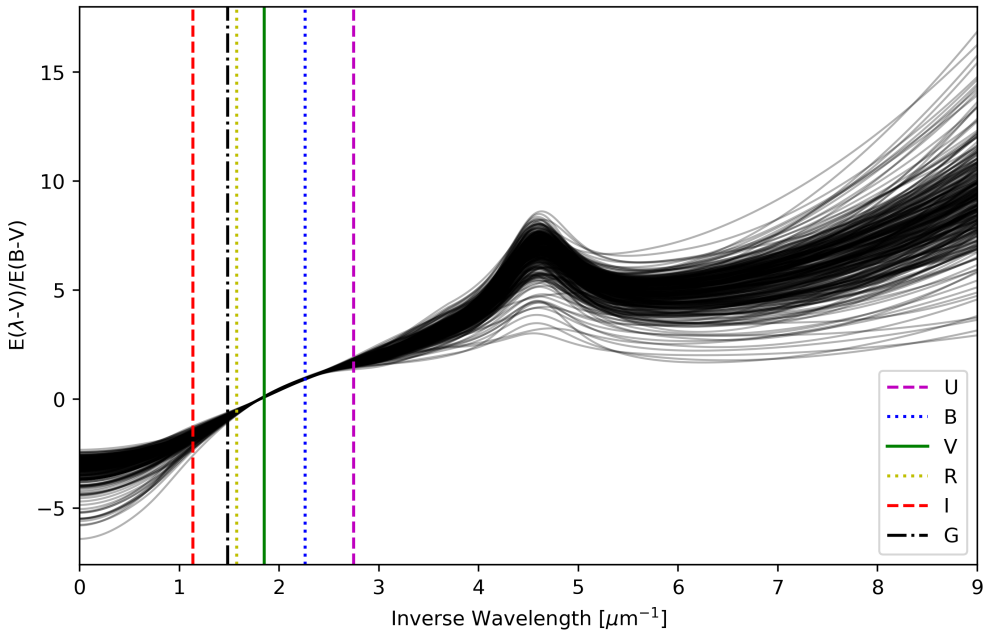


Figure 4.4: Interstellar curves modelled for 312 stars. Displayed are also the central positions of several photometric filters. Apparent is the absorption bump at ~ 2175 Å, believed to be produced by small carbonaceous dust grains. Based on the results from Fitzpatrick & Massa (2007).

where the coefficients c_i depend on λ – values were calculated for Gaia's passbands G, B_P and R_P.

Overall, the interstellar extinction is the result of the interaction of light waves with the dust grains. The electronic structure of these particles will scatter and absorb the light. The evidence for the scattering effect can be seen in the colour of the reflection nebulae (Henyey & Greenstein 1938).

Polarisation

Hiltner (1949) provided possibly the earliest observational evidence for the polarisation of stellar light. It was shown that this polarisation occurs at distant stars but was not observed for unreddened stars. Gold (1952) speculated that anisotropy of the ISM must be the source of this effect. While it was first believed that the polarisation is independent of the wavelength, it was soon found that this is not the case. Davis & Greenstein (1951) discussed the wavelength-dependence and also pointed out the possibility that polarisation originates from interstellar dust grains which have elongated shapes and are (at least partially) aligned with the interstellar magnetic field.

The relation between polarisation and wavelength was further studied by Serkowski et al. (1975) who found an empirical relation (today called the Serkowski law) between the two quantities in typical regions of the ISM:

$$p_\lambda = p_{\max} \exp\left(-K \ln^2\left(\frac{\lambda_{\max}}{\lambda}\right)\right), \tag{4.2}$$

where $p_{\max}$ is the maximum polarisation observed at wavelength $\lambda_{\max}$ and K is a scaling factor which depends linearly on the wavelength (Wilking et al. 1980). Typically, the maximum amount of polarisation is observed near $\lambda_{\max} \approx 0.5\ \mu$m. This law seems to hold in the wavelength range of about $0.3\ \mu\text{m} < \lambda < 3.0\ \mu$m.

Magnetic fields play an important role in many aspects of astrophysics. As such, it is important to study the strength and the orientation of the magnetic field of a galaxy. For example, Leroy (1999) showed that the local magnetic field is far from being uniform or smooth. Investigating small-scale structures of the Galactic magnetic field may give us some information about star-forming regions and molecular clouds (Vrba et al. 1976; Planck Collaboration et al. 2016). Constructing large-scale maps such as those provided by Planck Collaboration et al. (2015) can give information about the large-scale Galactic magnetic field – the observed structure of the global field is quite complex, as can be already extrapolated from the small-scale observations. It is also worth mentioning that it is possible to study the structure of the magnetic fields of other galaxies. Comparison with the CO observations of spiral arms can inform us about the mechanisms that lead to the formation of giant molecular clouds (Patrikeev et al. 2006).

One of the long-standing problems regarding magnetic fields and interstellar dust is concerning the mechanism which is responsible for the alignment of the grains. Two of the conclusions that can be reached by all researchers are that the process must involve the interaction between the spin of the grains and the magnetic field, and that the long axis of a grain is being aligned perpendicular to the magnetic field vector. However, different works do not agree on the specifics:

- Davis-Greenstein mechanism (Davis & Greenstein 1951; Roberge & Lazarian 1999; Weingartner et al. 2021): This model has been among the most favoured explanations for the grain alignment for a long time. The main idea behind the mechanism is the paramagnetic dissipation of energy in a rotating grain – unpaired electrons can be oriented in the direction of the external magnetic field. Assuming that the magnetic field vector is not precisely parallel with the rotation axis, the magnetisation of a paramagnetic grain will change over time. This will lead to the dissipation of energy and to the (partial) alignment of the grain with the magnetic field. Since the theory was proposed, it was shown over time that it can only explain the alignment of smaller grains.

- Radiative alignment torque theory (Dolginov & Mitrofanov 1976; Lazarian 2007; Andersson et al. 2015): One of the most successful theories, capable of providing the most precise predictions for the grain alignment when compared with the observational data. The radiation flux transfers momentum to the grain and since the actual shape of the grain is irregular, this will result in a non-zero net torque, spinning up the grain. This model of spin-alignment is most capable of explaining the alignment of the larger-sized ($> 0.1\ \mu$m) dust grains.

- Other suggested processes: See the review by Lazarian (2003).

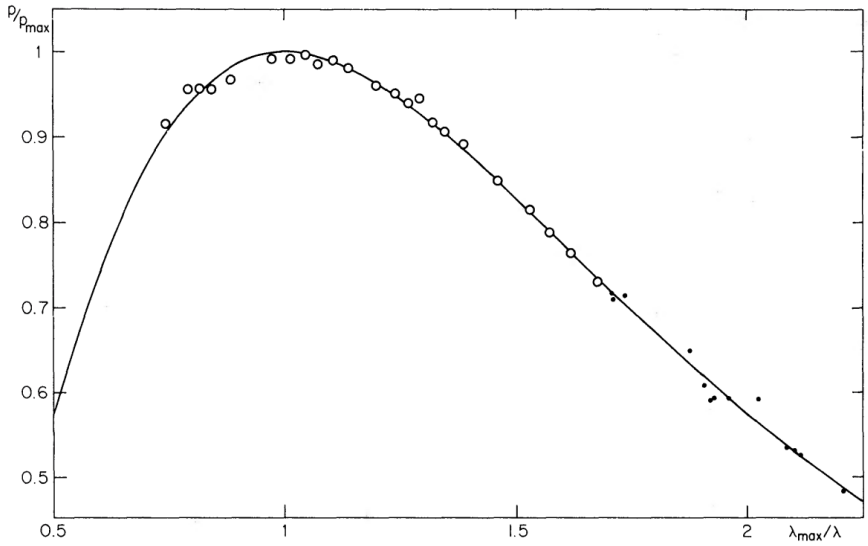


Figure 4.5: Observed polarisation of light produced by elongated dust grains and the corresponding analytical fit (Serkowski et al. 1975).

Depletion of Elements

Another very important observational fact is that some interstellar elements (e.g., Fe, Si, C) seem to be under-abundant in the gas phase (see Field 1974; Savage & Sembach 1996). This depletion can be explained by the accretion of the material onto dust grains as proposed by Weingartner & Draine (1999). It should be noted that this leads to the change in grain population and size distribution. To find whether this really can explain the depletion from the gas phase, the destruction and accretion of the material by the grains needs to be well described. However, since interstellar conditions are hard to recreate in laboratories, our understanding of these processes is still quite poor.

X-ray Halos

Interstellar dust is also responsible for the X-ray halo effect around distant X-ray sources with sightlines obscured by the dust clouds. Similar to the extinction, this is the result of the scattering by the grains with only small characteristic scattering angles which simplifies the work when calculating successive scattering (Overbeck 1965). This can be used for testing the theoretical models of the grains as was mentioned by Draine (2003).

Meteorites

Unlike many aspects in astrophysics, the dust grains may be studied directly by investigating meteorites (e.g. Anders & Zinner 1993). Whether the studied material originates from the interstellar medium or our Solar System can be determined by looking at the isotopic ratios of elements that vary from place to place across the Galaxy. Most of the grains found in the meteorites show sizes < 1.0 μm and the material they are made of seems to originate from the atmospheres of evolved stars (Mostefaoui & Hoppe 2004). Furthermore, the orbit of the Sun (around the Galaxy) passes through the local interstellar medium, which can

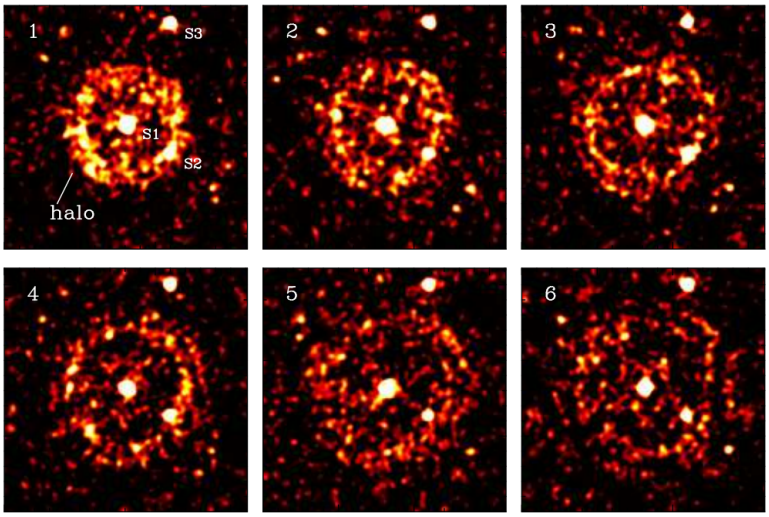


Figure 4.6: X-ray halo around the gamma-ray burst GRB 031203 (Vaughan et al. 2004). The ring-like structure is produced by layers of dust in the line of sight between us (the observers) and the source of the burst (located in a distant galaxy).

also be used to study dust grains (Frisch et al. 1999). In this case, the grain particles were detected as part of the interstellar wind by several spacecraft: Ulysses (Gruen et al. 1992), Galileo (Altobelli et al. 2005), Stardust (Westphal et al. 2014) or Cassini (Altobelli et al. 2016).

4.2.2 Structure of Dust Grains

Chemical Composition

Our understanding of the chemical properties of dust grains is based mostly on spectroscopic observations. There are several absorption bands present in the NIR and MIR, as well as the absorption bump at 2175 Å. But how do we know that they are related to grains?

Interstellar dust grains have to be made of matter for which we have observational data which constrain the maximum abundance of a given element, therefore only the more abundant species can be important for the chemical composition, such as hydrogen and carbon. In fact, it is almost surely the electronic transition of C−C bonds that leads to the formation of the absorption bump (e.g. Mathis 1994) but the specific source is still unknown. However, because of the strength of this feature, it is obvious that carbon makes a significant fraction of the dust.

The idea that carbon is very important when dealing with the interstellar extinction is also supported by the detection of emission features corresponding to the polycyclic aromatic hydrocarbons (PAHs) in the observed spectra. The most apparent features are located at 3.3, 6.2, 7.7, 8.6 and 11.3 μm and were first associated with PAHs by Leger & Puget (1984) who studied the laboratory IR emission of coronene which resulted from UV irradiation. Unfortunately, to this date, no specific PAH has been identified in the ISM.

Another example of a spectral feature pointing to the carbon is the 3.4 μm feature. Sandford et al. (1991) showed that the O−H or C−H stretching vibrations are present

at ~ 3.0 μm and ~ 3.4 μm, respectively. It is natural, that the mentioned interstellar absorption feature is usually attributed to the C−H bonds. According to Sandford et al., aliphatic hydrocarbons are the most likely candidates. They point out that the 3.4 μm feature is quite similar to the laboratory observed features of irradiated ice mantles.

The presence of the 3.1 μm feature in the spectra of the ISM is generally associated with the O−H bonds of H_2O molecules (Draine 2003). This suggests that the mantles of grains are at least partially coated by ice. Weaker features of other ice-forming molecules can also be detected, like CO (see Tielens et al. 1991). Whittet et al. (1988) studied the Taurus dark clouds complex and the relation between the strength of this feature and A_V absorption. They found that the ice coating is only present if $A_V > 3.3 \pm 0.1$ mag – it appears that ice mantles cannot exist outside of the regions of the clouds shielded from the radiation. Ice features present at ~ 4.6 μm, originating from CO and unidentified species with C=N bonds, have been used for measuring polarisation (Chrysostomou et al. 1996) but the fact that they overlap may cause difficulties when calculating abundances of individual species.

Silicon is undoubtedly the most important part of the interstellar dust, alongside carbon. The very strong absorption band located at 9.7 μm is commonly associated with the vibrational bands of the molecules containing Si−O bonds (Draine 2003). Another feature located at 18 μm has been attributed to the O−Si−O bending mode of silicates. The profile of the 9.7 μm band does not contain additional obvious features. However, laboratory observations of the spectra of crystalline silicates have shown complicated structures – this tells us that interstellar silicates are mostly amorphous, although the crystalline fraction may be non-negligible (Draine 2011). Unlike in the case of the carbon material, the most likely carrier of these bands has already been determined .- the profiles are very similar to those of olivines $Mg_{2x}Fe_{2-2x}SiO_4$ (Draine 2011). Such material would contain, besides Si, a large fraction of the interstellar Fe and Mg, which means that the models working which such grains can be immediately tested with abundance constraints.

The possibility of the crystalline silicate $CaMgSi_2O_6$ (diopside) being the carrier of a feature at 65 μm was proposed by Onaka & Okada (2003) who detected this feature in emission from the Carina Nebula and Sharpless 171 (star-forming regions). Verification of their results is still needed but it should be pointed out that this silicate would contain a significant number of interstellar calcium.

Size Distribution

The interstellar dust grains are small solid particles with sizes between $\sim 0.001 - 10$ μm. Assuming a fixed shape (and orientation) of a grain particle, the interaction with light will depend on its size. The observed extinction curves are, therefore, the result of a summed-up effect of all grains of various sizes. To be able to determine the extinction curve from theoretical models, one has to also model the size distribution function (number of grains as a function of the grain size).

Mathis et al. (1977) published one of the first empirical models of the dust grain size distribution. This so-called MRN model describes the size distribution by a function

$$\frac{dn(a)}{da} = Na^{-\gamma}\,, \tag{4.3}$$

where N is a scaling factor and $\gamma \approx 3.5$. Such function is capable of producing a very precise shape of the extinction curve in the optical part of the spectrum. The model was found to work also in the IR (Draine & Lee 1984). However, it is incapable of describing the UV region properly.

A more sophisticated distribution model was submitted by Weingartner & Draine (2001). The population of silicate and graphitic particles is described by somewhat different functions. Moreover, a population of PAHs is also included, which is needed for explaining the extinction at shorter wavelengths and to account for the emission bands in the IR (see Leger et al. 1989).

Grain Shapes

The final parameter which can affect light extinction is the shape of the dust grains. First models (such as MRN) have used the spherical approximation of the grain shapes. While calculations based on this approximation can give informative theoretical and observational results, they cannot account for all of the observed properties of the dust. For example, elongated particles are required in order to reproduce polarisation.

The next approach is to take an elongated shape of the grains. While elliptical grains are a possibility, it was found that spheroidal shapes are sufficient to reproduce most of the observations (Purcell 1969). But what is the actual shape? Are these shapes closer to flat disks or to long needles? It is generally agreed that grains have an oblate structure rather than prolate, with a typical ratio of axes $\sim 2/3$ (Hildebrand & Dragovan 1995). Of course, it is reasonable to assume that not all grains have the same shape. Instead, it is advised to take into account various elongations in the theoretical models. Draine & Hensley (2021) analysed the most likely "shape-distribution" functions for ellipsoids.

To complete this short review of the grain shapes, an emphasis must be put on the fact that dust grains are not necessarily solid particles. Since a dust grain grows by coagulation (Dullemond & Dominik 2005; Ormel et al. 2007), it is easy to imagine that its surface will be far from smooth. Instead, it was proposed that the actual structure of dust grains is best modelled by porous solid particles (e.g. Shen et al. 2008). Perets & Biham (2006) showed that porous grains can explain why molecular hydrogen can effectively form also in the photon-dominated regions, which was not possible to explain with "smooth" grains.

4.2.3 Modelling the Properties of Interstellar Dust

The final output of our theoretical knowledge of interstellar dust requires us to create models that can be compared with observations (extinction curves, polarisation curves, IR emission, and so on). In order to do that, very complicated theoretical models need to be constructed. The core physical problem requires the solution to the Maxwell equations for a chosen shape, size and chemical composition of the particle. Even for spheres, this is not a trivial task.

One possible approach is to make use of the Mie solution to the problem (Mie 1908). It consists of several steps:

1. Pick the shape and the composition of the body. The body has to be chemically homogeneous, although multilayered shapes can also be studied. Quasi-analytical

solutions can be found if the shape of the body is a sphere, spheroid, ellipsoid, toroid, cylinder, or another similar shape. The simplest task is to solve the problem for a sphere – this is what Mie did.

2. Construct the boundary conditions. Four conditions must be set for each boundary layer (only one layer in the case of a simple sphere).

3. The problem is simplified to two Helmholtz equations. With the use of the separation of variables technique, a solution to the problem is found in the form of vector harmonics (spherical harmonics in the case of a sphere).

4. To reconstruct the incident radiation (plane wave), one has to calculate the solution as the sum of a combination of the spherical wave functions. Each term of this sum is multiplied by an expansion coefficient.

5. To calculate the scattered and internal fields, the same approach is required, although the expansion coefficients have a much more complicated form than in the case of the incident field.

6. Since the electric and the magnetic vectors are used to calculate all necessary quantities (Poynting vector, scattering cross-section, polarisation), the main task is to determine the expansion coefficients. These are calculated by applying the boundary conditions.

The detailed solution to the problem can be found for spheres (see Bohren & Huffman 1983) or even spheroids (Voshchinnikov & Farafonov 1993).

Another way of solving the problem is to apply the discrete dipole approximation (or DDA, see Purcell & Pennypacker 1973; Draine 1988; Draine & Flatau 1994). Instead of solving the full problem, the whole grain is separated into a certain number N of oscillating dipoles. The electric field vector is found by summing the contribution from the different dipoles. We end up with a set of $6N$ equations that need to be solved. This, of course, requires us to find the dipole moments for each of the components. The strength of this approach is apparent – we can calculate the scattering, extinction and absorption by a grain of any shape. On the other hand, the problem scales with the number of included dipoles. Furthermore, DDA is useful only in the cases when the grain is smaller or similarly large when compared with the wavelength of the incident light. As a consequence, this approach has been often limited to studies of smaller grain (although Draine & Flatau have been working on improving the freely available code DDSCAT, `http://ddscat.wikidot.com/`, for details see Draine & Flatau 2013).

What has been neglected so far is the fact that at some point the dielectric function of the medium enters the calculations. This is not too much of a problem for silicates. One can either work with various types of silicates for which the dielectric function is known, or it should be possible to use the function for "astro-silicates" derived by Draine & Lee (1984).

For graphite, the main complication lies in the fact that it is a conductor. As such, the dielectric function must be replaced by a tensor. To get any results, the problem can be treated by assuming two situations – when the incident wave is parallel or perpendicular

to the vector normal to the "basal plane" of the grain (see Draine & Lee 1984). Another approach is to use an "effective" dielectric function which is a combination of the two functions corresponding to the two described situations. The most precise results are found when using the combination based on the effective medium theory mentioned in Draine (2016, namely MG_1).

Finally, the most complete models of the interstellar dust must also include PAHs. To include them, one of the best options is to include them using the tabulated cross-sections (`https://www.astro.princeton.edu/~draine/dust/dust.diel.html`).

4.3 Mysterious DIBs

DIBs were discovered around 100 years ago by Heger (1922). They are easy to observe in the lines of sight toward O- and B-type stars due to the absence of many stellar lines in their atmospheres. Since the 1920s, the unidentified features were analysed multiple times. Merrill & Wilson (1938) studied their wavelengths in the spectra of spectroscopic binaries and found that DIBs are staying stationary, proving their interstellar origin. Moreover, they found a correlation between the interstellar reddening and the strength of the DIBs (usually measured by the equivalent width). Based on this, they proposed the idea that DIBs may originate from solid dust particles.

4.3.1 Is Dust a Carrier?

This possibility was investigated further. Duley (1968) suggested that DIBs represent spectral fingerprints of atomic species trapped on the surfaces of dust grains. Nandy & Seddon (1970) reported detection of polarisation in the profile of the broadest DIB located near 4430 Å. However, these findings were not confirmed by later observations (Martin & Angel 1975; Adamson & Whittet 1992, 1995). This would suggest that either DIBs originate from grains that cannot be aligned with the magnetic field or they must be of gaseous (molecular) origin.

As was pointed out in the discussion by Herbig (1995), the correlation between the reddening and the strength of the DIBs is not very strong. Scatter around the predicted linear relation is quite appreciable. The main source of this scatter is most likely the variation in the environment along the line of sight. For example, Snow & Cohen (1974) showed that if dense clouds are introduced into the line of sight, the strength of the DIBs does not increase by much (unlike the value of the reddening). This "weakening" of the bands in dense clouds was also supported by Adamson et al. (1991) and Cami et al. (1997). This would suggest that the formation of DIBs depends on the presence of a UV radiation field. Hence, it is expected that the carriers will be more likely located in the outer layers of the dense clouds. This further proves the molecular origin of the DIBs.

4.3.2 Emission Features

Another interesting finding was the observation of emission features in the Red Rectangle which may be connected to the DIBs (Scarrott et al. 1992). However, at least some of the observed features are likely shifted significantly away from the positions of DIBs (Van

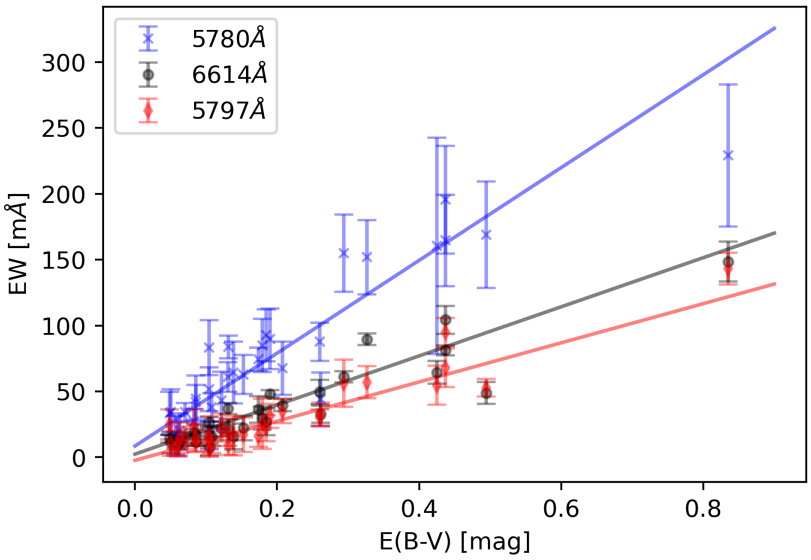


Figure 4.7: Correlations between the EW of three different DIBs and the interstellar reddening $E(B-V)$. Based on the data from Puspitarini et al. (2013).

Winckel et al. 2002). Lai et al. (2020) recently revisited this issue but found no evidence that would support the idea that the emission features are connected to the DIBs. This topic requires further investigation.

4.3.3 Specific DIBs and Their Profiles

Some of the strongest and most studied DIBs are those located near: 4430 Å, 5780 Å, 5797 Å, 6196 Å, 6284 Å, 6379 Å and 6614 Å. The band located near 4430 Å is quite different from the rest – it is very broad ($>$ 10 Å, but other broad DIBs are also detected, see Galazutdinov et al. 2020). It should be noted that the DIBs may form families of features, as suggested by Krelowski & Walker (1987). It was suggested (e.g. Galazutdinov et al. 2002) that the bands at 6196 Å and 6614 Å are very well correlated. Later study by Krelowski et al. (2016) confirmed this strong correlation but also suggested that the two react differently to the changes in the conditions in the ISM.

A very important aspect of the DIBs is their profiles. While some of the features have very symmetric profiles, most of them are quite asymmetric (e.g. Jenniskens & Desert 1993). At lower spectral resolution, the profile shapes do not display any significant changes (Porceddu et al. 1991). However, small variations in the positions of the sub-profile structures were hinted at by observations at very high resolutions (Cami et al. 2004). While it is generally accepted that the structure of the DIB-profiles is dominated by the temperature, suggestions were made that also the isotopic effect (for detection of the effect in the CH^+ molecule, see Hawkins & Jura 1987) may play a role in the profiles variations (Webster 1996; Galazutdinov et al. 2008).

When investigating profiles of DIBs, one must take into account several effects (e.g. Westerlund & Krelowski 1988). First of all, most of the spectra are produced in ground-based observations. In such a case, telluric lines produced in the Earth's atmosphere must

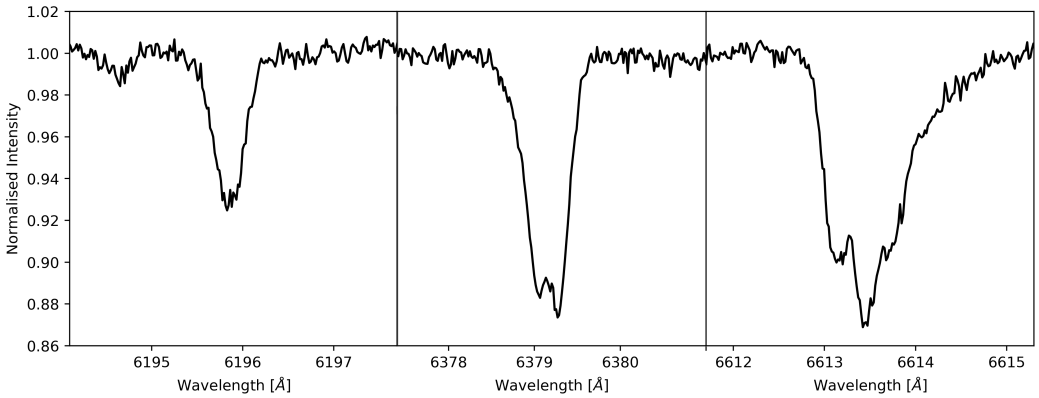


Figure 4.8: Some of the DIBs seen in the line of sight towards HD 163800. The spectrum was observed by ESO VLT with the UVES instrument. Data downloaded from ESO archives (archive ID: `ADP.2020-06-28T20_02_50.053`, program ID: `081.C-0475`).

be treated correctly. Furthermore, DIBs have also been studied in the spectra of stars of spectral type cooler than B, where stellar lines may overlap with the interstellar features (for possible treatments of the problem, see Destree et al. 2007; Puspitarini et al. 2015). Some problems may remain even in the case of hot stars. For example, the 6614 Å band is located in the red wing of the prominent H_α line. Stellar and telluric features dominate the region between 9000 Å and 10 000 Å, where several important lines are present (see below). Atmospheres of cool stars may heavily influence the detection of the DIBs at 5780 Å and 5797 Å (together with the telluric lines). However, the most clearly affected band is the one located at 6284 Å. If observed from the surface of the Earth, this region is tainted by the strong signature of the O_2 molecule. To deal with this problem, a comparison star is also observed and its (scaled) spectrum can be subtracted from the spectrum of the studied star. This should almost completely remove the influence of the telluric lines on the measurements and also possibly reduce the strength of the stellar features, depending on the choice of the comparison star.

Of interest are also the recent studies concerning the DIBs in the IR. Kos et al. (2014) prepared maps of the ISM based on the strength of the band at 8620 Å. Such maps may prove to be invaluable tools for investigating the spiral structure, although the connection between the spiral arms and the DIBs has not been established yet. The DIB at 8620 Å has been of a great interest in the last years because of its location within the spectral coverage of Gaia's Radial Velocity Spectrometer instrument. Zhao et al. (2021) presented their study of the kinematics of the carrier and they argued that the carrier may be distributed in the regions significantly closer to the Sun than the observed targets. Another IR band, located at 15 273 Å, was studied by Elyajouri & Lallement (2019) with the use of the APOGEE spectra. Their efforts resulted in an unprecedented number of DIB measurements, providing the strength of the DIB in over 100 000 lines of sight. The focus was put on the study of the dense regions in the sky (such as in Orion and Taurus), where the DIB measurements were compared with the Planck's dust map. The authors confirmed the relative weakness of the DIB within/near the clouds.

4.3.4 Other Galaxies

Is it possible to observe DIBs outside of our Galaxy? To answer this question, we have to realise that other galaxies have (typically) a non-zero radial component of the velocity vector. When we observe such a galaxy, we should notice absorption features originating from two components – the ISM of our Galaxy and of the studied galaxy. Due to the radial velocity difference, the interstellar features of the components will be shifted from each other. For this reason, the detection of the DIBs is feasible, although an overlap of an extra-Galactic DIB with a strong feature originating from the Galaxy may still cause some problems.

Vladilo et al. (1987) presented one of the first reports of extra-Galactic DIBs, namely in the Large Magellanic Cloud. They found that the DIBs in this galaxy appear to be weaker when compared with our Galaxy. Another detection of DIBs in Magellanic Clouds was mentioned in Ehrenfreund et al. (2002). Results were also obtained for NGC 1448 (Sollerman et al. 2005), M 31 (Cordiner et al. 2008a) or M 33 (Cordiner et al. 2008b). The only differences between these works are found in the different values of the DIB-strength per unit reddening.

4.3.5 Origin

The ultimate question related to the DIBs is concerning their origin. This means that we not only need to discover which specific molecules are causing the specific observed DIBs, we also have to figure out how these molecules can exist in the ISM. It must be emphasised that a typical ISM is affected by the interstellar radiation field which can destroy all of the known interstellar molecules.

For a molecule to survive, it must be able to withstand high-energy radiation for a significant amount of time. This would require complex molecules that could redistribute the absorbed energy into multiple vibrational and rotational modes. One example of a group of such molecules is represented by PAHs (Leger & D'Hendecourt 1985; Salama et al. 2011). However, a specific identification based on laboratory studies is still required (Maier et al. 2004; Huisken et al. 2014).

Another group of molecules often discussed together with PAHs is represented by fullerenes. These are large molecules, where atoms form rings and the rings form a closed cage-like structure. One of the fullerenes, C_{60}, has been suggested as a carrier of the DIBs already during its discovery by Kroto et al. (1985). It was theorised that the molecule can form in the circumstellar material (or perhaps in the winds) of the carbon-rich stars (e.g. Cherchneff et al. 2000) in the ejecta of supernovae (Clayton et al. 2001). A huge step forward was made when Sellgren et al. (2010) confirmed the presence of the C_{60} in the ISM.

Foing & Ehrenfreund (1994) were among the first to show that two bands in the 9000Å region may originate from the C_{60} molecule. It took around 20 more years until laboratory experiments showed that the two (and possibly more) DIBs at 9577 Å and 9632 Å are located at wavelengths corresponding to laboratory observed features of C_{60}^{+} cation (Campbell et al. 2015). These findings were later confirmed by follow-up investigations (Cordiner et al. 2019; Nie et al. 2022).

Despite this incredible discovery, over 99 % of the DIBs remain unidentified. Clearly, a lot of work remains for observational and laboratory groups. Moreover, we are still nowhere close to understanding how such molecules are formed and find their way to the ISM. Top-down formation model seems to be one of the possibilities (Berné et al. 2015). However, this requires the existence of sizeable molecular species – one feasible pathway is the formation of such molecules on the surfaces of dust grains (Cami et al. 2018; Bernal et al. 2019).

4.4 Concluding Remarks

DIBs represent interesting features that have a lot of potential in future astrophysical studies. However, there is only little that we know about these absorption bands. Leaving aside the identification of the carriers, the maps of our Galaxy (or of other galaxies) based on the DIBs remain quite poor-quality. Even the detailed study by Elyajouri & Lallement (2019) leaves many of the Galactic regions uncovered. Could global maps DIBs in our Galaxy provide new insights about the ISM?

Moreover, the connection between the dust and the carriers of DIBs is still not fully explored. For example, we know that the positions of the most reddened regions correlate with the positions of the spiral arms. This is observed in our Galaxy, but becomes even more apparent when looking at other galaxies. Can a similar study be done with the use of the DIBs? Would the results be consistent with other studies (e.g. based on open clusters)?

Chapter 5

Structure Of Open Clusters - Gaia DR2 And Its Limitations

***Piecka, M.**, Paunzen, E., 2022, Bulgarian Astronomical Journal, Volume 36, 27 pp.*

During an investigation of tidal tails around open clusters, one will notice in the 3D space that almost all of the observed clusters are greatly elongated along the line of sight. Moreover, the elongation will be smaller towards the Sun than in the direction outwards. This is the well-known problem with the transformation of parallaxes to distances.

It would be interesting to explore how the situation changes if a Bayesian approach is used for the calculations. From the simple fact that we have some observational uncertainties, we should not expect that the elongation will disappear, as the measurement uncertainties are the prime candidate for this effect. Such an effect has a direct impact on our studies of open clusters. Not only will it create a discrepancy between the photometric (CMDs) and the astrometric (parallaxes) distances, it will also affect our ability to study clusters in 3D and 6D.

Since the Bayesian approach is often non-trivial (often, multiple free parameters have to be included in the prior estimate), it would be interesting to see whether there is a simple measure of the distance that would provide reasonable results (at least to such degree that it would agree with a Bayesian estimate within the error range). Open clusters represent interesting objects where we can study hundreds or even thousands of stars. Since we expect that a cluster's dimension is significantly smaller than its distance, we can view each star as an independent measurement of the parallax. The main problem with estimating distances using $d = 1/\varpi$, ignoring the effect on the estimated distance uncertainty, is the fact that this relation is correct only if the "true" parallax is known. Of course, this is not possible in reality. However, an analysis of the distribution of the parallax measurements should yield a good estimate of the "true" value. For this reason, we should expect that the inverse of the mean (or median) parallax of cluster members may yield a fairly precise value of the cluster's distance.

Another important aspect of studying clusters in 3D space is their projected size. While the radii of clusters can be studied in multiple ways, we should realise that two major factors have an impact on the calculations. First, the radius is linearly proportional to the

distance of the cluster. Any issue with the latter will directly affect the former. Secondly, the number of derived cluster members has a huge effect on the radius. Hence, the value of the size will somewhat vary, depending on the choice of the procedure used during the membership classification process.

Finally, a display of the elongation of clusters due to the parallax (or distance) uncertainties is missing in the literature. While the effect has been known for a long time, a good number of modern works still ignores its existence. Providing a diagram of what the effect actually looks like (in 3D) may give the needed push to carefully analyse cluster distances. This is a prerequisite for studying the kinematics and dynamics within the observed clusters – such studies are very important for the comparison of the models with observations.

MP analysed the parallax distributions from the cited catalogue of open clusters, prepared the mathematical analysis of the parallax problem, calculated the projected widths of open clusters, and was responsible for writing the larger part of the paper.

Structure of Open Clusters - *Gaia* DR2 and its limitations

Martin Piecka[1], Ernst Paunzen[1]
Department of Theoretical Physics and Astrophysics, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic
408988@mail.muni.cz

(Submitted on 29.04.2021; Accepted on 01.07.2021)

Abstract. Very precise observational data are needed for studying the stellar cluster parameters (distance, reddening, age, metallicity) and cluster internal kinematics. In turn, these give us an insight into the properties of our Galaxy, for example, by giving us the ability to trace Galactic spiral structure, star formation rates and metallicity gradients. We investigated the available *Gaia* DR2 catalogue of 1229 open clusters and studied cluster distances, sizes and membership distributions in the 3D space. An appropriate analysis of the parallax-to-distance transformation problem is presented in the context of getting distances toward open clusters and estimating their sizes. Based on our investigation of the *Gaia* DR2 data we argue that, within 2 kpc, the inverse-parallax method gives comparable results (distances and sizes) as the Bayesian approach based on the exponentially decreasing volume density prior. Both of these methods show very similar dependence of the line-of-sight elongation of clusters (needle-like shapes resulting from the parallax uncertainties) on the distance. We also looked at a measure of elongations of the studied clusters and find the maximum distance of 665 pc at which a spherical fit still contains about half of the stellar population of a cluster. It follows from these results that the 3D structure of an open cluster cannot be properly studied beyond $\sim$ 500 pc when using any of the mentioned standard transformations of parallaxes to distances.

Key words: open clusters astrometry galactic structure gaia satellite mission

Introduction

Galactic star clusters are most important objects not only when it comes to describing the Milky Way and its structure, but also for studying the individual stellar members. Various star groups, such as variables and binaries, can be studied in star clusters in a statistical way. This is based on the idea that the cluster reddening, age, distance and metallicity can be assumed to be the same for each of the cluster members. These cluster parameters can be deduced by fitting proper isochrones, for example.

In the recent years, most open clusters were photometrically studied in a (semi-)automatic way using 2MASS $JHK_{\rm S}$ and *Gaia* $G_{\rm BP}$, $G_{\rm RP}$, and G data. The traditionally photometric systems like the Johnson-Cousins $UBVR_{\rm C}I_{\rm C}$ and Strömgren $uvby\beta$ ones are hardly used any more. Especially critical is the lack of observations in the ultraviolet region – this makes it difficult to deredden individual stars or to get membership probabilities by using a classical $(U-B)$ versus $(B-V)$ diagram (Yontan et al. 2019), f08 for example.

With the launch of the *Gaia* satellite the hopes were high to get precise membership probabilities using parallaxes, proper motions, and radial velocities for a statistically sound sample of star clusters. For the first time, even the intrinsic kinematics was hoped to be investigated. The latter is important for the understanding how the angular momentum of the initial molecular cloud is conserved during the formation and evolution of star clusters. We basically have no knowledge about the initial conditions when it comes to the rotational characteristics. But for our understanding of kinematics and dynamics

of a cluster (for example, modelling the first stages of cluster evolution, studying the kinematic evolution of clusters and dynamical effects), this information is vital – see, for example, Küpper et al. (2010) and Parker & Wright (2016).

Cantat-Gaudin et al. (2018) presented a status report for 1229 open clusters on the basis of the *Gaia* DR2 release. They established a list of members and derived cluster parallaxes and distances within a given error range. The other three cluster parameters (age, reddening, and metallicity) were neither derived nor taken into account (especially the reddening). They also reported the discovery of 60 new open clusters. These were identified on the basis of consistent proper motions, parallaxes, and concentrations on the sky. Using this method, Cantat-Gaudin et al. (2019) detected 41 additional new star clusters. Later on, Monteiro et al. (2019) also used *Gaia* photometry to get all four cluster parameters for the above mentioned discovered aggregates using a cross-entropy global optimization algorithm to fit theoretical isochrones. However, their analysis showed that 80 candidates are likely not real open clusters. This already shows that kinematical data alone are not sufficient and photometric data have to be taken into account when analysing star clusters. Bossini et al. (2019) derived ages, reddening, and distances (for a fixed metallicity) for 269 open cluster from the sample by Cantat-Gaudin et al. (2018). They have used an automated Bayesian tool for fitting stellar isochrones to *Gaia* photometry using the membership probabilities from Cantat-Gaudin et al. (2018) for selecting the cluster sequences. Their sample is biased because they selected only low reddening objects and discarded very young clusters. One of their main results is that 90% of the clusters have a sigma of the absolute distance modulus smaller than 0.037 mag (median is 0.025 mag). However, as they have shown, the errors increase by about one order of magnitude when metallicities are taken into account. If all four cluster parameters (distance, extinction, age, and metallicity) are considered, the differences of the derived values from different independent sources and data sets are quite large. Fitting isochrones to an open cluster population is a complex procedure and depends, for example, on the turn-off point and the location of the red giant population. In the literature compilation by Netopil et al. (2015), the dispersion between different data sets amount to about 0.2 dex for the age, 0.08 mag for the reddening, and 0.35 mag for the distance modulus. Similar or even larger discrepancies can be seen in Figs. 9 and 10 of Bossini et al. (2019).

In this paper, we investigate limitations of the *Gaia* DR2 data when it comes to the investigation of Galactic open clusters and their parameters. The paper is organized as follows: in Sect. 1 we present the basic characteristics of the data set and the target cluster selection; in Sect. 2a summary of the problem of transforming parallaxes to distances is given; in Sect. 3 we analyse in detail cluster distances and width.

1. Target cluster selection

For our work, we used the data set based on the analysis by Cantat-Gaudin et al. (2018) who presented an unsupervised membership assignment procedure to determine lists of cluster members based on the *Gaia* DR2 catalogue. They provided the membership and mean parameters for a set of 1229 clusters and 401 448 individual stars. Their analysis is based on the membership assignment

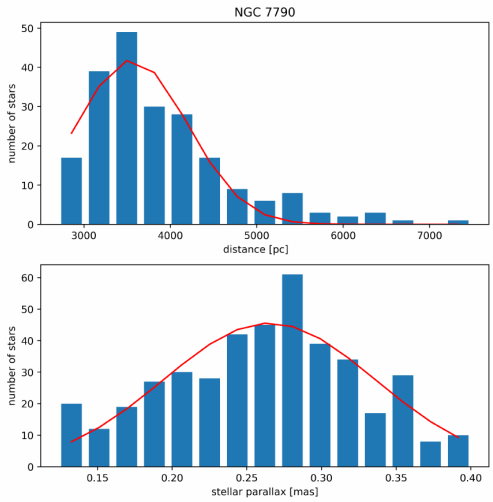


Fig. 1: The distribution of the distances (upper panel) and parallaxes (lower panel) for all members with probabilities higher than 50% for NGC 7790 taken from Cantat-Gaudin et al. (2018). The distances were calculated by inverting parallaxes. This approach is the reason why assembling a symmetric parallax distribution always means to have a tail of much more distant members.

code UPMASK (Unsupervised Photometric Membership Assignment in Stellar Clusters, Krone-Martins & Moitinho 2014) which does not rely on physical assumption about stellar clusters, apart from the fact that its member stars must share common properties, and be more tightly distributed on the sky than a random distribution. The analysis by Cantat-Gaudin et al. (2018) is based only on kinematical and astrometrical data.

One has to keep in mind that working in the parallax space is different than working in the distance space (see Sect. 3.1). Basically, Cantat-Gaudin et al. (2018) used a maximum likelihood procedure, maximising the probability of measuring the parallax of each individual star and the likelihood for the cluster distances to be the product of the individual likelihoods of all its members. They neglected any correlations between parallax measurements of all stars such as suggested in Schönrich et al. (2019), for example. The resulting distribution should be symmetrical for the parallaxes, but is not for the distances. In Fig. 1, the example of this effect for NGC 7790 is shown. This is not unexpected - Luri et al. (2018) already showed this effect in their analysis of the *Gaia* DR2 data. Working with a symmetric parallax distribution always means to have a tail of much more distant members for the investigated cluster (when using inverse parallax approach). In the case of NGC 7790 (distance of about 3700 pc) this means that members are distributed from 2700 to 7500

pc with about 15% of stars more distant than 4700 pc. These distances were calculated using $r = \varpi^{-1}$ for the individual cluster members. The conversion from parallaxes to distances is clearly problematic. However, it will be shown in Sect. 2. that there is a simple statistical solution when dealing with open clusters.

In the following, we define two different samples for which we did our analysis. These samples are defined as:

- "Loose sample": number of stars in a cluster > 50, individual parallax (or distance) error $< 50\%$, and individual membership probability $> 50\%$; 938 aggregates
- "Strict sample": number of stars in a cluster > 300, individual parallax (or distance) error $< 5\%$, and individual membership probability $> 70\%$; 181 aggregates

The individual parallax/distance error refers to the fact that we will begin our analysis with both, starting in the parallax space and in the distance space. The mentioned errors are taken from the used data sets.

For the purpose of this work, we have chosen to work with the data from Cantat-Gaudin et al. (2018) and Bailer-Jones et al. (2018). The former presents one of the most recent compilations of parallaxes for a larger number of open clusters. The latter data set gives the largest sample ($\sim 10^9$) of distances for Galactic stars – we have used those which coincide with the catalogue from Cantat-Gaudin et al. (2018). Although these distances may not be the best for analysing open clusters, they present a good starting point and an option for a comparison with a different approach.

The discrepancy of the number of clusters for the loose sample and the total number is explained by the fact that almost 300 open clusters have less than 50 members when applying the individual distance errors and membership probabilities. We have chosen the upper limit of 50 members because otherwise the distance distributions in the histograms are mostly dominated by noise which creates problems for the fitting procedure described in the next sections.

We also have to emphasize that 1795 stars were found to be members of at least two open clusters (49 individual ones in total), 579 of them with a membership probability of higher than 50% for both clusters. Although this number is insignificant compared to the overall number of investigated stars, it still shows that there are shortcomings in the numerical procedure for deriving the cluster memberships.

2. Calculation of distances from parallaxes

As was mentioned, calculating the distance r by inverting the parallax ϖ is a problematic approach (for more details, see Luri et al. 2018). This is due to the fact that the measurement is accompanied by an uncertainty. If we assume that the probability density function (PDF) for a parallax measurement is a normal distribution it will not transform to a normal distribution by assuming ϖ^{-1}. Instead, it will produce a longer tail towards the larger distances and the maximum of the distribution will be located at somewhat shorter distance when compared with the true distance. This effect will increase with the value

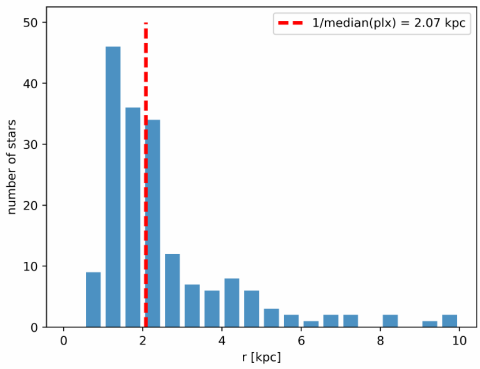


Fig. 2: Simulation of an open cluster located 2.00 kpc away. The red dashed line indicates the determined distance of a cluster calculated as the inverse of the median of all parallaxes of the cluster members.

of the relative error $f = \frac{\sigma_\varpi}{\varpi}$ of the parallax measurement. Obviously, for some (small) values of f the differences become quite negligible.

In this section, we aim to examine how this affects the determination of distances toward open star clusters. Although inferring precise values of distances from parallaxes can be quite complicated for the field stars (due to the transformation problems discussed above), we believe that the situation gets much better for open clusters (although not for the individual cluster members).

2.1. Getting distances toward open clusters

Due to the fact that the observed number of members of a cluster represents a number of distinct measurements of the centre of the cluster (in terms of parallaxes), the distribution of parallaxes should represent the PDF in parallax space. Finding the maximum of this distribution (usually close to a normal distribution) is very simple and can be achieved by calculating the median or fitting a Gaussian function. Since the centre of the distribution is assumed to be the true parallax $\varpi_{\rm true}$, the inverse of this value should yield the true distance $r_{\rm true}$ towards the cluster.

Similar to Luri et al. (2018), we have simulated a cluster of 200 stars at the distance $r_{\rm true} = 2.0$ kpc using a normal distribution with scale $s = 5$ pc (representing the cluster size). Then we transformed these 200 true distances to true parallaxes using $\varpi_{\rm true} = r_{\rm true}^{-1}$. The observed parallaxes were generated using normal distribution with the centre located at the true parallaxes and assuming the width of the distribution to be $\sigma_\varpi = 0.3$ mas. From those, we could construct the distribution of distances of the individual cluster members using a Bayesian approach. However, if we are only interested in calculating the distances toward the centres of the studied clusters, this may be unnecessary.

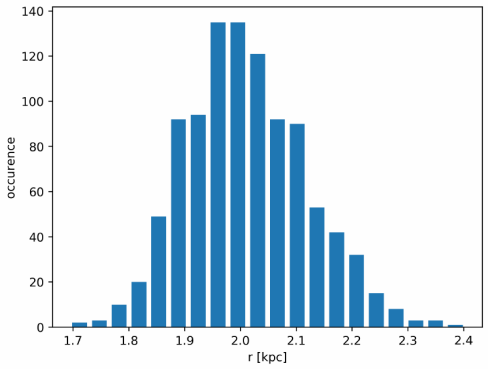


Fig. 3: Based on 1000 simulations of the same cluster, we can generate a histogram of the distances found by the inverse of the median(ϖ) (as in Fig. 2). The centre of this distribution is located at $R = 2.00$ kpc with $\sigma_R = 0.11$ kpc.

Instead, we are going to look at the distribution of parallaxes. Assuming that the distribution is normal, we can fit the data with a Gaussian which will give us the parameters of the fit, ϖ_0 and σ_ϖ, together with their uncertainties. If the number of the observed stars of a cluster is high enough and the observational uncertainty is low enough for the most of the members, we can safely assume that $\varpi_{\rm true} = \varpi_0$ and determine the distance towards the centre of the cluster (Fig. 2).

To test the results of this method, we can generate the same cluster (with the same input parameters mentioned above) a thousand times and can construct a histogram of the found distances. Analysis of this histogram (Fig. 3) yields us the variance of the distances found using this method. The variance depends on the input parameters – the number of cluster members and the observational uncertainties. For our test case, we find the values of median $R = 2.00$ kpc and the standard deviation $\sigma_R = 0.11$ kpc. It is worth pointing out that this gives about an order of magnitude smaller relative error than we have for the assumed measurement uncertainty.

2.2. Determining the distance errors

In reality, we do not have a thousand of observations of all members for the same cluster. However, we can still learn much from the Gaussian we used to fit the observed parallaxes of a cluster. All of the information about the variance of the final distance is found in the parameter which describes the centre of the Gaussian. It is therefore a good first approximation to use the uncertainty of this parameter to determine the uncertainty of the final value of the distance. This can be calculated from the simple approach as $e_r = \frac{e_{\varpi_0}}{\varpi_0^2}$, assuming that $\frac{e_r}{r} \ll f$. If we assume that $r \approx R$ and $e_r \approx \sigma_R$ (r represents

Table 1: Comparison of the true distances with the distances r and R found by inverting ϖ_0 for six different simulated clusters of size $s = 5$ pc. As expected, R very closely matches the true distances of a given cluster. Although the values of r typically differ from R, the difference $|r - R|$ is usually lower than e_r.

Cluster	N	$r_{\rm true}$ [kpc]	f	r [kpc]	e_r [kpc]	R [kpc]	σ_R [kpc]
1	200	2.0	0.60	2.12	0.13	2.00	0.11
2	50	2.0	0.60	2.22	0.41	2.00	0.22
3	50	2.0	0.20	2.05	0.05	2.00	0.07
4	200	4.0	0.60	3.82	0.12	3.99	0.21
5	50	4.0	0.40	4.26	0.46	3.99	0.27
6	100	5.0	0.25	5.21	0.15	5.01	0.16

the distance we get from one simulation of a cluster, R is the distance we get from one thousand simulations of the same cluster), then the term on the left hand side tends to be about an order of magnitude smaller than f, so this should be a good assumption. We present the values of distances r and R for 6 simulated clusters in Table 1. We have used different numbers of members and different observational uncertainties.

