**Surveying young stars with Gaia: Orion and the Solar neighbourhood**
Zari, E.M.

**Author**: Zari, E.M.
**Title**: Surveying young stars with Gaia: Orion and the Solar neighbourhood
**Issue Date**: 2019-10-22

# 1

# Introduction

## 1.1 Star formation in the *Gaia* era

Studying how stars form is at the core of contemporary astrophysics research. It is not only interesting in itself, but it is also essential in understanding the formation and early evolution of planetary systems, and the structure and the evolution of galaxies.

The last stage of the massive star formation process, and the context in which new stars are formed, are the so-called *OB associations*, groups of young and massive stars of spectral type O and B (Ambartsumian 1947). By noting that the spatial densities of stars in OB associations are well below the threshold necessary to prevent their disruption by Galactic tidal forces, Ambartsumian calculated that associations must be young ($< 25\,\mathrm{Myr}$), a conclusion that was supported by ages derived by colour-magnitude diagrams and by theory of stellar structure and evolution. This agrees well with the fact that these groups are usually located in or near star-forming regions, and hence are prime sites for the study of star formation processes and of the interaction of early-type stars with the interstellar medium (see Blaauw 1964; de Zeeuw et al. 1999). Although O and B stars are mostly found in associations, some of them do not seem to be associated with any group or cluster. A fraction of those moves at high velocity: these are the so-called *runaway stars* (Blaauw 1952; Blaauw & Morgan 1954; Ambartsumian 1955).

Since the work of Ambartsumian, much progress has been made in our knowledge of OB associations. At the end of the 20th century, the *Hipparcos* mission allowed for an extensive census of stellar content of the nearby OB associations (de Zeeuw et al. 1999). This was complemented, in the past two decades, by an unprecedented stream of new observational information and a parallel renaissance in theoretical investigation and numerical modelling of the star-formation process (see the reviews by McKee & Ostriker 2007; Kennicutt & Evans 2012). Yet, some questions remain unanswered. How are associations formed and how do they disperse in the field? What causes the distinction between the formation of bound open clusters and unbound associations? What are the characteristics of the stellar populations within single associations in terms of age sequences and kinematics? What are the properties of the ensemble of OB associations? What is their disposition in space and how does it compare with what is observed in other galaxies?

The data of the *Gaia* satellite are crucial to address these questions, as they allow to study the spatial structure, kinematics, and ages of OB associations with unprecedented precision. In this thesis we obtained a detailed census of the young stellar populations in the solar neighbourhood, focusing in particular on the Orion OB association. We found that both single associations and the ensemble of OB associations in the solar vicinity present a high degree of sub-structure in physical space, kinematics, and ages. The star formation history of the solar neighbourhood is complex, and it does not quite follow sequential star formation scenarios. This calls for a revision of our theories of the propagation and triggering of star formation. Data from the future releases of the *Gaia* satellite and from upcoming spectroscopic surveys will also contribute in exploring in more detail the kinematic and physical sub-structure of large star formation complexes.

In the remainder of this introduction, I will discuss the main features of OB associations. I will focus in particular on their spatial arrangement in the solar neighbourhood and on the properties of the Orion OB association, and I will describe the characteristics of O- and B-type runaway stars. I will give a short overview of the data products of the *Gaia* satellite and I will finally summarise the Chapters of this thesis and present some prospects for future research.

## 1.2 OB associations

OB associations were first recognised as loose groups of O- and B-type stars, but they contain members across the mass spectrum, including intermediate-mass A/F stars and lower-mass G/K/M stars, which are still in the pre-main sequence (PMS) phase of stellar evolution. Though the lower mass ($< 1.5\,\mathrm{M}_\odot$) stars blend in with the Galactic field population and are therefore much more difficult to identify than the OB stars, they comprise the dominant stellar component of OB associations (Briceño et al. 2007b).

The members of OB associations can be singled out using a combination of instruments and techniques, summarised for example in Brown et al. (1999) and Briceño et al. (2007a). Methods based on single-epoch photometry and on proper motions (and on their combination) were applied in this thesis. Low-mass PMS stars are located in the colour-magnitude diagram (CMD) above the zero-age main sequence (ZAMS). For this reason it is relatively straightforward to separate them from main sequence sources located at similar distances (see for instance Sherry et al. 2004; Kenyon et al. 2005; Bouy et al. 2014). Proper motion surveys allow to identify members of OB associations based on their kinematics (see for example de Zeeuw et al. 1999; de Bruijne 1999a; Hoogerwerf & Aguilar 1999). Indeed, associations are gravitationally unbound, however they have small internal velocity dispersion (a few kilometres per second), and thus they form coherent structures in velocity space. The streaming motion of the association as a whole, as well as the Solar motion, is reflected as a motion of the members towards a convergent point on the sky. An example of this is shown in Fig. 1.1, for the nearest OB association, Scorpius-Centaurus.

Precise proper motions allow to study the internal kinematic properties of OB associations, which provide clues for the understanding their formation. To explain the origin of OB associations two main competing models have been proposed. According to the first model (Lada & Lada 2003), OB associations are expanding remnants of star clusters. Star clusters are formed embedded within molecular clouds, where the gravitational potential of both the stars and the gas holds them together. When feedback disperses the gas left over from star formation, the cluster becomes super-virial and will expand and disperse, thus being visible for a short time as an OB association.