As we can see, the results are, for the most part, quite similar. The exceptions are such clusters where the number of the observed cluster members is lower than $N \sim 100$ and the observational relative error is higher than $f \sim 0.50$. Moreover, the distance towards the cluster plays a crucial role. It is also worth mentioning that the calculated uncertainties of the fit parameters of a single cluster will slightly vary due to the randomness included in the cluster generation procedure.

Generally, if the number of cluster members is $N > 50$ and the observational relative error is $f < 0.50$ then we can use the described procedure to determine distances (and their uncertainties) toward open clusters quite precisely up to $r_{\rm true} \sim 4$ kpc. However, it should be possible to use this approach also for the more distant clusters with $N > 100$ if the observational error is lower than $f \sim 0.25$.

2.3. Variations in the observational uncertainties

Unfortunately, the parallax measurement uncertainties are not the same for all members of a cluster. Let us take a look at what happens when we assume a distribution of uncertainty values.

We can find in the data from Cantat-Gaudin et al. (2018) that for distances below 3 kpc the distribution of uncertainties can be well described by a combination of a Gaussian distribution together with an exponential distribution. The probability distribution we used can be seen in Fig. 4. The Gaussian in the distribution is required to produce the short tail towards the smaller values. Finally, it must be mentioned that for clusters beyond 3 kpc the position and width of the Gaussian term increase with the distance. We have decided to ignore this small discrepancy at larger distances since we are only interested in the effect that such a distribution has on the determined distance errors.

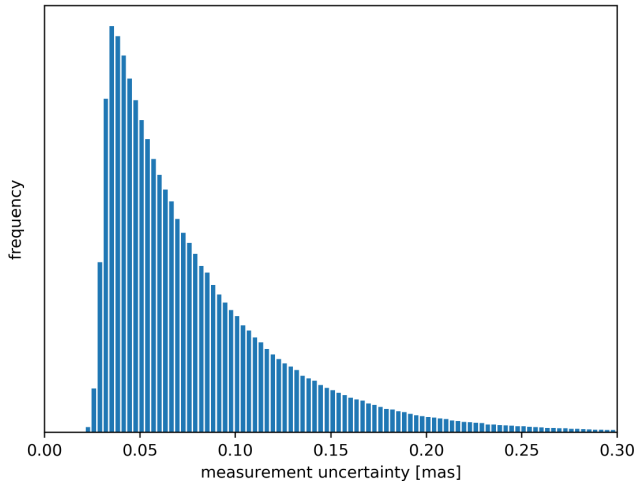


Fig. 4: Estimated probability density function for the uncertainty of the parallax measurement. The function is composed of two parts – the normal distribution creates a short tail toward the lower values and defines the position of the peak of the PDF (does not coincide with the peak of the normal distribution), the exponential distribution creates the long tail toward the higher values of uncertainty. The probability that the uncertainty will be lower than 0.32 mas is about 99.7 %.

Let us simulate two very different situations ($N = 50$ and $N = 200$) to study the effect of the distribution of uncertainties. We have generated 80 clusters between 0.05 kpc and 4.00 kpc (with equidistant steps) and for each cluster we calculate $\frac{\sigma_R}{R}$. We want to know how this ratio varies with the number of cluster members and with distances. If the distribution of σ_ϖ plays any important role, it should present itself as a spread in the curve of the plot of the relative error of R against the true distances.

As we see in Fig. 5, the ratio $\frac{\sigma_R}{R}$ gets (statistically) smaller for larger values of N. Also, it increases with the true distance of the cluster, as was expected. However, it is apparent that the plot is dominated by the variation of the observational uncertainty which produces the scatter around the dashed lines (the same as solid lines, but with fixed uncertainty 0.07 mas). It should be noted that only rarely $\frac{\sigma_R}{R} > 0.10$ with the assumed distribution of σ_ϖ.

We can also display the deviations of the values of R from the true distances $r_{\rm true}$ using the same approach but putting $R - r_{\rm true}$ on the y-axis. This is shown in Fig. 6. Statistically speaking, these differences are very small, therefore we assume that the designed procedure should work very well assuming that the

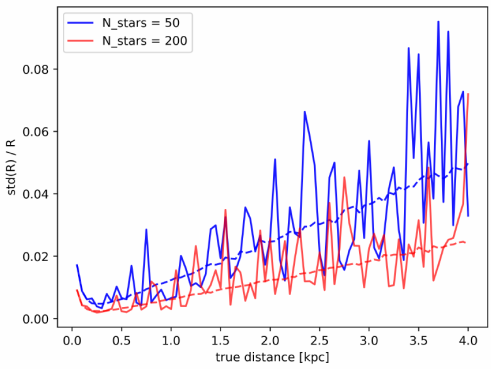


Fig. 5: The influence of the distribution of measurement uncertainties on σ_R/R. Solid lines represent the values calculated with the assumed distribution of uncertainties, dashed lines represent the same cluster but assuming constant $\sigma_\varpi = 0.07$ mas.

typical observational uncertainty is low enough and that the number of cluster members is sufficient (this is a very good assumption for our "strict sample").

2.4. Individual cluster members

The last question that remains to be answered – can we calculate the distances toward the individual stars of a cluster with the procedure described above? Unfortunately, the answer is no. To calculate the individual distances one must rely on the Bayesian approach assuming a reasonable prior probability distribution and then determine the posterior distribution. Possible approaches to this problem were described by Bailer-Jones (2015). Another solution would be to observe the parallax of the same star multiple times but this is not practical.

Since the goal of this work is not the calculation of the distances but rather the statistical analysis of the effects of observational uncertainties on the results derived from *Gaia* DR2, we will look at this problem from a different perspective. As we already have an acceptable estimate of the distance and of its error (r, e_r), we only need to get the variances of distances of the individual members of clusters. To find these, we will use the distances $r_{\rm bay}$ resulting from a Bayesian approach (taken from Bailer-Jones et al. 2018) and the inverse-parallax distances $r_{\rm inv}$.

Before we start our analysis, let us estimate what we should expect from $r_{\rm bay}$ and $r_{\rm inv}$. We will simulate a cluster of 1000 stars at the distance $r_{\rm true} = 2.0$ kpc with cluster size $s = 5$ pc. Inverse-parallax distances toward the individual stars are calculated simply as $r_{\rm inv} = \varpi^{-1}$. The Bayesian distances can be calculated as described in Bailer-Jones (2015), with the help of the exponen-

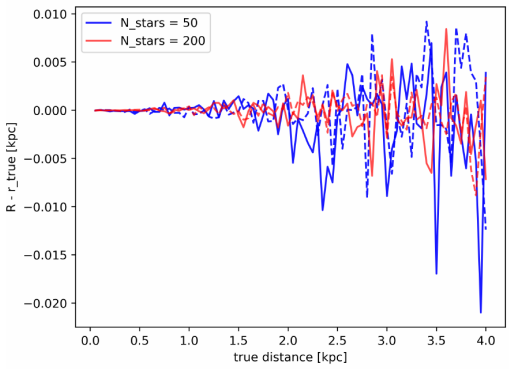


Fig. 6: The influence of the distribution of measurement uncertainties on $R - r_{\rm true}$. Full lines represent the values calculated with the assumed distribution of uncertainties, dashed lines represent the same cluster but assuming constant $\sigma_\varpi = 0.07$ mas.

tially decreasing volume density prior – the distance of a star is found as the maximum in the PDF assuming the characteristic length scale $L = 1.2$ kpc.

To characterize the cluster, we would like to fit the distributions of both distance measures with Gaussians (for simplicity) and use their parameters for describing the results. However, in the case of $r_{\rm inv}$ the distributions may be quite asymmetrical at large cluster distances (or with large relative observational uncertainties). Therefore, we have decided to fit $r_{\rm bay}$ with a Gaussian directly, while in the case of $r_{\rm inv}$ we first stay in the parallax space. Parallax distribution is fitted by a Gaussian (as was discussed in Sects.2.1 and 2.2), and we use the width of this distribution as the measure of the width of the distribution in distances as $\sigma_r = \frac{\sigma_\varpi}{\varpi_0^2}$. This measure should not be confused with the mentioned calculation of the distance errors in Sect. 2.2 – in this case, the relative errors are clearly very similar ($\frac{\sigma_\varpi}{\varpi_0} = \frac{\sigma_r}{r}$), therefore we need to be very careful when assigning a meaning to the quantity σ_r.

For $\sigma_\varpi = 0.10$ mas, the widths of these distributions are almost identical ($\frac{\sigma_{\rm inv}}{\sigma_{\rm bay}} \sim 1.10$, see Fig. 7). On the other hand, when we take $\sigma_\varpi = 0.30$ mas, the Bayesian approach starts giving significantly smaller Gaussian widths ($\frac{\sigma_{\rm inv}}{\sigma_{\rm bay}} \sim 1.45$, see Fig. 8). However, it should be noted that such high observational uncertainties are rare in our two data samples and the typical relative uncertainty never gets close to 50 %.

Clearly, Bayesian approach is superior to the naive inverse-parallax approach. However, both methods give very similar values of distribution widths (in terms of the widths of fitted Gaussians), significantly deviating only in the most extreme cases. Based on this, we expect the two approaches to give approximately the same values of cluster widths.

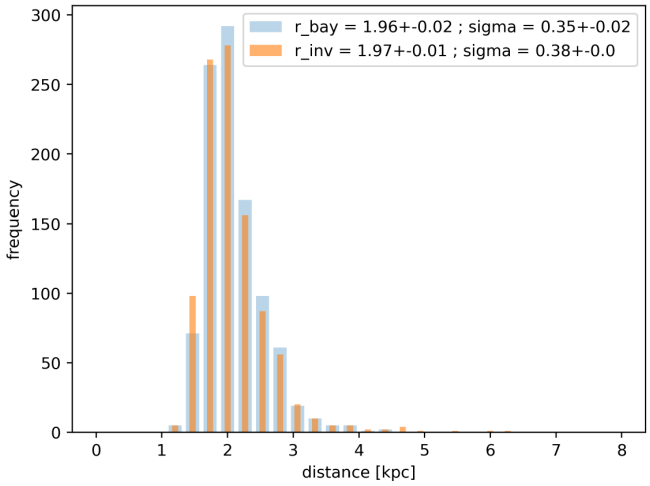


Fig. 7: Comparison of the Gaussian parameters derived from Bayesian distances $r_{\rm bay}$ and inverse-parallax distances $r_{\rm inv}$, assuming $\sigma_\varpi = 0.10$ mas. In the approach of inverse-parallaxes, the distribution of parallaxes was used instead of the distribution of $r_{\rm inv}$. The used method is described in Sects. 2.1 and 2.2.

3. Analysis

In the following subsections we present a detailed astrometrical and kinematical analysis of the targets. Notice that throughout the paper the errors in the final digits of the corresponding quantity are given in parentheses.

3.1. Distances and widths

To start with, we will define the rectangular Galactic coordinates $[X\,Y\,Z]$ using the spherical Galactic coordinates r, l, and b

$$X = r\,\cos l\,\cos b\,, \tag{1}$$
$$Y = r\,\sin l\,\cos b\,, \tag{2}$$
$$Z = r\,\sin b\,. \tag{3}$$

Assuming spherical symmetry on the sky, we need to pick only two coordinates to characterize a cluster in 3D space. The first of these coordinates is the projected distance d_1 which can be calculated by projecting the distance r of the star on the vector oriented in the line-of-sight towards the centre of the cluster

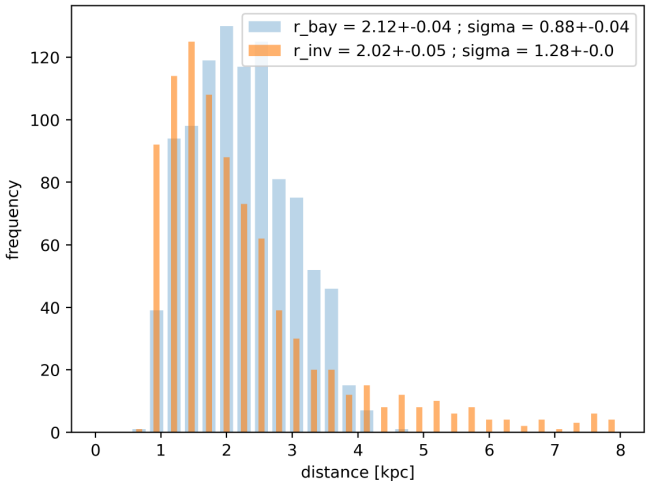


Fig. 8: Same as Fig. 7 but assuming $\sigma_{\varpi} = 0.30$ mas.

$$d_1 = r\ \cos(\phi), \tag{4}$$

where ϕ is the angle between the reference vector (centre of the cluster) and the vector oriented towards the given star. It should be noted that the value of d_1 is almost indistinguishable from the value of r, therefore we will use r as the measure of the distance throughout this work. The second coordinate is the projected width d_2, which can be calculated similarly

$$d_2 = r\ \sin(\phi). \tag{5}$$

As was mentioned in Sect. 2, we cannot easily analyse a cluster with the distances toward individual cluster members. Instead, if possible, we would like to use the whole collection of stars to determine some properties of a given cluster. To do this, first we have to calculate the distance towards the cluster r. Then we can analyse the distribution of the Galactic coordinates l and b which are determined quite precisely, and determine the angular size of a cluster ϕ_{clust} as the mean of the standard deviations of the two coordinates. The absolute cluster size can then be determined as the projected width d_2. If this is done for each of the clusters in our samples, we can find the typical spatial diameter of a cluster by analysing the distribution of the projected widths (which can be done by calculating the median and the median standard deviation, for example).

We have applied the procedure discussed in Sects. 2.1 and 2.2 to our samples and calculated their distances and projected widths, together with the corresponding errors. For the most part, we will focus our attention on the

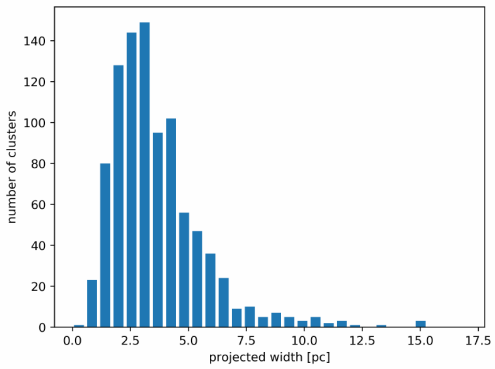


Fig. 9: The histogram of the projected widths. We can interpret these widths as a measure of the cluster size or diameter.

loose sample. The only aggregate from the complete sample that we were not able to fit in the distance space is Tombaugh 2, which is a distant (6 to 8 kpc from the Sun) old open clusters in the direction of the Galactic anti-centre (Cantat-Gaudin, 2016). The available studies agree that its metallicity is sub-solar with a widespread of given values (-0.07 to -0.44 dex). Looking into WEBDA[1] there is no doubt that this is a true open cluster, but the colour-magnitude diagram presented in Fig. 2 by Cantat-Gaudin et al. (2018) shows several different main sequences probably due to foreground stars.

As the next step, we can analyse the distribution of projected widths. It should be noted that many clusters in this data set display different behaviour than normal. However, a Gaussian fit is still a viable option. It is also interesting to note that in the histogram of those values (Fig. 9) we see that the projected width of the clusters peaks at about 2.7(1.6) pc. When using the strict sample, the peak is located at 3.0(1.3) pc. This result is comparable to the 2D analysis of the cluster sizes by Kharchenko et al. (2013) which tells us that the typical size of a cluster is of the order of several parsecs. Moreover, the upper limit of absolute cluster diameters of about 25 pc (van den Bergh 2006) is also nicely supported. This upper limit is caused by the dimension of the initial molecular cloud from which the clusters are formed and the dissipation due to the differential rotation of the Milky Way (Joshi et al. 2016).

We should keep in mind that the size of the cluster, determined from projected widths, depends heavily on the procedure that was used to determine the cluster membership probabilities. A good example of this can be seen if we compare our results for NGC 2682 with another work which focuses on studying this particular cluster. While our cluster width turns out to be about 3 pc, results from Carrera et al. (2019) show that the size of the cluster is about one order of magnitude larger. This can be expected since it is known that

[1] https://webda.physics.muni.cz

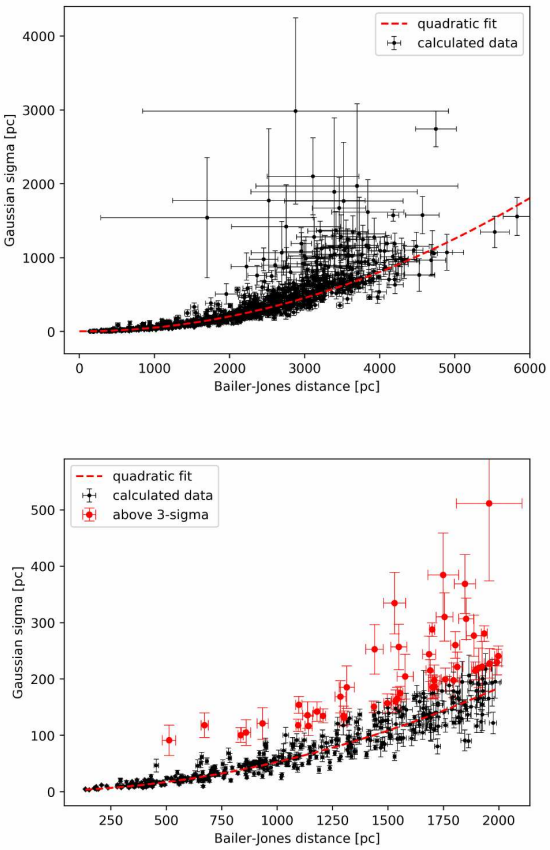


Fig. 10: The dependency of the sigma parameter on the Bailer-Jones distances for the loose sample (upper panel) and limited to 2 kpc (lower panel), respectively. The red circles in the lower panel are the 50 clusters deviating more than 3σ from the quadratic relation (red dashed line).

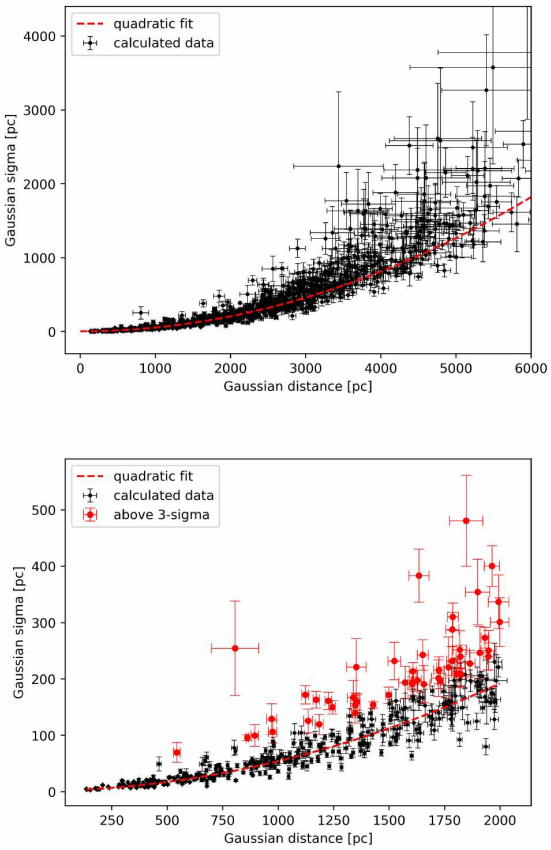
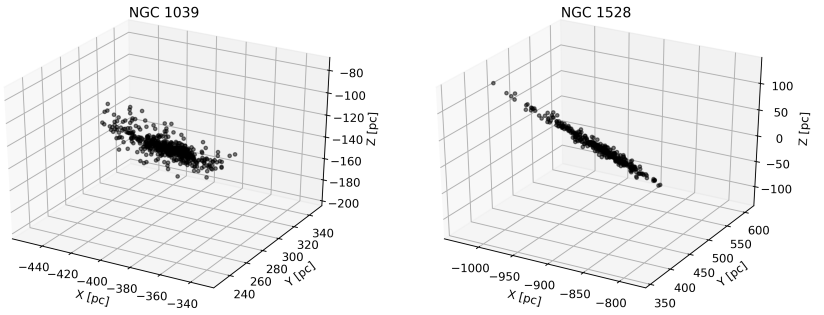
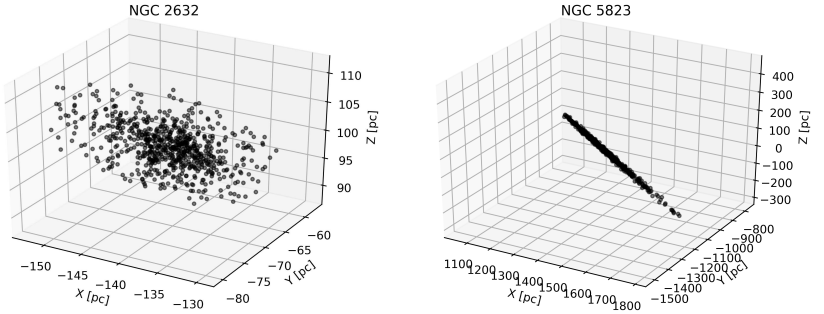


Fig. 11: The dependency of the sigma parameter on the inverse-parallax distance for the loose sample (upper panel) and limited to 2 kpc (lower panel), respectively. The red circles in the lower panel are the 50 clusters deviating more than 3σ from the quadratic relation (red dashed line).



(a) NGC 1039 at $d = 505(2)$ pc. (b) NGC 1528 at $d = 1021(4)$ pc.



(c) NGC 2632 at $d = 186(< 1)$ pc. (d) NGC 5823 at $d = 1813(11)$ pc.

Fig. 12: The 3D structures of the clusters indicated. The expected needle-like structure in the line-of-sight is clearly visible for more distant clusters.

this cluster is very old ($\log t \sim 9.5$, Bossini et al. 2019) and has experienced a significant amount of dynamical evaporation (Carrera et al. 2019). The radius of a cluster should increase with time. We conclude that our calculated widths of clusters contain some systematic errors which should be negligible for the youngest clusters and get significantly larger for much older clusters. We expect that this would affect the distribution in Fig. 9 by slightly enhancing the size of the tail toward larger projected widths at the expense of lowering the peak at lower values of the distribution.

Finally, we have also calculated the values of the sigma parameter σ_r from σ_ϖ using both inverse-parallax approach and the distances from Bailer-Jones et al. (2018), who included two parts into the prior of their analysis – the exponentially decreasing volume density term and a Galactic model term. Although their approach is not best suited for studying open clusters, it still gives us a different look at the sigma parameter (especially for the comparison with the very different inverse-parallax approach). In Fig. 10, we present the dependency of the sigma parameter σ_r on the distance r for the loose sample (upper panel), derived using the Bailer-Jones distances $r_{\mathrm{B-J}}$ (the cluster distance is calculated as median of $r_{\mathrm{B-J}}$). The most distant cluster is Teutsch 106,

with a distance of about 6 kpc from the Sun. A closer inspection yields that the data up to 2 kpc (lower panel of Fig. 10) allow to study the outliers (in those plots) in more details. A quadratic fit of the sigma parameter (SGP) in the closer inspection yields

$$\mathrm{SGP}_{\mathrm{B-J}} = 1.48(46) + 0.011(2)\, r_{\mathrm{B-J}} + 0.0000400(18)\, r^2_{\mathrm{B-J}} \ , \qquad (6)$$

with a standard error of 14.9 pc. This transforms to a SGP of [3, 7, 17, 53, 184 pc] for distances of [100, 250, 500, 1000, 2000 pc] not taking into account the derived standard deviation. In total, we found 50 open clusters which exceed 3σ above the standard line. These aggregates are good candidates for either hosting two populations in the same line-of-sight or not being true star clusters. Within 1 kpc, we find five of them: Alessi 44 (ASCC 106; most deviating case), NGC 1579, NGC 2183, NGC 6178, and vdBergh 80. For these open clusters, we find no conspicuous features. However, these 50 aggregates have to be investigated in more details using photometric data and the available results from the literature to shed more light on the inconsistencies.

On the other hand, we find several clusters which are 3σ below the standard line and are therefore very well defined. In principle, these clusters could be the best candidates for studying the individual three-dimensional structures. However, it is advisable to first analyse the colour-magnitude diagrams which should help to lower the field-star contamination.

When using the inverse-parallax approach, the situation does not change significantly (Fig. 11). We find the quadratic fit of SGP (again, for the plot in the closer inspection)

$$\mathrm{SGP}_{\mathrm{inv}} = 1.94(56) + 0.009(3)\, r_{\mathrm{inv}} + 0.0000429(22)\, r^2_{\mathrm{inv}} \ , \qquad (7)$$

with a standard error of 15.3 pc, and SGP of [3, 7, 17, 54, 192 pc] for distances of [100, 250, 500, 1000, 2000 pc]. These results are very similar to the case when we used the Bailer-Jones distances. Although the numerical results are somewhat different, the elongation of the clusters (measured by SGP) does not significantly differ from the previous method used to derive the distances. The values of $\mathrm{SGP}_{\mathrm{inv}}$ start to notably deviate from $\mathrm{SGP}_{\mathrm{B-J}}$ only at distances starting from about 2.0 kpc and beyond.

3.2. Characterizing clusters in three spatial dimensions

The three dimensional spatial structure of open clusters based on observations is very much needed for all cluster formation and evolution models (Kroupa 1995). Open questions, like the internal kinematical and spatial distributions of the members and their evolution, can only be answered by detailed observations of open clusters of different ages. Up to now, there are only very few of such investigations on the basis of *Gaia* DR2 data available (Franciosini et al. 2018, Karnath et al. 2019). This motivated us to investigate the 3D characteristics on the basis of the currently available data and their errors. Particularly, we are interested up to which distances such an analysis is meaningful.

For a given cluster, the coordinates d_1 (or r) and d_2 of its individual members form a distribution which can be displayed in histograms. These can give us insights about the spatial structure of the cluster. As we have seen, if a

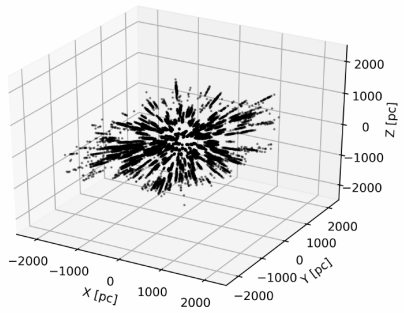


Fig. 13: The shape of the star clusters from the strict sample in the Galactic $[X\,Y\,Z]$ coordinate system.

given cluster is located within 2.0 kpc, the SGP does not depend too much on the method we use for deriving the distances. Although such approach does not allow us to study clusters in detail at large distances, we can still say something about the effect of the SGP on the 3D structure (from a statistical point of view).

Using the data from *Gaia* DR2 and the distances $r_{\mathrm{B-J}}$, we are able to create scatter plots using the $[X\,Y\,Z]$ coordinates of cluster members. For distant clusters, we would expect the 3D structure to be needle-like, because of the absolute values of the errors in distances. It is interesting to see, that we can see this structure quite clearly even at distances <750 pc. Nearby highly-populated clusters (like NGC 2632) with $r < 250$ pc, on the other hand, show only a very weak elongation in the line-of-sight. This tells us something about how even the best data (we currently have) limit the investigation of 3D structures of open clusters. In Fig. 12, we show the situation for four clusters: NGC 1039 (distance of 505(2) pc), NGC 1528 (1021(4) pc) NGC 2632 (186(< 1) pc), and NGC 5823 (1813(11) pc). The expected needle-like structure in the line-of-sight is clearly visible. One has to keep in mind, that the apparent members of NGC 5823 are spread about 160 pc around the Galactic disk, for example. To investigate this topic further, we have plotted all star clusters from the strict sample in the Galactic $[X\,Y\,Z]$ coordinate system (Fig. 13) which corresponds to Fig. 2 shown in Ward et al. (2020). The described limitations might be one of the reason why we are not able to precisely trace the spiral arms with open clusters (see Fig. 11 in Cantat-Gaudin et al. 2018).

We would like to show the impact of those results (using the loose sample) on the estimates of 3D radii of the individual clusters. We have fitted each cluster with a sphere which contains 50% of the population of a given cluster (the centre of the cluster is taken to be the median value of the $[X\,Y\,Z]$ coordinates). These spheres can be then compared with the reference sphere of radius 12 pc and we can search for the sphere which is smaller than (or

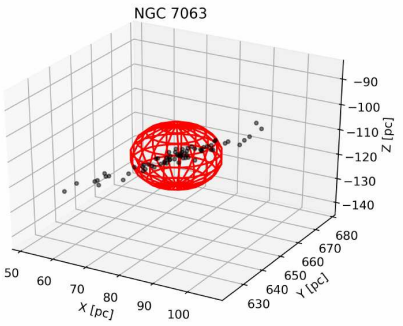


Fig. 14: The 3D structure of NGC 7063 at $d = 665(2)$ pc for all stars with a membership probability larger than 50%. The elongation is the result of uncertainties in the observed parallaxes. Included is a sphere with a radius of 12 pc centered at the star cluster.

approximately equal to) 12 pc and is the most distant. In Fig. 14 we show the case of NGC 7063 with a distance of 665(2) pc from the Sun. All stars with a membership probability larger than 50 % are included in this figure. Again, the distribution of the members is not spherical but more needle-like. This should serve as an upper limit to the distance up to which we can still somewhat fit spheres to open clusters. However, it has to be emphasized that the true internal 3D structure of such an open cluster cannot be studied. This is mainly due to the elongation resulting from the observational uncertainties. However, there is also a secondary effect that has to be taken into account – the bias introduced by the search of members in the line-of-sight. In the future, algorithm searching for members of more distant star clusters should transform the astrometrical data in the $[X\,Y\,Z]$ space and not using direct line-of-sight distances. This would guarantee to search for members in a three-dimensional space around the cluster centre and overcome possible (although small) selection effects.

In this respect we have to think about the definition of a star cluster and the differences to a moving group. Recently, Faherty et al. (2018) presented an analysis of a co-moving catalogue including 4555 groups of stars (10606 individual objects). Questions arise like what is the lowest number of members and the lowest total mass of a star cluster, for example. And how can we distinguish between a moving group and a star cluster? It seems that the distribution of the stars (central agglomeration) is only a poor criterion to do so. Here, new comprehensive methods are needed.

It is expected that star clusters should have a finite lifetime resulting from the dynamical evaporation process (Chumak et al. 2010).

Another well known effect are tidal tails of star clusters. It is expected that star clusters should have a finite lifetime, resulting from the dynamical

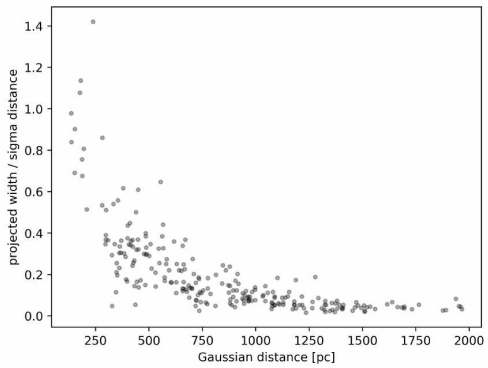


Fig. 15: The plot of a measure of ellipticity of clusters against the distance of the clusters (strict sample).

evaporation process (Chumak et al. 2010). Especially for old open clusters, the total mass at birth is difficult to establish because of the member loss over several hundreds of millions of years. But estimating this parameter is important for putting constraints on the star forming rate in the Milky Way. Recently, tidal tails were investigated for the close open clusters Gamma Velorum (Franciosini et al. 2018), Hyades (Röser et al. 2019), and Praesepe (Röser & Schilbach 2019), for example. The search for such tidal tails is not straightforward, because according to models they could reach up to a length of about 800 pc, as in the case of the Hyades (Ernst et al. 2011). To find these tails, one has to use a method calculating the space velocities to a convergent point. However, because measured radial velocities are missing for the vast majority of stars, the before mentioned works rely on criteria solely based on tangential velocities. If we investigate our sample, in the case of most clusters the distance errors are far too high for performing such an analysis. A good example is M 67 (distance of about 800 pc from the Sun), for which Carrera et al. (2019) reported an extended halo of up to 150 pc which transforms to 10^{deg} on the sky. Such large extended searches for cluster members meet the limits for any automatic method. We also have to emphasize that the needle-like structures seen in Figs. 12 and 13 are not due to tidal tails but due to the uncertainties in observed parallaxes of the individual members of star clusters.

In order to further analyse the limitations of the data for studying clusters we may need to find a measure of the ellipticity of clusters. For this, we have chosen to take the ratio of projected widths and distance sigma parameters (based on r_{inv}). In Fig. 15 we see how this measure behaves as a function of the distance. We have used the strict sample to get the best possible results. It can be clearly seen that the clusters from our sample significantly differ from spherical symmetry already at 500 pc.

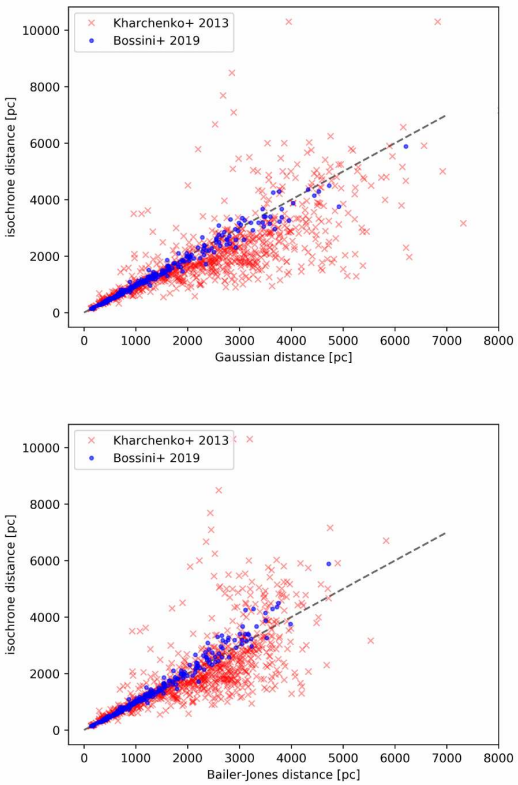


Fig. 16: Comparison of the cluster distances derived using two different isochrone fitting techniques with the distances calculated from the inverse values of the central locations of the parallax distributions (upper panel) and from the median of the Bailer-Jones distances (lower panel). The black dashed line represent one-to-one correlaton.

3.3. Comparison with isochrone fitting techniques

In order to fully understand the quality of *Gaia* DR2 astrometric data, we would like to compare our results with the distances from literature that were derived using isochrone fitting techniques. Isochrones present us an option of comparing distances calculated from two independent methods. For this purpose, we have taken the data from Kharchenko et al. (2013) and Bossini et al. (2019) and compared their distances with the distances we found using $r_{\rm inv}$ and $r_{\rm B-J}$.

In the upper panel of Fig. 16 we see that the cluster distances determined from the inverse-parallax approach very well correlate with the distances from isochrone fitting. However, there are some apparent differences. First of all, there is an apparent offset between the distance values from Bossini et al. (2019) and those we calculated. When compared with Kharchenko et al. (2013), our distances seem to be somewhat over-estimated, which is especially clear at distances larger than 2.0 kpc.

The lower panel of Fig. 16 shows us that the Bailer-Jones distances are better correlated with those from Kharchenko et al. (2013) than in the previous case. The offset is gone when plotted against the data from Bossini et al. (2019), but the correlation at larger distances appears to be worse – here $r_{\rm B-J}$ seems to be somewhat underestimated.

Finally, we would like to verify our suspicion that the values of cluster diameters (discussed above) are underestimated. The main reason for our assumption is the dissipation of clusters – for older clusters, we would expect much higher values of diameters. We can check this by looking at the diameters and ages derived by Kharchenko et al. (2013). The problem is that we cannot simply look at the differences of diameters since the definitions of the cluster radii (and diameters) in Kharchenko et al. (2013) differ from the approach we used in this work (projected widths, discussed in previous subsections). Instead, we want to see how the standard deviations of the diameter differences at a given range of ages depend on the logarithmic age. We have plotted this relationship (Fig. 17, upper panel) for all three radii defined in Kharchenko et al. (2013). This result seems to confirm our suspicion – it seems that the cluster members taken from Fig. 11 represent only the core population of the studied clusters. However, the disagreement between the distances (Fig. 16, upper panel) is going to affect this result. In Fig. 17 (lower panel), we have plotted the angular diameters of the cluster in the same way as before. In this case, there is no clear scatter at $\log\,(\mathrm{Age}) > 8.5$.

3.4. Simulating SGP

We predict that the shape of the function SGP(r) is determined by the parallax uncertainty. This can be easily verified by simply simulating a number of clusters at a random distance. We have chosen to simulated 500 clusters (containing between 100 and 300 members) at distances between 100 pc and 2000 pc, which corresponds to the region shown in the lower panels of Fig. 10 and Fig. 11. The cluster radii were chosen to be 5 pc.

When the clusters are created ("real" parallaxes are found by inverting the simulated "real" distances), "observed" parallaxes are simulated using a random (normal) function based on two possibilities – either the absolute or

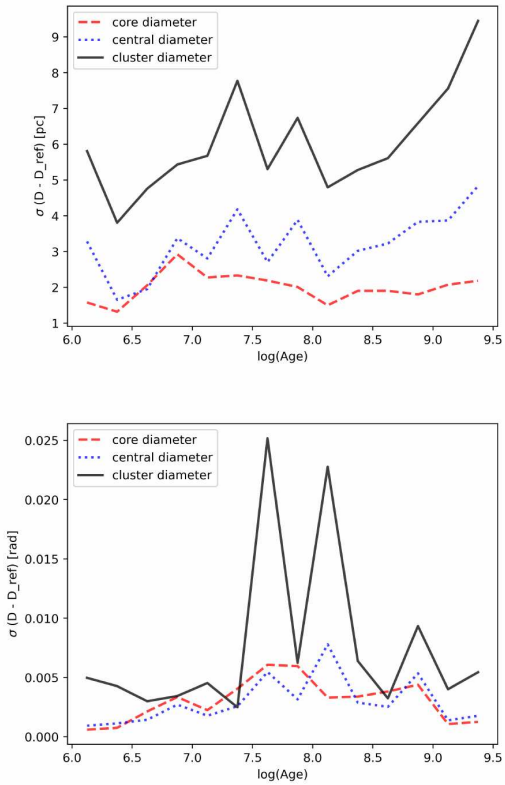


Fig. 17: Scatter of the differences between diameters determined in this work and those from Kharchenko et al. (2013). The scatter is measured by the standard deviation of diameter differences at a given range of logarithmic ages (step $\Delta \log(\mathrm{Age}) = 0.25$). The upper panel shows absolute diameters. In the lower panel, angular diameters are displayed.

the relative error is set to be a constant. Afterwards, the distances and SGP values are determined by fitting Gaussian functions, as described previously.

The results for $\sigma_{\varpi} = 0.05$ mas and $\sigma_{\varpi}/\varpi = 0.05$ are displayed in Fig. 18. The first situation coincides (almost exactly) with the results of our analysis of the clusters from Cantat-Gaudin et al. (2018) – the quadratic relation is clearly the result of the absolute parallax uncertainties. The larger spread in the observed data, when compared with the simulated data, is most likely the result of variations in the value σ_{ϖ} (an estimate of the distribution of these values is shown in Fig. 4).

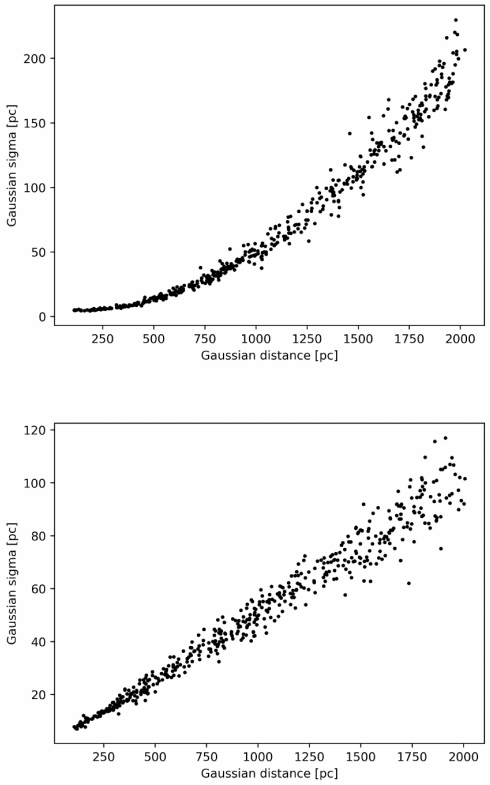


Fig. 18: The relation between the SGP and the distance of a cluster based on a simulation. The upper panel shows the example of a constant absolute parallax error ($\sigma_{\varpi} = 0.05$ mas). The lower panel assumes a constant relative parallax error ($\sigma_{\varpi}/\varpi = 0.05$).

Conclusions

With the most recent *Gaia* DR2, it is now possible to study nearby open clusters in more details. Especially the internal structure and kinematical characteristics are still only known for a very few clusters like the Hyades and Pleiades. But these characteristics are very important as input parameters for models dealing with the formation and evolution of star clusters.

Cantat-Gaudin et al. (2018) studied 1229 open clusters and derived membership probabilities of stars as well as cluster distances and diameters based on astrometrical and kinematical data. We used the cluster members from these data to study the limitations of the *Gaia* DR2 when it comes to study open clusters. The distances in this work were determined by using the most typical procedures – the inversion of parallaxes and the Bayesian method with decreasing volume density prior. For the second procedure, we used the values presented by Bailer-Jones et al. (2018) who also included a Galactic model in their prior which influences the distances of the open clusters by slightly underestimating these values. The comparison of the two data sets of distances with the isochrone fitting methods shows that the calculated distances are in a good agreement. Together with the simulations of clusters presented in Sect. 2, this shows that the distances used in this work are quite reliable, at least in a statistical sense.

Due to the uncertainties in observed parallaxes, most of the clusters have needle-like shapes and are not even close to being spherical, which can be expected when comparing with the results from Luri et al. (2018). We conclude that this affects the determination of distances not only when using inverted parallaxes but also when the Bayesian approach with decreasing volume density prior is applied. The use of the Galactic model in the prior in Bailer-Jones et al. (2018) seems to have little to no effect on the apparent elongation of clusters along the line of sight. It is possible that the situation will improve when a better prior is used, like the one mentioned in Carrera et al. (2019).

With the current available data, the diameters of open clusters can be well studied up to about 2 kpc (using a statistical approach). The results of the overall distribution are in line with the current models showing that all clusters have diameters less than 20 pc with a peak value lying between 2 and 4 pc. However, this result depends critically on the method used to determine the cluster membership probabilities. Furthermore, we find that individual open clusters beyond 500 pc should not be considered for 3D studies with the most widely used parallax-to-distance transformation methods.

Comparison of the derived cluster distances with isochrone fitting methods shows that both approaches give statistically very similar results (except when we try to compare distances from *Gaia* with isochrones derived from older data sets). Looking at the comparison of the derived projected widths with the diameters from Kharchenko et al. (2013), we find no evidence that would show an expected systematic increase of the cluster diameters with the increasing cluster ages. The most likely explanation is that both sets of cluster members, those from Kharchenko et al. (2013) and Cantat-Gaudin et al. (2018), fail to include the outermost members.

The work by Cantat-Gaudin & Anders (2020) provided additional clusters when compared to Cantat-Gaudin et al. (2018). However, the previous clusters remain unchanged. For this reason we argue that the inclusion of the updated

data set should not significantly change the statistical results of this analysis. On the other hand, the data for the individual stars in clusters have slightly changed in the recently released *Gaia* EDR3 (Gaia Collaboration 2021). From the statistical point of view, we do not expect anything to change, although this prediction has to be verified once the new set of clusters (based on the new data) has been released.

With the new data sets (e.g. *Gaia* DR2 or EDR3), the definitions of an open cluster and of a moving group have to be revised. Quantities like the lower limit of the number of cluster members and total masses have to be assessed anew.

Acknowledgements

This research has made use of the WEBDA database, operated at the Department of Theoretical Physics and Astrophysics of the Masaryk University, the SIMBAD database, operated at CDS, Strasbourg, France and NASA's Astrophysics Data System. This work presents results from the European Space Agency (ESA) space mission Gaia. Gaia data are being processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia MultiLateral Agreement (MLA). The Gaia mission website is https://www.cosmos.esa.int/gaia. The Gaia archive website is https://archives.esac.esa.int/gaia.

References

Bailer-Jones C.A.L., 2015, *PASP*, 127, 994
Bailer-Jones C.A.L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, *AJ*, 156, 58
Bossini D., Vallenari A., Bragaglia A. et al., 2019, *A&A*, 623, A108
Cantat-Gaudin T., Anders F., 2020, *A&A*, 633, A99
Cantat-Gaudin T., Donati P., Vallenari A., Sordo R., Bragaglia A., Magrini L., 2016, *A&A*, 588, A120
Cantat-Gaudin T., Jordi C., Vallenari A. et al., 2018, *A&A*, 618, A93
Cantat-Gaudin T., Krone-Martins A., Sedaghat N. et al., 2019, *A&A*, 624, A126
Carrera R., Pasquato M., Vallenari A. et al., 2019, *A&A*, 627, A119
Chumak Y.O., Platais I., McLaughlin D.E., Rastorguev A.S., Chumak O.V., 2010, *MNRAS*, 402, 1841
Ernst A., Just A., Berczik P., Olczak C., 2011, *A&A*, 536, A64
Faherty J.K., Bochanski J.J., Gagné J. et al., 2018, *ApJ*, 863, 91
Franciosini E., Sacco G.G., Jeffries R.D., Damiani F., Roccatagliata V., Fedele D., Randich S., 2018, *A&A*, 616, L12
Gaia Collaboration, Brown A.G.A., Vallenari A., Prusti T., de Bruijne J.H.J., Babusiaux C., Biermann M., 2021, *A&A*, 649, A1
Joshi Y.C., Dambis A.K., Pandey A.K., Joshi S., 2016, *A&A*, 593, A116
Karnath N., Prchlik J.J., Gutermuth R.A. et al., 2019, *ApJ*, 871, 46
Kharchenko N.V., Piskunov A.E., Schilbach E., Röser S., Scholz R.D., 2013, *A&A*, 558, A53
Krone-Martins A., Moitinho A., 2014, *A&A*, 561, A57
Kroupa P., 1995, *MNRAS*, 277, 1522
Küpper A.H.W., Kroupa P., Baumgardt H., Heggie D.C., 2010, *MNRAS*, 401, 105
Luri X., Brown A.G.A., Sarro L.M. et al., 2018, *A&A*, 616, A9
Monteiro H., Dias W.S., 2019, *MNRAS*, 487, 2385
Netopil M., Paunzen E., Carraro G., 2015, *A&A*, 582, A19
Parker R.J., Wright N.J., 2016, *MNRAS*, 457, 3430
Röser S., Schilbach E., 2019, *A&A*, 627, A4

Röser S., Schilbach E., Goldman B., 2019, *A&A*, 621, L2
Schönrich R., McMillan P., Eyer L., 2019, *MNRAS*, 487, 3568
van den Bergh S., 2006, *AJ*, 131, 1559
Ward J.L., Kruijssen J.M.D., Rix H.-W., 2020, *MNRAS*, 495, 663
Yontan T., Bilir S., Bostanc'i Z.F. et al., 2019, *Ap&SS*, 364, 152

Chapter 6

Aggregates Of Clusters In The Gaia Data

***Piecka, M.**, Paunzen, E., 2021, Astronomy & Astrophysics, Volume 649, id.A54, 12 pp.*

When studying the parallax distributions of cluster members provided by Cantat-Gaudin & Anders (2020), one will notice that the parallaxes provided by the authors do not match the distributions for some of the clusters. An example of this discrepancy can be seen for clusters RSG 7 and RSG 8. When we combine the members of both clusters, the resulting parallax distribution can be well-characterised by a double-Gaussian distribution. However, the mean cluster parallaxes provided by Cantat-Gaudin & Anders (2020) differ substantially from what such an analysis provides. The same can be seen for several other clusters.

The issue lies in the approach that was used for deriving membership probabilities. At least in the current version, the UPMASK tool provides very good results only when a single over-density is located among the population of field stars. By looking at the example of the two RSG clusters, an issue may arise when two clusters are grouped together. Based on the results provided below, it should be recommended that such groups of clusters should be analysed using a different tool. A big problem can be seen, from the mathematical point of view, in the membership probabilities for a single chosen star – one should expect that $p \leq 1$, but it is easy to show that a substantially high number of the derived cluster members have $p > 1$.

It must be noted that this issue should appear only if the two (or even more) clusters overlap in the phase space (usually only the reduced 5D phase space is analysed, as the radial velocities as mostly unknown). If two clusters overlap in the parallax distribution, they may still be distinguishable and the cluster should share no (or only a few) assigned members.

MP studied the reduced phase space of open clusters from the recently published catalogue, searched for the aggregates of clusters, analysed the CMDs and the other diagrams, was responsible for writing the larger part of the paper, and submitted the manuscript.

A&A 649, A54 (2021)
https://doi.org/10.1051/0004-6361/202040139

Astronomy & Astrophysics

Aggregates of clusters in the *Gaia* data

M. Piecka and E. Paunzen

Department of Theoretical Physics and Astrophysics, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic
e-mail: 408988@mail.muni.cz

Received 16 December 2020 / Accepted 10 February 2021

ABSTRACT

Context. The precision of the parallax measurements by *Gaia* is unprecedented. As of *Gaia* Data Release 2, the number of known nearby open clusters has increased. Some of the clusters appear to be relatively close to each other and form aggregates, which makes them interesting objects to study.
Aims. We study the aggregates of clusters which share several of the assigned member stars in relatively narrow volumes of the phase space.
Methods. Using the most recent list of open clusters, we compare the cited central parallaxes with the histograms of parallax distributions of cluster aggregates. The aggregates were chosen based on the member stars which are shared by multiple clusters.
Results. Many of the clusters in the aggregates have been assigned parallaxes which coincide with the histograms. However, clusters that share a large number of members in a small volume of the phase space display parallax distributions which do not coincide with the values from the literature. This is the result of ignoring a possibility of assigning multiple probabilities to a single star. We propose that this small number of clusters should be analysed anew.

Key words. open clusters and associations: general – astrometry

1. Introduction

Assuming that the stellar members of a given open cluster are born from a molecular cloud which is chemically almost homogeneous, one can determine the distance, reddening, age, and metallicity for these stars with the help of statistical methods (for example, isochrone fitting techniques). The latter two quantities are especially difficult to determine for single stars. This shows that galactic star clusters represent an important tool for studying the evolution and properties of stars and the Milky Way itself.

Binary clusters are especially interesting for studying the formation and evolution on a global scale. The pairs could either be formed at the same time (Priyatikanto et al. 2016) or sequentially caused by stellar winds or supernova shocks generated by one cluster inducing the collapse of a nearby cloud, thus triggering the formation of a companion cluster (Goodwin 1997). Furthermore, they can be formed completely separately and then captured either by tidal forces (van den Bergh 1996) or resonant trapping (Dehnen & Binney 1998).

Since its first data release, the *Gaia* satellite presented an improvement in the measured parallaxes of stars and expanded the list of known stars by several orders of magnitude. With the use of *Gaia*'s most recent accurate astrometric and photometric measurements (Gaia Collaboration 2018), it is not only possible to study clusters at larger distances but also to refine the list of nearby clusters. Cantat-Gaudin et al. (2018, CG18 from now on) present a status report for 1229 open cluster based on the *Gaia* Data Release 2 (*Gaia* DR2). They establish the parallaxes, the proper motions, and the most likely distances of the clusters, together with the membership probabilities of the individual stars (based on UPMASK). These data were later expanded and currently describe 1481 clusters (Cantat-Gaudin & Anders 2020, CG20).

Soubiran et al. (2018) examined the kinematical properties in the 6D phase space of open clusters derived from *Gaia* DR2 data. They have confirmed that the velocity distribution of clusters coincides well with the velocity distribution of stars at given Galactic locations. Furthermore, they have shown that several clusters seem to be confined to the same (small) volume of the phase space. This suggests that there may be a physical connection between the clusters.

Any procedure used for detecting clusters and deriving cluster membership probabilities is expected to have some limits. Using only astrometric measurements (stellar coordinates, proper motion, and parallax) necessarily leads to the assignment of non-member stars to the clusters, although Bayesian statistics may help to constrain the number of such false identifications. For the parallax space, this is discussed in Luri et al. (2018), for example. Additional information can be extracted from the colour-magnitude diagrams of the clusters, which may further constrain the membership probabilities. Overall, the final list of cluster members will never completely resemble the true host cluster – the quality of such a representation will depend on the number of detected cluster members (signal), the number density of the surrounding field stars (noise), and instrumental errors (random and systematic).

Based on the arguments above, it is possible for an automated procedure to fail under special circumstances. An example would be a case of two apparently (or physically) neighbouring clusters which may overlap in the position, parallax, and proper motions space – this was already predicted by the creators of the UPMASK method (Krone-Martins & Moitinho 2014). The only assumption used by this procedure is that the distribution of stars in the phase space is more tight than in the case of a random (uniform) distribution. The procedure is composed of two main steps – the identification of clusters in the data (based on the *k*-means clustering method) and the comparison of the groupings

Table 1. Aggregates used for the detailed analysis.

Aggregate	Cluster	l [deg]	b [deg]	ϖ [mas]	μ_α [mas yr^{-1}]	μ_δ [mas yr^{-1}]
Agg02	ASCC 19	204.914	−19.438	2.768	+1.152	−1.234
	Gulliver 6	205.246	−18.138	2.367	−0.007	−0.207
	UBC 17a	205.335	−18.019	2.753	+1.582	−1.200
	UBC 17b	205.142	−18.179	2.376	+0.078	−0.163
Agg07	Alessi 5	288.058	−1.966	2.501	−15.411	+2.503
	BH 99	286.585	−0.592	2.225	−14.494	+0.919
Agg17	COIN-*Gaia* 40	174.048	−0.794	0.470	+0.393	−2.762
	Gulliver 53	173.226	−1.141	0.384	+0.401	−2.837
	Kronberger 1	173.106	+0.049	0.443	−0.050	−2.199
	NGC 1893	173.577	−1.634	0.267	−0.231	−1.410
	Stock 8	173.316	−0.223	0.446	+0.094	−2.249
Agg32	FSR 0686	156.835	−2.212	1.087	−1.088	−2.556
	UBC 55	156.823	−2.190	1.079	−1.067	−2.604
Agg42	Gulliver 56	185.664	+5.915	0.458	+0.534	−3.240
	UBC 73	185.661	+5.918	0.448	+0.455	−3.210
Agg53	RSG 7	108.781	−0.320	2.337	+4.927	−1.865
	RSG 8	109.150	−0.484	2.210	+5.343	−1.654

Notes. The presented values are the mean values for the whole clusters taken from CG20.

with random distributions (returning a binary value which states whether the grouping is a cluster or not). The process is applied several times, and the membership probability of each star is determined based on the ratio $\frac{\text{number of positive assignments}}{\text{total number of iterations}}$.

If some of the cluster members are assigned to more than one cluster, one must be cautious with the calculation of the cluster parameters. We find it important to re-examine the data presented by CG20. Our goal is to investigate the number of duplicates appearing in the data set and determine their impact on the derived cluster parameters. However, we do not aim to search for new cluster groups.

2. Data selection

The data set from CG20 consists of two data tables. The list of clusters (1481 entries) includes the number of members for each cluster, derived positions, proper motions, and parallaxes. The list of cluster members (435 833 entries), for which the measured *Gaia* values of astrometric and some photometric measurements are available, presents the derived membership probabilities and IDs of the host clusters.

Based on the IDs of member stars, we have found 133 clusters containing stars which are also assigned to at least another cluster. The occurrence of these 'duplicates' (2148 individual stars) is the sign of an overlap in the reduced phase space (leaving out radial velocities) in which the determination of the membership probabilities was executed. This overlap is mostly attributed to the measurement uncertainties. Very rarely, this could also be a hint that the host clusters may be physically close to each other and/or even overlapping. In such cases, the determination of the membership probability requires a more complex approach.

For the probability of a given star being a member of at least one of the assigned clusters, the sum of the probabilities for all assigned clusters must be considered. We find that 1011 of the 2148 duplicate stars have summed probabilities of $p_{\rm sum} > 1.0$ (of these, 456 stars with $p_{\rm sum} > 1.5$ and one star with $p_{\rm sum} > 2.0$). Clearly, the approach used for the membership probability evaluation does not take a possible assignment to several clusters into account. For the purpose of our analysis of some of these cluster aggregates, we simply assume that all stars assigned to the aggregates (without duplicity) are members.

We estimate 60 aggregates from the sample of 133 clusters (Table A.1). This was done with the use of transitivity – if cluster A shares a member with cluster B, and if cluster B shares a member with cluster C, then all three clusters form an aggregate. The found aggregates can be distinguished based on the number of included clusters. In 52 instances, the aggregate consists of only two clusters. There are also four aggregates composed of three clusters, three aggregates containing four clusters, and one case of a five-cluster aggregate.

We have chosen an example of double-Gaussian distributions in the parallax space (Agg07), which serves as a reference point. There are at least four aggregates (Agg02, Agg32, Agg42, and Agg53) for which the constituent clusters overlap in the whole phase space in such a way that they are very hard to distinguish from each other. These were chosen in such a way that at least two clusters follow these rules: The difference between the median values of their proper motions (each component) is less than half the sum of the standard deviations; the same was applied for the positions; and for parallaxes, the full sum of the standard deviations was used. Moreover, the five-component aggregate (Agg17) was also considered. The chosen clusters are summarised in Table 1.

Some of the stars presented in the data were assigned to the same cluster twice – such cases are highlighted by our procedure but not considered to be actual duplicates in our analysis. This situation occurred in the case of UBC 3, UBC 19, and UBC 31. For example, *Gaia* `DR2 4505874688732436992` is presented twice for the host cluster in the `VizieR` catalogue `J/A+A/633/A99/members` (CG20).

3. Case analysis

In this section, we present an analysis of the chosen aggregates. We are especially interested in comparing the mean parallaxes derived in the data source with the parallax histograms of the whole aggregates. We predict that the derivation of parallaxes

should be severely affected for aggregates in which multiple stars are being shared by the clusters.

The histograms (points located at the centre and maximum height of each bin) were fitted by a sum of two Gaussian functions – this type of model has six free parameters (peak height, central position, and width of both peaks). This was achieved with the use of the basic `scipy.optimize.curve_fit` function available in `Python`.

For the comparison of cluster parameters (age, extinction, and metallicity) of the individual objects, we used the results from Dias et al. (2002), Kharchenko et al. (2013), and Bossini et al. (2019). Most of the figures related to the analysis are presented in Appendix A.

3.1. Alessi 5 and BH 99

Alessi 5 was first mentioned in the work by Dias et al. (2002). It is a cluster of intermediate age ($\log t = 7.723$) with a good amount of extinction being present in the line of sight ($A_V = 0.592$ mag). The cluster appears to be somewhat metal-poor, with [Fe/H] = −0.133.

BH 99 is almost as old as Alessi 5 ($\log t = 7.908$). However, the estimated extinction appears to be somewhat lower ($A_V = 0.203$ mag). The metallicity of this cluster has been assumed to be solar, [Fe/H] = 0.000.

The comparison of the parallaxes from Table 1 with the histogram (fitted by a double-Gaussian) of this aggregate (Fig. 1, top panel) clearly shows that the parallaxes were derived correctly. The clusters share only one of the total number of 722 stars.

Based on the parameters listed in Table 1, it seems that these clusters are fairly close to each other. Using the Galactic coordinates listed in `VizieR`, their relative distance is found to be about 50 pc. However, when compared with the clusters analysed below, these two are fairly distant in the position on sky and noticeably differ in their mean proper motion (Fig. 1, middle panels).

This aggregate is mentioned in Soubiran et al. (2019). Due to the large difference in metallicities, the most likely assumption is that these clusters represent two distinct groups of stars, despite their proximity in the phase space. The difference can be clearly seen in the colour-magnitude diagrams of the aggregate (Fig. 1, bottom panels, adopting the membership probabilities from the source data). However, differential extinction may also play a role in the determination of the cluster parameters – given their relative distance, the difference between the derived extinction values seems to be quite large ($\Delta A_V \sim 0.25$ mag over ∼50 pc).

3.2. FSR 0686 and UBC 55

The age of FSR 0686 is not specified in Bossini et al. (2019). Dias et al. (2002) provided a value of $\log t = 8.610$, making it a fairly old cluster. Curiously, it is discussed in CG18 that the angular diameters of FSR clusters differ when looking at the catalogues from Dias and Kharchenko (Kharchenko et al. 2013).

UBC 55 is one of the newly discovered clusters, which was first presented in the source data used in this work. For this reason, its cluster parameters have not been determined yet.

The histogram of parallaxes is displayed in the top panel of Fig. A.1 – the source parallaxes slightly differ from those determined here. However, given the relatively small stellar count, fitting a single Gaussian function is also possible, leading to a parallax value which matches those from the literature.

The distance between these clusters was calculated to be less than 10 pc. Their position and proper motion appear to be almost identical (Fig. A.1, middle panels). The situation is the same when we look at the colour-magnitude diagrams (Fig. A.1, bottom panels). In principle, this makes the aggregate a good candidate for studies of binary clusters. However, 42 out of 80 stars (all have $p_{\rm sum} > 1.0$) are being shared between the clusters, which makes it difficult to prove whether this aggregate truly consists of two different groups.

3.3. Gulliver 6 and UBC 17b

The stars in Agg02 seem to form two different subgroups, which can be seen in all planes of the phase space and even in the colour-magnitude diagrams (Fig. A.2). ASCC 19 and UBC 17a seem to be fairly similar when looking at the different plots of the clusters. However, they remain relatively distinguishable in the coordinate and proper motions space. This subgroup is significantly far away in the sky from the second subgroup and has a different proper motion (similar to the case of Alessi 5 and BH 99).

Gulliver 6 and UBC 17b share a very small volume of the phase space. For this reason, we have focused our attention on this subgroup. Literature provides no age or metallicity estimates for these two clusters. They share 75 out of 412 stars, with 103 being assigned to UBC 17b. Similar to the previous aggregate, most of the members of one cluster were also derived as members of the other cluster, all of them having a summed probability larger than 1.0. We do not see a clear distinction between the clusters in the parallax distribution (Fig. A.2, top panels) – the two clusters can be easily fitted with a single Gaussian function.

UBC 17a and UBC 17b are very close to each other in the coordinates space. However, they become quite distinguished when looking at their parallaxes and proper motions.

Gulliver 6 has been previously compared to ASCC 19, and several other clusters, in the phase space. Soubiran et al. (2018) determined that the aggregate consists of five clusters. We must note that our requirements for the phase space were more strict – we only searched for the duplicates in the data. This led to the exclusion of ASCC 16, ASCC 21, and NGC 2232 from our list.

3.4. Gulliver 56 and UBC 73

Once again, there are no cluster parameters available for these clusters. Gulliver 56 has already been recognised in CG18, while UBC 73 is another newly discovered cluster. They share 50 out of 104 stars, with UBC 73 containing no star which has not been previously assigned to Gulliver 56.

The clusters occupy the same (extremely small) volume of the phase space (Fig. A.3), although they do slightly differ in the proper motion component μ_α. The histogram in the top panel of Fig. A.3 shows that the previously derived parallaxes differ from what a fit of a double-Gaussian function predicts. However, the histogram fitting procedure may be inaccurate because of the small number of cluster members.

3.5. RSG 7 and RSG 8

The RSG clusters were discovered by Röser et al. (2016) who estimated $\log t = 8.3$ for RSG 7 and $\log t = 8.5$ for RSG 8. The reddening value was found to be quite low, $E(B-V) < 0.1$ mag. No additional sources are available for the cluster parameters of these objects.

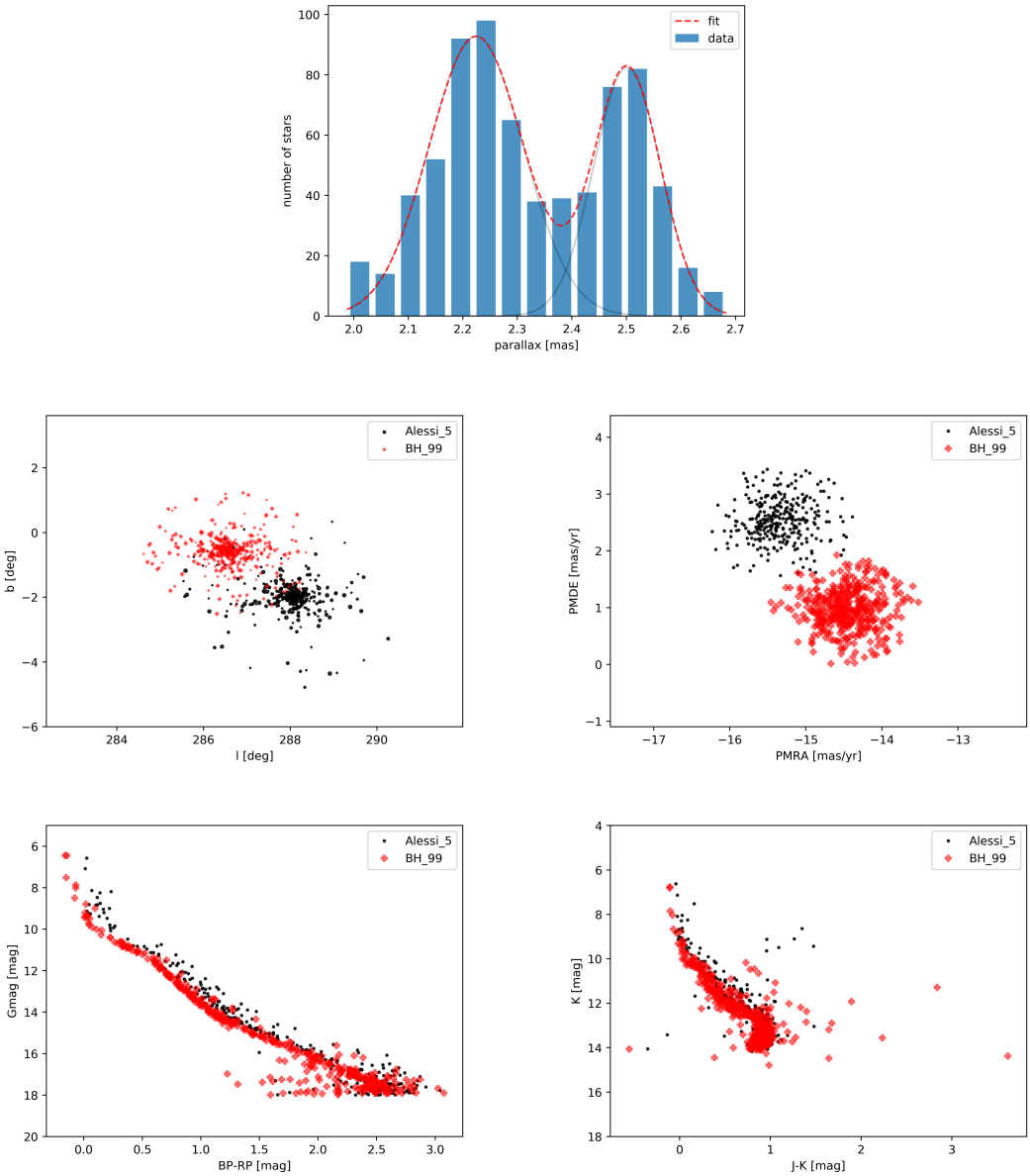


Fig. 1. Different slices through the phase space of Agg07 (Alessi 5, BH 99), together with the complete colour-magnitude diagrams (based on *Gaia* and 2MASS photometry). *Top*: histogram of parallaxes, excluding the duplicate cluster members. The best fit was achieved with the following double-Gaussian function parameters: $\varpi_1 = 2.225 \pm 0.007$ mas, $\sigma_1 = 0.087 \pm 0.007$ mas, $\varpi_2 = 2.501 \pm 0.006$ mas, $\sigma_2 = 0.061 \pm 0.006$ mas. *Middle-left*: coordinates of the stars in Agg07. Size of the points indicates values of the observed magnitude G. *Middle-right*: proper motion diagram of Agg07. *Bottom-left*: colour-magnitude diagram based on the observed values of $B_P - R_P$ and G). *Bottom-right*: colour-magnitude diagram based on 2MASS photometry. The stars were located using the coordinates in *Gaia* database. The two clusters remain distinguishable in this photometric system.

From the total of 404 stars, 143 are shared among the two clusters (117 stars have summed probabilities larger than 1.0). CG18 lists parallaxes $\varpi = 2.337 \pm 0.008$ mas for RSG 7, and $\varpi = 2.210 \pm 0.012$ mas for RSG 8. They are very well distinguished in the parallax distribution (Fig. A.4, top panel). However, we find that at least one of the clusters has a parallax ($\varpi = 2.297 \pm 0.012$ mas and $\varpi = 1.984 \pm 0.010$ mas) which does not coincide with the values provided by CG18.

In the sky, they overlap and are difficult to distinguish from each other (Fig. A.4, middle-left panel). More interesting is the proper motion space (Fig. A.4, middle-right panel), where we can see at least two concentrations of stars – a similar structure can be seen in the proper motion diagram of the tidally disrupted cluster Coma Ber in Tang et al. (2019). We find that the two concentrations are not represented well by the determined membership probabilities. However, they also do not correspond to the parallax distribution well.

Looking at the colour-magnitude diagrams (Fig. A.4, bottom panels), the two clusters are almost indistinguishable. In Soubiran et al. (2019), this aggregate was paired with another

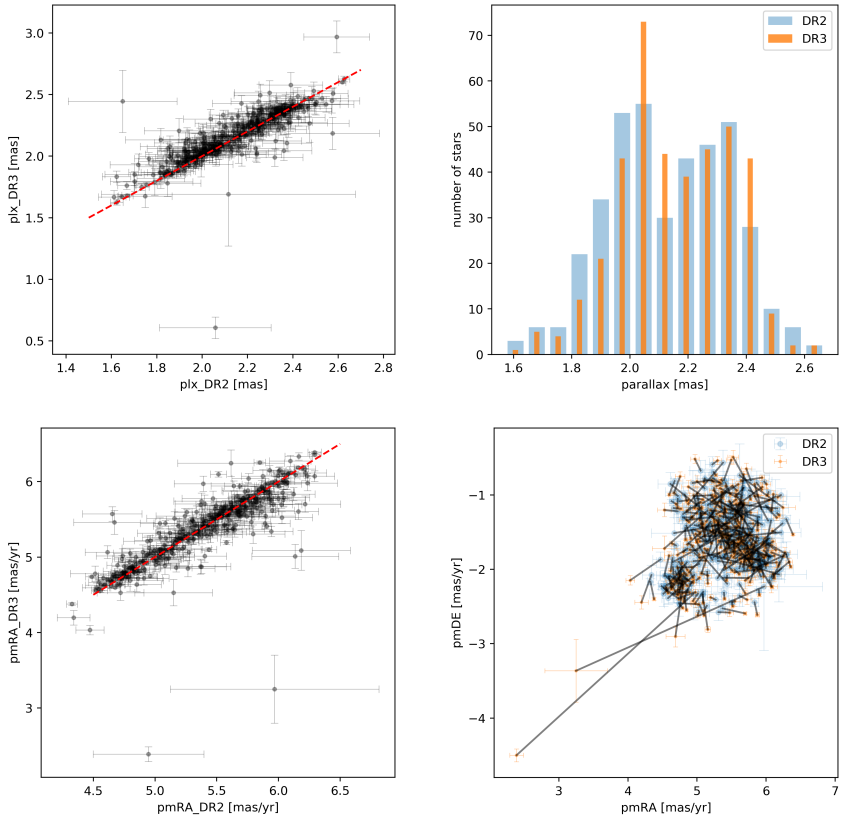


Fig. 2. Comparison of the Agg53 (RSG 7, RSG 8) cluster members determined by CG20 based on the DR2 and the EDR3 data; duplicates were filtered out. *Top-left*: parallaxes of the clusters members. The red line represents the one-to-one relation. *Top-right*: histogram of parallaxes. *Bottom-left*: μ_α component of the proper motion. The red line represents the one-to-one relation. *Bottom-right*: the proper motion diagram. The shifts of the positions of points is indicated by the black lines.

cluster, ASCC 127. It should be noted that the RSG clusters significantly differ from ASCC 127 in the colour-magnitude diagram. In our analysis, this cluster was most likely excluded due to the differences in the occupied volumes of the phase space.

3.6. Kronberger 1 and Stock 8

Three subgroups can be identified in the aggregate composed of five clusters. The overall distribution of parallaxes of this aggregate (Fig. A.5, top-left panel) is even more complicated than in the case of Agg02. However, two clusters (Fig. A.5, top-right panel) seem to deviate from the other ones.

COIN-*Gaia* 40 is a new cluster which does not share much in the phase space with the other clusters (Fig. A.5, middle panels). It has been assigned 404 stars of which only 12 are shared with the rest of the aggregate.

Gulliver 53 and NGC 1893 share only four stars with each other (none with $p_{\rm sum} > 1.0$). They have similar central coordinates but differ in parallaxes and proper motion.

Information about Kronberger 1 has been presented by Sujatha et al. (2016), who derived its age ($\log t = 7.65$) and the reddening ($E(B-V) = 0.45$ mag) in the line of sight. Cluster parameters are also available from SAI Open Clusters Catalog based on 2MASS photometry ($\log t = 8.10$, $E(B-V) = 0.43$ mag). However, the derived distance (800 pc) is significantly different from the one derived using *Gaia* photometry (>2000 pc).

Stock 8 is a young open cluster that has been known since Dias et al. (2002) determined its parameters ($\log t = 6.300$, $E(B-V) = 0.400$ mag). Kharchenko et al. (2013) present slightly different results, with a higher cluster age, $\log t = 7.050$, and reddening, $E(B-V) = 0.604$ mag. None of the works provided an estimate of the metallicity.

Kronberger 1 and Stock 8 can be somewhat distinguished in the coordinates space but CG18 give almost identical parallaxes. They share most of the members assigned to Kronberger 1, 37 out of 778 stars available in this sub-group (29 stars with $p_{\rm sum} > 1.0$). Their relative distance is estimated to be less than 20 pc. The mismatch in the derived cluster ages suggests that these clusters are quite different objects. However, it is possible that the age estimation is inaccurate for at least one of the clusters – even for the much older Kronberger 1, the spread in the colour-magnitude diagrams is fairly large (Fig. A.5, bottom panels).

Figure A.5 shows that the distribution of parallaxes does not coincide with the published values of parallaxes. Instead, our analysis puts the two clusters at much larger distances from each other, which would better explain the discrepancy between the cluster ages.

4. Comparison with *Gaia* EDR3 and 2MASS

At the beginning of December 2020, *Gaia* Early Data Release 3 (*Gaia* EDR3, Gaia Collaboration 2021; Lindegren et al. 2021) became available. G-magnitudes are now systematically shifted towards higher values (less bright). The values of parallaxes and proper motions have slightly changed, as well. However, they should still represent the same statistical result, in a given subset of stars, as before.

The lack of reliable radial velocity data means that it will be difficult to establish the binary fraction in a given subset of stars. The binaries may affect the search for the clusters – we should

expect that at least some of the members will occupy a different volume of the phase space than previously assumed.

Clearly, the membership probabilities have to be assessed anew. However, it would be interesting to look at the stars from CG20 and compare the situation between *Gaia* DR2 and *Gaia* EDR3. We searched for the cluster members of RSG 7 and RSG 8 and display the results in Fig. 2. We can see that the parallaxes are systematically shifted from the one-to-one relation towards higher values than previously believed. On the other hand, the histogram of parallaxes did not change too much – only the systematic shift towards higher values stands out. We see no systematic changes in the proper motion of the stars.

There are only a few outliers in the subplots. Therefore, we should expect that only a negligible number of cluster members is going to change (the most striking cases are seen in the proper motion diagram). Statistically, RSG 7 and RSG 8 now look similar to what we have seen in DR2. We conclude that the results of the analysis presented in this work should also remain valid in the new data release.

It is possible to search for the *Gaia* sources in other photometric catalogues based on the coordinates of the stars. We carried out this exercise for the 2MASS catalogue, where we searched for the stars of the six analysed aggregates. As we can see in the bottom-right panel of Fig. 1, Agg07 still shows a clear distinction in this photometric bands. Two populations also appear in the colour-magnitude diagram of Agg53 (Fig. A.5). For the other aggregates, the situation is much more complicated – at least for Agg32 and Agg42 (Figs. A.1 and A.3, respectively), there seems to be no distinction between the two constituent clusters.

5. Conclusions

Based on our analysis of the aggregates (derived from the stars which were assigned to multiple clusters), we conclude that the automatic procedure used in CG18 and CG20 failed to correctly determine parallaxes of some of the clusters listed in Table 1:

1. Alessi 5 and BH 99: The central parallaxes from the source work were correctly reproduced by analysing the distribution of parallaxes. We propose that the match results from the fact that the clusters are only partially overlapping in the phase space. Another important contributing factor is that only one of the stars is shared among them. Assuming that their relative distance is about 50 pc, it is interesting to note that the metallicity and the extinction (taken from the literature) differ for these clusters – this can be seen as a population difference in the colour-magnitude diagram of this aggregate.
2. FSR 0686 and UBC 55: Our fitting method gives different parallaxes than presented in the literature – this is most likely due to the large amount of duplicate stars in this aggregate. However, the total number of members is fairly low, which seriously affects the usefulness of the method we used. The distribution can be well fitted by a single Gaussian function which would yield a value similar to those from CG20.
3. Gulliver 6 and UBC 17b: In this case, the cited parallaxes differ only a little from what we see in the histogram. The shape of the parallax distribution is slightly asymmetrical. The two objects are also difficult to distinguish from each other in the coordinates space. We suggest that the aggregate could be somehow physically distorted, producing a significant deviation from normal distribution in the parallax space.
4. Gulliver 56 and UBC 73: The fitting method gives parallaxes that are quite different from those presented in the source work. However, this case is very similar to Agg32 – many stars are shared and the total number of stars is low.
5. RSG 7 and RSG 8: This aggregate is very interesting. Two populations are apparent in the parallax space and the proper motion space, but they do not seem to correspond to each other. On the other hand, there seems to be no evidence in the coordinate space or the colour-magnitude diagram that this aggregate is composed of multiple clusters. Based on our analysis, we propose an update to the parallax values presented by CG18.
6. Kronberger 1 and Stock 8: Most of the stars assigned to Kronberger 1 are also present in Stock 8. Our analysis of parallaxes puts the clusters at a larger relative distance than previously derived. This would better explain the difference between the ages found in the literature.

Most of the cluster parallaxes listed in CG20 were derived using a very robust method. However, a problem with deriving the parallaxes seems to occur when at least two clusters overlap in a small volume of the phase space and when they share most of the members. This results from ignoring a possibility of encountering more complex aggregates of clusters in the phase space. In such cases, an identification of such aggregates followed by a separate analysis of the parallaxes is required.

We propose that the membership probabilities of stars in cluster aggregates, which seem to occur less frequently than single clusters, should be calculated anew. Most importantly, the probability of a star being a member of at least one of the assigned clusters (the summed probability) should be restricted by the condition $p_{\rm sum} \leq 1.0$.

Acknowledgements. This paper is dedicated to Harald Pöhnl who died during its preparation. It was supported by an internal grant of the Masaryk University (MUNI/IGA/1530/2020). E. P. acknowledges support by the Erasmus+ programme of the European Union under grant number 2020-1-CZ01-KA203-078200. This work 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. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. This research has made use of the WEBDA database, operated at the Department of Theoretical Physics and Astrophysics of the Masaryk University.

References

Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, A&A, 623, A108
Cantat-Gaudin, T., & Anders, F. 2020, A&A, 633, A99
Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93
Dehnen, W., & Binney, J. J. 1998, MNRAS, 298, 387
Dias, W. S., Alessi, B. S., Moitinho, A., & Lépine, J. R. D. 2002, A&A, 389, 871
Gaia Collaboration (Brown, A. G. A., et al.) 2018, A&A, 616, A1
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Goodwin, S. P. 1997, MNRAS, 284, 785
Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2013, A&A, 558, A53
Krone-Martins, A., & Moitinho, A. 2014, A&A, 561, A57
Lindegren, L., Bastian, U., Biermann, M., et al. 2021, A&A, 649, A4
Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9
Priyatikanto, R., Kouwenhoven, M. B. N., Arifyanto, M. I., Wulandari, H. R. T., & Siregar, S. 2016, MNRAS, 457, 1339
Röser, S., Schilbach, E., & Goldman, B. 2016, A&A, 595, A22
Soubiran, C., Cantat-Gaudin, T., Romero-Gómez, M., et al. 2018, A&A, 619, A155
Soubiran, C., Cantat-Gaudin, T., Romero-Gómez, M., et al. 2019, A&A, 623, C2
Sujatha, S., Krishna Kumar, K., & Komala, S. 2016, Acta Astron., 66, 333
Tang, S.-Y., Pang, X., Yuan, Z., et al. 2019, ApJ, 877, 12
van den Bergh, S. 1996, ApJ, 471, L31

Appendix A: Additional material

In this section, we display the coordinate space map, the proper motion diagram, the parallax distribution, and the colour-magnitude diagram of each of the six analysed aggregates. Included are all stars belonging to all clusters of the whole aggregates (except for the parallax space, where duplicates were excluded). Moreover, the complete list of all aggregates we found is presented in Table A.1.

Table A.1. Complete list of aggregates of clusters.

Aggregate ID	List of clusters
Agg01	ASCC 105, UPK 82
Agg02	ASCC 19, Gulliver 6, UBC 17a, UBC 17b
Agg03	ASCC 88, Gulliver 29
Agg04	Alessi 2, UPK 312
Agg05	Alessi 43, Collinder 197
Agg06	Alessi 44, UBC 14
Agg07	Alessi 5, BH 99
Agg08	Alessi Teutsch 5, BDSB30
Agg09	BH 121, IC 2948, Ruprecht 94
Agg10	Barkhatova 1, Gulliver 30
Agg11	Berkeley 58, NGC 7788, NGC 7790
Agg12	Berkeley 62, COIN-*Gaia* 29
Agg13	Berkeley 81, NGC 6735
Agg14	Biurakan 2, FSR 0198, NGC 6871, Teutsch 8
Agg15	COIN-*Gaia* 16, COIN-*Gaia* 17
Agg16	COIN-*Gaia* 24, NGC 2168
Agg17	COIN-*Gaia* 40, Gulliver 53, Kronberger 1, NGC 1893, Stock 8
Agg18	COIN-*Gaia* 5, COIN-*Gaia* 6
Agg19	Collinder 106, Collinder 107, NGC 2244
Agg20	Collinder 135, UBC 7
Agg21	Collinder 220, IC 2581
Agg22	Collinder 272, NGC 5168
Agg23	Czernik 20, NGC 1857
Agg24	Czernik 31, NGC 2421
Agg25	Czernik 39, NGC 6755, NGC 6756
Agg26	Danks 1, Danks 2
Agg27	Dias 1, King 16
Agg28	FSR 0306, NGC 7086
Agg29	FSR 0448, FSR 0451
Agg30	FSR 0534, UBC 39
Agg31	FSR 0542, SAI 14
Agg32	FSR 0686, UBC 55
Agg33	FSR 0826, Teutsch 10
Agg34	FSR 1207, NGC 2345
Agg35	FSR 1315, NGC 2453
Agg36	Feibelman 1, Gulliver 23
Agg37	Gulliver 12, NGC 4103
Agg38	Gulliver 15, NGC 6561
Agg39	Gulliver 16, NGC 581
Agg40	Gulliver 3, Gulliver 4
Agg41	Gulliver 52, Trumpler 15
Agg42	Gulliver 56, UBC 73
Agg43	Haffner 18, Haffner 19
Agg44	Hogg 10, NGC 3572
Agg45	Hogg 17, NGC 5617, Pismis 19, Trumpler 22
Agg46	Kharchenko 1, Koposov 63
Agg47	King 14, NGC 146
Agg48	NGC 2194, Skiff J0614+12.9
Agg49	NGC 2318, Ruprecht 8
Agg50	NGC 436, NGC 457
Agg51	NGC 5269, SAI 118
Agg52	NGC 659, NGC 663
Agg53	RSG 7, RSG 8
Agg54	Ruprecht 100, Ruprecht 101
Agg55	Ruprecht 43, Ruprecht 44
Agg56	SAI 4, Stock 20
Agg57	Teutsch 11, Teutsch 12
Agg58	UBC 10a, UPK 169
Agg59	UPK 219, UPK 220
Agg60	UPK 429, UPK 431

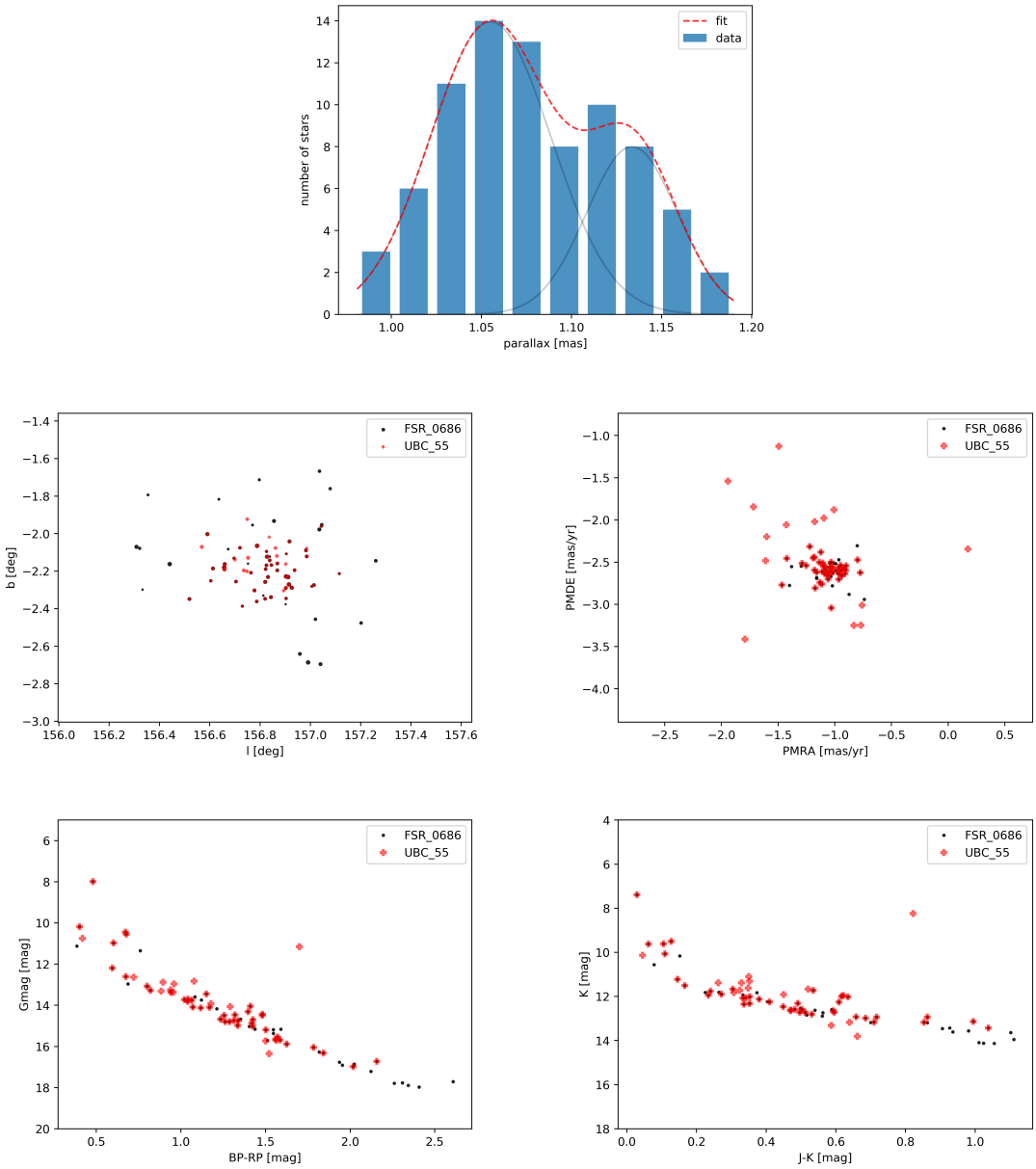


Fig. A.1. Different slices through the phase space of Agg32 (FSR 0686, UBC 55), together with the complete colour-magnitude diagrams (based on *Gaia* and 2MASS photometry). *Top*: histogram of parallaxes, excluding the duplicate cluster members. The best fit was achieved with the following double-Gaussian function parameters: $\varpi_1 = 1.055 \pm 0.005$ mas, $\sigma_1 = 0.033 \pm 0.004$ mas, $\varpi_2 = 1.134 \pm 0.007$ mas, $\sigma_2 = 0.025 \pm 0.005$ mas. *Middle-left*: coordinates of the stars in Agg32. Size of the points indicates values of the observed magnitude G. *Middle-right*: proper motion diagram of Agg32. *Bottom-left*: colour-magnitude diagram of Agg32, based on the observed values of $B_P - R_P$ and G). *Bottom-right*: colour-magnitude diagram of Agg32 based on 2MASS photometry. The stars were located using the coordinates in *Gaia* database.

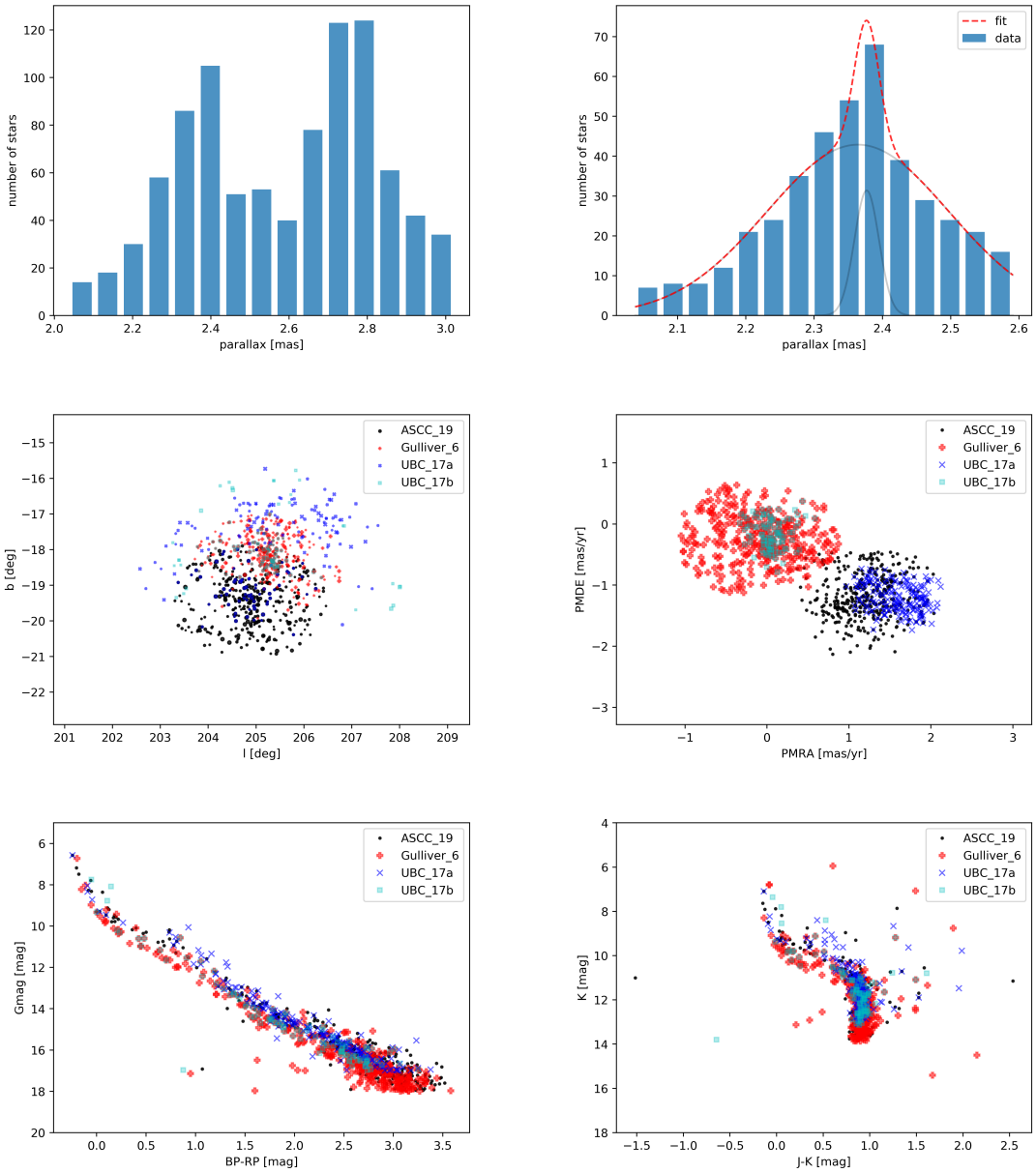


Fig. A.2. Different slices through the phase space of Agg02 (ASCC 19, Gulliver 6, UBC 17a, UBC 17b), together with the complete colour-magnitude diagrams (based on *Gaia* and 2MASS photometry). *Top-left*: distribution of parallaxes of all individual stars in Agg02. Duplicates are excluded. *Top-right*: the histogram of parallaxes, excluding the duplicate cluster members. The best fit was achieved with the following double-Gaussian function parameters: $\varpi_1 = 2.364 \pm 0.007$ mas, $\sigma_1 = 0.134 \pm 0.009$ mas, $\varpi_2 = 2.377 \pm 0.008$ mas, $\sigma_2 = 0.018 \pm 0.011$ mas. *Middle-left*: coordinates of the stars in Agg02. Size of the points indicates values of the observed magnitude G. *Middle-right*: proper motion diagram of Agg02. *Bottom-left*: colour-magnitude diagram of Agg02, based on the observed values of $B_P - R_P$ and G). *Bottom-right*: colour-magnitude diagram of Agg02 based on 2MASS photometry. The stars were located using the coordinates in *Gaia* database.

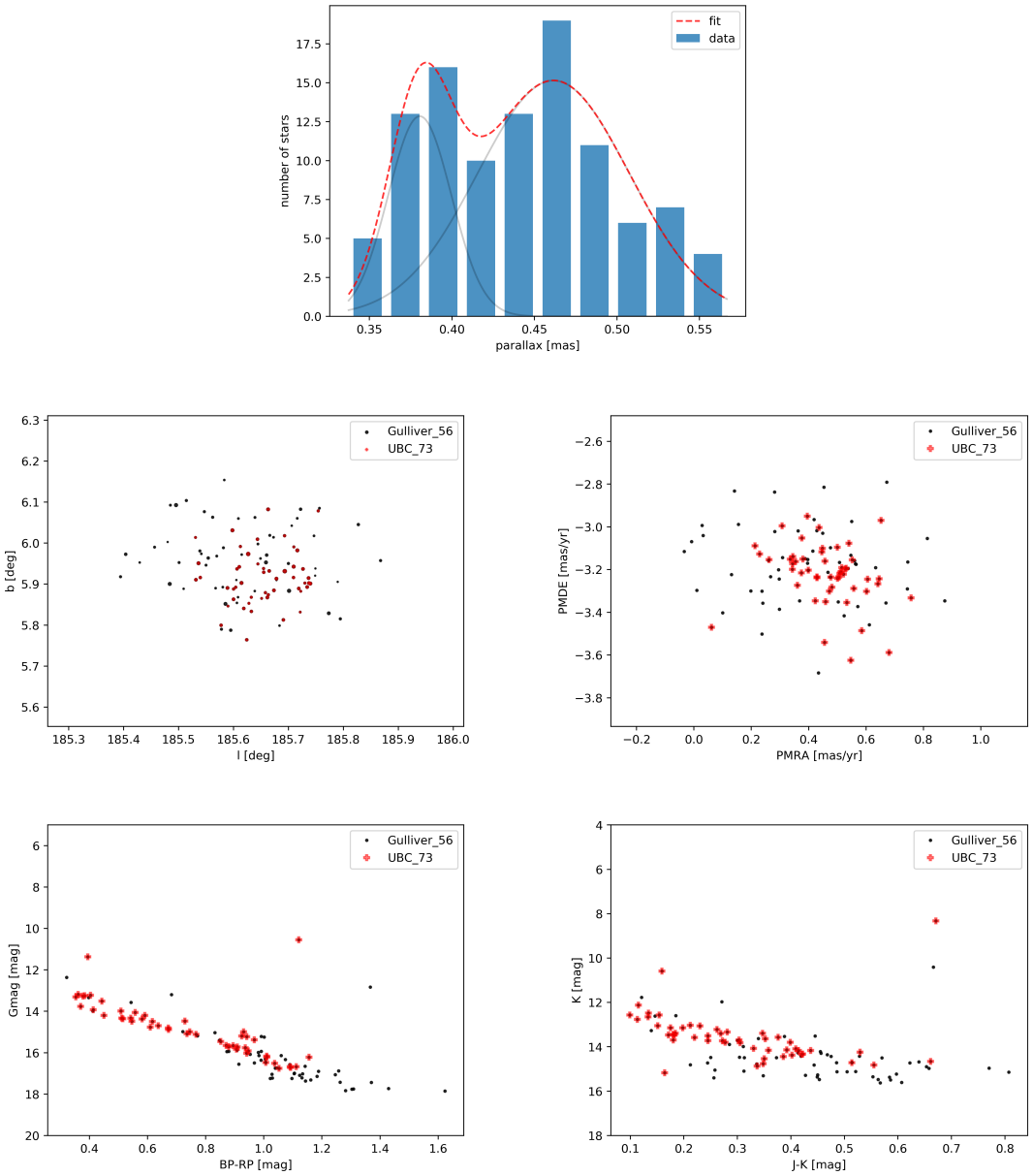


Fig. A.3. Different slices through the phase space of Agg42 (Gulliver 56, UBC 73), together with the complete colour-magnitude diagrams (based on *Gaia* and 2MASS photometry). *Top*: histogram of parallaxes, excluding the duplicate cluster members. The best fit was achieved with the following double-Gaussian function parameters: $\varpi_1 = 0.381 \pm 0.007$ mas, $\sigma_1 = 0.019 \pm 0.009$ mas, $\varpi_2 = 0.462 \pm 0.014$ mas, $\sigma_2 = 0.046 \pm 0.014$ mas. *Middle-left*: coordinates of the stars in Agg42. Size of the points indicates values of the observed magnitude G. *Middle-right*: proper motion diagram of Agg42. *Bottom-left*: colour-magnitude diagram of Agg42, based on the observed values of $B_P - R_P$ and G). *Bottom-right*: colour-magnitude diagram of Agg42 based on 2MASS photometry. The stars were located using the coordinates in *Gaia* database.

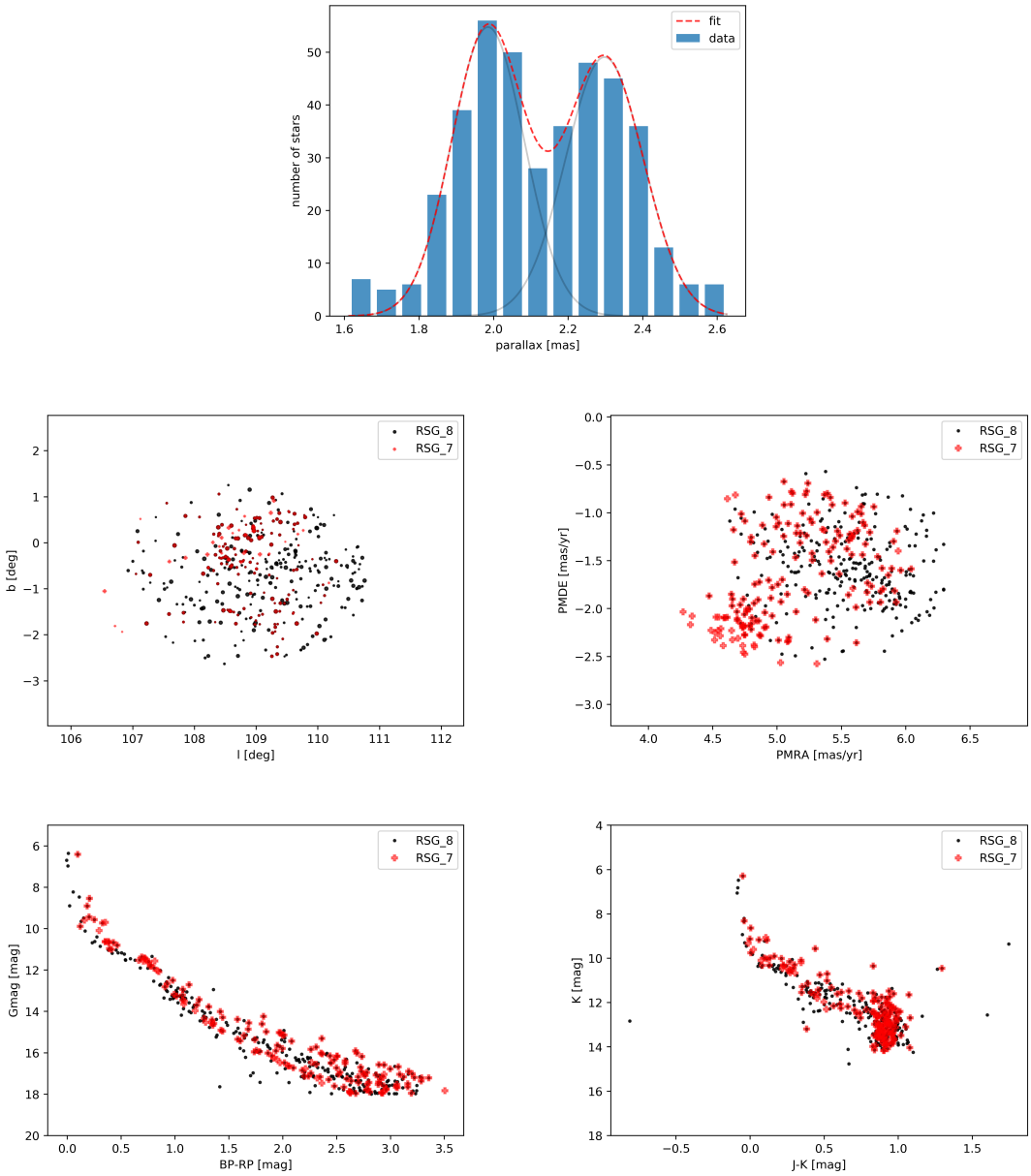


Fig. A.4. Different slices through the phase space of Agg53 (RSG 7, RSG 8), together with the complete colour-magnitude diagrams (based on *Gaia* and 2MASS photometry). *Top*: histogram of parallaxes, excluding the duplicate cluster members. The best fit was achieved with the following double-Gaussian function parameters: $\varpi_1 = 1.985 \pm 0.010$ mas, $\sigma_1 = 0.098 \pm 0.010$ mas, $\varpi_2 = 2.297 \pm 0.012$ mas, $\sigma_2 = 0.105 \pm 0.012$ mas. *Middle-left*: coordinates of the stars in Agg53. Size of the points indicates values of the observed magnitude G. *Middle-right*: proper motion diagram of Agg53. *Bottom-left*: colour-magnitude diagram of Agg53, based on the observed values of $B_P - R_P$ and G). *Bottom-right*: colour-magnitude diagram of Agg53 based on 2MASS photometry. The stars were located using the coordinates in *Gaia* database.