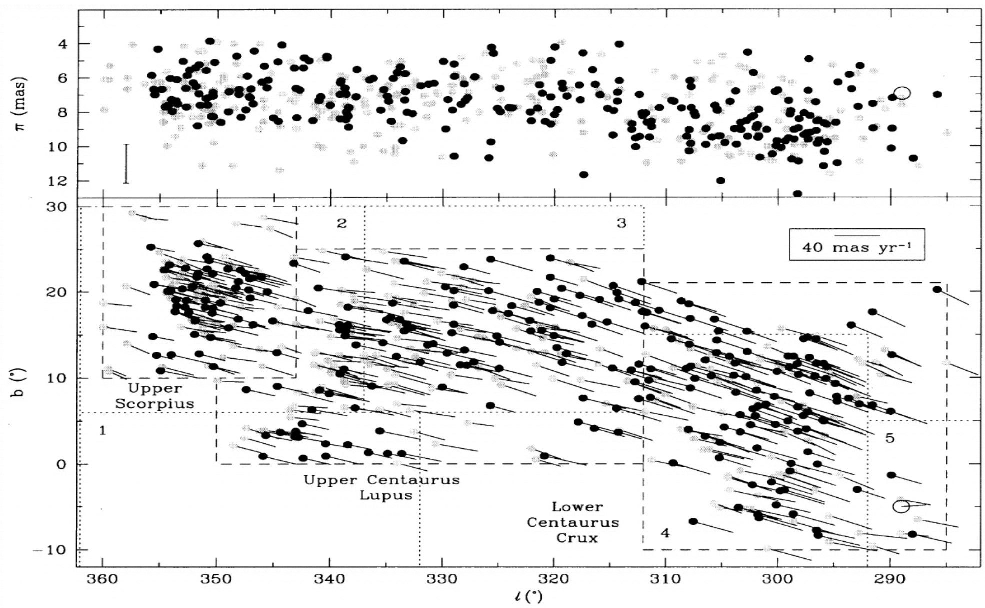


**Figure 1.1:** Positions and proper motions (bottom), and parallaxes (top), for 521 members of the Scorpius-Centaurus association (Sco OB2) selected from 7974 stars in the *Hipparcos* catalogue in the area bounded by the dashed lines (de Zeeuw et al. 1999). The vertical bar in the top panel corresponds to the average $\pm 1\sigma$ parallax range for the stars shown. The dotted lines are the schematic boundaries of the classical subgroups Upper Scorpius (2, US), Upper Centaurus Lupus (3, UCL), Lower Centaurus Crux (4, LCC), and the candidate subgroups (1 and 5) defined by (Blaauw 1964). The large open circle represents the open cluster IC 2602. The figure and the caption are from de Zeeuw et al. (1999).

The second model (Clark et al. 2005) instead predicts that OB associations are born in highly sub-structured, multiple small-scale star formation events that take place in long and filamentary molecular clouds. The kinematics of OB associations would keep memory of the parental gas sub-structure where they originated. The results reported by Wright et al. (2016), Wright & Mamajek (2018), and in this thesis (Chapter 3) seem to confirm the latter view.

A problem that both models need to explain is the star formation history of OB associations. Indeed, although OB associations as a whole occupy large regions in physical space ($\sim 100\,\mathrm{pc}$), they can be divided in smaller sub-groups, that can be distinguished on the basis of the ages of their members, their degree of association with interstellar matter (Blaauw 1964), and on the basis of their kinematics (see for example Cantat-Gaudin et al. 2019; Kounkel et al. 2018). Simple triggered star formation scenarios (see Preibisch & Zinnecker 2007, and references therein) struggle in explaining the lack of regular age sequences and the apparent coordination of star formation on large spatial scales, and more complex models are required to explain the observations (see for instance Krause et al. 2018).

## 1.3 The Gould Belt

OB associations in the solar vicinity seem to be arranged in a ring-like structure, inclined by $\sim 20^\circ$ with respect to the plane of the Milky Way. This structure was recognised by by Herschel (1847) and Gould (1874) and became known as the *Gould Belt*. This huge ring of bright stars and gas, up to 700 pc in diameter, seems to link a number of the closest associations, some of which fit a coherent pattern of expansion and rotation (Lindblad et al. 1997; Torra et al. 1997). The Gould Belt was also found to be associated with young stars (Guillout et al. 1998) and interstellar material (Lindblad 1967), the latter interpreted as an expanding ring of gas (Olano 1982; Elmegreen 1982). Various scenarios have been proposed to explain the formation of the Belt, which include the passage of the Carina spiral arm near the Sun (Elmegreen 1993; Elmegreen & Efremov 1998), the impact of an high velocity cloud on the stellar disc (Comeron et al. 1998), a cascade of supernova explosions (Olano 2001), and the collision between a dark matter clump and a gas cloud (Bekki 2009). Elmegreen (1993) in particular proposed that the passage of the Carina spiral arm $\sim 60$ Myr ago triggered the formation of the Cas-Tau association. The Lindblad's ring could have been then generated by feedback and supernova explosions from high-mass stars in Cas-Tau. The Scorpius-Centaurus, Orion, Perseus, and Lacerta OB associations would have formed around 20 Myr ago from Lindblad's ring and constituted a second generation of star formation. The present star formation seen in Taurus and in Ophiuchus is regarded as the third generation. Figure 1.2 shows locations of the OB associations studied in de Zeeuw et al. (1999) projected onto the Galactic plane. de Zeeuw et al. (1999) concluded that the physical arrangement of the ensemble of OB associations was in qualitative agreement with Elmegreen (1993) picture, but called for a reassessment of the star formation history of the solar neighbourhood, as they observed that there was not a clear difference between bound open clusters and unbound expanding associations and that the total mass of young stellar groups might have been underestimated.