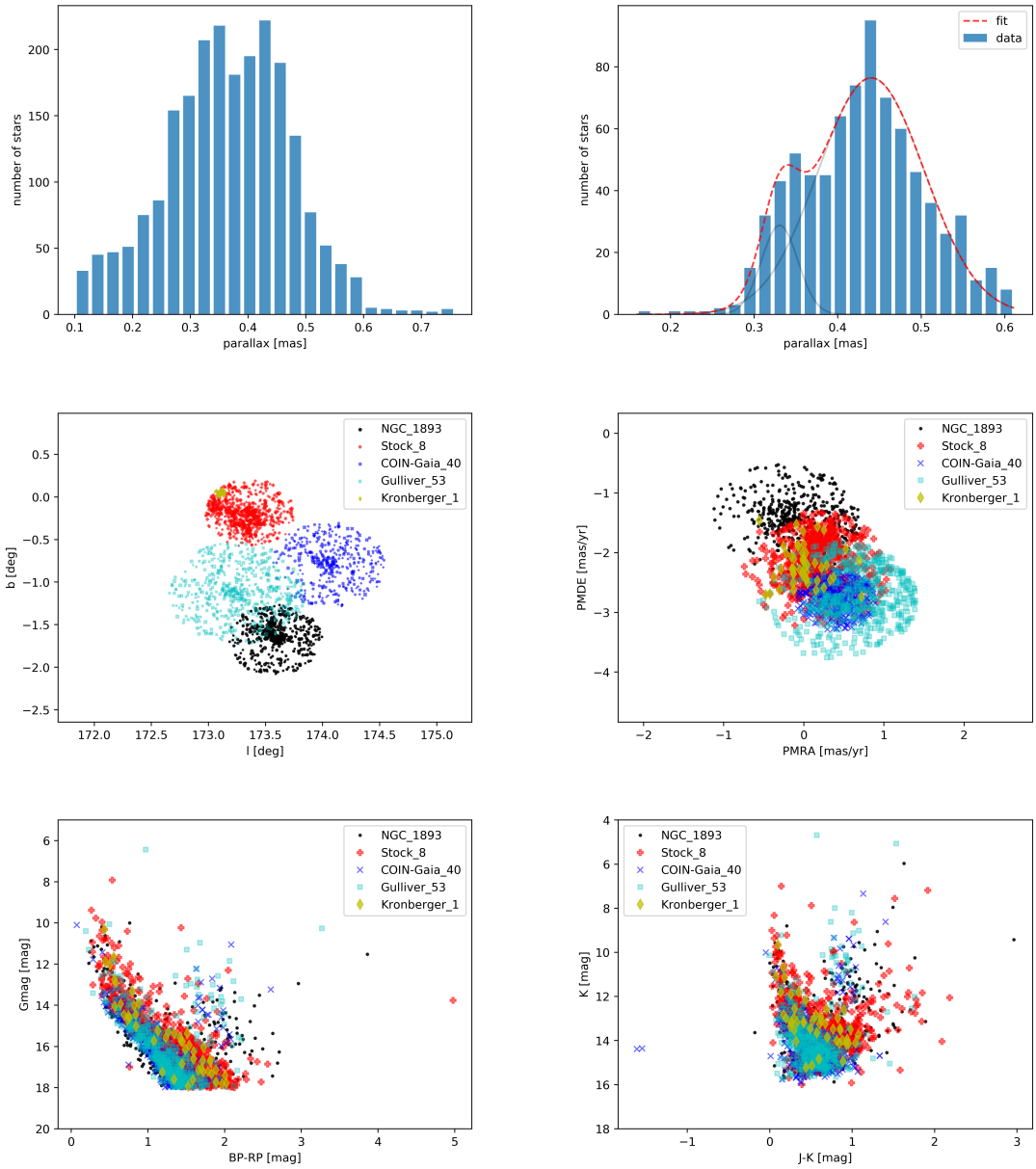


Fig. A.5. Different slices through the phase space of Agg17 (COIN-*Gaia* 40, Gulliver 53, Kronberger 1, NGC 1893, Stock 8), together with the complete colour-magnitude diagrams (based on *Gaia* and 2MASS photometry). *Top-left*: distribution of parallaxes of all individual stars in Agg17. Duplicates are excluded. *Top-right*: histogram of parallaxes, excluding the duplicate cluster members. The best fit was achieved with the following double-Gaussian function parameters: $\varpi_1 = 0.331 \pm 0.006$ mas, $\sigma_1 = 0.022 \pm 0.007$ mas, $\varpi_2 = 0.440 \pm 0.005$ mas, $\sigma_2 = 0.064 \pm 0.005$ mas. *Middle-left*: coordinates of the stars in Agg17. Size of the points indicates values of the observed magnitude G. *Middle-right*: proper motion diagram of Agg17. *Bottom-left*: colour-magnitude diagram of Agg17, based on the observed values of $B_P - R_P$ and G). *Bottom-right*: colour-magnitude diagram of Agg17 based on 2MASS photometry. The stars were located using the coordinates in *Gaia* database.

Chapter 7

Mapping Local Interstellar Medium With Diffuse Interstellar Bands

***Piecka, M.**, Paunzen, E., 2020, Monthly Notices of the Royal Astronomical Society, Volume 495, Issue 2, 18 pp.*

As a follow-up to the study of open clusters presented in my bachelor thesis, I have prepared an extended ($10^2 \rightarrow 10^3$ lines of sight) and more deep analysis of the maps of DIBs.

The maps show the structure in the x-axis view that was already shown in the bachelor thesis. It is argued that this structure is unlikely due to the errors in the EW measurements. Instead, the DIBs seem to trace something in the ISM. Another explanation could be an unknown observational bias.

Next, the diagram showing EW vs l tells us that the strength of the DIBs varies when looking in different directions (even when a constraint is put on the distance of the observed stars). A global minimum can be seen in the range of 200 deg $< l <$ 300 deg, although not many objects were observed at large enough distances ($d >$ 500 pc) in this region in the sky. Still, global changes in the strength of the DIBs in our Galaxy can be seen in the lines of sight toward the Magellanic Clouds or even in the Baade's Window. Furthermore, the data used in this study cover a good portion of the sky (except for the region near $l = 50$ deg, where only $\sim$ 20 measurements are provided).

Even more importantly, the correlation between the EWs and the interstellar reddening was studied in multiple directions, which were chosen based on the EW vs l diagram. Based on this analysis, we can notice that the strongest correlation (from the global point of view) can be found in the region 40 deg $< l <$ 125 deg. On the other hand, no real correlation can be found in 150 deg $< l <$ 200 deg. The most likely explanation is the presence of multiple nearby molecular clouds in this region (the complex of clouds in/near Taurus).

MP prepared and analysed the different sources of the EWs of DIBs, created the maps of DIBs, analysed the figures and separated the data into several regions, was responsible for writing the larger part of the paper, and submitted the manuscript.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **495,** 2035–2052 (2020) doi:10.1093/mnras/staa1112
Advance Access publication 2020 April 24

Mapping local interstellar medium with diffuse interstellar bands

Martin Piecka★ and Ernst Paunzen

Department of Theoretical Physics and Astrophysics, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic

Accepted 2020 April 8. Received 2020 April 8; in original form 2019 August 26

ABSTRACT

With the use of the data from archives, we studied the correlations between the equivalent widths of four diffuse interstellar bands (4430, 5780, 5797, 6284 Å) and properties of the target stars (colour excess values, distances, and Galactic coordinates). Many different plots of the diffuse interstellar bands and their maps were produced and further analysed. There appears to be a structure in the plot of equivalent widths of 5780 Å DIB (and 6284 Å DIB) against the Galactic x-coordinate. The structure is well defined below $\sim$ 150 mÅ and within $|x| < 250$ pc, peaking around $x = 170$ pc. We argue that the origin of this structure is not a statistical fluctuation. Splitting the data in the Galactic longitude into several subregions improve or lower the well-known linear relation between the equivalent widths and the colour excess, which was expected. However, some of the lines of sight display drastically different behaviour. The region within $150^\circ < l < 200^\circ$ shows scatter in the correlation plots with the colour excess for all of the four bands with correlation coefficients $R < 0.58$. We suspect that the variation of physical conditions in the nearby molecular clouds could be responsible. Finally, the area $250^\circ < l < 300^\circ$ displays (from the statistical point of view) significantly lower values of equivalent widths than the other regions – this tells us that there is either a significant underabundance of carriers (when compared with the other regions) or that this has to be a result of an observational bias.

Key words: dust, extinction – ISM: lines and bands – ISM: molecules – ISM: structure.

1 INTRODUCTION

Diffuse interstellar bands (DIBs) are absorption spectral features, usually observed in the lines of sight towards hot stars (but also seen in spectra of other astronomical objects). To this date, several hundreds of different DIBs have been observed (Hobbs et al. 2008), most of which are present in the visible part of the spectrum. Their discovery was made by Heger (1922) while studying the sodium D lines, although the proper study of these features began in 1930s. The interstellar nature of the DIBs was discovered by looking at the spectra of binary stars (Merrill 1936) where the position of the DIBs does not significantly change at the relevant time-scales. Furthermore, it was also found that the DIBs are correlated with the colour excess $E(B - V)$ of the observed stars (Merrill 1936; Snow, York & Welty 1977) which points to the fact that they are somehow related to the interstellar dust. The next step here would be the identification of the carrier of these bands. However, most of the ideas for carriers have been rejected. Atoms and diatomic molecules simply cannot fit the structure of the DIBs, while the studies of the polarization at those wavelengths (Cox et al. 2007) show that the carriers are most likely not part of the dust population which is responsible for most of the polarization of the starlight.

After almost 100 yr, our knowledge of these mysterious interstellar bands has slightly improved. It was discovered that there are probably several families of the carriers (Krełowski & Walker 1987) and that some sort of structure can be identified in the profile of several individual bands (Sarre et al. 1995). Since the discovery of fullerenes (Kroto et al. 1985), it was theorized that these molecules are going to be abundant in space and maybe also responsible for the formation of narrow absorption bands in the visible part of the spectrum. Although their presence in the universe was detected several years ago (Cami et al. 2010), it was only recently that the improvements in the mass-spectrometry allowed for an identification of at least two (and possibly three more) of the DIBs in the spectrum of the buckminsterfullerene C_{60}^+ (Campbell et al. 2015) which is supported by space observation (Cordiner et al. 2019).

Our work is focused on the usage of available data and spectra in the data bases that can be used for mapping the DIBs in the near vicinity of the Solar system, and possibly beyond. It is possible that a comparison of these maps with other maps of the interstellar medium (ISM) can tell us a bit more about the nature of the carriers of the DIBs. Furthermore, these maps could be used as another diagnostic tool for the ISM, once these carriers are discovered. In the following sections, we will discuss the data sets used in this work which were retrieved from archives and the results which can be obtained from the plots of the equivalent widths (EWs) of the DIBs against different Galactic coordinates.

★ E-mail: 408988@mail.muni.cz

To this date, there have been only several attempts to map the DIBs on the global scale – we will quickly mention some of the latest works. Bailey et al. (2016) created a map of the DIBs based on observations of 670 nearby hot stars. On larger scale, Zasowski et al. (2015) used the data from APOGEE and analysed over 60 000 lines of sight with the DIB at 1.527 μm which resulted into good-quality maps covering about two-thirds of the whole Galactic plane. Finally, Lan, Ménard & Zhu (2015) analysed SDSS data and studied the lines of sight towards a significant number of stars ($\sim$ 250 000), most of which are of cool spectral types. Although their map is very detailed, it is limited only to high Galactic latitudes (mostly $|b| > 30^{\circ}$). On the other hand, this part of the sky complements very well the map from Zasowski et al. (2015).

Regarding the globular clusters, tracing DIBs is mostly used for studying the ISM between us and the targets. One of the earliest works was done by van Loon et al. (2009) who looked at the lines of sight towards ω Centauri cluster and studied the structure of the diffuse ISM. Afterwards, more works follow – for example, Monreal-Ibero et al. (2015) studied the relation between the DIBs and the reddening towards the M4 cluster. Damineli et al. (2016) also studied the relation between extinction and the DIBs – they observed the DIB at 8620 Å towards Westerlund 1 and derived an empirical relation between the DIB and the extinction in the near-infrared part of the spectrum. More recently, Wendt et al. (2017) showed by observing the NGC 6397 cluster using VLT/MUSE that we now have instruments capable of studying the variations of the strength of the DIBs on quite small scales.

On the other hand, stellar associations and clusters can be used to study DIBs themselves and their carriers. Hamano et al. (2016) observed spectra of stars in the Cygnus OB2 association and studied the relation between the carriers of the near-infrared DIBs and the C_2 molecule. They found that these carriers and the C_2 molecule are not correlated that most likely means that the carriers are located, for the most part, in the diffuse part of the ISM.

2 ARCHIVAL DATA

The critical part of the problem when studying the DIBs is the lack of a significant number of publicly available observations. In the case of studying global properties of the DIBs, having a homogeneous sample of data (in terms of lines of sight) is quite important as well. Since most of these bands are very narrow (FWHM is typically below 1 Å, although most of the stronger DIBs are somewhat wider) and some of them have low intensity, one needs a good spectral resolution and high signal-to-noise ratio – we find the spectra with $R > 15\,000$ and S/N > 100 to be the best for studying the DIBs, although it is possible to use data with lower quality, as well. However, this will have a small effect on the error estimates of the EW measurements. Moreover, some DIBs may be difficult (or impossible) to measure in low-quality spectra.

Unfortunately, the data present in archives usually contain information about only a few of the strongest DIBs, namely 4430, 5780, 5797, and 6284 Å. For the purposes of our work, we have chosen the data sets from the VizieR data base which are listed in Table 1. The displayed uncertainties of the values from EWs were calculated as the median of all values for a given DIB. It should be noted that not all of the data sets contain the information about all four chosen bands but all of them contain the information about reddening towards given lines of sight.

From now on, the referenced data sets will be abbreviated according to the Table 1. There are several facts about these data sets which need to be pointed out. First, S77 contains no information about the uncertainties of determination of the EWs. Since this set contains almost half of the combined data, we adopted the median values of uncertainties of the EW measurements in the other data sets. Secondly, there are two sets of values of the 6284Å EWs in C13 – we note that for this work we used the EW values after correction. Furthermore, the target stars in C13 are objects in the Baade's Window and are of cool spectral types, unlike most of the other target stars. For this reason, we should be careful when comparing these data with the ones from other data sets. Finally, there is an outlier point in S77 in the 5797Å DIB data which is probably just a bad measurement. We have decided to remove this value and replace it with the one from X11 (for the same target star).

Since we also aim to explore the relation between DIBs and the rectangular Galactic coordinates, we need to know the distances/parallaxes assigned to the target stars. These can be taken from SIMBAD using an automatic procedure which uses the names of the target stars that are listed in the VizieR data base. However, not all of the stars have had their parallax measured, so the amount of data used for this part of the work is somewhat reduced. For example, data from C13 are completely absent in SIMBAD and we were unable to obtain specific values of the distances in any other way (although in this case the knowledge of the distance would give us little to no additional information). Furthermore, it should be pointed out that the parallaxes obtained from SIMBAD were measured by both, *Hipparcos* and *Gaia* (*Gaia*-DR2) (Perryman et al. 1997; Gaia Collaboration 2018) space observatories. Since presently *Hipparcos* data form only an extremely small fraction of the whole set used in this work, we do not have to worry as much about the significant error of measurements present in the data provided by *Hipparcos*. It is also worth mentioning that most of the target stars are located within 1 kpc from the Sun – the ISM mapped using the data can be, therefore, considered as 'local'.

In order to investigate the astrophysical properties of our targets in a colour–magnitude diagram (CMD), we use the homogeneous *Gaia*-DR2 photometry. We have excluded stars with no parallax and photometric measurements. Furthermore, we restricted the analysis to stars with parallax errors less than 20 per cent in order to avoid a significant bias to the statistics of the sample (Francis 2014). This left us with 531 objects that we used to investigate biases in the CMD (the rest of the analysis in the following sections was based on the complete sample).

The transformation of the reddening values was performed using the following relations:

$$A_V = 1.1\,A_G = 3.1\,E(B-V) = 2.1\,E(\mathrm{BP}-\mathrm{RP}). \quad (1)$$

The absolute magnitudes were calculated using the distances from Bailer-Jones et al. (2018). In Fig. 1, we present the CMD of our targets together with the PARSEC isochrones (Bressan et al. 2012) for a solar metallicity of $Z = 0.02$. We favour this value because it is consistent with the recent results of Helioseismology (Vagnozzi 2019). A smaller value, as also suggested by Bressan et al. (2012), shifts the main sequence slightly to the blue. The overall estimated reddening values seem very reliable because only eight objects (BD + 40 4220, HD 1810, HD 30112, HD 36960, HD 175803, HD 181963, HD 191639, and HD 202349) are located (within $3\,\sigma$) below the zero-age main sequence. The absorption A_G for HD 30112 is 5.6 mag. A wrong extinction value from the literature would explain its location in this diagram. For all other objects, the absorption is only small.

It is worth mentioning that there are some differences between inverting parallaxes to obtain distances from *Gaia*-DR2 and the distances calculated by an alternative procedure from Bailer-Jones

Table 1. Data sets and the related errors of the measurements of the EWs. Third and fourth display the referenced spectral resolution and the number of measured lines of sight, respectively. We labelled a field N/A if no value is present in the respective work.

	Reference	R	LoS	Err (4430 Å)	Err (5780 Å)	Err (5797 Å)	Err (6284 Å)
C13	Chen et al. (2013)	22500	219	N/A	N/A	N/A	8 %
G88	Guarinos (1988)	N/A	753	11 %	12 %	20 %	14 %
P13	Puspitarini, Lallement & Chen (2013)	48000	129	N/A	35 %	>50 %	50 %
R12	Raimond et al. (2012)	48000	150	N/A	30 %	N/A	15 %
S77	Snow et al. (1977)	N/A	2798	N/A	N/A	N/A	N/A
X11	Xiang, Li & Zhong (2011)	N/A	797	N/A	3 %	6 %	9 %

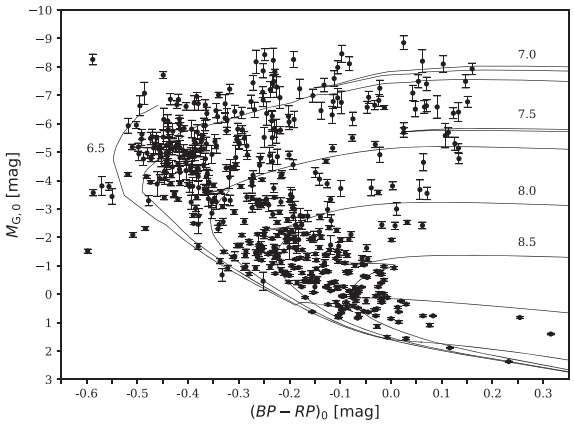


Figure 1. The $(BP - RP)_0$ versus $M_{G,0}$ diagram of our targets together with the PARSEC isochrones (listed are the logarithmic ages) for a solar metallicity of $Z = 0.02$.

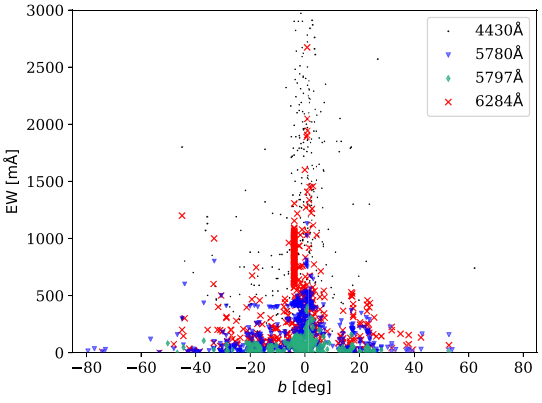


Figure 2. Distribution of the strength of the DIBs in Galactic latitude b. Black dots represent 4430 Å, blue triangles 5780 Å, teal diamonds 5797 Å and red crosses 6284 Å.

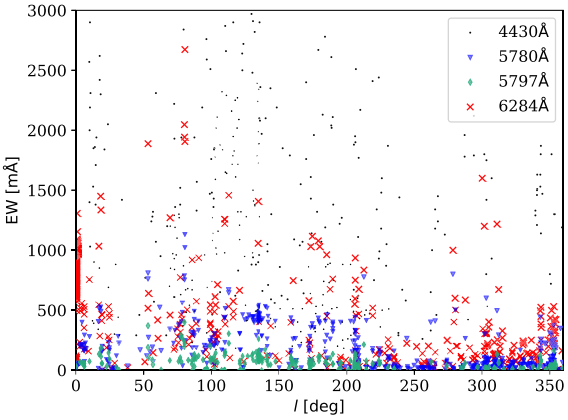


Figure 3. Distribution of the strength of the DIBs in Galactic longitude l. Black dots represent 4430 Å, blue triangles 5780 Å, teal diamonds 5797 Å, and red crosses 6284 Å.

et al. (2018) (which is also using data from *Gaia*-DR2). Although these differences are mostly minor (especially for distances below 2 kpc the correlation is almost 1.0), there are still some outliers and so we have decided to use this catalogue since it promises more accurate distances. Inversion of parallax was used only in the case when Bailer-Jones et al. (2018) do not give the distance for a star.

If we investigate the distribution of stars in Fig. 1, we find that one sample is concentrated for ages below 40 Myr and one above 100 Myr. This characterizes the stars within OB stellar associations and the Galactic field, respectively. Furthermore, there is a significant fraction of stars which have already left the zero-age main sequence. These objects may have significant stellar winds that have been connected to the formation or destruction of the carriers of DIBs (Law et al. 2017).

3 MAPS OF THE LOCAL ISM

We define the rectangular Galactic coordinates as per usual – we choose the x-axis to be the direction towards the Galactic centre, the y-axis is perpendicular to x-axis and is oriented towards the direction of Galactic rotation and the z-axis is perpendicular to the x–y plane and orientated towards the Galactic north pole. This coordinate system (x, y, z) can be easily transformed into the right-handed spherical Galactic coordinates (l, b, r). The coordinate r represents the distance between us and the observed object, Galactic longitude l represents the angle between the Galactic centre and the observed object and the Galactic latitude b represents the angle between the x–y plane and the observed object.

Using the joint data from the VizieR data base, we have created maps of the four DIBs at 4430, 5780, 5797 and 6284 Å. Combined data of all four bands are displayed in Figs 2 and 3. The rest of the plots show mostly the 5780Å band as a prototype of the results. The

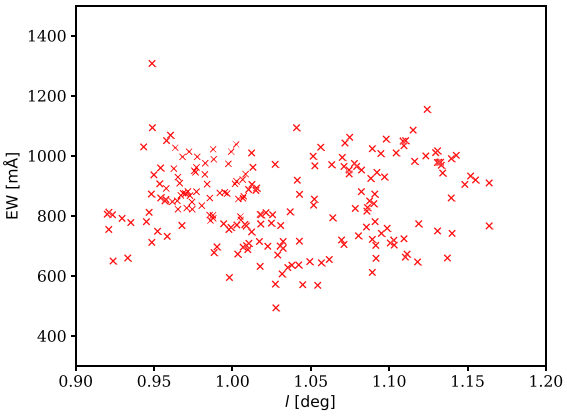


Figure 4. Strength of the 6284 Å DIB in the lines of sight towards the Baade's window.

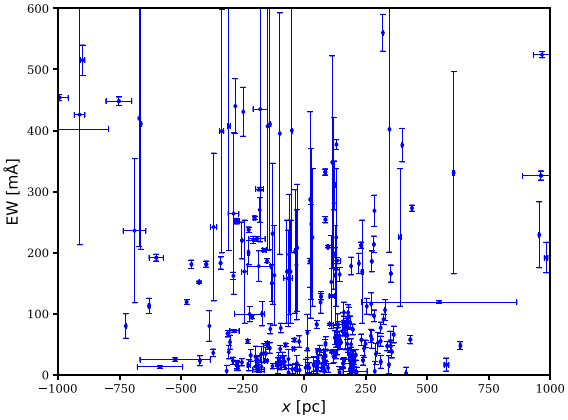


Figure 5. Relationship between the strength of the 5780 Å DIB and the Galactic x-coordinate of the target stars, with displayed error bars. The structure in the lower part of the plot seems to be well constrained within the uncertainties which means that it is most likely real and not a product of random distribution.

rest of the plots are shown in the appendix but the comments on these are presented in this section.

The EW plotted against the Galactic latitude can be seen in Fig. 2. We can clearly see that the strength of the DIBs falls down as we look further up. The same can be seen for the Galactic coordinate z that is displayed in Fig. A1, although the uncertainties in the z coordinate can be quite high and affect what we see. The values of the EWs seem to decrease up to about 500 pc which would correspond to the fact that the carriers of the DIBs are confined to the Galactic disc. However, it should be noted that this can also be the result of the bias caused by the lack of stars observed at $|z| > 500$ pc. Although it is, in principle, possible to determine the shape of the structure seen in Fig. 2, we advise against it due to the fact that there are many points without known uncertainties. It is also worth mentioning that for the 5780 Å DIB there are data points (EW $\sim$ 400 mÅ) in the latitude plot which seem to have shifted zero-point value. Further discussion of this issue is presented in the next section, after we have had a look at another important plot.

There seems to be a very complex behaviour in the plots of the Galactic longitude (Fig. 3). Generally for all the DIBs, there seems to be a maximum somewhere between $l = 0°$ and $l = 150°$, while a minimum can be located at about $l = 250°$. However, there is much more going on if we focus on just one of the plots. There appear to be at least two local maxima ($l = 0°$, $l = 80°$) and at least two local minima ($l = 50°$, $l = 250°$). On the other hand, the region $80° < l < 250°$ displays a behaviour that cannot be displayed in the simple terms of minima and maxima alone. We have chosen the 5780 Å band as a prototype since the behaviour for the 5797 and 6284 Å bands is quite similar. The data from C13 give us the opportunity to study the 6284 Å plot in more detail in a small area of the sky. As can be seen in Fig. 4, there is an apparent structure (upper outline of the distribution of points) in the Galactic longitude plot even on the scales of 0.1° – it needs to be pointed out that the typical value of uncertainties in Fig. 4 is about 8 per cent. The detailed structure of the last DIB at 4430 Å appears to significantly differ from the other ones. Unlike in the case of the other DIBs, we do not observe any striking variation in the apparent structure except of a wide minimum in the right half of the plot.

The existence of the mentioned structure is questionable. Since we expect the number of carriers to increase as we look farther away, there must be some (likely very complicated) relation between the apparent surface values of the distribution of the EWs and the distances towards the observed objects. This could be explored in detail by looking at different distances if more measurements of DIBs in different lines of sight at different distance were available in the archives.

Fig. 5 shows the plot of the EW of the DIB at 5780 Å against the Galactic coordinate x that represents the direction towards the Galactic centre (see also Fig. A2). In this direction, if we concentrate on the area below the horizontal line at EW = 150 mÅ, we will find a structure that can be defined at $|x| < 250$ pc – this is apparent also in the plot for 6284 Å band but not for the other two DIBs. This structure reaches a maximum at the distance $x = 170$ pc. We see a scatter of points at values EW > 200 mÅ. Before jumping to conclusions, we need to consider whether what we see can really be there. First, we have two sources of uncertainties which need to be taken into account. The error in the x coordinate is taken to be the actual error in the distance towards the target star – this is the upper limit. Next, since there are many measurements from S77, we need to estimate uncertainties for the EWs. For this, we take a look at the typical values mentioned in Table 1 and take 50 per cent as the uncertainty for all EW measurements from S77. Although this approach is somewhat arbitrary, it may be the only way of estimating the uncertainties for S77. The result is shown in Fig. 5 which displays only the points with errors equal or lower than 30 per cent in the x coordinate and 0.0 mag $< E(B-V) <$ 4.0 mag. If we focus on the structure we defined in this paragraph, we can see that the error bars allow its existence. There may still be some sort of bias due to the amount of observations made towards positive and negative direction of x, but we see a lot of measurements on both sides and conclude that this is unlikely, mostly due to what we see in Fig. 5. However, we cannot dismiss the argument that a bias may still actually be there – for example, the zero-point shift (discussed above) may play some role in this.

We have taken the same approach with the Galactic coordinate y, defined by the direction of Galactic rotation. Once again, we can see some sort of structure in the plot of the EW of the DIB at the 5780Å DIB (Fig. A3) in the area EW < 150 mÅ, −500 pc $< y <$ 0 pc. This time it seems obvious that the result is biased – the plot

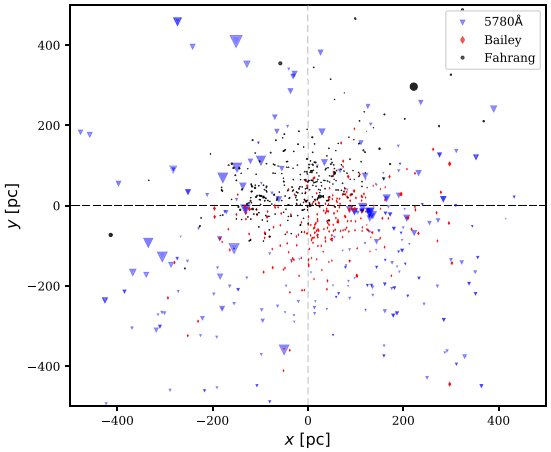


Figure 6. Distribution of the target stars of the used sample for the band at 5780 Å. Blue triangles show the data we present in this paper, black circles and red diamonds are based on the data taken from VizieR – Farhang et al. (2019) and Bailey et al. (2016), respectively. Size of a symbol refers to the strength of the DIB in the spectrum of a given star – symbols are larger as the DIB gets stronger.

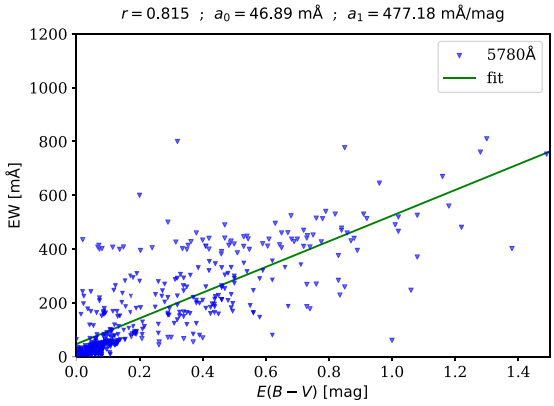


Figure 7. The correlation between the EW of the 5780 Å and the colour excess $E(B - V)$.

shows only a few measurements in the positive direction of y. In order to check this, we have created simple plots of the Galactic disc plane with the size of the points indicating the value of the EW – this is displayed in Fig. A4. Looking closely at this map, we see that there is a relative lack of lines of sight containing higher amounts of the species producing the 5780 Å feature in the plane ahead of us compared with the plane behind us. This map can be directly compared with the one from Bailey et al. (2016) where we see that both maps are quite complementary. Moreover, we can see in Fig. 6 the comparison with the data from Bailey et al. (2016) and the recent work by Farhang et al. (2019) who studied mostly objects within 200 pc. Data from this work complete our picture of the map within 100 pc radius area where data from Table 1 is lacking in the number of observations.

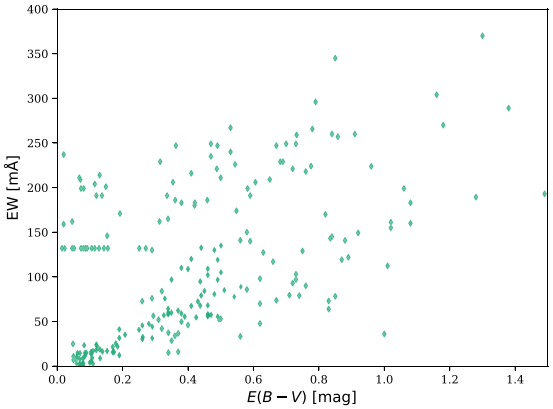


Figure 8. The correlation between the EW of the 5797 Å DIB and the colour excess $E(B - V)$. Some points display an obvious shift of their zero-point value (∼130 mÅ), similar to the 5780 Å DIB (∼400 mÅ). These data points correspond to the same objects for both DIBs, with EWs presented in S77.

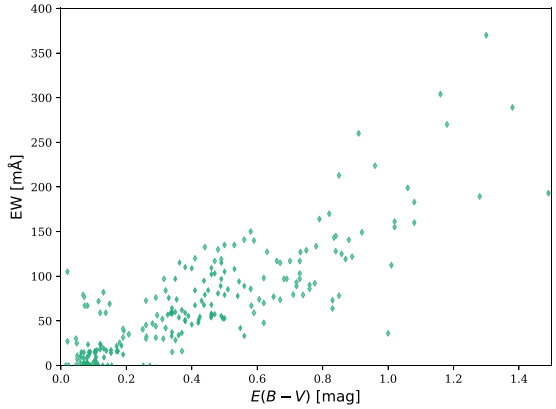


Figure 9. The correlation between the EW of the 5797 Å DIB and the colour excess $E(B - V)$. Some of the points where shifted downwards by 132 mÅ in order to correct for the zero-point shift present in the data.

4 TRENDS IN THE REDDENING CORRELATIONS

Let us now take a look at the correlations between the EWs and the colour excess $E(B - V)$. As can be seen in Fig. 7, there appears to be an almost linear relation which was already established by Merrill (1936). However, this trend is quite broad that means that the relationship between the carriers and the reddening is a bit more complex. With the use of the results shown in the previous section, we attempted to search for more narrow trends in these plots. In this section, we will only look at the plots for the 5780 Å DIB – the rest of them are presented in the Appendix.

Before we start analysing the data, we need to take a closer look at the 5797 Å DIB in Fig. 8. It can be clearly seen that there are two different trends in the data. If we pay attention to the details of the plot we can see that there is a line of points at EW ∼ 140 mÅ. This seems like a zero-point value part of the upper trend in the plot – it appears that some of the data are shifted towards higher values, possibly by the same amount as the zero-point value part.

Table 2. The correlation coefficients of linear fits of the data in different lines of sight. Asterisk is added for regions where the total number of data points is lower than 10.

Region	Name	r (4430 Å)	r (5780 Å)	r (5797 Å)	r (6284 Å)
	Original (unseparated) data	0.675	0.815	0.850	0.704
$-60^\circ < l < 40^\circ$	Galactic-central region	0.612	0.779	0.942	0.774
$40^\circ < l < 75^\circ$	First-minimum region	0.782	0.923	0.876*	0.653
$75^\circ < l < 125^\circ$	First-peak region	0.633	0.871	0.897	0.740
$125^\circ < l < 150^\circ$	5780 Å double-trend region	0.681	0.553	0.616	0.780*
$150^\circ < l < 200^\circ$	Scatter-field region	0.577	0.338	0.299	0.553
$200^\circ < l < 250^\circ$	Unknown-trend region	0.743	0.628	0.718	0.918
$250^\circ < l < 300^\circ$	Wide-minimum region	0.746	0.745	0.477*	0.821

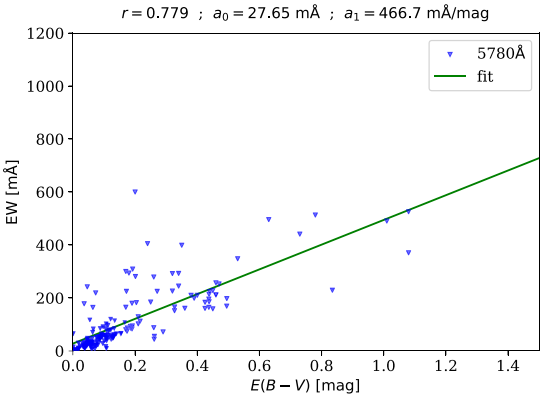


Figure 10. The correlation between the strength of the 5780 Å DIB and the colour excess $E(B-V)$ in the Galactic-central region.

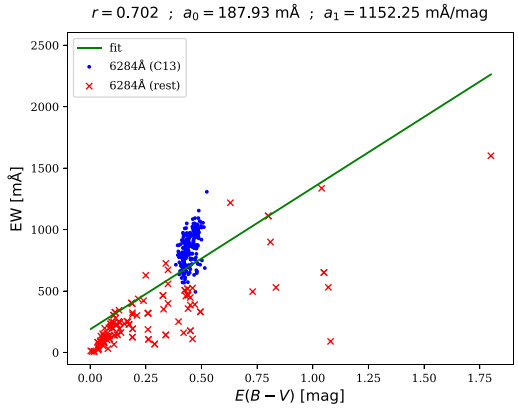


Figure 11. The correlation between the colour excess $E(B-V)$ and the strength of the 6284 Å DIB in the Galactic-central region. The second trend (blue crosses) corresponds to the stars in the Baade's Window.

For this reason, we have decided to shift the data points in the upper trend by this value downwards (Fig. 9) which improves what we see in all of the plots that include the 5797 Å DIB. We did this by looking at points above the line $\mathrm{EW}_{5797} = 154.3\,E(B-V) + 73.0$ and shifting them by the value of the upper zero-point value (132 mÅ). Unfortunately, we do not know which points are part of this shifted trend and which are not, which means that such procedure will insert a possible bias in the data set. Furthermore, the data can only be shifted by separating the trends by a line that is chosen almost arbitrarily – there is no reason to choose the parameters of the line that we used over slightly different values. We would also like to point at the 5780 Å DIB where we also see a similar shift (Fig. 7). However, in this case we cannot make any corrections due to the fact that we cannot precisely define the zero-point value. Moreover, the spread of points in Fig. 7 is somewhat larger than in Fig. 8. Finally, it should be mentioned that this shift is exclusive to S77 (not all measurements display this behaviour) and cannot be seen in the other data sets.

We have separated the data into seven different sets, described in Table 2, based on the details that can be seen in the plots of the Galactic longitude. In the table, the values highlighted in green show improvement (in terms of correlation coefficients) from the original correlation, the red colour indicates lower value for the correlation coefficient and asterisk denotes regions where the total number of data points is less than 10. The names of the regions defined by these different values of longitude are related to what we see in the plots. For example, the first minimum region is related to the first obvious minimum in the longitudinal plot. Also, for each of the plots we fitted a linear function and determined the correlation coefficient for the given region. In principle, we should expect each region to have a higher value of this coefficient than we get if we do not distinguish between the longitudinal regions. However, it must be emphasized that there is an error present in the measurement of the EWs and that we do not have the amount of measurements required for this 'detrending' method to work perfectly. On the other hand, we can expect that some of the regions should display at least slight improvement, even with the data we use.

By looking at the results of the separation of regions by longitude and comparing with the original correlation coefficients, we can safely say that there really are many different trends in the original (unseparated) reddening plots. The increase in the value of the correlation coefficient can be best seen for the 5797 and 6284 Å DIBs (in some lines of sight). On the other hand, there is only a small improvement for the 4430 Å DIB in terms of the coefficients. Finally, the 5780 Å DIB seems to be well correlated with the reddening only in several lines of sight and possibly displays a very complex behaviour in the direction towards the region $125^\circ < l < 150^\circ$.

Let us discuss these results in detail. To start with, the Galactic-central region (Fig. 10) seems to be the dominant part of the original data since the correlation line does not significantly differ from the original one for all of the DIBs (see Figs A5 and A6). This is an important result because if there are other trends in the unseparated data, observing mostly in these directions will produce a bias. Moreover, we can also take a closer look at the 6284 Å DIB (Fig. 11) and see the behaviour of the data from C13 – these clearly follow a

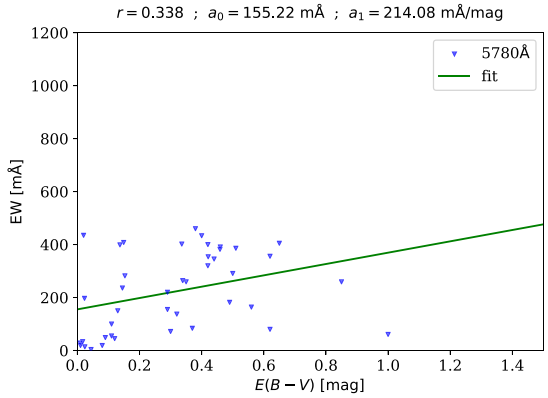


Figure 12. The correlation between the strength of the 5780Å DIB and the colour excess $E(B-V)$ in the scatter-field region.

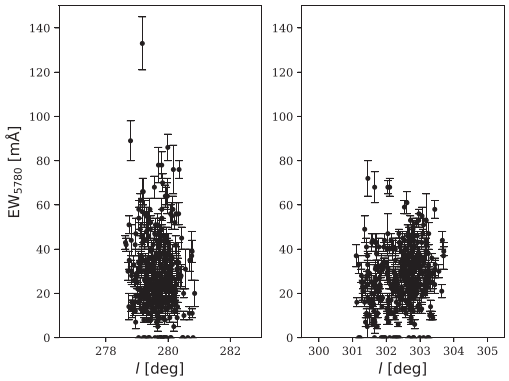


Figure 13. EWs of the foreground 5780 Å DIB in the lines of sight towards the LMC (left) and the SMC (right), taken from Bailey et al. (2015). The low value of the strength of the DIB should be noted. There is an interesting behaviour of the dependence in the direction towards the SMC, with an apparent rising towards the higher values of l.

very different trend than the rest of the data in this plot (if we had separated them) but also perfectly fit to the line in the plot which means that the correlation would not change much if we had not used the data from C13.

The first-minimum region improves the correlation for all of the DIBs but the one at 6284 Å (Fig. A7). However, not much else can be said about this region since there are only a few measurements in this region for all of the DIBs we work with.

The first-peak region contains a slightly larger amount of data. We can see some improvements from the unseparated data in terms of the coefficients. However, the shape of the correlation must be considered, as well. Although the correlation value is smaller for the 4430 Å DIB, the correlation can be clearly seen in the plot and it appears to be quite consistent, although broad (Fig. A8, upper left panel). On the other hand, the 6284 Å DIB gives a better value of the correlation coefficient but the behaviour of the points in the plot suggests that we should be careful about this result (Fig. A8, bottom right panel).

There seem to be broad correlations for the 4430 and 5797Å DIBs in the 5780 Å double-trend region. As the name of the region suggests, the 5780 Å DIB shows two different trends (Fig. A9, upper right panel). Although uncertainties may play a certain role in this result, it is very unlikely that a random error would produce such clear differences between two parts of the plot. Moreover, we note that the width of this region is relatively small (compared to the other regions). Comparing with the latitude plot, it seems that the most likely explanation is that this trend is the result of the zero-point shift. It is possible, that this could be corrected in a similar way as we corrected the data for 5797 Å band, but as was mentioned before, in this case the value of the shifted zero-point is not easy to find.

The scatter-field region (Fig. 12) shows essentially the same picture for each of the DIBs. The correlation is either missing or the separation by longitudes needs to be finer in this region (Fig. A10). This behaviour is quite unique in the picture of the DIBs.

In the region of the unknown trend, we see a significant improvement of the correlation coefficients for the 4430 and 6284 Å DIBs (Fig. A11). For the other two DIBs, the correlation is much weaker, although still apparent in the plots. The name of the region was chosen based on the high value of the correlation coefficient of the 6284 Å DIB (although slight improvement can be seen also in the case of the 4430 Å DIB), which seems striking especially when we compare these results with those of the next region.

Finally, the wide-minimum region shows that the correlation for 4430 and 6284 Å DIBs improves. For the 5797 Å DIB, there are simply not enough points to say anything about its behaviour. On the other hand, we have enough data for the 5780 Å DIB which shows a weaker correlation with the reddening (Fig. A12, upper right panel). However, when we look at the plot we see that the values of EWs are very low – it is a local minimum. It is possible that the correlation for the 5780 Å DIB cannot be found at such low values. Although we were able to find a trend in the first-minimum region, we note that the DIBs are much stronger there and the number of measurements is lower. To analyse this further, we checked with the data from Bailey et al. (2015), who provided measurements of the 5780 and 5797 Å bands towards the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC). They were able to resolve DIBs originating from both, the foreground ISM of our Galaxy and the Magellanic Clouds themselves. The data fit well together with the longitudinal plots of the 5780 Å band in the wide-minimum region. Moreover, when we look more closely (Fig. 13), we can clearly see some sort of trend going on at larger scale, which is most likely connected to the ISM of our Galaxy. However, we need to point out the important fact that Magellanic Clouds are located at higher Galactic latitudes than most of the other target stars used in this work. Finally, it is possible to compare our results from the wide-minimum region with the data from Bailey et al. (2016) and Farhang et al. (2019) which show basically the same result.

5 COMPARISON WITH EXTINCTION MAPS

It might be helpful to compare the result we derived with the available extinction maps. We have used the data from VizieR – one of the best maps, that are available there, are from Gontcharov & Mosenkov (2018). Since these maps are already restricted in the Galactic height (-600 pc $< z <$ 600 pc), we only require the A_V or $E(B-V)$, X and Y data. From these, we can derive the Galactic longitude plot of the extinction – see Fig. 14. With careful inspection of the comparison with the longitudinal plots of the four bands, we see that the DIBs seem to follow the pattern of the interstellar extinction, at least on the global scale. However, there are a few parts of the extinction map at which we should take a closer look.

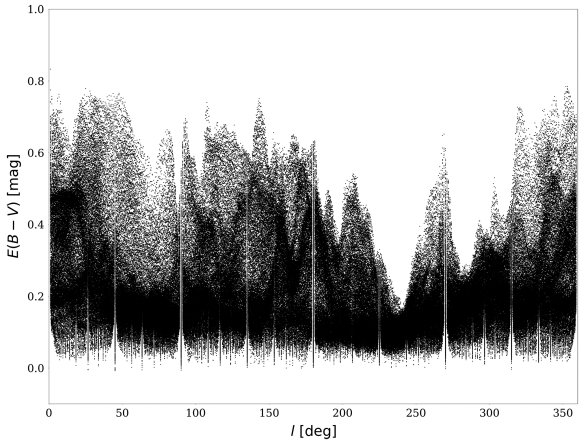


Figure 14. Colour excess $E(B-V)$ in different lines of sight, taken from Gontcharov & Mosenkov (2018). There appears to be slight resemblance of the upper outline of the distribution of points with the one that we see in Fig. 3.

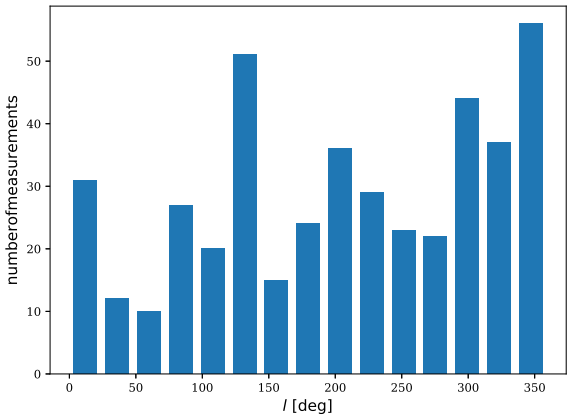


Figure 15. Histogram that shows us the number of non-zero measurements of the EWs for the 5780 Å DIB.

First of all, the region between $l = 40^\circ$ and $l = 75^\circ$ does not quite match with the DIBs. This may be due to the fact that we do not have many observations in these lines of sight – this can be seen in the histogram displayed in Fig. 15. It must be noted that the behaviour in the DIB maps and in the extinction corresponds very well in the Galactic-central, first-peak and 5780 Å double-trend regions.

The wide-minimum region seems to match the behaviour in the extinction map quite well. On the other hand, when we compare the relative strength of DIBs in the wide-minimum and the first-peak regions with the ratio in extinction, the strength of the DIBs appears to fall down a bit more quickly in this region than the extinction map would suggest, although the EW uncertainties are too high and number of measurements too low for us to be sure.

Finally, we would like to discuss the scatter-field region. We have seen that for the DIBs it is probably impossible to find a correlation with the colour excess here. If we take a closer look at this region in the extinction map (Fig. 16), we can see that there is a very complicated and strong structure on top of a bottom structure which is the continuation of the behaviour at smaller and larger longitudes. It is very likely that this has some relation to

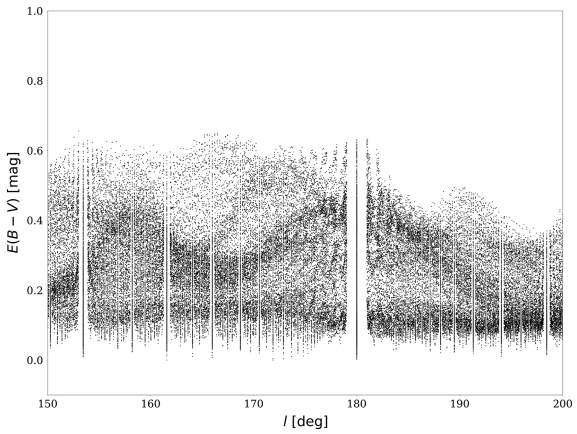


Figure 16. Appearance of different distributions of reddening in the lines of sight that coincide with the scatter-field region. Presence of such structures is explained by the existence of dense ISM that is related to the nearby molecular clouds between us and the target stars.

the number of nearby molecular clouds located in this region, like the prominent California (Lada, Lombardi & Alves 2009), Orion (Jenniskens, Ehrenfreund & Foing 1994), Perseus (Ortiz-León et al. 2018), and Taurus (Adamson, Whittet & Duley 1991) clouds. It was shown that DIBs are seen weaker towards dark regions (e.g. Snow & Cohen 1974) and stronger towards regions with strong UV radiation fields (e.g. Orion nebula, Jenniskens et al. 1994). The only DIB that appears to be an exception (in our set of bands) is the 6284 Å DIB that displays a very similar strength towards both, the Orion nebula and Taurus dark clouds (Jenniskens et al. 1994). The results of our analysis show that all of the four DIBs seem to be less correlated with the colour excess in the 'scatter-field' region when compared with the original (unseparated) data set. We would like to point out that this does not contradict the results of the previous works. Let us assume that the observed lines of sight are distributed uniformally across the sky and that the absolute dimension for different clouds are almost the same. It is clear that we should get much more data points related to the clouds that are closer – this is due to the fact that the apparent size of a cloud depends on the distance from this cloud. Therefore, it is less likely that more than one line of sight passes through a distant cloud than it is in the case of clouds which are closer. Let us now analyse the case of getting multiple measurements related to a single cloud. We will assume that for two distinct segments of a cloud we can find identical values of the correlation coefficient. The conditions (such as temperature, UV radiation field, density) in the Orion nebula, and other similar clouds, change as we look towards different regions of the whole complex – generally, the relation between the EWs and the $E(B-V)$ should vary (in terms of different parameters of the fitting function) as we study different parts of the cloud. The net effect for the whole cloud would then be a much lower value of the overall correlation coefficient when compared with the individual segments.

A second effect may also take place with varying contributions to the scatter in the plots. It is a result of the fact that, given a line of sight, the number of absorbing particles depends on the distance where we look. When looking further away and comparing with shorter distances, it can be expected that the column densities are going to be dominated mostly by the diffuse parts of the ISM, if such regions are present in the line of sight.