Such reassessment came ten years later, when Elias et al. (2009) studied the distribution of young open clusters in the solar neighbourhood, using again the *Hipparcos* catalogue. They proposed that the position with respect to the galactic plane and the kinematics of the two associations dominating the inclination of the Gould Belt, Orion and Scorpius-Centaurus, can be explained in terms of their relative position to the density maximum of the Local Arm in the solar neighbourhood. They therefore concluded that the Gould Belt could be explained by the result of the internal dynamics of the Galactic disc.

This conclusion has been further corroborated by Bouy & Alves (2015). Bouy & Alves (2015) re-analysed the distribution of O- and B-type stars in the solar neighbourhood, and by making use of a three-dimensional kernel estimation, they studied their spatial density and produced the three-dimensional density map shown in Fig. 1.3. They suggested that the distribution of O and B stars in the solar neighbourhood would be better described by stream-like structures, similarly to what is observed in other spiral galaxies, and concluded that there is no evidence of a ring-like structure such as the Gould Belt in the three dimensional configuration of young, bright stars in the solar neighbourhood. Bouy & Alves (2015) results were based on the *Hipparcos* data, and motivated us to perform the study presented in Chapter 4.

## 1.4 Orion

The figure of Orion the Hunter is a familiar sight in the winter sky of the Northern hemisphere (see Fig. 1.4). The area is an extraordinarily active site of star formation. Over the years, no similar region has received such intense astronomical scrutiny, or has been studied with such a variety of observational tools (see the reviews by Stahler & Palla 2005; Bally 2008).

The *Hipparcos* census of nearby OB associations (de Zeeuw et al. 1999) represented a major step forward in terms of determining the membership of OB associations, however the data was not accurate enough to make significant progress in Orion. The main reasons for this are that a) the distance to Orion is $\sim 400\,\mathrm{pc}$, thus the *Hipparcos* parallax uncertainties were large, and b) Orion's motion is mainly directed away from the Sun, thus the observed proper motions are small. Thus, a detailed characterisation of the stellar population of Orion in terms of kinematics, ages, and spatial structure was still missing: this constitutes one of the main topics of this thesis. In the following, we will describe the features of the Orion region relevant to this thesis.

The Orion OB association (Ori OB1) is divided in several groups and clusters, partially super-imposed along the line of sight (Blaauw 1964; Brown et al. 1994). Blaauw (1964) suggested that star formation sequentially propagated in the association. The members of the oldest sub-group, located north-west of the Belt stars (Ori OB1a, 8-12 Myr, Bally 2008) may have triggered the formation of the Ori OB1b sub-group (3-6 Myr) towards the Orion's Belt, from which star formation seemed to have propagated further south in the Ori OB1c region (2-6 Myr). The youngest sub-group is the Orion Nebula Cluster (ONC, see for instance Da Rio et al. 2014), located at the northern tip of the Orion A molecular cloud. Within these four groups, many clusters have been identified, such as 25 Ori (Briceño et al. 2007b), $\sigma$ Ori (Walter et al. 2008) and $\lambda$ Ori (Mathieu 2008). Spectroscopic data allowed to analyse the kinematic properties of

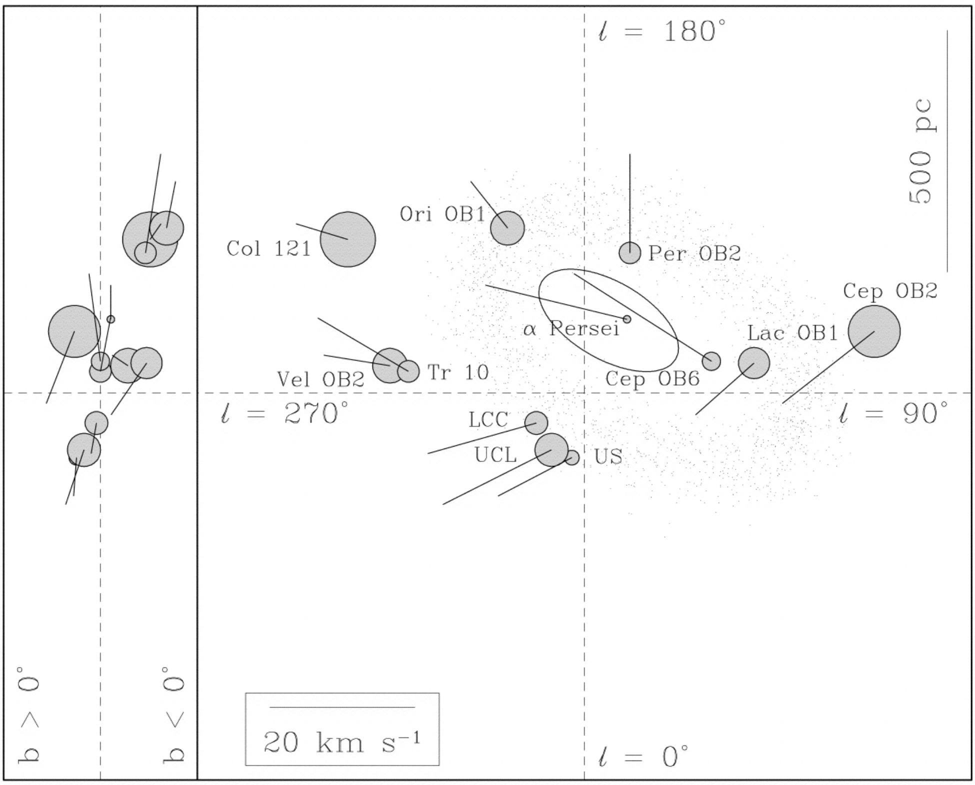


**Figure 1.2:** Locations of the OB associations studied in de Zeeuw et al. (1999) projected onto the Galactic plane. The gray circles indicate the physical dimensions as obtained from the angular dimensions and mean distances, on the same scale. The lines represent the streaming motions, derived from the average proper motions, mean distances and median radial velocities of the secure members, corrected for "standard" solar motion and Galactic rotation. The ellipse around the $\alpha$ Persei cluster indicates the Cas-Tau association. The small dots schematically represent the Olano (1982) model of the Gould Belt. The figure and caption are from de Zeeuw et al. (1999).

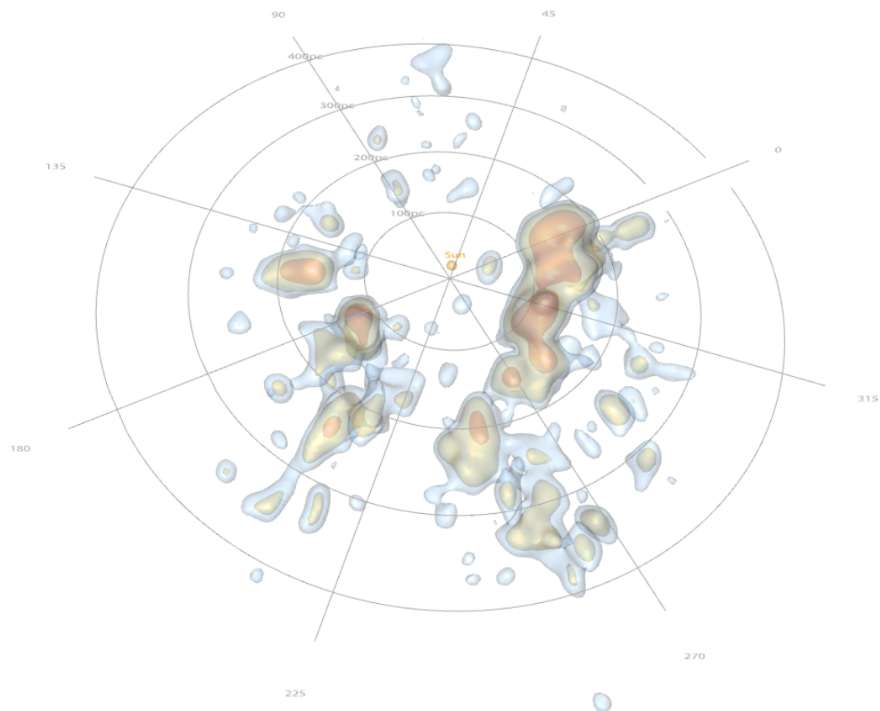


**Figure 1.3:** 3D map of OB star density iso-surface (1.0, 1.38 and 2.76 $\times 10^{-4}$ OB star per $\mathrm{pc}^3$ (Bouy & Alves 2015). The circles have 100, 200, 300, and 400 pc respectively. The radii represent longitude values. The figure and the caption are from Bouy & Alves (2015).

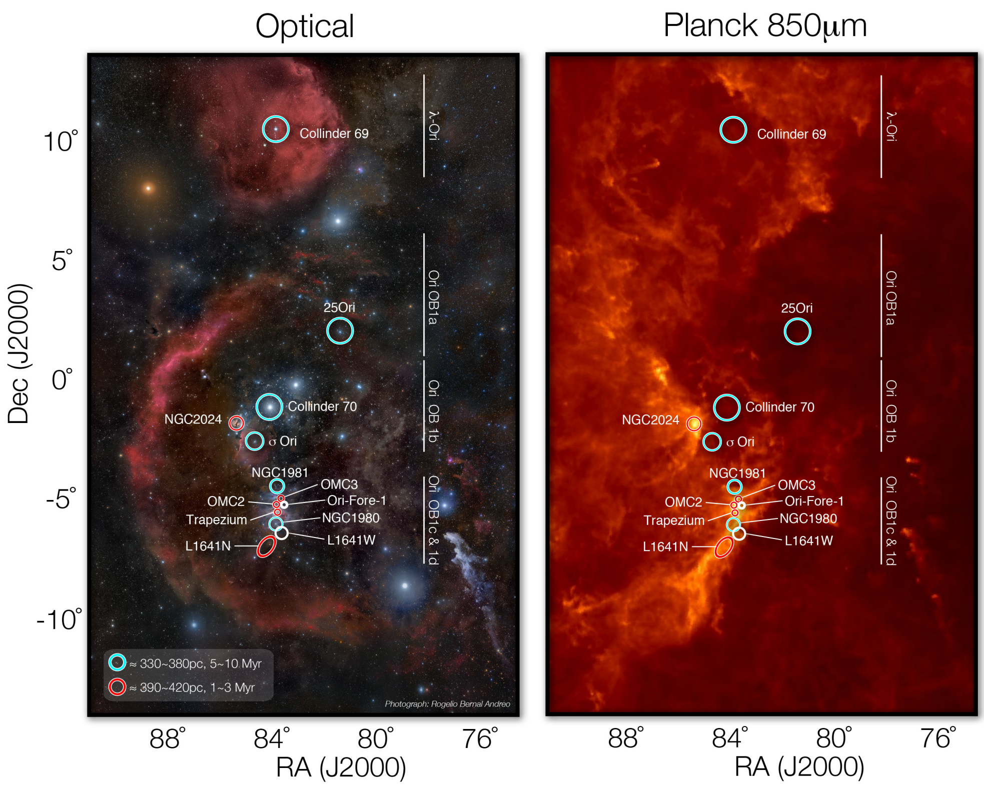


**Figure 1.4:** Left panel: distribution of groups over-plotted on an optical photograph of the Orion constellation (courtesy of Rogelio Bernal Andreo - *DeepSkyColors.com*). Right panel: same as left panel, but over-plotted on a far-infrared (850 $\mu$m) *Planck* map. The figure and the caption are from Bouy et al. (2014).