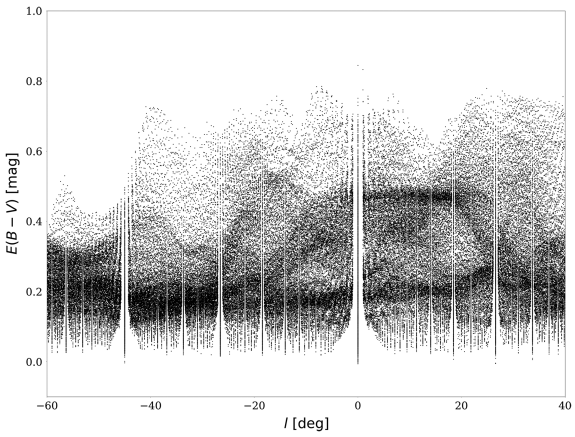


Figure 17. Appearance of different distributions of reddening in the lines of sight that coincide with the Galactic-central region. Correlations between the DIBs and the colour excess is quite good, despite the presence of regions of denser ISM in these directions (towards the Galactic centre).

There is only one other region which displays a similar structure as can be seen in Fig. 16 – the one at the centre of the Galaxy (see Fig. 17). It is puzzling why we see an improvement in the correlation coefficients of the 4430 and 5780Å DIBs when we would expect the opposite to happen (as for the other two DIBs), similarly to the scatter-field region. It is possible that this can be the result of different distances of the observed stars. However, when looking at Figs A13 and A14, we see that these two regions are quite similar from the statistical point of view.

6 DISCUSSION

Overall, we have found the following new results:

(i) There appears to be a structure in the Galactic coordinate x seen in the 5780 Å (and maybe 6284 Å) DIB within $|x| < 250$ pc. The structure appears to be non-symmetric and is not of statistical origin, based on the relatively small errors of EWs and a good sampling of data points.

(ii) The correlation coefficients between the EWs of the individual DIBs and the colour excess $E(B - V)$ should change if we split the total data into longitudinal regions. This was done with the help of the overall longitudinal plot and resulted in different changes for different DIBs (most noticeable is the comparison of the 5780 and the 6284 Å DIBs). The region defined by the interval of longitudes $150^\circ < l < 200^\circ$ displays a significantly lower value of correlation coefficients for all of the studied DIBs.

(iii) Scatter seen in some of the regions is most likely caused by the presence of nearby molecular clouds. This is a result of the fact that physical conditions change as we look at different parts of a cloud. Another effect that must be considered is coming from different contribution of the diffuse regions along the line of sight when the line of sight crosses a cloud.

(iv) For the 5780, 5797, and 6284 Å DIBs, the region located in the interval $250^\circ < l < 300^\circ$ appears to contain a smaller amount of carriers (inferred from the lower typical values of EWs) when compared with the other regions. The result for the 4430 Å DIB is most likely biased in the Galactic plane by the lack of observations in the fourth quadrant.

We have shown that with the use of publicly available data, we can construct maps that can be readily used to extract information about DIBs. We have seen that the strength of the DIBs is centred around the Galactic plane while, on the other hand, there is a much more complicated structure in the longitudinal plots. Moreover, it appears that the 4430 Å DIB displays a different behaviour in these plots than the other three studied DIBs.

With the use of the recently released data from *Gaia*-DR2, we were able to get the distances towards our target stars from Bailer-Jones et al. (2018) (if available) with, for the most part, relatively low uncertainties. This made it possible to construct plots that show different behaviour of the EWs of the DIBs in the individual rectangular Galactic coordinates. It appears that there is a small-scale structure in the near vicinity of the Sun in the interval of x-values (−250 pc, 250 pc). Due to the lack of data in the negative y direction, we cannot be certain whether there is a structure in the plots of this coordinate. Finally, we have also shown very simple maps of the different DIBs in the Galactic x–y plane. These show that we can see that the strength of a DIB may significantly vary towards different lines of sight even in the case when the observed stars have similar distances. Making maps such as those using much larger amount of data points could point towards the global distribution of the carriers around the Sun (within 1 kpc).

Using the results from the longitudinal plots and the fact that the correlation plots between the EWs and the reddening show very wide spread, we 'detrended' the plots to several subregions that all display different behaviours. Most interesting are the 'Galactic-central' and 'scatter-field' regions. We see multiple trends in these two regions even when looking at the extinction maps. We argue that this behaviour is explained by the complicated structure of the ISM towards the Galactic centre and the presence of a nearby complex of giant molecular clouds in the 'scatter-field' region. However, the results shown in the 'detrended' longitudinal plots are somewhat inconsistent – while the presence of molecular clouds can likely be the source of the scatter that we see in the related plots, there are obvious trends in the 'Galactic-central' region which seems to point to the fact the complicated nature of the ISM in these lines of sight (despite containing molecular clouds) does not affect the behaviour of the DIBs as much as the presence of a big complex of molecular clouds.

However, we can expect that there are clouds present also in other lines of sight than towards the 'scatter-field' region, and these do not seem to be a source of a significant scatter in other plots. We suspect that the distances toward the probed clouds, varying conditions through the clouds, and the distribution of lines of sight on the sky are the reason behind this different behaviour. Moreover, distance together with the column density of the carriers can have another effect – looking in the same direction (same column density) gives a different number of carriers (and therefore different EWs) at various distances. Therefore, looking in a direction of a cloud which is, for example, 300 pc away will result in carriers which are close to (or within) the cloud to be a more dominant source of absorption than it would be in the case of this cloud being more distant. Due to the lack of knowledge about the distribution and the structure of the carriers, it is impossible to estimate how important this effect can become.

To search for other clues regarding the 'Galactic-central' region, we looked in the literature for works focused on the DIBs and molecular or dark clouds. According to Adamson et al. (1991), there could be a strong connection between some of the DIBs as they vary across the cloud and their strength in general seems to be very low in the sightlines towards dark clouds. Moreover, Jenniskens

et al. (1994) were unable to find a good correlation between the EWs of DIBs and the extinction in the Orion molecular cloud. Our results appear to be consistent with their findings. Finally, studies of the Sco OB2 near the Galactic centre (Vos et al. 2011) show that the correlation of the DIBs with the reddening is, on average, very similar to the 'general' correlation in the diffuse ISM. This is slightly in conflict with our findings since we have shown that there is not going to be a 'general' correlation, but the idea of averaging of the correlations towards the regions near the Galactic centre could explain why we do not see scatter in the EW versus $E(B-V)$ plots. This can be supported by the fact that the total correlation differs only slightly from the 'detrended' correlation in the 'Galactic-central' region (see Fig. 10). However, we need to point out that this region in our study is still very broad compared to the one in the last mentioned work.

We also point our attention to the 'first-peak' region. We can clearly see a small improvement in the correlation between the EW and the reddening as a result of 'detrending'. However, this is not true for all of the DIBs. This can either be explained by the uncertainties of the determination of the EWs, or by the existence of an absolutely different behaviour of the DIBs in these lines of sight. Moreover, the strength of all of the four DIBs seems to reach maximum values in these directions. We cannot explain why this is so – for comparison, we looked at the maps of the local ISM from Lallement et al. (2014, 2018). There are both, high-absorption (i.e. high-density) regions and diffuse medium regions in the interval of longitudes (75°, 125°). This seems to point against our assumption that clouds are a source of a scatter in the plots of EWs against the colour excess but the clouds in this region are typically more distant. On the other hand, we also see in the maps from the mentioned works that there are several high-absorption regions (typically less distant than those seen in the 'first-peak' region and located within $125^\circ < l < 150^\circ$) where we find that the correlation coefficients are somewhat lower for the 5780 and the 5797Å DIBs. Again, this seems to indicate that the distances toward clouds (within the line of sight) seem to play some role in the study of the correlations.

The difference between the behaviour in the 'double-trend' region and the 'scatter-field' region can be explained by assuming the second effect of the distance mentioned above – according to the map in Lallement et al. (2018), the 'double-trend' region contains more high-absorption regions within 1 kpc than diffuse regions when compared with the 'scatter-field' region.

In general, the 'wide-minimum' region seems to contain the least amount of the high-absorption material. This may be the result of a lack of carriers in these regions of the ISM or an effect produced by an observational bias. Looking at the $x-y$ map of the 4430Å, the small number of observations in the fourth quadrant (where large part of this 'wide-minimum' region is located) will greatly affect the results. For the other three DIBs, the number of data points in the fourth quadrant are satisfactory. Although almost all observed stars are within a 500 pc radius, we can see larger values of EWs even at shorter distances. This points to the explanation that the carriers of 5780, 5797, and 6284Å DIBs are relatively underabundant in this region.

The complexity of the longitudinal plots was shown in greater detail in the data provided by Chen et al. (2013) and Bailey et al. (2015). The OGLE stars in C13 tell us that there is an interesting behaviour towards the Galactic bulge which can be studied if we have enough data in $\Delta l = 0.1^\circ$ intervals. On the other hand, some sort of behaviour on slightly larger scale appears to be also present in the lines of sight towards the SMC and LMC. With sufficient amount of high-quality data, it should be possible to connect small-scale structure across the whole sky into one distribution that could be used to analyse global properties of carriers of the DIBs.

Obviously, our 'detrending' process reduced the plots to subplots that now consist of only small numbers of points. More data and observations are required to verify our findings. In the future, we aim to create an automatic procedure that would search the spectra available in the data bases (for example, ESO data base), search for the presence of the DIBs and determine their EWs. This would give us more information to work with and we could improve the quality of the maps (especially with the help of *Gaia*-DR2). Furthermore, we could study the 'wide-minimum', 'Galactic-central', and 'scatter-field' regions in much higher detail. It is quite possible that studies of the regions (on small and global scale), where the behaviour of the DIBs significantly differs from the 'normal trends', will pinpoint the birthplace of the carriers of the DIBs.

REFERENCES

Adamson A. J., Whittet D. C. B., Duley W. W., 1991, MNRAS, 252, 234
Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, AJ, 156, 58
Bailey M., van Loon J. T., Farhang A., Javadi A., Khosroshahi H. G., Sarre P. J., Smith K. T., 2016, A&A, 585, A12
Bailey M., van Loon J. T., Sarre P. J., Beckman J. E., 2015, MNRAS, 454, 4013
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Cami J., Bernard-Salas J., Peeters E., Malek S. E., 2010, Science, 329, 1180
Campbell E. K., Holz M., Gerlich D., Maier J. P., 2015, Nature, 523, 322
Chen H.-C., Lallement R., Babusiaux C., Puspitarini L., Bonifacio P., Hill V., 2013, A&A, 550, A62
Cordiner M. A. et al., 2019, ApJ, 875, L28
Cox N. L. J. et al., 2007, A&A, 465, 899
Damineli A., Almeida L. A., Blum R. D., Damineli D. S. C., Navarete F., Rubinho M. S., Teodoro M., 2016, MNRAS, 463, 2653
Farhang A., van Loon J. T., Khosroshahi H. G., Javadi A., Bailey M., 2019, Nat. Astron., 3, 922
Francis C., 2014, MNRAS, 444, L6
Gaia Collaboration, 2018, A&A, 616, A1
Gontcharov G. A., Mosenkov A. V., 2018, VizieR Online Data Catalog, 2354
Guarinos J., 1988, Bulletin d'Information du Centre de Donnees Stellaires, 34, 141
Hamano S. et al., 2016, ApJ, 821, 42
Heger M. L., 1922, Lick Obs. Bull., 10, 146
Hobbs L. M. et al., 2008, ApJ, 680, 1256
Jenniskens P., Ehrenfreund P., Foing B., 1994, A&A, 281, 517
Krełowski J., Walker G. A. H., 1987, ApJ, 312, 860
Kroto H. W., Heath J. R., Obrien S. C., Curl R. F., Smalley R. E., 1985, Nature, 318, 162
Lada C. J., Lombardi M., Alves J. F., 2009, ApJ, 703, 52
Lallement R., Vergely J.-L., Valette B., Puspitarini L., Eyer L., Casagrande L., 2014, A&A, 561, A91
Lallement R. et al., 2018, A&A, 616, A132
Lan T.-W., Ménard B., Zhu G., 2015, MNRAS, 452, 3629
Law C. J. et al., 2017, MNRAS, 470, 2835
Merrill P. W., 1936, ApJ, 83, 126
Monreal-Ibero A., Lallement R., Puspitarini L., Bonifacio P., Monaco L., 2015, Mem. Soc. Astron. Ital., 86, 527
Ortiz-León G. N. et al., 2018, ApJ, 865, 73
Perryman M. A. C. et al., 1997, A&A, 323, L49
Puspitarini L., Lallement R., Chen H.-C., 2013, A&A, 555, A25

Raimond S., Lallement R., Vergely J. L., Babusiaux C., Eyer L., 2012, A&A, 544, A136
Sarre P. J., Miles J. R., Kerr T. H., Hibbins R. E., Fossey S. J., Somerville W. B., 1995, MNRAS, 277, L41
Snow T. P. J., Cohen J. G., 1974, ApJ, 194, 313
Snow T. P. Jr, York D. G., Welty D. E., 1977, AJ, 82, 113
Vagnozzi S., 2019, Atoms, 7, 41
van Loon J. T., Smith K. T., McDonald I., Sarre P. J., Fossey S. J., Sharp R. G., 2009, MNRAS, 399, 195
Vos D. A. I., Cox N. L. J., Kaper L., Spaans M., Ehrenfreund P., 2011, A&A, 533, A129
Wendt M. et al., 2017, A&A, 607, A133
Xiang F. Y., Li A., Zhong J. X., 2011, ApJ, 733, 91
Zasowski G. et al., 2015, ApJ, 798, 35

APPENDIX: ADDITIONAL PLOTS

Most of the plots used in the text are displayed in this section. Additionally, figures not mentioned in the text are presented as well. All of them display the relation between the EWs of DIBs and Galactic coordinates or colour excess $E(B-V)$.

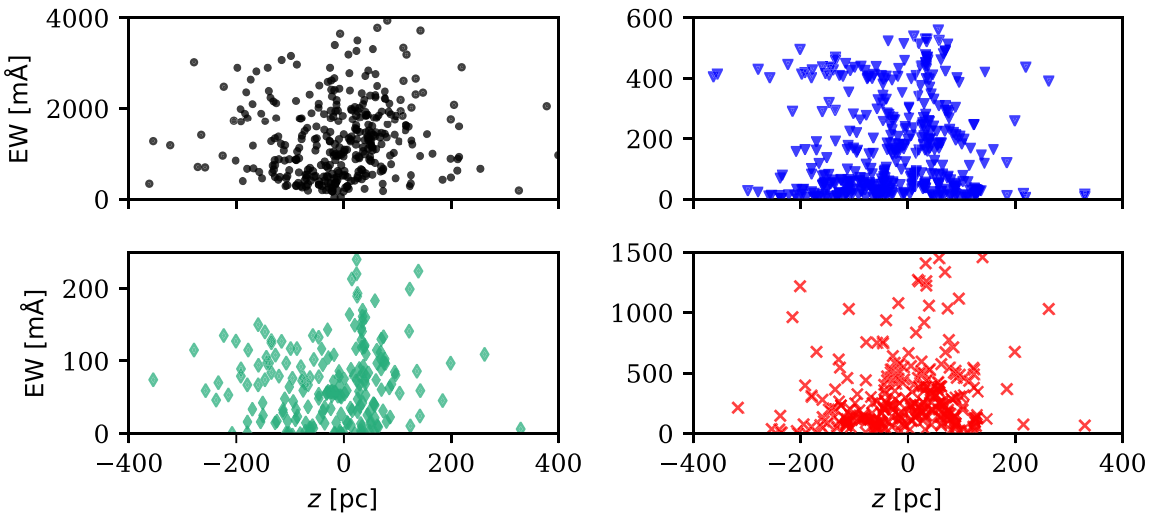


Figure A1. Correlations between the strength of the DIBs and the Galactic z-coordinate. Black dots show 4430 Å DIB, blue triangles 5780 Å DIB, teal diamonds 5797 Å DIB and red crosses 6284 Å DIB.

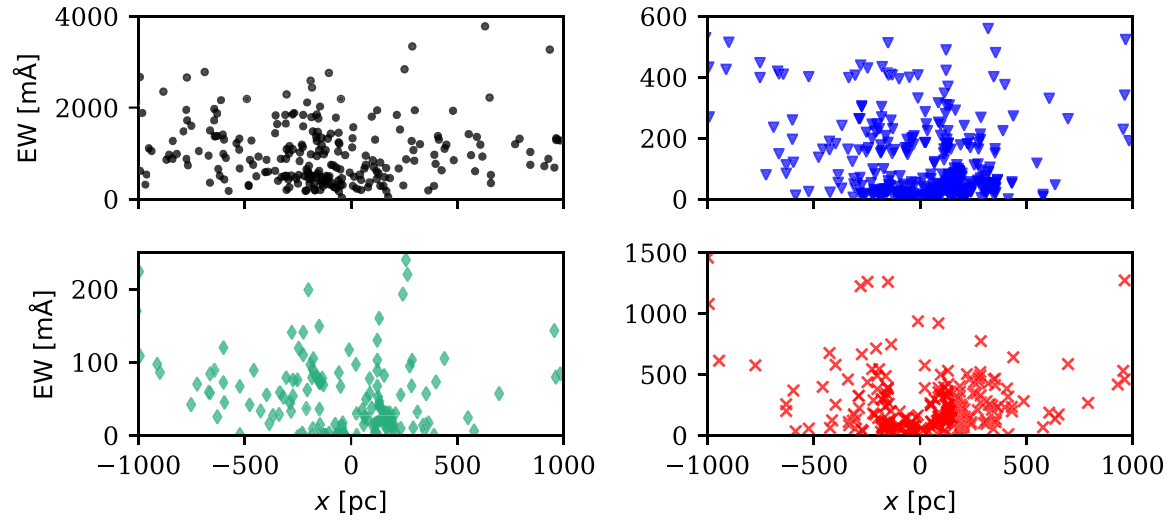


Figure A2. Correlations between the strength of the DIBs and the Galactic x-coordinate. Black dots show 4430 Å DIB, blue triangles 5780 Å DIB, teal diamonds 5797 Å DIB and red crosses 6284 Å DIB.

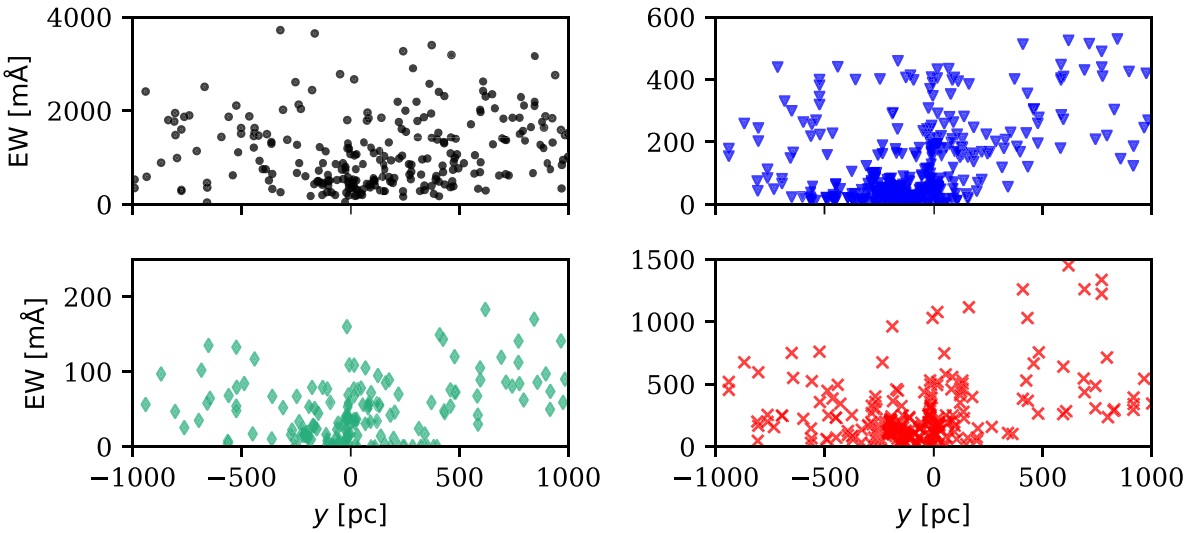


Figure A3. Correlations between the strength of the DIBs and the Galactic y-coordinate. Black dots show 4430 Å DIB, blue triangles 5780Å DIB, teal diamonds 5797Å DIB, and red crosses 6284Å DIB.

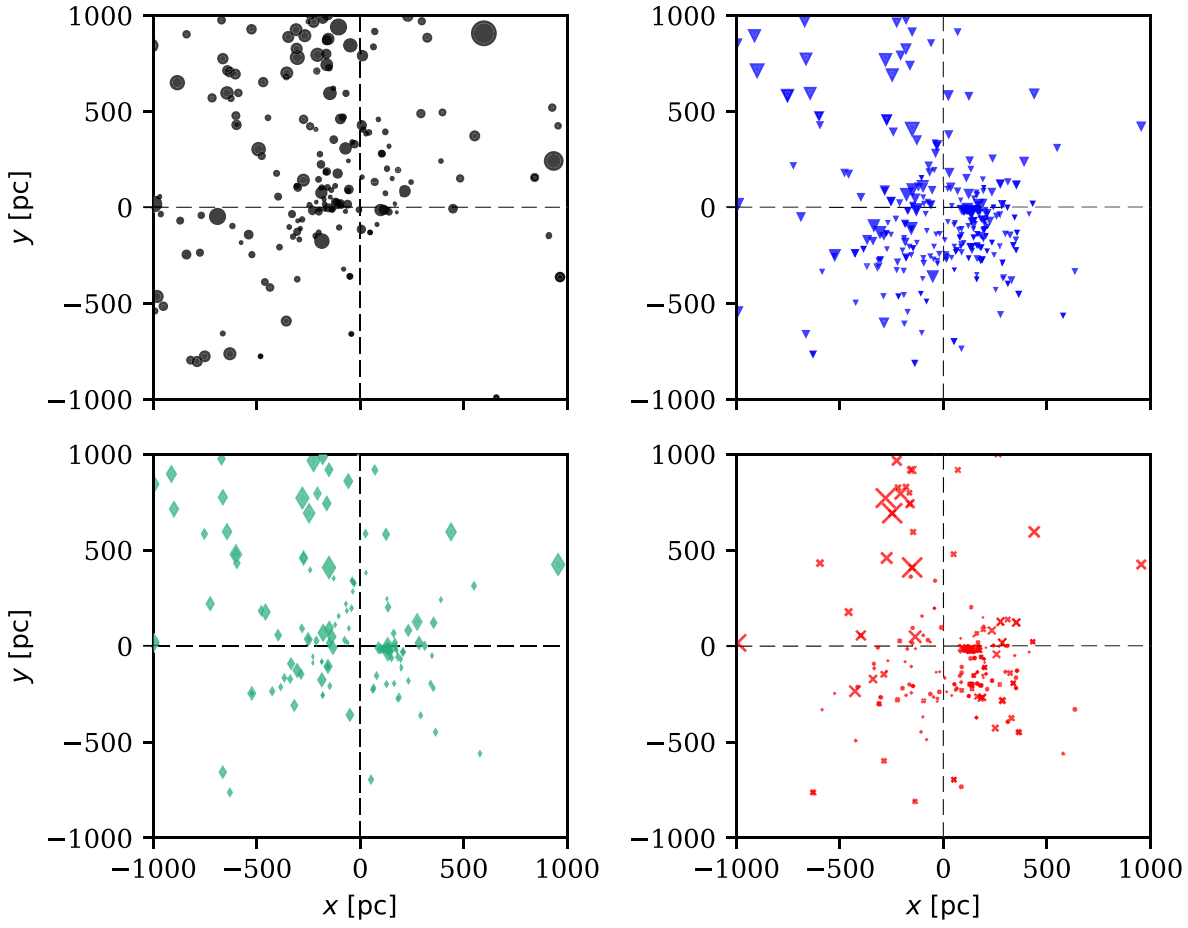


Figure A4. Positions of the target stars projected to the Galactic plane. Size of the points depends on the strength of the DIB. Black dots show 4430 Å DIB, blue triangles 5780Å DIB, teal diamonds 5797Å DIB, and red crosses 6284Å DIB.

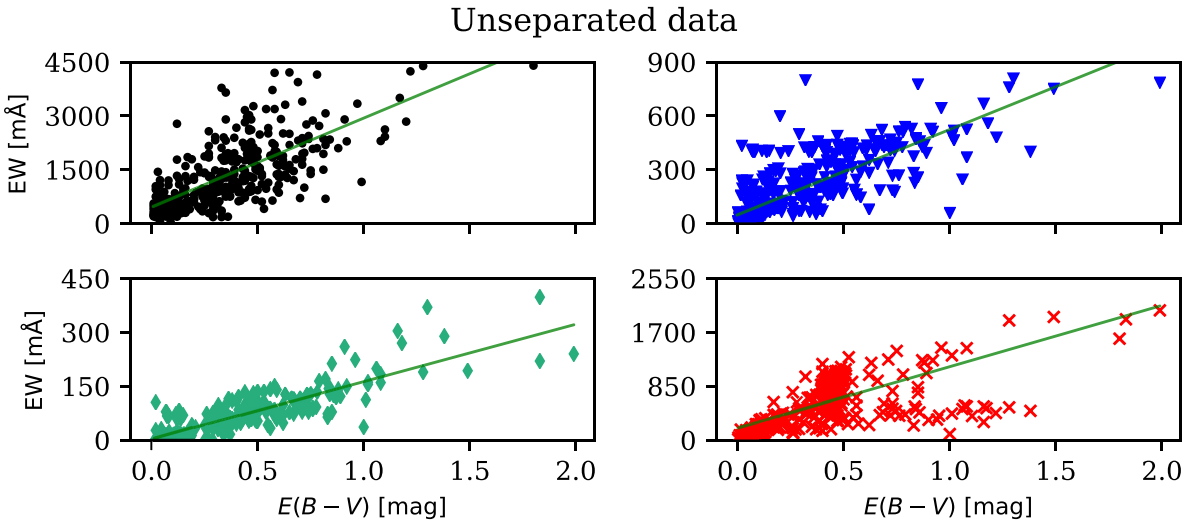


Figure A5. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the unseparated data. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430Å DIB ($r = 0.675$, $a_0 = 465$, $a_1 = 2473$), blue triangles 5780 Å DIB ($r = 0.815$, $a_0 = 47$, $a_1 = 477$), teal diamonds 5797 Å DIB ($r = 0.850$, $a_0 = 3$, $a_1 = 160$) and red crosses 6284 Å DIB ($r = 0.704$, $a_0 = 190$, $a_1 = 968$).

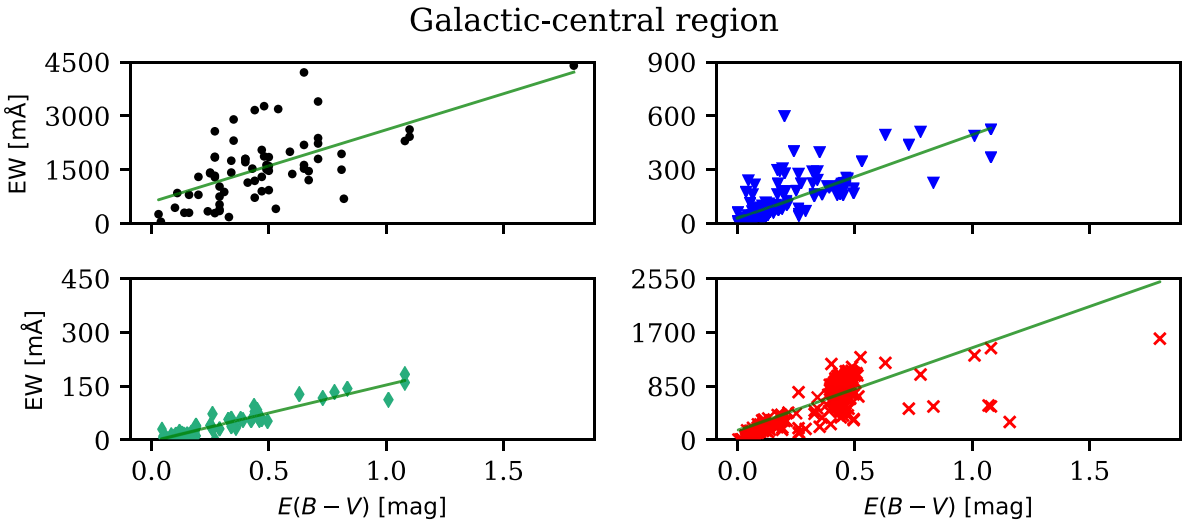


Figure A6. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the Galactic-central region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.612$, $a_0 = 601$, $a_1 = 2010$), blue triangles 5780A DIB ($r = 0.779$, $a_0 = 28$, $a_1 = 467$), teal diamonds 5797A DIB ($r = 0.942$, $a_0 = -3$, $a_1 = 156$) and red crosses 6284A DIB ($r = 0.774$, $a_0 = 154$, $a_1 = 1302$).

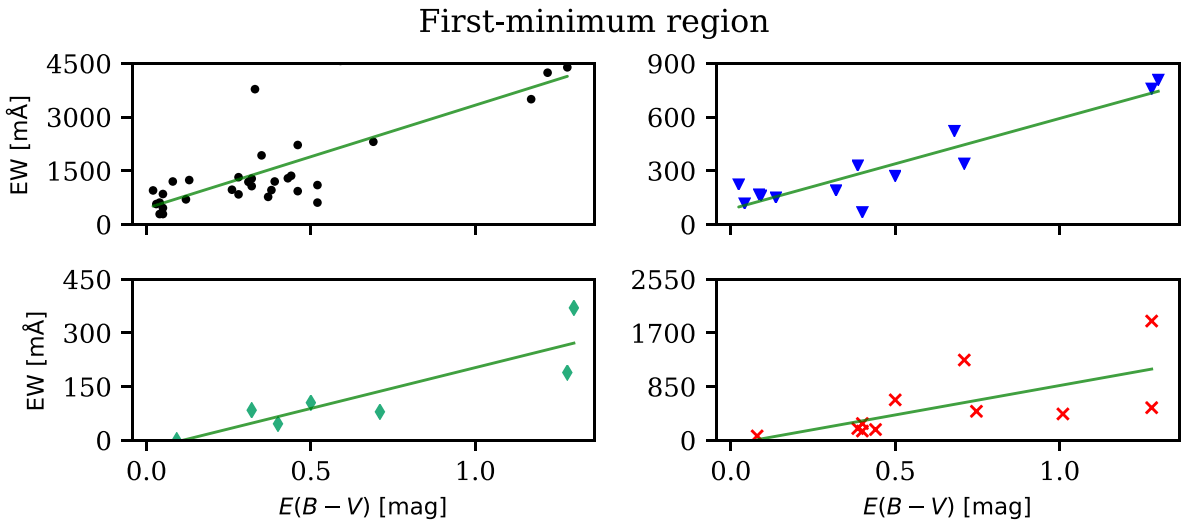


Figure A7. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the first-minimum region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.782$, $a_0 = 453$, $a_1 = 2879$), blue triangles 5780A DIB ($r = 0.923$, $a_0 = 85$, $a_1 = 507$), teal diamonds 5797A DIB ($r = 0.876$, $a_0 = -25$, $a_1 = 228$) and red crosses 6284A DIB ($r = 0.653$, $a_0 = -64$, $a_1 = 931$).

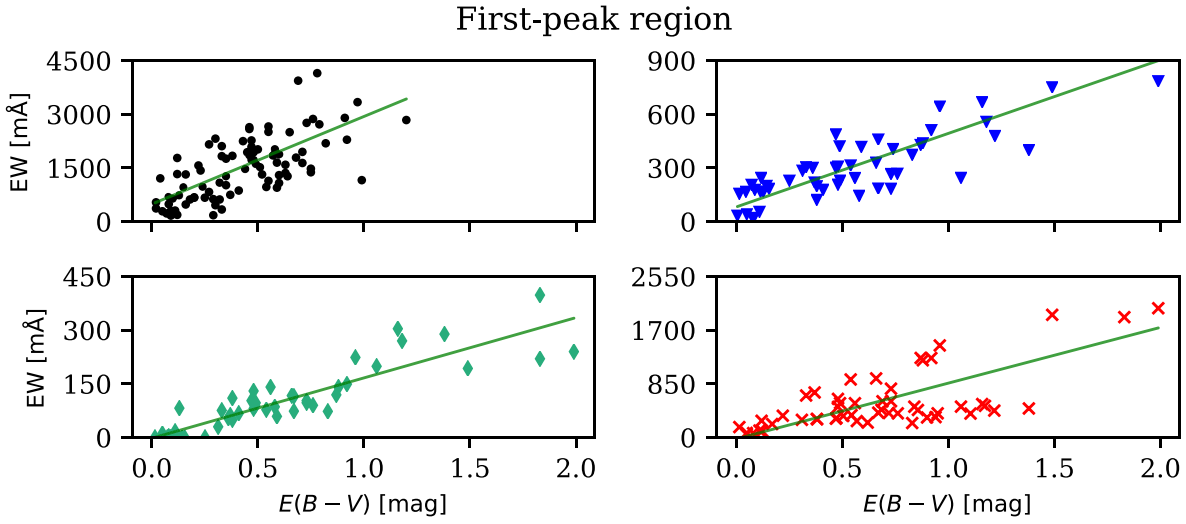


Figure A8. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the first-peak region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.633$, $a_0 = 491$, $a_1 = 2448$), blue triangles 5780A DIB ($r = 0.871$, $a_0 = 82$, $a_1 = 411$), teal diamonds 5797A DIB ($r = 0.897$, $a_0 = -2$, $a_1 = 169$), and red crosses 6284A DIB ($r = 0.740$, $a_0 = -22$, $a_1 = 883$).

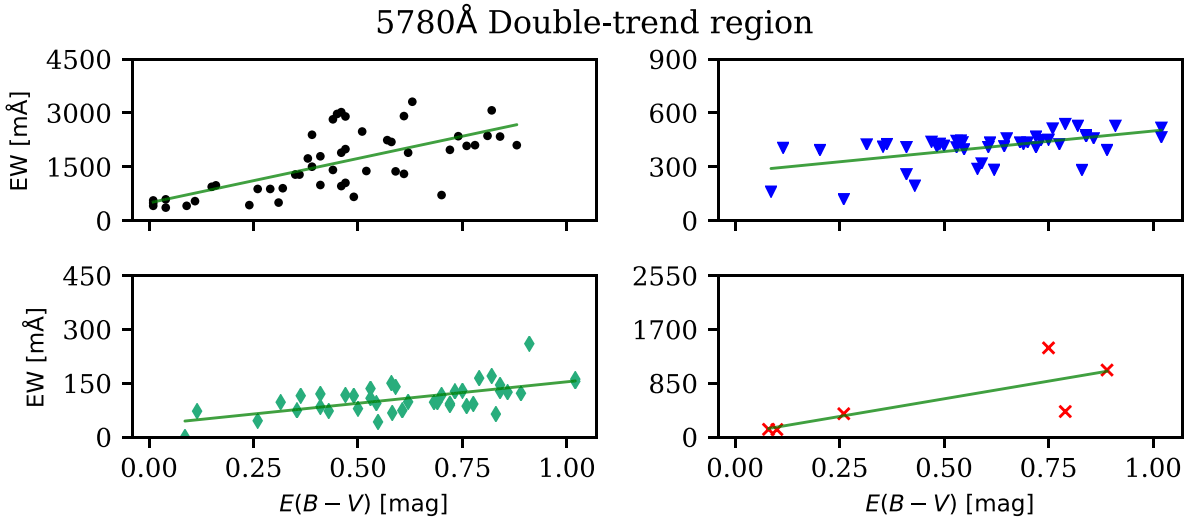


Figure A9. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the 5780A double-trend region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.681$, $a_0 = 492$, $a_1 = 2476$), blue triangles 5780 DIB ($r = 0.553$, $a_0 = 270$, $a_1 = 229$), teal diamonds 5797A DIB ($r = 0.616$, $a_0 = 34$, $a_1 = 120$) and red crosses 6284A DIB ($r = 0.780$, $a_0 = 45$, $a_1 = 1116$).

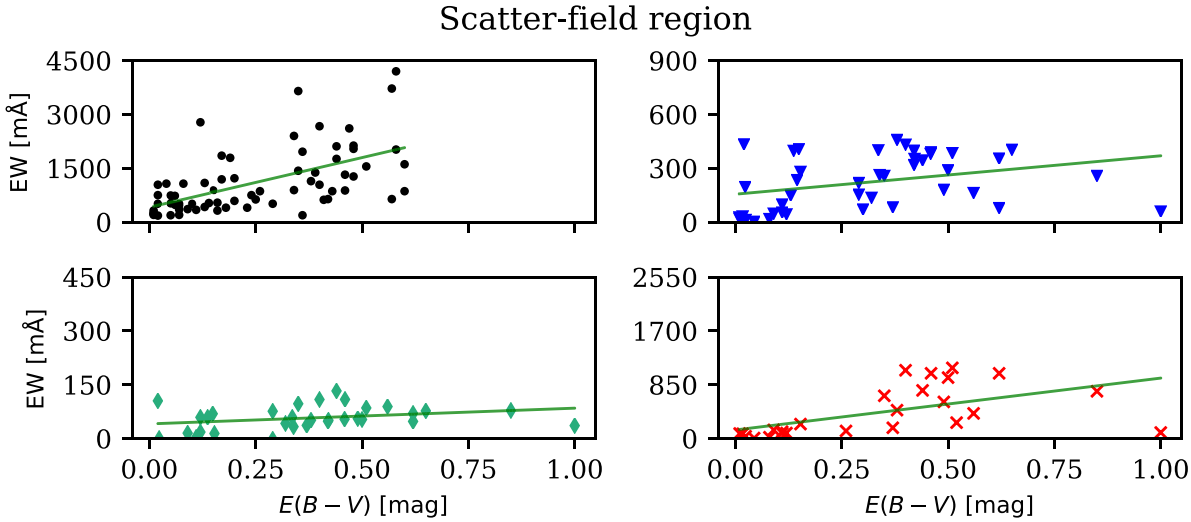


Figure A10. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the scatter-field region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.577$, $a_0 = 412$, $a_1 = 2769$), blue triangles 5780A DIB ($r = 0.338$, $a_0 = 155$, $a_1 = 214$), teal diamonds 5797A DIB ($r = 0.299$, $a_0 = 41$, $a_1 = 44$), and red crosses 6284A DIB ($r = 0.553$, $a_0 = 135$, $a_1 = 819$).

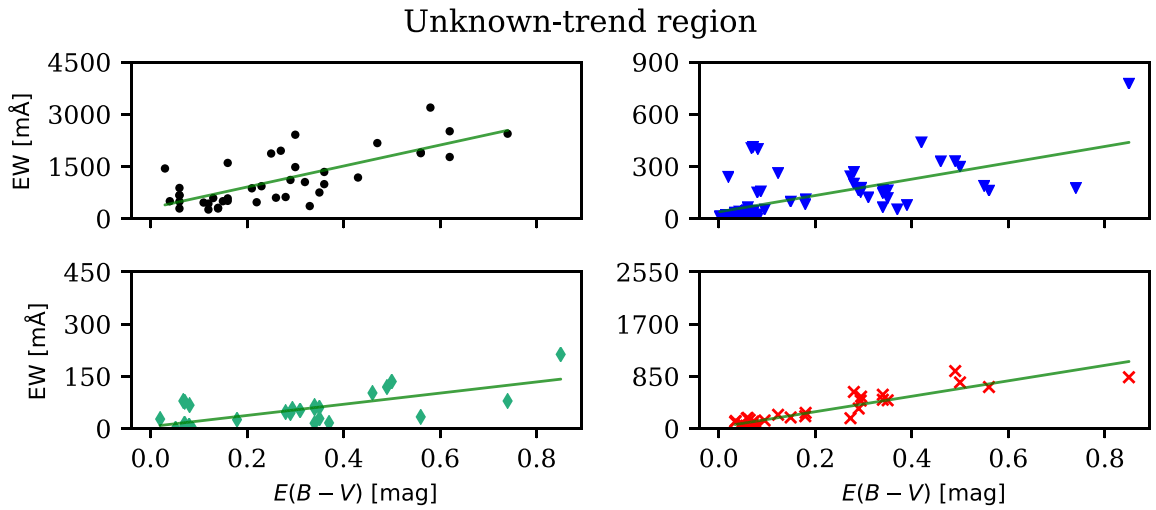


Figure A11. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the unknown-trend region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.743$, $a_0 = 303$, $a_1 = 3018$), blue triangles 5780A DIB ($r = 0.628$, $a_0 = 39$, $a_1 = 470$), teal diamonds 5797A DIB ($r = 0.718$, $a_0 = 5$, $a_1 = 161$), and red crosses 6284A DIB ($r = 0.918$, $a_0 = 21$, $a_1 = 1258$).

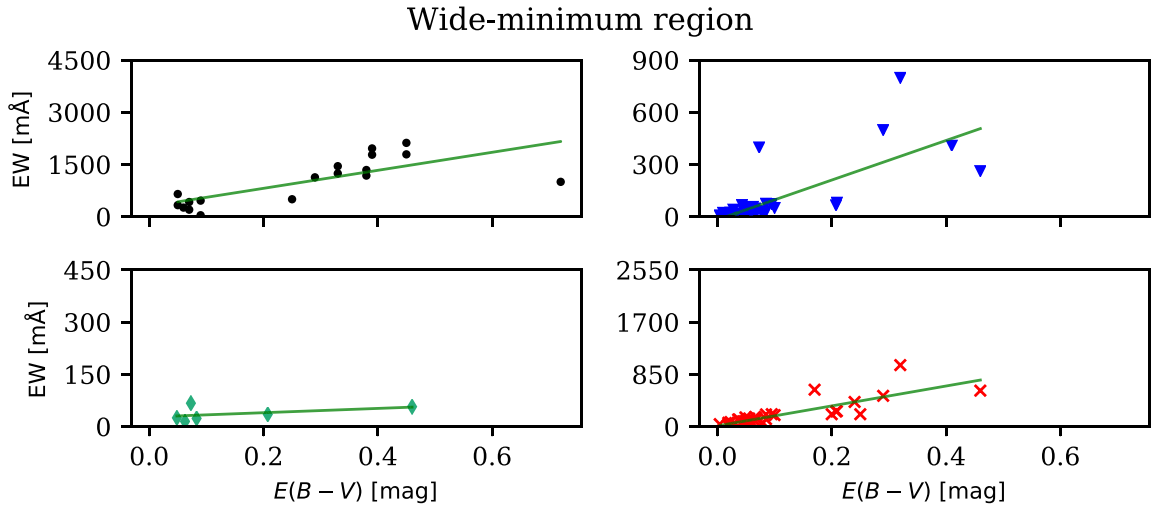


Figure A12. Correlations between the strength of the DIBs and the colour excess $E(B-V)$ in the wide-minimum region. Green line represents a linear fit to the data with correlation coefficient r and fit parameters a_0 and a_1. Black dots show 4430A DIB ($r = 0.746$, $a_0 = 296$, $a_1 = 2589$), blue triangles 5780A DIB ($r = 0.745$, $a_0 = -17$, $a_1 = 1137$), teal diamonds 5797A DIB ($r = 0.477$, $a_0 = 27$, $a_1 = 62$), and red crosses 6284A DIB ($r = 0.821$, $a_0 = 11$, $a_1 = 1621$).

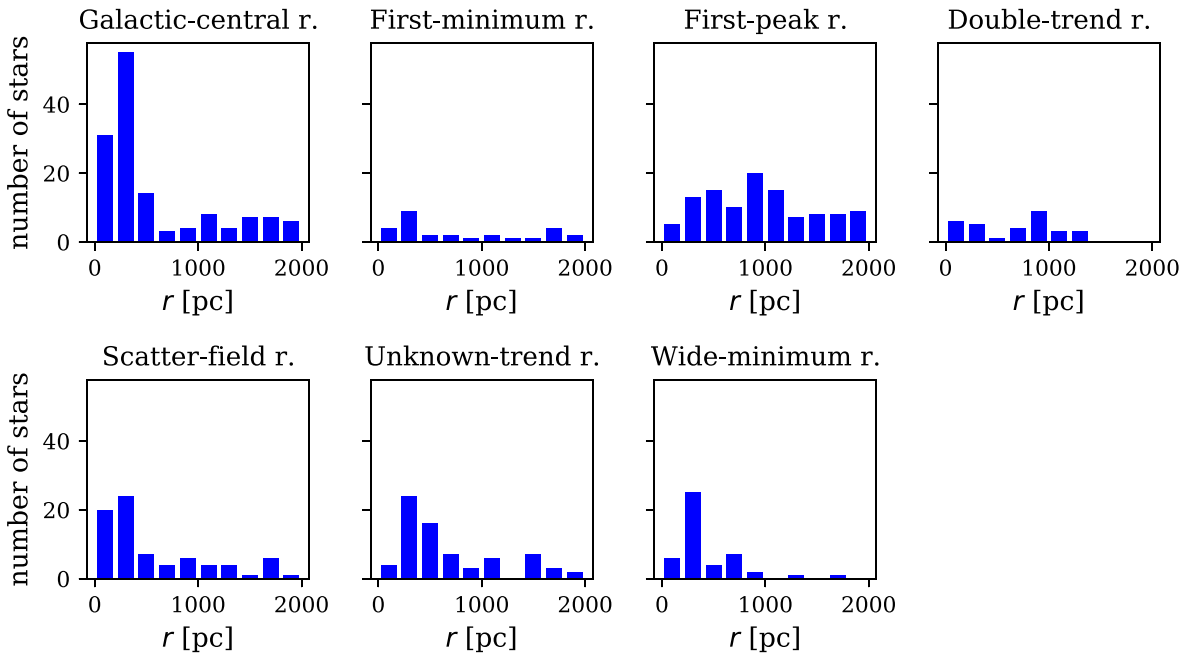


Figure A13. Distribution of distances (up to 2 kpc) in all regions defined in Table 2.

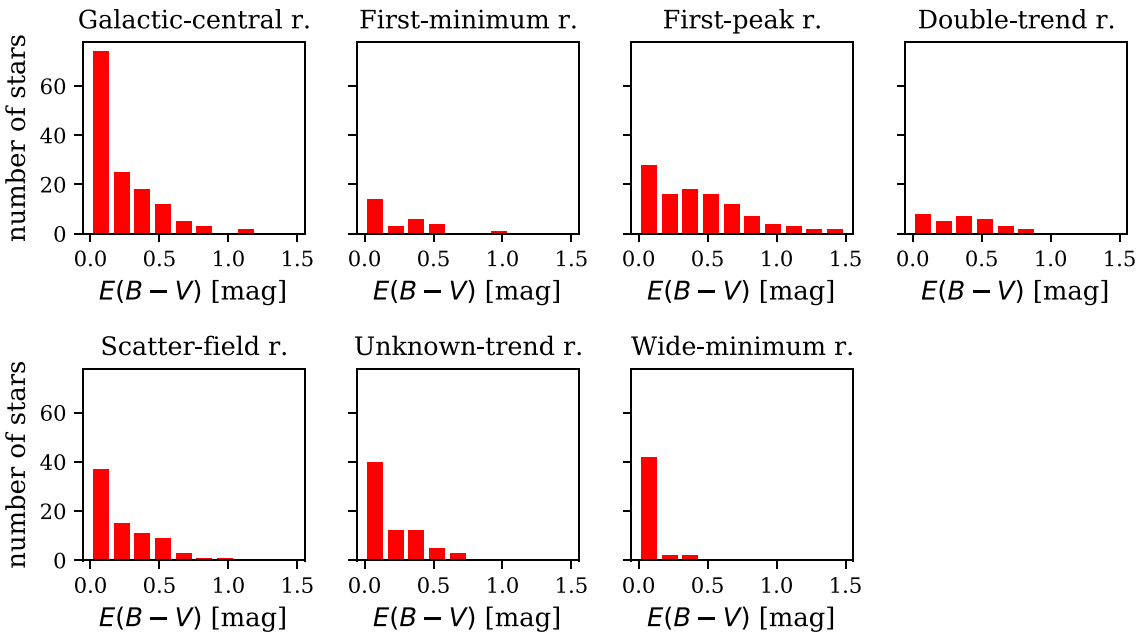


Figure A14. Distribution of colour excess values (for distances up to 2 kpc) in all regions defined in Table 2.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Chapter 8

A Comparison Of The Simulations And Observations For A Nearby Spiral Arm

***Piecka, M.**, Paunzen, E., 2021, Frontiers in Astronomy and Space Sciences, Volume 8, id.89, 20 pp.*

One of the problems with tracing spiral arms of galaxies with open clusters is that the ages are often not taken into account. For example, the map of open clusters from Kharchenko et al. (2013) does not seem to suggest that the distribution of clusters follows the modelled spiral arms. However, once the ages are considered, the younger clusters do show a trace of a structure of the nearby spiral arm (Castro-Ginard et al. 2021). Still, other spiral arms in our Galaxy cannot yet be studied.

It is difficult to know how the age of a cluster affects its ability to trace a spiral arm. It is known that the spiral arms orbit a galaxy with an angular velocity which, generally speaking, differs from the rotational curve of that galaxy. Clearly, the clusters should become separated from their point of origin within a spiral arm (assuming that the arms are long-lived structures). This is a well-known fact for stars and clusters. However, only complex numerical models were generally used to study how the positions of these objects vary with respect to the spiral arm within which they were born.

Interestingly, simple analytical and numerical models suffice to show the drift of clusters from the arms. Let us assume that we want to study a spiral galaxy for which we know its rotational curve, pattern speed associated with the spiral arms and velocity dispersion. Furthermore, let us approximate that the orbit is circular. In such a case, the drift of the clusters (separation distance of a cluster from the spiral arm) becomes a relatively simple function of time. Such function can be modelled for our Galaxy (and several other observed spiral galaxies), and the drift for clusters of various ages can be shown as a function of the galacto-centric distance. The resulting curve can be very simple (e.g. M 51) but can also become quite complicated (e.g. NGC 3310).

Velocity dispersion has one of the largest effects on the drift. Even the simplest model of the motion of clusters in a galaxy will show a drift once the dispersion is included. In the more realistic kinematic models, the drift may be dominated by the rotational curve. The difference in the orbiting velocity for different clusters in the arm will result in the

formation of a disk-like (or ring-like) distribution after about 1 Gyr, leaving almost no trace of the spiral arms in the distribution. Further investigation shows that the spiral arms of our Galaxy can be traced by clusters younger than about 100 Myr. This is in agreement with the results provided in other (more sophisticated) works.

MP studied the distribution of clusters, prepared the analytical and a simple numerical method to analyse spiral arms, was responsible for writing the larger part of the paper, and submitted the manuscript.

frontiers in Astronomy and Space Sciences

ORIGINAL RESEARCH
published: 17 June 2021
doi: 10.3389/fspas.2021.642776

A Comparison of the Simulations and Observations for a Nearby Spiral Arm

*Martin Piecka and Ernst Paunzen**

Department of Theoretical Physics and Astrophysics, Masaryk University, Brno, Czechia

The analysis is focused on the ability of galactic open clusters to trace the spiral arms, based on the recent data releases from Gaia. For this, a simple 1D description of the motion of spiral arms and clusters is introduced. As next step, results are verified using a widely accepted kinematic model of the motion in spiral galaxies. As expected, both approaches show that open clusters older than about 100 Myr are bad tracers of spiral arms. The younger clusters (ideally < 30 Myr) should be used instead. This agrees with the most recent observational evidence. The latest maps of the diffuse interstellar bands are compared with the spiral structure of the Milky Way and the Antennae Galaxies. The idea of these bands being useful for studying a galactic structure cannot be supported based on the current data.