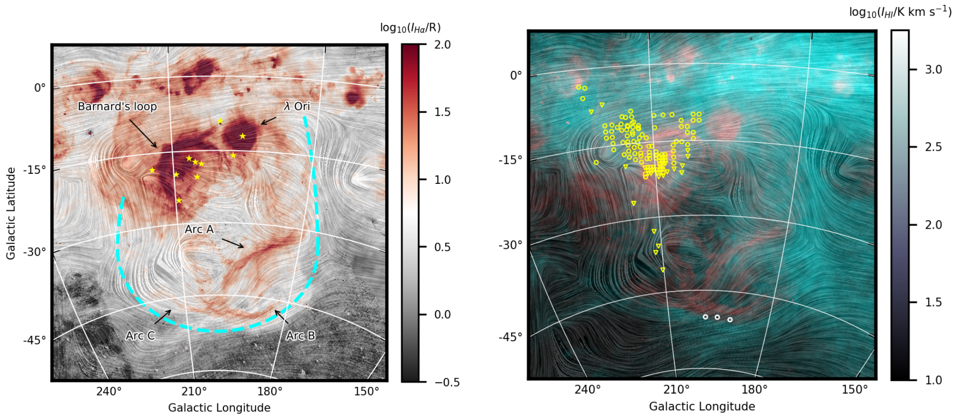


**Figure 1.5:** The drapery pattern corresponds to the plane-of-the-sky magnetic field orientation inferred from the Planck 353 GHz polarisation observations. Left. Total integrated H$\alpha$ emission map. The dashed line indicates the approximate location of the edge of the super-bubble. The yellow symbols correspond to the main stars in the Orion constellation . Right. Total integrated H$\alpha$ emission and HI 21 cm emission integrated between - 20 and 20 $\mathrm{km\,s^{-1}}$ shown in red and teal colours, respectively. The yellow symbols correspond to the line-of-sight magnetic field directions derived from the HI emission-line Zeeman splitting observations. The circles and triangles correspond to magnetic fields pointing toward and away from the observer, respectively. The three white circles in the bottom are the regions analysed in Soler et al. (2018), from which these figure and caption are taken.

these groups. Briceño et al. (2007b) and Jeffries et al. (2006) found in particular that the 25 Ori and $\sigma$ Ori clusters have different kinematic properties than the sub-groups in which they are located (Ori OB1a and OB1b, respectively). Alves & Bouy (2012) and Bouy et al. (2014) recently reported the discovery of a young population of stars in the foreground of the ONC, which was however questioned by Da Rio et al. (2016), Fang et al. (2017) and Kounkel et al. (2017a), while Kubiak et al. (2016) identified a rich and young population surrounding $\epsilon$ Ori.

The combined effects of UV radiation, stellar winds, and supernova explosions from the Orion OB1 association have created a bubble that spans $\sim 40^{\circ}$ in the sky (or 300 pc at a distance of 400 pc): the Orion-Eridanus super-bubble. Ochsendorf et al. (2015) studied in detail the structure and the evolution of the Orion-Eridanus super-bubble, concluding that it consists of a series of nested shells. They also found that Barnard's Loop is a part of a complete bubble, probably associated with a supernova remnant. Both the Barnard's Loop bubble and the $\lambda$ Ori Bubble are expanding within the Orion-Eridanus super-bubble. By using polarization observations by the *Planck* satellite, Soler et al. (2018) characterised the magnetic field in the Orion-Eridanus super-bubble, finding that the large-scale magnetic field in the region was primarily shaped by the expanding super-bubble (see Fig. 1.5).

Orion contains two giant molecular clouds ($M \sim 10^5\,\mathrm{M_{\odot}}$): the Orion A molecular cloud, located in the southern portion of the constellation, and the Orion B cloud, that lies at the east of the Orion's Belt (Bally 2008). Both the clouds are thought to be located within the walls of the Orion-Eridanus super-bubble. Schlafly et al. (2015) presented 3D maps of dust reddening, tracing the total column density towards the Orion clouds (reported in Fig. 1.6). They found that the Orion A and B clouds are

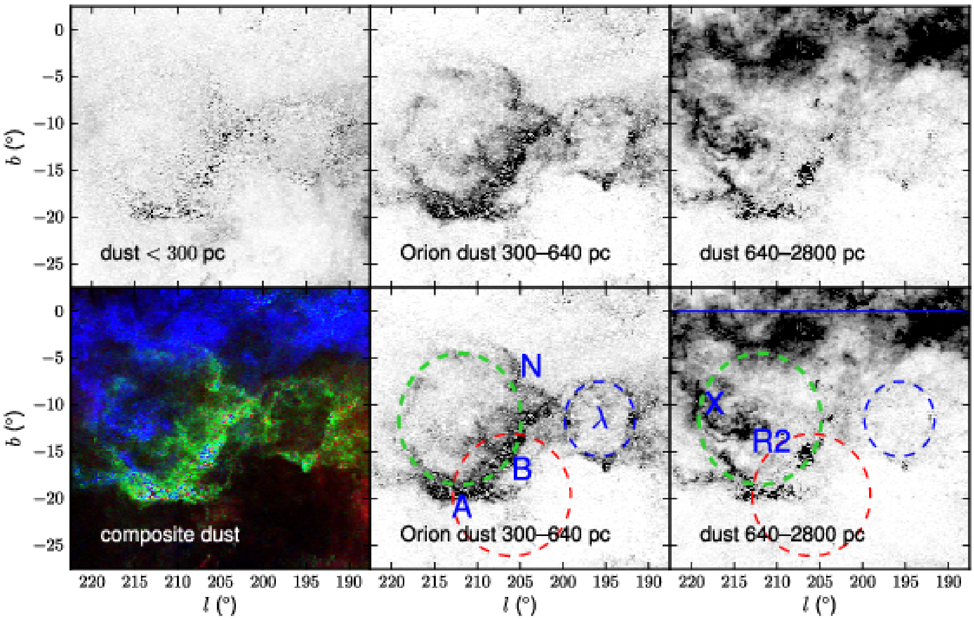


**Figure 1.6:** The 3D distribution of dust towards the Orion Molecular Complex. The top panels show the column density of dust with distance < 300 pc, 300-640 pc, and 640-2800 pc, respectively. The fourth panel (bottom left) shows a 3-colour composite image of these three slices, illustrating the 3D distribution of dust in the region. Finally, the fifth and sixth panels again show the Orion and more distant dust, this time over-plotting circles tracing the various bubble-like structures in the region. The green dashed circle shows the Orion dust ring; the blue dashed circle shows the $\lambda$ Ori molecular ring; and the red dashed circle approximately aligns with Barnard's Loop (see Figs. 1.4 and 1.5). The last two panels also label the Orion A (A) and Orion B (B) molecular clouds, the Northern Filament (N), the star $\lambda$ Ori, Monoceros R2 (R2), the Crossbones (X), and the Galactic plane (horizontal line). Differential extinction and an insufficient number of well-observed stars lead to artefacts in the far distance slice through particularly dense clouds in Orion A and B. White to black corresponds to 0-0.7 mag E(B-V). This same scale is used for each of the colour planes in the lower left panel. The figure and the caption are from Schlafly et al. (2015).

part of a "dust ring", which may have implications on the triggering of star formation in the region.

## 1.5 OB stars on the run

As mentioned in Section 1 of this introduction, not all O- and B-type stars are found in OB associations and clusters. A large fraction of these field objects moves at very high velocities: these are referred to as "runaway" stars (Blaauw 1952; Ambartsumian 1955). Orion has been the source of several well known runaway stars, including the $150\,\mathrm{km\,s^{-1}}$ runaway star AE Auriga, and the $117\,\mathrm{km\,s^{-1}}$ $\mu$ Columbae which is moving exactly in the opposite direction (Blaauw 1991). Hoogerwerf et al. (2001) used new *Hipparcos* proper motion data to show that these two stars, and the colliding wind X-ray binary $\iota$ Ori were at the same location in the sky $\approx 2.6$ Myr ago. Gualandris et al. (2004) argue that the two runaway stars and $\iota$ Ori suffered a four-body interaction in which two binaries in the same cluster underwent an exchange. The two most-massive members became the tight $\iota$ Ori binary; the gravitational energy released kicked the two less massive stars out of the region at high velocity. The process explained above describes one of the two runaway production channels, and it is usually referred to as *dynamical ejection scenario* (DES, Poveda et al. 1967; Leonard 1991). The second scenario, the *binary supernova scenario* (BSS, Blaauw 1961; Zwicky 1957; Boersma 1961), predicts that a runaway star might originally have been a member of a close binary pair consisting of two massive stars. If the companion exploded as a supernova, the star of interest could escape with a speed equal to the orbital value. Runaway stars have been identified in the *Hipparcos* catalogue by Hoogerwerf et al. (2001) and Tetzlaff et al. (2011). Hoogerwerf et al. (2001) selected 56 sources of spectral type from O to B5 with total peculiar velocities higher than $30\,\mathrm{km\,s^{-1}}$, and, by studying their orbit identified the parent associations for a sub-set of them. Tetzlaff et al. (2011) identified young stars ($< 50$ Myr) of any spectral type, and selected those with large peculiar velocities, finding in total 2547 candidate runaway stars. The present and upcoming *Gaia* data releases are expected to drastically increase the available sample of stars with precisely known velocities, allowing for the construction of more complete samples of runaway star candidates. This is the goal of Chapter 5. Such samples will be then compared with the results of numerical simulations that predict the fraction of runaway stars produced by the BSS or the DES, such as those by Renzo et al. (2019b) and Ryu et al. (2017). This will be useful to determine the relative importance of the two formation mechanisms, which in turn will provide more clues on massive star formation and evolution (see Renzo et al. 2019b; Gvaramadze et al. 2009; Portegies Zwart et al. 2007).