Keywords: galaxy: kinematics and dynamics, galaxy: disc, galaxy: evolution, galaxy: structure, open clusters and associations: general, local interstellar matter

OPEN ACCESS

Edited by:
Jacques R. D. Lepine,
University of São Paulo, Brazil

Reviewed by:
Antonella Vallenari,
Osservatorio Astronomico di Padova (INAF), Italy
Vadim Vadimovich Bobylev,
Pulkovo Observatory, Russia

****Correspondence:***
Ernst Paunzen
epaunzen@physics.muni.cz

Specialty section:
This article was submitted to Fundamental Astronomy, a section of the journal Frontiers in Astronomy and Space Sciences

Received: *16 December 2020*
Accepted: *14 May 2021*
Published: *17 June 2021*

Citation:
Piecka M and Paunzen E (2021) A Comparison of the Simulations and Observations for a Nearby Spiral Arm. Front. Astron. Space Sci. 8:642776. doi: 10.3389/fspas.2021.642776

1 INTRODUCTION

Galaxies are among the largest objects which are still bound by the gravity of the constituent matter (just after the galaxy groups and clusters). Typically, we distinguish between several classes of galaxies (e.g., see Binney and Tremaine, 2008)—we will be focusing on the spiral class of galaxies. Specifically, the current ability to trace the spiral arms with a specific set of objects is reviewed and compared with simple simulations.

There are many different objects useful as tracers of the spiral structure of galaxies. Spiral arms are thought to represent the bulk of the star-formation in their host galaxies. It is reasonable to assume that the distribution of the stars of spectral class O or B will be very useful, as well as the associated molecular masers (Reid et al., 2019), H II regions (Moffat et al., 1979) and molecular clouds (Loinard et al., 1999; Hou and Han, 2014). We can also extrapolate that young stellar associations, which did not yet have the time to move away from their birthplace, can also be regarded as an effective tool for tracing the spiral structure. Classical Cepheids form another group of stars that helped with the identification of the Galactic spiral structure (Fernández et al., 2001; Bobylev et al., 2021). Finally, H I radio emission represents a standard tool for tracing the spiral structure of galaxies (Russeil, 2003; Hou and Han, 2014).

Open stellar clusters are regarded as an excellent tool in astrophysics. They almost always form in the galactic disks and follow the galactic rotation—this distinguishes them from the globular clusters, which belong to the population of the galactic halo (or the bulge). The ages of the constituent stars are usually taken to be the same, since an open cluster is formed within about a few mega/million years (Myrs) after the start of the collapse of the giant molecular cloud. If we consider that all of the member stars were born from the same material, we can expect that the metallicity must be almost the same for each and everyone of them. Moreover, cluster diameters are usually much smaller than their distances from us. In principle, this means that we need to evaluate the distance from only one of the members. However, the whole sample of cluster members should be used to find distances with

higher precision. Finally, the extinction of the cluster can be also assumed to be the same for all of its stars, but differential reddening (important usually only for young clusters) must be ruled out first.

Regardless of the choice of the spiral arms tracer, the positions of these objects must be specified. In some situations, for example when observing Cepheids, there exist relations which can be used to precisely determine the distances. However, here we will mostly be interested in star clusters, for which a different method must be applied (although some clusters do contain Cepheid variables). One good approach is to study the photometric data of the cluster members in the colour-magnitude diagram by applying the isochrone fitting procedure. The output of this procedure is the distance, the reddening, the age, and sometimes also the metallcity of the cluster members (Jørgensen and Lindegren, 2005; Pöhnl and Paunzen, 2010; Netopil et al., 2015). Presently, a very popular alternative is to study the parallaxes (and therefore also the distances) of the clusters, based on the data from the Gaia satellite mission (Cantat-Gaudin et al., 2018, 2020).

An interesting topic for the researchers of the interstellar medium is concerning the diffuse interstellar bands (DIBs, for a recent review see Krełowski, 2018). Almost each of these unidentified absorption features is believed to trace a quite unique set of physical conditions. Many of these bands seem to originate from the regions with an appreciable interstellar UV-radiation field (Jenniskens et al., 1994), and the measured equivalent widths are shown to correlate with the extinction (e.g., see Herbig, 1995; Raimond et al., 2012; Kos and Zwitter, 2013; Zasowski et al., 2015). These properties suggest that there could be some connection between the appearance (and strength) of these bands and the local distribution of the UV stellar sources. Is it possible to use these interstellar tools to map the spiral arms?

In **Section 2**, we will explore the current observational knowledge of the spiral arms and review the latest large-scale survey of the cluster distances. **Section 3** is focused on studying how clusters become less useful, over time, for tracing the spiral structure of galaxies; comparison of the simulations based on simple kinematic models is made with the observations of open clusters. Finally, we study the maps of diffuse interstellar bands in **Section 4**, where we review their ability to trace spiral arms in the Milky Way and in other galaxies.

2 OBSERVATION OF SPIRAL ARMS

For the most part, we will be ignoring the influence of the galactic-bulge on disk kinematics. It may be appropriate to distinguish between the orbits of physical objects, such as stars, and the motion of the spiral arm. In the first approximation, these two motions do not influence each other since there is no collision between two physical bodies—the spiral arms are physically disconnected from the orbital motion in the galactic disk. Instead, the density wave (as we shall discuss, spiral arms are the result of density waves) simply passes through a medium (the interstellar matter or the stellar disk), locally enhances the density of the medium, which then returns into its initial state (ignoring the effect of star-formation and any stellar evolution and outflows) after the wave has moved away again.

In this section, we shall first briefly mention the kinematic situation in spiral galaxies. Afterwards, a review of the most interesting observational aspects of the spiral arms will be presented.

2.1 Rotation Curves and Spiral Structures of Galaxies

One of the key observational features of galaxies one wishes to study is the rotation curve. This curve represents a relation between the galacto-centric radius and the circular component of the orbital velocities of the observed objects. In the Milky Way (and in other galaxies), this curve deviates from the curve predicted by the models based on the observed matter outside of the bulge—the observed rotation curve is actually above the theoretical curve. The most accepted explanation is the presence of the dark matter, which interacts with other particles only *via* gravity (Chrobáková et al., 2020).

In addition to the rotation curve, the motion of the objects in a galaxy consists also of a dispersion of velocities σ. The magnitude of the dispersion may differ based on which component of the velocity we are looking at. In spiral galaxies, we generally find that $\sigma_r > \sigma_t > \sigma_z$, corresponding to the radial, tangential and vertical component (in the cylindrical coordinate system), respectively.

A circular orbit cannot describe the motion of an object in a galaxy. However, under the assumption of an axisymmetric galactic potential, we may derive a solution to the equations of motion if we let the radius of the studied object to slightly vary. The solution to such problem is called the epicyclic approximation, and it is very useful for studying the orbits of stars in a galaxy. The final result of this approximation is a constructed ellipsoid (or an ellipse, if we ignore the vertical motion). The studied object is located always on the surface of the ellipsoid (the motion is oscillatory), while the center of the ellipsoid moves on a circular orbit around the galaxy. Ignoring the presence of vertex deviation (Smith et al., 2012), the orientation of the ellipsoid is fixed if we follow the circular motion of its center—the x-axis usually points in the radial direction, the y-axis is tangent to the circle, and the z-axis coincides with the vertical axis in the galacto-centric cylindrical coordinate system. In principle, the size of the ellipsoid in the epicyclic approximation results from the shape of the velocity ellipsoid constructed from the velocity dispersion.

Since more than a century ago, it has been very clear from the observations that some of the galaxies display a distinct spiral structure. Astronomers have been trying to study the physics behind this structure since the beginning—as we do not plan to consider the dynamical aspects of the problem, we redirect the reader to:

- Toomre (1977), a review of the theories behind the spiral arms
- Sellwood et al. (2019), one of the most recent studies of the origin of the spiral structure, based on the Gaia data

The most important observational fact to our analysis is that the spiral arms are long-lived. The most accepted theory is that the arms must result from a density wave travelling around the galaxy, which are excited by a rotating (pattern) potential. Assuming that the orbits result from the epicyclic approximation, resonances appear at certain radii (called Lindblad resonances), which dynamically force the existence of the density waves. Currently, the origin of this pattern is still poorly understood (e.g., Sellwood, 2012).

The velocity of the orbiting pattern has been analyzed multitudes of times. In the Milky Way, the pattern velocity is found to be roughly constant over several kpc in the radial direction.

There are several simplifications which we assumed in this section. For example, the spiral structure of different galaxies may vary, and no single theoretical model can describe all of these structures (Seigar and James, 1998; Hou and Han, 2014; Díaz-García et al., 2019). Furthermore, we have completely ignored the time-evolution of the galactic structure (which may be affected by mergers, for example), although this effect should be negligible on the time-scales considered in **Section 3**.

2.2 Interstellar Medium Tracers

As we have already mentioned, most of the star-formation occurs in the spiral arms. This means that the high-density medium (molecular clouds), which contains a significant amount of dust when compared with the other regions of the galaxy, must be a good indicator of the presence of spiral arms. Indeed, optical images of spiral galaxies show dark regions coiling around the galactic centers. Historically, this was one of the first hints of the spiral structure of galaxies. The dust itself can be best observed by studying atomic and molecular signatures in the spectra (Bouwman et al., 2019).

Most of the volume of a galaxy is dominated by a very hot plasma (referred to as the hot ionized medium, (HIM) alternatively the coronal gas). Both, in terms of the abundance and the mass, ionized hydrogen dominates these regions. This medium mostly fills the galactic halo but is also present in the galactic disk. Once we get inside a denser part of the disk, hydrogen becomes shielded from the ionizing radiation by the surrounding layers of the medium—this results into primarily neutral hydrogen inside such regions, which can be called H I clouds (Bekki et al., 2005). Clearly, neutral hydrogen prevails mostly near, or inside, the spiral arms. The most useful tracer of such medium is the 21-cm radio emission line, originating from the spin-flip transition of the electron in the hydrogen atom at the ground-state.

If we want to study the molecular clouds, and therefore some of the densest interstellar regions in a galaxy, we must look for the most abundant molecule—the molecular hydrogen. The density of the medium affects the abundances of H_2 in two ways: 1) the most likely process of the formation of H_2 is on the surface of dust grains, which we may find in the denser medium, 2) the molecular hydrogen can be dissociated by the UV radiation ($h\nu < 11$ eV), so a substantial protective layer must be present above the molecular region, which happens only in the molecular clouds.

However, since the molecular hydrogen regions have to be surrounded by a dense layer, observations in the optical part of the spectrum become obscured. To get an image of the inner parts of a molecular cloud, we must look, again, at the longer wavelengths. The second most abundant molecule in the Universe is CO, which has a very prominent line at 0.26 cm. In practice, this line is the most useful when probing dense regions of galaxies (Tang et al., 2016).

We have already mentioned that H II regions are very important when studying spiral arms. These regions are very common around hot young stars (Conti and Crowther, 2004). Due to the state of an H II region, electrons interact with the ionized hydrogen in a free-free process—bremsstrahlung. In the first approximation, such region is in a local thermodynamics equilibrium (unlike in the case of HIM) and the motion of the particles is described by the Maxwell-Boltzmann distribution. However, this is only useful if the probed medium is dense enough for this to be true. In any case, a very common approach is to study the continuum radio emission in order to probe the H II regions.

Finally, we can study also the regions associated with the early stages of star-formation. These can be observed looking at the masers from CH_3OH and H_2O at 4.48 and 1.36 cm (Beuther et al., 2002; Zhang et al., 2019), respectively.

2.3 Embedded Clusters

An embedded cluster is a group of young stellar objects (YSOs) that is still embedded in its natal molecular cloud, i.e., in dust and gas. It is typically not fully observable at optical wavelengths due to large extinction caused by the dust grains in the cloud, but it can be seen in the near infrared. There YSOs emit a significant amount of radiation (Robitaille et al., 2006), and the dust is more transparent (Fitzpatrick, 1999).

The loss of gas defines the end of star-formation in an embedded cluster and may also cause the young cluster to dissolve (Lada et al., 1984). Connected with that the time-scale for a cluster to clear enough material becoming optically visible is still a matter of debate. Leisawitz et al. (1989) conducted a CO survey of open cluster regions and found a value of about 5 Myr whereas Morales et al. (2013) analyzed several young clusters with molecular material, and proposed an upper limit of the embedded phase of 3 Myr, respectively. However, the time-scale is sensitive to the initial mass function (IMF). More massive stars develop H II regions that are much more efficient in dispersing the cloud material than the outflows from low-mass stars (Matzner, 2002) and are therefore very efficient to clear their surroundings.

The time-scale on which gas is removed from the embedded cluster not only affects the mass and number of stars in the surviving cluster, but also its degree of mass segregation and the density profile (Er et al., 2009). The parameters that strongly affect the outcome of the out-gassing phase are the star-formation efficiency (SFE), the population of the most massive stars, and the efficiency of the radiative coupling (He et al., 2019). The typically observed SFE of about to 0.4 can be explained by using radiative magneto-hydrodynamic simulations with self-consistent star-formation and ionizing radiation (Geen et al., 2017). They

showed that the SFE can even approach unity for very dense clouds.

From an observational point of view, the number of embedded clusters is too high with respect to the number of observed gas-free clusters for a given age (Lada and Lada, 2003). Within the solar neighborhood the observed distribution of cluster ages suggests that more than 95% of embedded clusters dissolve into the field within 100 Myr. There are several disruption processes suggested such as the loss of the remnant gas and tidal forces (Elmegreen and Hunter, 2010). The latter incorporates tidal shocks from passing gas clouds—this becomes important only at later ages (~1 Gyr, Gieles et al., 2006). The aspects of these effects have been studied in more details using N-body simulations (Baumgardt and Kroupa, 2007; Smith et al., 2011).

The search for new Galactic embedded clusters is still ongoing, mainly based on Infrared (IR) photometry. For example, the European Southern Observatory (ESO) public survey VISTA variables Vía Láctea Survey (VVV) presented a list of 88 new candidates (Solin et al., 2014). Some follow-up investigations (for example, Borissova et al., 2020) confirmed their characteristics as embedded clusters.

2.4 Gaia Era—Revisiting Open Clusters

For estimating the membership probability of a star in a field of an open cluster, in the ideal case, the parallax, the proper motion, the radial velocity and a colour information is needed. These parameters have then to be compared with the mean cluster ones and a color-magnitude diagram constructed. For this purpose, several, mostly independent, methods have been developed already in the pre-Gaia era and updated since then (for example, von Hippel et al., 2006; Krone-Martins and Moitinho, 2014; Perren et al., 2015; Balaguer-Núñez et al., 2020).

With the successful launch and the first data release of the Gaia satellite, a new era in star cluster research began. Until then, the most precise parallax measurements came from the Hipparcos satellite (van Leeuwen, 2007) which was limited to approximately 12th magnitude and therefore to star clusters in the solar neighborhood, only. For the proper motions, the United States Naval Observatory CCD Astrograph Catalog (UCAC, Zacharias et al., 2017) and the PPM Star Catalogue (Bastian and Röser, 1993) which was later extended to the PPMX (Röser et al., 2008) and PPMXL (Roeser et al., 2010), were two available independent data sources. A comparison of these two sources showed offsets and systematics which were later investigated in more details on the basis of the Gaia DR2 (Shi et al., 2019).

The Gaia DR1, which is also based on Hipparcos and Tycho-2 data, was validated with open clusters and other methods (Arenou et al., 2017; Gaia Collaboration et al., 2017). The reconstructed mean cluster parallaxes and proper motions were generally in very good agreement with earlier Hipparcos-based determination. The problem of the discrepant distance of the Pleiades was finally solved, reconciling astrometric results with other observational methods.

The first catalogue based purely on Gaia data (Gaia DR2) was released relatively shortly afterwards (Arenou et al., 2018; Gaia Collaboration et al., 2018b). In the series of initial papers from the Gaia consortium, Gaia Collaboration et al. (2018a) investigated the Hertzsprung-Russell diagram using 46 open clusters. Still, the difference between the observed and theoretical main sequence of the Pleiades remained.

Later on, the significant discrepancy of the main sequence of Pleiades was solved using a cross-correlation of the Gaia catalogue with large-scale public surveys to complement the astrometry of Gaia with multi-band photometry from the optical to the mid-infrared (Lodieu et al., 2019).

The first comprehensive study for 1,229 star clusters including a list of members and cluster parameters was presented by Cantat-Gaudin et al. (2018). They used an updated version of the UPMASK algorithm (Krone-Martins and Moitinho, 2014) which makes a membership assessment based on an iterative process, principal component analysis, clustering algorithm, and kernel density estimations. They compared the estimated distances with those for 38 open clusters from the Bologna Open Cluster Chemical Evolution Project (BOCCE) project (Bragaglia and Tosi, 2006). Later on, Bossini et al. (2019) published a cluster parameter determination for 269 aggregates using an automated Bayesian tool (von Hippel et al., 2006). This method uses both, the Gaia DR2 astrometry and photometry for fitting isochrones to the high probability member stars. They also presented a comparison with literature values taken from Dias et al. (2002) and Kharchenko et al. (2013), Kharchenko et al. (2016). The differences of the cluster parameters showed a huge spread (see **Figures 9**, **10**, therein) which cannot be explained by the superiority of the Gaia data alone. A similar study was published by Monteiro and Dias (2019) who investigated 150 poorly studied open clusters from which 80 turned out to be non-physical aggregates. We have to emphasize that the determination of the cluster parameters themselves are sometimes severely constricted by the choice of the isochrones, the metallicity, and even the photometric system (Netopil et al., 2015; Netopil et al., 2016; Dias et al., 2021). Cantat-Gaudin et al. (2020) used a set of objects with available well-determined parameters to train an artificial neural network. Thus, as a next step, they estimated the cluster parameters from the Gaia photometry of high probable members and their mean parallax for 1,867 aggregates. Recently, Monteiro et al. (2020) determined cluster memberships using a maximum likelihood method applied to Gaia DR2 astrometry for 45 aggregates. They presented an improved isochrone fitting code taking into account the interstellar extinction using an updated extinction polynomial for the Gaia DR2 photometric bandpasses and the Galactic abundance gradient as a prior for metallicity.

Almost all studies about individual star clusters since 2018 include the usage of either Gaia photometry or astrometry (for example, Yontan et al., 2019; Baratella et al., 2020; Straižys et al., 2020; Niu et al., 2020).

It can be concluded that different applied methods resulted in intrinsically consistent cluster parameters which are not always compatible with previous published values. What is still missing, for example, is a comprehensive study including also available Johnson UBV (Mermilliod, 2006) and 2MASS JHK$_s$ (Skrutskie

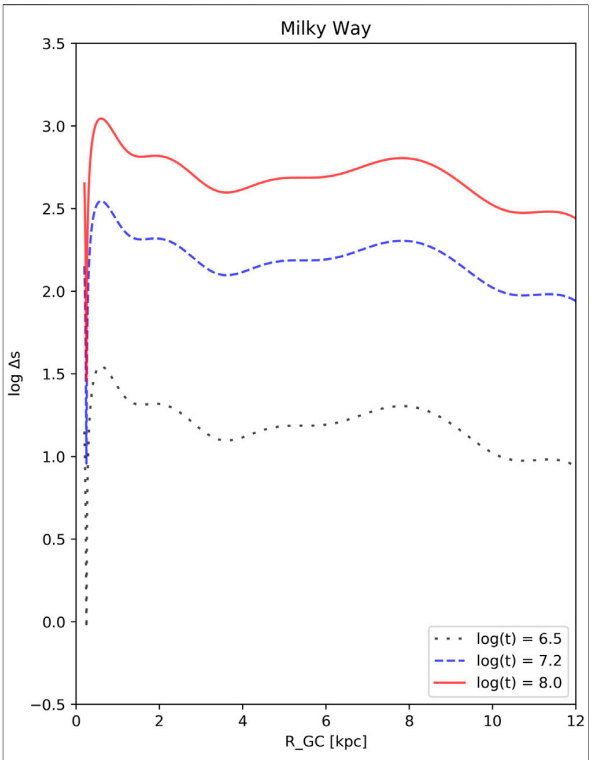


FIGURE 1 | The orbital drift of the open cluster in our Galaxy from their host spiral arm as a function of the galacto-centric radius and time. The values of Δs under the logarithm are in parsec. Used parameters: Ω_p = 25 km/s/kpc, ΔR = 0.1 kpc, rotation curve from Sofue et al. (1999), approximated by a high-degree polynomial (deg > 15).

et al., 2006) photometry for isochrone fitting. Furthermore, such a study should be based on more than one numerical method for the cluster parameter estimation.

Due to the high accuracy and the completeness of the Gaia DR2, the search for previously unknown Galactic open clusters was intensified. As a result, several hundreds of apparently new open clusters has been published in the last 3 years (for example, Castro-Ginard et al., 2018; Liu and Pang, 2019; Castro-Ginard et al., 2019; Castro-Ginard et al., 2020). Although, in general, a thorough cross-check with the already published catalogues of open clusters has been performed in the corresponding papers, still a certain percentage of newly announced aggregates has been already known before or are just a sub-population of a larger one. One has to keep in mind that due to the huge amount of data, only automatic routines are capable to process all the needed information. Therefore, a careful inspection of the results in a graphical form is strongly advised.

When radial velocities are included, the characterization of open clusters is still not satisfactory. This is mainly caused by the fact that a large database of reliable and homogeneous radial velocities is still very much missing. Although the Gaia mission also includes the measurements of radial velocities, the chosen spectral window (8,450–8,700 Å) is optimized for cool type stars. In these regions, the spectra of upper main sequence stars are dominated by the Paschen lines. Together with the, in general, moderate to high rotation rate of these stars, the radial velocities cannot be accurately measured. The current available data set from the Gaia DR2 includes stars in the effective temperature range from 3,550 to 6,900 K (Katz et al., 2019) which roughly transforms to spectral types from M2 to F2, respectively. The complete upper main sequence is missing which is essential when studying young open clusters (Liu et al., 1991). This whole situation is reflected in the paper by Soubiran et al. (2018) who investigated the spatial and velocity distribution of 861 open clusters. From their sample, for 406 aggregates (the so-called high-quality sample) the mean radial velocity relies on at least three members, only. If one keeps in mind the different causes of radial velocity variability and the often large amplitudes (Percy, 2007), a much larger number of individual measurements are needed to come to a statistically sound result (Mermilliod et al., 2009). This problem becomes even more evident when studying the internal kinematics and dynamics of a star cluster (Gaburov et al., 2008).

Finally, the different Gaia data releases even challenge the classical definition of a star cluster itself. Oh et al. (2017) and Faherty et al. (2018) found many thousands of star groups in the solar neighborhood with up to 10 members. These moving groups share a common kinematic characteristics. Naturally, questions arise like: are these groups dissipated star clusters? What is the minimal number of members for an open cluster? What is the minimal total mass for an open cluster?

The future Gaia releases will result in more precise photometric, spectroscopic, kinematic, and astrometric data. They will challenge our current theories about the formation and evolution of stellar clusters on the basis of observational data.

3 COMPARING SIMULATIONS WITH OBSERVATIONS

3.1 Choice of Kinematic Models

In this section, we will mention several simulations of clusters in spiral arms. To begin with, several observational values need to be taken from the literature.

One of the most fundamental quantities necessary for Galactic studies is the knowledge of the position of our Sun in the Galaxy. For this, we need to know the direction of Galactic rotation and our distance from the Galactic center R_G. The value of this distance varies in the literature. For the purpose of our analysis, we have chosen the value of R_G = 8.3 kpc, measured by Reid et al. (2014). In the same work, the spiral structure of our Galaxy was analyzed. Presented was an equation of the spiral arms and the parameters which well represent the Scutum, Sagittarius, Local, Perseus and Outer arms. Their typical width was found to be of the order of hundreds of parsecs.

Another important aspect of the Galactic kinematics (and dynamics) is the rotation of the Galaxy. There are three main

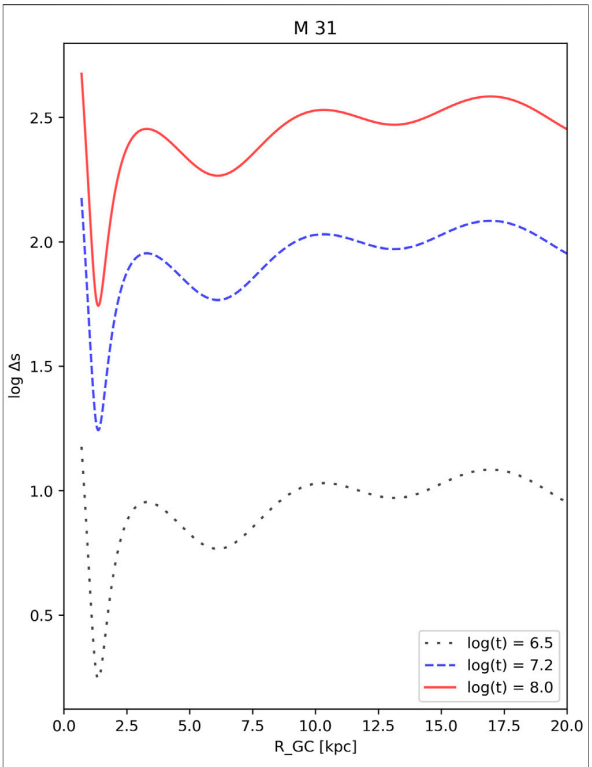


FIGURE 2 | The orbital drift of the open cluster in M 31 from their host spiral arm as a function of the galacto-centric radius and time. The values of Δs under the logarithm are in parsec. Used parameters: Ω_p = 15 km/s/kpc, ΔR = 0.1 kpc, rotation curve from Sofue et al. (1999), approximated by a high-degree polynomial (deg > 15).

components of the motion which are of interest: the rotation curve, pattern velocity and velocity dispersion.

Rotation curves of several galaxies were presented by Sofue et al. (1999) which are based on CO observations. Most of our focus is aimed at our Galaxy but other galaxies will be explored as well in the following.

The estimated pattern velocity for the spiral arms depends on the method used to derive this value. For our Galaxy, this problem was described by Gerhard (2011). The values seems to fall in the interval between 15 km/s/kpc and 40 km/s/kpc. Dias and Lépine (2005) found from the analysis of open clusters that the best matching value is 25 km/s/kpc. This is close to the value determined by Dias et al. (2019) (Ω_p = 28 km/s/kpc), who studied the motion of fairly young open clusters. It must be kept in mind that, in the standard notation for the Galacto-centric coordinates (used also in this work), pattern velocity and angular orbital velocities of clusters and field stars must have a negative sign ($\Omega_p < 0$, $\Omega < 0$).

One of the most recent analysis of cluster kinematics was presented by Soubiran et al. (2018). In the data set they published, all components of the 3D velocity vectors are included. However, before we start computing any statistical values, we must realize that the velocity dispersion depends heavily on the cluster age. For our analysis, it is best to exclude all clusters with log(age) > 7.5. For this, we have used ages from Bossini et al. (2019) and, if a cluster was not included in this work, we have also considered the data from Kharchenko et al. (2013). After the exclusion, we are only left with 18 clusters from Soubiran et al. (2018). The standard deviation of the cluster velocities represents a good measure of the velocity dispersion, which we found to be $\sigma_t \approx 7.2$ km/s (15.0 km/s, considering also the excluded clusters) for the tangential component of the velocity vector. The value of the radial component is similar, $\sigma_r \approx 6.8$ (10.5) km/s.

3.2 Pattern Breaking—Relative Orbital Drift

First, let us choose the coordinate system and the frame of reference. Since we want to understand how open clusters follow spiral arms across their host galaxies, we have decided to pick the stationary arm as the frame of reference. For the spiral arm, a galacto-centric coordinate system is possibly the best choice. We will start by defining the rectangular coordinates X and Y as they are commonly used. The center of the galaxy is located at the point $(0, 0)$ while Sun can be found at $(-R_G, 0)$, where $R_G \sim 8.3$ kpc. The galacto-centric radius R is then simply the distance of the point (X, Y) from the center of the coordinates system. The azimuthal angle ϕ is zero when we are looking in the direction towards the Sun, and it increases in the direction of the galactic rotation. The conversion between the rectangular and the polar coordinates is given by

$$\begin{aligned} X &= -R\cos(\phi) \\ Y &= R\sin(\phi) \end{aligned}$$

We expect that the distribution of clusters born in a spiral arm will deviate over time from the initial (spiral) distribution. For simplicity, we shall assume that the motion of the spiral arm resembles the rotation of a rigid body with a constant angular velocity Ω_p (therefore, Ω_p is independent of R). Moreover, we expect that there is no momentum transfer between the density wave and the star-forming regions. Assuming a circular orbit of the cluster in the galaxy, the distance along an arc of a circle between the spiral arm and a given cluster is

$$\begin{aligned} s = \Delta v\, T &= \left[V_{rot}(R) - \Omega_p R\right] \cdot 10^{\log(age)} \\ &\approx 1.02 \cdot \left[V_{rot}(R) - \Omega_p R\right] \cdot 10^{A-6}\ (pc), \end{aligned} \tag{1}$$

where R is the galacto-centric radius of the cluster in (kpc), $V_{rot}(R)$ is the orbital velocity of the cluster given by the rotation curve in (km/s), Ω_p is the pattern velocity of the spiral arm in (km/s/kpc), and A = log(age) is the logarithm of the cluster age in (yr). Since the motion of the pattern is strictly different from the rotation curve (except for the corotation regions), it is clear that most clusters born in a spiral arm will drift away from the position of this arm. However, the situation is quite different for each individual cluster since this drift depends strongly on the cluster age. In the later sections, we will analyze the populations of clusters based on their ages and use more realistic orbits.

Unfortunately, we cannot make a proper use of the quantity *s*. We would like to use the distribution of clusters to find the

location of the spiral arm they were born in. This prohibits us from knowing the location of the arm in advance, therefore *s* bears no real information. To deal with this, we can use the number of available clusters in our advantage. We can calculate the relative drift of two clusters located (radially) some distance apart, assuming that they were born at the same time (in the same phase of the arm's orbit). We find that

$$\Delta s \approx 1.02 \cdot \left[V_{rot}\left(R + \frac{\Delta R}{2}\right) - V_{rot}\left(R - \frac{\Delta R}{2}\right) - \Omega_p \Delta R \right] \cdot 10^{A-6}\ (\mathrm{pc}) \quad (2)$$

where ΔR is the radial distance of the clusters. The clusters will drift away from each other simply because their orbits do not follow the kinematics of a rigid body. Furthermore, we may assume that V_{rot} = const, which leads to the simplification of **Eq. 2**

$$\left|\Delta s\right|_{V_{rot}=\mathrm{const}} \approx 1.02 \cdot \Omega_p \Delta R \cdot 10^{A-6}\ (\mathrm{pc})$$

This should, approximately, hold for the region in the Solar neighbourhood. For illustration, we shall consider the pattern velocity Ω_p = 25 km/s/kpc. Then we will find that the relative orbital drift for two cluster ΔR = 0.1 kpc apart will be about 8 pc for A = 6.5, 80 pc for A = 7.5 and 250 pc for A = 8.0. Again, this is simply the result of the fact that the cluster orbits do not follow the spiral arm but rather the Galactic rotation. For comparison, the width of the spiral arms of our Galaxy tends to be around 300 pc.

If we were to assume that the clusters follow the spiral arm perfectly, the term $[V_{rot}(R) - \Omega_p R]$ in **Eq. 1** would be zero, and therefore also *s* would be identically zero. However, there is still a possibility for the drift to occur. So far, we have not considered the fact that the velocity field of a galaxy also consists of a velocity dispersion. If we were to include it in **Eq. 1**

$$s \approx 1.02 \cdot \left(V_{rot}(R) - \Omega_p R + \sigma_t\right) \cdot 10^{A-6}\ [\mathrm{pc}]$$

then the term in the brackets can become very small only if the dispersion of velocities is also very small (σ<1.0 km/s). The term σ_t represents the component of the velocity dispersion tangential to the circular orbit at the given galacto-centric radius. We note that σ_t does not cancel out in Δs because it does not represent a value but rather a function generating random values with a given dispersion.

The situation gets even more complicated when we assume that the clusters can also drift in the radial direction. Obviously, the radial drift would be fully dominated by the velocity dispersion since no other significant terms would be present in the bracket in **Eq. 1**. For this reason, such drift is purely statistical and can only enhance the total value of *s*. Since the orbits are expected to be bound, no linear motion is expected in the radial direction anyway. Later on, we will return to the two-dimensional drift dominated by the velocity dispersion.

The problems, which we are about to study, have already been discussed and observationally confirmed in the past (for example, see Roberts, 1969; Mathewson et al., 1972). Our intention is to look at the comparison of our simulations with the newest available data for star clusters. The relevance of using such objects as tracers of spiral arms must be tested with each major observational leap forward, which Gaia definitely represents.

3.3 Pattern Breaking—Rotation Curve and Age Dependence

For start, we will assume zero velocity dispersion. Let us compare the observed rotation curves of different galaxies (Sofue et al., 1999) with the model described in the previous section. This will be done using three different clusters ages: log(age) = (6.5, 7.2, 8.0). It is worth noting that many galaxies (even our own) display sharp changes in the rotation curves in the inner regions—we should expect that at some galacto-centric radius R (near the inflection point) the bracket in **Eq. 2** may becomes zero. Around this point, Δs changes sign, although we only care about the magnitude (the absolute value) of the relative drift.

For our Galaxy, we have used Ω_p = 25 km/s/kpc and ΔR = 0.1 kpc (**Figure 1**). At a distance below ~1 kpc, the function Δs peaks with the values 35, 350, and >1000 pc for the three different ages, respectively. Afterwards, the curve slowly falls down, until it reaches a secondary maximum (with a height smaller by a factor of 0.57 of the primary maximum) at around 8.0 kpc. It seems that the drift becomes significant enough in the Solar neighbourhood for clusters with log(age) > 7.5—specifically, the drift starts to reach values comparable with the widths of spiral arms (~300 pc for the Local Arm).

Braun (1991) studied the structure and kinematics of the Andromeda Galaxy (M 31). One of the results of this work is the spiral pattern velocity, Ω_p = 15 km/s/kpc. This helps us to explore the Andromeda Galaxy in the same way as our own Galaxy (assuming ΔR = 0.1 kpc). **Figure 2** shows the resulting curves. It is immediately apparent that there is no global maximum (although we have cut out the region R < 0.5 kpc). The relative orbital drift varies only slightly with the galacto-centric radius, averaging at about 9, 90, and 285 pc, respectively. These are the smallest values among the four galaxies analyzed in this section. Indeed, the older clusters [8.5 > log(age) > 8.0] appear to trace the spiral arms quite well (Caldwell et al., 2009). However, there is another point to be made about the older clusters, which we shall explore later on.

The next galaxy we want to inspect is M 51. It was discussed in Meidt et al. (2008) that there are quite obvious variations in the spiral pattern velocity of this galaxy at different galacto-centric radii. The pattern velocity starts from around 90 km/s/kpc in the inner regions of the galaxy and reduces down to about 50 km/s/kpc in the outer regions. For our analysis, we have used the latter value and ΔR = 0.1 kpc. **Figure 3** shows that the situation is quite different from what we saw in the Milky Way. If we disregard the region R < 0.5 kpc, we see that there is no apparent global maximum and that Δs oscillates around a well defined value of 34, 340, and 1,050 pc, respectively. The mean value of Δs appears to be somewhat larger than in Milky Way. However, the situation would change considerably if we assumed that $\Omega_p \neq$ const. Let us try a function $\Omega_p = \frac{108}{R^{0.8}}$, which has a value 90.0 km/s/kpc at R = 1.2 kpc and 50.0 km/s/kpc at R =

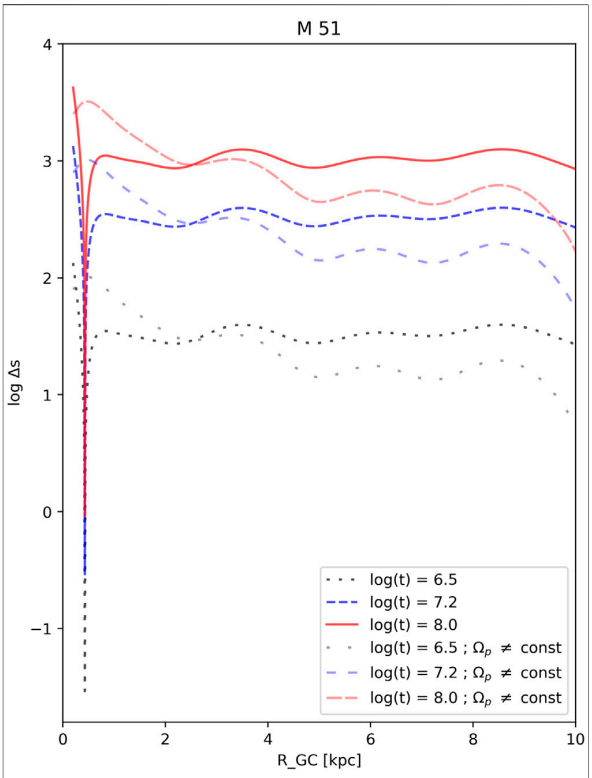


FIGURE 3 | The orbital drift of the open cluster in M 51 from their host spiral arm as a function of the galacto-centric radius and time. The values of Δs under the logarithm are in parsec. Used parameters: Ω_p = 50 – 90 km/s/kpc, ΔR = 0.1 kpc, rotation curve from Sofue et al. (1999), approximated by a high-degree polynomial (deg > 15).

2.5 kpc—the plot of $\Delta s(R)$ curves calculated based on this function is overlaid in **Figure 3**. The results for different functions Ω_p clearly vary—the power law function seems to bring $\Delta s(R)$ of M 51 closer to what we see in our Galaxy (where we, however, assumed constant Ω_p).

Finally, Mulder and Combes (1996) showed that Ω_p = 14 km/s/kpc provides the best fit to the observations of NGC 3310. Once again, let us use the value ΔR = 0.1 kpc. The rotational curve of this galaxy was taken from Epinat et al. (2008). The function $\Delta s(R)$ (**Figure 4**) is much more interesting than in the previous cases. Although the local maxima reach about the same values as we saw before, the typical values of the relative drift is generally smaller. Moreover, it is evident that there are multiple changes in the sign of Δs. It seems that NGC 3310 could be a bit more resistant to the breaking of the spiral distribution of clusters. However, there are no estimates of the width of the spiral arms of this galaxy and it is also worth mentioning that there is evidence of a recent merger (Miralles-Caballero et al., 2014), which could slightly complicate the situation.

Another interesting approach is to set the rotation curve to be constant. In such case, the only remaining term in **Eq. 2** is $\Omega_p \Delta R$. For spiral arms rotating as rigid bodies, this will be a constant, therefore $\frac{d(\Delta s)}{dR} = 0$ and the relative drift $\Delta s(t) \sim 10^t$ will grow exponentially as a function of time.

3.4 Pattern Breaking—Velocity Dispersion

The velocity dispersion of an aggregate of bodies (we are interested solely in the bodies within individual galaxies) plays a significant role in the kinematic evolution of that aggregate. The magnitude of the dispersion varies from one system type to another. For stars in open clusters, one typically finds values of the order of ~1 km/s (e.g., Kim et al., 2019, the value tends to increase with the age of the cluster). The velocity dispersion of young open clusters in their host galaxies reaches somewhat higher values (5–10 km/s in our Galaxy, Soubiran et al., 2018). More intermediate values are typically found in the stellar population of the disks of spiral galaxies—for our Galaxy, Gaia Collaboration et al. (2018c) found 10–50 km/s (different values for the radial, tangential, and vertical components). It is worth keeping in mind that the results are affected by the binary fraction, although the extent of this effect is currently unknown. Finally, an emphasis must be put of the fact that

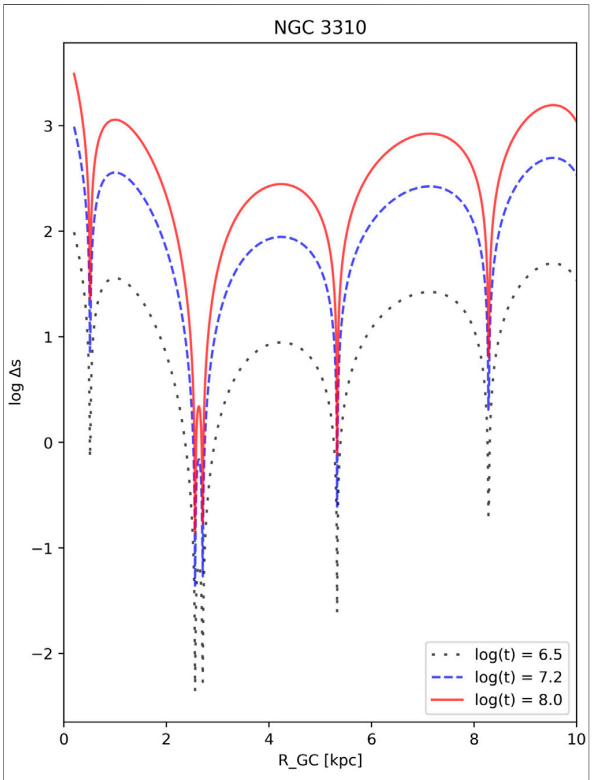


FIGURE 4 | The orbital drift of the open cluster in NGC 3310 from their host spiral arm as a function of the galacto-centric radius and time. The values of Δs under the logarithm are in parsec. Used parameters: Ω_p =14 km/s/kpc, ΔR = 0.1 kpc, rotation curve from Epinat et al. (2008), approximated by a high-degree polynomial (deg > 15).

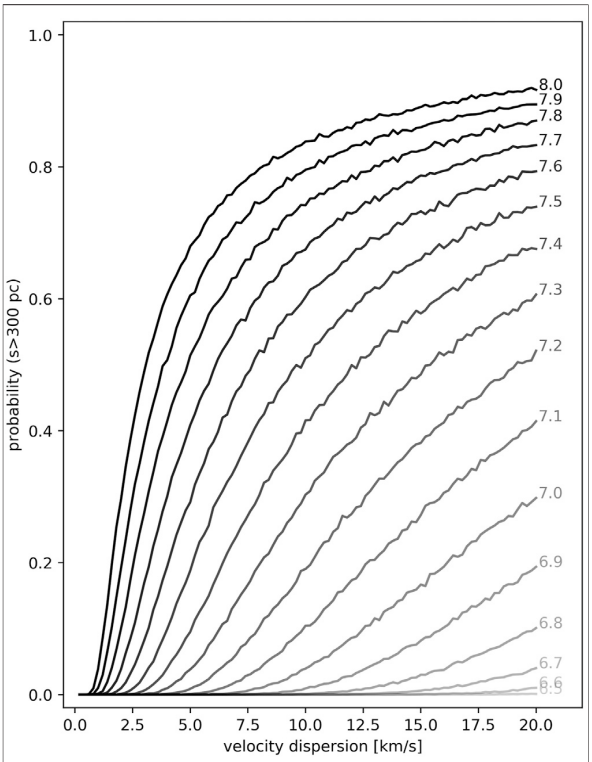


FIGURE 5 | The estimated probability, that a randomly located cluster in a galaxy drifts from the spiral arm (its birthplace), as a function of the velocity dispersion. The simulation was done for multiple cluster ages. The cut-off value for Δs was chosen to be 300 pc. The pattern velocity is Ω_p = 25 km/s/kpc.

the velocity dispersion is a time-dependent quantity—it slowly increases over time (Yu and Liu, 2018).

If we set a cluster to follow the spiral arm perfectly (the term in the brackets of **Eq. 1** is zero), the only remaining quantity is the velocity dispersion. In this example, we will be following the arm in its frame of reference and assume that the velocity dispersion is defined in the rotating frame. The random velocity vector is generated from a normal distribution, based on the dispersion, for each individual cluster—this vector is set to be constant. This is not physically acceptable and we shall return to this point in a moment. Right now, we would like to statistically examine, what is going to happen to the distribution of clusters if their orbital velocities randomly deviate from the pattern velocity.

For this, we can generate (for example) N = 20,000 values of Δs for a given value of σ_t and log(age). Afterward, we can calculate the ratio M/N, where M is the number of values for which the resulting drift $\Delta s > s_{lim}$. Parameter s_{lim} is set to be either 300 or 500 pc, which is comparable with the spiral arm widths in our Galaxy. Statistically, this should represent the probability that a cluster of given age is going to be able to drift beyond s_{lim}, assuming that the kinematics is dominated by the given value of σ_t. The results for $6.5 \leq \log(\text{age}) \leq 8.0$ are presented in **Figures 5**, **6**.

When assuming pure velocity dispersion (σ_t = 10.0 km/s) for the relative velocities of clusters compared to the motion of the spiral arm, the probability of a cluster drifting away from the arm becomes significant starting at ages between log(age) = 7.5 and log(age) = 8.0. We can see in **Figures 5**, **6** that the result depends strongly on the considered velocity dispersion value—this value is not very well constrained for the young clusters (Gieles et al., 2010).

The combination of the rotation curve velocity term (Milky Way) with the dispersion term is shown in **Figure 7**. While the velocity term of the rotation curve simply moves the clusters away from the spiral, the velocity dispersion destroys (over time) any information about the original distribution. This conclusion is comparable to what has been observed for M 51. It is clear in **Figure 1** from Chandar et al. (2017) that clusters up to about log(age) = 8.0 qualify as good tracers of the spiral arms. On the other hand, the observed older clusters in M 31 should not be good tracers of the spiral arms, assuming a velocity dispersion of $\sigma_t > 10$ km/s (Collins et al., 2011). This contradicts what has been indicated by observations—see **Section 3.3**. An observational bias

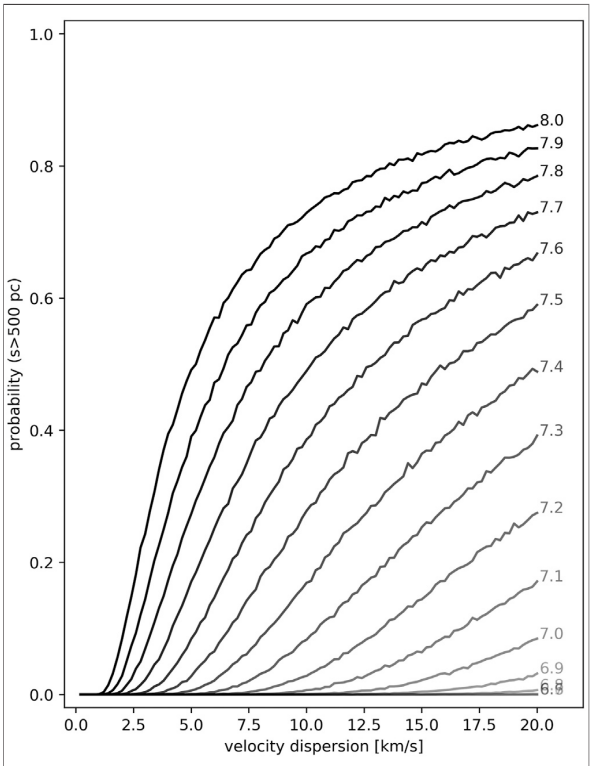


FIGURE 6 | The estimated probability, that a randomly located cluster in a galaxy drifts from the spiral arm (its birthplace), as a function of the velocity dispersion. The simulation was done for multiple cluster ages. The cut-off value for Δs was chosen to be 500 pc. The pattern velocity is Ω_p = 25 km/s/kpc.

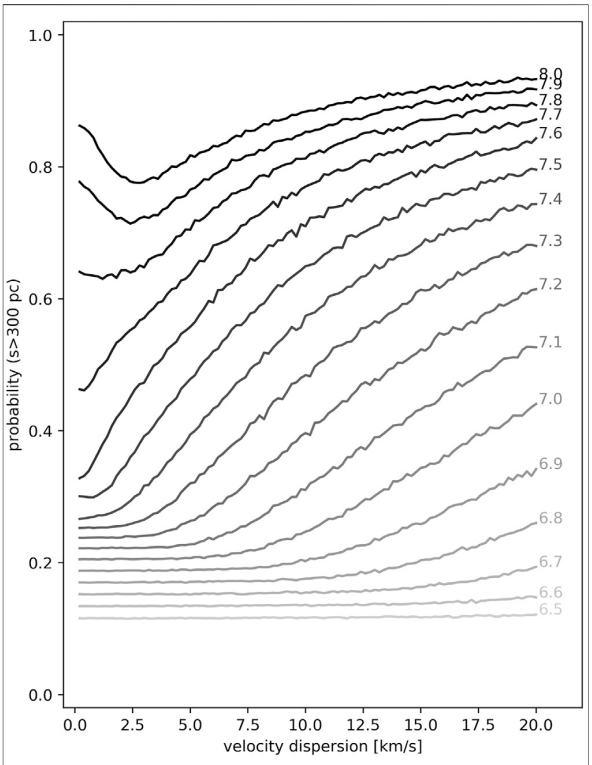


FIGURE 7 | The estimated probability, that a randomly located cluster in a galaxy drifts from the spiral arm (its birthplace), as a function of the velocity dispersion. The simulation was done for multiple cluster ages. Orbital motion based on the rotation curve of the Milky Way is also included. The cut-off value for Δs was chosen to be 300 pc. The pattern velocity is Ω_p = 25 km/s/kpc. The decrease in the probability for higher ages is due to the fact, that rotation is enough to destroy the spiral structure—the velocity dispersion then slightly increases the chance that some of the clusters will stay near the arm (when compared with the zero-dispersion).

is one possible explanation (not many clusters were used by Caldwell et al., 2009, Figure 10), but a more sophisticated model (Dobbs and Pringle, 2010) could also solve this problem.

The pattern of the spiral arm in the distribution of open clusters does not break precisely in the way that was described here—this approach was used solely to estimate the effect of a random deviation of the orbital velocities from the pattern velocity. The velocity vector of clusters (or field stars) cannot be considered to be constant since the orbits are bound. Instead, a different approach must be taken. For example, the epicyclic approximation can be used to describe non-circular orbits. If we were to think about the velocity dispersion in the X and Y coordinates, the approach introduced here would not work—orbits need to be bound, which means that fictitious forces have to appear in the equations of motion.

Hydrodynamic simulations serve as a more precise tool for studying this problem (although they can become more time-consuming). For example, Dobbs and Pringle (2010) showed that different excitation models for the spiral arms result into different age-distributions of star clusters. In combination with observations, such approach can be used to determine the possible excitation mechanisms for different galaxies.

3.5 Simulating Local Spiral Arm

As a next step, we would like to compare a simulation of the breaking of the spiral pattern of the Local Arm with observations. For this, we shall use all of the kinematic parameters of our Galaxy mentioned in **Section 3.1**, together with the Gaia data. In what follows, we outline the overall procedure.