## 1.6 Gaia

*Gaia* is an ESA mission, launched at the end of 2013 (Perryman et al. 2001; Gaia Collaboration et al. 2016a). The main aim of *Gaia* is to measure the three-dimensional spatial and the three-dimensional velocity distribution of stars and to determine their astrophysical properties, such as surface gravity and effective temperature, to map and understand the formation, structure, and past and future evolution of our Galaxy. *Gaia*'s astrometry delivers absolute parallaxes and proper motions. Complementary photometry and radial velocities are also provided by *Gaia* so that astrophysical parameters and six dimensional phase space information can be derived.

Two years and half after the launch, the first release of data was presented (hereafter *Gaia* DR1). *Gaia* DR1 (Gaia Collaboration et al. 2016a,b) is based on the first 14 months of mission and consists of three components. The first component consists of a primary astrometric data set which contains the positions, parallaxes, and mean proper motions for about 2 million of the brightest stars in common with the *Hipparcos* and Tycho-2 catalogues, the Tycho-Gaia Astrometric Solution (TGAS), and a secondary astrometric data set containing the positions for an additional 1.1 billion sources. The second component is the photometric data set, consisting of mean G-band magnitudes for all sources. The third component is formed by the G-band light curves and the characteristics of $\sim 3000$ Cepheid and RR Lyrae stars, observed at high cadence around the south ecliptic pole. For the primary astrometric data set the typical uncertainty is about 0.3 mas for the positions and parallaxes, and about 1 mas/yr for the proper motions. A systematic component of $\sim 0.3$ mas should be added to the parallax uncertainties. For the subset of $\sim 94\,000$ Hipparcos stars in the primary data set, the proper motions are much more precise at about 0.06 mas/yr. For the secondary astrometric data set, the typical uncertainty of the positions is $\sim 10$ mas. The median uncertainties on the mean G-band magnitudes range from the mmag level to $\sim 0.03$ mag over the magnitude range 5 to 20.7 mag.

The second *Gaia* data release (*Gaia* DR2, Gaia Collaboration et al. 2018a), which is based on the data collected during the first 22 months of the nominal mission lifetime, was made public on the 25th of April 2018. *Gaia* DR2 represents a major advance with respect to *Gaia* DR1, making the leap to a high-precision parallax and proper motion catalogue for over 1 billion sources, supplemented by precise and homogeneous multi-band all- sky photometry and a large radial velocity survey at the bright ($G \sim 13$ mag) end. *Gaia* DR2 contains celestial positions and the apparent brightness in $G$-band for approximately 1.7 billion sources. For 1.3 billion of those sources, parallaxes and proper motions are in addition available. This data release contains four new elements: broad-band colour information in the form of the apparent brightness in the $G_{\rm BP}$ (330-680 nm) and $G_{RP}$ (630-1050 nm) bands for 1.4 billion sources; median radial velocities for $\approx 7$ million stars; for between 77 and 161 million sources estimates of the stellar effective temperature, extinction, reddening, and radius and luminosity; variability information for 0.5 million stars; epoch astrometry and photometry for a pre-selected list of 14 000 minor planets in the solar system.

## 1.7 This thesis

In Chapter 2, we use *Gaia* DR1 to explore the three-dimensional arrangement and age ordering of the many stellar groups toward the Orion OB association, aiming at a new classification and characterisation of the stellar population not embedded in the Orion A and B molecular clouds. We find evidence for the presence of a young population at a parallax $\varpi \approx 2.65\,\mathrm{mas}$, which is loosely distributed around the following known clusters: 25 Ori, $\epsilon$ Ori, and $\sigma$ Ori, and NGC 1980 ($\iota$ Ori) and the Orion Nebula Cluster (ONC). The low mass counterpart of this population is visible in the colour-magnitude diagrams constructed by combining *Gaia* DR1 G-band photometry and 2MASS, and in the density distribution of the sources on the sky. We estimate the ages of this population using a Bayesian isochronal fitting procedure assuming a unique parallax value for all the sources, and we infer the presence of an age gradient going from 25 Ori (13-15 Myr) to the ONC (1-2 Myr). Finally, we provisionally relate the stellar groups to the gas and dust features in Orion. These results represent the first step toward using Gaia data to unravel the complex star formation history of the Orion region in terms of the various star formation episodes, their duration, and their effects on the surrounding interstellar medium.

In Chapter 3, we present a study of the three dimensional structure, kinematics, and age distribution of the Orion OB association, based on *Gaia* DR2. The goal of this Chapter is to obtain a complete picture of the star formation history of the Orion complex and to relate our findings to theories of sequential and triggered star formation. We select the Orion population with simple photometric criteria, and we explored its physical arrangement by using a three dimensional density map. The map shows structures that extend for roughly $150\,\mathrm{pc}$ along the line of sight, divided in multiple sub-clusters. We separate the different groups by using a density-based clustering algorithm, and we studied their kinematic properties first by inspecting their proper motion distribution, and then by applying a kinematic modelling code based on an iterative maximum likelihood approach, which we use to derive their mean velocity, velocity dispersion and isotropic expansion. By using an isochrone fitting procedure we provide ages and extinction values for all the groups. We confirm the presence of an old population ($\sim 15$ Myr) towards the 25 Ori region, and we find that groups with ages of $12 - 15\,\mathrm{Myr}$ are present also towards the Belt region. We notice the presence of a population of $\sim 10$ Myr also in front of the Orion A molecular cloud. Our findings suggest that star formation in Orion does not follow a simple sequential scenario, but instead consists of multiple events, which caused kinematic and physical sub-structure. To fully explain the detailed sequence of events, specific simulations and further radial velocity data are needed.

In Chapter 4, we study the three dimensional arrangement of young stars in the solar neighbourhood using *Gaia* DR2 and we provide a new, original view of the spatial configuration of the star-forming regions within 500 pc of the Sun. By smoothing the star distribution through a Gaussian filter, we construct three dimensional density maps for early-type stars (upper-main sequence, UMS) and pre-main sequence (PMS) sources. The PMS and the UMS samples are selected through a combination of photometric and astrometric criteria. A side product of the analysis is a three-dimensional, G-band extinction map, which we use to correct our colour-magnitude diagram for extinction and reddening. Both density maps show three prominent structures, Scorpius-Centaurus, Orion, and Vela. The PMS map shows a plethora of lower-mass star-forming regions, such as Taurus, Perseus, Cepheus, Cassiopeia, and Lacerta, which are less visible in the UMS map due to the lack of large numbers of bright, early-type stars. We estimate ages for the PMS sample and we study the distribution of PMS stars as a function of their age. We find that younger stars cluster in dense, compact clumps, and are surrounded by older sources, whose distribution is instead more diffuse. The youngest groups that we find are mainly located in Scorpius-Centaurus, Orion, Vela, and Taurus. Cepheus, Cassiopeia, and Lacerta are instead more evolved and less numerous. We conclude that the 3D density maps show no evidence for the existence of the ring-like structure which is usually referred to as the Gould Belt.

In Chapter 5, we search for early type runaway stars within 1 kpc from the Sun by using *Gaia* DR2 and the stellar parameters provided in the `StarHorse` catalogue (Anders et al. 2019). We select upper main sequence (UMS) sources by applying simple photometric cuts. Our sample consists of O-, B- and early A-type sources. We study the tangential velocity, and, when possible, the total velocity distribution of our sample, and we classify as candidate runaway stars those sources that have tangential velocities significantly different from the rest of the population ($2\sigma$) or total velocities higher than $30\,\mathrm{km\,s^{-1}}$. We study the orbits of the candidate runaway stars with literature radial velocities, and we find that around half of our sources were produced further than 1 kpc. We focus on the runaway star candidates in the Orion and Scorpius-Centaurus (Sco-Cen) regions. In Orion, we confirm previously known runaway stars and we enlarge the sample by adding 6 new runaway candidates. In

Sco-Cen we identify two runaway star candidates that likely share the same origin. The analysis of the entire sample is on-going. Finally, we discuss our findings in the context of other studies, end we estimate the completeness of our sample. To further study the candidate runaway stars, more radial velocities are needed. These could be obtained from planned surveys, such as SDSS-V, WEAVE, and 4MOST, but also from dedicated proposals.

## 1.8 Outlook

In this thesis we have found many clues indicating that OB associations are complex entities, and we provided a description of their properties in terms physical structure, kinematics, and ages. We used early-type massive stars and pre-main sequence sources to trace the structure of the solar neighbourhood within 500 pc from the Sun and we studied the kinematics and dynamics of some of the fastest young stars in the Milky Way.

We did not however answer many questions, starting from: how are OB associations formed? The scenarios proposed to model the formation of OB associations assume that radiation and winds from massive stars disperse the gas surrounding them, locally terminating the star formation process and driving shocks in other regions, which cause cloud collapse and new star formation episodes. The models make different predictions for the observations, however none of them seem to completely explain the data. Recent studies on the Scorpius-Centaurus (Sco-Cen) association by Pecaut & Mamajek (2016) and Krause et al. (2018) show complex star formation histories, indicative of multi-stage formation processes, and not consistent with simple triggered star formation scenarios. Future *Gaia* data releases complemented with spectroscopic surveys, such as SDSS-V, WEAVE, and 4MOST, will enormously increase our knowledge of the formation and evolution of OB associations, both in the solar vicinity and in distant regions of the Milky Way. At the same time, detailed simulations of large scale star formation events will be needed to interpret the data. Another way to test theories of triggered star formation is to compare the kinematics of past and present star formation episodes. For this purpose it will be possible to combine the data from future *Gaia* releases with proper motions data in the infra-red such as those of the VISIONS survey. VISIONS, the VISTA star formation atlas, is a survey that aims to construct a sub-arcsec near-infrared atlas of all nearby ($< 500\,\mathrm{pc}$) star formation complexes from the southern hemisphere. The survey will provide multi-epoch, H-band observations that will be used to derive proper motions for the sources observed, with precision of $1-2\,\mathrm{mas/yr}$. By using VISIONS, it will be possible to relate the motions of embedded sources, invisible to *Gaia*, with those of evolved young stars that have already dispersed the gas surrounding them.

Going beyond the solar neighbourhood, O and B-type stars can be used to trace the structure of the spiral arms, and can probe the spiral arm features in remote regions of the Milky Way. Indeed, our position in the disc of the Milky Way does not allow to capture the global picture easily. For example, the number of spiral arms is still somewhat debated, although it is considered to be either 2 or 4. This has implications on the structure of our Galaxy: a large number of arms would support the view that the Galaxy better resembles a flocculent, rather than grand design spiral. An alter-

native interpretation is that the Galaxy has two main spiral arms, with the other two arms perhaps only present in gas and young stars (Drimmel 2000) By combining future *Gaia* data releases with, for instance CO and dust extinction maps we will be able to study in detail the connection between the spatial configuration and the different kinematic properties of gas and stars in the disc of the Galaxy. By doing so we will address two fundamental questions in large scale star formation studies:
1) which are the mechanisms triggering and propagating star formation in the Galaxy?
2) how is the interstellar medium shaped and transformed under the influence of young massive stars?