1. Simulate the Local Spiral Arm. For this we have used the model from Ringermacher and Mead (2009), using the parameters $par_A = 4.1$, $par_B = 0.3$ and $par_N = 3.7$. This gives us a fairly good representation of the Local Arm when compared with the model from Reid et al. (2014).
2. Create clusters at random positions along the arm. In our simulation, we have arbitrarily chosen 7,000 clusters to be produced. The width of the arm is simulated by randomizing position in the rectangular coordinates, reaching standard deviation of ~300 pc at R~8 kpc. This scatter can be scaled with the Galacto-centric radius but the change would be quite small in the spatial region we are interested in. Most of the open clusters studied in Gaia data are located within 4 kpc from the Sun, and one has to keep in mind that the distance error scales with the distance to the cluster.
3. For each cluster, simulate its members (a random number between 100 and 300 stars) using its location and a normal distribution with the width of 10 pc (corresponding to an upper value of typical cluster sizes, van den Bergh, 2006). This resembles the actual positions of cluster members in our simulation (although the kinematics will be used just for the clusters themselves, not for the individual member stars).
4. A good estimate of the absolute parallax error in Gaia DR2 is about 0.3 mas. The procedure converts the true distances into true parallaxes and applies the uncertainty values. For each star, the value of parallax is generated using a normal distribution centered at the true parallax, with the widths corresponding to the uncertainty. This simulates the process of measuring the parallaxes of the individual stars.
5. We shall assume that the parallaxes of cluster members form a normal distribution—the true parallax can be fairly precisely estimated by finding the center of the ϖ distribution of the cluster members. The procedure estimates the cluster distances d by inverting the central parallaxes. This approach is much more robust than simply inverting all of the parallaxes and finding the center of that distribution (however, not as good as using a Bayesian approach). For a more comprehensive insight into the conversion of parallaxes into distance, we redirect the reader to the interesting work by Luri et al. (2018).
6. Polar Galactic coordinates (from the perspective of the observer) are transformed again into rectangular coordinates, assuming that the Sun is located at (−8.3, 0.0) kpc. An arbitrary Gaussian probability distribution function (using the observed distance as a variable) is applied to simulate the fact that the more distant clusters can be

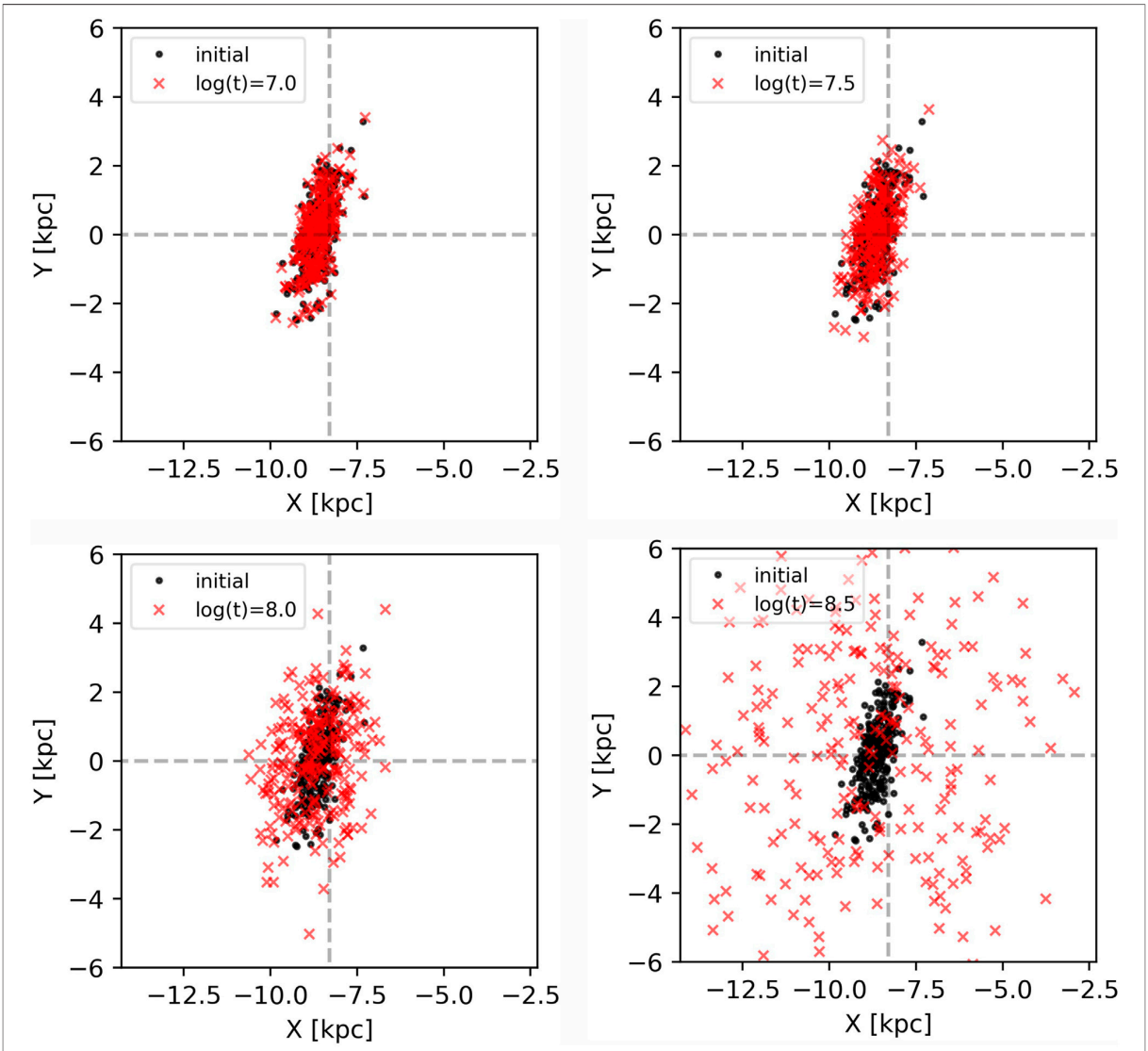


FIGURE 8 | Breaking of the spiral pattern in the distribution of open clusters due to the dispersion of velocities introduced in the rotating frame of reference. As we discussed in the text, this situation is not physical. However, it still provides us with statistical information—even relatively small deviations from the orbital velocity (~10 km/s) of the spiral arm would cause a breaking of the pattern.

missed. The cut-off between the revealed and the hidden clusters is set to be about 2 kpc away from the observer. We have to point out that only clusters along the spiral arm are simulated—we should still be able to recognize the spiral pattern in the distribution of clusters at this point.

7. A velocity vector is generated to each of the revealed clusters. It consists of the orbital velocity generated from the rotation curve of the Milky Way, and the x and y components of the velocity dispersion. Motion of each cluster can be described by the epicylic approximation. It is also worth mentioning that the uncertainty in the measurement of radial velocities should play a role similar to the one played by the velocity dispersion.
8. The clusters are allowed to move for a period of time A, after which the distribution of the clusters is checked again.

We shall assume that the kinematic effect of the tidal dissipation of clusters can be neglected, and clusters are allowed to follow orbits resulting from the epicylic approximation. Standard notation is implemented.

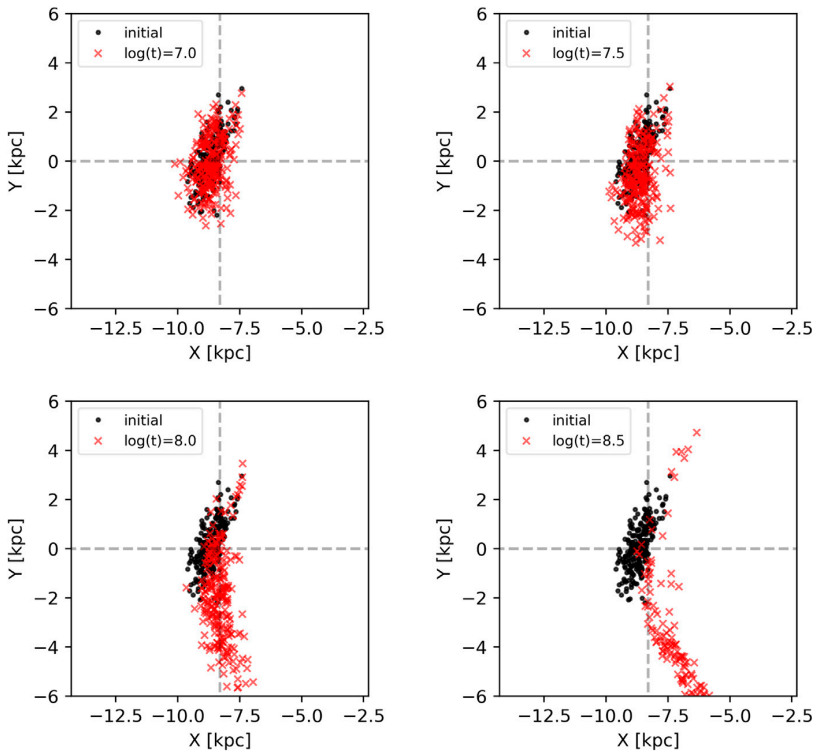


FIGURE 9 | Breaking of the spiral pattern in the distribution of open clusters due to the difference between the orbital motion of the clusters and the pattern velocity of the rotating spiral arm. Velocity dispersion is introduced in the non-rotating frame of reference and describes the epicyclic motion. If the difference between the orbital velocity and the pattern velocity was zero, the cluster would not deviate (in this simple approximation) far from the spiral arm, it would rather oscillate around a point within the arm. The pattern breaking is dominated by the rotation curve—most of the clusters in the Local Arm are located beyond the corotation radius of this model, which results in the clusters trailing behind the spiral arm.

$$\begin{aligned} x(t) &= x_0 \sin(\kappa t + \psi) \\ y(t) &= y_0 \cos(\kappa t + \psi) \end{aligned} \tag{3}$$

where x is oriented in the radial direction and y is oriented in the tangential direction of the circular part of the motion, x_0 and y_0 are the amplitudes of the epicyclic motion, κ is the epicyclic frequency, and ψ is the initial phase of the epicyclic motion. We have used the velocities generated by the velocity dispersion as the initial values of $\dot{x}$ and $\dot{y}$. If we calculate the time-derivatives of **Eq. 3**, we can eliminated ψ from the equations

$$\frac{\dot{x}^2}{x_0^2} + \frac{\dot{y}^2}{y_0^2} = \kappa^2 \tag{4}$$

and κ can be found from the relation $\left[R\frac{d\Omega^2}{dR} + 4\Omega^2\right]$ evaluated at the radius of the circular part of the motion, R_0. Moreover, x_0 and y_0 are linked by another relation

$$\frac{x_0}{y_0} = \frac{\kappa(R_0)}{2\Omega(R_0)} \tag{5}$$

We can combine **Eqs. 4, Eqs. 5** to determine the unknown values of the amplitudes, and then return to the relations for $\dot{x}$ and $\dot{y}$ to find

$$\tan(\psi) = \frac{\dot{y}/y_0}{\dot{x}/x_0} \tag{6}$$

The motion in the polar Galacto-centric coordinate system is then described by the following two equations

$$\begin{aligned} \phi(t) &= \phi_0 + \Omega t - \frac{y_0}{R_0}\sin(\kappa t + \psi) \\ R(t) &= R_0 + x_0 \sin(\kappa t + \psi) \end{aligned} \tag{7}$$

which can be easily transformed into the rectangular Galacto-centric coordinates. Finally, the spiral arm is also rotating in the Galaxy. To characterize the motion of the clusters with respect to the motion of the spiral arm, the positions of the individual clusters must be rotated by the angle $\alpha = -\Omega_p t$. We emphasize that the pattern velocity Ω_p is taken to be a constant.

We have simulated two situations—using only velocity dispersion, and assuming the full motion with the use of the epicyclic approximation. The resulting distributions are presented in **Figures 8**, **9** respectively. Pattern breaking, in the case of the epicyclic approximation, is dominated by the rotation curve. On the other hand, the velocity dispersion could not break the spiral pattern on its own—this is the result of the assumption of bound orbits

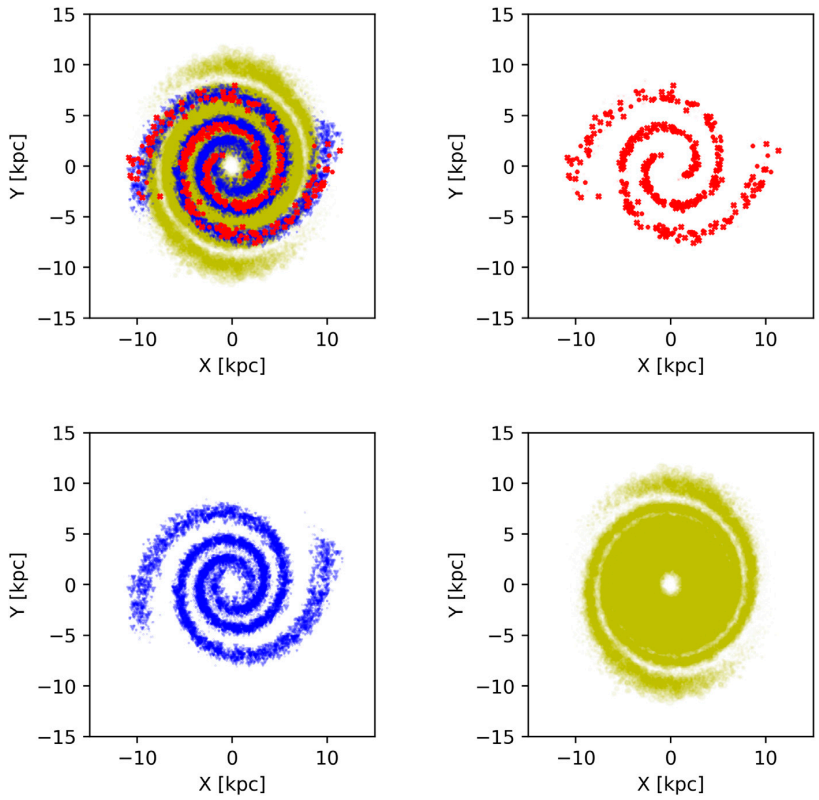


FIGURE 10 | Maps of the cluster distribution estimated for a kinematic model of our Galaxy. The clusters were separated into three groups: G6.5, G7.5, and G8.5. **(A)**: the composite picture of the distribution; red crosses and blue triangles follow the spiral arm quite closely. **(B)**: clusters with ages around 3.1 Myr. **(C)**: clusters with ages around 31 Myr. **(D)**: clusters with ages around 310 Myr.

(which is not considered in the pure-dispersion model). In reality, velocity dispersion vector increases with time, which would enhance the pattern breaking. Therefore, $\dot{\sigma}$ would be the driving mechanism for the spiral pattern breaking in terms of the dispersion.

It is worth mentioning that the uncertainties in the observed distances played only a minor role in the simulation (unless clusters with distances $d > 5.0$ kpc are taken into account). Much more dominant would be the uncertainties in the measurement of the velocity vector, if they were taken into account (we may assume that the uncertainties are hidden in the velocity dispersion term).

Clusters younger than $\log(\text{age}) \sim 7.0$ follow the spiral arms very closely, as was expected. These objects represent a good tool for tracing spiral arms, even in the dynamically warmer galaxies. Unfortunately, most of such objects were missed in Cantat-Gaudin et al. (2018) who used primarily clusters with $\log(\text{age}) > 7.0$. Nevertheless, their distribution of the clusters younger than $\log(\text{age}) \sim 8.0$ seems to represent the position of the spiral arms quite well, as was highlighted by the authors.

The main goal of this simulation was to study the situation after $\log(t) \sim 8.0$. Clearly, most of the clusters have drifted away from the spiral arm in both simulated cases. Although the simulation in **Figure 8** is quite non-physical, it still provides a simple look at what happens with the distribution if only velocity dispersion is taken into account—the structure slowly dissipates into the surrounding area. The more realistic case is the assumption of nearly circular orbits (epicyclic approximation, **Figure 9**). Here we see that the distribution of the clusters changes from a spiral structure into a ring-like distribution, corresponding to a randomized position of the clusters within the Galactic disk.

3.6 Spiral Pattern in the Distribution of Open Clusters

We have seen that the distribution of open clusters changes based on their ages. However, when we observe clusters, the resulting data set almost always contains objects representing a mixture of ages. In principle, we can determine the cluster ages from the photometric data (isochrone fitting) but even if we ignored uncertainties, the resulting subsets of clusters would contain only a small number of objects. This is especially true for Cantat-Gaudin et al. (2018), where we see a lack of the younger clusters due to an observational bias. A small number of objects means that whatever statistical analysis we want to perform (for example, young

clusters tracing a spiral arm), the results will turn out to be quite unreliable.

Fortunately, we can make use of everything we studied in the previous sections. We know that clusters younger than log(age) ~ 8.0 should follow the spiral arm they were born in quite tightly. On the other hand, older clusters should represent a background population. But how to display this situation? If we were to show all of the clusters, the map of their distribution would very messy. This results from the fact that a lower number of clusters are being born every 1 Myr than there are clusters older than them.

We find that breaking the distribution of clusters into three groups consisting of objects with ages around 3.1 (G6.5), 31 (G7.5) and 310 Myr (G8.5) works quite well. These groups are represented in **Figure 10** by different symbols and colours. The assumed kinematics are the same as in the previous section (epicyclic approximation, Ω_p = 25 km/s/kpc, rotation curve of the Milky Way, the same spiral arm as described above), but there are two spiral arms (the second one is created by rotating the original one by 180° around the Galactic center). The simulation starts at t = 0 Myr and continues with a time-step Δt = 2 Myr. At the beginning of each step, 100 clusters are born randomly across the whole arm, and the simulation stops at t = 1000 Myr. At this point, the distribution should be stable, and the positions of the clusters are plotted into **Figure 10**.

In the subplots, the data points are scaled using a Gaussian distribution centered at the age of the given group, with a width of 20% of this value. This means that if we were to follow one of the clusters in time, it would:

- First, appear at an age of about 6.5 as a red cross, then it would very quickly fade away
- It becomes visible again as a blue triangle at an age of about 7.3, afterwards fades at about 7.6
- Lastly, the cluster appears as a yellow circle at 8.4 for a significantly longer amount of time and disappears at an age of 8.6

The third group consists mostly of clusters just below the upper bound of the life-span of a typical open cluster in our Galaxy (~1 Gyr, Spitzer, 1958).

The distribution of clusters in **Figure 10** looks just as expected. The younger clusters from the groups G6.5 and G7.5 closely follow the spiral arms, with only a small angular lag between them (depends of the Galacto-centric radius, the lag is close to zero near the corotation). The group G8.5 shows that the distribution of clusters in the inner region of our Galaxy randomizes while the outer region should retain its spiral structure under our kinematic assumptions. However, this older distribution lags significantly behind the actual position of the spiral arm.

4 DIFFUSE INTERSTELLAR BANDS AS TRACERS OF SPIRAL ARMS

Diffuse interstellar bands (DIBs) is a designation of a several 100 mysterious absorption features, most likely originating from the interstellar medium. They appear in the lines of sight toward stars of different spectral types, although they are most prominent in the spectra of reddened hot stars. Most of the DIBs are quite narrow (FWHM < 0 nm), with only several exceptions (e.g., the strong bands at 443.0, 578.0, and 628.4 nm). These bands have been observed not only within Milky Way but also in other galaxies:

- Magellanic Clouds (Ehrenfreund et al., 2002)
- M 31 (Cordiner et al., 2008a)
- M 33 (Cordiner et al., 2008b)
- NGC 1448 (Sollerman et al., 2005)
- NGC 4038/NGC 4039 (Monreal-Ibero et al., 2018)
- NGC 1614, NGC 1808, NGC 2146, NGC 3256, NGC 6240, M 82, and IRAS 10565 + 2448 (Heckman and Lehnert, 2000)

Although many DIBs are located near regions contaminated by other lines/bands (for example, the overlap of the telluric band with the DIB at 628.4 nm), some remain easily resolved in the most of the spectra (e.g., 661.4 nm band). Given their characteristically narrow profiles, these bands could be quite useful for studying the interstellar medium of galaxies, depending on the spectral region, resolution and on the choice of the DIBs.

One of the common properties of the individual DIBs is that they are somehow correlated with the interstellar reddening (Merrill and Wilson, 1938). Although this correlation is often described as linear, the relation between the reddening [typically $E(B-V)$] and the strength of the DIBs (usually described by their equivalent widths, EWs) is far more complicated. For start, the distribution of the points in the reddening-EW diagrams is usually quite broad. Moreover, if we just use only one DIB, we will find that the correlation differs between different lines of sight in our Galaxy (see, for example, Piecka and Paunzen, 2020).

Some of the DIBs can be shown to posses quite complicated profiles using high-resolution spectroscopy (ideally R>100,000). We mention two examples, the bands at 579.7 and 661.4 nm, the profiles of which differ in many ways. Curiously, the origin of these structures is not yet well understood. According to Cami et al. (2004), the most likely explanation is that 1) the structure is the result of unresolved ro-vibrational structure of the molecules, or 2) the isotope effect dominates in the structure. The latter option is typically dismissed due to the observed variations in the relative positions of the individual peaks in the bands profile.

4.1 Origin of the Diffuse Interstellar Bands

Discovered about 100 years ago (Heger, 1922), vast majority of the DIBs remain unidentified to this date. Only quite recently, several bands were attributed to the molecular ion C_{60}^+ by Campbell et al. (2015). This assignment was later supported by the observations (Cordiner et al., 2019). There are many other molecules which were considered as carriers of some of the DIBs, for example: different polycyclic aromatic hydrocarbons (PAHs, e.g., see Salama et al., 2011), carbon chains (Maier et al., 2004), propadienylidene l-C_3H_2 (Maier

et al., 2011). However, these DIBs were found to be located outside the regions of bands of the corresponding molecules. The verification is based on the high-precision laboratory spectroscopy.

As was mentioned above, correlation between DIBs and the reddening varies across the sky (and most likely even in the radial direction). Not all of the DIBs are always present in a given line of sight and the EW-ratios of a pair of DIBs are usually inconsistent, except for the well correlated pair at 619.6 and 661.4 nm (Krełowski et al., 2016). Together with the variations in the profiles of DIBs, this would suggest that the number of carriers depends heavily on the conditions within the probed medium. Moreover, there have to be multiple different species in order to explain the observed EW-ratios. Jenniskens et al. (1994) showed that one of the most important properties of the medium surrounding the DIB carriers is the intensity of the UV radiation field. Generally, carriers seem to be more abundant in the regions with high intensity UV fields (such as the Orion Nebula, M 42, mentioned in the cited paper) than in the dense molecular regions. However, this is not a strict rule—Jenniskens et al. (1994) pointed out that the 628.4 nm DIB behaves differently than described.

One topic which remains seriously unexplored is the formation of the carriers of DIBs. Presently, there are no specific theories about the processes leading to the creation of the DIB-carriers in the interstellar medium (ISM). Very important questions are still left unanswered:

- Are these species formed *in situ*, or do they exist for an extended amount of time (albeit in several different states)?
- Do these species form during the destruction/formation processes of interstellar dust grains, or is their creation unrelated to dust?
- Can these species be born in stellar outflows?

We can learn much from the observed spectra. For example, studies of the planetary nebulae have shown that fullerenes (such as C_{60}) can be found in circum-stellar medium or in stellar outflows (García-Hernández and Díaz-Luis, 2013). Such relatively simple and very stable molecules are thought to be the product of larger molecules being broken down (Berné et al., 2015). The reason why C_{60} appears as the most abundant of the fullerenes is its stability. If we accepted this model (top-down formation) to be the most realistic, we could study the details of the formation processes. However, this is only usable for this single molecule as (at least) dozens of other carriers remain unidentified. Can the story behind the other molecules be the same as for the fullerenes?

The interstellar environment puts strict physical constraints on the structure of the carriers—they are most likely moderately sized organic or carbonaceous molecules, $N(C) > 20$, with highly symmetric structures providing resistance to the UV photo-dissociation. Recently, Omont et al. (2019) brought into attention a subclass of PAHs, the polyacenes, which easily fulfill all of the required properties. Unfortunately, high-quality laboratory spectra are only available for a small number of these molecular species.

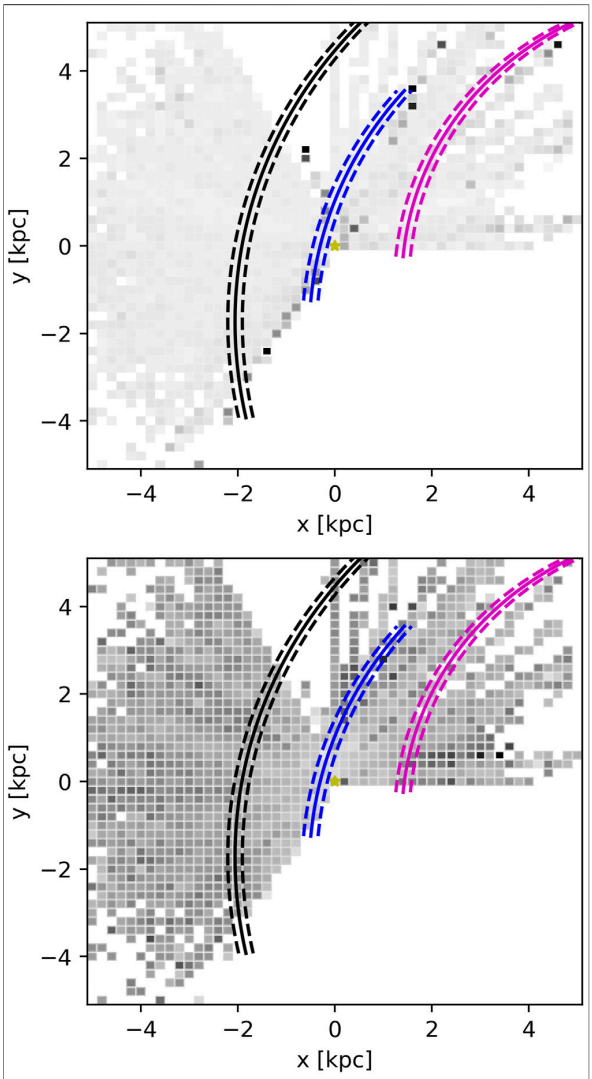


FIGURE 11 | The maps of the DIB at 1.5273 μm, based on the data from Elyajouri and Lallement (2019). The upper plot displays the distribution of the ratios EW/A_V, the lower plot shows the distribution of EWs. The darker the cell, the higher the value. The coordinates are centered at the position of the Sun (not the Galacto-centric coordinates). The nearby spiral arms from Reid et al. (2014) are displayed with different colours (black—Perseus Arm, blue—Local Arm, magenta—Sagittarius Arm); the corresponding dotted lines emphasize their widths.

4.2 Mapping the Interstellar Medium

It might seem that these bands could, in principle, be quite good tracers of the spiral arms. This is due to the much larger star-formation rates in the arms than in the space between them. Such regions can become rich in the nebulae such as M 42, which should produce larger absorption in the DIBs. However, we should keep in mind the following: although a given DIB is a good tracer of a specific set of conditions, the specifics of these conditions are poorly understood.

An interesting map of the infrared DIB at 1.5273 μm was presented by Zasowski et al. (2015). Their work is based on the spectra of ~60,000 stars from SDSS/APOGEE. Unfortunately, no spiral arm is visible in their map.

Elyajouri and Lallement (2019) presented over 100,000 measurements of the 1.5273 μm DIB (also SDSS/APOGEE). The lines of sight cover much of the region between $l = 0^\circ$ and $l = 270^\circ$. We wanted to use these results to construct maps similar to those from Zasowski et al. (2015). When mapping diffuse interstellar bands, it can be rewarding to also look at EW/A_V, the ratio of the EW and the extinction value in that direction (authors adopted values from Starhorse database). In **Figure 11**, we show the maps based on the results from Elyajouri and Lallement (2019). The ratios are shown in the upper plot, the EWs in the lower plot. Distances were calculated by inverting parallaxes ϖ from Gaia DR2, restricting ourselves to $|z| < 100$ pc. We would like to point out that the parallax inversion may lead to distortions in **Figure 11** at distances (from the center) larger than about ~1 kpc. To improve the situation, we excluded all lines of sight for which the stellar parallaxes from Gaia DR2 have relative errors $e_\varpi/\varpi > 0.20$. The maps were created by binning targets in cells (200 × 200 pc) and using the median of the mapped quantity in for each cell. Finally, the intensity of the cells reflects the magnitude of the quantity—larger value means darker cell. For the ratios, we have introduced a cut-off value $EW/A_V = 1$, meaning that all cells with larger values than this limit are equal in the map (darkest black). The intensity then drops and is equal to the ratio value. For the EWs map, the intensities are scaled by the maximum value of median (EW), which in our case is 491.92 mÅ.

The results are comparable to what Zasowski et al. (2015) showed. In the near vicinity of the Sun, the use of the DIBs for tracing Galactic arms is fairly limited. The biggest problem with the maps of EWs is that their values are always integrated in the line of sight. There is a possible solution to this problem. If we have two targets which are very close to each other on the sky but distant in the parallax, we could subtract the EW of the closer target from the more distant one, leaving potential only the amount of absorption from the region beyond the closer target. However, there are several difficulties. Firstly, the observed quantities (EWs, parallaxes) are accompanied by uncertainties, which may affect the result. Secondly, we have no prior knowledge about the variation of the EWs across the sky in the observed area around the target stars—these variations may have a large impact on the suggested subtraction procedure. Finally, if we look at the distance-EW diagrams, we shall see a significant scatter. The EWs do not seem to follow any curve in such plots, instead they are somehow distributed from zero up to an upper limit value, which depends on the distance. We can either attempt to extract these upper limits per distance, or median values, but neither of those provides a precise measure which is required for our procedure.

The ratios EW/A_V can be somewhat more useful when looking at the spatial maps of DIBs. The idea is to use the mentioned infrared DIB, because of its tight correlation with the extinction (Zasowski et al., 2015). Indeed, the darkest cells in the upper plot of **Figure 11** seem to follow the Local Arm. However, this may be a result of some unknown bias. However, this approach can be a bit problematic. This is due to the introduction of the division of by A_V (which is also accompanied by an uncertainty).

In **Figure 2**, Monreal-Ibero et al. (2018) presented maps of the 578.0 and 579.7 nm DIBs in the Antennae Galaxies (NGC 4038/NGC 4039). Their maps cover a large part of the northern spiral arm. Moreover, the carriers of the studied DIBs seem to be fairly abundant in some regions of the arm. Regardless, a closer inspection of the map of 578.0 nm DIB shows that the EW varies across the arm—the arm passes through regions with both, relatively strong and weak absorption in the DIBs. The situation is the same outside of the arm. In this case, the chosen band appears as a poor choice of a tracer of the spiral structure of these galaxies. The DIB at 579.7 nm covers somewhat smaller regions than the one at 578.0 nm. Its map shows that the strongest absorption is located within the region where the strongest optical absorption is observed. Again, no spiral structure can be distinguished in the map.

To our knowledge, no other detailed maps of other galaxies have been published. Theoretically, the ability of DIBs to probe interstellar medium with a narrow range of possible properties can be very handy. The practicality of using these tools is limited by the spectroscopic resolution, the signal to noise ratio (which becomes a very important factor when studying distant galaxies), and finally the ability to derive the parameters of the band-profiles. It remains to be seen whether DIBs will become useful as tracers of the galactic structure.

5 CONCLUSION

We have based our simulations of motion of open clusters on some of the newest data from Gaia (velocity dispersion) and on the rotation curve of the Milky Way. We have constructed a simple mathematical apparatus for the analysis of the distribution of open clusters when compared with the spiral arm they originate from. As expected, the older clusters seem to move away from the arm—this makes such clusters poor tracers of the spiral structure of galaxies.

Although the investigation of spiral arms was not the main contribution of the work Cantat-Gaudin et al. (2018), they have used older clusters for demonstrating the possibility of using Gaia for this purposes. We should remind ourselves that such results must be viewed upon with care—only the young open clusters (younger than about 100 Myr, but even younger should be preferred) can be used to get reliable results. It is worth mentioning that the number of clusters younger than about 30 Myr is very low in the cited work. Despite this, the nearby open clusters are found near the relevant spiral arms near the Sun.

Figure 8 and **Figure 9** display how the distribution of open clusters breaks away from the spiral structure after some amount of time. The former shows how this would happen, if the clusters had moved away from the arm with random velocities—although an non-physical solution, this demonstrates how the difference between the motion of physical objects of the host galaxy and the motion of the spiral arm must result in pattern breaking. The second plot is based on the epicylic approximation of the cluster orbits. After some time (>300 Myr), the distribution turns from being a spiral into a ring-like structure.

In **Section 3.6**, we have demonstrated purely kinematic simulation of open clusters being born in the spiral arms. Dynamical processes could further enhance the effect of pattern breaking in the distribution of open clusters. The model of the simulated galaxy is close to our present view of the Milky Way. We can see that, at some point, the distribution of the older clusters winds up.

Although greatly inferior to the N-body and hydrodynamic simulations, we have shown that the described approach proves to be useful for taking an initial look at the observed data. It makes it possible to analyze multiple different galaxies in a short time. However, more sophisticated methods should be preferred when attempting to study the problem at hand in greater detail.

Finally, we have studied the map of the DIB at 1.5273 μm (**Figure 11**). There does not seem to be any clear spiral structure present in the map, although something appear in the vicinity of the Local Arm—it is difficult to distinguish whether this is a result of a bias (possibly caused by the inversion of parallaxes). The lack of a larger number of measurements in the other DIBs prevents us from analyzing the map at different wavelengths. Based on this result and comparing with maps of other galaxies in the literature, we conclude that DIBs appear to be ineffective when tracing spiral arms. However, we do not rule out the possibility that a different approach or additional data could change this status.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MP: compilation of the literature and available data, programming and implementation of the simulation, interpretation of the results, generating of the figures, analysis of the DIBs, writing the article. EP: compilation of the literature and available data, interpretation of the results, writing the article.

FUNDING

This work was supported by MUNI/A/1482/2019 and MUNI/A/1206/2020, Masaryk University, Faculty of science.

ACKNOWLEDGMENTS

This research has made use of the WEBDA database, operated at the Department of Theoretical Physics and Astrophysics of the Masaryk University, the SIMBAD database, operated at CDS, Strasbourg, France and NASA's Astrophysics Data System. This work presents results from the European Space Agency (ESA) space mission Gaia. Gaia data are being processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia MultiLateral Agreement (MLA). The Gaia mission website is https://www.cosmos.esa.int/gaia. The Gaia archive website is https://archives.esac.esa.int/gaia. The rotation curves were based on a polynomial fit to the data from http://www.ioa.s.u-tokyo.ac.jp/~sofue/RC99/rc99.htm.

REFERENCES

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., Muraveva, T., et al. (2018). Gaia Data Release 2. *A&A* 616, A17. doi:10.1051/0004-6361/201833234

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., Robin, A. C., et al. (2017). Gaia Data Release 1. *A&A* 599, A50. doi:10.1051/0004-6361/201629895

Balaguer-Núñez, L., López del Fresno, M., Solano, E., Galadí-Enríquez, D., Jordi, C., Jimenez-Esteban, F., et al. (2020). Clusterix 2.0: a Virtual Observatory Tool to Estimate Cluster Membership Probability. *MNRAS* 492, 5811–5843. doi:10.1093/mnras/stz3610

Baratella, M., D'Orazi, V., Carraro, G., Desidera, S., Randich, S., Magrini, L., et al. (2020). The Gaia-ESO Survey: a New Approach to Chemically Characterising Young Open Clusters. *A&A* 634, A34. doi:10.1051/0004-6361/201937055

Bastian, U., and Röser, S. (1993). *PPM Star Catalogue. Positions and Proper Motions of 197179 Stars South of -2.5 Degrees Declination for Equinox and Epoch J2000.* Heidelberg, Germany: Spektrum Akademischer Verlag.

Baumgardt, H., and Kroupa, P. (2007). A Comprehensive Set of Simulations Studying the Influence of Gas Expulsion on star Cluster Evolution. *MNRAS* 380, 1589–1598. doi:10.1111/j.1365-2966.2007.12209.x

Bekki, K., Koribalski, B. S., Ryder, S. D., and Couch, W. J. (2005). Massive H I Clouds with No Optical Counterparts as High-Density Regions of Intragroup H I Rings and Arcs. *Monthly Notices R. Astronomical Soc. Lett.* 357, L21–L25. doi:10.1111/j.1745-3933.2005.08625.x

Berné, O., Montillaud, J., and Joblin, C. (2015). Top-down Formation of Fullerenes in the Interstellar Medium. *A&A* 577, A133. doi:10.1051/0004-6361/201425338

Beuther, H., Walsh, A., Schilke, P., Sridharan, T. K., Menten, K. M., and Wyrowski, F. (2002). CH3OH and H2O Masers in High-Mass star-forming Regions. *A&A* 390, 289–298. doi:10.1051/0004-6361:20020710

Binney, J., and Tremaine, S. (2008). James Binney and Scott Tremaine. ISBN 978-0-691-13026-2 (HB). *Galactic Dynamics*. Second Edition, Princeton, NJ: Princeton University Press. doi:10.1515/9781400828722

Bobylev, V. V., Bajkova, A. T., Rastorguev, A. S., and Zabolotskikh, M. V. (2021). Analysis of Galaxy Kinematics Based on Cepheids from the Gaia DR2 Catalogue. *MNRAS* 502, 4377–4391. doi:10.1093/mnras/stab074

Borissova, J., Kurtev, R., Amarinho, N., Alonso-García, J., Ramírez Alegría, S., Bernal, S., et al. (2020). Small-scale star Formation as Revealed by VVVX Galactic Cluster Candidates. *MNRAS* 499, 3522–3533. doi:10.1093/mnras/staa3045

Bossini, D., Vallenari, A., Bragaglia, A., Cantat-Gaudin, T., Sordo, R., Balaguer-Núñez, L., et al. (2019). Age Determination for 269 Gaia DR2 Open Clusters. *A&A* 623, A108. doi:10.1051/0004-6361/201834693

Bouwman, J., Castellanos, P., Bulak, M., Terwisscha van Scheltinga, J., Cami, J., Linnartz, H., et al. (2019). Effect of Molecular Structure on the Infrared Signatures of Astronomically Relevant PAHs. *A&A* 621, A80. doi:10.1051/0004-6361/201834130

Bragaglia, A., and Tosi, M. (2006). The Bologna Open Cluster Chemical Evolution Project: Midterm Results from the Photometric Sample. *Astron. J.* 131, 1544–1558. doi:10.1086/499537

Braun, R. (1991). The Distribution and Kinematics of Neutral Gas in M31. *ApJ* 372, 54. doi:10.1086/169954

Caldwell, N., Harding, P., Morrison, H., Rose, J. A., Schiavon, R., and Kriessler, J. (2009). Star Clusters in M31. I. A Catalog and a Study of the Young Clusters. *Astronomical J.* 137, 94–110. doi:10.1088/0004-6256/137/1/94

Cami, J., Salama, F., Jiménez-Vicente, J., Galazutdinov, G. A., and Krełowski, J. (2004). The Rotational Excitation Temperature of the λ6614 Diffuse Interstellar Band Carrier. *ApJ* 611, L113–L116. doi:10.1086/423991

Campbell, E. K., Holz, M., Gerlich, D., and Maier, J. P. (2015). Laboratory Confirmation of C60+ as the Carrier of Two Diffuse Interstellar Bands. *Nature* 523, 322–323. doi:10.1038/nature14566

Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., Jordi, C., Romero-Gómez, M., Soubiran, C., et al. (2020). Painting a Portrait of the Galactic Disc with its Stellar Clusters. *A&A* 640, A1. doi:10.1051/0004-6361/202038192

Cantat-Gaudin, T., Jordi, C., Vallenari, A., Bragaglia, A., Balaguer-Núñez, L., Soubiran, C., et al. (2018). A Gaia DR2 View of the Open Cluster Population in the Milky Way. *A&A* 618, A93. doi:10.1051/0004-6361/201833476

Castro-Ginard, A., Jordi, C., Luri, X., Álvarez Cid-Fuentes, J., Casamiquela, L., Anders, F., et al. (2020). Hunting for Open Clusters in Gaia DR2: 582 New Open Clusters in the Galactic Disc. *A&A* 635, A45. doi:10.1051/0004-6361/201937386

Castro-Ginard, A., Jordi, C., Luri, X., Cantat-Gaudin, T., and Balaguer-Núñez, L. (2019). Hunting for Open Clusters in Gaia DR2: the Galactic Anticentre. *A&A* 627, A35. doi:10.1051/0004-6361/201935531

Castro-Ginard, A., Jordi, C., Luri, X., Julbe, F., Morvan, M., Balaguer-Núñez, L., et al. (2018). A New Method for Unveiling Open Clusters in Gaia. *A&A* 618, A59. doi:10.1051/0004-6361/201833390

Chandar, R., Chien, L.-H., Meidt, S., Querejeta, M., Dobbs, C., Schinnerer, E., et al. (2017). Clues to the Formation of Spiral Structure in M51 from the Ages and Locations of Star Clusters. *ApJ* 845, 78. doi:10.3847/1538-4357/aa7b38

Chrobáková, Ż., López-Corredoira, M., Sylos Labini, F., Wang, H.-F., and Nagy, R. (2020). Gaia-DR2 Extended Kinematical Maps. *A&A* 642, A95. doi:10.1051/0004-6361/202038736

Collins, M. L. M., Chapman, S. C., Ibata, R. A., Irwin, M. J., Rich, R. M., Ferguson, A. M. N., et al. (2011). The Kinematic Identification of a Thick Stellar Disc in M31★†. *MNRAS* 413, 1548–1568. doi:10.1111/j.1365-2966.2011.18238.x

Conti, P. S., and Crowther, P. A. (2004). MSX Mid-infrared Imaging of Massive star Birth Environments - II. Giant H Ii Regions. *MNRAS* 355, 899–917. doi:10.1111/j.1365-2966.2004.08367.x

Cordiner, M. A., Cox, N. L. J., Trundle, C., Evans, C. J., Hunter, I., Przybilla, N., et al. (2008a). Detection of Diffuse Interstellar Bands in M 31. *A&A* 480, L13–L1. doi:10.1051/0004-6361:20079309

Cordiner, M. A., Linnartz, H., Cox, N. L. J., Cami, J., Najarro, F., Proffitt, C. R., et al. (2019). Confirming Interstellar C60 + Using the Hubble Space Telescope. *ApJ* 875, L28. doi:10.3847/2041-8213/ab14e5

Cordiner, M. A., Smith, K. T., Cox, N. L. J., Evans, C. J., Hunter, I., Przybilla, N., et al. (2008b). Diffuse Interstellar Bands in M 33. *A&A* 492, L5–L8. doi:10.1051/0004-6361:200810906

Dias, W. S., Alessi, B. S., Moitinho, A., and Lépine, J. R. D. (2002). New Catalogue of Optically Visible Open Clusters and Candidates. *A&A* 389, 871–873. doi:10.1051/0004-6361:20020668

Dias, W. S., and Lépine, J. R. D. (2005). Direct Determination of the Spiral Pattern Rotation Speed of the Galaxy. *ApJ* 629, 825–831. doi:10.1086/431456

Dias, W. S., Monteiro, H., Lépine, J. R. D., and Barros, D. A. (2019). The Spiral Pattern Rotation Speed of the Galaxy and the Corotation Radius with Gaia DR2. *MNRAS* 486, 5726–5736. doi:10.1093/mnras/stz1196

Dias, W. S., Monteiro, H., Moitinho, A., Lépine, J. R. D., Carraro, G., Paunzen, E., et al. (2021). Updated Parameters of 1743 Open Clusters Based on Gaia DR2. *MNRAS* 504, 356–371. doi:10.1093/mnras/stab770

Díaz-García, S., Salo, H., Knapen, J. H., and Herrera-Endoqui, M. (2019). The Shapes of Spiral Arms in the S4G Survey and Their Connection with Stellar Bars. *A&A* 631, A94. doi:10.1051/0004-6361/201936000

Dobbs, C. L., and Pringle, J. E. (2010). Age Distributions of star Clusters in Spiral and Barred Galaxies as a Test for Theories of Spiral Structure. *MNRAS* 409, 396–404. doi:10.1111/j.1365-2966.2010.17323.x

Ehrenfreund, P., Cami, J., Jiménez-Vicente, J., Foing, B. H., Kaper, L., van der Meer, A., et al. (2002). Detection of Diffuse Interstellar Bands in the Magellanic Clouds. *ApJ* 576, L117–L120. doi:10.1086/343731

Elmegreen, B. G., and Hunter, D. A. (2010). On the Disruption of Star Clusters in a Hierarchical Interstellar Medium. *ApJ* 712, 604–623. doi:10.1088/0004-637X/712/1/604

Elyajouri, M., and Lallement, R. (2019). Updated Extraction of the APOGEE 1.5273 μm Diffuse Interstellar Band: a Planck View on the Carrier Depletion in Dense Cores. *A&A* 628, A67. doi:10.1051/0004-6361/201834452

Epinat, B., Amram, P., Marcelin, M., Balkowski, C., Daigle, O., Hernandez, O., et al. (2008). GHASP: an Hα Kinematic Survey of Spiral and Irregular Galaxies - VI. New Hα Data Cubes for 108 Galaxies. *MNRAS* 388, 500–550. doi:10.1111/j.1365-2966.2008.13422.x

Er, X.-Y., Jiang, Z.-B., and Fu, Y.-N. (2009). A Numerical Simulation Study of Mass Segregation in Embedded Stellar Clusters. *Chin. Astron. Astrophys.* 33, 139–150. doi:10.1016/j.chinastron.2009.03.012

Faherty, J. K., Bochanski, J. J., Gagné, J., Nelson, O., Coker, K., Smithka, I., et al. (2018). New and Known Moving Groups and Clusters Identified in a Gaia Comoving Catalog. *ApJ* 863, 91. doi:10.3847/1538-4357/aac76e

Fernández, D., Figueras, F., and Torra, J. (2001). Kinematics of Young Stars. *A&A* 372, 833–850. doi:10.1051/0004-6361:20010366

Fitzpatrick, E. L. (1999). Correcting for the Effects of Interstellar Extinction. *Publ. Astron. Soc. Pac.* 111, 63–75. doi:10.1086/316293

Gaburov, E., Gualandris, A., and Portegies Zwart, S. (2008). On the Onset of Runaway Stellar Collisions in Dense star Clusters - I. Dynamics of the First Collision. *Monthly Notices RAS* 384, 376–385. doi:10.1111/j.1365-2966.2007.12731.x

Gaia Collaboration, Babusiaux, C., van Leeuwen, F., Barstow, M. A., Jordi, C., Vallenari, A., et al. (2018a). Gaia Data Release 2. Observational Hertzsprung-Russell Diagrams. *A&A* 616, A10. doi:10.1051/0004-6361/201832843

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Babusiaux, C., et al. (2018b). Gaia Data Release 2. Summary of the Contents and Survey Properties. *A&A* 616, A1. doi:10.1051/0004-6361/201833051

Gaia Collaboration, Katz, D., Antoja, T., Romero-Gómez, M., Drimmel, R., Reylé, C., et al. (2018c). Gaia Data Release 2. Mapping the Milky Way Disc Kinematics. *A&A* 616, A11. doi:10.1051/0004-6361/201832865

Gaia Collaboration, van Leeuwen, F., Vallenari, A., Jordi, C., Lindegren, L., Bastian, U., et al. (2017). Gaia Data Release 1. Open Cluster Astrometry: Performance, Limitations, and Future Prospects. *A&A* 601, A19. doi:10.1051/0004-6361/201730552

García-Hernández, D. A., and Díaz-Luis, J. J. (2013). Diffuse Interstellar Bands in Fullerene Planetary Nebulae: the Fullerenes - Diffuse Interstellar Bands Connection. *A&A* 550, L6. doi:10.1051/0004-6361/201220919

Geen, S., Soler, J. D., and Hennebelle, P. (2017). Interpreting the star Formation Efficiency of Nearby Molecular Clouds with Ionizing Radiation. *MNRAS* 471, 4844–4855. doi:10.1093/mnras/stx1765

Gerhard, O. (2011). Pattern Speeds in the Milky Way. *Memorie della Societa Astronomica Italiana Supplementi* 18, 185.

Gieles, M., Sana, H., and Portegies Zwart, S. F. (2010). On the Velocity Dispersion of Young star Clusters: Super-virial or Binaries? *MNRAS* 402, 1750–1757. doi:10.1111/j.1365-2966.2009.15993.x

Gieles, M., Zwart, S. F. P., Baumgardt, H., Athanassoula, E., Lamers, H. J. G. L. M., Sipior, M., et al. (2006). Star Cluster Disruption by Giant Molecular Clouds. *Monthly Notices R. Astronomical Soc.* 371, 793–804. doi:10.1111/j.1365-2966.2006.10711.x

He, C.-C., Ricotti, M., and Geen, S. (2019). Simulating star Clusters across Cosmic Time - I. Initial Mass Function, star Formation Rates, and Efficiencies. *MNRAS* 489, 1880–1898. doi:10.1093/mnras/stz2239

Heckman, T. M., and Lehnert, M. D. (2000). The Detection of the Diffuse Interstellar Bands in Dusty Starburst Galaxies. *ApJ* 537, 690–696. doi:10.1086/309086

Heger, M. L. (1922). The Spectra of Certain Class B Stars in the Regions 5630A-6680A and 3280A-3380A. *Lick Observatory Bull.* 10, 146–147. doi:10.5479/ads/bib/1922licob.10.141h

Herbig, G. H. (1995). The Diffuse Interstellar Bands. *Annu. Rev. Astron. Astrophys.* 33, 19–73. doi:10.1146/annurev.aa.33.090195.000315

Hou, L. G., and Han, J. L. (2014). The Observed Spiral Structure of the Milky Way. *A&A* 569, A125. doi:10.1051/0004-6361/201424039

Jenniskens, P., Ehrenfreund, P., and Foing, B. (1994). Diffuse Interstellar Bands in Orion. The Environment Dependence of DIB Strength. *A&A* 281, 517–525.

Jørgensen, B. R., and Lindegren, L. (2005). Determination of Stellar Ages from Isochrones: Bayesian Estimation versus Isochrone Fitting. *A&A* 436, 127–143. doi:10.1051/0004-6361:20042185

Katz, D., Sartoretti, P., Cropper, M., Panuzzo, P., Seabroke, G. M., Viala, Y., et al. (2019). Gaia Data Release 2. *A&A* 622, A205. doi:10.1051/0004-6361/201833273

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., and Scholz, R.-D. (2013). Global Survey of star Clusters in the Milky Way. *A&A* 558, A53. doi:10.1051/0004-6361/201322302

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., and Scholz, R.-D. (2016). Global Survey of star Clusters in the Milky Way. *A&A* 585, A101. doi:10.1051/0004-6361/201527292

Kim, D., Lu, J. R., Konopacky, Q., Chu, L., Toller, E., Anderson, J., et al. (2019). Stellar Proper Motions in the Orion Nebula Cluster. *AJ* 157, 109. doi:10.3847/1538-3881/aafb09

Kos, J., and Zwitter, T. (2013). Properties of Diffuse Interstellar Bands at Different Physical Conditions of the Interstellar Medium. *ApJ* 774, 72. doi:10.1088/0004-637X/774/1/72

Krełowski, J. (2018). Diffuse Interstellar Bands. A Survey of Observational Facts. *PASP* 130, 071001. doi:10.1088/1538-3873/aabd69

Krełowski, J., Galazutdinov, G. A., Bondar, A., and Beletsky, Y. (2016). Observational Analysis of the Well-Correlated Diffuse Bands: 6196 and 6614 Å. *MNRAS* 460, 2706–2710. doi:10.1093/mnras/stw1167

Krone-Martins, A., and Moitinho, A. (2014). UPMASK: Unsupervised Photometric Membership Assignment in Stellar Clusters. *A&A* 561, A57. doi:10.1051/0004-6361/201321143

Lada, C. J., and Lada, E. A. (2003). Embedded Clusters in Molecular Clouds. *Annu. Rev. Astron. Astrophys.* 41, 57–115. doi:10.1146/annurev.astro.41.011802.094844

Lada, C. J., Margulis, M., and Dearborn, D. (1984). The Formation and Early Dynamical Evolution of Bound Stellar Systems. *ApJ* 285, 141–152. doi:10.1086/162485

Leisawitz, D., Bash, F. N., and Thaddeus, P. (1989). A CO Survey of Regions Around 34 Open Clusters. *ApJS* 70, 731. doi:10.1086/191357

Liu, L., and Pang, X. (2019). A Catalog of Newly Identified Star Clusters in Gaia DR2. *ApJS* 245, 32. doi:10.3847/1538-4365/ab530a

Liu, T., Janes, K. A., and Bania, T. M. (1991). More Radial-Velocity Measurements in Young Open Clusters. *AJ* 102, 1103. doi:10.1086/115936

Lodieu, N., Pérez-Garrido, A., Smart, R. L., and Silvotti, R. (2019). A 5D View of the α Per, Pleiades, and Praesepe Clusters. *A&A* 628, A66. doi:10.1051/0004-6361/201935533

Loinard, L., Dame, T. M., Heyer, M. H., Lequeux, J., and Thaddeus, P. (1999). A CO Survey of the Southwest Half of M 31. *A&A* 351, 1087–1102.

Luri, X., Brown, A. G. A., Sarro, L. M., Arenou, F., Bailer-Jones, C. A. L., Castro-Ginard, A., et al. (2018). Gaia Data Release 2. *A&A* 616, A9. doi:10.1051/0004-6361/201832964

Maier, J. P., Walker, G. A. H., Bohlender, D. A., Mazzotti, F. J., Raghunandan, R., Fulara, J., et al. (2011). Identification of H2Ccc as a Diffuse Interstellar Band Carrier. *ApJ* 726, 41. doi:10.1088/0004-637X/726/1/41

Maier, J. P., Walker, G. A. H., and Bohlender, D. A. (2004). On the Possible Role of Carbon Chains as Carriers of Diffuse Interstellar Bands. *ApJ* 602, 286–290. doi:10.1086/381027

Mathewson, D. S., van der Kruit, P. C., and Brouw, W. N. (1972). A High Resolution Radio Continuum Survey of M51 and NGC 5195 at 1415 MHz. *A&A* 17, 468.

Matzner, C. D. (2002). On the Role of Massive Stars in the Support and Destruction of Giant Molecular Clouds. *ApJ* 566, 302–314. doi:10.1086/338030

Meidt, S. E., Rand, R. J., Merrifield, M. R., Shetty, R., and Vogel, S. N. (2008). Radial Dependence of the Pattern Speed of M51. *Astrophysical J.* 688, 224–236. doi:10.1086/591516

Mermilliod, J.-C., Mayor, M., and Udry, S. (2009). Catalogues of Radial and Rotational Velocities of 1253 F-K Dwarfs in 13 Nearby Open Clusters. *A&A* 498, 949–960. doi:10.1051/0004-6361/200810244

Mermilliod, J. C. (2006). VizieR Online Data Catalog: Homogeneous Means in the UBV System (Mermilliod 1991), *VizieR Online Data Catalog*, II, 168.

Merrill, P. W., and Wilson, O. C. (1938). Unidentified Interstellar Lines in the Yellow and Red. *ApJ* 87, 9. doi:10.1086/143897

Miralles-Caballero, D., Díaz, A. I., Rosales-Ortega, F. F., Pérez-Montero, E., and Sánchez, S. F. (2014). Ionizing Stellar Population in the Disc of NGC 3310 - I. The Impact of a Minor Merger on Galaxy Evolution. *MNRAS* 440, 2265–2289. doi:10.1093/mnras/stu435

Moffat, A. F. J., Fitzgerald, M. P., and Jackson, P. D. (1979). The Rotation and Structure of the Galaxy beyond the Solar circle. I. Photometry and Spectroscopy of 276 Stars in 45 H II Regions and Other Young Stellar Groups toward the Galactic Anticentre. *A&AS* 38, 197–225.

Monreal-Ibero, A., Weilbacher, P. M., and Wendt, M. (2018). Diffuse Interstellar Bands λ5780 and λ5797 in the Antennae Galaxy as Seen by MUSE. *A&A* 615, A33. doi:10.1051/0004-6361/201732178

Monteiro, H., and Dias, W. S. (2019). Distances and Ages from Isochrone Fits of 150 Open Clusters Using Gaia DR2 Data. *MNRAS* 487, 2385–2406. doi:10.1093/mnras/stz1455

Monteiro, H., Dias, W. S., Moitinho, A., Cantat-Gaudin, T., Lépine, J. R. D., Carraro, G., et al. (2020). Fundamental Parameters for 45 Open Clusters with Gaia DR2, an Improved Extinction Correction and a Metallicity Gradient Prior. *MNRAS* 499, 1874–1889. doi:10.1093/mnras/staa2983

Morales, E. F. E., Wyrowski, F., Schuller, F., and Menten, K. M. (2013). Stellar Clusters in the Inner Galaxy and Their Correlation with Cold Dust Emission. *A&A* 560, A76. doi:10.1051/0004-6361/201321626

Mulder, P. S., and Combes, F. (1996). Dynamical Modeling of Two Nearby Disc Galaxies. *A&A* 313, 723–732.

Netopil, M., Paunzen, E., and Carraro, G. (2015). A Comparative Study on the Reliability of Open Cluster Parameters. *A&A* 582, A19. doi:10.1051/0004-6361/201526372

Netopil, M., Paunzen, E., Heiter, U., and Soubiran, C. (2016). On the Metallicity of Open Clusters. *A&A* 585, A150. doi:10.1051/0004-6361/201526370

Niu, H., Wang, J., and Fu, J. (2020). Binary Fraction Estimation of Main-Sequence Stars in 12 Open Clusters: Based on the Homogeneous Data of LAMOST Survey and Gaia DR2. *ApJ* 903, 93. doi:10.3847/1538-4357/abb8d6

Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton, T. D., and Spergel, D. N. (2017). Comoving Stars inGaiaDR1: An Abundance of Very Wide Separation Comoving Pairs. *AJ* 153, 257. doi:10.3847/1538-3881/aa6ffd

Omont, A., Bettinger, H. F., and Tönshoff, C. (2019). Polyacenes and Diffuse Interstellar Bands. *A&A* 625, A41. doi:10.1051/0004-6361/201834953

Percy, J. R. (2007). Understanding Variable Stars doi:10.1017/cbo9780511536489

Perren, G. I., Vázquez, R. A., and Piatti, A. E. (2015). ASteCA: Automated Stellar Cluster Analysis. *A&A* 576, A6. doi:10.1051/0004-6361/201424946

Piecka, M., and Paunzen, E. (2020). Mapping Local Interstellar Medium with Diffuse Interstellar Bands. *MNRAS* 495, 2035–2052. doi:10.1093/mnras/staa1112

Pöhnl, H., and Paunzen, E. (2010). A Statistical Method to Determine Open Cluster Metallicities. *A&A* 514, A81. doi:10.1051/0004-6361/200810855

Raimond, S., Lallement, R., Vergely, J. L., Babusiaux, C., and Eyer, L. (2012). A Southern Hemisphere Survey of the 5780 and 6284 Å Diffuse Interstellar Bands: Correlation with the Extinction. *A&A* 544, A136. doi:10.1051/0004-6361/201219191

Reid, M. J., Menten, K. M., Brunthaler, A., Zheng, X. W., Dame, T. M., Xu, Y., et al. (2019). Trigonometric Parallaxes of High-Mass Star-forming Regions: Our View of the Milky Way. *ApJ* 885, 131. doi:10.3847/1538-4357/ab4a11

Reid, M. J., Menten, K. M., Brunthaler, A., Zheng, X. W., Dame, T. M., Xu, Y., et al. (2014). Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way. *ApJ* 783, 130. doi:10.1088/0004-637X/783/2/130

Ringermacher, H. I., and Mead, L. R. (2009). A New Formula Describing the Scaffold Structure of Spiral Galaxies. *MNRAS* 397, 164–171. doi:10.1111/j.1365-2966.2009.14950.x

Roberts, W. W. (1969). Large-Scale Shock Formation in Spiral Galaxies and its Implications on Star Formation. *ApJ* 158, 123. doi:10.1086/150177

Robitaille, T. P., Whitney, B. A., Indebetouw, R., Wood, K., and Denzmore, P. (2006). Interpreting Spectral Energy Distributions from Young Stellar Objects. I. A Grid of 200,000 YSO Model SEDs. *Astrophys J. Suppl. S* 167, 256–285. doi:10.1086/508424

Roeser, S., Demleitner, M., and Schilbach, E. (2010). The PPMXL Catalog of Positions and Proper Motions on the ICRS. Combining USNO-B1.0 and the Two Micron All Sky Survey (2MASS). *Astronomical J.* 139, 2440–2447. doi:10.1088/0004-6256/139/6/2440

Röser, S., Schilbach, E., Schwan, H., Kharchenko, N. V., Piskunov, A. E., and Scholz, R.-D. (2008). PPM-extended (PPMX) - a Catalogue of Positions and Proper Motions. *A&A* 488, 401–408. doi:10.1051/0004-6361:200809775

Russeil, D. (2003). Star-forming Complexes and the Spiral Structure of Our Galaxy. *A&A* 397, 133–146. doi:10.1051/0004-6361:20021504

Salama, F., Galazutdinov, G. A., Krełowski, J., Biennier, L., Beletsky, Y., and Song, I.-O. (2011). Polycyclic Aromatic Hydrocarbons and the Diffuse Interstellar Bands: A Survey. *ApJ* 728, 154. doi:10.1088/0004-637X/728/2/154

Seigar, M. S., and James, P. A. (1998). The Structure of Spiral Galaxies - II. Near-Infrared Properties of Spiral Arms. *MNRAS* 299, 685–698. doi:10.1046/j.1365-8711.1998.01779.x

Sellwood, J. A. (2012). Spiral Instabilities Inn-Body Simulations. I. Emergence from Noise. *ApJ* 751, 44. doi:10.1088/0004-637X/751/1/44

Sellwood, J. A., Trick, W. H., Carlberg, R. G., Coronado, J., and Rix, H.-W. (2019). Discriminating Among Theories of Spiral Structure Using Gaia DR2. *MNRAS* 484, 3154–3167. doi:10.1093/mnras/stz140

Shi, Y.-Y., Zhu, Z., Liu, N., Liu, J.-C., Ding, C.-Y., and Cheng, Y.-T. (2019). Comparison of the PPMXL and UCAC5 Catalogs with the Gaia DR2. *AJ* 157, 222. doi:10.3847/1538-3881/ab17d7

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., et al. (2006). The Two Micron All Sky Survey (2MASS). *Astron. J.* 131, 1163–1183. doi:10.1086/498708

Smith, M. C., Hannah Whiteoak, S., and Evans, N. W. (2012). Slicing and Dicing the Milky Way Disk in the Sloan Digital Sky Survey. *ApJ* 746, 181. doi:10.1088/0004-637X/746/2/181

Smith, R., Slater, R., Fellhauer, M., Goodwin, S., and Assmann, P. (2011). Formation Rates of star Clusters in the Hierarchical Merging Scenario. *MNRAS* 416, 383–390. doi:10.1111/j.1365-2966.2011.19039.x

Sofue, Y., Tutui, Y., Honma, M., Tomita, A., Takamiya, T., Koda, J., et al. (1999). Central Rotation Curves of Spiral Galaxies. *ApJ* 523, 136–146. doi:10.1086/307731

Solin, O., Haikala, L., and Ukkonen, E. (2014). Mining the VVV: star Formation and Embedded Clusters. *A&A* 562, A115. doi:10.1051/0004-6361/201322890

Sollerman, J., Cox, N., Mattila, S., Ehrenfreund, P., Kaper, L., Leibundgut, B., et al. (2005). Diffuse Interstellar Bands in NGC 1448. *A&A* 429, 559–567. doi:10.1051/0004-6361:20041465

Soubiran, C., Cantat-Gaudin, T., Romero-Gómez, M., Casamiquela, L., Jordi, C., Vallenari, A., et al. (2018). Open Cluster Kinematics with Gaia DR2. *A&A* 619, A155. doi:10.1051/0004-6361/201834020

Spitzer, L. J., and Lyman (1958). Distribution of Galactic Clusters. *ApJ* 127, 17. doi:10.1086/146435

Straižys, V., Boyle, R. P., Raudeliūnas, S., Zdanavičius, J., Janusz, R., Macijauskas, M., et al. (2020). Open Cluster IC 1369 and its Vicinity: Multicolor Photometry and Gaia DR2 Astrometry. *AJ* 159, 95. doi:10.3847/1538-3881/ab67b5

Tang, N., Li, D., Heiles, C., Wang, S., Pan, Z., and Wang, J.-J. (2016). Physical Properties of CO-dark Molecular Gas Traced by C+. *A&A* 593, A42. doi:10.1051/0004-6361/201528055

Toomre, A. (1977). Theories of Spiral Structure. *Annu. Rev. Astron. Astrophys.* 15, 437–478. doi:10.1146/annurev.aa.15.090177.002253

van den Bergh, S. (2006). Diameters of Open Star Clusters. *Astron. J.* 131, 1559–1564. doi:10.1086/499532

van Leeuwen, F. (2007). Validation of the New Hipparcos Reduction. *A&A* 474, 653–664. doi:10.1051/0004-6361:20078357

von Hippel, T., Jefferys, W. H., Scott, J., Stein, N., Winget, D. E., DeGennaro, S., et al. (2006). Inverting Color-Magnitude Diagrams to Access Precise Star Cluster Parameters: A Bayesian Approach. *ApJ* 645, 1436–1447. doi:10.1086/504369

Yontan, T., Bilir, S., Bostancı, Z. F., Ak, T., Ak, S., Güver, T., et al. (2019). CCD UBV Photometric and Gaia Astrometric Study of Eight Open Clusters-ASCC 115, Collinder 421, NGC 6793, NGC 7031, NGC 7039, NGC 7086, Roslund 1 and Stock 21. *Astrophys Space Sci.* 364, 152. doi:10.1007/s10509-019-3640-y

Yu, J., and Liu, C. (2018). The Age-Velocity Dispersion Relation of the Galactic Discs from LAMOST-Gaia Data. *MNRAS* 475, 1093–1103. doi:10.1093/mnras/stx3204

Zacharias, N., Finch, C., and Frouard, J. (2017). UCAC5: New Proper Motions UsingGaiaDR1. *AJ* 153, 166. doi:10.3847/1538-3881/aa6196

Zasowski, G., Ménard, B., Bizyaev, D., García-Hernández, D. A., Pérez, A. E. G., Hayden, M. R., et al. (2015). Mapping the Interstellar Medium with Near-Infrared Diffuse Interstellar Bands. *ApJ* 798, 35. doi:10.1088/0004-637X/798/1/35

Zhang, B., Reid, M. J., Zhang, L., Wu, Y., Hu, B., Sakai, N., et al. (2019). Parallaxes for Star-forming Regions in the Inner Perseus Spiral Arm. *AJ* 157, 200. doi:10.3847/1538-3881/ab141d

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Chapter 9

Summary

During the last six years, Gaia data became very frequently used in astrophysical studies. The provided astrometric and photometric data are invaluable. Distances can be measured toward ~ 1 % of stars in our Galaxy. However, many issues are often ignored. Furthermore, the topic of DIBs has not been gaining on popularity since the 2000s, despite the fact that they represent an important problem that still needs to be solved.

The presented works show that more research needs to be done regarding both, the open clusters and the DIBs. Gaia data releases provide unique opportunities to study both subjects. First of all, the four main cluster parameters can be determined more precisely than ever before. If we focus our attention on the clusters with $\log \mathrm{Age} \leq 8.0$, we should be able to trace Galactic spiral arms more precisely than in the case when age is not taken into account. Such research has not been yet been successful prior to the Gaia era.

When studying open clusters when using large catalogues of cluster members, one must keep in mind that the catalogues which make use of the current version of UPMASK method may contain improperly derived membership probabilities. This mostly affects such clusters that significantly overlap in the phase space. Curiously, such clusters may be of great interest when searching for binary clusters. An example of this kind of study was very recently published by Angelo et al. (2022).

Furthermore, it must be brought to attention that the distribution of cluster members in 3D space will become elongated along the line of sight. This is simply the result of the existence of the (observational) parallax uncertainties, which are (in relative values) significantly larger than the uncertainties in the positions in the sky. This does not hugely affect the determination of the cluster distances but rather influences the studies of the cluster shapes in the 3D space. Such studies are still largely missing in the literature.

Finally, the DIBs can be studied in much greater detail than ever before. One of the main reasons is that the stellar distances can be determined very precisely thanks to the Gaia data. The maps of DIBs can be taken more seriously and can provide important knowledge about the structure of the ISM. Unfortunately, the quality of the maps is still not good enough for tracing Galactic spiral arms. Hopefully, this will change in future studies.

Chapter 10

List of Publications

- *Mapping local interstellar medium with diffuse interstellar bands*

  Piecka, M., Paunzen, E., 2020, Monthly Notices of the Royal Astronomical Society, Volume 495, Issue 2, 18 pp.

  MP prepared and analysed the different sources of the EWs of DIBs, created the maps of DIBs, analysed the figures and separated the data into several regions, was responsible for writing the larger part of the paper, and submitted the paper.

- *White dwarf-open cluster associations based on Gaia DR2*

  Prišegen, M., **Piecka, M.**, Faltová, N., Kajan, M., Paunzen, E., 2021, Astronomy & Astrophysics, Volume 645, id.A13, 19 pp.

  MP helped during the initial phases when searching for the white dwarfs, prepared the CMDs and wrote a larger part of the subsection about isochrones.

- *Aggregates of clusters in the Gaia data*

  Piecka, M., Paunzen, E., 2021, Astronomy & Astrophysics, Volume 649, id.A54, 12 pp.

  MP studied the reduced phase space of open clusters from the recently published catalogue, searched for the aggregates of clusters, analysed the CMDs and the other diagrams, was responsible for writing the larger part of the paper, and submitted the paper.

- *A Comparison of the Simulations and Observations for a Nearby Spiral Arm*

  Piecka, M., Paunzen, E., 2021, Frontiers in Astronomy and Space Sciences, Volume 8, id.89, 20 pp.

  MP studied the distribution of clusters, prepared the analytical and a simple numerical method to analyse spiral arms, was responsible for writing the larger part of the paper, and submitted the paper.

- *Structure of Open Clusters – Gaia DR2 and its limitations*

  Piecka, M., Paunzen, E., 2022, Bulgarian Astronomical Journal, Volume 36, 27 pp.

  MP analysed the parallax distributions from the cited catalogue of open clusters, prepared the mathematical analysis of the parallax problem, calculated the projected widths of open clusters, and was responsible for writing the larger part of the paper.

Bibliography

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, Phys. Rev. Lett., 119, 161101

Adamson, A. J. & Whittet, D. C. B. 1992, ApJ, 398, L69

Adamson, A. J. & Whittet, D. C. B. 1995, ApJ, 448, L49

Adamson, A. J., Whittet, D. C. B., & Duley, W. W. 1991, MNRAS, 252, 234

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21

Aiello, S., Barsella, B., Chlewicki, G., et al. 1988, A&AS, 73, 195

Alonso, A., Arribas, S., & Martinez-Roger, C. 1996, A&A, 313, 873

Alonso, A., Arribas, S., & Martínez-Roger, C. 1999, A&AS, 140, 261

Alonso-Santiago, J., Frasca, A., Catanzaro, G., et al. 2021, A&A, 656, A149

Alonso-Santiago, J., Negueruela, I., Marco, A., et al. 2017, MNRAS, 469, 1330

Altobelli, N., Kempf, S., Krüger, H., et al. 2005, Journal of Geophysical Research (Space Physics), 110, A07102

Altobelli, N., Postberg, F., Fiege, K., et al. 2016, Science, 352, 312

Anders, E. & Zinner, E. 1993, Meteoritics, 28, 490

Andersson, B. G., Lazarian, A., & Vaillancourt, J. E. 2015, ARA&A, 53, 501

Angelo, M. S., Corradi, W. J. B., Santos, J. F. C., J., Maia, F. F. S., & Ferreira, F. A. 2021, MNRAS, 500, 4338

Angelo, M. S., Santos, J. F. C., & Corradi, W. J. B. 2020, MNRAS, 493, 3473

Angelo, M. S., Santos, J. F. C., Maia, F. F. S., & Corradi, W. J. B. 2022, MNRAS, 510, 5695

Anguiano, B., Majewski, S. R., Freeman, K. C., Mitschang, A. W., & Smith, M. C. 2018, MNRAS, 474, 854

Baba, J., Kawata, D., Matsunaga, N., Grand, R. J. J., & Hunt, J. A. S. 2018, ApJ, 853, L23

Bailer-Jones, C. A. L. 2015, PASP, 127, 994

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58

Balaguer-Núñez, L., Jordi, C., Galadí-Enríquez, D., & Zhao, J. L. 2004, A&A, 426, 819

Balaguer-Núñez, L., López del Fresno, M., Solano, E., et al. 2020, MNRAS, 492, 5811

Barton, E. J., Geller, M. J., & Kenyon, S. J. 2000, ApJ, 530, 660

Bash, F. N., Green, E., & Peters, W. L., I. 1977, ApJ, 217, 464

Belokurov, V., Erkal, D., Deason, A. J., et al. 2017, MNRAS, 466, 4711

Bennett, M. & Bovy, J. 2019, MNRAS, 482, 1417

Bernal, J. J., Haenecour, P., Howe, J., et al. 2019, ApJ, 883, L43

Berné, O., Montillaud, J., & Joblin, C. 2015, A&A, 577, A133

Bianchi, S. & Schneider, R. 2007, MNRAS, 378, 973

Bobylev, V. V. & Bajkova, A. T. 2014, MNRAS, 437, 1549

Boesgaard, A. M. 1989, ApJ, 336, 798

Bohren, C. F. & Huffman, D. R. 1983, Absorption and scattering of light by small particles (New York: Wiley)

Bonnell, I. A., Bate, M. R., & Zinnecker, H. 1998, MNRAS, 298, 93

Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, A&A, 623, A108

Bottinelli, L., Gouguenheim, L., Paturel, G., & Teerikorpi, P. 1986, A&A, 166, 393

Bottinelli, L., Gouguenheim, L., Paturel, G., & Teerikorpi, P. 1988, ApJ, 328, 4

Boubert, D., Guillochon, J., Hawkins, K., et al. 2018, MNRAS, 479, 2789

Bovy, J. 2017a, MNRAS, 468, L63

Bovy, J. 2017b, MNRAS, 470, 1360

Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127

Caldwell, N., Harding, P., Morrison, H., et al. 2009, AJ, 137, 94

Cami, J., Peeters, E., Bernard-Salas, J., Doppmann, G., & De Buizer, J. 2018, Galaxies, 6, 101

Cami, J., Salama, F., Jiménez-Vicente, J., Galazutdinov, G. A., & Krelowski, J. 2004, ApJ, 611, L113

Cami, J., Sonnentrucker, P., Ehrenfreund, P., & Foing, B. H. 1997, A&A, 326, 822

Campbell, E. K., Holz, M., Gerlich, D., & Maier, J. P. 2015, Nature, 523, 322

Cantat-Gaudin, T. & Anders, F. 2020, A&A, 633, A99

Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1

Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93

Cantat-Gaudin, T., Krone-Martins, A., Sedaghat, N., et al. 2019a, A&A, 624, A126

Cantat-Gaudin, T., Mapelli, M., Balaguer-Núñez, L., et al. 2019b, A&A, 621, A115

Carrera, R., Pasquato, M., Vallenari, A., et al. 2019, A&A, 627, A119

Castellani, V., Degl'Innocenti, S., Prada Moroni, P. G., & Tordiglione, V. 2002, MNRAS, 334, 193

Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, A&A, 635, A45

Castro-Ginard, A., Jordi, C., Luri, X., et al. 2018, A&A, 618, A59

Castro-Ginard, A., McMillan, P. J., Luri, X., et al. 2021, A&A, 652, A162

Cherchneff, I., Le Teuff, Y. H., Williams, P. M., & Tielens, A. G. G. M. 2000, A&A, 357, 572

Chevalier, R. A. & Oegerle, W. R. 1979, ApJ, 227, 398

Christian, C. A., Adams, M., Barnes, J. V., et al. 1985, PASP, 97, 363

Chrysostomou, A., Hough, J. H., Whittet, D. C. B., et al. 1996, ApJ, 465, L61

Churchwell, E. & Walmsley, C. M. 1975, A&A, 38, 451

Clayton, D. D., Deneault, E. A. N., & Meyer, B. S. 2001, ApJ, 562, 480

Cordiner, M. A., Cox, N. L. J., Trundle, C., et al. 2008a, A&A, 480, L13

Cordiner, M. A., Linnartz, H., Cox, N. L. J., et al. 2019, ApJ, 875, L28

Cordiner, M. A., Smith, K. T., Cox, N. L. J., et al. 2008b, A&A, 492, L5

Cordoni, G., Milone, A. P., Marino, A. F., et al. 2018, ApJ, 869, 139

Corsaro, E., Lee, Y.-N., García, R. A., et al. 2017, Nature Astronomy, 1, 0064

Crawford, D. L. & Perry, C. L. 1976, AJ, 81, 419

Cropper, M., Katz, D., Sartoretti, P., et al. 2018, A&A, 616, A5

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792

Dame, T. M., Hartmann, D., & Thaddeus, P. 2011, Replication data for: Taurus (DHT21), Harvard Dataverse

Darma, R., Arifyanto, M. I., & Kouwenhoven, M. B. N. 2021, MNRAS, 506, 4603

Davis, Leverett, J. & Greenstein, J. L. 1951, ApJ, 114, 206

de Juan Ovelar, M., Gossage, S., Kamann, S., et al. 2020, MNRAS, 491, 2129

de Zeeuw, P. T., Hoogerwerf, R., de Bruijne, J. H. J., Brown, A. G. A., & Blaauw, A. 1999, AJ, 117, 354

Dehnen, W. & Binney, J. J. 1998, MNRAS, 298, 387

Demarque, P. 1980, in Star Clusters, ed. J. E. Hesser, Vol. 85, 281–303

Destree, J. D., Snow, T. P., & Eriksson, K. 2007, ApJ, 664, 909

Dettmar, R. J. 1990, A&A, 232, L15

Dias, W. S., Alessi, B. S., Moitinho, A., & Lépine, J. R. D. 2002, A&A, 389, 871

Dias, W. S., Lépine, J. R. D., & Alessi, B. S. 2001, A&A, 376, 441

Dias, W. S., Monteiro, H., Moitinho, A., et al. 2021, MNRAS, 504, 356

Dolginov, A. Z. & Mitrofanov, I. G. 1976, Ap&SS, 43, 291

Dong, R. & Draine, B. T. 2011, ApJ, 727, 35

Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, ApJS, 178, 89

Draine, B. T. 1988, ApJ, 333, 848

Draine, B. T. 2003, ARA&A, 41, 241

Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton University Press)

Draine, B. T. 2016, ApJ, 831, 109

Draine, B. T. & Flatau, P. J. 1994, Journal of the Optical Society of America A, 11, 1491

Draine, B. T. & Flatau, P. J. 2013, arXiv e-prints, arXiv:1305.6497

Draine, B. T. & Hensley, B. S. 2021, ApJ, 910, 47

Draine, B. T. & Lee, H. M. 1984, ApJ, 285, 89

Dugan, Z., Bryan, S., Gaibler, V., Silk, J., & Haas, M. 2014, ApJ, 796, 113

Duley, W. W. 1968, Nature, 218, 153

Dullemond, C. P. & Dominik, C. 2005, A&A, 434, 971

Dzib, S. A., Loinard, L., Ortiz-León, G. N., Rodríguez, L. F., & Galli, P. A. B. 2018, ApJ, 867, 151

Eggen, O. J. 1960, MNRAS, 120, 540

Ehrenfreund, P., Cami, J., Jiménez-Vicente, J., et al. 2002, ApJ, 576, L117

Elmegreen, B. G. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 148, Origins, ed. C. E. Woodward, J. M. Shull, & J. Thronson, Harley A., 150

Elyajouri, M. & Lallement, R. 2019, A&A, 628, A67

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in (AAAI Press), 226–231

Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4

Fabricius, C., Bastian, U., Portell, J., et al. 2016, A&A, 595, A3

Fall, S. M. & Rees, M. J. 1985, ApJ, 298, 18

Famaey, B., Pont, F., Luri, X., et al. 2007, A&A, 461, 957

Ferlet, R., Vidal-Madjar, A., & Gry, C. 1985, ApJ, 298, 838

Fernández, D., Figueras, F., & Torra, J. 2001, A&A, 372, 833

Field, G. B. 1974, ApJ, 187, 453

Fitzpatrick, E. L. & Massa, D. 1990, ApJS, 72, 163

Fitzpatrick, E. L. & Massa, D. 2005, AJ, 130, 1127

Fitzpatrick, E. L. & Massa, D. 2007, ApJ, 663, 320

Fitzpatrick, E. L., Massa, D., Gordon, K. D., Bohlin, R., & Clayton, G. C. 2019, ApJ, 886, 108

Flower, P. J. 1996, ApJ, 469, 355

Foing, B. H. & Ehrenfreund, P. 1994, Nature, 369, 296

Francis, C. 2014, MNRAS, 444, L6

Friel, E. D., Janes, K. A., Tavarez, M., et al. 2002, AJ, 124, 2693

Frisch, P. C., Dorschner, J. M., Geiss, J., et al. 1999, ApJ, 525, 492

Gagné, J., Mamajek, E. E., Malo, L., et al. 2018, ApJ, 856, 23

Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, A&A, 616, A10

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, A&A, 616, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021a, A&A, 649, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016a, A&A, 595, A2

Gaia Collaboration, Katz, D., Antoja, T., et al. 2018c, A&A, 616, A11

Gaia Collaboration, Luri, X., Chemin, L., et al. 2021b, A&A, 649, A7

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016b, A&A, 595, A1

Galazutdinov, G., Bondar, A., Lee, B.-C., et al. 2020, AJ, 159, 113

Galazutdinov, G., Moutou, C., Musaev, F., & Krelowski, J. 2002, A&A, 384, 215

Galazutdinov, G. A., Lo Curto, G., & Krelowski, J. 2008, MNRAS, 386, 2003

Galli, P. A. B., Moraux, E., Bouy, H., et al. 2017, A&A, 598, A48

Gatewood, G., de Jonge, J. K., & Han, I. 2000, ApJ, 533, 938

Genzel, R. & Downes, D. 1977, A&AS, 30, 145

Gieles, M., Portegies Zwart, S. F., Baumgardt, H., et al. 2006, MNRAS, 371, 793

Gold, T. 1952, Nature, 169, 322

Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, ApJ, 680, 428

Gomez-Gonzalez, J. & Guelin, M. 1974, A&A, 32, 441

González Hernández, J. I. & Bonifacio, P. 2009, A&A, 497, 497

Goodman, F. O. 1978, ApJ, 226, 87

Gratton, R. G., Carretta, E., & Bragaglia, A. 2012, A&A Rev., 20, 50

Gruen, E., Fechtig, H., Kissel, J., et al. 1992, A&AS, 92, 411

Gutermuth, R. A., Megeath, S. T., Pipher, J. L., et al. 2005, ApJ, 632, 397

Habing, H. J. & Israel, F. P. 1979, ARA&A, 17, 345

Haffner, L. M., Dettmar, R. J., Beckman, J. E., et al. 2009, Reviews of Modern Physics, 81, 969

Hanson, R. B. 1975, AJ, 80, 379

Harper, G. M., Brown, A., Guinan, E. F., et al. 2017, AJ, 154, 11

Hawkins, I. & Jura, M. 1987, ApJ, 317, 926

Healy, B. F. & McCullough, P. R. 2020, ApJ, 903, 99

Heger, M. L. 1922, Lick Observatory Bulletin, 10, 141

Heiles, C. 1984, ApJS, 55, 585

Heiles, C. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 231, Tetons 4: Galactic Structure, Stars and the Interstellar Medium, ed. C. E. Woodward, M. D. Bicay, & J. M. Shull, 294

Heiter, U., Soubiran, C., Netopil, M., & Paunzen, E. 2014, A&A, 561, A93

Hennebelle, P. & Inutsuka, S.-i. 2019, Frontiers in Astronomy and Space Sciences, 6, 5

Henyey, L. G. & Greenstein, J. L. 1938, ApJ, 88, 580

Herbig, G. H. 1995, ARA&A, 33, 19

Hildebrand, R. H. & Dragovan, M. 1995, ApJ, 450, 663

Hiltner, W. A. 1949, ApJ, 109, 471

Hobbs, L. M. 1969, ApJ, 157, 165

Hobbs, L. M., York, D. G., Snow, T. P., et al. 2008, ApJ, 680, 1256

Hodge, P. W. 1979, AJ, 84, 744

Hoeg, E., Bässgen, G., Bastian, U., et al. 1997, A&A, 323, L57

Höfner, S. & Olofsson, H. 2018, A&A Rev., 26, 1

Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27

Holl, B. & Lindegren, L. 2012, A&A, 543, A14

Hong, J., Vesperini, E., Sollima, A., et al. 2015, MNRAS, 449, 629

Hoogerwerf, R., de Bruijne, J. H. J., & de Zeeuw, P. T. 2000, ApJ, 544, L133

Hoogerwerf, R., de Bruijne, J. H. J., & de Zeeuw, P. T. 2001, A&A, 365, 49

Huisken, F., Rouillé, G., Steglich, M., et al. 2014, in The Diffuse Interstellar Bands, ed. J. Cami & N. L. J. Cox, Vol. 297, 265–275

Iben, I., J. & Renzini, A. 1983, ARA&A, 21, 271

Janes, K. & Adler, D. 1982, ApJS, 49, 425

Janes, K. A. 1979, ApJS, 39, 135

Jenkins, E. B. 1978, ApJ, 219, 845

Jenkins, E. B. 2013, ApJ, 764, 25

Jenniskens, P. & Desert, F. X. 1993, A&A, 274, 465

Jenniskens, P. & Desert, F. X. 1994, A&AS, 106, 39

Jerabkova, T., Boffin, H. M. J., Beccari, G., et al. 2021, A&A, 647, A137

Jones, A. P. & Nuth, J. A. 2011, A&A, 530, A44

Jones, A. P., Tielens, A. G. G. M., Hollenbach, D. J., & McKee, C. F. 1994, ApJ, 433, 797

Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, A&A, 523, A48

Jørgensen, B. R. & Lindegren, L. 2005, A&A, 436, 127

Kanekar, N., Subrahmanyan, R., Chengalur, J. N., & Safouris, V. 2003, MNRAS, 346, L57

Katz, D., Sartoretti, P., Cropper, M., et al. 2019, A&A, 622, A205

Katz, N., Furman, I., Biham, O., Pirronello, V., & Vidali, G. 1999, ApJ, 522, 305

Kawata, D., Baba, J., Ciucă, I., et al. 2018, MNRAS, 479, L108

Kharchenko, N. V., Piskunov, A. E., Röser, S., Schilbach, E., & Scholz, R. D. 2004, Astronomische Nachrichten, 325, 740

Kharchenko, N. V., Piskunov, A. E., Röser, S., Schilbach, E., & Scholz, R. D. 2005, A&A, 438, 1163

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

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

Kim, D., Lu, J. R., Konopacky, Q., et al. 2019, AJ, 157, 109

King, I. 1962, AJ, 67, 471

Kos, J., Zwitter, T., Wyse, R., et al. 2014, Science, 345, 791

Kounkel, M. & Covey, K. 2019, AJ, 158, 122

Krelowski, J. 2018, PASP, 130, 071001

Krelowski, J., Galazutdinov, G. A., Bondar, A., & Beletsky, Y. 2016, MNRAS, 460, 2706

Krelowski, J. & Walker, G. A. H. 1987, ApJ, 312, 860

Kron, G. E. & Mayall, N. U. 1960, AJ, 65, 581

Krone-Martins, A. & Moitinho, A. 2014, A&A, 561, A57

Kroto, H. W., Heath, J. R., Obrien, S. C., Curl, R. F., & Smalley, R. E. 1985, Nature, 318, 162

Kroupa, P., Aarseth, S., & Hurley, J. 2001, MNRAS, 321, 699

Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, ApJ, 870, 32

Lada, C. J. & Lada, E. A. 2003, ARA&A, 41, 57

Lai, T. S. Y., Witt, A. N., Alvarez, C., & Cami, J. 2020, MNRAS, 492, 5853

Lallement, R., Babusiaux, C., Vergely, J. L., et al. 2019, A&A, 625, A135

Lallement, R., Welsh, B. Y., Vergely, J. L., Crifo, F., & Sfeir, D. 2003, A&A, 411, 447

Lamers, H. J. G. L. M., Gieles, M., Bastian, N., et al. 2005a, A&A, 441, 117

Lamers, H. J. G. L. M., Gieles, M., & Portegies Zwart, S. F. 2005b, A&A, 429, 173

Lazarian, A. 2003, J. Quant. Spectr. Rad. Transf., 79-80, 881

Lazarian, A. 2007, J. Quant. Spectr. Rad. Transf., 106, 225

Leger, A. & D'Hendecourt, L. 1985, A&A, 146, 81

Leger, A., D'Hendecourt, L., & Defourneau, D. 1989, A&A, 216, 148

Leger, A. & Puget, J. L. 1984, A&A, 500, 279

Lejeune, T. & Schaerer, D. 2001, A&A, 366, 538

Leonard, P. J. T. 1989, AJ, 98, 217

Leroy, J. L. 1999, A&A, 346, 955

Li, A. & Draine, B. T. 2001, ApJ, 554, 778

Li, F. & Ikeuchi, S. 1990, ApJS, 73, 401

Li, S., Frank, A., & Blackman, E. G. 2014, MNRAS, 444, 2884

Lindegren, L. & Bastian, U. 2010, in EAS Publications Series, Vol. 45, EAS Publications Series, 109–114

Lindegren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4

Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2

Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021b, A&A, 649, A2

Lindegren, L., Lammers, U., Bastian, U., et al. 2016, A&A, 595, A4

Lindegren, L., Lammers, U., Hobbs, D., et al. 2012, A&A, 538, A78

Lindenmann, E. & Hauck, B. 1973, A&AS, 11, 119

Lodieu, N., Pérez-Garrido, A., Smart, R. L., & Silvotti, R. 2019, A&A, 628, A66

Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9

Lutz, T. E. & Kelker, D. H. 1973, PASP, 85, 573

Mac Low, M.-M. & McCray, R. 1988, ApJ, 324, 776

Mackey, A. D. & Broby Nielsen, P. 2007, MNRAS, 379, 151

Maier, J. P., Walker, G. A. H., & Bohlender, D. A. 2004, ApJ, 602, 286

Malmquist, K. G. 1922, Meddelanden fran Lunds Astronomiska Observatorium Serie I, 100, 1

Malo, L., Doyon, R., Lafrenière, D., et al. 2013, ApJ, 762, 88

Marino, A. F., Milone, A. P., Casagrande, L., et al. 2018, ApJ, 863, L33

Martin, P. G. & Angel, J. R. P. 1975, ApJ, 195, 379

Martinez-Medina, L. A., Pichardo, B., Moreno, E., Peimbert, A., & Velazquez, H. 2016, ApJ, 817, L3

Martins, F. & Palacios, A. 2013, A&A, 560, A16

Masana, E., Jordi, C., & Ribas, I. 2006, A&A, 450, 735

Mathis, J. S. 1994, ApJ, 422, 176

Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425

Mayall, N. U. 1946, ApJ, 104, 290

McCall, M. L. 2004, AJ, 128, 2144

McKee, C. F. & Ostriker, E. C. 2007, ARA&A, 45, 565

McKee, C. F. & Ostriker, J. P. 1977, ApJ, 218, 148

McSwain, M. V. & Gies, D. R. 2005, ApJS, 161, 118

Meingast, S. & Alves, J. 2019, A&A, 621, L3

Mermilliod, J.-C. 1995, in Information & On-Line Data in Astronomy, ed. D. Egret & M. A. Albrecht, Vol. 203, 127

Merrill, P. W. & Wilson, O. C. 1938, ApJ, 87, 9

Mestel, L. & Spitzer, L., J. 1956, MNRAS, 116, 503

Mie, G. 1908, Annalen der Physik, 330, 377

Miller, G. E. & Scalo, J. M. 1978, PASP, 90, 506

Montargès, M., Kervella, P., Perrin, G., et al. 2016, A&A, 588, A130

Monteiro, H. & Dias, W. S. 2019, MNRAS, 487, 2385

Monteiro, H., Dias, W. S., & Caetano, T. C. 2010, A&A, 516, A2

Monteiro, H., Dias, W. S., Hickel, G. R., & Caetano, T. C. 2017, New A, 51, 15

Montes, D., López-Santiago, J., Gálvez, M. C., et al. 2001, MNRAS, 328, 45

Mostefaoui, S. & Hoppe, P. 2004, ApJ, 613, L149

Mucciarelli, A. & Bellazzini, M. 2020, Research Notes of the American Astronomical Society, 4, 52

Munari, U. & Zwitter, T. 1997, A&A, 318, 269

Murray, N. 2011, ApJ, 729, 133

Nakano, T. 1998, ApJ, 494, 587

Nandy, K. & Seddon, H. 1970, Nature, 227, 264

Netopil, M. & Paunzen, E. 2013, A&A, 557, A10

Netopil, M., Paunzen, E., Heiter, U., & Soubiran, C. 2016, A&A, 585, A150

Nie, T. P., Xiang, F. Y., & Li, A. 2022, MNRAS, 509, 4908

Nilakshi, Sagar, R., Pandey, A. K., & Mohan, V. 2002, A&A, 383, 153

Nozawa, T., Kozasa, T., Umeda, H., Maeda, K., & Nomoto, K. 2003, ApJ, 598, 785

O'dell, M. A., Hendry, M. A., & Collier Cameron, A. 1994, MNRAS, 268, 181

Onaka, T. & Okada, Y. 2003, ApJ, 585, 872

Ormel, C. W., Spaans, M., & Tielens, A. G. G. M. 2007, A&A, 461, 215

Overbeck, J. W. 1965, ApJ, 141, 864

Palla, F. & Stahler, S. W. 1999, ApJ, 525, 772

Patenaude, M. 1978, A&A, 66, 225

Patrikeev, I., Fletcher, A., Stepanov, R., et al. 2006, A&A, 458, 441

Paunzen, E., Heiter, U., Netopil, M., & Soubiran, C. 2010, A&A, 517, A32

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3

Pecaut, M. J. & Mamajek, E. E. 2013, ApJS, 208, 9

Perets, H. B. & Biham, O. 2006, MNRAS, 365, 801

Perren, G. I., Vázquez, R. A., & Piatti, A. E. 2015, A&A, 576, A6

Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, A&A, 369, 339

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1997, A&A, 500, 501

Petuchowski, S. J. & Bennett, C. L. 1993, ApJ, 405, 591

Pinsonneault, M. H., Terndrup, D. M., & Yuan, Y. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 198, Stellar Clusters and Associations: Convection, Rotation, and Dynamos, ed. R. Pallavicini, G. Micela, & S. Sciortino, 95

Piotto, G., Bedin, L. R., Anderson, J., et al. 2007, ApJ, 661, L53

Piskunov, A. E., Kharchenko, N. V., Röser, S., Schilbach, E., & Scholz, R. D. 2006, A&A, 445, 545

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, A&A, 576, A104

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 586, A138

Planesas, P., Scoville, N., & Myers, S. T. 1991, ApJ, 369, 364

Plummer, H. C. 1911, MNRAS, 71, 460

Poggio, E., Drimmel, R., Lattanzi, M. G., et al. 2018, MNRAS, 481, L21

Pöhnl, H. & Paunzen, E. 2010, A&A, 514, A81

Porceddu, I., Benvenuti, P., & Krelowski, J. 1991, A&A, 248, 188

Price, D. J. & Bate, M. R. 2008, MNRAS, 385, 1820

Price-Whelan, A. M., Nidever, D. L., Choi, Y., et al. 2019, ApJ, 887, 19

Purcell, E. M. 1969, ApJ, 158, 433

Purcell, E. M. & Pennypacker, C. R. 1973, ApJ, 186, 705

Puspitarini, L., Lallement, R., Babusiaux, C., et al. 2015, A&A, 573, A35

Puspitarini, L., Lallement, R., & Chen, H. C. 2013, A&A, 555, A25

Racine, R. & Harris, W. E. 1989, AJ, 98, 1609

Rand, R. J., Kulkarni, S. R., & Hester, J. J. 1990, ApJ, 352, L1

Randich, S., Tognelli, E., Jackson, R., et al. 2018, A&A, 612, A99

Rastello, S., Amaro-Seoane, P., Arca-Sedda, M., et al. 2019, MNRAS, 483, 1233

Rey-Raposo, R. & Read, J. I. 2018, MNRAS, 481, L16

Reynolds, R. J. & Cox, D. P. 1992, ApJ, 400, L33

Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3

Roberge, W. G. & Lazarian, A. 1999, MNRAS, 305, 615

Robichon, N., Arenou, F., Mermilliod, J. C., & Turon, C. 1999, A&A, 345, 471

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

Röser, S., Schilbach, E., & Goldman, B. 2019, A&A, 621, L2

Ryden, B. S. & Stark, A. A. 1986, ApJ, 305, 823

Salama, F., Galazutdinov, G. A., Krelowski, J., et al. 2011, ApJ, 728, 154

Salaris, M., Weiss, A., & Percival, S. M. 2004, A&A, 414, 163

Sandage, A. R. 1953, AJ, 58, 61

Sanders, W. L. 1971, A&A, 14, 226

Sandford, S. A., Allamandola, L. J., Tielens, A. G. G. M., et al. 1991, ApJ, 371, 607

Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6

Savage, B. D. & Sembach, K. R. 1996, ARA&A, 34, 279

Scarrott, S. M., Watkin, S., Miles, J. R., & Sarre, P. J. 1992, MNRAS, 255, 11P

Schultz, G. V. & Wiemer, W. 1975, A&A, 43, 133

Seabroke, G. M., Fabricius, C., Teyssier, D., et al. 2021, A&A, 653, A160

Sellgren, K., Werner, M. W., Ingalls, J. G., et al. 2010, ApJ, 722, L54

Sellwood, J. A. 2011, MNRAS, 410, 1637

Sembach, K. R., Howk, J. C., Ryans, R. S. I., & Keenan, F. P. 2000, ApJ, 528, 310

Serkowski, K., Mathewson, D. S., & Ford, V. L. 1975, ApJ, 196, 261

Sezima, T., Galadí-Enríquez, D., Paunzen, E., et al. 2015, in Highlights of Spanish Astrophysics VIII, 594–594

Shapley, H. 1918, ApJ, 48, 154

Shapley, H. 1930, Star Clusters, Vol. 2

Shaver, P. A., McGee, R. X., Newton, L. M., Danks, A. C., & Pottasch, S. R. 1983, MNRAS, 204, 53

Shen, Y., Draine, B. T., & Johnson, E. T. 2008, ApJ, 689, 260

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

Slettebak, A. 1949, ApJ, 110, 498

Snow, T. P., J. & Cohen, J. G. 1974, ApJ, 194, 313

Soderblom, D. R., King, J. R., Hanson, R. B., et al. 1998, ApJ, 504, 192

Soderblom, D. R., Nelan, E., Benedict, G. F., et al. 2005, AJ, 129, 1616

Sollerman, J., Cox, N., Mattila, S., et al. 2005, A&A, 429, 559

Soubiran, C., Cantat-Gaudin, T., Romero-Gómez, M., et al. 2018, A&A, 619, A155

Southworth, J., Maxted, P. F. L., & Smalley, B. 2004, MNRAS, 351, 1277

Southworth, J., Maxted, P. F. L., & Smalley, B. 2005, A&A, 429, 645

Stassun, K. G. & Torres, G. 2018, ApJ, 862, 61

Stassun, K. G. & Torres, G. 2021, ApJ, 907, L33

Tanaka, Y. & Bleeker, J. A. M. 1977, Space Sci. Rev., 20, 815

Telesco, C. M. & Harper, D. A. 1980, ApJ, 235, 392

Terlevich, E. 1987, MNRAS, 224, 193

Tielens, A. G. G. M. 2013, Reviews of Modern Physics, 85, 1021

Tielens, A. G. G. M., Tokunaga, A. T., Geballe, T. R., & Baas, F. 1991, ApJ, 381, 181

Todini, P. & Ferrara, A. 2001, MNRAS, 325, 726

Tompsett, M. F., Amelio, G. F., & Smith, G. E. 1970, Applied Physics Letters, 17, 111

Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rev., 18, 67

Torres, G. & Ribas, I. 2002, ApJ, 567, 1140

Tremblay, P. E., Gentile-Fusillo, N., Raddi, R., et al. 2017, MNRAS, 465, 2849

Tripicco, M. J., Dorman, B., & Bell, R. A. 1993, AJ, 106, 618

Trumpler, R. J. 1921, Lick Observatory Bulletin, 333, 110

Trumpler, R. J. 1925, PASP, 37, 307

Trumpler, R. J. 1930, Lick Observatory Bulletin, 420, 154

Tüllmann, R., Pietsch, W., Rossa, J., Breitschwerdt, D., & Dettmar, R. J. 2006, A&A, 448, 43

Uyanıker, B., Fürst, E., Reich, W., Aschenbach, B., & Wielebinski, R. 2001, A&A, 371, 675

van den Bergh, S. 1984, in Structure and Evolution of the Magellanic Clouds, ed. S. van den Bergh & K. S. D. de Boer, Vol. 108, 1–4

van Leeuwen, F. 1999, A&A, 341, L71

van Leeuwen, F. 2007, A&A, 474, 653

Van Winckel, H., Cohen, M., & Gull, T. R. 2002, A&A, 390, 147

Vande Putte, D., Garnier, T. P., Ferreras, I., Mignani, R. P., & Cropper, M. 2010, MNRAS, 407, 2109

VandenBerg, D. A. & Clem, J. L. 2003, AJ, 126, 778

Vasiliev, E. 2018, MNRAS, 481, L100

Vaughan, S., Willingale, R., O'Brien, P. T., et al. 2004, ApJ, 603, L5

Vázquez-Semadeni, E., Ryu, D., Passot, T., González, R. F., & Gazol, A. 2006, ApJ, 643, 245

Vladilo, G., Crivellari, L., Molaro, P., & Beckman, J. E. 1987, A&A, 182, L59

Vlasov, A., Vurm, I., & Metzger, B. D. 2016, MNRAS, 463, 394

Voshchinnikov, N. V. & Farafonov, V. G. 1993, Ap&SS, 204, 19

Vrba, F. J., Strom, S. E., & Strom, K. M. 1976, AJ, 81, 958

Webster, A. 1996, MNRAS, 282, 1372

Weingartner, J. C. & Draine, B. T. 1999, ApJ, 517, 292

Weingartner, J. C. & Draine, B. T. 2001, ApJ, 548, 296

Weingartner, J. C., Kolasi, E., & Woods, C. 2021, MNRAS, 504, 1164

Westerlund, B. E. & Krelowski, J. 1988, A&A, 203, 134

Westphal, A. J., Stroud, R. M., Bechtel, H. A., et al. 2014, Science, 345, 786

Wheeler, J. C. 1979, ApJ, 234, 569

Whittet, D. C. B., Bode, M. F., Longmore, A. J., et al. 1988, MNRAS, 233, 321

Wilking, B. A., Lebofsky, M. J., Martin, P. G., Rieke, G. H., & Kemp, J. C. 1980, ApJ, 235, 905

Williams, B. F. & Hodge, P. W. 2001, ApJ, 559, 851

Williams, J. P. & Cieza, L. A. 2011, ARA&A, 49, 67

Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278

Wouterloot, J. G. A. & Walmsley, C. M. 1986, A&A, 168, 237

Wu, Z.-Y., Zhou, X., Ma, J., & Du, C.-H. 2009, MNRAS, 399, 2146

Xu, Y., Hou, L. G., Bian, S. B., et al. 2021, A&A, 645, L8

Yan, Q.-Z., Zhang, B., Xu, Y., et al. 2019, A&A, 624, A6

Yi, S. K., Kim, Y.-C., & Demarque, P. 2003, ApJS, 144, 259

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579

Young, P. J., Westphal, J. A., Kristian, J., Wilson, C. P., & Landauer, F. P. 1978, ApJ, 221, 721

Zacharias, N., Finch, C. T., Girard, T. M., et al. 2013, AJ, 145, 44

Zacharias, N., Rafferty, T. J., & Zacharias, M. I. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 216, Astronomical Data Analysis Software and Systems IX, ed. N. Manset, C. Veillet, & D. Crabtree, 427

Zhao, H., Schultheis, M., Rojas-Arriagada, A., et al. 2021, A&A, 654, A116